A THEORY OF SAPIENCE:

Using Systems Science to Understand the Nature of Wisdom and the Human Mind

George Mobus

TABLE OF CONTENTS

Acknowledgements

The University of Washington Tacoma and the School of Engineering and Technology (SET) have graciously provided me with office space and computer support while completing this book. My thanks to the Dean of SET, Rajendra (Raj) Katti, for providing me with the support and very useful discussions on this subject. Thanks as well to my co-author on our previous project, “Principles of Systems Science”, Professor Emeritus, Michael Kalton, also for many great conversations delving into the nature of sapience and the human condition. Numerous other scholars have provided useful commentary and feedback on a number of the ideas presented in this work, unfortunately too many to list here.

Preface

Understanding the Human Condition

Consider this perplexing question: If human beings are so smart, why have we not solved the many social and physical problems that have plagued humanity throughout history? Are poverty, war, and genocide (to name a few) really going to be with us always? Moreover, why are we facing nearly imponderable problems that have developed due to our own inept actions, e.g. climate change?

An on-going major presumption about our species has been that we are supremely *intelligent*, able to solve difficult problems and create technologies to satisfy our needs and wants. We are a *clever* species, far more so than any other. We can communicate using language. We can learn how the world (meaning the universe) works through science. We can find and exploit numerous resources. We can occupy any environment on the planet and even off-planet, at least for short periods of time. From earliest times we have seen our species as fundamentally different from every other species. Indeed we have very often considered ourselves to be something different from other life forms, superior, in control of our destinies. The very foundations of both theological and secular humanism start with the presumption that human beings are exceptional in creationist or evolutionary terms respectively. We all seem to think we are pretty exceptional and of primary importance to the world.

Yet, as the first decade of the 21st Century came to an end the evidence that we had made some serious mistakes in judgments and choices with global negative impacts became effectively undeniable. Many very serious scientists, philosophers, historians, and other disciplinarians recognized the signs of dysfunction and questioned just how superior our intelligence might be under the circumstances.

Consider the effectiveness of various major governments in the world. How effective are they at identifying the causes of these problems? How effective are they at establishing safe, secure, and prosperous conditions for all of their citizens? Consider the political processes that determine who shall be put into decision-making positions in those governments. Are those processes really working very well?

We not only face problems that threaten us physically. We face problems with how we are even going to attack solving those problems.

A growing number of cross-disciplinarians, those who find relations between different areas of knowledge, began calling attention to an important aspect of all of these problems. It is clear that all of them are interconnected. For example, burning fossil fuels, which are a finite resource, pumps excessive carbon dioxide into the atmosphere and oceans where it contributes to global warming and ocean acidification. But it also reduces the reserves of energy that is needed to drive our economies. As the fuels are depleted they become more expensive[1]. The cost of everything goes up proportionately because everything we do, everything in the economy is based on energy inputs to do real work. As a result the economies suffer. At the same time, increasing costs threaten our ability to mitigate the warming or adapt to problems like sea level rising. Those activities will likely be very expensive. Can we afford them? The implications are disturbing. Yet, so far, our governments seem incapable of any meaningful actions that would reduce the threats and prepare us for adaptation to the changes in climate we fully expect to affect our lives[2].

Human minds, through creativity and intelligence, have developed the ability to extract and burn fossil fuels to increase the comfort of our lives; we possess a remarkable cognitive ability in the animal world. But those very capabilities have produced a situation that is threatening our very existence if not addressed soon. The issues are not just complex, they are systemic. But even while many similarly clever minds recognize the dangers wrought by our inventions and pell-mell consumption of the comfort afforded, we, as a species, seem unable to exercise the wisdom needed to preserve our future on the planet. "*If human beings are so smart, why do we find ourselves in this predicament*?"

The Missing Element – Wisdom

Why do these problems exist in the first place? As stated, humans are incredibly intelligent and creative. They came up with technologies that obtain and use this energy to make our lives more comfortable and to create multiple forms of what we call 'wealth.' On the one hand we are better off physically for burning the fuels. On the other we are facing a dangerous future as a result of that burning. What is going on here?

There is a paradox regarding the human condition. The evidence suggests that we are too clever for our own good[3]. Is it possible that, as Shakespeare's Cassius claims "*The fault, dear Brutus, is not in our stars, But in ourselves, that we are underlings*[4]"? Is it possible that there is something within us that prevents us from being able to use our cleverness to recognize unintended consequences of our clever inventions, or of our various, supposedly intelligent behaviors? Are we "underlings" in the sense that we are incapable of making the right decisions in pursuing life[5]? If so, can we identify the "fault in ourselves"? Can we consider possible remedies?

---

[1] Hall & Klitgaard (2011).

[2] Klein (2014) has done an admirable job of documenting the fundamental mismatch between economics based on traditional (neoliberal) economics ideologies and the failures of governments to develop effective plans for reducing carbon emissions. Politicians (at least the conservative variety in the United States) find themselves between a rock and a hard place. No matter what they might actually 'believe' about the science of global warming and climate change, they cannot be seen to be in favor of anything that would be seen as counter-capitalism.

[3] Dilworth (2009) has developed a compelling thesis regarding human nature that parallels some of the ideas that I will be presenting in this book. His perspective is quite different – a philosophical form of argument – and I am not quite in accord on some details. Nevertheless it is a reasonably considered viewpoint.

[4] Shakespeare, Julius Caesar , Act 1, Scene 2.

[5] See Sternberg (2002) to get an idea of just how paradoxical our situation is.

In fact, psychologists and neurobiologists have started to look more closely at the cognitive capacity we generally recognize as *wisdom*. One motivation for doing so is that while there are distinguishing characteristics to the construct, its presence, or its strength in many, if not most individuals seems to be wanting[6]. People, even more mature people, too often make unwise decisions as evidenced by the consequences of those decisions. Now it seems society as a whole is having issues with respect to making wise decisions regarding the very future of our species.

There is a growing belief that wisdom, both on an individual basis and on a societal basis, is lacking when it comes to how we manage our affairs, particularly as they impact the rest of our world. We act hastily. We act without due consideration of consequences, especially those that arrive years from the present. We too often act with selfish materialistic motives rather than doing what is best for society. Indeed, in our current neoliberal, neo-classical model of economics, selfishness has become an institutionalized norm. There are even those who argue that selfishness is a moral good! Are these behaviors and beliefs in accord with our long tradition of viewing wisdom as taking careful consideration and moral unselfishness into our judgments? Have we become culturally foolish in the name of material wealth? Or, is there something in us, a "fault", a characteristic or failing, which causes us to act and think foolishly in spite of our singular cleverness? Are we destined to make unwise choices, especially as the world becomes increasingly complex?

That we even know that there is something that we call wisdom is the result of some individuals displaying uncommon good judgment in difficult morally complex social problem solving. There *are* wise people in this world[7]. It just seems that they are rare. Is there an explanation for this observation?

There may be. We need to explain the rarity. We need to explain why people as individuals and as collective social units all too often make unwise decisions. We need to explain why there are some people who others recognize as wise, but also why most of us are not in that category. Perhaps most of all we have to explain why those to whom we do attribute wisdom are so rarely attended.

The Cognitive (Brain) Basis for Wisdom

The study of wisdom as a psychological phenomenon is relatively recent. Psychology has long been interested in questions regarding the areas of intelligence, creativity, and affect as they relate to cognition. Over the last several decades these "constructs"[8] have also been studied by

---

[6] I will be presenting more evidence and analysis later for why this is the case. What we would generally recognize as wisdom (or rather, "wise-ness") in people is thought to be a rarely occurring quality. More finely, I will consider "degrees of wise-ness", or *levels* of wisdom, akin to levels of intelligence, in the chapters ahead.

[7] Goldberg (2005); Kitchener & Brenner (1990); Labouvie-Vief (1990)

[8] I will be explaining the concept of a psychological construct more in chapter 1. But basically the idea of a construct is a constellation of measurable attributes that collectively define a major psychological phenomenon. For example, the intelligence construct involves several aspects such as memory (speed of learning, capacity, etc.) that can be measured using various forms of intelligence tests. Of course there are many different definitions of what

neuroscientists or neuropsychologists using new technological methods for probing the brain as they attempt to connect the phenomena to brain activities to localize what is going on in the brain as people solve puzzles (intelligence), think about unicorns (creativity), or react to emotion-rich images (affect). Additional work has probed the brain basis of early perception and how that processing leads to abstract thinking (concepts and their manipulations in working memory).

More recently some psychologists have tackled wisdom as a construct[9]. While the concept of wisdom remains fuzzy at the boarders, there are some general agreements about various aspects of wise-ness that are broadly held across cultures and between scientists who attempt to categorize and measure its attributes.

What is clear from the accumulating literature on the subject, wisdom is a major cognitive facility that has a major impact on decision making. When it is lacking, bad decisions even from seemingly rational analysis might ensue. When it is present, superior decisions can lead to good life outcomes. That is where wisdom is most notable – in making decisions regarding complex social problems. The wise elders (chapter 5) give guidance to the young tribe members based on their accumulated tacit knowledge (wisdom).

This book is devoted to the exploration of the human mind from the perspective of systems science with special interest in these questions about how did we get into what appears to be very difficult situations resulting from making unwise decisions. Can the use of systems science principles applied to understanding the human condition provide some answers, or at least point in the direction of answers? The systems science approach may offer some insights that are not currently being considered in understanding human cognition and behaviors, insights that might help us resolve the paradox.

Human beings are *autonomous decision agents*[10] in a complex web of social interactions where the decisions taken by any one person ripple effects through the web. No man or woman is an island unto themselves[11]. What every individual decides determines their actions or at least their speech and consequently their impact on others. The magnitude of such impacts is in proportion to the cleverness exercised by the mind taking the decisions. For example the decision to burn carbon-based fuels in rotary engines to make machines go faster with greater power first

---

goes into these various constellations! There appears to be some general agreement on the major issues, but it gets fuzzy at the boundaries. In this book I attempt to approach the problem with a bit more definitive circumscription of the several constructs in order to keep the analysis tractable.

[9] See Sternberg (1990a), the introduction for a short history.

[10] I will be using the term ‘agent’ in various contexts in the book. The basic notion is that an agent makes decisions and takes actions that affect the future of not only themselves but others as well. Autonomy, as used throughout the book, will mean that the agent is not constrained to make decisions by a strict set of ‘rules’ as is the case for a computer running an algorithm. Brains, as we will see, are non-deterministic processors and no two brains will necessarily take the same exact decision given the exact same information due to path-dependent memories making each individual uniquely different.

[11] Paraphrasing John Donne, *Devotions upon Emergent Occasions*, 1624.

impacted individual human lives by producing greater wealth, but then impacted us all in dealing with the effects of carbon in the atmosphere.

Cognition, including conscious thinking, subconscious thinking, learning concepts and generating behaviors, and generally interacting with the world, is the job of the brain. Intelligence, creativity, and affective modulation of thinking are all mediated by various brain regions and modules (chapter 4). So too, the capacity for being wise must be mediated by specific brain areas/functions. The answers to questions about why wisdom is rare or in short supply in most people will come from a better understanding of the brain basis for the capacity for wise-ness. That is what I propose to develop by using systems science to understand sapience.

Toward Understanding Sapience

The extant human species is given the taxonomic name *Homo sapiens*[12]. The system of naming, called Binomial Nomenclature was developed by Carl Linnaeus (1707 – 1778) who also named many species including humans. The names chosen for species was meant to express the characteristic that differentiated them from cousin species in the same genus, or, failing to find a distinctive feature name, to specify either a locale (e.g. *Homo neanderthal* named for the Neander Valley in Germany where a sample skeleton was first discovered), or to honor a person (e.g. *Aptenodytes forsteri*, the emperor penguin is named after the German naturalist Johann Reinhold Forster[13]). The word 'sapiens' is generally translated as 'wise', so Linnaeus named our kind "man the wise.[14]" It is believed that he intended that interpretation given the dominant view in Christian Europe of human beings as having been created in the image of God and the highest beings on Earth (just short of angel status). He considered the great distinguishing feature of humanity as their capacity for wisdom. Throughout the rest of the animal kingdom no other animal shows a capacity for acquired wisdom; no animal seems to give wise advice to others. Humans seem to be able to do so.

But what is wisdom? Why do humans seem to be capable of it and other animals, even our closest primate cousins are not? Until very recently in the science of psychology the first question was not easily approached let alone answered. The concept of wisdom has always been in the domains of philosophy and religion. Since so much has been written about the idea of wisdom from these perspectives I do not propose to linger over their treatments. Accounts of

---

[12] Throughout the book I adopt the traditional italicization of the formal binomial nomenclature designation of an animal name with the genus spelled with uppercase first letter – like a proper noun – and the species in all lowercase. The formal name of the animal is given in Latin, or Latinized terms.

[13] See the Wikipedia article about the Emperor penguin: section on taxonomy where the naming by George Robert Gray, an English naturalist. http://en.wikipedia.org/wiki/Emperor_penguin

[14] Sapience derives from the Latin, *sapere*: to be wise.

how these disciplines have treated the subject are abundant[15]. Instead I will turn attention to the most modern perspective of wisdom held by psychology.

In this book I take the position that, indeed, as Linnaeus and the Greek philosophers, and religious people have asserted, human beings are quite different from all other animals[16] and that a brain function/capacity that we possess gives rise to a unique form of cognition that includes the potential for developing wisdom. I call this *sapience* in respect to the name Linnaeus chose for our species.

Sapience, as I will lay it out in these pages, is not quite the same as wisdom per se, but it is the basis for acquiring wisdom in a person's lifetime. There is a collection of brain functions that give rise to this capacity. Several of these functions pre-existed to various degrees in earlier mammals and primates especially. But at least one function seems to have emerged in later hominins and expanded greatly (dramatically) in our species so as to constitute a difference in kind and not just in magnitude. Its influence on the other functions in the collection was profound; it appears to have boosted the power of the others considerably to produce the array of cognitive features that clearly separate us from the animals. That new function was what I am calling *strategic thinking*. This is the ability to cogitate over the past and the current situation, to think about the future, and to formulate plans for actions to be taken in that future to improve the situation in that future. In fact, in systems science, strategic "management" is a recognized aspect of what is called the *hierarchical cybernetic system* at the heart of sustainable adaptive systems like living animals (Mobus & Kalton, 2014, chapter 9). Every living entity has within it a network of sensors and actuators along with a complex of decision processors that keep the whole entity in balance and reactively interactive with its environment so as to keep that internal balance, which is the basis of life itself. Strategic decisions are at the epitome of a hierarchical control system. Most living entities do not have a strategic decision processing capability built into their brains; this kind of processing takes a considerable investment in "hardware". The reason is that evolution is a process that "decides" what strategies any given species will use to live. All that most animals, for example, need is just some amount of tactical decision making in order to interact intelligently with their environment. One of the things I will argue that make human truly different from all other animals is that they have evolved to make strategic decisions. Their brains include some special circuitry that "computes" strategic decisions and this circuitry is the basis for sapience. It takes wisdom to have flexible, adaptable strategic plans. Thus strategic thinking acted as a kind of integrator with pre-existing functions and some resulting new functions that make *Homo sapiens* the unique species that it is.

---

[15] For a cursory example refer to Sternberg (1990). He provides some background on the concept's treatment prior to the advent of psychological investigations.

[16] The differences, especially the cognitive aspects, are explored quite well by Thomas Suddenforf (2013). The nature of human cognition, and how it differs from the rest of the animal kingdom, can be found in, Tomasello, M. (2014); Deacon, T. (1997); De Waal, F. (2005); Marcus, G. (2004); Donald, M. (1991, 2001); Gangestad, S.W. & Simpson, J.A. (eds. 2007); Mithen, S. (1996) to name just a few.

For example, strategic thinking coupled with moral sentiment-driven judgment and a capacity to think systemically created a mental space that produced the capacity for generative, recursive, symbolic language to dramatically boost communications between individuals, thus pushing our genus from mere eusociality to hyper-sociality. It produced the ability to weave stories, not only from the past but looking into the future. Every individual can construct models of entities and processes (systems) in the world as well as models of themselves (self-reflection). They can construct models of others (theory of mind). They can play "what-if" games with these models and propose hypotheses to be tested. In short, evolution produced a completely new kind of being, building upon cognitive capabilities already in development, by adding a new kind of cognition to the mix and coupling it with pre-existing abilities that completely changed what an animal could do. And the new ability had an impact on other major cognitive capacities - intelligence, creativity, and affect were impacted as well.

Intelligence is a capacity to solve problems using various forms of reasoning, a generally mechanical process. It depends on memory systems, their speed of encoding and ease of access/recall, and is particularly well developed in the mammals, some birds, and especially primates. Intelligence always depended on judgment to modulate its processes and make it more efficient. Creativity includes an ability to deal with novelty in the environment as well as generate novel behaviors. It is necessary to keep animals doing some form of exploring their environments in a search for new resources. And affect is the oldest motivational and responsive mode of interacting with the world. Affective cognition goes back to the beginning of brain evolution (even further by some accounts). All three of these psychical modalities have been the main mechanisms resulting from brain evolution. They are responsible for low-level operational control of the body and behavior (motor control), logistical optimization of resource usage, and tactical management of interactions with the environment over time.

In other words animal brains up through and including the hominins contained everything necessary to manage operationally, logistically, and tactically to fulfill the fundamental mandate of biology - stay alive and reproduce as much as possible. Staying alive required fitness of the phenotype. Reproducing was the payoff. The fittest produced the greater number of offspring, which then, due to inheriting fitness repeated the cycle. Evolution took care of the strategic management issues. Animals evolved to continue to be fit in ever changing environments, which included some genera giving rise to more complex brains for handling greater amounts of information. Their survival and thriving strategies were molded by the species' experiences in life.

But the human brain acquired a capacity for strategic thinking, that is, an ability to generate possible futures for individuals and not just reaction of a species to what the environment did. They acquired the ability to willfully change the environment to conform to their desires. And that, it seems to me, is the real meaning behind the myth of Prometheus giving fire to humans. That story was an attempt to explain (in a pre-scientific age) how it is that humans are so significantly different from the rest of the animal kingdom. We represent a breakthrough in evolutionary terms, in mental "technology." We are to evolutionary history what the computer revolution was to societal history. Things are different now.

I will make an attempt to explicate what, how, and even why this revolution in cognition emerged. However I should point out that though we may represent a breakthrough it does not mean that we are "better off." As I will shortly explain, in fact, we are, as a species, worse off in some important ways[17]. We may be the first species on Earth to acquire responsibility for our own strategic management, both as individuals and for the groups, but what we acquired seems to be too little sapience to do a good job of managing our affairs strategically. Therein lay the rub.

The Systems Science Approach to Understanding the Human Condition

The Systems Science Approach to Deep Understanding

In the sciences there are two basic ways to claim an understanding of a phenomenon. One can make repeated observations of a phenomenal behavior, carefully measuring relevant variables of both the environment and the phenomenal response to changes in it, and then analyze the data for patterns (ideally some kind of function). One may infer future behavior of the phenomenon given particular environments, and inputs to whatever system may be doing the behaving. One may also infer what kinds of internals are responsible for the observed behavior based on any previous understanding of similar phenomena. But this knowledge is shallow in that it does not account for any hidden variables that could under conditions not previously observed (or not deemed relevant) and which could cause a completely unexpected behavior in the presumptive system. Many of the sciences have been until recently primarily observational. This includes, for example, astronomy and astrophysics, but also the social sciences. The latter have been hampered by an inability to apply reductionist explanations to their subjects since ultimately much of the behavior of such systems is based on what happens in the human brain and that has been terra incognito since the beginnings of their scientific endeavors.

Contrast observational sciences with those that have found ways to deconstruct their subjects. Sciences like physics, chemistry, and somewhat more recently biology, use reductionist methods to delve more deeply into their subjects' inner workings. Physics and chemistry have profited from the relative simplicity of their subjects and the reduced relational aspects therein. For example, the atomic weight of hydrogen does not depend particularly on the interactions that atom might have with an oxygen atom. Its measurement can be accomplished rather straightforwardly assuming one has the right instrumentation.

---

[17] I am reminded of the story of Adam and Eve and the forbidden fruit from the tree of knowledge, both good and evil. There is a benefit to taking over our own strategic management, possibly, but there is also a cost. And there is also a risk - that we won't be very good at it. The evidence I will be covering seems to suggest our species will pay a high price for being the pioneer into this new level of organization.

In biology we start to have problems because elements of biological interest are much more related to one another and understanding living mechanisms requires a deeper understanding of how these relations work in the dynamic whole. It wasn't until biologists had access to theories of evolution and the structure of DNA (and its role in genetics, the role of RNA molecules, ribosomes, and the like) that they began to acquire much deeper understanding that could begin to provide the basis for predictions based on knowing how the mechanisms inside worked.

This is deep understanding. Biology, in particular, has made incredible strides in recent years by adopting this more holistic form of analysis, what I call "deep systems analysis"[18]. And, in fact, 'systems biology' is now a dominant pursuit in the life sciences.

The social sciences are starting to go down this path as well. Computer modeling of complex systems of agents and better grasps of neuropsychology and behavioral correlates has put many of the social sciences in a position to probe much deeper into areas, such as decision-making (see Chapter 2) that affect individual and group behaviors by looking inside (i.e. deconstructing) mental states. This book represents one such attempt.

We need to be deeply understanding the human social system, but we also need to understand deeply the Earth system in which the former is a subsystem.

The whole of planet Earth is a system. Everything is ultimately connected to everything else. The major problems that our world faces right now, such as global warming and climate change, sea level rise, ocean acidification, and related issues such as soil erosion, and drinking water deterioration, just to name a few, are systemic and can only be fully understood if we attack their understanding with systems thinking.

The subjects in this series of books will be highly multi-disciplinary. It is not possible to address, for example, climate change, without grasping its relations to energy usage and the economic system. The systems approach demands that connections to all causal inputs be examined and made part of the model. This first book will address a subject that is actually at the core of all others. The world is what it is because of human beings living their lives and making the decisions they have. Many people who study the deep history of Earth have taken to calling our current age the "Anthropocene" or age of human impact on the Earth systems. By being who and what we are, we are changing the Earth environment in ways that we now realize are problematic. Who and what we are is fundamentally a question of what is the human mind and how does it interact with the world.

Using the principles of systems science to understand complex systems can be summarized simply. It consists of two basic stages, analysis to find all the relevant "dots" and synthesis or

---

[18] In a forthcoming work I will be describing the methods and procedures for doing this kind of analysis based on the principles of systems science. In brief, the analysis proceeds in a top-down recursive fashion to deconstruct a system (reductionist-wise) yet maintain the inter-subsystem relations discovered as going deeper into the system-subsystem-sub-subsystem hierarchy.

connecting the dots and motivating the network to see how it works (through modeling). Of course the details are a bit more complex.

Complex systems are generally considered as identifiable entities that have some kind of boundary through which substances (including information) flow in and out. From the outside the system might appear to be an opaque object; we see it transform inputs into outputs but do not see how it does that. However, we can infer that the system as a whole has strong internal organization, functionality, and is sustainable over long periods of observation. That sustainability is owing to an internal structure of subsystems which work together and are maintained in relations that support the final function. Systems have behaviors, and by the fact that subsystems are also systems, those subsystems have the same characteristics as well. The whole is dynamic yet organized such that it remains a whole under varying conditions in its environment. The principles of systems science (just summarized quite succinctly) tell us how to go about analysis of systems and how to consider capturing their relational integrity, usually in the form of a dynamical model. That analytic method is both reductionist and holist at the same time. It delves into the internals of the opaque object but preserves knowledge of interactions between subsystems in a way that allows us to build functional models of the systems, thus demonstrating out deep understanding.

I will not try to replicate the details of the principles in this and subsequent books, but rather refer to them in the *Principles* book (Mobus & Kalton, 2015) by chapter and section. I will make an effort to make this book somewhat independent for those who are already familiar with systems science in general, however. In chapter 1 I will provide a quick review of the twelve principles with notes on their applicability to the subject of this book, the mind.

A Broad Systems View of the Human Condition

A systems analysis of humans, their behaviors, their psychologies, their societies, and their interactions with the global ecology suggests quickly that the nature of human decision making is the core issue. There is a distinction to be made between decisions that are merely clever – that is they solve an immediate problem – and those that are wisely considered. For example, technologies, like burning carbon-based fuels for power, have always proven to be double-edged swords. Wisdom might cause us to pause and ask: Just because we *can* do something doesn’t mean we *should* do that thing? Consider the case of nuclear fission. We can split atoms, but should we?

In this book we explore the nature of human decision making in the context of how our actions impact each other and our environment. It comes down to explaining why being merely clever is not enough to ensure that our actions are in accord with nature’s (the Ecos’) processes as evolved on this planet, or positively supportive of the Earth system. And it comes down to explaining why so many decisions being made by humans are ending up having a negative consequence on the planet.

Make no mistake. In extremely complex adaptive and evolvable systems[19] like the Ecos, subsystems have to play nice with the other subsystems or they will be selected against. The human social subsystem is not doing so at present for reasons that will be demonstrated in this book. The Ecos will respond, and indeed, climate change is exactly that kind of response. It will set up conditions to eliminate any subsystem that is not contributing its fair share to the good of the whole.

The highest level systems view of our condition is shown in figure P.1. The system of interest here is the whole world. The human species is now a single global population of about seven and a half billion people. Every individual being and every kind of aggregate of individuals constitute decision agents that affect the whole. The Earth system receives energy input from our sun, Sol and dissipates waste heat into space[20]. The flow of energy through the planetary system drives all of the biophysical work that results in the organization of complex systems on the planet, such as the biosphere[21].

Humans take resources, material and energy, from the Ecos and deposit waste materials back into it. They generate a fair amount of waste heat too, which adds to that which the planet radiates into space. The decisions made by human agents, in aggregate, greatly affect the balance of materials and energies that cycle through their subsystem. The human subsystem has grown to become an almost overwhelming aspect of the Earth system[22]. The implication of this is that those decisions will have a significant impact on the whole Earth system.

Before humans came on the scene in their present form and capacity for impact, the Earth had evolved a complex set of material and energy cycles that, with variations within limits, maintained an on-going basic balance. There were extremes that occurred from time to time. For example there were at least five major die-offs where a majority (or significant minority) of species went extinct in very short periods of geological time. Most people are familiar with the great dinosaur die-off that left the world to mammals and birds. Yet within those dramatic episodes, the Earth's biosphere continued to produce new life and new species took their places. This is evolution writ large.

---

[19] Complex adaptive systems (CAS) are able to adapt to a limited range of environmental parameter values to remain viable. Complex adaptive and evolvable systems (CAES) are also able to modify their internal workings to become adaptive to values outside of their normal ranges. Biological evolution achieves this through mutations in individuals in large populations that provide enhanced fitness in changing environments. Our brains achieve this through adopting new behaviors.

[20] Of course there is a contribution from geothermal energy that, for example, drives continental drift and volcanism.

[21] Morowitz (1968, 2004); Schneider & Sagan (2005).

[22] As an example of the impact of human population size, see this article, "Can Earth's Plants Keep up with Us?" by Stephanie Renfrow, *NASA Earth Observatory* Web site, at http://earthobservatory.nasa.gov/Features/HANPP/hanpp.php , accessed 3/04/2015.

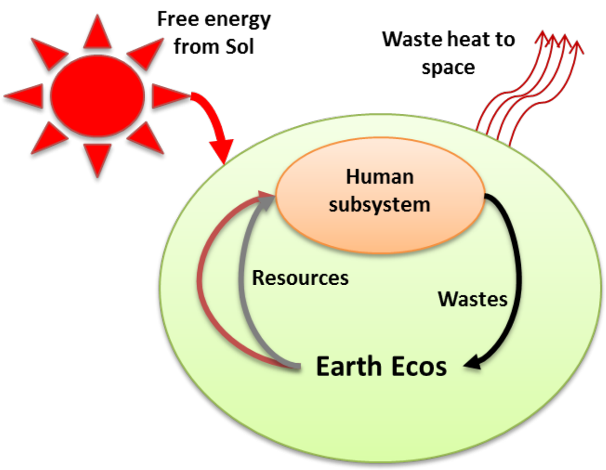


Fig. P.1. The Earth is a system in which the human species constitutes a significant subsystem. The planet receives high-grade energy from the sun, which is used to do biophysical work and ultimately is re-radiated to space as low-grade, waste heat. The system is effectively closed with respect to material flows, however it is extraordinarily complex with internal cycles (e.g. hydrological cycle) driven by the energy flows.

These die-offs had major impacts on the major material cycles, such as the carbon cycle and the hydrologic cycle. But the cycles continued and eventually achieved more steady state conditions. Life continued and recovered. Biodiversity expanded again.

The human subsystem (our species) started out small, a product of some of the fluctuations in the other cycles. As with any previous species they were the result of changes in the environment (in Africa). Their population was circumscribed by natural boundaries. But with the advent of strategic thinking that would not last long. An early human species, *Homo erectus*, broke out of Africa and migrated into Europe, the Middle East, and Asia. And they were uncommonly successful as species go. Later, *Homo sapiens* did the same, but with even greater success[23].

Any understanding of the human condition will start with a basic understanding of where the human subsystem fits into the whole Earth as a system.

The human subsystem of Earth is comprised of a huge number of smaller subsystems; societies and their cultures. Human beings are social animals, indeed we are *eusocial* or even *hyper-social*. The units of human interaction are groups, each tightly bound by strongly coupled interactions between individuals (figure P.2). Early human groups were small tribal units of fewer than 200 individuals. Modern humans participate in large social units that involve not only actual interpersonal relations, but also virtual relations that come into existence when strangers meet and recognize that one another have things in common. The largest units of human sociality are nations, states, ethnic networks and such.

---

[23] So much so that they probably played a major role in the extinction of all earlier species of the genus!

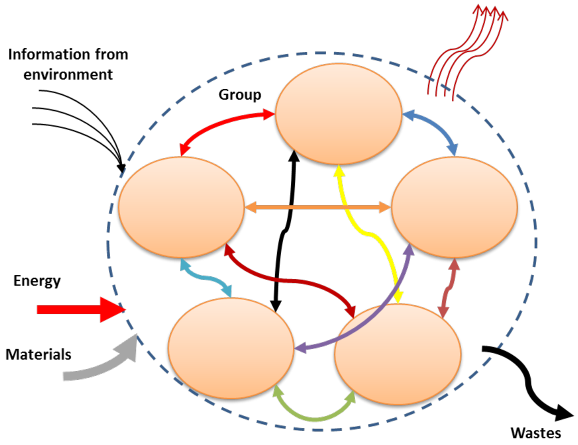


Fig. P.2. Human individuals are actually subsystems in a social group. All members of the group share common characteristics that constitute what it means to be a human being of the species *Homo sapiens*. But each individual also has unique personalities that give rise to somewhat varying interrelations with other members of the group (various colored bi-directional arrows). The group as a whole extracts energy and materials from the environment using information that they collectively receive from that environment. The group also produces waste heat and materials that need to be extruded from their community.

Individuals are effective units of decision making. The human brain has evolved to be the most autonomous decision making processor on the planet. The rest of the book is devoted to showing the veracity of this claim. But we cannot ignore that those decisions are conditioned and shaped by human sociality. Though each individual experiences their own personal perspective of the world and their relations with others, the power of social (and cultural) influences on those decision processes cannot be left out of the analysis[24].

Eusociality in species such as ants (or bees) is based on non-autonomy of the individual. But in humans eusociality exists in spite of the high degree of autonomy of the individual. Thus we are faced with an immediate complication when considering the system of interest in the case of human decision making. Is the unit of decision the individual or the social context? Or might it be both?

Fortunately the principles of systemness provide us with a way to answer this seeming conundrum. We can start our analysis with the individual as a decision maker (system of interest or SOI) and examine the condition of the individual that leads to “the fault in ourselves.” This will inevitably lead to unanswerable questions regarding the inputs to our decision processes. And that, recognized in the course of systems analysis, tells us that we need to expand the scope of the SOI to become the group (society and culture).

---

[24] Tomasello (2014) provides an excellent analysis of human cognition as being impossible except in terms of a social fabric.

In this book, I propose to start with the individual and touch on the larger scope as it impacts how individuals make their autonomous decisions. The subject of group decisions will be saved for a later work on the system of governance.

In figure P.3 we see a sketch of the SOI for the moment, the individual, one of the ovals in figure P.2. Actually our more particular SOI is the brain of the individual shown in the figure. It is necessary to make some preliminary observations about the individual as a system in order to then focus specifically on the brain/mind. Specifically we need to recognize that the brain is not an isolated system on its own. It depends completely on the body for its existence. Indeed an important function of the brain is to manage the body so that it successfully supports its brain[25]. However, for this book we will not dwell much on the functions of the body (e.g physiology) but will assume them as we focus in on the way the brain works, especially in support of higher cognitive functions, and most especially the function of sapience.

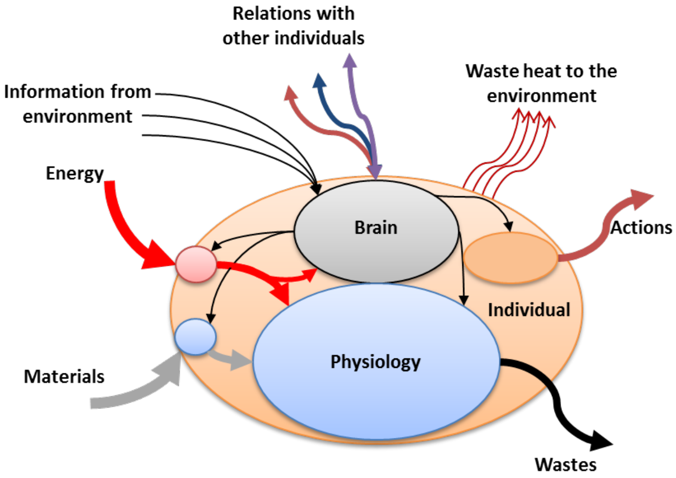


Fig. P.3. The human individual is a semi-independent agent as a component of the human social subsystem. Each makes autonomous decisions (what the brain does) that affect the flows of energy and materials and produce actions that impact others and the environment. Brains within a group are strongly coupled and contribute to the boundary of the group. The influx of energies and materials to the group have to be fairly distributed among all individuals in order for the group to function biologically.

The focal SOI for this work, then, is the human brain and its purpose/function in terms of systems principles. Figure P.4 brings us down to the scope desired.

---

[25] The works of Antonio Damasio (1994, 1999, 2010) make clear how interdependent the body and brain are. In evolutionary terms, early brains were mostly about managing the body, including managing its tactical relations with the animal environment. This relation will be explored more in chapter 4.

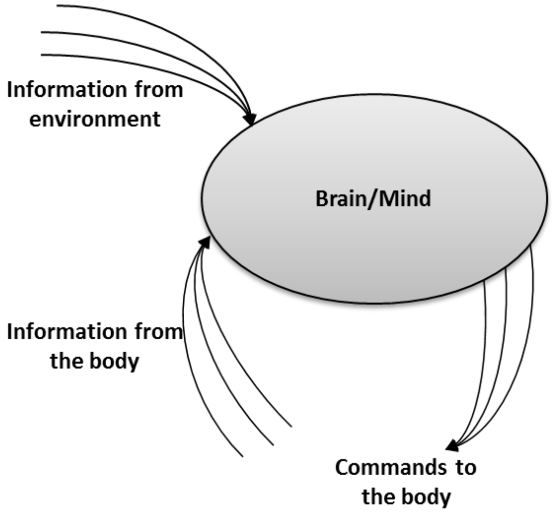


Fig. P.4. Here is a summary view of the system of interest with respect to the major flows of information into and out of the brain/mind.

The underlying thesis of this first book is to claim that human beings are, in fact, extremely intelligent and creative, but are we sufficiently wise to make veridical and strategic decisions that are for the good of all for the long haul? The evidence from the state of the world based on our historical behaviors underscores this basic question.

I have tackled this subject first because there are findings I will be reporting that are essential keys to understanding the larger issues and problems that face humanity. The nature of human cognition is central to understanding all other systems on this planet and so bears a deeper understanding in itself. For example, the next book in the pipeline will delve into the systems of governance for society. The problems that our various governments are experiencing are very much due to cognitive issues in decision-making agents (e.g. office holders, politicians, the public, and the like). There can be no grasp of the nature of governance (as it is practiced and as it could be accomplished) without a first grasp of how humans actually think and make decisions. This book will focus on human cognition as it pertains to making wise decisions in order to get a better grasp of what our social institutions require in terms of decision making nodes in those complex systems.

Toward a Better Understanding of the Human Mind

Human beings are a work in progress, both as individuals and as a species.

As individuals every person has some capacity to learn about the world, other people, and themselves; knowledge that they can use to interact effectively with their environment. Humans are able, at least under good conditions, of learning long into old age, but certainly for most of their lives. The brain remains highly malleable even as some of its functions seem to diminish with age. Many humans make individual progress in coming to understand the world they live in. However, they are limited by a number of factors, including their genetic endowment, to the extent of their knowledge and their ability to use it effectively. Moreover, we humans carry a substantial amount of cognitive baggage from our animal origins. Our brains are guided often by affective reactions to situations in the world rather than reasoning and knowledge. They have leftover circuitry that biases decision making that once served to make reactions more quickly in an often hostile environment, enhancing general fitness, but now preventing rational behaviors from correctly dealing with situations of our own making.

As a species our situation is less open. While learning and brain plasticity allow a form of mental evolution over an individual life span, the species is constrained by the slow process of biological evolution which operates on the whole population. The species is "learning" but only over very long time scales and only about those aspects of the environment which are changing in ways that require modifications in our genetic codes to remain fit.

And, for the species, the progress has not been as great as we might have hoped given what we now understand about our mental shortcomings. Our sciences show us that we need to think and behave in ways more befitting of a complex social and biophysical world. Yet our genes prevent the majority of us from doing so. Witness the irrational reaction of so many people to the scientific recognition of global warming and climate change who insisted that the scientists were pulling a hoax, or got their science wrong or any number of excuses just because they did not want to believe such a phenomenon possible.

The world that we humans have created demands that our species have brains that can interact with that world and make veridical decisions about what to do to have a sustainable culture and society. Our brains were evolving in that direction in the pre-agricultural, pre-historic world. Had there been adequate time before we discovered and began exploiting denser energy sources, such as fossil fuels, it is conceivable that we would have continued along that path toward more sapient minds. That is, we might have continued to evolve those cognitive capacities that lead individuals to acquire wisdom over their lifetimes. It is wisdom that is needed to manage the complexities of large-scale human societies and their interactions with the Ecos. Sadly, as we look around at the state of the world today what we see is the abhorrent lack of wisdom that contributes to our behavior as mankind proceeds to destroy the Ecos and consume every natural resource.

Even so it seems that the majority of humans take for granted that we are at the epitome of intelligence as a species and that no further improvement is possible or even necessary. The average person seems to accept the notion that we are what we are and that is just the way it is; angelic and devilish propensities in one package. Some might think we are inherently evil and that we have morals handed down from a higher power in order to "tame" us into behaving rightly toward one another. Others believe we are inherently good and kind and are just led astray by evil forces to end up treating others badly. Still others tend to think we are essentially neutral on the issue of good or evil and that most of our problems stem from the fact that we are ignorant and tend to make egregious errors that have the effect of evil. In any case few have asked if there might not be a human mind that could be the result of further evolution and that might result in individual behaviors that act to form more coherent social structures, especially structures that are in consonance with the Ecos such that our long term sustainability might be more assured.

There is some talk, in this era of genetic manipulation, of engineering higher intelligence in our offspring, but that is largely in the context of designer babies where prospective parents could afford the expensive medical procedures needed to accomplish that end. Surprisingly there is very little discussion about what “higher” intelligence actually means. Most people probably assume it means having more capacity to learn mathematics or become “rocket scientists” or be smart enough to earn high salaries; higher intelligence being roughly equated with a higher IQ, perhaps. What is missing is a larger discussion of what the mind really is and what role intelligence plays in what the mind does insofar as human behavior is concerned.

This book is an attempt to open up that conversation.

The perspective I want to take, as in my other books, is from the point of view of systems science. In our textbook, with co-author Michael Kalton, we explicate a set of principles that constitute the concept of systemness and that can be found operative in every kind of complex system. That the Universe appears to be organized as nested systems and subsystems really means that everything we observe is some kind of system. This includes human beings, species, and the human mind. In particular in this volume I want to explore the mind from that approach and point out some new ways of seeing what the mind is and how it operates. My central claim is that the traditional psychological and philosophical perspectives are necessary but insufficient for a better understanding of the human mind. Today we add the neurobiological perspective as well and the new insights we are gaining through all of these avenues, I will argue, show the systemness of the human mind.

Chapter 1 - The Concept of Sapience

Homo sapiens: A New (?) Kind of Animal

What makes the human species different from all other animals that have ever existed on Earth? As I stated in the Preface, we possess the beginnings of sapience and sapience is a breakthrough in brain evolution[26]. Sapience is what makes our species quite different from all others, including our nearest relatives, the great apes[27]. Scientifically we carry the name *Homo sapiens*, man the wise.

I have adopted the term 'sapience' to describe a constellation of attributes, capabilities, and capacities of cognition that are unique to our species. This constellation goes a long way toward explaining a number of phenomena in human experience that have both mystified and intrigued our greatest thinkers and scientists through the ages. The specifics of what is entailed also goes a long way to explain the human condition, particularly certain 'failings' of normal human cognition that have left us vulnerable to our own successes. And, finally, a careful reading of the subject suggests a surprising future for humanity, assuming we can take our limitations seriously and make a concerted effort to act wisely in the future.

The characterization of this constellation comes out of the application of systems science principles to analyze the human condition from a number of contributing disciplines, psychology, neurobiology, sociology, and paleoanthropology among many. Subsequently this is an attempt to synthesize findings from these disciplines by using the principles of systems science as an integrating framework. This framework has been explicated in my (with co-author Michael Kalton) 'guide book' *Principles of Systems Science*[28]. Throughout this book I will often reference the chapters and sections of that work where the principles being applied are explained in greater depth.

Using the approach of systems science some interesting results emerge. In some ways this work represents an original perspective on the human mind and society. The two of these cannot be understood separately but the whole conception would be beyond the possibility of covering in one volume. Therefore this is the first volume in a series of books in which I propose to investigate the human condition as an outcome of the evolution and workings of an extremely complex system. This book will focus primarily on human cognition since that, in the social context, is the main controlling force behind the whole human social system (HSS). In future work I will then situate humanity within its social system framework, and that within the whole Ecos. Indeed, the actual 'meta-system,' as explained in the Preface, is the whole Earth because,

---

[26] Harari, (2015) describes what he calls the "Cognitive Revolution" as the major transition from mere hominid apes, e.g. *Homo ergastor*, to the species that is capable of our "kind" of cognition, *Homo sapiens*.

[27] See Suddendorf (2013) for a treatment of what he calls the "gap" between humans and other species with respect to cognitive capabilities - both in kind and magnitude.

[28] Mobus & Kalton (2015). We prefer the term 'guide book' to textbook as it is designed to be used in general seminar and research-oriented courses in general systems science.

as things stand, the human species is a substantial subsystem within that larger system and the human condition is now, and will be, conditioned by what goes on in that larger system.

Evolution of the New Kind of Animal

The overarching principle in all of this is that of evolution. Human beings are the product of a long history of evolutionary processes and nothing can be understood about the human condition except in light of evolution[29] – in particular, the evolution of cognition and the brain structures that produce it.

Sapience, as I will be describing it in this work, is an emergent function arising from underlying brain structures/functions that were already evolving in the hominin line[30]. Traces of functions that would develop in capacity in humans, such as intra-specific communications, can be found in lower primates and even mammals in general. But in humans those functions have leapt to higher levels of performance and complexity. Human language, for example, is qualitatively different from the kinds of communications that other great apes have. Our ability to think about the future and the distant past is unknown among all other animals.

The emergence of sapience is the result of a process of auto-organization[31] among existing sub-functions (e.g. utterances for communications, existence of working memory, self-awareness, and others) that was selected by factors in a changing climate in Africa some 180 to 200 thousand years ago and by the increasing need for more cooperative social interactions with band and tribe members. The emergence may have created a situation that favored the expansion of a particular brain module, Brodmann area 10 (BA10), the patch of prefrontal cortex just behind the eyebrows (see chapter 4). This patch, in turn took on a brand new role in cognitive function, acting as a grand coordinator for all of those preexisting functions (essentially other patches within the prefrontal cortex). Once this new function emerged it was actively selected for by the success it gave humans in evolutionary terms; the success engendered by working in groups via a new kind of coordination capability[32]. Thereafter not only did the brain undergo refinements

---

[29] Borrowing Theodosius Dobzhansky's famous observation, see the Wikipedia article: https://en.wikipedia.org/wiki/Nothing_in_Biology_Makes_Sense_Except_in_the_Light_of_Evolution. Accessed, 4/6/2019.

[30] Mobus & Kalton (2014) explain the nature of emergence and how it is related to auto-organization in chapter 10 of the book. Note that what I refer to as 'auto-organization' is what is generally meant by the term 'self-organization.' In the Principles book we explain why we think the latter term is ill-advised and why we prefer the former term to describe a universal process.

[31] As explained in Mobus & Kalton (2014) the term auto-organization is preferred over the more commonly used term, 'self-organization.' The latter term, I have found in talking to students who are not technically sophisticated, implies some sort of mysterious spirit-like intentionality in otherwise inanimate systems. We may have mounted a losing battle in wanting to suggest a change to 'auto,' which is less fraught with mysterious baggage, but there it is. If the reader is more comfortable with the more established self-organization, then feel free to translate whenever I write about auto-organization.

[32] Tomasello (2014) identifies three stages of cognitive evolution for humans starting with our last common ancestor with the other great apes, proceeding through the major groups of Australopithecines and early Homo, and ending in our own species, modern humans. The three stages, to be further explained in later chapters, were 1) individual intentionality (looking out for #1, even in extended family groups), 2) joint intentionality (cooperating

through coevolution of all the sub-functions, but the kind of knowledge that could be constructed within the brain, specifically the neocortex, became essentially open-ended; any kind of knowledge could be constructed and modified with experiences[33]. Thus knowledge itself could evolve within an individual brain.

Nothing like this existed in the animal kingdom before. This was new.

Those of us who have believed we are ‘above’ the animals have some justification, it seems. But that doesn’t mean we are ‘better’ than the animals. We are just different. And it will be up to evolution to decide if we are, in fact, competent in the biological sense. Are we fit as a species?

That is where the real issue is. Sapience may be a great new feature of life in terms of providing cognitions that appear superior to those of other animals but is it making humans more fit in their environment and therefore more likely to survive as a species? Those who hold a religious view that humans are ‘better’ simply assume this is the case. But the judge and jury, nature and natural selection, have not yet determined the situation. Only time and circumstances will tell. With the realization that we humans, with the increased cognitive powers we express in, for example, producing new technology, are changing the very environment in which we will be selected comes an even larger question: Are we creating the conditions that will select against us in the long run?

Because sapience is a new feature it is also not a well-developed feature (or, alternatively, a strongly expressed feature). In evolution when new features appear they are not necessarily optimal in terms of contribution to fitness. They are opportunities not guarantees. Think, for example of the early lobe finned fish that were transitional to tetrapods[34]. Their fins still functioned as fins for swimming, but they also had a primitive, possibly clumsy capacity to be used for ambulatory-like propulsion on the sea floor. So it is with sapience. Our species is the first sapient species and, as I will be pointing out in these pages, our level of sapience, and therefore its capacity to increase our species fitness, is actually quite low. It is nascent. It hasn’t started “walking” and, in fact, appears barely able to “crawl”. The problem is that we are very clever and have invented a whole new world in which to exist. We have not, however, shown great wisdom in the choices we’ve made in creating that world. The environment we have produced (e.g. climate change and over-reliance on fossil fuels) may prove to be unfavorable to

---

with another for a specific purpose, e.g. hunting) and 3) collective intentionality (culturally imposed normative thinking for the good of the group). This last stage corresponds with Harari’s (2015) Cognitive Revolution and the major transition to sapient consciousness.

[33] The idea that knowledge is “constructed” in the mind comes from the school of constructivism (see this Wikipedia article for background: https://en.wikipedia.org/wiki/Constructivism_%28philosophy_of_education%29). In this book, especially in chapter 4, I provide a more holistic account of how knowledge comes to be encoded in the neural fabric of the brain, especially the neocortex. The process is one of constructing networks through reinforced and meaningful experiences. In other words, learning, in all its various forms, depends on how the brain builds specific networks of neural stimulations that come to represent things, actions, and relations in the world.

[34] For example see the Wikipedia article on Tetrapodomorpha: http://en.wikipedia.org/wiki/Tetrapodomorpha

our species in the long run. The human condition (sometimes characterized as a tragi-comedy) is that we started out ignorant but full of capabilities; we started out clever and inventive but unknowing of consequences. And we had just the bare start of sapience with which we might have moderated our eagerness to conquest. Now we have gained significant knowledge of how the world works but appear not to be able to moderate our eagerness to dominate even so. Is something not quite right with us?

A New Kind of Consciousness

There is general consensus among scientists and philosophers (and agreed by religious and lay persons alike) that human beings are conscious in a way that animals in general are not. The notion of consciousness has been explored for ages and that exploration is itself part of the human condition. We cognize our cognition and that demands attention (no actual pun intended).

With the advent of advanced methodologies in neuroscience the questions revolving around the nature of consciousness are beginning to be answered, though not always to everyone's satisfaction. I take the position *that what we call human consciousness is an explicable phenomenon that is co-extensive with the phenomenon of sapience* as covered in this book. Specifically, the form of consciousness that we experience is a result of the emergence of sapience and therefore must be explored concurrently with it. We are conscious in the way we are owing to being sapient to the extent we are.

Human consciousness is most often described as something like 'awareness of being aware,' or a kind of second-order awareness. Awareness itself is generally described as having the capacity to attend to environmental conditions (states of the world) as a prelude to making decisions to act accordingly and be successful in interacting with the world. All animals are aware. It is thought that only humans are aware of being aware, but developing evidence from cognitive science and, for example clever probing of animal consciousness, is calling that assumption into question. It now appears that many animals, and not just our closest evolutionary cousins the great apes, have cognitions that include some forms of being aware of their own awareness.

Consciousness of being conscious is what we call second-order consciousness. It is not, by indications from animal studies, an all-or-nothing phenomenon but has degrees, with the human version being the most extreme form. Regardless, it has become increasingly clear that the human form of consciousness isn't just second-order. As I will be arguing in this book, the capacities of sapience are actually the 'symptoms' of a higher-order consciousness, which I have chosen to call 'two-and-a-half' order consciousness ($2^{1/2}$-order). The reason for not calling it 3rd-order consciousness will become much clearer as I proceed through the concept of sapience. However, what stands between our kind of consciousness and a real 3rd-order form (presumably a higher order) turns out not to be an impossible hurdle. As I will argue, the problems that we humans currently face are precisely a result of not being quite conscious enough – not sapient enough – for current conditions. But the potential for transcending whatever seeming barrier there might be is high.

The subject of consciousness, particularly human consciousness and its relation to sapience will be covered in chapter 3 in much greater detail. For now it is important to recognize that sapience is the evolved phenomenon of human brain function and $2^{1/2}$-order consciousness is the resulting cognitive experience.

But it is quite a bit more than just that. Individual consciousness is a prerequisite for the human condition but there is an even larger phenomenon of consciousness for which to account. Humans are completely dependent upon, and absorbed into a group or social system. The consciousness of an individual is nothing except in relation to the consciousness of the group as a whole. Each individual is conscious in the human form because of being in a group, not just about the group (in the way a bison is aware of being surrounded by other bison). It turns out that $2^{1/2}$-order consciousness is a merging of individual consciousness with a group mentality (the physical manifestation of which is what we call culture, in the wide sense). This is no mysterious collective consciousness in some surreal or ethereal plain of existence. This is a simple aspect of sapience and the emergence of hyper-socialization as a potential new 'major transition' in system organization. The "major transitions in evolution" concept has become an incredibly insightful organizing theme in understanding phenomenon in biological evolution[35]. See chapters 10 and 11 in Mobus & Kalton (2014).

The Principles of Systems Science and Application to Understanding the Mind

In *Principles of Systems Science*, co-author Michael Kalton and I explicate a set of principles that apply to understanding all complex adaptive systems (CASs) and an expanded concept of complex adaptive *and evolvable* systems (CAESs). There are twelve principles (and a few sub-principles) that we cover, not meant to be an exhaustive list, but to show how all systems, regardless of kind, share certain fundamental properties that allow us to develop deeper understanding of specific systems by seeing how they embody those principles. So it is with the *mind*. The mind-brain can be studied as a system, applying the principles in order to grasp a better understanding of what that system does and how it works.

In this section I will outline the principles, providing brief summaries of what they are about and how they can be applied to the study of the mind-brain as a system. Throughout the rest of the book I will show how they apply in greater detail and what their application can tell us about that system. The main thesis of the book, however, is that there is an important aspect of cognition that has been overlooked until recently that is more fully explicated by the systems approach. Specifically, the brain's capacity to develop and use that background knowledge we call wisdom through higher order judgment is revealed through this approach. I call the collection of

---

[35] Smith & Szathmáry (1995) provide a grand perspective on how biological evolution has progressively integrated diverse, previously competitive systems into new cooperative systems through auto-organization and emergence. See also, Morowitz (2004), Bourke (2011), and Volk (2017) for additional perspectives on the evolution of complexity in systems.

components (subsystems) and their interactions with each other and the brain as a whole "sapience." Sapience is the system in the brain that gives humans the capacity to make good judgments in complex social problems and is the basis for human *eusociality*[36]. One finding that is suggested from this analysis is that humans have evolved sapience only recently in geological time (with the emergence of *Homo sapiens*, in fact) and as a consequence it is yet a weak facility. This is the basis of a further suggestion that this is the cause of many of the problems that the species faces today. However, further application of some of the principles suggests that sapience need not remain a weak facility. The last chapter will explore future evolutionary possibilities.

Principles of Systems Science[37]

These principles seemed, to my co-author and me, to capture the major attributes of systems, in particular those systems that are highly complex and have behavior that can adapt to changing environments. Such systems are called complex, adaptive systems (CAS) in the literature. In our work we expanded the notion of a CAS to include a distinction between regular adaptability and something called "evolvability" or the ability to evolve new structures and functions while maintaining an essential quality or characteristic or "mission". Systems like species, mammalian and avian brains, cities, and ecosystems are examples of systems that are both adaptive to short-term changes and evolvable to accommodate long-term changes. We call such systems complex, adaptive, evolvable systems (CAES). Note that the concepts associated with phenomena such as pre-life (pre-cell), biological, and social organizations occupy the CAS and CAES domains. Various living organisms and supra-organisms (societies) demonstrate all of these principles. The principles listed below apply to all such systems. Some apply universally to even simple systems.

The principles include:

1. *Systemness*: Bounded networks of relations among parts constitute a holistic unit. Systems *interact* with other systems, forming yet larger systems. The universe is composed of systems of systems.
2. Systems are *processes* organized in *structural and functional hierarchies.*
3. Systems are themselves, and can be represented abstractly as, *networks of relations* between components.
4. Systems are *dynamic* on multiple time scales.
5. Systems exhibit various kinds and levels of *complexity.*

---

[36] The term 'eusocial' ('true' sociality) is still somewhat controversial when applied to human beings. Its very definition is debated. E. O. Wilson and several others who studied the phenomenon in social insects like ants have asserted that the term can apply to human beings (and to the naked mole rats of Africa). See: Wilson (2013). Also see the Wikipedia article: https://en.wikipedia.org/wiki/Eusociality for background.

[37] This section will be found in many of my books as a brief background in the principles. For more extensive treatment of the principles the reader is always directed to the main guidebook, Mobus & Kalton (2015) *Principles of Systems Science*.

6. Systems *evolve* to accommodate long-term changes in their environments.
7. Systems encode *knowledge* and receive and send *information.*
8. Systems have *regulation* subsystems to achieve stability.
9. Systems contain *models* of other systems (e.g. simple built-in protocols for interaction with other systems and up to complex anticipatory models).
10. Sufficiently *complex, adaptive* systems can contain *models of themselves* (e.g., brains and mental models).
11. Systems can be *understood* (a corollary of #9) – Science.
12. Systems can be *improved* (a corollary of #6) – Engineering.

Systemness is a kind of overarching concept that establishes the framework for the other principles. Implicit in the idea that systems interact with one another are both the concepts of network organization at a given level of complexity and a hierarchical structuring, systems inside larger systems (also called meta-systems). Because the Universe is comprised of energy and matter in constant flux all systems are dynamical, constantly moving both, their internal components and they in relation to other systems.

This "big picture" view of systems entails a wonderful property that makes understanding systems (by human minds) possible, perhaps even necessary. Principle 11 concerns the way in which some CAES may evolve models of the other systems in the Universe (principle 9) as well as models of themselves (principle 10). These models are, as the name implies, systems of dynamic representations that can be "run" in fast forward to generate anticipatory scenarios about what might happen in the future. In the human mind artificial or "imaginary" models take on as much of a role in anticipation of possible states of the world as simple models based on actual historical experience.

Information, knowledge, computation, and regulatory (management) subsystems are all involved in how models get constructed and run. These, collectively, will be a major focus of this book as they relate to the mind, although it should be noted that the principles involved are not necessarily restricted to human minds. An interesting application of these principles to artificial agent models is discussed in my various works on artificial adaptive autonomous agents[38]. Figure 1.1 shows a conceptual mapping of these principles in order to see the functional relations between them.

---

[38] See, for example, my Adaptive Agents Laboratory page and particularly my work with the MAVRIC robot at: http://faculty.washington.edu/gmobus/AdaptiveAgents/ for background. Accessed 4/6/2019.

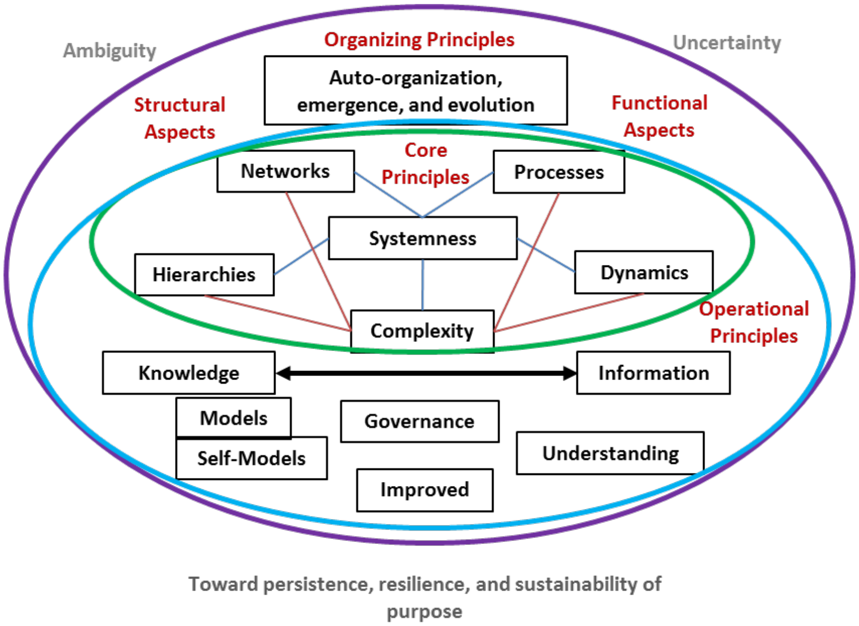


Fig. 1.1. This shows a rough map of relations between the principles. Structure (left side) refers to principles like networks, hierarchies, and complexity that are involved in how a system is organized. Function (right side) refers to the principles like process and dynamics that are involved in what systems do. See text for more explanation.

The outer purple oval contains the system of principles as a whole and represents the "Organizing Principles" of systems. The box, "Auto-organization, Emergence, and Evolution," at the top, is an overarching principle in the sense that it applies to all systems principles at all times. Parts of that overarching principle are issues of uncertainty and ambiguity, e.g. seemingly random mutations in genes. The light blue oval contains the main "operational" principles (toward the bottom) and the "core" principles. The core principles describe the main aspects of all systems whereas the operational principles apply variously and to different degrees to all systems.

"Systemness" is the center of the core principles is the integration of all of the core principles. It is named as a principle to provide a point of integration as discussed below. The blue connections between the other core principles and systemness reflect this centrality.

The core principles are grouped into two major categories, structure and function. Network structure refers to the fact that all systems are constructed of networks of subsystems and, in relation to the principle of knowledge, can be represented abstractly as networks of relations. Structures are also hierarchical in nature. Systems are composed of subsystems, which are, in turn, composed of still "smaller" and generally less complex, sub-subsystems. These hierarchies have "components" at the lowest level, but exactly what these components are is open to definition in order to avoid infinite regress, as described below.

Systems are fundamentally processes internally. They receive inputs from the environment upon which they "operate" and transform the inputs into outputs. This means that things are happening and that means the system is dynamic. But given the hierarchical structure of subsystems, the dynamics of each level of the hierarchy will be operating on different time scales.

Finally, the issue of complexity applies to both the structural and functional aspects. Hierarchies introduce one kind of complexity. Nonlinear dynamics introduces another kind. But the two kinds of complexity are interrelated. The interactions of dynamical processes on multiple levels of a hierarchy introduce many forms of nonlinearity.

The intermediate, light blue, oval includes the "operational" principles. These provide something like semantics to the interactions of the core principles. These are all interactional and describe important relations among the core principles. The information-knowledge duality (represented by the two-directional arrow) is a major interaction that connects the functional and structural principles. Mainly, a system, by its actions (behavior) produces messages that are potentially receivable by other systems. Those messages may contain information (news of difference) for the receiving system. This means the receiver did not have a priori knowledge that the message would be of a specific form, i.e. was ignorant of the state of the sending system that resulted in the message. The information value of a message can be "used" by the receiving system to alter its own knowledge (learning), which is embodied in the dissipative structure of the system. Messages that do not result in any actual change in the structure of a system are not informational. Thus the relation between both network and hierarchical structure can be changed by dynamics in another system resulting in changes in processes in the receiving system.

A more particular aspect of knowledge involves the inclusion of what can be called "models" of other systems. This comes into play mostly in very complex and adaptive systems (CAS) that learn and construct knowledge structures and processes. A model is a knowledge structure that is specifically processed with message inputs to extract information that guides the modification of the model as it learns.

Particularly sophisticated systems may contain models of themselves. The major benefit of this is that a system can simulate both the other systems in their environment and their own responses to changes in those other systems. Simulations of this kind allow a system to anticipate future situations and plan actions as a result.

Every system exists by virtue of a governance subsystem that can be as simple as the set of physical laws that govern the interactions (e.g. electromagnetic law governing atomic bonds) or more complex in the form of algorithms or heuristics that guide behaviors. Every subsystem has its own governance sub-subsystem. The latter involves the models just discussed. These models are processed with current information input to produce behavior outputs. The more efficacious the models the more successful a system will be existing within its environment. Effectively this means the system is fit (in the evolutionary sense) and will be persistent and sustainable within its environment[39].

A governance subsystem that produces successful behavior means that the system's models of the environment and itself constitute "understanding." The system understands its world, how it works and what it should do to continue into the future. More complex, adaptive, and evolvable systems (CAES) also understand possible weaknesses in their capabilities as well as threats and opportunities in the environment. This leads to the last operational principle which is that systems can be improved, or at least have an opportunity to become more fit. Some CAESs with intentional governance subsystems (explained below) have the capacity to guide their own improvement or improvement in constructed systems (e.g. machines). This is what we call engineering. In all other cases system improvement falls back on the evolution overarching process. For example a biological species (and populations) is a system in which random variations and natural selection do the job of "engineering".

The concept of "improvement" strictly means that a system is undergoing long term adaptation to environmental changes and by doing so is maintaining or improving fitness relative to that environment. Environments are forever changing (see the evolution section) and so the systems need to change their internal structures and functions in order to meet that change. Human-built artifacts undergo directed changes over time as engineers seek improved performance or even style design. These are tested in the environment of customer desire/satisfaction taking the place of natural selection (this process is the same as breeding new varieties of plants and animals).

The Brain-Mind as a System

I use the term brain-mind to connote a single system comprised of the physical system (brain) and its behavior as a system (mind). I will explore the sticky issue of consciousness in chapter 3 and beyond. Sapience and our human level of consciousness are co-extensive so that one cannot be treated without the other.

The Brain-Mind as a "Black Box"

No brain-mind is an isolated system. All brain-mind complexes are nodes in a far more complex network of relations, particularly with other mind-brain systems. I'll get to the larger social systems later. For now, to start the systems analysis we have to situate the brain first as a bounded system of interest (SOI). This is a somewhat artificial step in that we treat an SOI as being an independent entity called a "black box." That is, we first look at the SOI from the outside with no particular knowledge of what goes on inside. But what we do look at is the kinds and volumes of "things" that pass through the boundary, going in and coming out. Our intent is to characterize the system in terms of its function or purpose. That is we start by trying to discern

---

[39] See (Mobus, 2015, 2017). Available online: http://journals.isss.org/index.php/proceedings59th/article/view/2497 and https://onlinelibrary.wiley.com/doi/10.1002/sres.2482 respectively. Accessed 4/6/2019.

what the system is doing to produce the observed outputs given the observed inputs (the realm of psychology). Of course for anything as complex as a human brain we have some serious work to do (in the realm of neurobiology). Nevertheless, to treat the brain-mind as a system and attempt to understand it as a whole phenomenon requires we start from this perspective.

Figure 1.2 shows a more explicit systems map of the brain-mind situated within its environment, the rest of the body and the world external to that, compared with figure P.4 in the Preface. Every system is analyzed by first considering the inflows and outflows of matter, energy, and messages. The latter are special cases of low level matter/energy flows that are modulated in such a way as to encode the message. In the figures in this book, the flows of messages are represented by thin black arrows. In figure 1.2 I have omitted any material or energy flows which would include nutrients carrying both building materials for brain cells and energy with which to do the work of processing accomplished by those cells. Here I am mostly concerned with the informational aspects of the brain-mind system so the figure only shows the major aspects of messages into and out of the brain-mind.

The key point of a systems analysis is that the SOI must be treated as having a boundary through which inputs and outputs *flow*, passing through recognizable interfaces, i.e. through specific channels with some form of regulation. I start by observing the brain-mind as just such a black box system[40]. Shown in the figure above, the message flows into and out of the brain-mind from other body subsystems (open rectangular shapes representing un-modeled entities in the environment of the SOI). These other body subsystems are actually the interfaces between the brain and the external environment in that changes in the environment affect the subsystems but, as in the case of the skeletal muscles (voluntary movement), they can affect the environment leading to larger feedback loops. The black arrows through the Environment show the loop where the brain-mind causes muscles to produce actions in the body that, in turn, affect the environment. For example, the brain makes the hand pick up a glass and the glass's relation to the eyes and body change. In turn that causes the perceptual system to register the change and further affect the mental state (e.g. is the glass close enough to my mouth?)

---

[40] An unfortunate disjoint between the terminology, 'black box' and my preferred figural representation of a system as an oval, I hope, will be ignored. The term comes from the engineering tradition of treating a mechanism, regardless of its shape, as an unknown, as if an opaque box were constructed around it. It could be characterized only by the nature and quantitative aspects of inputs and outputs. I've retained the terminology since it is widely understood, but I persist in using grey ovals to represent systems for which we have no immediate knowledge of internal workings. My apologies for any confusion!

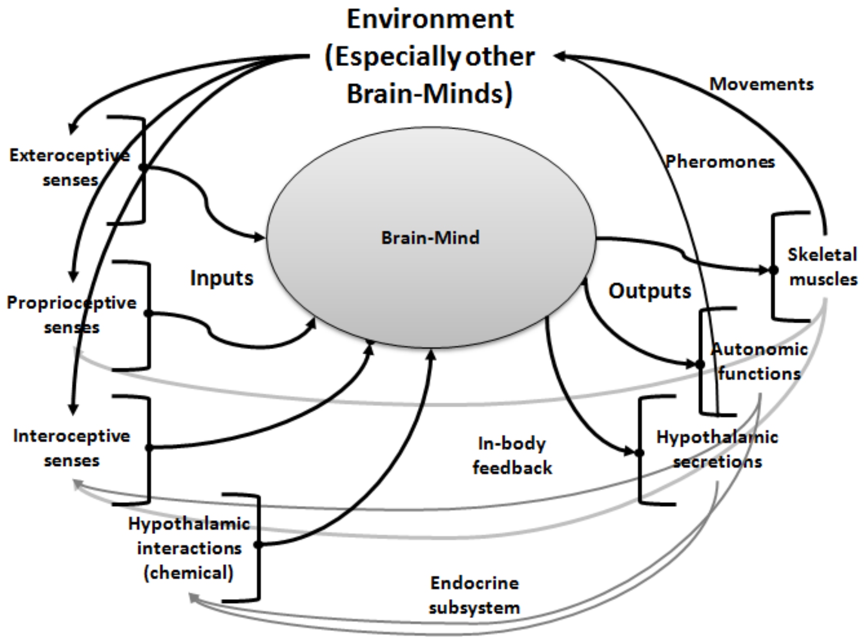


Fig. 1.2. The brain-mind is a subsystem of the whole person but can be analyzed as a system on its own via the principle of systemness. Here we see inputs and outputs to/from the brain-mind system. See text for explanations.

Exteroceptive senses are the "ordinary" ones, the standard five senses plus some extra support senses that bring messages in from the external world. Proprioceptive senses are those that sense the dynamic state of the skin, muscles, tendons, and such. They provide the brain-mind with information about the status of the body parts relative to one another, e.g. the angle of the elbows. Interoceptive senses provide information about the physiological state of the body, such as the extent of the stomach, the heart rate, etc. The hypothalamic interactions provide another physiological channel of information in which the chemical status (of the blood) is monitored. Oxygen levels, pH, and other such measurements are converted into neural impulses to be used in the brain-mind processing. But on the output side, the brain-mind can cause the hypothalamic subsystem to secrete chemicals into the blood to signal other parts of the body (viscera) to behave accordingly. The details of these channels are beyond the scope of this work so I will leave it at that.

On the output side, besides the chemical signaling from the hypothalamic subsystem, the autonomic nervous system and the skeletal muscle (voluntary) outputs are the major ones. The latter is part of the tactical management system (below is provided a summary of the hierarchical cybernetic model and explains the concept of strategic, tactical, logistical, and operational management as applied to the brain-mind system) whereas the former is part of the logistical management system. Both logistical and tactical management keep the organism as a whole functioning and persisting as an organized whole over extended time. I'll be getting back to this later.

A creature's (or human's) behaviors are the observable aspects of input and output. From measuring and modeling these behaviors it is possible to make inferences about what is inside the black box, the mechanisms that might be responsible for taking the inputs and processing them to produce the outputs.

The Brain-Mind as Information-Knowledge Processor

The brain is a kind of computer, just not the kind most people think about. That is, brains do computations to transform input message streams into output message streams that affect bodily functions and movements (including language acts). Computation is as applicable to what brains do as it is what digital (algorithmic) computers do. They do what they do differently, and indeed work best on different kinds of problems. But computation is a broader concept than most realize[41]. Fundamentally, computation involves data processing to extract information from incoming messages (see chapter 4, the section titled "Complexity in the Brain Reflects Complexity in the World" for a more complete definition of information and knowledge). That information is used to alter or modify the internal structure of the computing system itself. So, for example, in a digital computer the input data might be used to calculate some result which is then stored in the memory of the machine, perhaps to be output to a human or used in a later computation. In the brain-mind system information is extracted to cause responses to real-time situations and learning (knowledge construction in sufficiently complex brains) to occur.

Information drives decisions (chapter 2). Decisions need to be made at many different levels of complexity in brain-mind systems. Most of them are made under the radar of consciousness. For example, how quickly and deeply to breathe need not be decided at a conscious level, but the brain does make such decisions. At the highest levels of complexity (in the brain-mind system) decisions are brought to conscious attention for review and sometimes for modification or even veto. Ultimately the hierarchy of decision processing and knowledge construction that goes on in the brain-mind system results in output behaviors which can be observed and most often measured. This is the job of psychology and physiology sciences.

Information vs. Knowledge

There is a terrible confusion in almost every field that uses the concept of information in even the most technical ways. In Mobus & Kalton (2015), chapter 7, we go to some effort to try to clear up this confusion as it does cause something of a 'tower of Babble' effect in too many situations.

The problem stems from the indiscriminate use of the concept of information interchangeably with that of knowledge. All too often authors will talk about knowledge as if it is 'stored information.' But this is actually not the case. Information and knowledge are reciprocally related to one another but the nature of that relation is more subtle than seems to be realized.

---

[41] In chapter 8 of Mobus & Kalton (2014) we explicate a general theory of computation to show how digital and biological computations are accomplished.

Space here does not allow a full description of that relation so I will attempt a simple explanation and refer the reader to Mobus & Kalton, chapter 7 for more details.

The brain is an information processing subsystem, much as a computer is often used to process data to derive information. So the concepts of information and knowledge are critical aspects of understanding how the brain works and what it means for a person to learn something about the world.

All brains receive messages from the body and the outside world. Those messages are coded to measure ‘levels’ of a parameter, such as heart rate from the body or light intensity in a specific spot in the visual field. All neural cells have evolved to react to the messages based on their changes in the codes (e.g. changes in the frequency of firing of upstream sensors or cells – see chapter 4). Information is basically the property (a quantitative property) of a message that registers a difference in the code value from the *a priori* expected value. In other words, if a brain cell receiving a signal from a retina cell measuring light values has been receiving a value of, say, x number of lumens, and the next value received shows y lumens then the difference x-y is a measure of the information conveyed.

Most people think of information as being about something, and indeed the message is about something, namely in this example the amount of light at that spot on the retina. That spot is in relation to other spots and the collection of light values on all of them do in fact convey meaning. So changes in the values at different spots when correlated do mean something, possibly important. But the information value is just a comparison between something expected and something actually experienced vis-à-vis a message received in a particular ‘channel.’

The confusion between meaning and information value is quite deep in our common use of the words in vernacular language. Another way to think about what information actually is would be to consider it as a level of surprise that a receiver gets from receiving a particular message (that is about something important). For example suppose you are a sports fan who follows the scores of games (for your own edification). Suppose you missed the report of the outcome of a game. Also suppose you expected team A to win because you ‘thought’ they were the better team. So you have an a priori expectation. Then a friend tells you team B won by a score. You are surprised! In your mind you may start wondering if maybe team B is better than you thought.

You were just *informed*. See the Mobus & Kalton (2015, Chapter 6) for more on the nature of information as it plays a role in a cybernetic system (ibid, Chapter 8).

The Brain-Mind as an Evolvable System

Evolvability is a special characteristic of highly complex adaptive systems (CASs)[42]. The example of an evolvable system that most people have encountered is the Darwinian species or

---

[42] Evolvability is explained in chapter 11 of Mobus & Kalton (2014). See the Wikipedia article: https://en.wikipedia.org/wiki/Evolvability for background. Accessed 4/6/2019.

genera of living systems (though an argument for population level evolution or micro-evolution can be made). The evolution of a system is a form of longer term adaptation in which the system functions and structures may, themselves, be altered in order to accommodate a quite different set of physical aspects of an environment. In ordinary Darwinian (biological) evolution the changes happen in apparent random fashion (genetic mutations). In the evolution of, say, a business enterprise changes in things like procedures appear to be intentional, but on closer inspection are often seen to be trial-and-error explorations of 'what might work.' The altered system, however, retains its essential nature. Evolution allows complex systems to sustain themselves over very long periods of time while their environments themselves evolve. For example, every animal is genetically predisposed to accommodate a certain range of temperatures by physiological adaptation to whatever actual temperature within that range they find themselves. However if they are exposed to temperatures beyond that range the stresses will overpower their ability to adapt and the animal will succumb. Now, if there is a gradual change in the range of temperatures itself over many generations of the species, then it is possible for the species to adapt by way of natural selection for individuals with whatever genetic mutations allow them to accommodate the changed range. Individuals, in this sense, do not evolve; only the species (or genus) evolves.

Another example of an evolvable system is the enterprise organization such as a corporation. These entities are sometimes referred to as supra-biological (and even supra-psychological) systems in that they are built, as it were, on top of biological systems, namely us humans. A corporation maintains a certain fundamental quality (e.g. it generates equity for shareholders) even while it develops new products or services, new production methods, and so on. It is constantly evolving its internal structures to accommodate the changing external market environment.

Learning is an adaptive process (see chapter 4 for an explanation of learning in neural networks as adaptivity). Neurons in the brain have a normal range of adaptive capacity for encoding engrams (activation traces through the network) in synapses. However, in cortical structures, and especially in the neocortex of higher mammals, the ability to encode almost any combination of traces produces a capacity for both knowledge construction and information extraction that resembles evolutionary processes more than mere adaptivity[43].

The Brain-Mind as a Tightly Coupled Nexus in a Social System Network

We now come to what is a special consideration of the human brain-mind system that is highly relevant to the concept of sapience. That is the way in which every brain-mind system is actually a node in what, in systems terminology, we call a tightly coupled network of similar nodes. It is,

---

[43] In fact a very compelling theory of learning was proposed by Gerald Edelman (1987) called "Neural Darwinism." He posited that multiple concepts and thoughts are generated that need to compete with one another for neural hardware space. Those most "fit" survived in memory.

in fact, a nexus of connections to other similar nodes. I speak, of course, of the eusocial character of human beings.

In the next section I will address a society as a system. Here I want to call attention to what it is about the brain-mind system that makes the social system possible and functional.

One of the principles we discuss in Mobus & Kalton (2014) is number 9 above, that a sufficiently complex, adaptive system can contain models of other systems. In the current context this means that a human brain-mind can contain models of other human brain-minds; what is called the "theory of mind" in the psychology literature[44]. Sapience, as we will see, involves a capacity for the brain to build models of other minds and to connect to those minds through language communications, body language, and empathy. This is the essence of a tight coupling that acts as a force of attraction between human brain-minds. I've tried to represent this relation in figure 1.3 below, a recasting of figure P.2 in the Preface. Each oval represents a different brain-mind system that is automatically capable of making connections with other brain-mind systems and, in doing so, constructs knowledge (models) of those other systems.

What is the purpose of this network of nexii? As I will be arguing shortly, it is the social system that actually constitutes the human condition. Sapience is the capacity that allows human beings to form such complex networks that have a higher organization purpose than is realized by any one brain-mind by itself – the basic notion of emergence again.

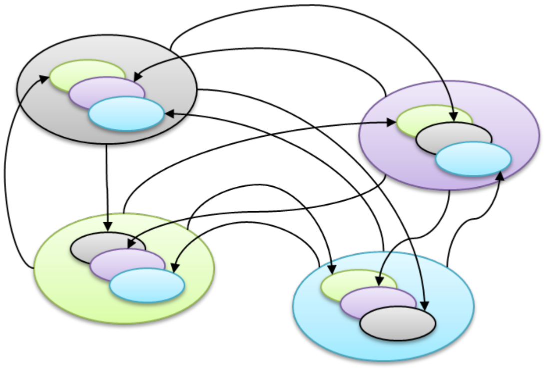


Fig. 1.3. Human brain-minds are able to construct internal models of other brain-minds. These models are the basis for connections between brain-minds that give rise to a social organization that is the basis for evolutionary success of the human species.

The brain-mind interconnections illustrated in Fig. 1.2 does not do justice to the effects of such coupling. For example, each of the larger brain-mind ovals also contains models of the world, each a unique construction based on the particular experiences of each individual. Not only do individuals have their connecting models of others in the society, they also have the capacity to share their models of the world through the facility of language. Each individual uses their model

---

[44] There is an extensive literature on this concept, "theory of mind." See: Cacioppo, et al. (2006); Dennett (1991); DeSalle & Tattersall (2012); Donald (1991); Mithen (1996); Pinker (1997); Suddendorf (2013); and Tomasello (2014) for various treatments.

of the other particular person to shape their communications based on what they think the other can understand (well ideally anyway). As always there is no guarantee that either the individual's model of the world or their model of the other is sufficiently veridical to make the communication successful. Figure 1.4 shows a diagrammatic version of interpersonal communications via language.

The brain's ability to construct models of the world and other people is the basis for eusocial interactions. The strength of sapience in mutually communicating parties determines how veridical the communications acts are and the ultimate success of forming a social unit that is successful as a system. Indeed, it is the weakness in sapience in modern humans that is at the root of mutual misunderstanding. Models of the world, which are as influenced by social embedding (e.g. ideologies and beliefs) as by evidence from the real world are wildly different between individuals and, I claim, increasingly non-veridical as a result of lower levels of sapience becoming the population norms. I will have much more to say on that aspect in the last chapter.

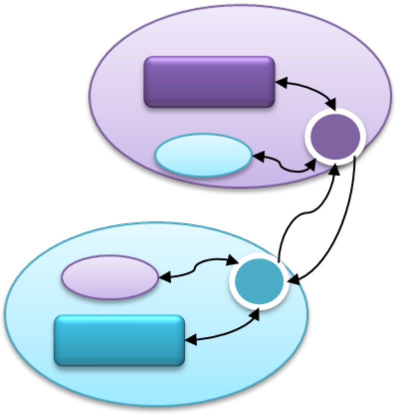


Fig 1.4. Every individual has a model of the world (knowledge) which they share with one another. The language facility (white bordered circles) allow the transfer of messages that convey the speaker's model of the world (rounded rectangles) with another person based on the speaker's model of that other person (colored ovals).

Application of the Principles and the Emergence of the Sapience Concept

The Motivating Question

I introduced what was for me the motivating question in the preface. "*If human beings are so smart, why do we find ourselves in this predicament*?" By predicament I mean the fact that many of our choices in how we pursued our lifestyles have led us to the possible brink of extinction, or at the very least portend physically difficult times ahead[45]. We are certainly causing major extinctions in many species and lowering biodiversity radically by having chosen to expand so aggressively throughout the planet. But more of a predicament, and a deeper aspect of the question, is that we seem to be unable to change our basic course even in light of substantial scientific knowledge about the consequences. That goes beyond stupidity.

Over the course of history many various intellectual endeavors have attempted to analyze the human condition and make sense of the what, how, and why of our insistence on what appears often as self-destruction. Philosophy, political science, economics, psychology, and now brain science have all puzzled over the diverse and often fuzzy patterns of human behavior that gives rise to people making decisions that, in the short-term, seem best for the individual or family group, but can be detrimental to the larger society and then the individuals themselves in the long run. Take the case of global warming and subsequent climate chaos due to individuals making decisions about what kind of vehicle they needed to drive and how far and when they would drive it (such as deciding to live in the suburbs and work in the city resulting in a long commute). Each person apparently can only see the scope of their immediate situation and they seek to maximize their utility as they see it. They also make familial decisions about child bearing (how many) that impacts the growth rate of the population. Collectively all of these individual self-serving decisions integrate into one huge dump of carbon dioxide and other pollutants into the atmosphere. No one individual can see that their contribution amounts to much, and yet the effect of the masses consuming all of that fossil fuel is literally threatening our very existence as a species.

But here is the real puzzler. Our cleverness produced the knowledge garnering process of science. We accumulate knowledge of how things work in this world. And among those bits of knowledge we have figured out the impact of burning fossil fuels on the environment. We have realized the impacts and consequences of continuing to do so. And yet, as of this writing, the response has been nearly negligible compared with what we would need to do to actually prevent the worst possible scenarios from unfolding. How can this be? I suppose it is not any different from our inability to avoid purified sugars or excess salt. Most of us know those things are potentially harmful to our health and yet we succumb to temptation all too often.

Why is this so?

I propose that by applying the principles of systems science to the various levels of organization of the human condition, Earth ecology, species, social, brain-mind, and biological, we might better understand the phenomenon of humanity and perhaps be able to formulate an answer to this question. It is not feasible to tackle all of these simultaneously in great detail in this book even though one of the principles of systems science tells us that systems must be understood as

---

[45] For just a sampling of the existential threats that humanity faces owing to past and current unwise decisions see: Klein (2014);

wholes. So my general approach is to start with what seems to me to be the core, the mind-brain, and work out from there. Even so, the mind-brain cannot be understood in isolation from either the biology (a lower level of organization) or the social (the group). So there will be numerous "hooks" to these throughout the book. For example the neuroscience in chapter 4 is deeply rooted in biology just as the psychology in chapters 2 and 3 are deeply rooted in the social. These links will not be ignored but also not explained to the degree the reader might wish. Systems biology is already a well-established discipline that can address the former area. Some fields approaching "systems psychology" or "systems sociology" are still nascent and are not as clearly organized as is systems biology. I plan to address this in future work. For example the next book planned is on the application of hierarchical cybernetics to the structure and functions of social governance (for large populations of humans). This project will end up relying heavily on this current book since the decisions in a governance organization are made by human agents and their efficacy depends on human cognitive capabilities.

If there are 'flaws' in human cognition then a careful analysis of the brain-mind system should tell us what they are and how they are affecting our interactions with the Ecos. Let me start by outlining three perspectives in order to orient the further discussion.

Species as a System

The concept of a species is actually quite messy in biology. We have a rough idea of what we mean by a species; the differences between cats and dogs are clear enough. Operationally the definition of species involves reproductive isolation between populations that have anatomical and/or behavioral differences. It may be physically possible to cross breed between two groups and even produce viable offspring (hybrids) but left to their own breeding preferences the two groups will sort themselves into isolated populations.

There are several different mechanisms by which this separation and subsequent divergence can take place. The most easily understood is physical isolation, as when some subpopulation moves to a remote region and is later cut off from the parent population. After many generations of responding to local climate and ecological changes, then, each population diverges in anatomy and/or behavior. This is called allopatric speciation. However it is also possible that two co-extensive populations may emerge simply from changes in behaviors that favor assortative mating (birds of a feather) and subsequent divergence. This is called sympatric speciation.

Species can be looked at as subsystems within a larger meta-system called a *genus*, at least from a genetics viewpoint. They can also be seen as components in one or more ecosystems. Insofar as the application of the principles it should be relatively easy to see how all, but #10 (containing models of self) pertain. For the tenth principle we must invoke the fact that a biological species contains within its own DNA a model of itself. That is, its genetic and epigenetic endowments contain all of the information for the construction of new members of the species.

Individual as a System

Perhaps seeing an individual organism as a system following all of the above principles is the easiest. Or, at least we thought it would have been until we've begun to understand the many mutualisms that allow multiple species to co-exist in what appears to be a single organism. The human gut, we now know, is the home for many varieties of bacteria, some of which assist with aspects of our digestion and physiology. Nevertheless, if we take this into account we can still treat an individual as a system, extending the notion of component subsystems to possibly many other kinds of organisms tightly bound to the nominal host.

When we get to complex animals with brains then the benefits of using the systems approach becomes more apparent. The diagram in figure 1.2 hints at this. If we wrapped all of the external sources and sinks in that figure into a larger system (the whole body and brain-mind) and then represented all of the sensory inputs coming from environmental entities and outputs (motions, sounds, excrements, etc.) going into the environment to various sinks then we would be able to analyze the individual as a system. In fact that is what much of biology and medicine are about. Field biology, in particular, seeks to situate an animal in its natural habitat and record all of its inputs and outputs, its dynamical relation with its environment, its ability to adapt to changes, etc. In reality all of biology is aligned with systems science[46].

The Social Group as a System

The emergence of sapience has its origins in the nature of social groups. Among primates (and other mammal and bird species) there is a biological need to form and maintain groups of related or near-related individuals. There are a number of reasons from an evolutionary fitness point of view why groups are favored (see chapter 5 for the evolutionary story of human group formation and group selection). For humans the principal reason is economics[47]. Human individuals are highly variable in their various talents and capabilities. As a group they can specialize in what they are good at and because they are in a group where other individuals are good at other work, they can cooperate to produce more life support than they could if every individual had to do everything on their own. Throughout the evolutionary history of the genus *Homo* and especially in the evolution of the more modern species, e.g. *sapiens*, groups that cooperated for the good of all did better, meaning had greater reproductive success, than groups that tended to be less cooperative within the group[48].

---

[46] Indeed, the main ideas of systems science came from biologists or those aligned with biology who recognized that living organisms and other units, like species and societies, were systems.

[47] Actually the 'economics' argument is true for all of biology. If we recognize that energy flow is the true currency of all systems then all systems solve income, allocation, and investment problems in the same framework as we see in the study of economics in human society. In the next book on governance I will be presenting the argument that human economics is really no more than an extension of the more general biology economics model. See Odum (2007) for the basic theory.

[48] Tattersall (2012); Wilson & Wilson (2008).

And what caused individuals within a group to cooperate more than compete with one another turns out to be what I am calling sapience, or at least a precursor mental capability. As I will argue in this chapter sapience in individuals is the glue that binds them together into groups that act as whole systems, optimizing their behavior as a group to achieve greater species fitness. In its more advanced form, sapience is the basis of individual capacity for wisdom, or having tacit knowledge that enables them to make wise decisions.

The Systems Approach in Practice

In the Wild State

The brain-mind of an individual person is a system. It is embedded within a biological organism - its immediate environment. The body and brain-mind are embedded, in turn, within a larger environment composed most directly of other human beings. The social organization is, itself embedded within the larger natural environment[49]. Each individual within the social organization has a unique identity and personality that governs the kinds of interactions each can have with the other individuals. Those interactions form a dense network of relations that are realized in the forms of verbal, body, and emotional communications.

The boundary of a social organization is very porous. Each individual interacts directly with the embedding environment. Each is an interface with that environment. Each processes material, energy, and message inputs from the environment and produces effects that impact the environment as well as wastes that must be absorbed by the environment.

The brain-mind governs the body and behaviors of the person to accomplish these interactions. It mediates internal operations and logistical coordination. It manages the tactical decisions that result in overt behaviors and coordination with the body's environment, especially the other people in the social unit. And it manages the strategic processing that results in building knowledge of all with which it interacts so as to have and use models of the rest of the known world to guide the tactical management in the long run.

The brain-mind can improve. As the person gains experiences interacting with the rest of the world it learns. It constructs and improves models of how the world works. It corrects earlier mistakes as they become obvious.

In all that it does it is guided by whatever level of sapience it possesses. It has some capacity to gain and use wisdom; some more than others, perhaps.

---

[49] In the views put forth here I claim that the human built world is, itself, part of the natural world. That is, we humans are natural phenomena and what we do is every bit a part of nature (as opposed to "supernatural"). Thus the so-called human-built world is really just the natural environment, as much as a bird's nest or an ant hill.

In the Domesticated State

We no longer live in small tribes. Our social groups are varied in composition and scale. We behave by following usually very complex sets of rules that pertain to what is appropriate in each social setting. What we learn tends to be more what other people tell us than what we experience directly ourselves. So our models are at the mercy of what we are told. What we believe is what our society tells us is truth. Very few of us every really experience the world directly. Some of us don't even experience a small part of it without some part of our social matrix mediating our interpretations[50].

The brain-mind is still a system, a sub-system of a body, that a sub-system of a social system, and so on. But now it thinks differently than did our Paleolithic predecessors'. The modern brain-mind is literally insulated from the "real" world by our culture and the way in which we educate our young brain-minds.

Humans have not had to depend on developing strong sapience for many millennia now. Our societal mind has taken over and its purpose does not reflect the same kind of wisdom that our species would have developed in the wild. Our models now are about making more material stuff and having more pleasure and having more convenience. We could seemingly care less about the rest of the Ecos except as it can be exploited for those ends. And we could seemingly care less about the long run.

Humans as Rational Agents

One of the most common conceptions of human mental capability is that a person can be a rational agent or decision maker. Chapter 3 will examine the nature of mental decision making in depth and chapter 4 will cover the nature of decision processing in the brain. Here I would like to briefly cover a systems view of rationality and situate human beings within a typological framework of rational processes. When I say humans *can* be rational I need to be careful because there are many kinds of rationality and there are different ways in which humans can be rational.

To illustrate consider the case of a thief who gets caught, tried and convicted, and ends up in jail. To most people the choices one makes to go into the thievery business seem irrational since getting caught and going to jail are hardly outcomes one would want. But what if you found out that the thief was a father of four who had lost his job and couldn't find work? Suppose he had pursued every reasonable avenue to be gainfully and honestly employed and had ended up desperate. Does his decision look so irrational now?

What is one person's rational is another person's stupid. We generally equate rationality with making good decisions and we most often compare rational thought with something like making choices that lead to optimal outcomes. The school of neo-classical economics went so far as to

---

[50] Tomasello (2014) tells a compelling story about the uniqueness of human thinking being the result of the evolution of what he calls "collective intentionality." I will have much more to say about this in chapter 5.

declare that when it comes to economic decisions humans are ‘utility maximizers’; their decisions always play to their own self interests. They created the vision of man as *Homo economicus*, a completely rational agent who would always make the right choices for their own good. As I pointed out earlier in this chapter, we now have a much clearer psychological view of human thinking and decision making that shows decisively how non-rational we are. But that only means we are non-rational in the sense of maximizing utility per the economic model.

Herbert Simon (1998) made a start on re-defining rationality in human decision makers when he realized that people don’t actually maximize utility (or profit) in real life. Real life decisions are always far more complex than can be handled by simple optimization computations, for example, and there are generally a large range of OK outcomes that are good enough to satisfy the general objectives (especially in the case of biological objectives). He called this process ‘satisficing,’ indicating that the decisions are made in constrained time and without complete information, but as long as they satisfy the objective function, that is provide a satisfactory solution even if it is not the absolute optimum, then that is good enough[51].

The problem with understanding what is rational and what is not, in terms of behavior, is that it depends entirely on the environment (context) and the objectives (mission) and what we call values. The term ‘rational’ is related to ‘reason,’ as in reasoning soundly. The ultimate in rationality we find in mathematics and formal logics. The objective of reasoning is to construct an object that has a truth ‘value’, such as a new, more complex mathematical object that is validated through a theorem and its proof. The atoms of composition are axioms and these are accompanied by a set of internally consistent rules for combining (or using) axioms to construct the object. The reasoning process is called deduction and it is the most veridical form in that if followed consistently (and carefully) one is guaranteed that the outcome reflects properties asserted for the object.

Somewhat less ‘rigid’ and therefore prone to occasional failures is the realm of physical world architecture, such as the design and construction of buildings, bridges, and computers. Architecture is based on the possession of some component atoms, namely the ‘building blocks’ used in constructions. Each of these has properties that include how it can be interfaced with other components. The rules of combining these components are based on the physics of ‘what can be done.’ Architecture is less rigid and somewhat less guaranteed in outcome because both the components and the rules of combination can be altered when a new objective or new value is asserted. For example, a building can be designed with much larger window apertures than are customary by creating larger header beams. The larger windows may be required because there is more sunlight available, or because the owner wants to take advantage of a grand view. The environment plays a large role in shaping the objectives and values that dictate alterations in the components and rules. In math and logic, in a sense, there is no real environment. Math deals

---

[51] Simon (1998). See chapter 2, page 28, for the description of satisficing.

strictly with symbols that are not really grounded in meaning coming from the environment. Thus there is no change in values (all that is wanted is the truth) or objective.

Reasoning in both of these realms involves the careful application of the rules to the given components to construct a more complex object (a theorem or a building). What makes this even possible is the relatively non-complex system of components and rules. In biological agents, however, the gloves come off and anything can happen. Indeed, everything will be tried, tested in the crucible of natural selection, and the failures will be disposed of, the successes will get to keep making decisions. It is almost like theorem-proving by trial and error, except there is no final “it is proven” endpoint. The reason is that the natural environment is forever changing and genera have to continue evolving in order to continue to survive. Evolution makes most of the decisions through trial-and-error. But as brains evolved, and especially with the advent of the neocortex in mammals, much decision making function passed to individuals and we enter the realm of psychological reasoning.

Components are forever emerging and morphing in response to new environmental forces. The rules of combination are forever changing, just slowly enough to provide some stability to the system, but relentlessly. Reasoning, and therefore rational choices, cannot be compared with mathematical or logical reasoning. The brain might, at times, approximate aspects of mathematical reasoning, but it does not work by doing mathematical reasoning.

Humans are rational in the biological-psychological sense but that doesn’t always conform to the more rigid, mathematical-logical forms of rationality. In fact it rarely does.

That is why there is sapience and wisdom.

Unlike intelligent decision making, which is based on a reasoning engine (see chapter 4), sapience is based on the construction of malleable models of the world and the self. Those models are not guaranteed to be right or perfect. They are not absolutely ‘true.’ Nor are they ever ‘complete.’ They are always being modified as the result of actual experience with the modeled object in the real world. Sapience operates to ensure that models are refined in the sense that they become *truer* over time and experience. That means that truth, in this sense, is not a yes or no proposition. Rather there is a range of truth values from completely false to completely true. And most models of reality fall on that range. Strong sapience means that those models are closer to the completely true end, but can never actually get there. And that is because ‘there’ is forever itself moving.

Sapience constructed models of reality do not mechanically cause our decisions to be good. Rather they act as an influence on the decision process and they do so at a subconscious level (see below). They work through judgments and intuitions. Greater sapience produces better models, which produce better judgments and intuitions. The collection of models is what we call wisdom.

Sapience and its produced wisdom are much less about reasoning and rationality and more about feeling, of being a part of that world which is modeled. Some part of the world moves in a certain way and our minds move with it. Moreover, and much more importantly, when it moves our minds, because of our wisdom, can anticipate the future effects of that motion. We are not mere responders to changes, we are predictors of what comes next and that is the key to our ability, when fully developed, to decide what to do next. We don't so much experience changes (the dynamics) of the world as intellectual, reasoned, thoughts as we *feel* we know what is going to happen.

Wisdom as a Psychological Construct

Intuitions and judgments come from deep within the mind, from the subconscious levels of thinking. They arise, as it were, to conscious awareness and have the role of guiding conscious decisions. The mind is doing an incredible amount of work in sorting out all of the variables and conditions that constitute a complex situation. Those men and women who seem to be able to effortlessly provide a recommendation (guidance) in complex social problems that proves to be worthwhile in the long run are noted as 'wise' individuals, so long as the others are able to recognize the advice and take it[52]. Most often, and this seems to be a universal characteristic, such individuals are more elderly. The wise elder, in spite of the trend in western cultures to favor youthfulness, is an enduring meme in our collective consciousness.

Wisdom involves the capacity to make good moral judgments in complex social problems that are life supporting for the majority of a society for the longest time into the future[53]. Wisdom also involves the behaviors of individuals who are competent in learning life experiences in a manner that gives rise to veridical intuitions later in life. It must be based on a brain function and competency level that increases the likelihood of survival and reproduction of the species or it would not have become a feature of human mentation. It can be argued, from the anthropological record, that early human groups that possessed at least one very wise leader were more fit than other groups in surviving the exigencies of life. Indeed, an argument can be made that the wisdom of the older members of a tribe contributed to the group fitness and thus to the differential success of those groups[54]. Wisdom in the ways of the world was the evolutionary basis for the success of *Homo sapiens*. Grandparents, the locus of wisdom, could transmit that

---

[52] Note that in eastern cultures there has been a reverence for the elders as being wise, especially in agrarian regions. In western cultures, especially in non-native North American (i.e. European imports) communities the reverence for elder wisdom is generally quite diminished. The cult of youth seems to have become the more dominant influence in the culture. In the west we tend to relegate our elders to 'living centers', which are actually 'dying centers.'

[53] This aspect comes up in a number of chapters in Sternberg's edited volume (1990). See: Robinson, "Wisdom through the ages" (chapter 2), Labouvie-Vief, "Wisdom as integrated thought: historical and developmental perspectives" (chapter 4) for just a few examples.

[54] Group selection is advanced to explain the evolution of a number of cooperative traits in human small groups. See, for example: Smith (1964), Wilson & Wilson (2008). & Wilson & Sober (1994). It is hypothesized that in a small group, say 50 to 100 individuals, at least one elder would have achieved cognitive capabilities that would be identified as wise counsel.

fitness to their children and especially their grandchildren thus providing a differential reproductive success advantage to their kin group. In the late Pleistocene era, wisdom contributed to the fitness of the human species.

As the genus *Homo* evolved there was apparently a progressive orientation toward two-partner mating (or some would argue toward mild polygamous mating) as child rearing became more expensive energetically (as compared with other apes, for example). Along with this arrangement, humans were living longer past their reproductive ages. It seems that these grandparents took an interest in their children's children (perhaps because in a monogamous or semi-monogamous relationship it is easier to trace one's progeny) and their upbringing. An interesting correlated hypothesis, the Grandmother Hypothesis[55] postulates the continued capacity for postmenopausal women to care for youngsters led to stabilization of the family units and contributed to the success of the species.

The development of wisdom depends on the emergence and differential development of brain areas that contribute to capturing, organizing, and recall of tacit knowledge. In chapter 4 I will delve more deeply into the neuroscience of sapience as it pertains to the brain, and in chapter 5 I will further develop the evolutionary ideas just mentioned, but also consider the future potential evolution of sapience.

I have, several times, associated the idea of sapience with that of wisdom. I have asserted that sapience is essentially the brain basis for what we observe as wise-ness in individuals. To begin a deeper understanding of this relation I start by examining the psychology of wisdom that has been developed over the past several decades. This is basically looking at the behavioral aspects of wisdom (from a systems perspective) and identifying the modal phenomena that contribute to what researchers have identified as wise-ness.

Wisdom is Deep Understanding

The word ‘understand(ing)’ is curious. It is used in many different contexts as a noun (an understanding), a verb (I understand), and as an adjective (an understanding mind). In all cases the word refers to some kind of *knowledge*. A dictionary definition of understanding (literally derived from ‘standing under’) includes terms like ‘superior power of discernment’ and implies a capacity to ‘handle’ something, meaning an ability to work with that thing. One can say they understand something when they are able to, for example, make predictions about what that something will do or become under different circumstances. For example, you understand the nature of a force on an object when you can predict that the object will be moved or resist motion depending on the mass and the magnitude and direction of the force.

---

[55] See: Williams (1957). He produced a hypothesis that might help explain an evolutionary paradox in the reproductive life of women, that they undergo menopause and yet continue to nurture the young of their children, a form of altruism. Also: http://en.wikipedia.org/wiki/Grandmother_hypothesis

Certainly some understanding is a result of mere intelligence. By that I mean that a very intelligent individual can learn the ins and outs of something (and by that I mean a system) and build some kind of model, mental or mathematical, that can predict future behavior of that something. Such a person understands that something intellectually. But that kind of understanding comes from the limits of complexity on that something. It has to be a relatively simple system in order for an intellectual understanding to work. For example, a physicist can readily understand the dynamics of a pendulum such that she can write a set of equations that completely describe the behavior of the pendulum-gravity system. Those equations may be used to predict the position and angular velocity of the pendulum given the starting state of the system. In this sense the physicist completely understands the pendulum.

The kind of understanding that is wisdom is quite different. The kind of system that is involved in wisdom is extraordinarily complex. Social problems involving relationships, emotions, values and any number of human attributes are complex in this sense. There is no act of intellect or set of equations that could capture the essence of these kinds of systems. The number of variables involved is staggering, the dimensionality of the state space effectively infinite. Yet wise people are noted for bringing to bear “advice” or “pointers” to resolutions of problems in these complex systems. And the reason that the rest of us recognize a person as wise is that more often than not the advice is good – it works. Moreover, the wise person could not necessarily tell you how they came to understand the problem or generate the solution. The mystery is buried deep in their subconscious minds.

This is a qualitatively, if not quantitatively, very different kind of understanding, deep understanding. The wise person has a wealth of tacit knowledge about how the world works that they have acquired over their lifetime. That knowledge is a kind of model, based on the encoding of systems in their brains that can be run (subconsciously) to produce a prediction, or more precisely, an anticipated scenario.

Wise people understand the world more deeply than most others ever could. That is generally what we mean by wisdom – knowledge of how very complex things work and having a sense of what to expect under different sets of inputs (contingencies). And sapience is the brain capacity to build that tacit knowledge and run those subconscious models. The more sapient an individual is, by genetic propensity and luck in their developmental environment, the more wise they can become in their later years. That brain basis and its results in competencies is what I want to explore. The starting point will be to look very closely at the system of sapience from a sort of black box perspective, i.e. psychologically.

Psychological Constructs

A psychological construct is basically a set of measurable cognitive functions and the framework in which they operate. The most commonly known construct would be intelligence. The function of the construct can be explained. For example intelligence (or at least one of its sub-functions[56]) is the way in which the brain solves "problems." Moreover, the attributes of the construct have quantitative measures. The 'famous' (or infamous) IQ measure is a way to quantify 'school' intelligence. But there are a large number of other measures of intelligence that have been worked out and generally correlate well with life successes (especially in school). While many controversies and open questions remain about the defining qualities of intelligence, there is general agreement that the construct is real and generally predictive of the competency of the individual in terms of problem solving abilities. Creativity is another similar construct[57]. Psychology researchers often tend to specialize in one of these constructs in the sense that they attempt to find ways to expose their finer structure and functions. In recent years some of these researchers have begun to explore the idea of wisdom (clearly a cognitive process) as a construct, seeking attributes and parameters that they can probe and measure

I think this is a good thing, generally, and well overdue in terms of understanding *Homo sapiens* better.

To begin an explication of this hypothesis we need to start with the current understanding of the psychological theories of intelligence, creativity, wisdom, and affect. A considerable amount of work has been done regarding these various constructs and their interactions[58]. In figure 1.5 I have attempted to delineate these constructs in a kind of Venn diagram. Fundamentally, the four aspects of mental life have both individual characteristics, but also share some characteristics that account for the correlations that a number of researchers have noted in the above referenced works.

---

[56] By sub-function I am referring to something somewhat equivalent to Martin Gardner's notion of multiple intelligences. See Gardner, 1999 for his latest views on this topic.

[57] In Sternberg (2003) the author clarifies the three psychological constructs, wisdom, intelligence, and creativity while showing their relationships. See esp. part 2, pp 89 - 143, on creativity.

[58] As in note #13, Sternberg (2003) provides an excellent framework and overview of the subject. In chapter 6 he provides a more detailed account of the interrelations between intelligence, creativity, and wisdom showing quantitative correlations between psychological components of each as developed in the prior chapters. Also see Sternberg (1990), chapter 7, pp 142 - 159.

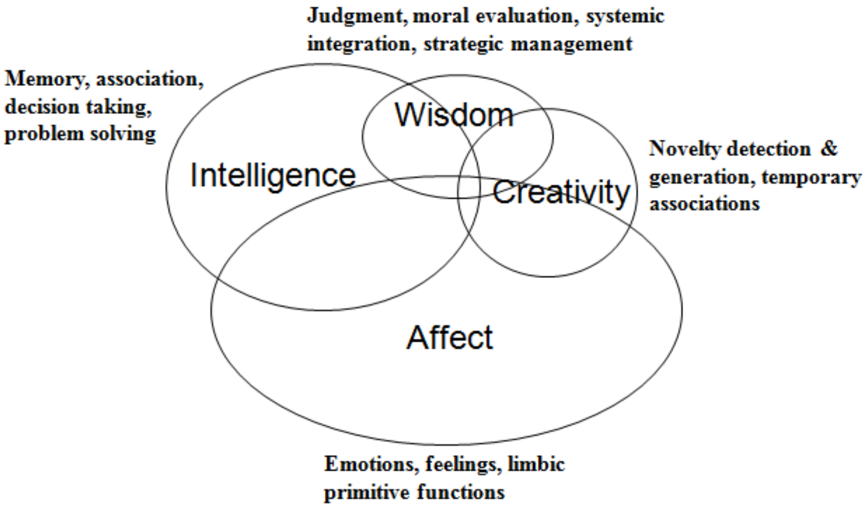


Fig. 1.5. There are four major features or constructs to the human mind. The affect system represents the emotional and autonomic systems that we inherited from the most primitive animals (particularly reptiles). Intelligence, creativity, and more recently, wisdom are basic constructs that have been studied by psychology. Of course many other areas, such as perception, are important, but are seen as contributory to these main constructs. The relative sizes of the ovals are meant to convey the sense in which each of these constructs seem to dominate human mentation. Wisdom (sapience) is envisioned as the newest and least developed capacity of the human mind.

Wisdom shares some aspects with intelligence and creativity[59] as well as having an emotional aspect[60]. This is why I used overlapping ovals to show the intersections of the four constructs. All four components of mentation interact and operate together to produce the complete, normal conscious person.

In the next chapter I will break down the relationship between wisdom and the other three constructs of mind in more detail. This overview has been meant to simply set up the framework of these relationships. Here I want to further dissect the nature of sapience to delineate it from the common concepts of intelligence and creativity and to disabuse the reader of thinking that sapience is just quantitatively more of these.

Constructs as Subsystems

The overlap of constructs in figure 1.5 would at first seem problematic from a strict systems point of view. If we insist that, for example, intelligence is a wholly enclosed subsystem then it would seem correct to show it as an independent oval with message arrows running between it and the other constructs. But this is actually not necessary. Rather we consider the constructs as sets of modules (each of which is itself a subsystem). Modules can be shared. That is a single module may perform a function that can be used by more than one other module in a system much like subroutines in computer programs can be called from multiple different points in a

---

[59] Sternberg (2003) as a whole first delineates the three constructs and then, as suggested in footnote #13, integrates them to show how they work together. Figure 1.1 only indicates a qualitative way in which the major constructs interact. Sternberg provides a quantitative mapping of those interactions.

[60] Kramer (1990).

larger program. In chapter 4 I will demonstrate how brain modules can serve multiple constructs, thus allowing the kind of overlapping represented in figure 1.5, which can be interpreted as one construct using modules from the other constructs.

Subsystems of Constructs

In chapter 3 I will be looking at what are effectively the subsystems within sapience in greater detail. Below I will outline these as sub-constructs to introduce them in terms of their major functions. The major point here is that if we examine these constructs as systems then it is possible to analyze their internal structures and function along the same lines as subsystems. At the same time this should give us guidance to thinking about these as constructs in their own right. That is, they should have measureable psychological (behavioral, cognitive) attributes that we can probe to determine their impact on overall psychology.

While this book is not about intelligence or creativity per se, the same thing is true for those constructs as well. That is we can tease out subsystems within the intelligence construct by systems analysis. In effect this is what Howard Gardener (1999) did to bring us the concept of multiple intelligences, but the approach traces back at least to Galton's methods of psychometrics (mostly about the senses and response times).

Introduction to Thinking

In chapter 4 I will provide a very detailed description of the process of thinking. That is, how do circuits in the neocortex operate in both conscious and subconscious actions that result in selection of action choices? In this section I will look at thinking from a high-level psychological perspective.

The four psychological constructs play unique but mutually interactive roles in the conscious experience of thinking. Our thoughts are shaped by the interplay between them. Actually, our conscious thoughts are, so to speak, only the tip of the cognition iceberg. Our brains carry on far more processing at a subconscious level than most of us might imagine[61]. Since it is, by definition, subconscious, and therefore not 'visible' to conscious awareness, our conscious minds can be fooled into believing that the thoughts we experience come from some effortful process, especially from our intelligence. But in fact, they are now thought to emerge as perceptually or conceptually initiated structures. Indeed evidence suggests that a number of structures emerge simultaneously in various parts of the brain, e.g. areas responsible for creatively combining basic concepts in new ways or areas responsible for activating beliefs. These multiple pre-thoughts are then tested and filtered in various ways until something like a fully formed thought emerges into

---

[61] Indeed unconscious thinking is probably significantly underappreciated by most people. Stanislas Dehaene (2014) provides a wonderful description of the amount of subconscious or subliminal thinking that goes on in the brain in chapter 2, *Fathoming Unconscious Depths*. It turns out that the evidence strongly favors the view, as I have presented it in the text, that most of our thinking goes on in the subconscious and that only a tiny, and final-form amount of thinking takes place in conscious work space.

the limelight of conscious awareness. That thought may then be incorporated into an intentional structure such as a sentence or an action[62].

Using the same construction of the mind as represented in figure 1.5, figure 1.6, below, shows the delineation of conscious and subconscious thinking. Note that most of the affective processing happens at a very low subconscious level owing to the fact that it comes from an evolutionarily older part of the brain, sometimes referred to as the 'limbic' system (more of which I will discuss later). As depicted, a proportionately larger amount of thinking is now thought to occur in the subconscious brain, with only selective thoughts rising to the level of consciousness. In chapter 4 I will explore this aspect much more thoroughly.

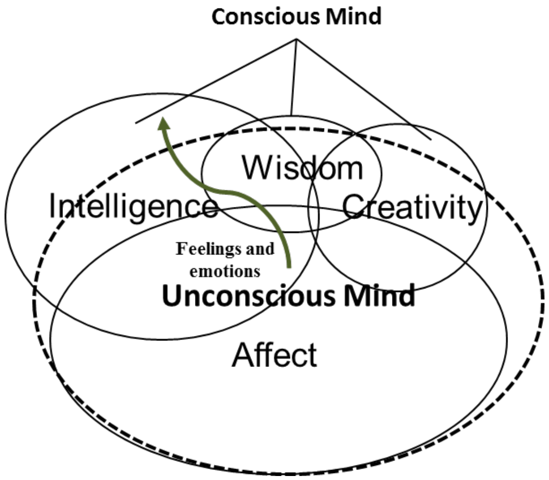


Fig. 1.6. The conscious mind involves a relatively small proportion of our total mental activities. Here the unconscious mind forms essentially a core (dashed oval) in which the mental activities of the four constructs operate to produce thoughts that eventually emerge into the outer conscious mind. Affect is shown as completely in the unconscious mind which is not quite correct. Our moods and emotions are 'perceived' in consciousness obviously. One way to think about it is that various affective states are brought to consciousness through the intelligent processing construct and through the wisdom (sapience) construct (green arrow). More explanation will follow.

It now appears that the subconscious mind does an extraordinary amount of thinking[63]. Most of this thinking is likely to be in the form of pattern recognition, matching, and categorization based on heuristic associations. If it walks like a duck and quacks like a duck then it must be a duck. These thoughts emerge as intuitions or, as we will see, judgments that guide decision making. Other forms of thinking correspond with what we normally think of as logical reasoning. This is thinking that is sequential and based on what I think of as 'tight' heuristics. These are rules that are nearly like rational reasoning, e.g. syllogistic or even predicate logic (but see above re: rationality and reasoning). They involve concepts that are essentially symbol-like (almost non-mutable), but that can be used to represent variables, and combine those symbols according to

---

[62] Baars (2007)

[63] Mlodinow (2012) explores this new view of the subconscious mind's power in doing the majority of work in all kinds of thinking.

the tight heuristics such that the derivation of a conclusion is very close to what we would call sound (veridical). Almost all minds can perform the simpler forms of formal logic, but fewer have the ability to perform the higher forms.

This process can be most readily seen in certain kinds of games, like chess. Master chess players use a combination of logic and pattern recognition/manipulation to decide on their moves. They examine the board and the pattern of pieces calls into their subconscious minds all similar patterns that have previously been encountered by that player. Those patterns are tagged with affective valence (see chapter 2 for more explanation) as well as tacit weighting having to do with prior wins and losses. The mind then filters all of the relevant patterns as well as uses the reasoning system to make a move decision. Part of the decision is rational and part intuition or judgment.

Several authors in psychology have posited the existence of two thinking systems as just described[64]. They view the use of these two systems in terms of both working on a problem simultaneously and then the mind essentially resolving any differences or choosing one solution over the other. In a very real sense this is the ‘heart versus head’ conundrum that most of us experience at various times in our lives. In chapter 4, “The Neuroscience of Sapience”, I will delve much deeper into the kinds of neural structures that might underlie these two, what we might think of as extremes, of thinking. I will argue there that I suspect the seeming dual systems of thinking is actually just the psychological observation of the two ends of a spectrum of neural processing structures that actually share some basic architectures. Combined with the conscious vs. subconscious processing, its apparent duality, this actually provides for a generally very smooth global thought processing. The same kinds of neural structures, composed of similar kinds of neurons, can simply produce more rule-like or more pattern-like processing. Indeed, in some instances it takes both working together to even represent things like the variables needed for predicate-style logic.

However, the twin dualities of conscious versus unconscious and logic-like versus pattern-like processing serve reasonably well in understanding some of the more mysterious aspects of the mind. Figure 1.7 shows a slightly different image of the four mental constructs, shifted around a little and with a line demarcating a (somewhat arbitrary) boundary between logical (that is sequential, rule-based) and pattern (parallel, statistical matching) processing. Note that the latter occupies a much greater fraction of all of the construct areas.

---

[64] c.f. Sloman (2002) for a review

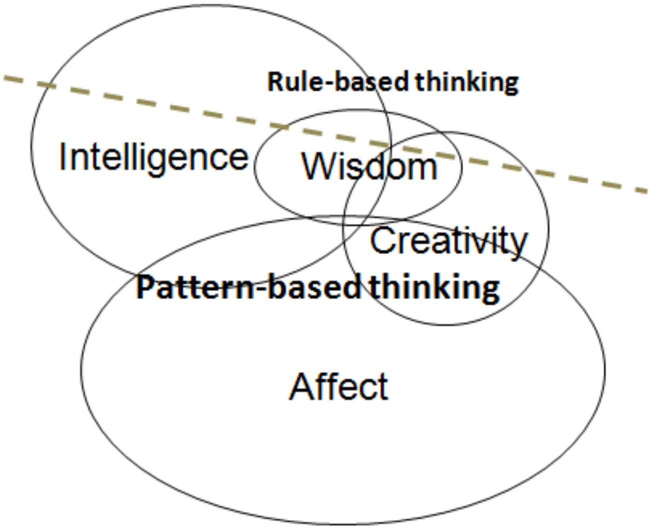


Fig. 1.7. There is a hypothetical division of the mind between two complementary processing systems, a logic-like system responsible for rule-based reasoning, and a pattern processing system responsible for similarity matching. The latter is shown as having a much greater amount of total mental capacity and accounts for almost all affective processing, most creativity, and a fair amount of intelligence and sapience. Logical reasoning is relegated to smaller portions of each. However, intelligence, which is what most people think of in terms of logical reasoning, carries the weight for thinking things through rationally.

The line dividing the logical (rule-based) from the pattern (associative) forms of thinking is placed to indicate that much more actual thinking takes place in the form of pattern-processing and associations than logic and rule-based inferences. Most humans, most of the time, deal with the world as they find it by pattern matching with memories of similar situations rather than thinking through the logical inferences[65]. This is because most life situations do not require thinking through of the sequences of premises to conclusions. If we have had life experiences that encoded in our memories as meaningful situations, we will preferentially draw upon those experiences and linkages to meaningful outcomes to guide action decisions in the here and now. Besides, thinking logically requires considerably more work and time and most people have difficulty thinking through the situation when too many variables are involved (seven plus or minus two variables seems to be the average size chunks that most people can handle).

Most of our thinking is associative, and most of it takes place in the subconscious. A small smattering, then, takes place in conscious awareness and involves sequences of rules applied to propositions and variables, what most of us think of as thinking. Indeed one very plausible explanation for the conscious mind (or evolutionarily speaking, what is consciousness for?) is that it evolved as an addendum to our ordinary supervisory functions (e.g. impulse control) specifically to orchestrate logical thinking. In fact several authors liken the conscious mind to a conductor of an orchestra. Some more elaborated supervisory functions need to selectively (and programmatically) activate specific other cognitive functions at just the right time in order to

---

[65] Daniel Kahnaman (2011) has produced a compelling theory about a dual system brain, a fast, reactive system that does most of the work of deciding by pattern recognition and automatic responses and a slow, deliberative system that has to apply reasoning.

produce what we experience as conscious thinking, especially the experience of talking ourselves through a situation in our heads (silently).

I will return to this whole subject and provide some more details of how this might be accomplished in neural tissues and brain regions given what we currently understand about these. But here I would like to point out a simple fact. Though I have referred to the conscious process of thinking as a logic-like system, in fact it is highly prone to many kinds of errors. It is not fully constrained to only use *a priori* true premises or axioms (in math and deductive logic). Nor is it constrained to only use valid inference rules. Indeed we invented mathematics and formal logics, done by writing the symbols and rules down in an external representation, precisely because our brains have greater or lesser degrees of competence when trying to think through more complex problems and situations. Unfortunately, the formal systems we invented don't scale to most real life situations and complexities. For that we humans rely on other mechanisms to convey patterns that can be combined and manipulated (as I inferred above) in what we call stories or narratives. These, some of them of universal experience, can be combined and recombined in ways to produce new, interesting stories. Our fictional prose and poetry reflect our attempts to capture almost-rule-like, language-based patterns that can be shared around, not just for entertainment, but to teach lessons and educate the young. Such story construction, involving all parts of the cognitive constructs, are examples of how the supposed two systems cooperate to produce hybrid thoughts, neither purely logical nor purely associative.

From Whence Cometh Wisdom

There may be, however, a few problems with understanding wisdom as purely a psychological construct, which may also be why it has been tackled only lately. Most obviously the word wisdom often denotes a kind of knowledge and not necessarily the mechanism by which that knowledge is acquired or used. To be ‘wise’ is to gain and use wisdom. Thus to say that one is studying wisdom one might tend to focus on the kind of knowledge that connotes wisdom or the behavior of one who supposedly possesses wisdom. Of course this must be done, just as the kind of knowledge that is gotten and used by intelligence must be well characterized in order to say anything about what intelligence is. However, it seems to me problematic to lump the knowledge and the processes that produce and use it together. The prior may be in the realm of social psychology while the latter is surely in the realm of neuropsychology. We need to link the behavioral phenomenon we call wisdom with the brain structures and functions that underlie it.

For one thing the nature of the knowledge attributed to wisdom varies across cultures and world views. Was Machiavelli wise or merely devious? Was Gandhi wise or merely political? Does one need to be fearful of a vengeful god to be wise? Can an atheist be wise? These questions reflect some of the cultural problems with pursuing a scientific study of wisdom, even as a psychological construct. We can list off a number of attributes of behavior of those others consider wise and construct a multidimensional space of characteristics, of course and that is what the psychologists have done. But which set of attributes shall be used? Can everyone who wishes to weigh in on the subject agree on them?

This situation, of course, reflects the nascent state of the study of wisdom, especially using only the tools of cognitive psychology. I prefer a different perspective for which I propose a different but obviously related moniker. That perspective is based on a more holistic systems approach to the type of cognition that is clearly more than mere intelligence and creativity. With the advances in neuroimaging available today, exciting new discoveries about what parts of the brain are engaged in various kinds of cognition for which we can establish correlates in behavior are possible. Combining this avenue with the evidence of human evolution from physical anthropology's discoveries regarding the evolution of prefrontal cortex from brain case endocasts and cultural paleoanthropology's discoveries regarding the advent of advanced symbolic processing we can then apply the approach of systems analysis to develop a more comprehensive theory of human cognition and especially the basis of wisdom.

My preference is to use the word 'sapience' to designate this more expanded notion of a construct. Sapience is the Latin word generally translated to mean wisdom but also meaning wise, as in *being wise*. I want to use it to connote the brain basis for a form of cognition that goes beyond intelligence, creativity, and affect, to provide the substrate for gaining and using the kind of knowledge that we call wisdom. More than just a psychologically defined construct, sapience encompasses the brain structures and their evolution from out of pre-sapient hominids.

This isn't without precedence. Neuropsychology is currently mapping the brain basis for a number of intelligent behaviors and also what parts of the brain are engaged in creative acts such as imagination and invention. The same is true for affective behaviors. Thus our understanding of our behavior and how it arises in the workings of the brain has progressed considerably. What makes the study of sapience particularly interesting is the correlation between the most dramatic and very recent development in the brain of *Homo* and the emergence of several behavioral traits that embody wisdom. Specifically the patch of tissue right behind your eyebrows, in the prefrontal cortex, known as Brodmann area 10 (BA10) underwent a significant expansion at about the same time that symbolic language appears to have emerged, or, at least, began to have clearly evident impact on human social organization. In chapter 4 I will explore the contribution of BA10 to the other brain structures that play a role in sapience. In chapter 5 I will explore how its expansion at the beginning of the Holocene resulted in the emergence of wise behavior.

Sapience, thus, encompasses the psychological construct of wisdom, the recently evolved brain structures involved, and the behavioral impacts that affect human social organization.

Introducing the Components of Sapience

I would now like to start deconstructing the functions of sapience in the cognitive framework established above via systems analysis. In this introduction I will outline the components of sapience that seem the most salient at present. In chapter 3 I will delve even deeper into each component as it were, peeling the layers of the onion to reveal more detail.

The components of sapience that I will present are based on the study of the wisdom literature[66]. I have identified four basic themes that seem to encompass much of what most people consider to be wisdom in many cultures. There may, of course, be more components than just these four. Or some of what I include in any one of these four might better be described as a separate component having equal footing with these. This model should only be considered as a start on the program of understanding the basis of wisdom and human cognition including wisdom.

In a manner similar to the diagram of the mind introduced above in figure 1.5, figure 1.8 below shows a diagram of the components (sub-constructs) of sapience. The four components are *judgment*, *moral sentiment*, *strategic perspective* (thinking), and *systems perspective* (thinking). Each will be described briefly along with their relationships to the other mental components of intelligence, creativity, and affect. In the figure they are shown overlapping because they all interrelate to one another in very complex ways in exactly the same way that intelligence, creativity, affect, and sapience interrelate.

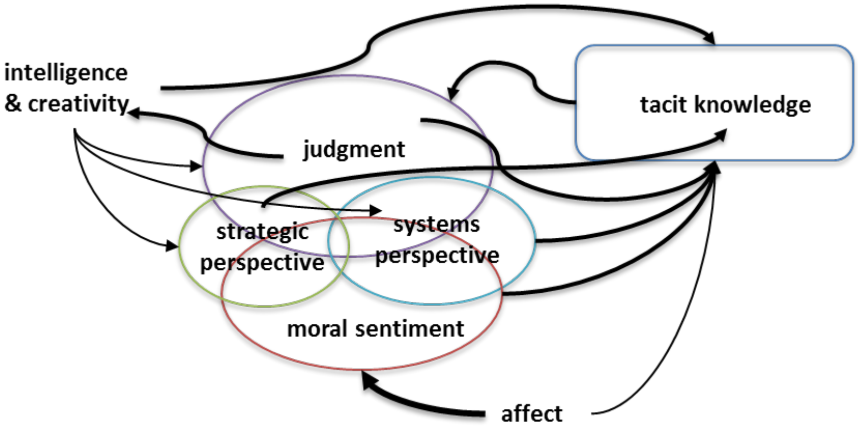


Fig. 1.8. Sapience decomposed into four basic subsystems (sub-constructs), judgment, moral sentiment, strategic perspective, and systems perspective. See the text for explanations.

The arrows, from and to external functions, in the figure are meant to roughly represent the recurrent messages that pass between functions (and between brain centers that are involved in

---

[66] A major source of this literature has been Sternberg's edited collection (1990). My method was to search through every relevant chapter doing a content analysis of terms that occurred frequently and seemed related across chapters. I noticed in this approach that terms like 'long-term perspective' and 'thinking about the group as opposed to self' appeared to cluster strongly (i.e. strategic perspective in this case). I found that four clusters stood out as thematic and I looked for titles that seemed to fit. I then re-read the chapters to verify the groupings and to assess the relative weighting of each. By far there were more mentions of judgment related cognition than any of the other three, which is why I weight it more heavily. Moral sentiment (e.g. cooperativity or empathy) was mentioned almost as often. Certainly these are very rough estimates of weightings (not an attempt to quantify them). However as I continued to delve into the literature on judgment and wisdom from other sources the same four themes in approximately the same weightings appeared consistently. Thus I present this as a viable model of sapience.

processing them). The overlapping of central functions in sapience means to convey the tight integration of these functions without explicitly representing it with arrows. Note the thick arrow from judgment to intelligence and creativity. This is meant to convey the effect judgment has on these processes (see below).

The box labeled 'tacit knowledge' represents the storehouse of implicit memory encodings that connect conceptual memories in functional ways. It is, in effect, our model (or models) of how the world works. All of the functions of mind (intelligence and creativity - I&C - and affect) and all of the functions of sapience affect tacit knowledge. I&C act to form concepts and their relations and link them to the tacit knowledge store. Affect tags each such linkage with valence, good or bad affect[67]. In chapter 3 I will delve deeply into this structure. For now I will just give brief description of the components.

Judgment informs what new knowledge should be incorporated and how it should be integrated into our existing tacit storehouse. Similarly, strategic and systems perspectives help guide this process (knowledge acquisition) as well as helping interpret stored strategic and systems knowledge for the judgment function. Finally, moral sentiment, our sense of rightness and wrongness, guide the integration with respect to our social mores and rules of conduct.

Judgment

The capacity for judgment has come under the study of psychology and neuroscience in the past several decades. Judgment is one of those ineffable capacities that seems somehow related to intuition and yet clearly is linked with rational thinking and decision making. On the one hand, our judgments come unbidden from somewhere in our minds to guide our decisions, yet most of us do not really have a conscious experience of forming a judgment[68]. It is just something we all do.

Judgment is an integral part of decision making. In chapter 2 I will delve more deeply into that relationship and show how the mechanics of decision making, a function of intelligence, is mediated and shaped by judgment. For now simply note that we make judgments about almost everything without consciously thinking about it. That is, judgment is not the same as rational decision making or thinking. It is something that emerges from our subconscious thinking (as above) and yet prevails upon our conscious thinking to direct our actions. Recent neurological evidence suggests that, in fact, our conscious "decisions" are already decided and that consciousness is really just an after-the-fact recognition of the decision and an illusion of conscious volition. I will take this up in chapter 4 in greater detail.

More primitive minds, and by this I mean those of animals whose brains are pre-sapient, are guided in decision making (presumably without conscious thought, or at least a language) more

---

[67] Damasio (1994) describes the way in which the affective system tags or marks neural links with valence. See also chapter 4 for more details.

[68] Hogarth, 1980, see esp. page 1

by affective drives than by intelligent reasoning, and those drives are built-in. I refer here to the primitive reptilian brain which is reactive to environmental cues that trigger emotion-like responses. The classic fight-or-flight response is a good example. In those minds there is no higher consideration of the circumstances or overriding of the initial response because there are no higher brain functions to evaluate the 'real' situation or broader context of the cues. The affective modulation of decision processing had worked very well through the evolution of the reptiles. The evolutionary hardwiring of responses to cues based upon temporally immediate survival considerations was sufficient as a control for behavior. We humans (and all mammals) have inherited this basic judgment system. And, all too often, it can take the upper hand in modulating our more rational sides, as when you lose your temper and behave badly as a result.

Human minds have an added advantage in being able to bring more complex learned knowledge to bear on decisions even when we are not consciously trying to do so. That does not mean we are completely rational in our choices. Quite to the contrary the system of acquiring tacit knowledge and bringing it to bear in judgments guiding decision making is far from perfect, as we will see. Our brains have many built-in biases that keep us from making completely rational decisions[69]. And, indeed, it seems that most humans tend to rely more on their affective system guidance in most day-to-day or ordinary sorts of decision problems (like what socks to wear) than the experience-based approach. The reason is actually quite rational in the sense of saving time and energy. The affective inputs to simple decisions work reasonably well most of the time. In a simpler world, where our major concerns involved conspecifics (both us and them) and nature (predators and prey) the kinds of judgments we needed to make were relatively simple. They needed lots of tacit knowledge, to be sure, but were about things that could readily be learned and understood in our model of the world. But as the world changes more rapidly and gets more complex our native judgment capacities are being put to the test. As Robin Hogarth puts it:

> *However, the increasing interdependency and complexity of modern life means that judgment now has to be exercised on matters with more important consequences than was ever the case in the past. Moreover, the frequency with which people are called upon to make important judgments in unfamiliar circumstances is growing.* (1980)

The modern world is putting our capacity to build an adequate storehouse of tacit knowledge and our capacity to make critical judgments from whatever knowledge we are able to obtain to the test, most severely. Complex problems and the decisions needed to solve them require more effort and time. The brain has to kick into a mode of decision making that requires more rigorous thinking, both conscious and subconscious.

---

[69] see: Gilovich et al (2002)

It is not unreasonable that we relegate much of our everyday decision processing to guidance from the limbic system and the paleocortical brain[70] more than from the neocortical brain (see chapter 4). As we will see the circuits for affective and more primitive experiential (e.g. routine patterns stored in the paleocortex) inputs to decisions are already in place and a very rapid activation of those circuits is all that is needed to produce reasonable decisions. But when the problems are really complex, full of uncertainty, and convoluted, emotions, feelings and simple patterns cannot provide adequate guidance. That is where deeper tacit knowledge comes into the picture. That is where sapience takes center stage, sometimes needing to down modulate the affective inputs, sometimes overriding them entirely.

Good (veridical) judgment is necessary for good (sound) decisions, those that lead to good (favorable) outcomes. Social problems are especially complex and involve a much larger scale in time and space. More factors are involved. More people will be affected. More time will be needed. One has to think about the future and what impact the current actions will have on that future. The scope of space and time is much greater and the possible variations are literally too numerous to work through in a purely logical sense. Such problems have a more ‘global’ scale of impact and that poses a serious problem for judgment.

The reason is relatively simple. Every problem is composed of a network of sub-problems that all affect one another (see below on Systems Perspective). Yet the conscious mind must focus on one local problem at a time. If all that is brought to bear on decisions regarding that local problem is intelligence (and a smattering of creativity) there will be a tendency to try to find what we call a ‘local optimum’ solution. The reason is that we typically only have local explicit information to use in forming a decision about what to do. That local information will not include the fact that just around the next bend, out of our local (i.e., conscious) view, is an obstacle or a precipice — other related sub-problems that might be made worse by solving the current problem for its optimum. We are forced to make decisions using our intelligence and the best local information we can muster. But it does not guarantee us that we are making the right decision in a global sense. In fact there are many examples of how solving a local sub-problem for an optimum will make the global problem much worse. Tacit knowledge, if it is relatively complete and relatively valid (if our models of the whole are good) can then come into play subconsciously to alter or shape the intelligent decision making to override local optimization if there is a chance of lowering the global optimization of the larger problem.

Hence, we higher mammals, and especially humans, have brains that allow us to build a storehouse of global knowledge over our lives (if we survive!) and use that to guide intelligence in making decisions based not just on local information but also on a history of experience that can be brought to bear. Fortunately many decision points have characteristics of situations we

---

[70] The paleocortex is the more ancient part of the cerebral cortex just below the neocortex (the outer rind). It evolved in late reptiles and early mammals to support memory encoding used to modulate purely reactive (affective-based) behaviors. In the modern mammalian brain it is still involved in early memory formations (e.g. the hippocampus).

have seen in the past. As such we can apply our implicit knowledge, gained in the past, to anticipate the results of a current decision. Even this does not guarantee an absolutely correct decision. But it is better than using local information alone. There is a category of computational algorithms that mimic this concept somewhat called dynamic programming. The basic idea is to build a table of prior decisions that were successful on the chance that the same decision point may be encountered later in the computation. Then rather than solve the decision again, the algorithm simply looks in the table to extract the solution. This is a major time saver in many kinds of decision problems (and a kind of machine learning). Unfortunately computers are constrained to working on only a small piece of information at a time whereas the brain can do a massively parallel search for the needed information (see chapter 4).

The more comprehensive a model (tacit knowledge) we have the more likely our decisions will prove adequate (i.e. satisficing). Comprehensive here means covering a larger scope of space and a longer time scale. The more and varied life experiences we have had and the more lessons we have learned about those experiences the more power we bring to bear on the present local situation. This is why the brightest people who have lived long seem to be the wisest in general. They have brains capable of storing large knowledge sets with reliable and ready access to memories. They have lived long and experienced more than average. And those experiences and their meanings are encoded in tacit form. It gets back to brain competency. The brain of someone who has a higher level of sapience has the competency to acquire the right kinds of tacit knowledge and has the competency to use that knowledge to maximize the likelihood of making good decisions in a complex, fast moving world.

Another critical aspect to sapient judgment is the capacity to judge one's own judgments, or meta-judgment.

One of the attributes of wisdom seems to be the capacity for a wise person to not pass judgments on some issues when their tacit knowledge does not include experiences that apply. In other words, a wise person knows when they do not know enough to offer a judgment (which has the form of an opinion). They, like everyone else, could still rely on the built in heuristic models (chapter 3 will offer details of how this works) and form opinions on these subjects. But their sapience includes the capability to recognize their own limitations and can prevent or override an urge to offer an opinion since one has no real, efficacious tacit knowledge with which to form such opinions. As we will see in chapter 4 the ability to have knowledge of one’s own knowledge (meta-knowledge) requires a more advanced knowledge structure than had been available to pre-sapient beings. The expansion and reorganization of the prefrontal cortex leading to modern *Homo sapiens* may hold the key to understanding this facility.

All ‘merely’ (nominally) sapient humans are strongly motivated to construct explanations about happenings in the world, even when they do not have adequate information with which to do so. They construct these explanations anyway as forms of casual or informal hypotheses with the possibility that gaining additional information might allow them to refine or modify their tentative explanation in the future. This is actually a facility of the acquisition process for gaining tacit knowledge. But it can work against most people when they do not recognize that their tentative hypotheses are just that. When they, instead, fail to recognize that they do not really have the tacit knowledge needed to form more sound judgments, they offer up opinions anyway! One can witness this phenomenon most starkly in people who hold nominal leadership positions in organizations or society. Their own self-image, as well as the expectations of their 'followers', puts pressure on them to have the answers. Yet, and especially in a culture that values youth, they tend not to have the tacit knowledge to form efficacious judgments. Even so they are compelled to form an opinion. The followers can only hope that either the leader is uncommonly brilliant and can apply rational logic using explicit knowledge or gets lucky and guesses a good solution. But they cannot count on the leader being wise. This is an issue I plan to take up again in my book about governance mentioned above.

Moral Sentiment

Altruism evolved in social mammals (and birds) as a means of increasing the fitness of the group over that of individuals[71] and the general fitness of the species. We evolved a sense of right and wrong behavior in ourselves and others. The specifics of many practices and social mores vary from culture to culture, but all cultures have rules of behavior that reflect the inner sense of moral and ethical sentiments. Moral sentiment appears to be a universal property of human cognition.

While many religious and conservative people believe that moral reasoning (the 'axioms' and rules) comes from a higher power, the scientific evidence that our brains are hardwired by evolution to base judgments on inherent, and subconscious, moral sentiments is now solid[72].

The drive to moral reasoning is built into us as social creatures that need to cooperate more than compete within our tribe. Higher moral sentiments provide guidance to our acquisition and use of tacit knowledge and our modulation of the limbic system's automatic responses to prevent unreasoned actions. We have the ability to inhibit our tendency to get even with someone who has hurt us. We have the ability to inhibit our baser desires. Higher sapience means that we will exercise this control over our primitive urges. Moral sentiment and higher judgment work together. They produce behaviors that we think of as being significantly different from mere animals. We call it humanism.

There is another, perhaps even more important, aspect of moral sentiment that needs to be mentioned. Moral sentiment is driven primarily from the affective component of the mind. And one of the most powerful aspects of affect is love (hate is potentially even more powerful, but only for the lower sapient mind). That is, the human sentiment to cherish other beings in various ways, as mates, as offspring, as neighbors and friends, etc. is one of the most important factors in

---

[71] Sober & Wilson, 1994

[72] (Tomasello, 2016; Wright, 1994)

social organization and the sense of moral motivation. When all is right with the world, we love one another. This attractive force runs deeper in sapience than is generally appreciated. It is not just gushing emotion. It is a real basis for caring and thus motivation for thinking about the good for all. Wise persons, throughout history, have often been described as loving their fellow beings and nature as well.

Loving relations are based on a capacity for empathy. We are not the only mammals that experience the emotions associated with caring for others. Mothers, in many species, show real grief when their babies go missing or die. Father gorillas show caring for their children; they will play with them or at least tolerate their shenanigans. Empathy is the ability to feel for another. It isn't necessarily feeling their pain or joy exactly, but rather an awareness of what they are feeling when they are in pain or joyous. We know that the other is feeling something like what we have felt ourselves and we care that they do. Every culture has a moral sentiment that has been expressed in many ways but is basically what we in the western Christian tradition call the Golden Rule: "Do onto others as you would have done onto yourself."

Systems Perspective

One of the main problems with a notion of higher sapience being dependent on more brain power in acquiring massive amounts of knowledge is that the brain is, after all, limited in its capacity to encode memories. However, the memory capacity of the brain depends on how those memories (knowledge) are encoded. We now know that our memories are not simple recordings of happenings or images. Rather the brain builds conceptual hierarchical codes to represent things, relationships, movements, and so on. The brain re-uses representations through complex neural networks which allow sharing low level features among many higher-level concepts. The brain also organizes concepts in a hierarchical classification scheme that allows ready associations of ontological categories. For example our hierarchy of 'mammal-dog-Fido' associated with 'dog-pet-Fido' allows us to relate other kinds of dog-like animals and compare features, such as 'mammal-wolf-teeth-aggressive' with 'mammal-dog-teeth-friendly(mostly)'.

It is the organization of concepts into networked hierarchies that allow us to not have to store every little detail with every instance of a thing, place, or action. Rather, we now know that our brains reconstruct memories from cues by activating a specific network of associated neural clusters. The brain is designed to store massive amounts of encoded 'engrams' but only because it knows how to organize the components in such a way that many engrams can share sub-circuits[73].

There is another trick to organizing knowledge to achieve maximum compression. That is to base the organization of knowledge on universal models that pertain to all aspects of life. The most general such model is systemness, or the nature of general systems[74]. No matter how

---

[73] (Abdou, et. al., 2018; Alkon, 1987; Brodt et. al., 2018; Seung, 2016; Sporns, 2016;)

[74] (Mobus & Kalton, 2014)

complex the world seems, it is resolvable into a hierarchy of systems within systems. That is everything is a system and a sub-system of some larger meta-system. Systems have universal properties even though their forms may seem significantly different. Some systems with fuzzy boundaries may not even be readily recognizable. Yet the world, indeed the universe, is organized as systems within systems with varying degrees of inter connectivity, complexity, and organization. Some systems are too small to be detected with the normal human senses (bacteria and single-celled organisms). Some systems are too huge to be readily detected (the galaxy) by ordinary sensory means. Some systems are so diffuse that they cannot be easily categorized as a system (the atmosphere). But as long as there are aggregates of matter and flows of energy there are systems. As an aside, if this were not true, then science could not work as it does!

The mammalian brain is wired so as to perceive and conceive of the world as systems of systems. That is to say, our perceptual systems are genetically organized so as to detect boundaries, coherencies, patterns of interactions and connectedness, and many other attributes of systemness. We see things and we see those things interact in causal ways. We literally can't help it. This built-in capacity is the basis for learning how the world works. Above I alluded to the idea that tacit knowledge was a form of model (or models) of how the world works. And here I claim that the encoding of tacit knowledge begins with the grasp of systems.

To see things as systems and to recognize things interacting with one another in causal ways is, however, not enough for what I am calling a systems perspective. All animals to greater or lesser degrees have the ability to encode systemness (see chapter 4). Humans have the added ability to compose models of their world using systemness as a guide to construction of those models in memory. This affords us an ability to play ‘What-if’ games or test possible outcomes by altering some variables and running the models in fast forward. In other words we can think about the future. Even so, as remarkable as this ability might be, most people do it more or less without conscious recognition. It is so natural to do we rarely even consider how marvelous a facility it is.

Systemness processing is actually part of general intelligence. Psychologists have recognized several template forms of thinking that contribute to the perception of systems based on a basic typology of systems. It has been known for some time that humans have built-in processing modules for what are called ‘folk-thinking’ such as folk-physics (or mechanics specifically), folk-biology, and folk-psychology. These are presumed to be a combination of inherited models, a native ability to recognize and work with subject content, and learned particulars. The native ability is built into the brain from the get-go. Then life experiences are incorporated and the templates guide the integration of those experiences in a form of inductive learning – that is construction of tacit knowledge. Folk-physics and –biology we no doubt share with other animals, at least birds and mammals. The brain has to be predisposed as to how the world works insofar as things like gravity and forces work on the self. Similarly, animals need to have a predisposition to recognize animate objects and the basic differences between them. The recognition of conspecifics is probably already highly developed, but the ability to learn other species that are either food or eaters is built on top of that. Folk-psychology, also referred to as ‘theory of mind,’ is seen to some degree in chimpanzees but is quite robust in humans. This is the cognition of other minds and what they might be thinking. For example having a belief that another believes such-and-such is part of what allows humans to achieve the level of eusociality they have.

Sapience goes a step further. It builds upon these basic intelligence modules to refine what is learned about the physical world, the natural world, and the social world. It guides the learning and construction of models that not only are more nuanced and veridical, but also extend into the future. We think about the future states of any and all of these models. But moreover it involves thinking about thinking about the future. And it also involves thinking about systemness itself. In other words, higher sapience involves more comprehensive systems thinking. This comes out in several ways. One of the first and most important is the natural tendency to ask questions such as: “Of what larger system is this sub-system a part?” Effectively sapience drives us to want to understand the context of what we perceive in the immediate area of interest. Or, we ask: “What are the sub-systems inside this system (of interest) that make it work the way it does?” Curiosity and willingness to probe deeper or outward are necessary ingredients to support increasing one's tacit knowledge. The more sapient person has the formula for how to answer these questions by following the properties of systemness. They know what they are looking for in terms of roles to be played in a systems organization and dynamics even if they don't know in advance what the specific ‘thing’ looks like.

The capacity to quickly organize new information on the basis of systemic principles is what allows some people to learn completely new cultures, jobs, or even careers. They can relate the specifics of a newly encountered system to the general principles of systemness and learn to manipulate the new system based on those principles applying. It is a strong perspective of systemness that allows some people, and especially the more sapient, to build a comprehensive storehouse of tacit knowledge and later to use that knowledge to rapidly adapt to new system particulars. Systems recognition and perspective is at the base of the aphorism: "There is nothing new under the sun." Or the expression that: "No matter how much things change, they stay the same."

Wisdom is often characterized by a person's ability to deal with ambiguity and uncertainty. This ability is greatly enhanced by the systems perspective. It permits one to be calm in the face of uncertainty, for example, knowing that systems dynamics may seem chaotic (in the vernacular sense) but are really part of the probabilistic nature of the cosmos. Systems thinking gives resolve to the notion that while there may be great ambiguity now, further investigation (gaining additional information) will reduce ambiguity since the systems principles hold universally. In other words, this is the source of faith for the wise. The world will become clearer in time!

Strategic Perspective

You may have noticed that the world is forever changing. Systems thinking helps one adapt to change by providing generic templates of systemness that can be used as scaffolding for learning new things. But if one is to do better than simply react to change and hope to adapt then one has to employ a more advanced form of thinking — strategic thinking. As with systems thinking most people are able to do some strategic thinking, at least from time to time. But the vast majority of people stick to logistical and tactical decision making which is why our species has a tendency to discount the future and make near-term decisions on profitability. Strategic thinking is the most advanced form of thinking about the future (mentioned above). It is more than just playing what-if games. It also involves incorporating important global objectives into the models and deriving plans that will be used to drive tactical and logistical thinking into the future. Strategic thinking involves developing a vision of what the future world will be like and then picturing yourself (or your group) integrated into that future world in a way that is both satisfying and sustainable.

The importance of strategic thinking as an individual capacity is just beginning to dawn on some psychologists and evolutionary psychologists. So there isn't yet a large body of literature on this. What exists comes, again, from the judgment literature where people have been studying the systemic biases in human judgment that seem to hinder people in thinking long-term, especially subconsciously. But it should be clear that a wise person is concerned with what will happen in the long run. A wise person will counsel for actions today that will have a positive impact on the future even when he or she will not be a part of that future. Average humans are relatively short-sighted. They cannot really imagine realistically what the distant tomorrow will bring. The default assumption is that it will be like yesterday only ‘more so.’ Or they cannot envision what they need to do in order to fit into that distant tomorrow.

My suspicion is that our species was just starting to evolve higher capacities for strategic thinking (which explains why we even know what strategic thinking is!) as an advancement of our evolving sapience. But with the advent of technology, and especially agriculture, the selection pressures that would have moved us further in that direction were removed as humans came more and more to rely on technical solutions to survival. The result is that the vast majority of people do not think very strategically, even about their own lives let alone the lives of their fellow beings and the lives of future generations.

In chapter 4, *The Neuroscience of Sapience*, I will describe the model of the brain that I think best describes the nature of human thinking and accounts for the functions of sapience. Specifically I will show how the hierarchical cybernetic model[75] describes brain architecture and

---

[75] “Cybernetics” is the science of control theory. The term ‘control’ carries some unwanted baggage in general parlance, where it can mean ‘command and control’ in a very top-down manner. Hierarchical cybernetics is a model that is applicable to complex, autonomous entities like people or organizations. Modern concepts of such entities cause many people to eschew the word control as having a negative connotation. Therefore I choose to use

functions as are being elucidated by neuroscience. Briefly, the hierarchical cybernetic model is a description of the management of a complex system based on operational level controls (e.g. the feedback controls used to regulate low-level work processes), logistical level coordination (e.g. the coordination of many operational subsystems and distribution of resources in order to optimize the global behavior of a complex system), tactical level coordination (essentially equivalent to logistical level but focused on coordination with external systems so as to obtain needed resources and avoid external threats), and, finally, strategic level management (as just described above). The human brain is a management and control system that regulates our bodies, our behaviors relative to the external environment, and to some extent tries to guide our futures.

The infancy of sapience has created a situation in which we can think about the future and can have aspirations for what our situation will be in that future but it also means that we have not yet developed the strength of strategic thought that allow us to better guide ourselves into those aspired futures.

Figure 1.9 is a schematic representation of the hierarchical cybernetic model as it might pertain to any governance system, be it the brain, an organization, or a government. The figure suggests that strategic control (or management) takes considerable processing power to be effective. This is because within its domain it must have extensive computational abilities, extremely complex models of both the system it controls and the external environment, and a massive knowledgebase of past experiences. The overhead of a highly effective strategic control capability is considerable (though not greater than the total overhead from all other parts of the system). I posit that in human evolution terms, we have just obtained a beginning capacity for strategic thinking – imagine a strategic control function 1/10th the size of the one if the figure.

We can do things like plan our future to the extent of wanting to achieve some goals, but our ability to break those goals down into logical sub-goals and formulate sub-plans (e.g. tactical plans) to achieve them seems to be extremely limited. I put forth a conjecture that if humans ever do evolve higher levels of sapience, the main gain will be an ability to systematically break long-term goals (some people prefer ‘objectives’) into sets of sub-goals in intermediate time scales, and near-term goal subsets and also plan the actions that have to take place to achieve them. The latter is the hard part and requires substantially more veridical models of how the world works in order to assess the plans before execution. We, of course, do this on a very limited scale when we daydream about what words we are going to say to our ‘focus of affection’ tomorrow to woe him or her. We try to imagine their reaction and do a little adjusting if we think it would be more effective. We do not routinely think much about what kind of world our grandchildren will live in and how they will be affected by future situations (like climate change).

---

the term cybernetics to help assuage any preconceived notions about what this model has to say. A preferred alternate terminology includes coordination, cooperation, and management.

Strategic thinking also involves honest assessments of one's own strengths and weaknesses. Making those plans involves learning how to overcome weaknesses and using our strengths most effectively. Most people tend to have somewhat unrealistic self-images buoyed by egoistic thinking (something left over from our limbic system). Many people have very unrealistic notions of their strengths. Walter Mitty was not an anomaly[76].

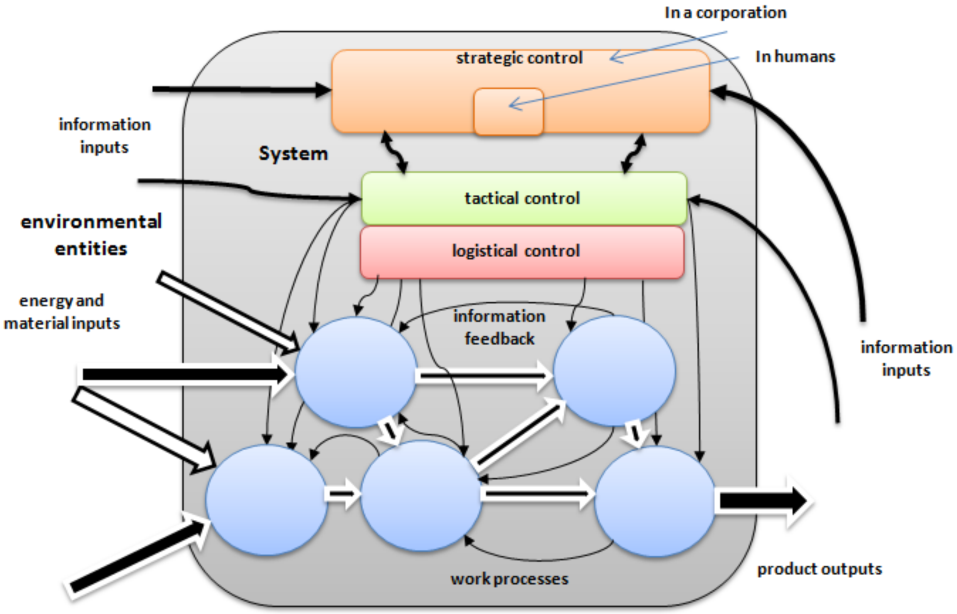


Fig. 1.9. The brain is a hierarchical cybernetic system that mirrors some of our organizational governance systems. Low level work processes need to get resource inputs from environmental entities. They interact internally by doing work on those resources and passing the intermediate products on to the next process in the system. Certain sub-systems (processes) are tasked with expelling the products (or waste products) out to the environment. This level in the hierarchy requires some feedback cooperative control (arrows with solid heads) but ultimately needs a higher, more integrated level of coordination control from a logistical coordinator that monitors the operations (not shown to avoid clutter) and provides directives to processes to assure the optimal distribution of resources. At the same basic level, the tactical coordinator monitors the immediate external world in order to coordinate the activities of the input and output processors with the availability of resources and product sinks. Monitoring the larger world, including entities not directly involved in input/output, is the strategic manager (actually a planner). This level of management also monitors the coordination levels and then provides goals and plans to that level. The time scale for activities of each level is greater as you go up the hierarchy from operations.

My central claim is that the hierarchical cybernetic theory best explains the current situation with respect to brain functions that give rise to human psychology and behavior. The strategic level of self-management was the most recent capability to emerge and constitutes a deep integral part of sapience. Sapience and the resultant wisdom that might obtain are ultimately dependent on strategic thinking but not just for the individual. Rather, the whole move toward group success

---

[76] James Thurber's *The Secret Life of Walter Mitty*, the titular character imagined himself much braver and stronger, etc. than he was in real life.

relied on the strategic perspective of the group leader(s), the wise elders, who thought ahead for the good of the whole group. I will revisit this in chapter 5, The Evolution of Sapience.

Other Psychological Attributes Associated with Sapience

Calm Contentment

The psychological literature on wisdom is full of descriptions of attributes most strongly associated with wisdom, but that do not seem to neatly fit into functional categories. Wise people are often described as having contentment, calmness, peace of mind, and so on, even in the face of stresses and turmoil in their world. The ability to live with ambiguity and uncertainty without becoming anxious is also often noted. And, perhaps most important, wise people have an ability to know when they don't know. That is they are aware of their own knowledge limitations and do not attempt to offer opinions regarding issues for which they have no basic understanding.

I would argue that the latter quality is still one of judgment. As I will discuss in chapter 3, we make second-order judgments subconsciously about the quality of our first-order judgments. This is especially important in judging, for example, the strength of evidence that our minds can bring to bear on an opinion[77]. Less wise people tend to be overly confident in their own judgments and this is compounded by an inability to recognize their mistakes and learn from them.

The feelings of contentment in the face of adversity, however, would not seem to be recognized as a function of judgment on first blush. But as I will argue in chapter 3 when we look at the components of sapience, what I call the judgment processor provides inputs down to the affect system, which includes damping down the outputs of the fear/anxiety processes (also see chapter 4, *The Neuroscience of Sapience* in which I discuss the relationship between the prefrontal cortex and the amygdala where such emotions are triggered.) Thus I think many of the attributes that are often ascribe to wise people, having to do with a sense of calmness and even satisfaction with the world as it is found are the result of the relationship between sapience and affect (implied by the overlap in figure 1.5).

It should be no surprise that as one acquires greater wisdom a certain quality of serenity also attends. The systems perspective provides an overall or holistic picture of the workings of the world and with that a sense of the dynamics following a natural course. Even when that course is negative or threatening, the wise person can often take solace in the idea that evolution is unfolding as it should and must. And with the long view of strategic perspective, one can accept the unfolding events as natural. This seeming acquiescence to negative forces may seem to the less wise, especially the young, as a kind of abrogation of caring about what happens. They would ‘fight the good fight’ to overcome any challenge. They might think the elders have just grown too tired to fight and have given up. But the wise elder can recognize the inevitable. They

---

[77] Griffin & Tversky, (2002)

may council against the fight, depending on the circumstances, or they might be content to let the young carry on the fight even when they suspect they will lose. The wise person will have a good sense of when it doesn't hurt to try! With wisdom comes acceptance of the way the world is. If there is an opportunity to make the world a little different such that everyone and everything is better off, so much the better.

Absence of Dominance Sentiment

Human relations are often structured by power of one individual over one or more other individuals. Possibly a hold-over from the dominance relations seen in our primate cousins, the chimpanzees, and assumed to have been operative in our last common ancestor[78], humans organize themselves into dominance hierarchies quite naturally when the social organization becomes relatively large and the work gets complex. In fact these hierarchies map onto the hierarchical cybernetic model of system management directly. The major difference between a systems model of hierarchical cybernetics and the human dominance hierarchy is that in the former the relations are based not on authority and coercion as much as on inter-level cooperation and clear delineation of decision processing. Decision agents at higher levels in the hierarchy can be seen as 'working for' those in the lower levels. The idea of 'lower' means closer to the work processes rather than lower in rank. The purpose of a mid-level manager in a strict hierarchical cybernetic system is to provide the operations processes with coordination to facilitate their interactions for the good of the whole. There is no sense of them being somehow more important than the operational decision makers. They are not.

In the human condition the imposition of a dominance hierarchy is likely a remnant of the primate heritage along with the fact that individual humans are still relatively autonomous and not fully sapient (i.e. cooperative naturally). People who occupy coordination level positions end up resorting to asserting their 'power' over the 'underlings' because they cannot cognize any other way to get their job done efficiently. Their emotional selves override their cooperative impulses and what emerges is the classical power structures of society.

A number of writers have noted that people deemed as wise by others often do not exercise coercion or 'bossing' to get others to behave in certain ways[79]. They generally use persuasion and patience. They recognize that individuals can be stubbornly ignorant or have selfish motivations that cause them to not want to cooperate for the good of the group, but do not resort to threats of sanctions if their directions are not followed. Very often such wise people rarely

---

[78] Although we appear to share this trait of dominance hierarchy or "pecking order" with the chimps (and it is seen in many other species of primates) our other close cousins, who split from the common chimps (*Pan troglodytes*), the Bonobos (*Pan paniscus*) about one million years ago, show little in the way of such a hierarchy, especially a male dominated one. It is not known if the bonobos simply lost this trait after diverging with the common ancestor with chimps or if the common ancestor of both humans and chimps may have been more egalitarian and chimps and humans acquired the dominance trait independently. However the fact that the trait is seen more generally in many other primate species suggests that we started out with the trait and it was lost in bonobos. See De Waal (2005, 2010, 2014) for a thorough analysis of this issue.

[79] See Robinson (1990) for a review.

take positions (titles associated with rank in the hierarchy) of ‘authority’ in their social organizations, preferring to act in advisory capacities so that they avoid temptation to try to assert their position over those of supposed lower rank.

The existence of dominance hierarchies in our social systems, which most people take for granted as ‘just the way it is.’ is a consequence of the lack of higher sapience in the general population. Their existence is necessary given that individuals do not generally see the good of the whole as more important than their own and so will tend to pursue selfish ends. Thus the exertion of power relations becomes necessary in order to enforce the form of coordination needed to keep the whole social endeavor progressing in an orderly way.

In a subsequent volume on the governance of social systems based on the principles of systems science I will be revisiting this issue in much greater depth[80]. At present I only wish to observe that what we commonly experience as politics is a natural consequence of our species possessing a ‘minimal’ level of sapience – our tendency to be eusocial while retaining individualistic tendencies at the same time.

Aesthetic Sense

Though the main thrust of my arguments is toward the role of sapience in terms of wisdom, there are other uniquely human traits that are likely to be products of the emergence of sapience that are not necessarily part of wisdom (though some psychologists might claim they are). Specifically I refer to aesthetics in the broad sense. Appreciation of symmetry, art, music, craftsmanship, and other purely sentimental feelings may very well be associated with judgment as I have been describing it. They may play a role in guiding decisions, as discussed in chapter 2, adding another dimension to the process. In figure 2.4 in that chapter I show an influence (information channel) arrow from the perception module to the affect module and then one from affect to the decision processor (all explained in that chapter). My best guess is that aesthetics arise from the overlap of all four constructs in figure 1.5 above.

At the present I am unaware of any definitive studies on where aesthetic sense arises in the brain. So I will not delve into the subject any further than to say that I suspect it is an important subject to explore since many of the other attributes of wisdom seem somehow linked to it. The study of aesthetics has been pursued by philosophers, which to me implies that it relates to wisdom in some strong sense.

---

[80] The working title is “The Systems Science Approach to Understanding Governance.” It will examine the use of the hierarchical cybernetic system model (chapter 9 in Mobus & Kalton, 2014) to grasp the ‘ideal’ of a governance system given that the decision agents are sufficiently sapient and then compare that with what we observe ‘in nature.’ It analyzes the differences between what we see in the world and what might be given that society were comprised of more sapient individuals. See chapter 5 in this volume for a discussion of how more sapient individuals might become the norm.

About all I am certain of is that since aesthetic sense is a part of unique human consciousness, then it must be a result of the evolution of sapience. Aesthetic sense is related strongly to yet another attribute of human cognition.

Spirituality

We now know that early humans such as Neanderthals had some sense of a spiritual life. Wisdom has been associated with ineffable feelings of oneness with the world (universe), a sense of belonging to something very much bigger than the individual. This is not to be confused with religious beliefs, which are derivative ideas that were developed by various spiritual experiencers in an attempt to "explain" their experiences (described further in chapter 2). The ineffable nature of these experiences makes it difficult to try to express in words what an emotional feeling is, in essence. Try describing the feeling of love sometime to see how hard it is to formulate "rational" descriptions of these kinds of experiences.

In no way can we discount the reality of spiritual experiences just because they cannot be described easily (or at all). However we should be extremely careful about ascribing these experiences to sources, especially outside of the brain. I have had spiritual experiences myself. These have ranged from 'out-of-body' perceptions to a complete voiding of any awareness while still presumably conscious (in meditation). I have no doubts that I had such experiences but I also do not attribute them to outside influences (an unseen spiritual world). I have come to suspect that such experiences are related to the subconscious elements of the moral sentiment and the systems perspective as described above. That is, as one begins to see the wholeness of the universe from a strong systems perspective the subconscious mind is busy formulating an explanatory model that does attribute causes to external sources. After all, we are all completely embedded in the fabric of the Universe and therefore components in it. Our minds, at least the subconscious parts that deal with highly abstract models of reality must incorporate this as a result of our evolving sapience.

The struggle to express these very difficult concepts, especially to those who are of "normal" sapience (as explained in chapter 5) force the use of language that preserves the mystery of the spirit experience as a means of motivating those whose experiences are not as vivid to attend to an important aspect of group cohesion. Hence religious doctrines and rituals help those who see help those who do not quite see. Religious traditions as core parts of cultures have grown from spiritual insights that needed to be codified in order to share with other humans. Religions per se are not directly spiritual embodiments, but in their best form do supply low sapient beings with access to important feelings that fulfill some of the important functions of sapience in shaping the collective intentionality of society[81].

---

[81] See Atran (2002) for a comprehensive treatment of the evolution of religious motivations both from the standpoint of the inherent psychology and from the standpoint of cultural development.

Today religions are starting to be seen as problematic because the doctrinaire elements are being taken far too seriously in the absence of truly wise individuals acting as shaman interpreters. There are few to no sages these days[82]. And so what was once a strong socializing influence and a source of great inspiration in motivating people to love their world has turned into one of the greatest sources of conflict that humans possess. In the next chapter I will return to the issue of spiritualistic thinking and consider the interactions between it and ideological thinking that has given rise to the dangerous side of religions. In chapter 5 I will return to the idea that the future evolution of greater sapience may once again supply humanity with a purer form of spirituality.

Conclusion

Sapience is the brain basis for what is unique in human cognition. It is significantly evolved in the current human species and demonstrably produces very important executive functions in this species. It is the basis for the development of whatever capacity for wisdom we see in humans. Wisdom develops over the life of an individual as a result of sapience functions obtaining tacit knowledge and using that knowledge to form moral, strategic, and systemic judgments.

Sapience is an inherent, that is genetically mediated, capacity of brains that have greater processing power in key regions of the prefrontal cortex (chapter 4). However, it is not sufficiently well developed in the vast majority of the population, which could explain why humanity is in the mess it is in today (see chapter 5). We've made some very unwise choices throughout history. We continue to fail to learn from our past mistakes, both individually and collectively. Our lack of wisdom on both fronts will doom us to make more serious errors in the future[83]. It could possibly lead to the extinction of the genus *Homo* as there are no other representatives of the only talking ape, *sapiens*.

This has been a basic overview of the thesis on sapience. In chapter 2 I will begin to delve much deeper into the relationships between sapience and the other constructs that have been studied most heavily by psychology and neuroscience, intelligence, creativity, and affect. This will, hopefully, establish sapience as a real and separate construct (as wisdom) that might be explored by those sciences in its own light. Chapter 3 will further develop the concepts described here as the components of sapience, judgment, moral sentiment, systems perspective, and strategic perspective. I will attempt to show how these components are derived from general psychological and neurological knowledge. Then, in chapter 4 I will explore the specific neurological basis for my claims regarding the nature of sapience. There I will attempt to bring together some of the most recent research on neuroscience, with my own theoretical work on intelligence, creativity, affect, and sapience as it may be realized in actual brain structures and

---

[82] As I write this the newest Catholic Pope, Francis, seems to have broken with many modern trends in Popeish behavior by eschewing the riches of the Vatican and ministering to the poorer members of societies in the manner of his God-son, Jesus, was said to do. Could we be witnessing wisdom in action in a powerful person?

[83] See: Catton 1982, 2009; Diamond, 2005; Homer-Dixon, 2006; Klein, 2014; Lovelock, 2006; Meadows, et al., 2004; Reese, 2003; Tainter, 1988, for a sample of literature on the likelihood of the collapse of human civilization, if not extinction of the species!

tissues. Finally, in chapter 5 I will explore the genetic, developmental, and evolutionary significance of sapience. This will include looking backward at how sapience evolved as a unique human capacity, how it became stunted by cultural evolutionary forces, and then observe some speculations about what might be in the future. This latter subject is motivated by the concern that modern humans are, in fact, inadequately sapient for the very world we have created from our own cleverness. Human beings will need to evolve a greater sapience in order to have a wiser species if there is to be a human presence in the far future of the Earth.

Chapter 2 – Making Decisions: Relations of Intelligence, Creativity, Affect, and Sapience

Being in the World

The world we live in is complex, dynamic, and forever changing. To be in the world means taking appropriate actions in whatever situation we find ourselves. Those actions will then affect the world in some ways and the world will then respond. The appropriateness of actions means that how the world responds will support our future existence and provide us with the resources we need to survive and thrive. What we call 'intelligence,' or rational decision making is simply inadequate to process the information associated with perceiving the state of the world, conceiving how it got to the current situation, choosing appropriate actions, and anticipating the results of those actions far enough into the future to gain advantage. The human brain has evolved several interrelated forms of mental processing besides mere intelligence to synergistically process decisions.

All mental processes contribute ultimately to the decisions we make for action. Most of those decisions are made in the subconscious mind but even so, a lot of mental machinery is at work to guide their making. Even our conscious decisions are subject to heavy influences from multiple conscious and subconscious processes. Intelligence may be taken to be the general machinery that processes decisions. But sapience, creativity, and affect play major, and generally preconscious, roles in shaping the pathways through decision space that intelligence will take. We are not especially rational, dispassionate, objective beings when it comes to making decisions. Affect, from an historical biological point of view, is probably the most influential and it is definitely the quickest to push our decisions. Creativity can have an effect, if nothing more than introducing some novelty into the process to generate exploratory behaviors. Sapience, in some sense, can be somewhat antithetical to affect in the sense of downplaying emotional responses. But it also acts as a subtle, quiet guide that brings a rich tacit knowledge of how the world works to bear and provides intuitions that can improve intelligent decisions. For example, delaying an action that at first glance seems to have an immediate payoff because one knows that the situation may have hidden costs is possible only if one's tacit knowledge of similar situations has prepared one to exercise caution. This is an element of wisdom.

Decisions, Decisions[84]

At every moment the human brain is making billions of decisions at many scales of time and effect. In the millisecond time scale, neurons are deciding whether to fire action potentials to send signals downstream to other neurons. In fractions of seconds clusters of neurons decide to fire in synchrony to assert a representation of something at the microscopic atomic feature level

---

[84] A small compendium of works on decision making: Ariely (2008); Burton (2008); Damasio (1994); Gilovich, et al (2002); Griffin & Tversky (2002); Hogarth (1980); Johnson-Laird (2006); Kahneman (2011); Lehrer (2009); Mlodinow (2012); Montague (2006); Schneider & Shanteau (2003); Schwartz (2004)

to that of whole complex concepts. It makes decisions in the second time scale regarding the sequencing of concepts that we know as thought. Speech acts operate in this time scale. Over ranges from seconds to days brain circuits make decisions on the employment of tactical behaviors that presumably put the organism in better situations relative to the rest of the environment.

On a much longer time scale, up to years, the brain makes decisions that are of a strategic nature. This is the realm of sapience. Every animal makes the same basic kinds of decisions that humans make, including some advanced tactical moves that we call intelligent. But humans have entered a new realm of decision types (see below) that involve complex strategic thinking. As far as we know, only humans consider questions like, “What do I want to be when I grow up?”

The brain is a decision taking machine. Unlike our mechanistic and deterministic machines (e.g. the computer), the brain deals with stochastic reality, that is an uncertain and noisy world. It is stochastic in its operations and is matched to the way the world works in this sense. It takes a stochastic machine to know a stochastic world. But the job of the brain is to make decisions that guide motor outputs and behavior that positions the organism for successful living. Organisms that are more complex live in more complex environments and have more complex and variable behaviors to choose from. Humans have the highest degree of flexibility and the greatest number of choices to make. Thus, the human brain decision making machinery is far more complex and sophisticated than any other animal on this planet. The four psychological constructs can only be understood in terms of their role in affecting decisions, specifically, decisions that are accessible to conscious realization. That is what this chapter is about. We will later examine the details of sapience’s role in this process, but here we look at how all four of the major constructs contribute to the making of decisions at multiple scales of time and space.

The Four Constructs and Their Processors

Intelligence, creativity, and wisdom have been characterized by Sternberg[85] and others as related but different constructs. Other psychologists have included affect (e.g. emotions) in this list of constructs. Each of these performs different cognitive functions that interact with one another as suggested in figure 1.1 in the prior chapter. I contend that the major form of these interactions is the effect each has on making decisions. Figure 2.1 shows the schematic relations in which a decision “processor” (explained later) takes in information on the state of the external world and state of the internal world (from the body) that is used to decide on an appropriate action. This information is shared with an intelligence processor, a judgment processor (as the main interface for wisdom), and a creativity processor. These three plus the influence from the affect processor act on the decision processor to help arrive at a decision. Not shown in this depiction is a recurrent signal from the higher brain functions back to the affect processor. That signal

---

[85] Sternberg (1990).

originates in the part of the brain responsible for sapience; it is more than just judgment per se. It can act to down modulate the emotional responses arising in the affect processor.

In using the term 'processor' I realize I am risking making what the brain does sound computer-like. I also risk leaving the impression that there are circumscribed 'modules' in the brain, each responsible for particular functions, as one would find in a man-made machine. But the term actually refers to the functional performance that arises by the interactions of many and often widely distributed areas of the brain, where often time a single area can participate in the processing of these different functions. Intelligence, as a process, for example, is the result of many brain regions working in concert to produce a final result. There is not a single area (module) in the brain that does all of the work of intelligence[86]. The same is true for the other constructs as well. So, when I use the term processor, just realize that I am using it in a strictly functional sense. In addition, as noted above, the brain is a stochastic process that works nothing like a computer. It computes but by a very different set of principles than those operating in an electronic computer in which actual physical modules can be recognized as processors.

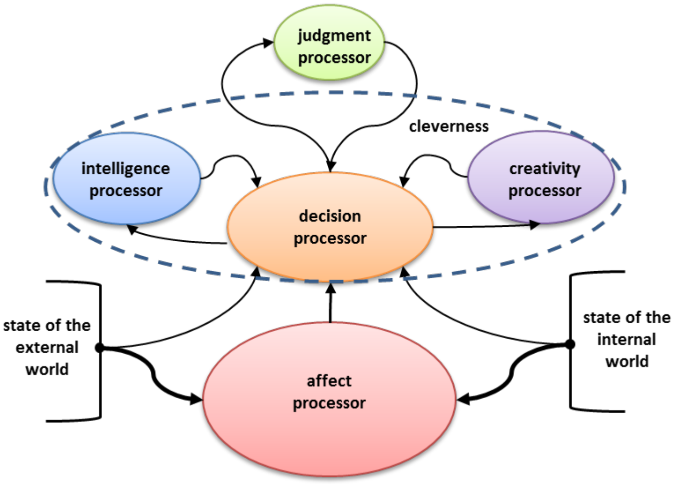


Fig. 2.1. A decision processor must take in information regarding the states of both the external (environment) and internal (body) worlds. The affect processor supplies emotion-laden motivation. But the three cognitive processors, intelligence, judgment, and creativity, act on the decision process as well. The cognitive processors draw on memories (not shown) to produce their effects. They are also kept informed of current state information as well as the progress of the decision process, hence the reciprocal arrows between the decision processor and the three. In this model communications between the three are assumed to go through the decision processor, however another model might well allow there to be direct communications between all three. The dashed blue oval represents the tighter coupling between intelligence and creativity that I call 'cleverness,' our inventive problem solving capability (see section below).

Each of the constructs described in psychological terms is the result of brain subsystems that can be considered as sub-processes in an overall cognitive process that produces behavior. Thus I use the term processor as a logical functional entity to identify the cooperating mechanisms that give

---

[86] See, for example, the descriptions of brain networks in Seung (2013) and Sporns (2011).

rise to a particular form of cognitive activity, i.e. intelligence, judgment (wisdom), creativity, and affect. Note that many psychological theories about these constructs include explicit components that have specific functional jobs. Sternberg's 'componential' theory of "successful intelligence," for example, posits three major sub-processes or systems that comprise the intelligence construct[87].

In this chapter I will examine what each of these processors is doing and how they influence the decision process. My approach is based on the extensive literature but also varies in certain details that will be noted. I have set the framework up around the concept of making decisions. Here a decision is a step along the way from a starting condition, say after recognizing a problem, to a goal condition, say solving the problem against some identifiable criteria. All of the mentioned constructs as processors are contributors to the making of a decision in their own ways and according to the nature of the specific decision point (step).

The world around us is dynamic, that is constantly moving, and forever changing in its characteristics and their statistical properties, or technically, *non-stationary*. All autonomous agents living in such a highly dynamic and non-stationary environment[88] must continuously make decisions about what to do next, given a particular situation. Simple animals living in less complex environments have fewer decisions to make. Humans appear to live in the most complex, dynamic, and non-stationary environments imaginable; this despite their every effort to construct a predictable, convenient environment with technology! The evolution of intelligence and creativity is the development of neural computation systems increasingly able to handle more complex environments requiring more elaborate decision making processes and more memory capacity[89].

For humans, only *some* decisions are made consciously after some form of mental analysis of the situation. Some decisions are the result of intentional thinking in this way. But the vast majority of decisions are made subconsciously or pre-consciously (before conscious awareness). This often comes as a shock to people who are unfamiliar with the research in the psychological and neurological basis of decision making.

---

[87] Sternberg's model includes 1) "metacomponents," which are 'executive processes for planning and monitoring, 2) "performance components," which are those that carry out the functions determined by the plan, and 3) "knowledge acquisition components," which, as their name implies, acquire the requisite knowledge for problem solving. See, Sternberg (2003) page 44.

[88] "Non-stationary" is a technical term used to describe the nature of stochastic (noisy) processes. A time series of measurements of some attribute(s) of the process show jitter which needs to be smoothed out, for example using statistical properties such as the mean and variance. For example climate data include the mean temperature calculated over many years. In a stationary process the mean would stay the same over this longer time scale. However for non-stationary processes the mean can actually change over time. For example it can be trending up or down. The current situation with global mean temperatures rising is an example of the climate being a non-stationary process.

[89] See Geary (2005) for a very comprehensive account of what is known about the evolution of the neocortex in mammals and humans.

Additionally we need to understand the scale of the situation demanding a decision with respect to both space and time. Most decisions are trivial, here-and-now-what-do-I-do types. Many are routine. We go through our daily lives hardly thinking about what to do next because we have habits that serve well under ordinary circumstances. It is when the circumstances are not ordinary (or common) that we need to engage conscious thinking to come up with a choice of actions.

Then there are the decisions that our limbic brain makes for us and before we are even aware that a decision is needed. The limbic system is the ancient portions of the brain (mid brain) that handles early sensory perception and motor signal relays to the body. It also is involved in automatic responses to semi-complex stimuli that have semantic value to the wellbeing of the animal. The conscious brain experiences these responses through various forms of emotion but only after the limbic brain, in the subconscious, has jumped into action.

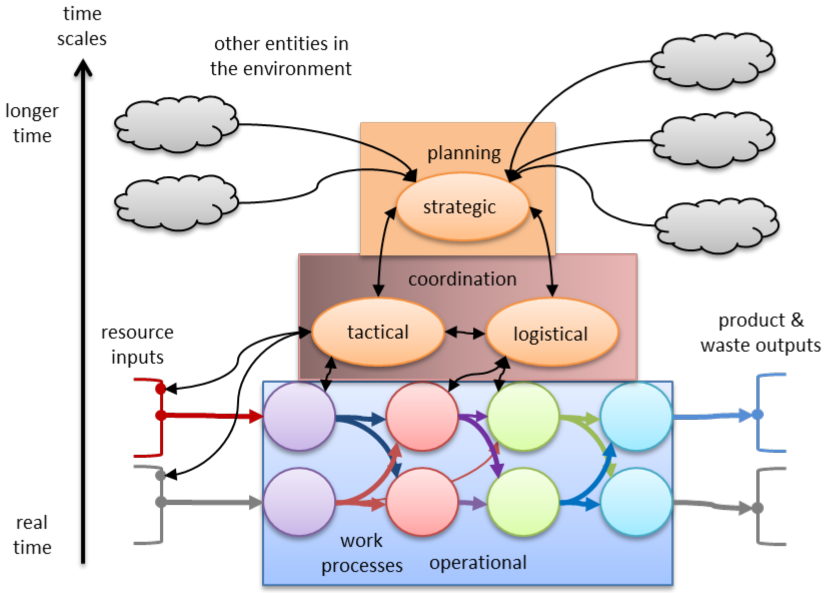
In what follows I want to identify each of these functions' role in decision making in general. All of the functions of the brain are integrated to produce the final behavior of the animal. But it helps to understand what is going on in the brain by teasing apart the core functions to see their primary responsibilities in the overall scheme of decision making.

Types of Decisions

In chapter 1 I introduced the hierarchical cybernetic model of governance with respect to the nature of sapience. Here I will provide a semi-formal presentation of this model and it is more deeply covered in Mobus & Kalton (2015, Chapter 9). Here I want to develop this model with respect to the nature of decisions. The decisions that an agent has to make are related to the regulation of activities in which it is engaged. They are hierarchically organized by types and time scales over which they are made (figure 2.2 below). There are basically three levels or layers of decisions in this model. At the lowest level decisions are about how to control the fundamental work processes in which an agent is engaged. Work processes are the various subsystems within the agent’s system that, working in concert, transform resource inputs into various products and behaviors. Each process has ‘built-in’ controls that require real-time decisions to keep the work flow going smoothly and properly. In simpler systems (e.g. single cells) these sub-processes will be found to cooperate with one another and directly communicate with each other to achieve this cooperation. However, in more complex systems, with many work processes, there needs to be some form of coordination control imposed to keep them working together for best effect. The coordination layer keeps the work processes operating optimally (logistical decisions) and the whole system working with the entities in its environment with which it interacts directly (tactical decisions). When a strategic/planning layer is present (e.g. in mammals) it is responsible for observing other entities in the environment which might have a causal relation with the interaction entities that might provide clues for anticipation of changes that the tactical coordinator would not have available.

All complex adaptive systems have the bottom two layers in this model. In the case of animals, those below the mammals have little need for a planning level or strategic decision making. Evolution has determined the strategies that they use, effectively the totality of their behaviors with respect to their eco-niches. Mammals, with their neocortices, are capable of learning and adapting tactical behavioral sequences that result from life experiences. Humans have evolved a greater capacity for planning and the ability to make strategic decisions that cover much longer time scales.

As you will notice, this hierarchical model applies to many kinds of complex adaptive systems that may be described as purposeful. It applies to organizations and to governments[90]. It applies much less to loosely organized systems like farming communities and ecosystems. The former depends on the cooperation among community members and open markets for distribution of goods and services. Such communities may transition to purposefulness if they develop some kind of regional competitive advantage, say in growing a food stuff that is traded with other communities, that requires more coordination among the farms to maximize the benefits of the advantage to the whole community as a system. Ecosystems never really develop a purpose as such. Though complex and adaptive, they do not develop cybernetic structures of the kind I am talking about here.



---

[90] Mobus (2015, 2017).

Fig. 2.2. The hierarchical cybernetic system (control and regulation) is found in all complex adaptive systems with the strategic level being found in complex adaptive and evolvable systems.. The thin black arrows represent the communications (mostly recurrent) that are needed in order for the system to work as a whole. The model presented here is basically the same as in chapter 1.

Operational Level Decisions

The most fundamental level of decision making involves low-level operational behaviors. These are the genetically given controls such as breathing, heart rate modulation, the chemical milieu of the blood, and so on. Cells and bodies operate on the principle of homeostasis or keeping the general milieu of the organism (cell or body) in a nominal operational state. This level operates in what we call real-time. The organism must respond to changes in its environment in the time scale of those changes. Operational decision work on positive and negative feedback loops (known as closed-loop control) that provide error information to an actuator that can counter the external influences and keep the organism's state in the desired one.

Since operational level decisions are a function of genetically determined mechanism and operate far below the level of conscious awareness, as a rule, we will have little more to say about them. You can find abundant information about *homeostasis*[91], the mechanism for this kind of control in the physiology literature.

Coordination Level Decisions

Metabolism, at the cellular level, and physiology at the organism level are incredibly complex interactions between many different, sometimes competing operational processes. Moreover, the organism as a whole must interact with entities and situations in the environment with which it must seek the most positive (for it) relations. The organism needs to coordinate its internal processes for optimal conditions and its interactions with external entities for maximal fitness. Coordination decisions are categorized in two basic forms, logistical, for internal processes, and tactical, for coordination with the rest of the world. Of course we find that logistical and tactical processing has to be cooperative since the body is responding to external conditions.

Logistical Decisions

The brain, especially the more primitive portions in the brain stem, has to monitor all of the bodily work processes and provide control signals that achieve balance between all of them. For example when the muscles are working extra hard (perhaps due to a tactical decision to run from a predator) the heart rate and breathing have to be increased accordingly. Much of the low-level body function control is based on coordinating the various functions by optimal allocation of resources (energy) to various tissues to keep the body as a whole functioning. As with operational controls, since these are largely automatic we have little else to say about them. However, it is important to note that the higher levels of brain function are influenced by body

---

[91] Or read this Wikipedia article: https://en.wikipedia.org/wiki/Homeostasis for background. Accessed 4/8/2019.

state information, so nuclei in the lower brain do provide signals that go up to the limbic system to provide body state information that might be used in making certain kinds of decisions.

Within the higher levels of brain function, however, there are several logistical decisions that, under the right circumstances, can be made consciously. For example the direction of attention can be controlled by conscious decision. The brain is allocating attentional resources to specific perceptions, concepts, and thoughts under the executive functions of the prefrontal cortex. This will be examined in chapter 4.

Tactical Decisions

How an organism manages to interact with its environment requires it to be able to coordinate its activities with those of entities and situations in that environment on a moment by moment basis, but also on a longer term basis in more complex environments.

There are numerous time and spatial scales for tactical decisions. At the lowest level the organism is merely negotiating the terrain over which it is moving, for example. How high should it lift its legs to get over a rock in the pathway? At the higher levels it involves decisions such as how to get to the water hole without drawing the attention of the local lion. At one of the highest levels for any biological entity (especially the males) are decisions about how to best attract the potential mate (either unconsciously or consciously). For humans such decisions also include things like, how can I impress my boss so I can ask for a raise!

Planning Level and Strategic Decisions

At the highest level of tactical decision making we see a transition into long-term planning for the future and strategic decision making. These arise in the brains of creatures that need to adopt a plan of tactical executions (a strategy) that will achieve a long-term goal or objective. There is some evidence that chimpanzees can plan as much as a day into the future, but only humans seem to have the ability (when used) to plan weeks, months, and even years into the future, assembling tactical programs in a sequence that will yield, if successful, a more favorable position for themselves relative to their environments.

Plans, however, are motivated by longer term goals and objectives. Strategic decisions include setting those goals based on a large number of criteria such as one's current model of themselves, the world that they have experienced, and values. Goals are actually distributed in a hierarchy based on time scales. That is, an overall long-term goal, say, getting a baccalaureate degree after high school with the desire to work in a particular field, should then generate a set of intermediate term goals, for example applying to various universities and looking into financial support or considering a part-time job to help pay tuition. In turn, each of these intermediate term goals should generate near-term goals, such as getting information about the various schools and what is required in the way of coursework that would lead to the right degree.

Each of these goals also generates tactical plans. For example the information gathering goal might be met by visiting the local library to see if they have the catalogues for the schools of interest.

Goals and plans are not necessarily fixed or rigid. Sapience is responsible for their construction but it also can deal with contingencies, uncertainty, and ambiguity. Plans are always contingent and goals are subject to revision as new information may alter one's model of how things really work. A highly sapient person will be constantly re-planning as necessary and is never tied to any specific goal or plan. When we are children we want to be cowboys or astronauts but as we mature we switch to doctors and scientists. Since a sapient mind is always in the process of maturing, switches can continue for much of life.

I characterize strategic decisions as those that determine to what in the environment one *should* attend, to what one knows of one's self, and lead to goal setting from the top of the hierarchy down to immediate tactical plans. Sapience determines where the top of the hierarchy starts, that is how much into the future the goals extend. A highly sapient individual will be thinking very long term and have a strong sense of what they want their situation to be many years in the future. But they never stop evaluating conditions, situations, and especially trends. They never stop re-planning as necessary. And they always have a sense of goals.

In other words, a stronger sapient mind is one in which strategic thinking is ever present.

Sapience is also present in the social mind. Social units such as tribes and nations act as a super-organism, under the best conditions, and those units require governance in the same form of a hierarchical cybernetic system. That means there are all three levels of decisions that must be made by the group and sub-units within the group. As a rule the higher-order tactical and strategic decisions are made by a smaller group of individuals who have the mental capacity. Unfortunately this does not necessarily mean that those making those decisions are particularly sapient. The types of decisions have to be made for there to be a functioning social unit, but that doesn't mean they will be made with any real wisdom. In fact (and as I argue in the governance book) all too often the people making the tactical and strategic decisions are merely the ones with power relations that put them in those positions and not chosen for their meritorious wisdom.

That said, sapience does not restrict itself to decision types for the benefit of a single individual. In concert with the moral sentiments discussed above, along with the notion that each individual constructs models of others and has a motivation based on empathy, sapience in individuals can be the basis for wisdom in groups as well.

Decisions Modulated and Shaped by the Functional Processors

So the central problem is: How do the psychological constructs of intelligence, creativity, affect, and wisdom operate on the decision machinery, especially at the highest levels of cognition. It is not too difficult to understand the importance of intelligence and creativity, for example, in decision processing. This has been the mainstay of psychology for nearly two centuries. But we are only just recently beginning to appreciate the roles of affect and wisdom on the process. In the case of the former we are discovering just how non-rational (in the sense of rational agent theory) a beast we are as a rule[92]. And for the latter, we are yet to fully appreciate the role of higher cognitive functions (sapience) in shaping human decision processing, and its consequences.

Certainty, Ideology, and Decision Taking

The Feeling of Being Certain

Making a decision and then acting on that decision requires the individual to have a sense of confidence that their mental process produced the correct decision. Were it otherwise, they might become mired in doubt and paralyzed in inaction. Such paralysis would be catastrophic for an animal that needs to respond to threats or act quickly on possible rewards. The affect system, in automatic mode, provides rapid responses to situations that do not require conscious thinking and decision taking. But for sentient, consciously thinking beings such as ourselves, where the intelligence process produces decisions based on analysis, with inputs from heuristic sources (i.e. feelings and intuitive judgments) the carrying through with action based on those decisions requires something more. The individual must also have a sense of being right about the quality of the decision. Moreover, for decisions that are strongly guided by explicit knowledge there has to be a sense that the knowledge itself is correct and reliable. If such a feeling were not imposed from somewhere in the affect system then, again, an individual would be faced with indecision and that would lead to inaction. In a stochastic world where one can never be truly certain of anything action is a necessity, even if occasionally wrong action is taken. Sufficiently intelligent and creative individuals can more often than not recover from mistaken actions based on mistaken beliefs, so the damage is often mitigated. But inaction would inevitably lead to the extinction of the species.

Thus, humans have evolved a cognitive capacity for holding their explicit beliefs about the world with a sense of certainty that those beliefs are right. Moreover, that same sense applies to their intuitive judgments. According to Robert Burton (2008) this is a feeling of knowing (whether what one thinks they know is correct or not) that is rooted deep in the subconscious limbic brain. We have to feel that we know, and are thus confident in our decisions, or we would never attempt actions based on beliefs for fear that those beliefs might be wrong.

Certainty has a benchmark in statistics. A probability of 100% means that a predicted event or state will happen no matter what. Degrees of certainty range from 0% up to 100%. In human affairs, given the way the brain works as a stochastic learning machine, absolute certainty is an

---

[92] Human thinking is highly subject to built-in heuristics and biases. See Kahneman (1982); Gilovich, et al (2002); Kahneman (2011).

impossibility. For most complex situations in life, certainty of prediction above 50% is very difficult.

“Pseudocertainty” is the cognitive bias identified by Tversky and Kahneman (1986) wherein people have a tendency to be confident in their own beliefs irrespective of the level of evidential support for those beliefs. In essence they believe they are right in their beliefs! But when tested against the statistical benchmark in various behavioral experiments it becomes clear that their certainty is, itself, a false belief. They make horrible judgment errors.

Strength of Sapience and Three Levels of Certainty

The advent of sapience in human evolution and our $2^{1/2}$ order conscious capacity raised the cognitive situation with respect to certainty in our own beliefs, Burton's "feeling of knowing." A first requirement for an increasingly eusocial sentient animal would be the ability to acquire beliefs from the social milieu itself, especially from parents. Children must first believe that their parents/guardians, their extended families (clans), and bands/tribes are producing, in words and actions, truth about the nature of the world in which they must survive and thrive. Three affect-based additional feelings emerged with sapience to bolster this social solidarity, trust, hope, and faith. These three motivations arose as a basis for collectivising human interactions allowing members of a group to believe in the same things and have positive feelings about how these beliefs underlie the potential for the future well-being of the collective.

But stronger sapience also brought with it a capacity to think more systemically and rely on stronger evidence regarding the workings of the world. The very quality that allowed humans to observe and exploit natural phenomena, such as fire and the invention of stone tools, is a higher version of faith. It is faith that observations are not perceptions of random or chaotic organization of nature. More explicit systems thinking leads to beliefs in reality based on reality itself and faith (a stronger version of the feeling of knowing) in the quality of those observations led to a stronger tendency to rely on observed relations. At the level of judgment it led to faith in one's own intuitions as guides for decisions and action.

Yet stronger sapience leads to a more sophisticated or nuanced version of trust, one that is provisional and always subject to revision in light of new observations. Truly wise people are noted to be able to work with ambiguity and uncertainty while learning from mistakes if made. They accumulate evidence and modify their models (both explicit and tacit) over their lifetimes.

Beliefs as Opposed to Knowledge

In chapter 4 I will be providing a more detailed account of how the brain builds knowledge, that is, models of the way the world works, as a basis for making decisions. There it will become quite clear that what the brain is doing is making a best estimate of the models based on information that it receives over time. Those models are under potential constant revision provided the feelings of certainty just described do not overwhelm the individual’s capacity to openly observe new evidence. Where sapience is weak the socially-motivated trust, hope, and faith, affective, versions of the feeling of knowing can override the sapient brain's tendency to rely on evidence-based trust. This leads to strong biases that result in ignoring (or denying) counter evidence and promoting seeming confirming evidence so as to reinforce the held beliefs.

Beliefs can never be more than mere approximations of true real knowledge (corresponding with the real systems themselves). Some will be better than others. Children will have constructed models that may have elements of reality in them but as often as not the "gaps" are filled in by fantasies. As children mature to adults, presumably their models get better at approximating reality. But there are good reasons to believe this process of piecewise closer approximation falls off logarithmically with age[93]. That is our ability to learn more realistic models shows diminishing returns. Of course the rate of fall off determines how "good" our models become as we mature. Stronger sapience seems to provide those possessing it with an ability to continue learning better approximations of truth even in later years.

Thus beliefs are always a matter of degrees of reality being encoded in our mental models. We should be careful in our use of the word "knowledge" in this regard. If pressed on the matter I prefer to reserve the term knowledge for the products of the sciences, a collective and tested set of models. Of course any individual may adopt the results of science for their own beliefs in how the world works but that doesn't quite mean their beliefs held are still true knowledge. Learning the results of science still involves a lot of interpretive work on the part of the recipient.

Beliefs and individual certitude about one's own beliefs interact in cognition in sometimes dangerous ways. The weaker sapience is, the more dangerous these interactions. It was not always so for humanity. Upon the early emergence of sapience in extended family band social units, these interactions served to solidify the social ties within the group, enhancing the fitness of the group. But in the social structures arising along with agricultural expansion of stationary tribes, and with the fact that sapience itself was still not terribly strong in the majority (see chapter 5) the dangers of those interactions began to become more apparent. That is they are apparent now in hindsight rather than to the people of the time.

Socially-motivated Trust, Hope, and Faith

Beliefs based on socially-motivated affective factors such as trust, hope, and faith, as described above, tend to be organized into systems that we, today, call ideologies. Ideologies are sets of ideational constructs that effectively provide predisposed answers to questions about what the possessors should do under varying conditions. They prescribe what beliefs to invoke to confront new situations. Essentially your parents and your tribe provide you with a readymade set of beliefs that appear to have some kind of internal consistency but more importantly come for people you have to trust in order to survive. The beliefs cover every aspect of decision making

---

[93] Mobus & Kalton (2014), chapter 7, Information and Knowledge, explains the limits to perfect knowledge.

that one needs as long as the external environment changes in ways that challenge the validity of those beliefs.

The strength of certitude about those beliefs is proportional to the eusocial bonding that gives fitness to the group and thus to the individual. Within this level of certainty about our beliefs there are two related sub-levels. These share common features as ideologies, which is why I put them in the same general category, but have different degrees of motivated certainty.

Religious Beliefs

I will address the issue of how it is that newly sapient brains came to entertain religious, or spiritualistic beliefs later. The fact that most people in this world do hold some forms of spiritualistic ideas is hardly contested, even by self-proclaimed atheists. So taking this as a given the question is: What is the nature of these beliefs and why are they held so strongly.

Religious and spiritualistic beliefs come in sets and so are, cognitively speaking, ideologies. But unlike political ideologies (below) most such beliefs are about supernatural phenomena such as gods and devils (see below, A Surprising Consequence). That is they involve ideas about things which cannot be verified by objective observations. There may be perceptual and conceptual experiences that are tied to the origins of such ideas but we now know that such experiences can actually be induced in individuals through magnetic stimulation of the temporal lobes of the brains of healthy individuals and certainly observed in some brain pathologies.

Religious traditions and spiritualistic explanations for mysterious phenomena have evolved over time[94]. They are a cultural phenomenon based in weak sapient psychology. But the evidence suggests that these were key forces in eusocialization of groups, so played a role in group fitness.

However, what some of these more established religious traditions have evolved into with respect to dictating norms of behavior, particularly in regard to the treatments of non-believers, has become dysfunctional for the whole of humanity. Religious and ethnic persecutions have become far too much the norm in the world. Terrorism in the name of some ideas of gods is now a fact of daily life for us all. Religions, with their potential to bring people closer together have instead turned into excuses for wars and genocide. This is, in part, because of their cognitive relation to the other sub-level of certainty based on trust and faith, political-economic beliefs.

Political-Economic Beliefs

There exist today academic fields called political science and economics. Ironically they purport to be social sciences covering the hard-to-quantify realms of human activities in these strongly related fields. To the degree that practitioners incorporate solid psychological and neurological sciences they may come reasonably close to being sciences, but that is because they study the reasons that human beings hold ideologies about group decision making and the production and distribution of wealth. The reason I say "ironically" is that the mainstreams of both disciplines

---

[94] Atranm (2002); Bulbulia, et al., (2008)

are based on explicit models of the decision making process and the wealth creation/distribution process that are largely based on beliefs. These beliefs, in themselves, are based on observations made over history with interpretations that were mostly founded on ideas that "sounded" logical and reasonable. They were, however, ideologies nonetheless.

Unlike in the natural sciences where experimentation and repeatability led to deeper understanding of mechanisms, the social sciences have been plagued by the largely not understood behaviors of the focus of study, human beings. Ethics prevents the experimental approach, and sheer complexity would seem to thwart the exploitation of simple mathematical relations upon which to base models. Rather, political scientists and economists have mostly invented their own theoretical frameworks. In a manner reminiscent of the invention of supernatural explanations for mysterious phenomena, these practitioners have, over the years, crafted what they firmly believe to be reasoned explanations for political and economic phenomena. They have, as best they could, developed more explicit mathematical models of these phenomena in hopes that they will have explanatory power so that predictions of future outcomes might be made and sustained by the real world phenomena. Unfortunately, as recent history attests, many of the assumptions upon which these models are based have proven misleading at best and more often faulty at worst.

Still the practitioners persist in some of their fundamental beliefs, their ideological premises. Political scientists in general and economists on the whole believe in the notion of progress in one form or another (neoliberalism). They believe in something akin to God-given rights for human beings which include unrelenting growth in the extraction of natural resources and production of goods and services without necessary consequences for the environment and human wellbeing. Indeed the prevailing belief is that growth is necessary to maintain if not improve wellbeing. Were these practitioners grounded in the natural sciences they might, perhaps, see the folly of this kind of thinking. They hold these ideas mostly because their intellectual predecessors schooled them to think this way in the same manner as preachers and parents school their children in the goodness of pure faith in the supernatural.

Today the reliance on political and economic ideologies is being played out across the world but especially, it seems to me, in the United States and several other neoliberal, capitalist, so-called democracies. The growing schism between self-proclaiming conservatives and similarly self-proclaiming liberals has now produced a deadlock in governance. Each side is cleaving to a hardline insistence on their ideas as the right ones. It doesn't even occur to them that there could be some kinds of problems that are best solved with an appeal to conservative principles and others best solved with liberal thinking. There is no room for any kind of compromise.

These are the kinds of results from the dominance of trust, hope, and faith-based certitude coupled with what has come to be called "magical thinking." This is the result of the majority of people possessing a low level of sapient cognitive capacity.

Evidence-Motivated Trust

Jurisprudence and science are two realms that depend much more on evidence than ideologies to motivate trust in explicating truth. Most human beings are sapient enough to realize that having the truth of the matter is important for successful living. It is just that most are not capable or not skilled in using observation and evidence to elicit such truth, and so rely on the easier cognitive style of holding ideological (and often unfounded) beliefs that appear to show them the "truth."

Scientists, lawyers, and judges, in general, seem to be not only more intelligent than the population mean, they also seem to understand that wisdom and the finding of truth relies on carefully gotten and examined evidence whether it confirms or invalidates propositions held and advanced (e.g. hypotheses).

Both kinds of professions, unfortunately, sometime succumb to locking onto beliefs in their own advanced propositions. They can be guilty of confirmation bias almost as much as the general population at times. Lawyers have an added incentive to ignore, for example, evidence of the guilt of their client, in favor of evidence that proves their client innocent (and vice versa for prosecutors). Their very purpose in life is to win their cases even when, sometimes, those cases are faulty. Judges are less likely to follow such patterns. They are supposed to be neutral and are dedicated to finding the truth wherever it lay. Scientists can fall in between these two predilections, sometimes falling in love with their own theories to the point of biasing experiments or even, on the rare occasion, falsifying data so that their "discoveries" appear to be validated. Their motivation, aside from pride, can be similar to the lawyers'; they need to win grants to continue to play in the game.

Even so the larger meta-systems of justice and science are designed to provide compensations for errors among the individual components, that is, where they are working. Justice is served as best possible by a system that puts great weight on discovering evidence by multiple parties and bringing it all before the bench and the jurors. Absent rigging and cultural biases such as racism in the Deep South, as an example, our western societies have held faith that the process largely worked. Science, for its part, based on quantitative methods and rigorous measurement and mathematics is probably much more reliable on the whole than the justice system. The latter still relies on things like eye witnesses, who we now realize have very unreliable memories. Science tends to be self-correcting because multiple parties are apt to make multiple experimental tests of particularly important claims (discoveries). A phenomenon looked at from multiple perspectives and through multiple lenses and subject to multiple interpretive minds is likely to be exposed rather well, as the history of science shows.

Unfortunately both institutions have come under increasing pressures and are showing signs of deterioration in terms of the goodness of their outcomes. For the justice system the sheer weight of increasing population and the complexity of modern society is stressing all aspects of the system. Neither judges nor lawyers are able to keep up with the latest technology to the extent it could improve the evidentiary procedures. Increasingly law enforcement is coming under siege due to their increasingly heavy handed approach to fighting crime - or perceived crime. Being one of the prime sources of evidence to the justice system, the truth of such evidence is becoming increasingly unreliable according to many.

Scientists are under pressure to produce results. Young scientists in the United States and other western cultures, under the current regime of staying in research universities, are increasingly being pushed to publish and get federal grants if they want to get tenured. Not terribly different from the pressures some students face at exam or paper writing time, they are tempted to cut corners where they think no one will catch them.

Thus even though the processes relying on evidence-based trust initially involved a level of sapience above the norm, the sheer sizes and internal pressures building up are tending to destroy the validity of the outcomes. Some of those pressures are resulting from the broadening of admissions to the professions, which used to be fairly exclusive intellectual clubs, to include increasing numbers of ordinary sapient students who are told they too can be lawyers and scientists if they just study hard. Indeed many might have the intelligence needed to master methodologies and techniques as well as memorize key facts, but this isn't the same thing as students having natural critical thinking abilities. They are not "trained" to think critically; their level of sapience includes that capacity.

Provisional Trust and Questioning

The third level of certainty in beliefs might better be called uncertainty in beliefs. A highly sapient mind is cognizant of its own knowledge being provisional and containing areas of uncertainty. Wise people are characterized in part by their ability to admit that they are not all knowing, or that their beliefs might be incomplete or even wrong. Somehow, sapience involves a damping down of certitude, which results in keeping open minds well into old age, even when life experiences have reinforced current beliefs over and over again. There is always the possibility that there is missing knowledge that would change the held belief. Even so, wise people will offer their best intuitions in the face of uncertainty. They also know that lack of certain knowledge cannot be an excuse for inaction where action is needed. If they turn out to have been wrong due to faulty knowledge, they learn from their mistakes and fill in the gaps in knowledge.

Ordinary people often think of wise people as the ones with all the answers. They naively believe wisdom is about knowing it all when, to the contrary, it is about knowing and acknowledging ones limitations when it comes to knowing. Sapience involves an additional capacity for self-insight that goes far beyond ordinary self-image. This too will be seen as a consequence of a stronger processing capacity of the prefrontal cortex (chapter 4) and a climb to consciousness above 2 $^{1/2}$-order.

The great advantage of provisional trusting is that it allows one to take decisions and actions based on what one currently holds as true, but allows for one to learn from mistakes or from gaining new information. The wise individual can act to the best of their knowledge but not be bound to always do so based on some static version of that knowledge. One does not need to hold to ones' beliefs if those beliefs are faulty or incomplete. Unlike an ideologue who clings to beliefs even in the face of contrary evidence, the highly sapient individual is able to change one's mind, but not willy-nilly. Some people are known to change their opinions based on the flimsiest excuses (like the herd mentality). Wise people change their views and beliefs after careful examination of the evidence and analysis of the meaning of the differences. They change their models of the world based on grasping closer approximations to reality, not on group-think.

I now switch from the language of psychological constructs to that of functional processors. Each of the types of processors described above have specific jobs to do with respect to influencing decisions at all levels in the hierarchy of decision types just described.

The Basic Relationships: A Functional Model

For much of the history of psychology and the study of intelligent behavior the focus has been on cognitive processing[95], and in particular for humans, rational thinking. The field of artificial intelligence, in computer science, mirrored this focus in its attempt to replicate the human ability to play games like chess. For quite a while the basic belief was that intelligence was best seen in the capacity to win such games. Today both psychologists and computer scientists have gained a much deeper understanding of the realities involved in making decisions. They no longer insist that intelligence, for example, is the result of a deep formal logic nor even a semi-formal logic such as classical heuristics (if-then rules) pre-programmed. Today's artificial intelligence (AI), informed more by real human psychology has swung to more stochastic-like processes[96].

Formally, a decision process is a temporally sequential set of stages in getting from a starting state to a goal state. At each stage the decision maker is faced with a set of options, i.e., moves in game language. Each option is tagged with some kind of 'objective' value that should help the decision maker select the best option. This is often complicated by the fact that the valuation is based only on local information that may lead to a less than optimal global outcome or cause a failure to reach the goal state when many more stages need to be traversed before reaching that state. We can represent a decision network graphically as a tree structure where the start state is the root and each stage is represented by some number of option nodes (figure 2.3).

By 'state' I mean a set of variables representing specific aspects of the body, the environment, and the memory of the agent. Specifically, there will be a set of variables, each of which has an

---

[95] "Cognitive" was once almost synonymous with conscious. But over the decades as it became clear that much of what was going on in thinking, that many people had assumed was "intentional," was actually non-conscious, the term has been expanded to encompass the entirety of brain processes.

[96] For example the use of statistical methods in machine learning and pattern recognition.

associated salience, which can take on a range of values, and each of which can be changed as a result of actions taken by the agent. The set is managed by reducing the dimensionality through abstraction and approximation measures. For example, the agent, as a perceiver, need not specifically observe the position of every hair on a dog's body as it moves, it only needs to recognize hair-ness (color, texture, etc.) and estimate the distance the dog is from the self and from other entities being observed simultaneously. Such a reduction minimizes the computational load on the brain while not losing important information.

At each instance the agent's memory is called upon to recall prior states and prior choices (more on this below). The current states of the world and the body (e.g. motivators like hunger) constitute a state situation that, along with the memory of prior state, becomes the basis for making a decision in the present situation. This is what each node in the tree represents. It can most easily be seen in the playing of board games like chess.

Each node represents the state of play on the board, the position of all the pieces and estimates of the strengths associated with those pieces and their positions. The player whose turn it is will be motivated by a desire to win (state of the body) and will draw upon recent memory or prior plays to analyze the situation and decide on the next move. Games of this sort are handled in consciousness (for the most part) with some contribution from intuitions derived from past experiences. The state information is relatively straight forward in these kinds of artificial games. In real life social interactions, however, the state situation is much more complex, multi-dimensional even with abstraction, and subject to much more noise (ambiguity and uncertainty). Even so the mechanisms for making decisions under such conditions are fundamentally the same as with the game playing.

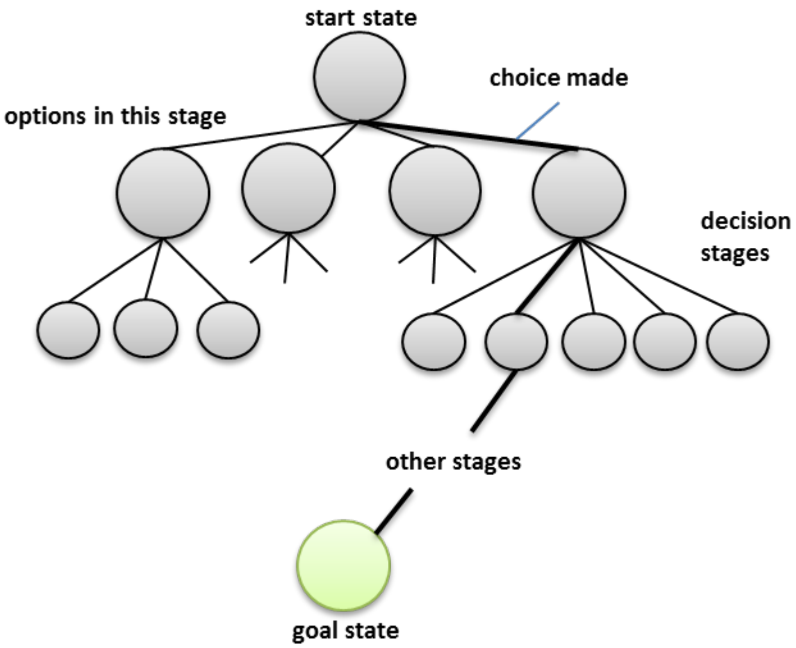


Fig. 2.3. A decision process can be represented formally by a tree structure. The nodes represent the 'state of affairs' and the links represent action choices that will lead to a new state of affairs. In this representation there is an assumed 'goal state', i.e. winning the game (or at least not losing badly!)

Many researchers hold that this formal model is an idealization of what takes place in the human brain. Though an idealization, the model of a network of decision nodes may not be far off the mark in real brains. My own work trying to build a primitive brain with simulated neurons that learn options and the weightings associated with their links has convinced me that the model is very valuable in understanding how decisions are made in the brains of humans and animals. I will cover this in chapter 4.

But the formalization is just a starting point in understanding what is happening. The brain's ability to encode situations in the environment into meaningful concept states (active neural networks) is the starting place for understanding decision making. In effect, the concepts form the option nodes in this tree structure (note that the network is artificially represented as a tree because many nodes may be replicated both within a stage and between stages so as to eliminate cross linkages that would complicate the analysis). The concepts need to be learned by experience and so do the links that lead from one concept node (state of affairs) to another under the considerations that a given action is taken by the decision maker. When a decision is taken (in the sense of a path from the current node to the next state is selected), it generates a physical action that in some sense will change the situation in the world, i.e., the new state is the concept that obtains from the selection and action. For example, in a chess game a piece is moved to a new location and that generates a new state of affairs.

For example, say you see a stranger dog sitting in front of you. Your concept of a dog is activated along with several specific features that are observed in this particular dog. One aspect of your concept of a dog is that if it is wagging its tail it is signaling friendliness and is therefore approachable. You are motivated to pet the dog, because we humans are susceptible to their cuteness. You have to make a decision as to whether or not to approach the dog. In your mind there are, let's say, two possible future states of the world if you do. In one situation the dog responds nicely to your advance and petting. In the other the dog doesn't appreciate it and tries to nip your hand! You already have these two versions of the concept of petting a stranger dog in your memory. If the dog is wagging its tail, you are likely to decide to approach it and attempt to pet it. If it is not, then you have to decide to look for any other cues in its behavior that would signal friendliness or otherwise, and then decide on approaching or not. The decision to look for more cues is, in fact, an action generator; namely it caused you to scan the situation, to collect more data, which is an action. Your decision process in this scenario can be mapped into a structure like the above tree.

There are, however, several unanswered difficulties with using this model. One has to do with the granularity or precision of decision-actions. The world appears to be a continuum rather than a set of discrete states. If it were encoded using discrete representations (i.e. neurons) it would be unreasonable to expect that the precision of encoding would be so great as to consider every little slight change in the world as a completely different state. A possible clue as to how to solve this problem comes from the brain's visual and auditory perceptual systems in which it appears that there actually is a sampling rate associated with capturing frames of visual and auditory information. In other words, our perceptual systems discretize the world for us. Given how most neurons operate — communicating with discrete pulses called action potentials — this actually makes sense. But the topic is beyond the scope of this work. For our purposes we will assume that there is some form of 'just noticeable difference[97]' function operating in the nervous system that discretizes the world into small enough chunks that we can approximate continuous dynamics without great error and yet not so small that the computational load is so high we could never keep up with it.

There is some evidence that the brain operates not on single discretized representations but on small populations of semi-independent representations that collectively provide the 'illusion' of continuousness by something akin to statistical approximations. Again this is beyond the scope of this work, but it should be clear from this argument that the decision stage model above is viable for understanding intelligence.

That is, it can be if we can explain how the link evaluations are instantiated in the first place. The decision processor is faced with selecting one option out of a set of options at each stage. Each possible option carries a value. How this value is attached to each possible link is a matter of learning. As in the case of choosing to approach or not approach the dog to pet it, the link to the

---

[97] A good account of this phenomenon can be read at: http://en.wikipedia.org/wiki/Just-noticable_difference, accessed 12-27-2014.

‘approach’ decision has a positive valence, but only if the dog is wagging its tail. Values may be numerical (i.e. strength of activation) or logical (present or not present).

But what happens when several values are the same and there is no clear higher valued single choice? If indeed all other factors were equal this is where the creativity generator comes into play (see figure 2.4 below). In the simplest possible version, this would be expressed as a random selection. Suppose you had not actually experienced friendly dogs wagging their tails previously, so had not learned to use this cue. You see a dog. You like dogs. You are not totally sure this one is friendly. Nevertheless you make a choice and approach it. Perhaps you are about to learn something you didn’t previously know about dogs and wagging tails. Later I will discuss a somewhat more sophisticated approach to making these kinds of decisions (in the face of uncertainty or ambiguity) that is not technically random per se, but does involve novelty.

Most often, however, all other factors will not be equal. It turns out that the evaluation value attached to each node in the tree is not a simple scalar but a complex function involving percepts, explicit concepts, affective inputs, and the subtle effects of judgment or intuitions. Figure 2.4 shows a functional map of these various components. It is an expansion from figure 2.1. This map reflects to some degree the psychological construct model from chapter 1. The model makes explicit how the various constructs overlap, i.e. how they interact with one another.

Intelligence, in the sense of analytical reasoning, is used by a central decision processor (see below). It constructs a network of options for action based on the current perceived situation (percepts) and explicit concept knowledge previously learned. These are limited in capacity and can never have perfect knowledge. Thus the decision will almost always be made in uncertain and ambiguous conditions. Evolution equipped our ancestors with affective responses triggered directly from perceptions (thick blue arrow from perception to affect) of situations that were inherently rewarding or damaging, with instincts that guided decision making from affect (thick red arrow). This is a very fast reactive system, but is not sufficient to deal with more complex environments encountered by more complex animals. The explicit ‘concept store’ (the majority of the paleo- and neo-cortices) evolved to provide more information to the decision processor. The higher levels of the perceptual system and the concept system interact directly (explained below). The tacit/intuitive knowledge (implicit) system evolved in higher mammals and made its greatest advances in humans. This system supplies tacit knowledge, which is learned from life experiences, to the decision processor in the form of judgment. Creativity is a kind of co-processor that generates novel links within the decision network. Note that there are more links between creativity and other modules. However, these are not as well understood and in the current context of decision making not as necessary for understanding.

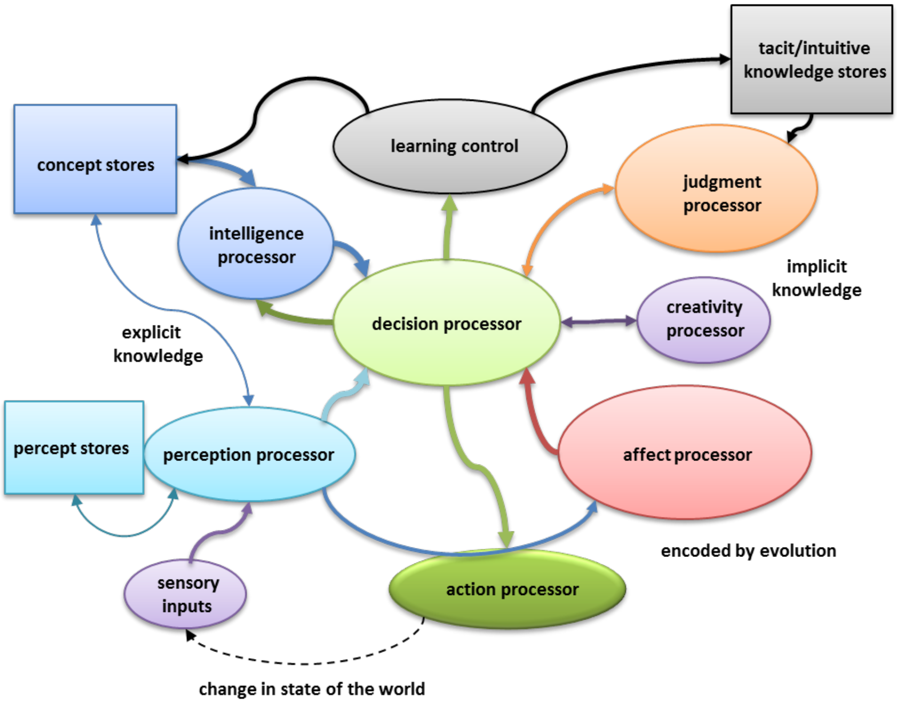


Fig. 2.4. The basic relationships between intelligence, creativity, affect, and sapience are shown in a functional diagram. Included also are memory stores for perception, concepts, and tacit knowledge (models) as well as additional processors in the brain. These include sensory inputs (exteroception only), action processor (e.g. control of muscles), and learning control. The latter is responsible for determining how learned knowledge is to be stored. See text for details. The thin blue two-headed arrow between the perception processor and the concept stores represents the fact that many higher level concepts influence what is perceived and that perceptions do feed upward to concepts.

Shortly I will tease out more details as to how the various processors and stores shown in figure 2.4 work. Here I want to describe the interrelations more holistically.

As already noted the decision processor's job is to traverse the decision tree (network) seeking the best sequence of actions that will lead to a desired goal state for the decision maker. It is, for the most part, a pretty mechanical operation, coming as close to what we normally think of as a computer as you will find in neural tissue. Unfortunately, for many Star Trek fans and as I will explain below, it is not an algorithmic machine as is a computer. The choices it makes at each node are guided by many factors. There is no truly objective function that is guaranteed to produce a correct result. The decision processor takes in whatever facts of the matter it can from the perceptual system (telling it the state of the world and self) and from the conceptual system, the store of explicit knowledge that has been acquired through learning. Due to limitations of size and speed, these systems cannot produce absolutely veridical information. They provide, at best, approximations of how the world is and what concepts bear on that state of the world. Thus uncertainty and ambiguity tend to predominate in almost all decisions. The perceptual system may have gotten it wrong somewhere in pattern recognizing. The conceptual system may not have a good representation of a concept that pertains. All things considered, the decision will not be straightforward. Under circumstances of high-stakes decisions and with considerable effort the intelligence processor may be called upon to invoke more careful reasoning in the form of deductive process. This is hard for people to do, in general, so is not used frequently by the average person[98].

That is where the affective and the tacit/intuitive knowledge systems come into play. It has been known for a while that humans make decisions more frequently based on something called 'judgment[99]'. Psychologists have been working hard to tease out the various forms this takes and how it works. Antonio Damasio (1994), a neurologist researcher, discovered the importance of the role of affective influence on our decision process as an addendum to the conceptual/rational reasoning process of intelligence. He discovered that decision making without affective input was nearly impossible. Patients with a brain defect that disconnected their limbic systems from the prefrontal cortex executive functions had remarkable difficulty arriving at decisions on even very simple problems (like when to schedule the next appointment). These patients got stuck in analysis paralysis, only able to evaluate the values attached to conceptual nodes without the benefit of any affective weightings (e.g. good/bad valences associated with pathways through the network). This brought home the difficulty that a character like Spock (Star Trek) would actually have being a purely rational being (the back story on Spock was that he was half human and so, from time to time, slipped back into the more 'primitive' thinking abhorred by his Vulcan contemporaries! But it was a great plot device to show how humans really do need their emotions and drives, as evidenced by Captain Kirk.)

Affective inputs to the process are important. But they can all too often be wrong. They come from a system that evolved to cause reactions to environmental conditions that could be rewarding (presence of food or mates) or punishing (predator nearby). For reptiles and earlier genera this was really about all that was needed to survive and thrive. We mammals inherited it because it is still often useful, especially when we were totally dependent on survival in the wild. Nevertheless, in matters of complex social nature, simple emotional reactiveness is not necessarily a good guide to appropriate behavior. For more complex and subtle situations the decision process needs some kind of knowledge that is global in scope, generalized and broadly applicable, and can be learned from experiences in life. This knowledge is implicit. It is held, manipulated, and retrieved subconsciously. It enters the conscious awareness only as it affects our decisions, and then only for those decisions that we are being conscious of making. Sapience is the processing system that manages the gaining and using of this tacit knowledge. It provides subtle but powerful inputs to the decision process (see below). Judgments, then, come in various degrees of tacit knowledge/affect ratios. Sapience facilitates the ratio by what we might call a second order judgment on the contributions of these two systems. For example, if the tacit knowledge system does not have a strong input to provide in some situation for which the

---

[98] Kahneman (2011) is an extraordinarily accessible account of what is currently understood about the differences between automatic (patterned and fast) and deliberative (slow) decision making.

[99] Schneider & Shanteau (2003); Hogarth (1980); and Johnson-Laird (2006) provide a good overall review of judgment.

decision maker has no prior experience, it may be a good idea to let affect decide! At other times, especially as the decision maker ages and acquires greater life experiences, the input from tacit knowledge may actually provide a different direction to the decision, in which case sapience needs to override or down-modulate the affective input. I will examine this in more detail in the next chapter.

This has been a general overview of the interrelationships between the various constructs from a functional perspective. I would now like to examine each of the constructs/functions in greater detail and then I will explicate the interrelationships at the micro (i.e., the decision node) level.

Circumscribing Intelligence, Creativity, Affect, and Sapience

In the first chapter I mentioned Robert Sternberg's work on the integration of intelligence, creativity, and wisdom (Sternberg, 2003). In the above I have indicated that intelligence and creativity are involved in making decisions and are modulated by both affect and judgment. In order to put a slightly finer point on the distinction between sapience and cleverness (the combination of intelligence and creativity) and to distinguish between judgment and affect-modulated decision making I want to de-integrate these functions and clarify what each contributes to the final decisions and behavior of the individual.

Intelligence

I'm going to take a somewhat non-conventional approach to describing intelligence in that I will limit it to information processing that can be recognized as reasoning. Psychologists may tend to lump various kinds of behavior into a category they call intelligent when it produces reasonable results[100]. But much of our behavior is the result of pattern recognition of inputs and patterned outputs. It is reactive and learned over time through experience. Procedural knowledge, such as how to ride a bicycle, is implicit knowledge of this sort. That we are able to learn to match input patterns to output patterned behaviors is not really what most of us think of as intelligence at work.

Information processing involves taking in raw data, extracting correlations in both space and time and detecting patterns that have semantic content. Next those patterns and their semantics are used to generate sets of options for action. If the options can be weighed (salience or importance) directly, that is local information is sufficient to make a choice, then the machinery of intelligence can do the weighing and select the 'best' option. More often, with complex and fast changing patterns the number of options generated is great and the weights attached may not be distinguishable purely from local information. Choices are ambiguous and outcomes are uncertain.

Roughly speaking, intelligence is responsible for the more rational approach to problem solving. Rational, here, includes inductive and abductive reasoning processes, not just deduction (at

---

[100] Gardner, 1999l; Sternberg, 1990 – chapter on intelligence

which humans are actually not very good in general). Given the information at hand, the machinery of intelligence assembles the components of interest in a decision.

Other factors which are often attributed to intelligence are things like memory capacity, speed of learning and recall, accuracy and appropriateness of encoding memories. These are the characteristics that can be measured (more or less) in tests (e.g., IQ). It might be better to restrict the concept of intelligence to the notion of rational decision processes and collect these background capacities under a general heading like 'memory management' competency. Doing it this way might help identify functional aspects of intelligence.

Fundamentally intelligence is the ‘rational-like’ processor based on an algorithmic-like process of computation. I say algorithmic-like because, strictly speaking, it is not really the same as an algorithm executed by a computer program. Still it has similarities that make it count as a form of computation that follows rules[101].

The job of the intelligence processor, however, is not to make the decision but to assemble the relevant information regarding the current state (node) in the decision tree so that the decision processor can make the right choice[102]. This is not an easy job. It involves searching through the space of memories (concepts), pattern matching where appropriate, and manipulating both concepts and relations among concepts in a rule-like way to produce a rational value on the various choice options. It is basically a memory retrieval function that organizes the memories in an appropriate network structure. The rules are essentially hard-coded into the neural circuitry of the neocortex and its associations with sub-cortical structures like the hippocampus (a kind of switching station and clearing house for memories). They include, for example, the way in which neurons encode causal relations between percepts and concepts. This subject will be covered in chapter 4, so I only mention it here to put it in the context of what intelligence does with respect to the decision process.

Rationality, when it can be invoked to guide decision making, is essential to making the ‘best’ decisions. But it relies on the accumulation of veridical explicit knowledge and is bound by rules for manipulating that knowledge. Unfortunately, not only do we humans start out with very little in the way of explicit (and veridical) knowledge, most of us do not really accumulate a sufficiently vast store of it over a lifetime. We either don’t encounter the instances in sufficient quantities, or our physical capacity is limited such that intelligence is not enough when we are younger, or is inadequate for all circumstances as we grow older. Below I will discuss the nature of built-in, affect-based influences on decisions when intelligence alone fails to provide the answers. But for now I want to discuss intelligence’s close ally in helping to make “nearly” intelligent decisions; in fact, these are decisions that many behavioral psychologists count as

---

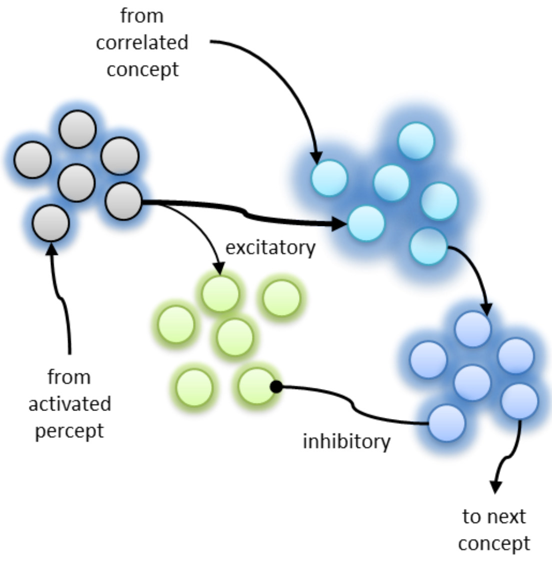
[101] See chapter 8, section 8.2.5, Biological Brain Computation (page 331) in Mobus & Kalton (2014).

[102] In fact the intelligence processor makes decisions internal to selecting relevant concepts. These are decisions on a smaller scale and are rule-like, i.e., are more like algorithms. These decisions are not accessible to introspection directly. The decision processor discussed here is at a higher level.

intelligent simply because they almost always lead to appropriate behaviors. But, as it turns out, they are not really rational, so in my use of the term, not truly intelligent. What sort of behavior do we usually count as intelligent but not necessarily rational?

The most basic challenge of a living organism is finding food (or other resources) in an environment where food sources are not distributed in space in an orderly fashion. René Descartes said, “I think therefore I am.” I would change this slightly: “I need to eat, therefore I think”[103]. Below I will show how thinking, and intelligent problem solving really require creative processes to work under a larger range of problems than can be solved by intelligence alone. But for now I will restrict the discussion to the above description of specifically what the intelligence processor does in terms of finding the “best” path through a decision tree.

The process is based on the way in which causally correlated concepts are encoded in the neocortex. Essentially each concept can be logically viewed as a cluster of component sub-concepts that when any one component is activated, will co-activate its neighbors in concept space so that the whole cluster is activated in synchrony. When such a cluster is successfully activated, say by the activation of a lower-level perception, it then sends signals out to other clusters. Those communications pathways exist because the concepts have been causally associated in past experience; they have been learned as going together in a temporally ordered way. Some signals may be stronger than others, and some may be inhibitory. Figure 2.5 shows a map of several related concepts that are sequentially activated through the learned communications links.



---

[103] This was my basic thesis in Mobus (1999).

Fig. 2.5. The grey concept cluster to the left is first activated by a perception. It, in turn sends activation signals to both the light blue and the light green clusters. However, the light blue cluster is also being activated by another concept that is causally correlated with the grey cluster so it is strongly activated. It then sends an activation signal to the dark blue cluster which gets excited. And finally this cluster sends a learned inhibitory signal to the light green cluster damping down its activity. In this way the sequence proceeds from the perception through the blue clusters and out to some other concept cluster further down the chain. The green cluster does not participate in this particular chain of activations.

The figure basically shows the kinds of rules involved in intelligent selection of relevant concepts. The links are learned (as are the concepts themselves). This is why those with stronger capacities to learn, in general, are also more intelligent. Their neocortices become elaborate networks of connected concepts, both excitatory and inhibitory, which collectively constitute the agent's models of the world. Waves of activation through these networks are the 'running' of the models. In chapter 4 I will go much deeper into the encoding of memory traces and how learning causal relations occurs in the first place.

Creativity

Much of problem solving is trying to find a path through a complex web of decision points as noted above. Most real life problems do not admit of a purely rational solution due to complexity, ambiguity, and uncertainty. So intelligence alone cannot do the solving. Moreover, there are times when it is advantageous for an organism to simply try something novel for the sake of exploration. That is, in the sense of evolutionary change, the organism can take a leap from a current situation node to some entirely unrelated part of the tree, or even a completely different tree as a seeming error, like a genetic mutation is to biological evolution. And on rare occasions, such a leap proves to be advantageous in generating a completely new solution.

The subject of creativity is much larger than I will be exploring here. It involves such leaps of intuition, "aha!" moments, and "out of the blue" insights, as are part of the creative process in human experience[104]. These are all fascinating aspects. However, what I prefer to focus on is the role the creativity processor plays in influencing single decision points in choosing the path through a decision tree. This processor is utilized when the intelligence processor is hard pressed to make the path choice. Such will be the case for most of the decision nodes in complex social behavior networks. Some social interactions may be 'instinctive,' e.g., responding to someone else's smile with a smile of your own. But far many more are situational and unique. When a person is young and inexperienced the uniqueness or novelty of social interactions is great. Only over a lifetime of learning (and acquiring wisdom) do situations tend to become prototypical and choices become more automatic. However, to get to that point the brain needs some mechanism for making choices even in the face of novelty and uncertainty of the outcomes.

If there are no clues at a decision node as to which next path to take then a random selection might seem as good as any other. However this isn't actually the case. It is likely that one finds one's self at a particular point as a result of the history of clues based on causal relations, even if

---

[104] c.f. Andreasen (2005); Csikszentmihalyi (1997)

weak ones, which means one has been on a 'right' path at previous nodes in the tree. In that case it would be best to pick a new out path that was not too far different from the direction one is already going (see the example of a labyrinth below). So the choice should not be completely random. Some novelty needs to be injected into the path selection process so that the chances of finding a good path are actually increased, say, over a systematic (straight) path. The reason is that in nature and the brain, resources are often chaotically distributed[105] so that a systematic search would tend to fail more often than not (see figure 2.6A). On the other hand, a random walk search could also fail since it could just as easily produce a clumped search before getting out into the territory (see figure 2.6B).

Life solved the problem of searching for resources in a fractal distribution by the development of a different kind of search mechanism, a “drunken sailor walk[106]” based on a chaotic process for path choice selection. This process is found in a special oscillatory neural circuit called a central pattern generator (CPG)[107]. CPGs are evolutionarily very old circuits in motile animal nervous systems. I suspect that some kind of chaotic CPG (or more likely many) is (are) at the base of creativity in the brain. This is, of course, highly speculative, but as neuroscientists bear down of understanding what circuits in the brain are involved in creative functions, I predict they will find CPG-like circuits! It would not surprise me if someone were to report on such a circuit (probably to be found in the basal ganglia) that modulates attentional search in the cortical tissues of the brain. Such a mechanism would go a long way to explain how seemingly novel but not strictly random thoughts are generated. More research is needed!

---

[105] A chaotic distribution refers to the idea that resources are often found in patches that have fractal patterns. A fractal pattern is one that has a property of self-similarity at many spatial scales. For example the coastline of Norway viewed from space is quite punctuated with bays and fjords. From an airplane window at 40,000 ft. each fjord has many inlets that have mini-fjord-like shapes. From 1,000 ft. each of those looks to have its own smaller set of inlets. Resources might be found in large, medium, and small patches with inter-patch distances between being short, medium, and long. Distributions of this sort are not random in the strict sense. But they are not regular either.

[106] I should probably find a better name for this dynamic. No offense meant to sailors. The fact is that I was reminded of several times when I was a sailor when I had to get back to the ship after a night of indulgence. I’m pretty sure if you watched me from above my pathway looked similar.

[107] See Mobus & Fisher (1999) for a full explanation of this novel search pattern and its biological generator.

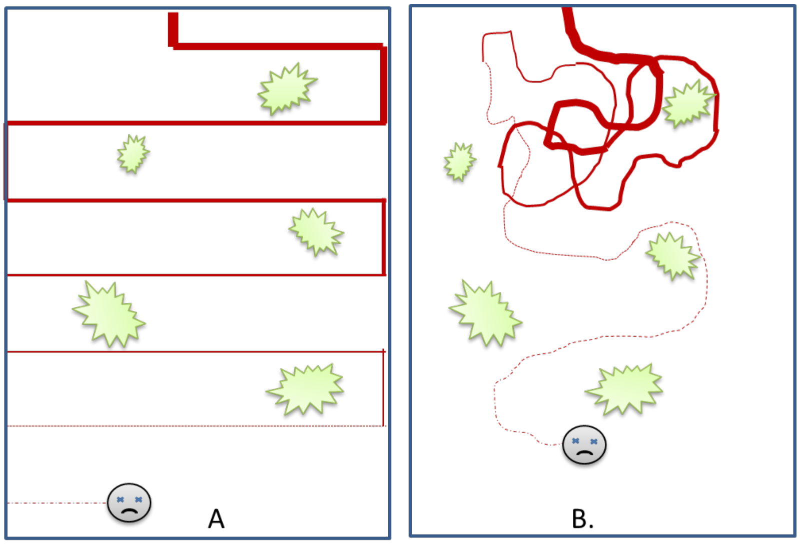


Fig. 2.6. The worst case scenarios with systematic (A) and random (B) searches through resource space would result in the agent running out of energy (declining thickness and solidity of the red lines) before finding its "food." The probability of success in these kinds of searches in spaces with fractal distributions of the resources is low compared with the "drunken sailor walk" search (explained below).

The drunken sailor walk path as depicted in figure 2.7 (with two instances of a search through the same space at different times) can be seen in the foraging behavior of many kinds of animals. In foraging, without other kinds of clues as to where a resource sits, the animal will wander over the terrain in a pattern just like shown below. This wandering is the result of a motor control output generated by a CPG circuit.

In any case, creativity can solve a big problem in searching for conceptual solutions to problems. By breaking ties, or simply causing a choice to be made irrespective of rational processes, creativity helps the brain keep from getting stuck in traps (local minima[108]). The right balance between intelligence and creativity solves the perennial problem of exploitation versus exploration in non-stationary environments. The truly intelligent agent has to strike a balance between the energy efficiency gains from exploiting a known resource and the potential discovery of new and better resources that would come by foregoing exploitation and spending some time exploring. The agent needs to invest some effort in finding new resources because there may be a better source out there somewhere and because current exploitable resources may run out.

---

[108] Simulated annealing is a computational approach to avoiding getting trapped in a local minimum (or maximum), meaning finding a solution that is not the best global solution. See: http://en.wikipedia.org/wiki/Simulated_annealing.

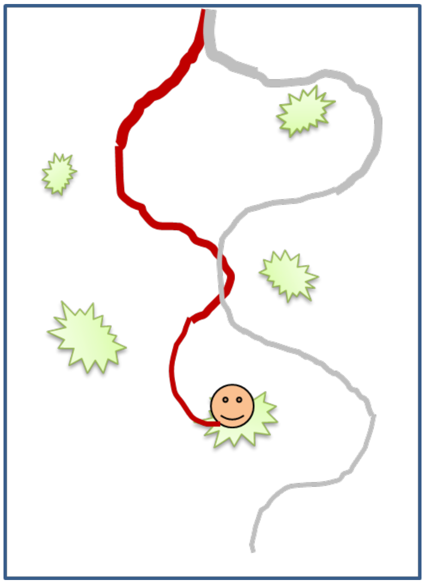


Fig. 2.7. A drunken sailor walk search employs a chaotic form of novelty to shape the paths. Note that it is not totally random but also not particularly systematic. Two paths are shown starting from the same point, one is successful (red) and the other is not successful (yet) over the same energy range (grey). It might yet succeed before the agent runs out of energy if the pattern of distribution of resources continues beyond the lower boundary.

Creativity, in general, is like this kind of search process. It cannot be random because random walks can as easily lead to getting stuck in a local area while continually using up precious energy. Similarly it can't be systematic for a similar reason. It has to be a generator of novel options but must be somewhat constrained by the probabilities of success. Creativity walks at the edge of chaos.

Creativity alone is not sufficient in the realm of decision making. Remember decision making is about goal states. Predominantly creative people are not necessarily concerned with a goal state as such. Intelligence, likewise, is insufficient because of the complexity and uncertainty of choices. But together, contributing their own capabilities to decision making, they are a powerful team. Their combination has been the reason for the success of the hominin genera in evolution.

Intelligence + Creativity = Cleverness

In fact it is difficult to completely separate creativity from intelligence. The two work so smoothly together to produce intelligent problem solving that most researchers do not really attempt to do so. However a strict functional analysis finds intelligence is responsible for capturing information and sifting through options looking for ways to exploit that information to solve the problem. Creativity, on the other hand, is responsible for constructing novel options, for putting concepts together that previously were not related (e.g. the unicorn).

When the number of options available to choose from gets too large the process of weighing and selection would become untenable[109]. It would take too much time to analyze the options relative

---

[109] Schwartz (2004) explains why more options (choices) is not a good thing.

to the time frame in which a choice needs to be made. As we will see below, the more primitive brain (the so-called limbic system) contains automatic pattern recognizers for life and death situations that work to simplify choices we have to make in such situations. Also, our affective brain works to bias choices according to prior emotional experiences in similar situations. These affective mechanisms work to trim the number of choices where emotional content is concerned and so make it easier for intelligence to do its job, with fewer choices to consider.

Similarly, we will see that sapience does the same kind of job, but from the storehouse of learned tacit knowledge that one develops over time. These systems act to keep the number of choices down to a manageable number but only if prior experience can be brought to bear. In other circumstances there is no 'precedent' by which to trim choices and no guidance available to intelligence in how to proceed. In those instances the brain invokes one of several methods to drive what might appear to be a random choice. It conducts an experiment!

This is where the central pattern generator discussed above can come into play. I posit that CPG-like circuits supply temporary weightings on pathways that are not random, but are chaotic or fractal in nature. The pathway through the decision tree might resemble the drunken sailor walk in figure 2.7 but through the neural topography of cerebral cortex and conceptual space looking for something like a resource concept. The search is neither systematic nor random and weightings tend to be strongest for pathways that tend to not deviate too much from the one leading into the node. In the section below, Making Decisions: Putting the Constructs Back Together, I will describe this process through an analogy.

Cleverness is behind the human propensity to invent. Invention covers a wide range. It can be the creation of a new or improved tool, or it can involve a new or improved process (procedure). The motive behind invention is generally always the same; how can one do this job better, faster, more efficiently? And this motivation comes from somewhere in the affect system.

Affect

Antonio Damasio (1994) writes about a patient who had lost the connection between his limbic (emotions) brain and the frontal cortex where decisions are made. One of the most dramatic effects on this patient was the loss of an ability to come to a final decision! His rational brain was intact, but had lost contact with his emotional brain. One would think that this would produce a Spock-like (Star Trek) super rational being. After all, with no more emotions clouding his reasoning he should have been able to make better decisions. In fact he had trouble making decision at all. Damasio has concluded from numerous such cases that emotions, or at least some kind of emotion-based valence (positive or negative) attached to aspects of decisions (the attributes of the world resulting from making a specific choice) act to reduce the size of the decision space[110]. The decision processor can prune out all negatively marked choices and focus on only the positively marked ones, thus making the decision process much more rapid.

The marking comes from prior similar experiences that evoked low-level emotions or feelings at the time and is generally subconscious. This means that decision nodes that were involved in obtaining a positive or negative outcome were marked with that valence by the experience. The decision pathway has to include a kind of short-term memory trace that etches an engram through the network. After an outcome is known, say the intelligence processor detects progress toward the goal, those nodes participating in the successful pathway are marked with a positive valence. If, on the other hand a negative outcome occurs, they will be marked with a negative valence. Those marks are used each time that particular decision path is traversed in future decision making. He speculates that virtually all life experiences are encoded with an emotional response tag that provides this valence factor. Then in current experiences when some choices need to be made, the valence tags can be used by the underlying processor to follow those choices that have had positive outcomes associated with them in the past.

My robot, MAVRIC[111], learned in this manner. It learned to associate a signal representing pain with objects in its environment that it should avoid. It learned to associate a signal representing reward with other objects. These “somatic markers” (using Damasio’s term for the marks) then compelled MAVRIC to avoid the potentially painful objects and seek out the rewarding objects.

Thus, ironically, good decisions really do depend on our emotions and are not the result of pure reason. This is something we inherited from our animal ancestors. Reptiles, for example, have basically limbic brains with a thin veneer of cerebral cortex to handle very simple learning functions. Early mammals had little better facilities. Most of the limbic pattern recognition, taking place in the amygdala, is largely based on genetically-controlled behaviors that proved useful in evolutionary terms. For most of animal evolution these limbic-based decisions (to approach positively marked or avoid negatively marked stimuli) have served well. As long as the eco-niche was relatively simple and non-changing over the course of many generations animals could rely on their limbic system to guide their decisions. It is when the environment changes and drives speciation toward higher use of learned patterns to modulate decisions that we see the cerebral cortices, and especially the frontal lobes, increase in relative mass and importance.

In humans this evolution has led to the preeminent place of cleverness and learned knowledge. It has also led to an expanded role for judgment in guiding decisions[112]. Sapience includes the capacity for down modulating, if not directly overriding, limbic signals. But it also includes the

---

[110] Damasio calls this tag a “somatic marker.” This actually means that the link is marked by its relation to how the body feels at the time of the encoding. See Damasio (1994) for a complete description.

[111] MAVRIC stands for *M*obile *A*utonomous *V*ehicle for *R*esearch in *I*ntelligent *C*ontrol. See Mobus & Fisher (1994).

[112] The affective tagging of links is another way to consider the ‘fast’ processing system that Kahneman (2011) discusses.

monitoring of limbic subsystems in order to provide an affective assessment to the current decision point.

Sapience

Sapience is the fourth and newest tool for living agents to use in making decisions. Put simply, once the environments of evolving animal life became sufficiently complex, ambiguous, and uncertain, cleverness and affect were not sufficiently reliable in guiding behavioral decisions. Something more was needed, something that could work based on acquired experiences to adapt behavior to the more complex worlds. That something was a reliance on learned tacit models of the world against which to judge current situations and bias decisions based on, essentially, what had worked in the past in similar situations. This goes far beyond simple conditioning, as in MAVRIC and animals. It is far more flexible in terms of being able to take into account many more situational variables. And it can coopt creativity to consider alternative models. Sapience is even able to coopt affect to give color and motivation to decisions. This is the basis for wisdom, for an elder who has many life experiences being able to bring those to bear on current situations. She does so not as one recalling episodes, but as one who intuitively knows the right things to do and the wrong things to avoid, no matter how complicated things seem. Choices are not merely tagged with valence based on simple somatic marking; they are filled with the moral sense of right and wrong.

The facilities of sapience are not just based on more or higher intelligence. Sapience is a self-management function that evolved out of the learning and recall control structures in the prefrontal cortex (recall figure 1.5). It is, in fact, the strategic manager in a hierarchical cybernetic system. More than just a learning controller, it is a planning system as well. This will be discussed in greater detail in the next chapter.

Degrees of Sapience

All mental capabilities admit to degrees of competence or power. Everyone has heard of the infamous IQ (intelligence quotient discussed above) which purports to measure a degree of intelligence. It is generally accepted that, however the measure of IQ is interpreted, there is a range of variation in terms of the power of problem solving (quickness of thinking, memory capacity and other factors that make some people better at solving certain kinds of problems). The same appears to be true of sapience as well. It might not show up until later in life, indeed much later, when some people seem to have developed uncommon wisdom about life while others are only marginally wise and still others seem to be just as foolish as when they were teenagers. I will return to this subject in chapter 5 (Evolution) in discussing the apparent distribution of sapience as a consequence of its evolutionary "newness." Unlike intelligence, which is generally thought to have a normal distribution around the population mean level, sapience appears to have a skewed distribution, toward the low end of the scale. This reflects the fact that sapience is a relatively recently acquired cognitive capacity and is still underdeveloped.

A Surprising Consequence

In the next chapter I will explore the various components of sapience in a manner similar to the treatment of the major psychological construct components of mentation in this chapter. There I will be concerned with the ideas related to "levels (or orders) of consciousness" and their relation to various "systems perspectives," e.g. the perspective of a worm vs. that of a reptile vs. that of a mammal, etc. I will be arguing that sapience represents a new(ish) level of consciousness that appears to be transitional between what I will call 2nd and 3rd order consciousness. Following from the recognition of varying degrees of sapience just above, I will introduce the notion of a super-sapience at the extreme high end of the distribution. But more importantly, I will show how a difficult to explain human propensity and resulting cultural phenomenon come about from the fact that the majority of humans have crossed a minimal threshold that makes them susceptible to a kind of thinking that scientists and philosophers have been unable to explain[113]. Yet the phenomenon clearly needs explaining.

In chapter 1 I introduced the four components of sapience that make us unique: judgment, moral sentiments, systems thinking, and strategic thinking. These will be further explored in the next chapter along with a deeper explanation of this surprising consequence. For now I will only mention that the last component, strategic thinking, involves a very remarkable facility, the ability to think about non-actual worlds. *This is what we call imagination*. It is the ability to consider how things might be in the future (a memory of the future) or how things might be if we modify objects or situations (invention)[114]. We have the ability to imagine worlds that are different from what is just in front of us. This obviously has an upside; we can invent and rework our world to our advantage because we can imagine "improvements" in the world, and then go about making them happen. But it also has a downside. Specifically our minds, with not quite enough sapience, can imagine *impossible* worlds. This can provide entertainment (fiction) but it can also provide an ability to deceive others with untruths.

Sapience brings more as well. Alongside this capacity to imagine our barely sapient minds have become aware of a terrible fact – the certainty of our own deaths. All of us share existential angst; we are forced to realize that one day there will be a world without us in it.

And we have become aware of our own consciousness through a facility sometimes called "folk-psychology" or theory of mind (chapter 5), in which we realize that other beings are intentional and causative agents like ourselves. Our brains evolved to be sensitive to agency (because agents can be good or bad for our survival) to such an extent that we have a capacity to assign agency to inanimate objects when certain patterns of behavior obtain. For example a person seeing a

---

[113] However, several cognitive scientists have come close or offered fairly consistent explanations that seem to me to fit in with my explication of sapience. See Atran (2002) for a very good analysis of the cognitive enabling of culturally-reinforced religious thinking.

[114] As I will show in the next chapter imagination is a result of strategic thinking using systems thinking in order to construct new "realities".

rockslide coming down a mountain slope toward them might attribute it to an agent attacking them (an ill-tempered mountain god).

Taken together these cognitive capacities, unchecked by a sufficiently high level of systems understanding (i.e. rational grasp of reality) and superior judgment (use of that grasp) has left us vulnerable and motivated to believe in supernatural worlds with beings, gods and demons who have causal influence over our lives, and a promise of life-after-death of some kind. Even if we are not taken in by beliefs in clearly supernatural worlds, we are still susceptible to believing unsubstantiatable claims about the world. That is we can readily accept ideas that sound somehow right to us even though we have no sources of information or true scientific verification to verify the veracity of the ideas. This is the formation of ideological-based beliefs as explored above (Decisions Modulated and Shaped by Functional Processors).

Further impetus for holding such beliefs comes from the awareness we get from systems thinking that we are but small parts of a larger world. In this transition state from animal consciousness to human consciousness, an ordinary human being has a sense that there is a larger “something” than themselves and their immediate environment that somehow gives meaning to their lives, even if they do not consciously have representations of what that bigger something is or how things really work. That sense, subconsciously present and only vaguely felt, essentially primes a person for accepting ideas that seem to satisfy this feeling. I call this mode of subconscious cognition that is only felt as a sense “spiritistic” (subconscious) thinking. I differentiate this from spiritualistic thinking which involves the actual content of thoughts arising and coming into conscious awareness – thoughts of supernatural powers and entities (which need explanation regarding origin, to be discussed in the next chapter). I also use the term spiritual*ism* to refer to a conscious mode of thought that essentially confirms the “rightness” of spiritualistic thinking, i.e. it affirms that having such thoughts is quite proper even when those thoughts have no basis in physical reality.

The resulting phenomenon in social systems is what we call, in their secular forms, ideologies and, in their supernaturalistic/existential assurance, religions.

Ideologies may be considered a mild form of spiritistic thinking. Consider, for example, the belief that so many people have in the so-called “free market” as a grand problem solver of economic problems. There is something awfully god-like about the “invisible hand” metaphor invoked by Adam Smith. No one can truly explain the mechanism that seemingly causes self-interested economic agents to nevertheless cooperate for the good of the whole society. Smith gave examples of the results of the “as-if” conjecture. Modern neoclassical economists had relied on an explanation heavily dependent on human beings being rational agents (*Homo economicus*), a notion which has been thoroughly debunked by psychology. Non-religious ideologies may not involve impossible realities as such, i.e. they involve seemingly “plausible” realities. But they still persist as beliefs mostly because they seem to promise solutions to complex existential or social viability problems (in a free market anyone can become wealthy). We want to believe they are true because they help us cope with otherwise seemingly intractable problems.

Religious beliefs simply take ideological mechanisms to a new level by admitting supernatural solutions.

Religions are the social constructs which attempt to organize the spiritualistic thoughts into an organized structure or belief system (e.g. dogma), complete with rituals and sacrifices needed to form social bonds with fellow believers. This, too, is a consequence of achieving just enough sapience to become eusocial animals. Religiosity, I define as the degree to which a person is willing or compelled to participate in religious practices and codes. On the surface religion and religiosity seem unique among human behaviors and cognitive styles but they are strongly related to ideologies. There are mental mechanisms that work to preserve beliefs and provide the believer with a sense of confidence in the veracity of the ideologies that are probably inherited from older brain systems. In weaker sapient brains, these work to prevent the thinker from questioning that veracity – a property we call "faith." This can sustain beliefs in supernatural things even in light of contrary evidence (e.g. the theory of evolution explains away special creation but the deeply faithful reject it as a valid theory).

Thus the advent of sapience in evolutionary history, while allowing great improvements in many aspects of decision making, has not yet significantly mediated the emotional attachments of affect, nor broken through to a higher order of consciousness where a person can perceive the world as a whole system and still function well in it. It is the rough, vague sense of there being more to life than just what you see, perhaps the sense of an intelligence, an unseen agent, observing from on high that predisposes people to accepting otherwise unbelievable ideas as part of an alternative reality. I will argue, however, that the trajectory of sapient evolution will eventually produce beings whose sapient strength is such that they will realize that they are that unseen agent after all. They are the one's monitoring themselves. It is, as it turns out, a matter of perspective. Humans could further evolve to 3rd order consciousness.

I will return to this topic in chapter 3 and delineate the components more carefully. Sapience is both a blessing and a curse. We need to more fully understand both aspects and why. Below I want to explore some of the "blessing" side of the equation – how sapience improves the human ability to make veridical decisions.

Making Decisions: Putting the Constructs Back Together

Now that we have dissected the individual components of the mind, the psychological constructs, it is time to look much deeper into how they all work together in the process of intelligent and wise decision making.

Having briefly considered each of the psychological constructs separately in their roles contributing to decision making, we should re-integrate them as a holistic unit–a system.

Where Do I Go From Here?

At any instance an organism has to size up the state of the world around it, those components of the world that can have some impact on its existence, and make choices about what to do next. Even the most primitive cells have to respond to their environments with behaviors that help ensure their continued existence. Most of us might not think of these choice points as constituting decisions since for very simple organisms the number of options to choose from are relatively low and even in cases of some uncertainty in the measurements of state variables the general strength of evidence is enough to determine what actions to take next. The paradigm example of this is the stimulus-response mechanisms so ubiquitous in living systems.

However, a simple stimulus-response mechanism is a single dimensional problem, essentially a single-valued function that doesn't take into account other dimensions, other stimuli and their relations to each other. The more complex an organism is, the more complex the decision process becomes as multiple stimuli have to be taken into account and various stimuli are correlated with others so that the functions of response become multi-variate. In essence organisms are generally faced with complex systems of interacting functions that lead to complex behaviors.

Evolution works to equip organisms with the machinery to make appropriate choices. Simply put, if an organism fails to make such a choice it will cease to exist. If it succeeds, it will pass on the genetic blueprint for the machinery that worked. If a small modification in the machinery improves its choice selection (e.g. speeds it up) then it will tend to outcompete its conspecifics and reproduce a bit more successfully, leading to a species improvement over time.

I will take this long-term dynamic as a given for the moment. What we are most interested in is the situation with human beings who live in a high dimensional world of stimuli and possible responses. Indeed, as I will be discussing later, the human capacity to invent behaviors in anticipation of novel stimuli combinations is what makes us what we are in the animal kingdom. But let me try to motivate the investigation of our unique decision making apparatus with a metaphor.

Into the Labyrinth

Picture yourself in a vast underground labyrinth composed of tunnels and chambers. The chambers are large enough so that, on average, ten tunnels lead in/out, including the one you just came through to get to this particular chamber. Some may have quite a few more, some fewer. The number doesn't really matter for this discussion; I just wanted to give a sense of definiteness to the 'allegory'.

Each chamber is well lit so you can make out any local features. Different chambers have different features, like wall color, or pictures, tables, etc. so that you can identify a chamber if you've been there before and you have a memory. The tunnels are long and possibly twisty so that you can't see what the next chamber has in it by looking down the tunnel. Chambers may contain food, water, a shower, or possibly an ogre that bites. Your life is wandering from chamber to chamber looking for resources and avoiding the ogres.

Here is a view from above some portion of a labyrinth.

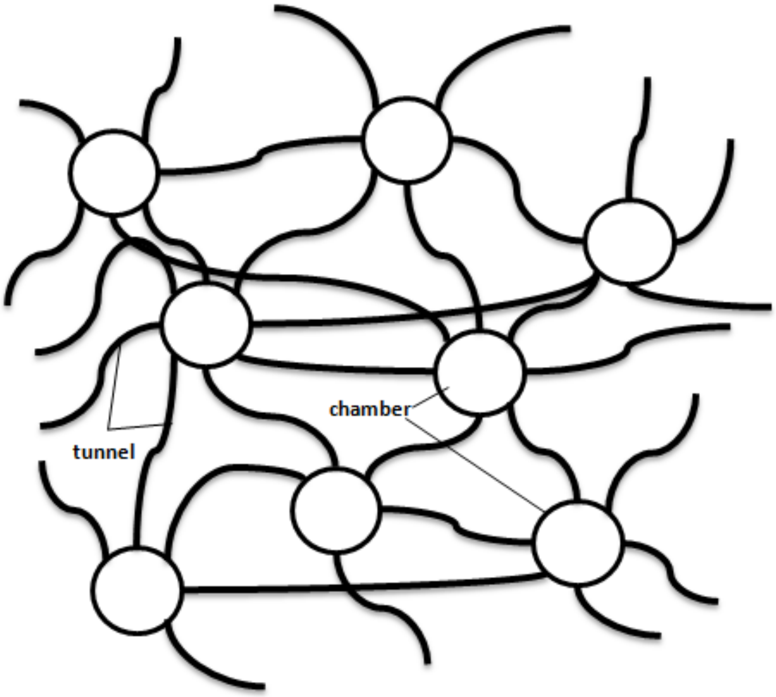


Fig. 2.8. A labyrinth is a model of decision processes. While inside a chamber, the searcher must decide which tunnel to take. How should the decision be made?

Note that tunnels can penetrate in three dimensions so that one tunnel can cross the path of another without intersecting. Thus one tunnel connects only two chambers.

Now picture yourself in a chamber. You got there by traversing a tunnel from another chamber. Currently there is a small amount of food in a plate on a table, which you gobble up because you are hungry. Now, you must decide which tunnel to go into to get to the next chamber. Say there are five tunnels to choose from. Which one should you choose? Just to make things interesting suppose you had never been in this particular chamber before (that you know of). You are still hungry. It would be nice to find something to drink also. So what is your decision?

You might be able to eliminate the tunnel you came in through since you were just in that chamber and left when there was no resource left. So you might be thinking you don't need to go back. But wait. This world is dynamic and non-stationary so it is possible that a resource (mysteriously) has appeared. Of course there is some likelihood that an ogre has entered that chamber too. So your decision looks like a random choice. Just pick one of the tunnels and take it to the next chamber.

Another possibility is that while there seems to be a lot of randomness associated with how the chambers are arranged and maintained, there might actually be some kind of organization involved. Indeed there might be some causal relations between chambers that, if you could learn them, could be exploited as cues. But until you have learned to read the “signs” you still need to decide on a tunnel to take now. Having no other clues to work with you decide to go straight, or, rather, go through a tunnel that is nearly opposite the one you came in through. Why? Well if there is some kind of larger, unseen causal process going on then maybe it is moving “ahead” as it were. So the best guess you can make is to find the tunnel closest to being positioned opposite the one you came in through. It most likely won’t be exactly opposite; your pathway as observed from above might then look like the drunken sailor walk! If this strategy is successful, that is you do end up finding food and drink sufficient to keep you going, then you will buy time to observe various artifacts in the chambers to see if they provide clues. You need a good memory and pattern recognition skills. If the arrangement of the artifacts were similar in other chambers they could be used to more quickly decide which route out to take.

For example suppose you discovered that if the chamber had a table with food, it almost always sat near a tunnel that led to a chamber with something to drink in it! Or perhaps the carpet next to a tunnel has scratch marks left by a transiting ogre. That tunnel might lead to a chamber where the ogre is in waiting. Thus there are possibilities for learning cues that might serve well to guide your travels. Your choices will be made based on a combination of knowledge and emotions (fear of an ogre) and drives (hunger).

This world is a model of the fundamental nature of decision making. We can characterize decision making in problem solving as a sequence of multiple choices. Making a choice at one juncture takes you to a new state. A choice leads to an action that changes your relation to the world, and hence, the world itself. Now you and your world are in a new state and find that you need to make another choice. Amazing as it may sound, all information processing boils down to such a sequence of choices. At first your choices are based on guesses but not necessarily random ones. Then as you learn more about the signs and their causal relations to outcomes (like when you took the route that was marked with the scratches on the rug not realizing that the ogre had made them and then found yourself in the ogre’s chamber; fortunately he was turned the other way and didn’t see you so you went back the way you came.) Before too long your decisions are being guided by those signs. The decisions might even become routine and you forget the original learning of signs, you just automatically, “intuitively” know which path to choose.

Now let's add a bit more motive to the chamber world. Suppose you know that there exists a chamber somewhere in this labyrinth that allows you to exit to the surface where your problems will be over. You have no real information provided in your current chamber so you don't have, say, a sign that says: "This way to the exit". Your task is to solve the problem of getting to the exit and it will involve using your memory and discovering cues (patterns that involve causal relations) that do point in the right direction. Or at least they point in a "better" direction.

The problem as posed is one of searching through a world of options, learning causal relations that can assist you in making increasingly better choices in the future, and eventually (it is hoped) finding a solution to the search problem. The payoff is not only solving this one instance, but using the accumulated knowledge of causal relations to generate general solutions to all similar circumstances; say you are thrust into another labyrinth in the future (Mobus, 1994)!

Structure of Decisions and How Cleverness, Affect, and Sapience Contribute

I'll get back to the labyrinth in a bit and try to introduce some more elements that make it more realistic. Meanwhile I want to delve more deeply into what it means to take a decision.

Figure 2.9, below, is another view of a segment of a decision tree. The node labeled 'current decision node' represents the current state of the world and all of what you can observe of the world around you. Below the current node are the set of choices that you can make and the resulting states of the world that obtain from making one of those choices. The down pointing arrow at the left indicates that these choices are in route toward a goal state — the solution to the problem that got you started choosing in the first place. Say for example that you are a carnivorous hunter looking for prey. And you are hungry! Your goal is to find food and your mode of operation is to hunt for said food. Thus your decisions are based on finding choices that lead to ever closer to your goal state. How should you choose such that the future state of the world, achieved by actions taken after making a choice, gets you closer to your goal? That is the fundamental problem in decision making.

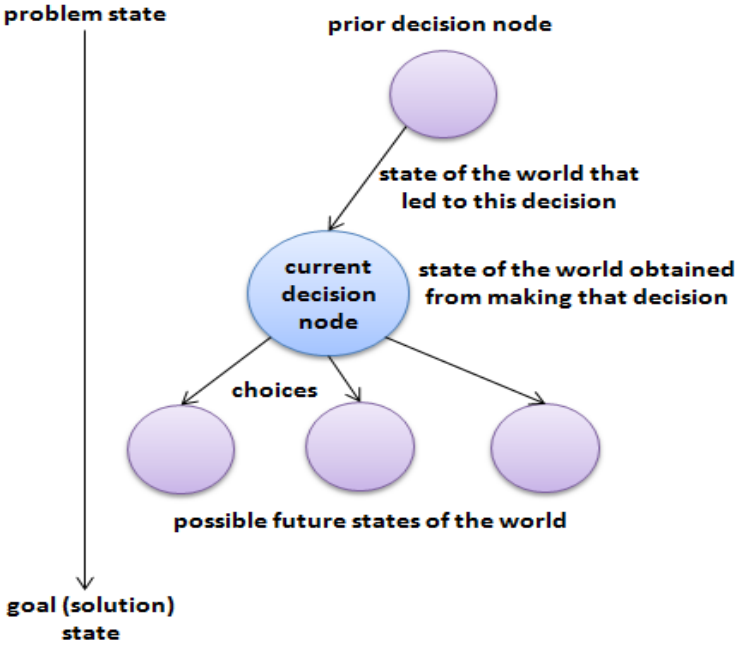


Fig. 2.9. The structure of decisions. At each node in a decision tree there is a set of possible future states of the world that include some state that will bring the decision maker closer to a goal state. Each decision is associated with an action output (not shown) that changes the state of the world. So the problem is to choose the next node such as to take an action that advances the agent toward that goal. Choosing the best option depends on having information at each node that will give a strong indication as to which option to take. In this figure, no information is indicated. In such a case a random choice would have to be made.

There is a whole science of decision making that is devoted to formal methods for solving the immediate problem of choosing. But unfortunately humans do not do very well in terms of thinking formally to make choices (c.f. the delightful Marcus, 2008). No animal does, and we are, after all, animals. Still we must make choices and carry on with our business.

Animals have evolved very clever mechanisms for carrying out heuristic decision making[115]. As with Damasio's recognition, above, that we mark (or tag) our experiences with valences so that we can use those experiences in the future to help guide our choices in similar circumstances, there are a number of pattern recognition 'tricks' that can be used to tag memories of states of the world such that we can use those to provide guidance as well.

Remember the problem with trying to achieve a global optimal outcome based on just local information? This issue is related to a well-known decision process called local optimization? As I pointed out, taking a decision based on a local optimum can lead to a global sub-optimization. The same kind of problem exists with respect to decision points. It may be the case that a prior experience (and its set of attributes and tags) has been coded by the limbic system with a negative valence. So our local information suggests that we should avoid that choice. But it is also possible that what was a temporary negative experience led to a later positive experience of much greater value. This is the classic 'face the danger for the greater reward' problem (it is also related to the reward postponement problem). A beast operating strictly on limbic signals will avoid the local negative situation. But a beast with a memory of pathways through the danger will be able to choose the dangerous selection on the off chance of reaping a bigger reward. Choosing the path with the highest immediate reward is called the ‘Greedy Method’ and there are instances in computer applications where local information is all one needs to make the right choice. But far more often local information is not enough (as in the above example). There needs to be more global information available at the decision point in order to make the right choice with respect to reaching the goal state.

To begin to understand this more elaborate mechanism of decision making in intelligent animals take a look at figure 2.10.

---

[115] See Gilovich, et. al (2002)

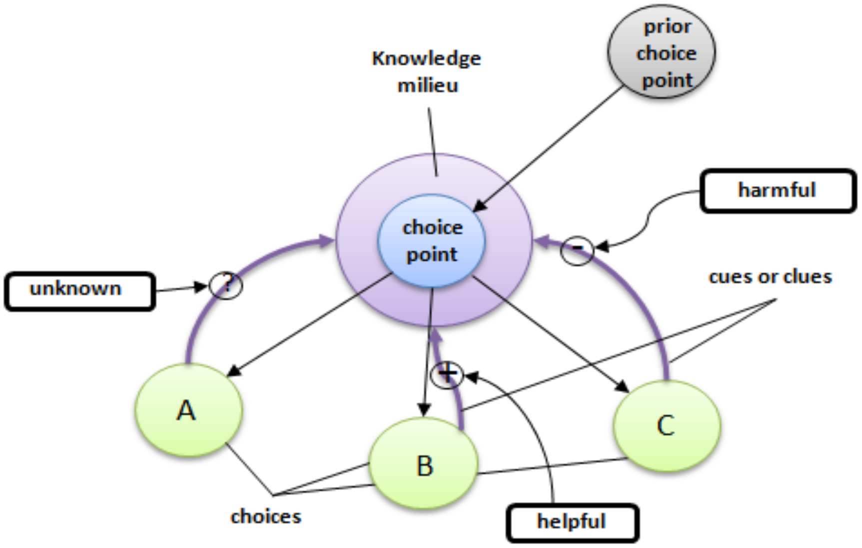


Fig. 2.10. The context surrounding a decision node in a creature with extensive memory contributes to decision making. The state of the decision process is shown where a prior decision had been made leading to the current node. This node is surrounded by a ‘knowledge milieu’ that will help interpret the various aspects of the state of the world to guide in the choice of one of the possible nodes (green). In this figure, the inputs from the affect system, produces the valence ‘tags’ or somatic markers, (per Damasio). Without any other information operating on the decision, the choice would be driven by the feeling that B would be the best one would be made. With a sufficiently rich knowledge milieu, especially from the base of tacit knowledge, the affective valences might be overridden

There is a great deal more going on in this figure (and in the brain) than most people might imagine. Remember these apply to every minuscule decision that the brain makes, especially unconsciously. The 'knowledge milieu' in the figure represents a host of background knowledge that can be brought to bear on every choice (decision point, see figure 2.11 below). This includes the tacit knowledge I've discussed previously, as well as the affective motivations (drives toward a goal state), and facts of the situation, meaning state of the world at that instant. The latter aspect is very confusing to most people who think we apprehend the world through our sensory perceptions moment to moment. The truth is we are conscious of the world through our memory systems. Our prior concepts are more responsible for our present perceptions than we realize. The so-called facts of the situation are really our subjective experience of facts and not objectively determined facts as we ordinarily think of them.

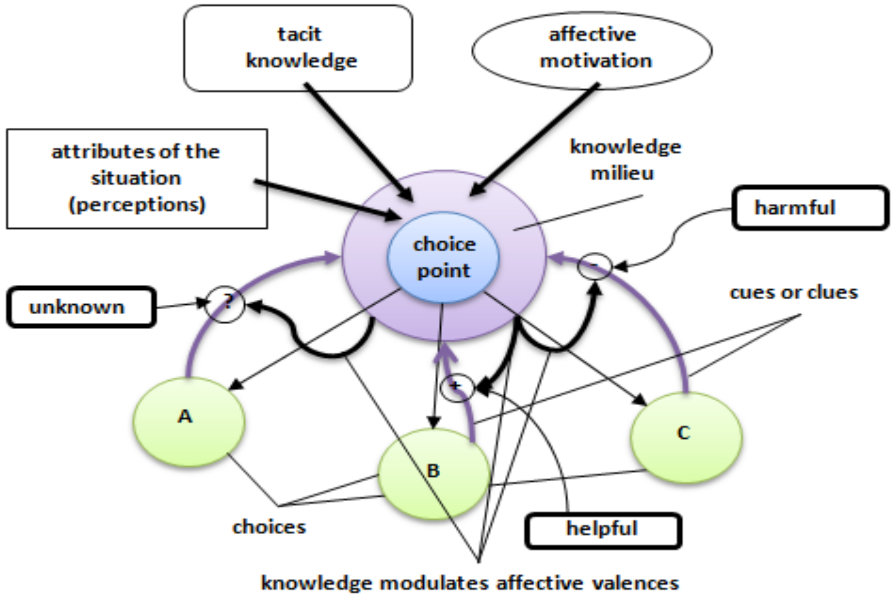


Fig. 2.11. Knowledge encoded in various areas of the brain can influence the decision through the knowledge milieu. At the very least, the knowledge milieu will modulate or override the affective valence effectors so that a decision may be made not on the basis of how good it feels, but on how beneficial it may be later on. Affective motivation feeding into the knowledge milieu includes drives such as hunger that would motivate such an override.

The figure includes the affective valences for choices regarding predicted states of the world. These states are encoded in memories based on experiences, in the past, of similar situations. Or they are conjured from roughly similar experiences. The brain, remember is a magnificent modeling machine that can project future states even when no direct historical experience was had. It does this by using creativity through analogical thinking. It finds past experiences that were sufficiently similar and uses those to estimate the likelihood of future states.

So the basic decision problem is to take all current information and background and affective knowledge into account while estimating the supposed best choice among the options presented. Now just to make things even a little more complicated, we add the role of creativity to the mix. This comes into the picture in three possible ways (at least). Suppose of all the choices presented two or more seem to have roughly the same affective and knowledge-based value. What to do? One obvious approach would be to choose one of the options at random. Little is known about how the brain resolves such choices but there is a rich literature on creativity suggesting that indeed the brain 'makes guesses' in some sort of quasi-random fashion and my candidate is the drunken sailor walk through conceptual space.

Even more interesting than choosing one of several equally attractive options is choosing an option that is not attractive. This could be done essentially in the same way as choosing by quasi-random selection. The choice may be made based on a need or drive to explore (more below). The third possibility is to add an option or two to the set of options that are not really part of the original (intelligent) construction. In other words, the creative brain adds a seemingly unrelated node to the mix of options on an off chance that it will lead to something creative and, hopefully, constructive. This may, again, be in response to a need or drive to explore; what we would call curiosity. However creative choices are made we do know that they can sometimes be completely irrational or, as we call it, outside the box. These mechanisms for creative choice might also go awry in mental diseases causing people to become permanently irrational (e.g., schizophrenic or psychotic).

The thing about creative choosing is that we only positively acknowledge it as creative if it works! Otherwise we write it off as a foolish mistake. Still there is a fundamental need for occasionally trying something not fully indicated by the information at hand. There seems to be a balance in dynamic autonomous systems between pure exploitation of a situation, choosing the best option, and pure exploration, choosing a non-best option to see what happens. Animals show a range of where this balance lies but in general they tradeoff between the two extremes. Exploitation is not betting. It is choosing a relatively certain outcome. But the problem is that there is no such thing, in nature, as a sure thing. The world is forever changing, even if just a little; remember the non-stationary process. That being the case pure exploitation is guaranteed to fail at some point as a life strategy. Every species thus incorporates variation in form and behavior of its members, always exploring, at the edges, new possibilities. This is especially important for a species that is dealing with a highly non-stationary eco-niche. The more complex that niche, the more non-stationary it is. Hence the species needs more exploration of what Stuart Kauffman (1996) calls the 'adjacent possible'. Darwinian evolution is such a long time scale exploration of genotypes, while over the short term species exploit their phenotypes. Animals that rely heavily on experience-driven adaptation (learning to behave!) such as humans require the tradeoff between exploitation and exploration in their individual lives.

Sometimes exploration will fail, or get the individual into trouble. Teenage and even younger boys, for example, seem to be very creative in the ways they get into trouble. Not all creativity leads to opportunities for new forms of exploitation. But when it does we gain new possibilities. And since some old situations that we have exploited in the past may disappear (because of non-stationarity again) it’s good to have new possibilities at hand.

Of course when people we recognize as creative are exploring their domains they are not randomly trying this or that. They are applying judgment to their created options, judgment based on the tacit knowledge they have accumulated over their lives. Truly outstanding individuals are often called geniuses.

And, at last, we come to the role of sapience in this process. In humans, tacit knowledge seems to have the greatest effect on decision making. Strong emotions can rule under the right circumstances, of course. Rational decisions can be made for very small problems (or through the exercise of external formal methods). But the vast majority of human decision making involves social, complex, and long-term problems with significant levels of ambiguity and uncertainty. Our models of the way the world works need to include moral sentiments (concepts of what is right and wrong behaviors), systemic knowledge (how things are interrelated and what effects will derive from what actions), and some number of scenarios for what consequences might be expected in the long run from different decisions.

There is reason to believe that creativity is always generating many more possibilities than we could ever hope to explore, at a subconscious level. One of the jobs that sapience does is help filter these possibilities at an early construction stage so that the processing load on intelligence in making a decision is reduced. Thus judgment plays a role in keeping out the most deviant forms of creative ideas so that what does reach the decision process is at least feasible. In all likelihood, much of the filtering is actually done within the tacit knowledge modeling process itself. If the model breaks down on incompatible situations then it probably dies. Many potential ideas may simply be too weak in activation to make it closer to the decision process. I suspect that the real role of sapience is to modulate the creation and filtering processes as well as provide some kind of final say in what gets through. This would help explain several anomalies we see with creativity in individuals. Some people are very creative in the sense of having interesting and different ideas that make it into the public sphere (through their consciousness). Their sapience may be promoting the generation of these ideas if the judgment is that the creativity processor generally does a good job (e.g. society rewards their creativity). Other people are dull and rarely have a creative thought. Their sapience may be underdeveloped when it comes to such a promotion function. Such people's creation generation might be weak or their pre-filtering may be too strong. Finally, there are a few people who are over possessed with wild creativity that reaches the public sphere that shouldn't. Some forms of extreme sports risk taking may be subject to this effect.

Sapience involves judgment of what to learn, what to attend to in life, how to organize it for most effective future use, how to access it when needed. It involves shaping current decisions in such a way that a good outcome is increased in likelihood; an outcome that is best for the greatest number and for the longest time. Such judgments are applied unconsciously, intuitively, even if they later come to conscious awareness after the fact. Sapience rounds out the toolkit of decision making methods that are feasibly implemented in neural tissues, as far as we know. Animal life started with simple stimulus-response behavior with some built-in adaptiveness. It developed nervous systems to provide the coordination level of hierarchical control as a response to the increasing exploitation of eco-niches by new species and more complex competition. The first versions evolved minimally modifiable reactive programs, automatic pattern recognition, and instinctive response repertoires. These served well for most of evolutionary history. But evolution toward increasing complexity continued to favor larger and more complex brains supporting more varied and modifiable responses. Affective response systems sufficed until evolution happened on the neural networks able to represent models of the more complex environments (cortical sheets with meta-cellular cortical columns). Learning and adaptive behavior took off. Intelligence and creativity were enabled at a new level of sophistication. They had obvious selective advantages and so generated species more quickly able to adapt to changing environments; essentially the birds and mammals.

These earliest 'learners' and 'thinkers' were also relying on a primitive kind of judgment just as they relied on affective input from the lower brain as described above. Simple judgment, as I have mentioned, is a guide to decision making in relatively simple situations but which are more complex or uncertain than rational decision making can handle. In social primates we see an expansion of judgment and integration with systems thinking and moral sentiments such that the earliest glimpses of what would develop into full-blown sapience, apart from raw intelligence. Finally, in the genus *Homo* we see the full basic model of sapient thinking applied to decision making. The domain of decisions humans have to operate in is vast compared to any other living primate. Language facilitates but also expands these domains. The capacity to use judgments to guide decisions has reached a significant level in *Homo sapiens*. But as we are beginning to learn, it is a level only able to deal with the measure of complexity and uncertainty experienced by pre-agricultural man. Moreover, it is far from reliable in the average individual. We all suffer lapses and biases that are genetically mediated. Our native capacity for judgment is limited to problem spaces much smaller than we encounter in the world we have created with our cleverness.

Figure 2.12 summarizes the relationships between intelligence, creativity, affect, and sapience. The latter two are directed at guiding decision processes in intelligence with the help of creativity to force exploration of new possibilities or simply keep the brain from getting stuck. The arrows represent major communications pathways and the direction of influence. Note that sapience (as processed mainly in the prefrontal cortex) has influence over the other three areas as well as monitors them. The biggest influence is over decision processing in the intelligence function. I have put the decisions processor of figure 2.1 back into the intelligence construct to conform more to the standard psychological treatment. In chapter 3 I will actually combine intelligence and creativity as I alluded to above as cleverness and thus abstracted let sapience interact with that single meta-construct. I will then decompose the functions of sapience in greater detail.

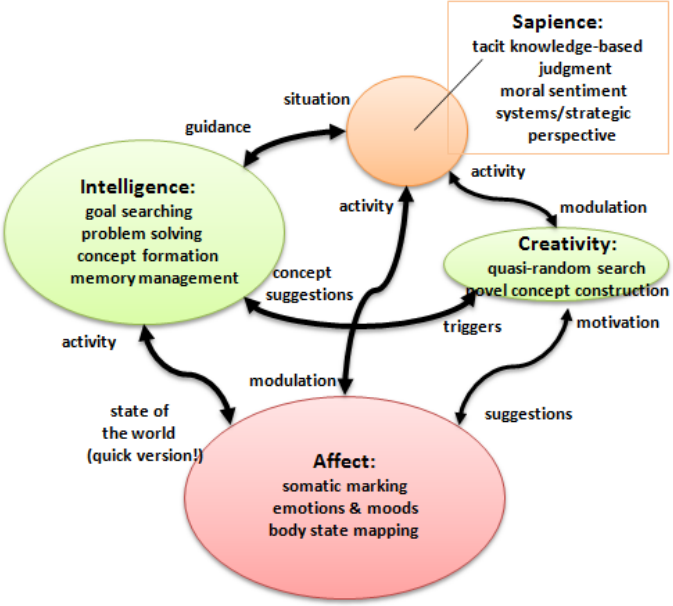


Fig. 2.12. The functional relationships between components of cleverness (intelligence and creativity), affect (emotions), and sapience are delineated. Affect and sapience provide biases and search "control" to the intelligent finding of solutions to problems. Creativity and intelligence interact to provide novelty to searches as needed.

Back to the Labyrinth

Now recall the labyrinth. You are in a chamber and need to decide which tunnel to take in your quest for an exit. Along the way you are also motivated by two basic things, the need for food and water to maintain your strength and the need to avoid ogres that have sharp teeth and consider you food.

I allowed that the chambers could be richly decorated. It turns out that the interior decorator had various but consistent themes that she used in various areas of the labyrinth. In other words there are actual patterns of decoration within chambers such that there is a conceptual connection among chambers. Moreover, the decorator left little clues at the entrances of some of the tunnels to suggest the theme that will be found in the next tunnel.

Your job is to learn these patterns and cues such that you can make a reasonable prediction of what will be found in the next chamber upon following a particular tunnel. If you go through chambers repeatedly, but each time take a different 'out' path you start to build up a model of the labyrinth including a pretty good ability to guess what is in a chamber that you had not been to before. You might also note associations of things like the kind or amount of food and drink with certain themes. You might even discover that ogres prefer to hang out in certain thematic regions, which would allow you to use the cues to avoid those. As you build up a knowledge base of these cues and themes you begin to get better at finding your way through the complexity. You may find an association between a sequence of themes that lead to the exit!

As you first wind your way through the labyrinth you will need to focus attention on details, and reason explicitly about the choices you have to make. But over time, the patterns that you begin to encode in your memory start to provide scaffolding for automatically adding to the knowledge with each new experience. Also, the decisions you make seem to come more easily and automatically, without reasoned thought. You are building a storehouse of tacit knowledge about the labyrinth that begins to guide your decision processing. You are using judgment more than logic to constrain your choices and generally (though by no means guaranteed) make better choices as time goes on.

Life is a very complex labyrinth for humans. It not only involves making choices for one's self, but making choices on behalf of your family, friends, and tribe. The number and kinds of choices are immense, even for more primitive people living in more natural settings. Imagine what it is for cosmopolitan humans surrounded by strangers and dealing with all of the rules, mechanisms, and bustle of modern life!

Conclusion

Life is dynamic and chaotic. Change is always in front of us. To be alive is to be making decisions all the time. Formal methods of decision making (like computer programming) have taught us that problems of reasonable size and complexity can be solved once all the necessary information has been gathered. But once you introduce huge scale, uncertainty, high risk, time constraints and other factors of real life, the ability to solve these problems with such formal methods dissolves. Instead, the human brain uses a variety of non-formal methods to provide approximate or satisfactory solutions under a wide variety of situations.

One method is the quasi-random or chaotic choosing of creativity. Another is the evolutionarily, tried and true, method of affective response. But for higher mammals and especially man, the method that has been instrumental in allowing us to adapt to extremely complex environments is knowledge-based judgment. In mankind we see the first glimmers of judgment based on moral sentiment, systemic and strategic thinking in social contexts. Indeed, I will be arguing that it is the latter element of strategic thinking that turns mere judgment into sapience. In *Homo sapiens*, sapience emerges in a nascent state. But as I shall also argue, that state is not yet able to handle decision making in the world that cleverness has wrought.

If, and how we might manage to transcend this latter point will need much thought.

Chapter 3 - The Cognitive Components of Sapience Explained

Sapience as a System

In this chapter I present a componential model of sapience not unlike that of the mind in chapters 1 & 2. And like the four component constructs of mind, sapience itself is composed of four constructs or subsystems. In this model I identify four basic "modes" of thinking that work together to produce the integrated function of sapience that operated to shape and guide decision processing as seen in the last chapter. Three of the components, judgment, systems perspective, and moral sentiments turn out to be expansions and elaborations of functions that existed in primate brains before humans became sapient. The fourth component, strategic perspective, appears to be of more recent vintage. It is unclear whether earlier hominins (the family *Hominidae* less the genus *Australopithecines*) possessed long-term, wide-scope thinking and planning though there is evidence that *Homo neanderthalensis* did[116]. I will present a model in which this latter component arose in concert with several other facilities, such as higher consciousness, abstract thought, and symbolic recursively syntactic language. These facilities, which are largely regarded as the distinctive qualities of humanness, appear to have coevolved with sapience, driven by the rapid expansion of strategic perspective thinking.

In chapter 1 I described, in brief, the four major components, or functions, of sapience. My early thinking about wisdom and sapience derived from the wisdom research literature in psychology, particularly the work presented in Sternberg (1990). But several lines of neuroscience, regarding the architecture of the brain, in particular the prefrontal cortex and its evolution have also contributed[117]. Chapter 4 will be devoted to this topic. In this chapter I want to provide more detailed descriptions of the four functions and how I think they interact with one another. The view presented here should not be taken as a neurological one. Even if it turns out that the prefrontal areas I think are most implicated in the central function of sapience, these four sub-functions should be understood as descriptions of functions only and not brain modules. My plan is to provide some intriguing evidence from neuroscience in the next chapter. Chapter 5 will provide some evolutionary background on sapience — how it may have been selected for, when and how it developed, and how it might have developed to provide better guidance to cleverness, as related to the neurological substrate.

---

[116] It is still not totally certain that Neanderthal man was a separate species or a sub-species of *sapiens*. The general view as of this writing seems to be that this group should be considered as a different species even though we now do know that *sapiens* genetic complement contains alleles from the *neanderthalensis* genome and vice versa. I should also point out that strategic-like thinking with very limited scope appears to be active in our more distant cousins, the chimpanzees and bonobos. See De Waal (2005).

[117] A small compendium of neuroscience literature for background, tying brain science to psychology: Barrs & Gage (2007); Calvin (1996); Calvin & Ojemann (1994); Damasio (1994, 1998, 2010); Gardner (1999); Gazzaniga (2005); Goldberg (2001, 2005, 2009). The last author is particularly concerned with linking brain and wisdom research.

Recall that the components are *judgment*, *moral sentiment*, *systems perspective*, and *strategic perspective*. I will take these in order. And just as I segregated functions of the mind (affect, intelligence, creativity, and sapience) in the prior chapter, I will attempt to circumscribe the functions of sapience in order to clarify the role of each separately. That is, I will apply boundary analysis to these components. Then I will attempt to knit them together to provide a more holistic view of sapience. As before, the way the four work together makes it difficult to identify a singular boundary for each. The boundaries are 'fuzzy' in the technical sense[118]. However this type of analysis should be useful in identifying functions even if they are shared between components.

As I tried to show in the last chapter, the key to understanding the role of sapience is how judgment interacts with cleverness to guide decision processing. This is how sapience affects behavior. Here I will try to show how the other three components of sapience work together to create the framework of judgment.

Figure 3.1 summarizes the basic interactions between the components of sapience and between those components and the construct I called cleverness (intelligence, including the decision processor and creativity in combination, from chapter 2).

In the figure we can visualize a rough map of how these components interact. The main processor I am calling "judgment" in keeping with the role that it plays vis-á-vis guiding cleverness in making decisions as covered in chapter 2. The judgment processor receives 'state of the world' information from cleverness as well as inputs from systems and strategic perspectives and moral sentiments. It has access to the storehouse of tacit knowledge which it will use to formulate its outputs, primarily of judgments and intuitions directed at the cleverness component. It also provides information to the three other sapience components as well as modulating input to affect. Affect should be seen as driving or motivating moral sentiments, but the latter are incorporated into sapience as explained later.

The judgment processor assembles the relevant tacit knowledge given the state of the world (and inputs from the other three components). It formulates the recommendation to cleverness, integrates any new knowledge into the previous model, and updates the three components as needed. Note that systems perspective and strategic perspective have to interact with one another directly as well. This is part of the hierarchical cybernetic structure discussed at the end of chapter 2. The systems perspective is also in communications with cleverness in order to facilitate some logistical, but mostly tactical controls as needed by the strategic controller. This is

---

[118] The term fuzzy comes from fuzzy set theory (Zadeah, 2000) and its companion fuzzy logic. A system may be mathematically defined as a set of subsets of components and relations. Members of a crisp set can also be members of a separate set, that is, they are in the intersection between sets. Fuzzy members are characterized as being partially in a set as defined by a membership function. Such a function can include a temporal factor. So for example a person can be in the set of all people who call themselves Republicans but then vote for a Democrat in any particular election. Their membership is not completely Republican nor are they, in practice, Republicans all of the time. The boundary of a system like this (the Republican party) is thus fuzzy.

less relevant to the present discussion but is included for completeness. I will further elucidate this loop in chapter 4, The Neuroscience of Sapience, when I cover hierarchical cybernetic structures in the brain.

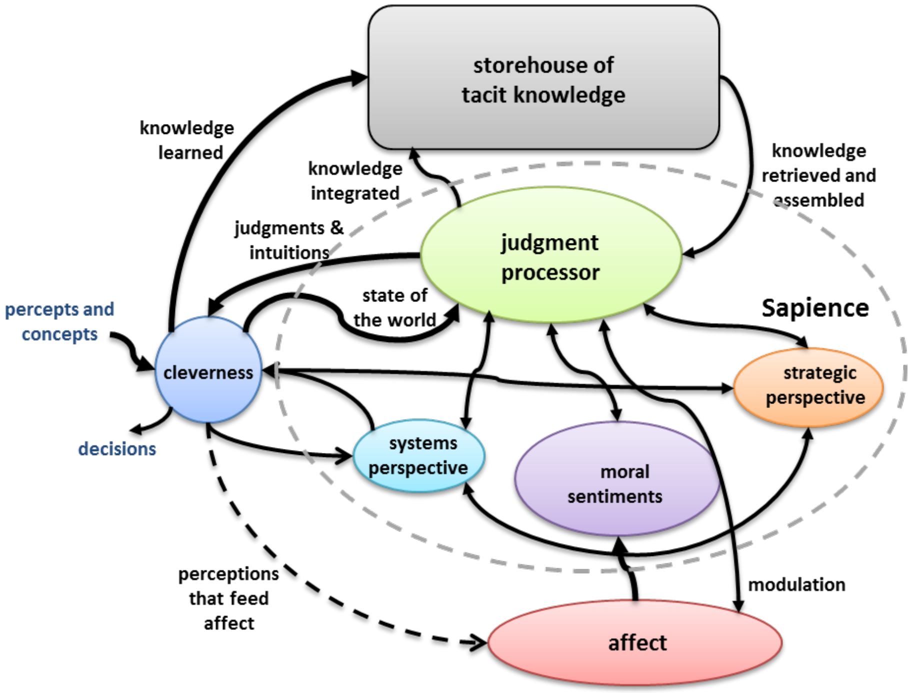


Figure 3.1. Sapience (grey dashed oval) can be viewed as four main, interacting functions; judgment, strategic perspective, systems perspective, and moral sentiment. All of these are supported by the storehouse of tacit knowledge generated by judgment-guided learning (supplied by cleverness). See text for explanation.

Tacit knowledge includes an elaborate set of mental models of how the world works, largely present in sub-consciousness but accessible to consciousness through working memory. These models are built up from experiences with the world over time through a presumed inductive process (see chapter 4 for an explanation of how mental models might be implemented in cortical tissues). Models are dynamic *concepts*. That is they are complex sets of sub-concepts that interact with one another over a temporal frame. One of the simpler models of a small part of the world is a script[119]. Scripts are learned as sequences of expectations of events and actions that are applicable in numerous situations. They are a generalization over many instances of experience in similar situations. For example, a common script given in the cited reference is that of going to a restaurant and ordering/paying for a meal. The general sequence of events is usually the same at any restaurant you go to, though there are variations between fast food and sit down styles. Once the knower has a reasonable script encoded in tacit knowledge that individual

---

[119] Shank & Abelson (1977)

no longer thinks about what to do next. It is automatically generated as part of the knowledge milieu for a decision point in a particular instance of going to a restaurant.

It might be a good idea to pause at this point and distinguish between two types of memories that are sometimes confused, namely between declarative memory and what I have been calling tacit memory (also a form of implicit memory). The former involves memories of actual events, places, faces, etc. The latter is the background knowledge of concepts, categories, and models. Specific memories of places, things, etc., may be the foreground of consciousness while meaning and context, unbidden to ordinary consciousness, is the background. The two memory types are not disconnected, though they are undoubtedly coded differently in the brain and have different locations in the cortices devoted to their management. One form of declarative memory called episodic memory is what most people think about as being 'memories'. For example, they might think that our heads are full of strings of episodes, like frames in a movie film, that somehow all connect, and that is all there is to memory. But memory doesn't work this way. What seems to make episodes encode in memory is a strong relationship to the background meaning, specifically, affective tagging or somatic marking. Briefly, our brains do have an ability to encode sequences of a sort such as the script described above (after all that is what motor programs are). So a specific episode may be encoded given sufficient motivation to do so. As an example we may have a vivid memory of going to a specific restaurant and proposing to our fiancé. We might not remember what we ate or how much we tipped the waiter, but we remember most of the actions and they are often indexed by the order of events (I proposed right after they served the dessert). But the pieces that make up the salient parts of an episode are probably not in discrete packages. Rather an episode is reconstructed from pieces of both episodic and tacit memories that are recalled in a specific sequence.

More importantly, there may be a deeper relationship between episodic and tacit memory in that temporarily stored episodes (from the day's activities) in working memory may get analyzed for semantic content which is transferred to tacit memory. At least some researchers suspect this is what is going on in REM sleep[120]. Episodes having unique relevance might be transferred to longer term episodic encoding during non-REM sleep periods. Much research is needed to say much more about the way memories get formed[121]. But the ways in which memory types are used has been revealed sufficiently to make note of these differences. Henceforth I will mostly be concerned with tacit knowledge (note also that tacit knowledge also involves procedural knowledge, as in how to perform some complex tasks without conscious attention or intervention, like riding a bike.)

While non-conscious mental activity is involved in processing tacit knowledge or mental models to inform judgments, eventually the results show up in conscious awareness. We are conscious of

---

[120] http://en.wikipedia.org/wiki/Rapid_eye_movement_sleep also http://en.wikipedia.org/wiki/Sleep_and_learning

[121] But see chapter 4 for some more specific thoughts about memory encoding.

our decisions even if we are not conscious of how we came to them. This raises a very difficult question regarding why we have consciousness at all. If our decisions are largely a result of subconscious processes (and that can actually include the more mechanical aspects of intelligence processing) then why is there consciousness at all? Why is it needed?

The answer may lie in the role that strategic perspective plays in sapience. I will describe this shortly. But first I want to explore the nature of consciousness more explicitly and provide some definitional background so that the reader will at least know what I am referring to as consciousness. You may have other thoughts about the subject; it is far from a settled concept. However we need a framework within which to work that provides some internal consistency. Consciousness and sapience, as I claimed in chapter 1, are just different aspects of the same mental phenomenon. That is, consciousness, as I describe it, is the subjective result of sapience.

Consciousness and the Mind Architecture

Every normal, healthy human individual experiences being aware of their surroundings and their own ‘feelings’ when they are awake and relatively alert. They have the capacity to communicate what they experience to others, most generally via vocal language. They have the capacity to interpret the vocal reports of other humans regarding their own experiences and understand what the other means to express. Consciousness is most commonly understood to mean this ability to be aware of awareness. There have been significant philosophical as well as practical questions raised about exactly what consciousness is in order to understand ourselves and our mental “states.” We often seek to compare ourselves with other animals in this regard, most recently animals such as the other great apes, dolphins, and elephants that show behavioral signs of being conscious in the above sense.

However in all of our more recent scientific searches for understanding the nature of consciousness, including its purpose biologically and evolutionarily, there definitely appears to be properties that are unique to the human version as I explained in chapter 1. Here I want to go much deeper into the subject as it relates to sapience and demonstrate how the two phenomena are actually part of the same brain development; sapience is the basis of the unique form of human consciousness.

Dangerous Territory

Perhaps fools do rush in where wise men fear to tread. The territory we call consciousness studies is fraught with dangers, intellectual as well as professional (for a scientist). Philosophers have never felt any danger (sometimes quite the opposite) because their job is to simply raise interesting questions about the phenomenon. They don't have to explain how it comes about[122]. René Descartes was content to just declare, “*Cogito ergo sum*,” and call it a done deal.

---

[122] Though Daniel Dennett (1991) seems to have been compelled to attempt an explanation. He even provides some testable hypotheses which put him dangerously into the camp of science!

Nevertheless the subject cannot but intrigue the scientist who contemplates how the brain works. After all, the brain, in its operation, produces mind and minds experience consciousness. At least I, like Descartes, think I do; the rest of you may be zombies for all I really know! This phenomenon, as I just described it, is extraordinarily problematic. What exactly does it mean, for example, to "*experience* consciousness?" What is doing the experiencing? Does this mean that consciousness is experiencing itself? If so, then we are led back to having to explain what consciousness is and how it goes about experiencing. I suspect these are as much problems with the language we use to describe the whole phenomenon as anything. For example, saying "Consciousness is an illusion," a comment heard in certain "explanations" contains an internal inconsistency. An illusion is experienced by a consciousness that recognizes something possibly amiss! So what is this consciousness that is experiencing the illusion of being conscious? As I said, this is dangerous territory.

I reject the idea that consciousness as a phenomenon will always remain inexplicable because we get caught up in circularities such as this – consciousness experiences (is conscious of) consciousness. Rather I assert that we need to approach this territory with a systems point of view as our compass.

As I will shortly explain, the term consciousness is, itself, problematic because it means slightly different things to different people. It is most often related to the concept of 'awareness,' which can mean that an organism is attending to a force, a chemical gradient, or information about objects in the environment through sensory inputs. The awareness of the environment is evidenced by the behavior of the creature. In the simplest form, even bacteria show awareness in the way they actively swim up a chemical gradient toward the food source that emits molecules associated with the 'taste' of that food[123]. Toward the human end of the spectrum we tend to think of awareness as in being awake, alert, and attending to external conditions, but we also think of awareness of internal thoughts/concepts and even the process of thinking itself – what we call awareness of being aware. This is what I refer to as second order consciousness. Some writers reserve the term consciousness to only describe this idea of awareness of being aware[124].

Yet another approach to consciousness derives from a recognition that our thinking can only be about a limited number of 'things' at a time, even while many 'things' are going on around us and within us. For example most people are oblivious (as they report verbally or behave as-if) to their own digestion (until something goes wrong). They are not aware of their blood chemistry even though their brains are actively monitoring it and sending out signals to various organs to keep that chemistry within ideal ranges of concentrations. They are not conscious of their low-level operations.

---

[123] See the Wikipedia article on chemotaxis: http://en.wikipedia.org/wiki/Chemotaxis

[124] See Block (1998) for a distinction between what he called P-consciousness (perception) and A-consciousness (access).

Nor are they aware, often time, of how they actually perform some behaviors. For example if you know how to ride a bike you do not have to be conscious of every muscle involved in adjusting your balance or your speed, etc. Your skill has been committed to tacit memory systems and your higher consciousness does not need to get involved (unless something goes amiss). Consciousness, in this sense, is strictly the phenomenon of thoughts of which we are aware. And it is what we mean by this phenomenon that constitutes the dangerous territory we enter when trying to explain it.

Modern theories of consciousness, as awareness of being aware, are being constructed on the basis of neurological evidence for what is actually going on in the human brain when a person reports verbally or behaviorally what they are aware of. Actually this is also the case for non-human animals as well, though they obviously cannot report verbally[125]! Below I will outline some aspects of these theories, especially as they relate to the uniquely human forms of consciousness that has evolved on our planet. I am interested in the relation between human consciousness and all that it entails, and sapience.

The So-called "Hard Problem"

Philosopher David Chalmers (1996) introduced the idea that there is an aspect of consciousness that is a hard problem[126], or rather, that some fundamental aspects of consciousness are too hard to explain by mechanistic models.

It seems that some of us need our "mysteries."

According to Chalmers there are "easy problems" associated with consciousness. For example the mere processing of external stimuli, recognizing what they are and where they come from is easy enough to explain from mere brain theory. For Chalmers and many other philosophers of mind the real problem is *subjective experience*. That is, how do the stimuli evoke subjective experiences such as "redness", what are called qualia, or "phenomenal experiences?"

This is where we run into significant rhetorical problems. As soon as we say an experience is subjective we are making a claim about our own experience, not a claim about another's experience, or what is objective. It is impossible to say that Carl experiences redness when looking at an object that I experience as red. At best Carl and I can agree that whenever looking at an object that I experience as red, he reports that he also experiences something he calls redness. We agree that some kind of visual experiences are consistent across objects. We both use the same name for it. And when I tell Carl that the object I just saw (which he did not see) is red, he understands what I mean in terms of his own experience. It is because of this property of consistency across shared experiences that we might readily conclude that redness is not actually a subjective experience only. There is some physical quality about the way a human brain

---

[125] See the many works by Frans de Waal (2005, 2010, 2014, & 2016).

[126] The Wikipedia articles http://en.wikipedia.org/wiki/David_Chalmers and http://en.wikipedia.org/wiki/Hard_problem_of_consciousness provide some background for most readers.

interacts with reflected light waves to see the same basic quality as almost all other human brains. I submit that while the issue of qualia may keep philosophers up at night it is not a real problem when considering the nature of and brain mechanisms for producing the phenomenon we call consciousness. I find myself in agreement with Stanislas Dehaene (2014, p 262) when he calls "red herring" to the idea of there being a "hard problem" at all. I hope to answer this problem in this chapter.

There are, however, additional significant semantic issues involved in grappling with the idea of consciousness. When I write, "I saw a red object," what exactly is the "I" (in both instances in this sentence)? There is a symbolic referent (I or me) that is used linguistically to identify the agency of a biological system. But more than that (and what is for me the truly hard problem) is that there is a locus of experience and thought that feels an identity and ownership of those experiences and thoughts as well as of the body in which it seems to reside. I can talk about "my body" as if it is a thing that does my bidding and is used to interact with the world. The "I" inside seems to be unique and, in a sense, somewhat isolated from the body. You will recognize this as the ancient mind-body problem so often argued by philosophers[127]. What I will attempt to demonstrate is that there really is a "mechanistic" explanation that solves this hard problem. The "I" is actually an agent processor located in the human brain (as well as some other mammals and possibly birds) that has evolved specifically to 'experience' the world and the results of actions taken so as to manage future behaviors. The problem of consciousness is explained within the evolutionary paradigm as an emergent capability to *strategically* manage the future.

The field of scientific consciousness studies has been making considerable progress in the last decade. Once something no respectable psychologist would touch (only respectable philosophers), the explanation of consciousness is starting to take shape. Neuroscience has had a lot to do with this. It is now possible to identify areas of the brain that are actively participating in conscious awareness in awake subjects using functional magnetic resonance imaging (fMRI) and other neuroimaging methods.

Famed neurologist and author, Antonio Damasio (2000) tackled this problem head on in his work, *The Feeling of What Happens*. Rather than ponder what consciousness must be from an armchair, Damasio has been examining the brain, its functions, and their correspondence with reported subjective experiences as well as behaviors. I have found his arguments (paraphrased below) quite convincing as far as they go. They do provide a more solid ground to start from than introspection alone. My own approach is, in a sense, similar to Damasio's but working from a kind of reverse engineering process (systems science analysis). My work on autonomous agents starts by attempting to emulate the brains of very primitive creatures such as a snail, paying particular attention to the critical role of memory trace encoding in neuronal synapses (Mobus, 1994). It is my contention that this is the first problem to be solved before attempting to

---

[127] For a somewhat comprehensive review of the Mind-Body problem for philosophers see: http://en.wikipedia.org/wiki/Mind%E2%80%93body_problem in Wikipedia.

emulate whole brains. It is absolutely essential to understand the dynamics of this encoding in order to solve certain critical problems in memory trace behaviors that we know affect long-term behaviors in all animals. My immediate goals are to build brains that are progressively closer to mammalian capabilities (not necessarily human). This will be demonstrated by their capacity to adapt to non-stationary environments and still succeed at a given mission objective.

I think the answer to the question of, "What is human consciousness," lay in the evolution of brains from those primitive versions up through mammals and to humans. Below I will review the work of Tomasello (2014) regarding the stages of kinds of thinking that correlate with the levels of consciousness model. In my own work I have elected to try to emulate the stages of brain evolution by simulating biological-like neurons and their dynamic interactions in brain-like structures (e.g. the hippocampus and its analogues in reptiles). Essentially I seek to grasp how the brain works by recapitulating its evolution. In the next two chapters I will come back to consciousness from, first, the neurobiological aspects, and then from the evolution of the brain aspects. Consciousness, the human kind, and sapience are likely just different perspectives on the same underlying mental system. Exploring sapience entails exploring consciousness. For now, however, I want to simply establish the form and functions of this phenomenon to show how the two are intertwined.

Mental Architecture

'Architecture' describes the components and their functional interconnections in a system. A building architecture is a description of building materials (bricks and mortar) and the rules for putting them together to achieve the objective of constructing a building that fulfills a desired function. A 'mind' architecture describes the various mental components, each with its relevant function, and how they are interconnected to produce mental experience. Since mental functions are the result of brain module processes there is a necessary overlay between mental and the physical brain architectures. In other words there is a deep and tight relationship between the psychological phenomena and the neurological structures.

Figure 3.2, below, shows a rough map of the mind in a slightly different way than in figure 2.4. Here the triangle (or pyramid) shows roughly the amount of neural machinery that is given to various functions. I've also identified the hierarchical cybernetic model represented by brains/minds as compared with figure 2.2. The diagram shows relative proportions of brain activity that is either available to conscious awareness (or semi-conscious as in aware of a mood or feeling without being able to say precisely why it is being experienced) compared to subconscious processing.

This representation is actually a little misleading in one sense. If read literally it implies that only a very small portion of the brain is involved in conscious mental activity – the top of the triangle. However the entire brain is *involved* in conscious thinking to one degree or another. That is the meaning of the arrows pointing upward from the labelled 'unconscious' areas. What is actually represented here is that a small portion of the brain, namely limited regions within the prefrontal cortex, are actually involved in ‘producing’ conscious experience but they do their work by ‘orchestrating’ the activities of the rest of the brain! This part of the brain (as I explain in the next chapter) acts as a conductor of an orchestra, the rest of the brain regions, each doing its own part but only vaguely, if at all, accessible for individual recognition.

Sapience is shown at the top as a relatively small part of what the brain is doing. But all activities of the brain do, indeed, eventually feed into this seemingly smaller activity. The flow arrow from the central region, labeled 'models of self & world', show that tacit knowledge about self and the world, which includes current states of both (short arrow pointing straight up in the middle), is available to the conscious/sapient mind which then uses this knowledge to perform executive functions, i.e. to guide intelligence and creativity in decision making.

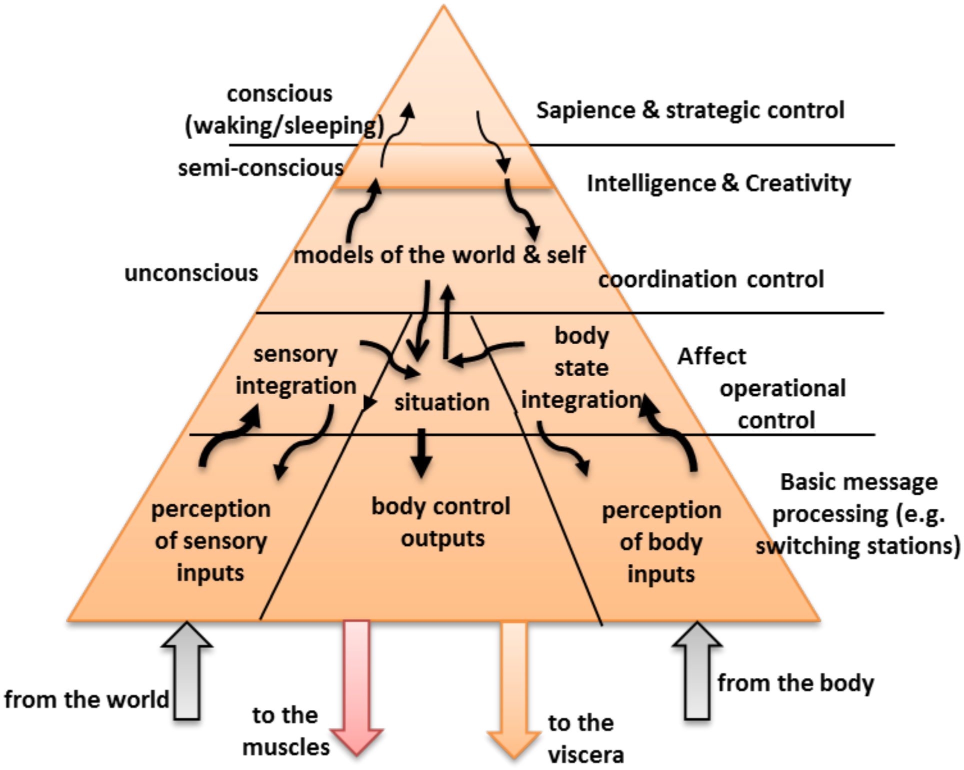


Figure 3.2. This is a rough overview of mind/brain functions that shows relative proportions of what is available to conscious awareness in the awake state. Interactions with the body and the outside world are shown at the bottom of the pyramid. Considerable processing power is given to receiving and processing sensory data from both the environment and from the body. The brain integrates this data to form a situation state report. This is sent upward informing the coordination control level and higher levels where motor and endocrine outputs are initiated and sent down for innervation of the body. The highest level of the pyramid is the conscious level where awareness of what is happening and what the body has done in response. At this level strategic controls operating over a much longer time frame formulate plans. Either consciously or unconsciously this level can adapt models in tacit knowledge based on long term results.

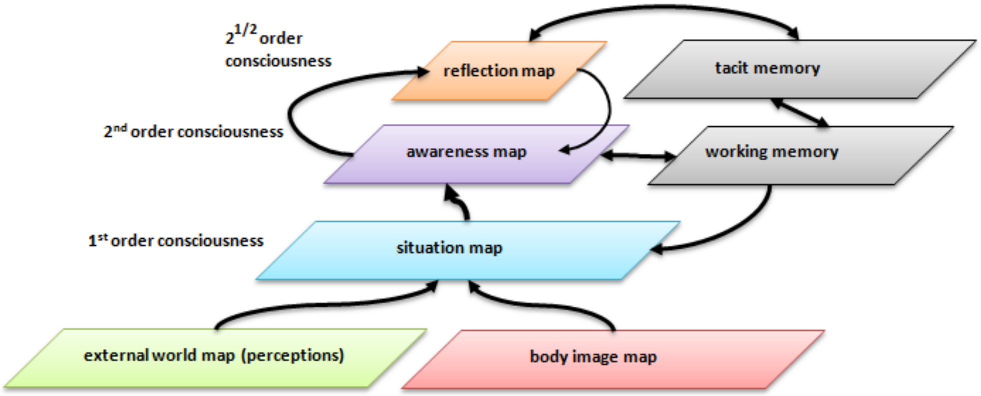
Most of our routine decisions do not require any awareness, or strategic control, and so are generated directly by the coordination level. You know the experience of driving a car. You hardly think about controlling it and only need to be keenly engaged (conscious) when a surprising event occurs. Indeed, most of our daily activities do not need supervision by conscious/sapient awareness. Most can be handled nicely by lower level intelligence(s).

When we are conscious we (our brains) are either involved in intentional thinking, attending to novelty or semantic salience in the environment, or are simply aware of spontaneous thoughts (as in day dreaming). The latter is arguable as far as consciousness is concerned, but most of us, when we snap out of a day dream, have a memory of what we were thinking. These conscious moments play an important role in sapience in that they trigger what I call a second-order judgment process. That is, the judgment processor is activated to make judgments about the judgments themselves. We are generally conscious of our decisions (made at the subconscious level) after the fact and we are aware of the effects those decisions have on the situation (the results in the environment and on ourselves). This awareness is fed back to the subconscious mind where it can be used to modify or alter the models if our results were less than favorable, for example.

Thus, sapience and consciousness are interrelated as operations at the highest level of the hierarchical cybernetic architecture of the mind. This, I think, is at least part of the explanation of what is unique about human beings as animals. Our minds have evolved this sapience/consciousness apparatus to achieve strategic control not only for a single individual, but for a social body as well. In a sense this mechanism implements a kind of distributed strategic control function among members of a tribe/family.

Humans are intensely social creatures. No one human is or can be an isolated entity without going mad. I posit that the evolution of language was driven by the need to share mind space among such beings. Language, which is a recursively generative symbolic information sharing mechanism, allows individuals to help others construct models in their minds that are similar to the ones held in the speaker's mind. The content of language reflects the thinking shaped by sapience. It is about internal sentiments, systems perspective, judgments, and strategic perspectives – the future. Language provides individuals with a tool for understanding others' thoughts and feelings in a way that mere body posture or facial expressions could never do. I will provide a brief discussion below regarding the nature of language and return to this aspect in chapter 5 when I look more closely at the argument for how sapience emerged and evolved.

One more note on the issue of consciousness before going back to a more mechanical explanation of sapience. Most people will accept that animals are, in some sense, aware of their environments (see below). Many will also accept that some higher mammals, e.g. chimpanzees and possibly dolphins, are self-aware. That is they have what I have called second-order consciousness. Self-awareness includes awareness of the self being aware of the environment. Such second-order consciousness, as outlined below, may have been the route to higher sociality (than just a herd or colony) in mammals. Chimpanzees recognize one another as individuals. They have unique personalities (some claim the same for dolphins and elephants). But sapience may add yet a higher order to consciousness, at least what I call a $2^{½}$-order consciousness wherein we are occasionally aware of being aware of awareness! There are first-party reports, from time to time, from people who have experienced internally observing their own thinking in action. This is often reported as a kind of disembodiment from the mind or as a "higher" mind observing the ordinary mind going about its business of being aware of the world. These reports are rare and sound esoteric. Nevertheless there may be something to this in the model, suggested by Damasio (1994, 1999) of a hierarchy of what he called maps (see also chapter 4)[128]. Figure 3.3 provides a quick summary of these ideas and a hint of what this sapient $2^{½}$-order consciousness might be.



---

[128] A map is essentially an array of neural elements through which inputs, say from perceptual areas in the cortex, are 'translated' into higher order representations. This concept is covered extensively in chapter 4. For now the word map can be thought of as something like a road map, in which roads and cities are 'represented' stylistically on a sheet of paper. There is a correspondence between what is on the map and the real things that it represents.

Figure 3.3. Damasio posits a hierarchy of what he calls maps or images of the states of things from moment to moment. There is a rough correspondence between this model and that in Fig. 3.2 in that there is a flow of information from the perceptual systems and body sensing through neural structures that map those states and present an integrated version to higher level maps whose job it is to figure out what is happening. The "situation map" is what we would call ordinary or 1st-order consciousness, or awareness of the environment and the body. This map is a convergence zone for all of the inflowing information and is related to lower level operational decisions (e.g. controlling the throwing of a rock). The "awareness map" monitors what is happening in the situation map and this is our ordinary experience of being conscious. We are aware of what is happening but also aware of our memories and objectives. This is where tactical control gets initiated. Working memory provides a scratchpad where current and recent states can be stored and called upon for generating actions, or focusing attention. Note the two-way communications between memory areas (also dynamic maps). Finally, the "reflection map" provides a kind of ultimate monitor able to control the use of tacit memories and working memory through the awareness map. While speculative, this model does help explain a number of phenomena associated with human consciousness and is very much conjoined with the thesis of sapience.

The Evolution of Thinking – a Preview

One of the most compelling accounts of the "Natural History of Human Thinking," comes from Michael Tomasello (2014), in his book by that name. What follows is a preview of the chapter on evolution (5) in that it examines what is essentially a notion of the accretive nature of progress in evolution. That is, as evolution proceeded, the complexity of the brain and behavior, as well as ways of thinking in animals, increased based on adding onto existing structures/functions by laying on new layers of neural capabilities over the existing ones. The human animal is not just a new kind of animal. It is an animal that has the most ancient structures/functions covered over and modulated by newer structures/functions. Nothing has been lost in terms of what the brain has done for primitive animals. What has happened is that newer capabilities have been added and the older capabilities are now regulated (to some degree) by the newer capabilities.

What follows is an amalgam of Tomasello's view of human thinking (and consciousness) with my own perspective of the orders of consciousness. This perspective is naturally mapped onto the evolution of consciousness so is difficult to not anticipate some of the ideas to be presented later in the chapter devoted to evolutionary considerations. But that is the nature of systems science. No single aspect of a CAS can be separated cleanly from any other aspect[129]!

Orders of Consciousness

Another way to think about levels of consciousness, from the level of mere perception to the level of human consciousness of being conscious, is that of a set of concentric rings wherein each outer ring adds various kinds of functionality to the process of the organism attending to its environment and using information to make increasingly complex decisions (as covered in the prior chapter). Figure 3.4 shows this idea along with flows of influence and information into and

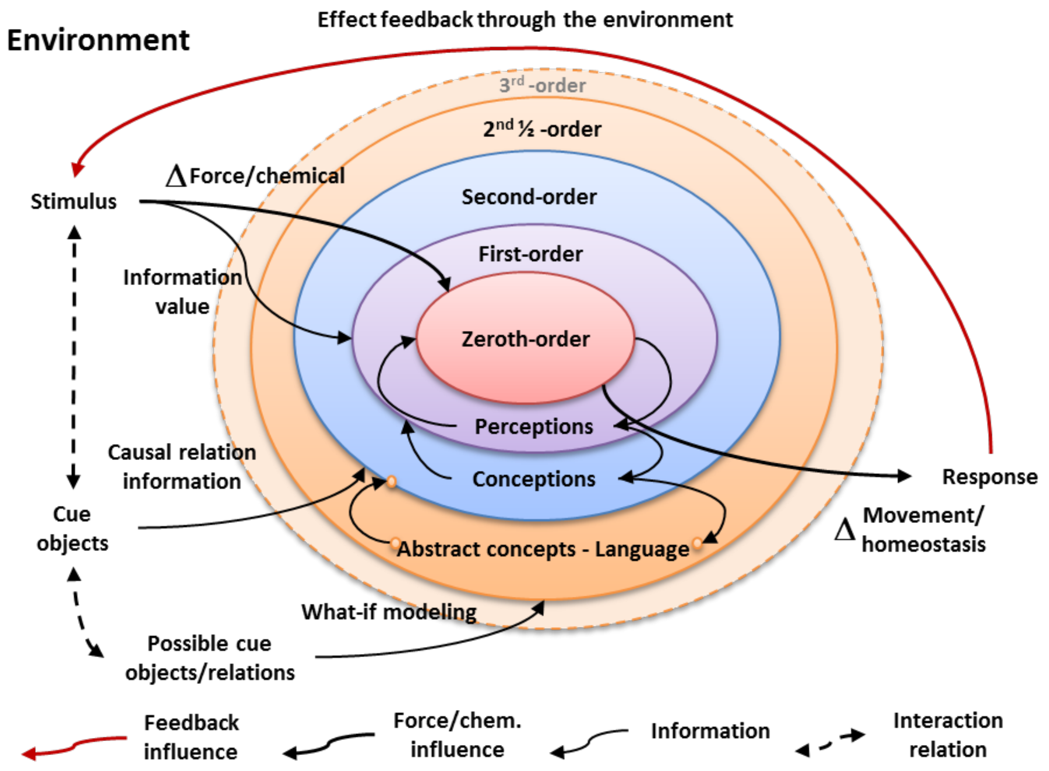
---

[129] This is the "curse" of the explication of systems principles that we found hard to keep simple in Mobus & Kalton (2014). In every chapter we have to reference topics covered in other chapters. The book is riddled with comments like: As you will see in chapter X… or As we covered in chapter Y…

between levels; it is just another version of figure 3.3 above and related to the pyramid diagram in figure 3.2.

Every aspect of our consciousness is about how we interact with the rest of our environment and particularly others of our kind.

This figure actually conforms to the way the brains of more advanced animals have evolved. Early animals (like worms and snails) had hard-coded responses in simple neural networks, responding directly to the influence of 'forces' and chemicals in their medium. This is what I have called Zeroth-order consciousness. Indeed, an argument can be made that all living systems have this capacity[130].



---

[130] This is a principal argument advanced by Maturana and Varela (1980) regarding the nature of *autopoiesis*. See the Wikipedia article: https://en.wikipedia.org/wiki/Autopoiesis.

Fig. 3.4. Orders of consciousness form a concentric set of "layers" of awareness, memory functions, communications, and intentionality. My claim is that humans have crossed from mere 2nd order consciousness into an intermediate order (2½ order) – not quite 3rd order. All organisms have zeroth order consciousness. All animals have 1st order consciousness. All mammals and most likely birds have 2nd order consciousness. Some apes show glimmers of 2½ order consciousness (e.g. shared intentionality shown in the act of grooming). But all humans of the species sapiens share advanced aspects of the later order (e.g. having language and the beginnings of strategic perspective).

As brains evolved more elaborate mechanisms for recognizing objects and encoding causal relations between those forces and those objects that were associated as being sources, brains were capable of not just responding after a force or chemical was experienced, but could anticipate what the environment was about to do based on perceiving the objects, First-order consciousness (earliest animals had these forms hard-coded into the neural networks but later animals, like amphibians and early reptiles probably had some capacity to learn new objects as well as causal relations).

With the advent of the mammals and birds[131], another layer was added (related to 2nd-order consciousness) that gave much more capability for encoding (learning) new objects and relations. This capacity expanded as mammals and birds evolved further. In the mammals, in particular, the advent of primates resulted in brains that could construct much more complex models of the world to use in anticipatory processing. The stage was set for yet another layer that would evolve in the human line (an advanced Second-order consciousness).

Zeroth-order consciousness

All organisms respond to stimuli of physical forces, temperature, and chemical interactions with genetically programmed responses. The very nature of life implies a primitive kind of awareness of environmental conditions even if there is no awareness of the sources or sinks associated with those conditions. In animals with nervous systems the situation is more complex but the basis of response to the environment, though instinctive, is no less. That is the genes that stipulate the hard-coded wiring of neurons in their primitive (or primitive parts of their) nervous systems provide an evolved automatic reaction system to factors in the environment that matter. We should call this stimulus-response awareness. The capacity to respond to these stimuli is built into the phenotype. In certain cases the stimulus-response mechanism is non-modifiable, those cases where the stimulus could directly impact the capacity to stay alive are of this kind.

Living systems are also mandated to conserve energy and materials whenever possible. There are a number of phenotypic mechanisms that can be up or down modulated as a function of longer-term experiences. These mechanisms are called adaptive response. They are maintained at a low level of response strength unless repeatedly stimulated in which case the organism invests more energy and material to build up the response mechanism so that it reacts more swiftly and with a stronger counter force. A good example of adaptive response is the way the muscles can increase

---

[131] Based on some inferred behavioral traits in dinosaurs, from which birds evolved, it is possible that these animals, though technically reptiles, had more than just instinct-driven behaviors.

in strength with repeated exercise. Over some period of time the body builds its capacity to respond to the increased stresses and the capacity will persist between episodes. Indeed, the capacity will persist over extended time, a form of memory, even when not repeatedly called upon. However, as with mental memory, the muscles will begin to revert to lower capacities if not reinforced with occasional episodes of increased stress.

Primitive neurons (and neurons in the most primitive parts of more advanced brains) show this same kind of minimal memory. Synapses (see chapter 4) encode memory traces in the form of increased response capabilities (called efficacy) when repeatedly excited with reinforcement. Animals like marine nudibranchs (like snails) have been shown to be 'trainable' in the convention of Pavlovian conditioning[132]. They have various stimulus-response mechanisms that have been shown to either habituate to repeated stimuli, or become hyper excitable if they represent a threat.

All biological basic drives are mandated to preserve life. Every animal must take in food (material and energy), water, oxygen, and live in an environment conducive to its existence. Under the appropriate circumstances all living systems will attempt to grow and reproduce (expand biomass of its particular kind). All have evolved stimulus-response capacities to fulfill these mandates under nominal or mildly stressful conditions.

Zeroth-order consciousness involves the most primitive form of self-other awareness. Every "nervous system" receives signals from the actuation of glands and muscles internal to the animal. These signals are part of what are called the proprioceptive senses. The external senses (sensing the environment) are called exteroceptive. The signals from both of these systems can be compared with one another to determine which external forces/chemicals are due to the self's actions and which are due to other non-self factors in the environment. The organism does not react to exteroceptive inputs that are matched with proprioceptive information. Figure 3.4 shows how a most primitive form of self/non-self determination (awareness) can be computed by neuronal networks.

---

[132] See (Alkon, 1987).

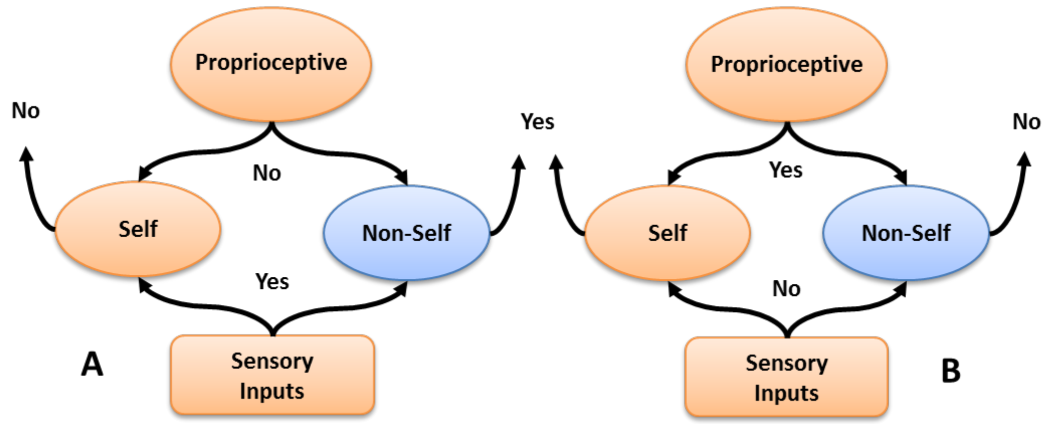


Fig. 3.5. Determining if an external (exteroceptive) input is due to the self's actions or not provides a primitive form of self vs. non-self awareness. The animal does not react to external stimuli that are the result of its own behavior.

The brains of zeroth-order conscious animals have no representations of objects per se. They just perceive forces/chemicals/temperature, etc. Their brains are effectively just mappings between stimuli (and combinations of stimuli) and responses, modified only by self-knowledge of self-actions. As explained below (see Systems Perspective) this corresponds with having a model of the self as the system.

First order consciousness - All animals

First-order consciousness became possible when new kinds of brain structures, generally cortical, or sheet-like, structures that are specialized to encode images and the causal relations between small modular regions containing the attributes of those images (more explanation in chapter 4). These structures are organized to receive direct inputs from various sensory modalities in topological order. That is, for example, a structure specializing in, say, touch will receive inputs from all of the skin regions across the sheet in one-to-one relation to the skin. In this way the sheet represents a map that encodes relations across a spatial domain (in any of the sensory modalities). Thus these structures have the potential to encode (memorize) "objects" (though the details of this are more complex than I can cover here).

In the most primitive vertebrate animals (e.g. fish) object representations are often hard-coded, e.g. that "thing" is food, danger, or a mate. However, if the brain structure (usually a globular module with a kind of rolled up sheet forming what is called a "nucleus") is complex enough it can encode new relations or learned perceptions.

In animals a little more advanced phylogenetically (e.g. amphibians and reptiles) a new version of a sheet structure, a true cortex emerged that formed a wrapper around the still-existing nuclei structures. These paleocortical structures proved to have even greater flexible encoding capabilities and could easily provide mappings between sensory inputs and object-relation memories. Recall figure 3.3. The lower two "maps" represent structures capable of encoding the current state of affairs in both the body and the external environment. Both of these feed into a higher level map (called the “situation map” in the figure). This is where the animal becomes aware of the relation between what is going on in the environment and its own internal states. In animals like amphibians and reptiles these maps are more than just real-time encodings. They involve memory (models) that allows the animal to anticipate that changes in the environment will produce changes in the internal states of the organism. This ability provides a form of anticipatory adaptive response, i.e. responding slightly before the environment changes to the point of causing a disturbance in the physiology of the organism.

The key characteristic of first-order conscious animals is something Tomasello (2014) calls “individual-intentionality.” Characterized in language, in human consciousness, it would be: "I" want *that* and "I" will get it. Of course animals much below humans in the phylogenetic hierarchy do not have language, however their brains process the equivalent of images and reactions to those images that carry the same basic semantic content. The animal recognizes objects that are beneficial to it (food, water sources, and potential mates) and is motivated by the biological mandates (zeroth-order) to obtain what it perceives. Alternatively, the animal might recognize a threat and “know” immediately what it needs to do to avoid it. The point is that each animal is a completely closed cognitive system and is concerned first and foremost with its own satisfaction of biologically-mandated needs.

First-order consciousness gives rise to the dominance of competition, both inter- and intra-specific, for recognized resources. The majority of the evolutionary progress in the animal kingdom, from fishes[133] upward has been in refining competitive competences (see discussion below re: tactical management). In order to accommodate the demands of increasing success in competitive behaviors brains needed to evolve more memory capabilities. The neurons that had always been able to encode short-term memory traces needed to be able to keep memory traces for much longer time scales, relative to the lifetime of the animal. Intermediate term memory trace encoding could be handled by introducing new adjunct and associative stimuli (see chapter 4 for more details). Some adjunct modules, for example nuclei that became the hippocampus and amygdala, helped the situation map increase its ability to commit memories of situations that could be called upon to predict (actually anticipate) future states of the world given the current perceived state.

The situation map also provides another cybernetic facility, the organism is capable of monitoring its own behavior relative to goals intended and make modifications if needed based on real-time feedback. Those modifications under the situations being mapped are then also available for learning and refining future responses.

---

[133] While fishes have very primitive brains, they do have some nuclei structures that provide a limited amount of learned object-relation encoding. It is a little problematic to fit them neatly into first-order consciousness, but since there are gradations in (e.g. 1, 1.1, 1.2,… etc. ordered) consciousness, it might be hard to distinguish between 0.99 and 1.01 order!

These anticipatory situation maps became the basis for improving tactical models, or models of what the organism should do in various situations to out-compete others. As evolution proceeded to produce more elaborate versions of animals with these capabilities, especially the mammals and birds, a kind of arms race ensued that drove the evolution of even more capable brains. The neocortex of mammals (and a similar structure called the palladium in birds) provide for the next stage of consciousness evolution. This involved more elaborate tactical models and more long-term conceptualization.

Second order consciousness – Mammals and Early Humans

The advent of a new layer of cortical structures overlaying the paleocortex, that is the neocortex, added a new capability in animals. Many more elaborate and complex kinds of perceptions could be encoded in the posterior portions (e.g. the occipital lobe) and more anterior portions of the new cortex could encode higher order representations, which I refer to generically as concepts. A principle difference between concepts and percepts (which are both encoded similar neural structures called cortical columns in very similar ways) is that the latter are dependent on spatial mapping, i.e. where in the environment outside the organism are the objects located, while the former are spatial location-free; they exist regardless of where in the perceptual field what they represent is located relative to the organism, including when they are not anywhere in perception. Moreover, concepts are formed in an increasing hierarchy of abstraction that allows the brain to form categories and archetypes or generalizations. All of the encodings in the neocortex are learned from experience. It (and possibly much of the paleocortex) is the only *tabula rasa* in the brain. But it should be understood that most of the lower-level sensory areas are bootstrapped by older brain structures (e.g. the thalamus relays sensory inputs to pre-specified regions of the primary sensory cortex).

In the next chapter I will provide more explicit models of what is entailed in perceptual and conceptual encoding and operations (e.g. thinking) using those memories.

The advent of conceptualization and the ability to use abstracted concepts mentally (thinking), under the management of the executive functions in the prefrontal cortex, allowed for a new mental state. Animals that tended to work in groups (not necessarily herds or flocks) such as wolves in packs found great advantage in what Tomasello (2014) describes as joint-intentionality. Two or more conspecifics will cooperate to obtain a resource that would be unobtainable by only one individual acting alone. What is needed is for all of the participants to have learned common concepts, such as: “THAT object with antlers is food.” They would have had to learn commonly experienced contexts and have some form of primitive communications that would trigger mental recall in other members of the concepts and their contexts[134]. For example, one wolf picks up the scent of a moose and howls to its pack members that a hunt is on;

---

[134] Tomasello (2014, chapter 3) lays out the presumed nature of communications that early hominins used which included pointing in a direction they wanted their companion to look, and pantomiming things and situations to convey the context and generate shared conceptions in those companions.

they all gather and enter the chase. In general shared intentionality might be characterized as: “We want THAT and will cooperate to get it - THAT is a great resource but more than either of us could get on our own.” Of course this cognitive process is not necessarily conscious, and it certainly does not involve explicit language constructions. In wolves the motivation to hunt as a group must be mostly instinctive. In chimpanzees that organize a monkey-hunt it is probably a conscious recognition of an intention (by others) to engage in a cooperative activity. But they do not use language to direct the activity. Early humans may have used pantomime and non-word vocalizations to initiate their intentions which were picked up by others. Modern humans can use language (vocal and body) to instantiate within the minds of their companions their intentions to hunt game.

The advent of a limited form of cooperation provided many evolutionary benefits to mammals (and birds). Cooperative parenting, for example, could expand to cooperating extended family models where the older siblings could assist with raising the newer offspring. The ability to learn models (higher-order concepts) of one’s own behavior as well as that of others recognized as individuals provided the basis for anticipating what others would do in various situations. This was the beginning of empathy in the sense that one individual could make estimates of what other individuals were thinking and seek behaviors that would maximize the benefits for the group.

Second-order consciousness and joint intentionality took mammals a very long way evolutionarily. In the primate line, the hominins consistently raised the bar, learning how to control fire, for example, and shape stones for special purposes. The nature of communications needed to coordinate among members of a group began to become more complex than could be handled by mere pointing and pantomiming. The environments of omnivores like *Homo ergastor* were becoming extremely complex. Utterances that conveyed concepts were needed to deal with the greater combinatorics of situations. A primitive language facility emerged, probably, in later hominids (e.g. *Homo erectus*) that may not have been able to generate complex combinatorial sentences as with modern syntaxes. But these animals could communicate immediate coordination requirements in carrying out ordinary daily living.

Parts of the brain in higher association and motor coordination cortices were already specializing for such a language facility when a dramatic evolutionary event took place. The frontopolar tip of the brain, what we now call Brodmann area 10 (BA10), underwent a rapid and extensive expansion and possibly a cytoarchitectonic alteration that triggered similarly rapid changes in the development of other areas of prefrontal, pre-motor, and higher association regions of the neocortex. Most notably, the primitive language areas evolved into their modern capabilities. We think this expansion event took place between 150 and 200 thousand years ago, and *Homo sapiens* emerged in Africa.

Second and a half order consciousness - Early sapience; modern humans

As Tomasello (2014, 2019) explains it what constitutes the major difference between modern humans and prior species of *Homo* (and all other animals) is that our brains produce a new capability of collective-intentionality, which is much greater than mere joint-intentionality of second-order consciousness. We might characterize collective-intentionality as: “We all agree that THESE are the things that we SHOULD want and THESE are the ways we agree we SHOULD go about getting them.” The “should” and the “agreement” are part of a collective knowledge that we call culture. All of the individuals exist in, and develop within, a commonly accepted milieu of artifacts, practices, beliefs, mores, and so on that constitute the community identity.

A community identity reinforces cooperative attitudes between members even though occasionally a “cheater” might attempt to not carry their weight in doing work in the group. Sapience and 2 ½-order consciousness includes strong moral sentiments (see below) in which members experience a range of emotions such as moral outrage and a desire to punish when a cheater is discovered, to feelings of gratitude to hard workers, and personal pride in one’s contributions. These kinds of emotional ties can only make sense in the context of an identifiable group to which one feels a belonging.

Therein, however, lay the roots of a potential problem. Group selection holds that groups that have higher levels of cooperation internally are more competitive against other groups in terms of access to and “ownership” over resources, which leads to conflicts. The group that better organizes to capture/defend/utilize resources will be the fittest and the genetic propensities that cause the brain to wire up for cooperation will be differentially passed on to the progeny leading to maintaining and perhaps increasing cooperativity. To get greater cooperation among group members would seem to require greater competition between groups. I will revisit this conundrum in chapter 5.

A big boost for the species as a whole, and what makes it more competitive (more fit) against other species of humans that emerged more or less contemporaneously with *sapiens* (e.g. *neanderthalensis*) is a capacity to think more deeply into the future as a result of the expansions of the management facilities of BA10. As I will show in the next chapter, this amounts to an ability to increase the anticipatory aspects of complex causal models. An ability to imagine a future state of affairs based on questions about what would happen if WE do such-and-such takes things like tool design and planning future migrations to new heights.

According to Tomasello’s (2014, chapter 4) view the advent of collective-intentionality led to each individual adopting a new perspective on the world, that of an objective reality. Because of the immersion of one’s development in a culture, which includes especially the specific characteristics of how the inherent moral sentiments are actualize (beliefs about good and bad), the individual internalizes sets of facts about the world that take on the perspective of being objectively true. An individual sees the world as being the same way other individuals see the world, which must mean that the facts exist independent of the individual's thinking them to be true.

All of this is made possible because humans evolved a language capability that provides a sophisticated mechanism for transmitting concepts from one person to the next and even constructing more complex concepts by hearing others' descriptions of concept relations. This is accomplished through a syntactic language that includes recursive generative structures (sentences with clauses), which allow the construction of any complex idea from the combinations and re-combinations of simpler constructs.

Second and ½-order consciousness allows humans to participate in the highest level of cooperativity yet achieved in the animal kingdom. And it has been a spectacular success insofar as humans have been able to adapt to living in almost every climate regime on the planet, extract numerous resources, especially energy, from the environment, and basically take over the planet.

But that spectacular success is beginning to look like the seeds of spectacular failure. Sapience was a major leap into a new realm of organization of living systems, the most recent "major transition" in Smith's and Szathmáry's (1995) terminology. Will it prove to be a failure of evolution, or might there be more transition to come?

Third order consciousness - Advanced sapience; future humans (?)

Let me anticipate some of what I will cover in chapter 5 on the future evolution of sapience. Collective-intentionality is the latest "wrapper" around the other forms of intentionality and cognitive capacities (as in figure 3.4 above). It does have some ability to manage the older and more primitive forms, from zeroth-order to 2nd-order but it does not eliminate the more "primitive" urges, impulses, emotions, and instinctive behaviors that do not particularly accord with the collective good. This is how it is possible to have "cheaters" show up in a group. The genetic propensities to take care of #1 first are still there and under the right circumstances poke their heads up into conscious decision making and take over. In today's world, those circumstances seem to have become the "normal" state of things.

If human beings are to achieve true hyper-sociality there must be a leap into a fully 3rd-order consciousness in which the managing role of consciousness becomes strong enough to keep our baser instincts in check, or applied to efforts that are for the good of the group. Following Tomasello's schema I propose that some form of improvement in the computational capacities of, in particular, BA10 (or an add-on module) could boost human cognition and consciousness fully to the 3rd-order; sapience would be strengthened considerably to what I would term "eusapience," true sapience.

I suggest the form that this would take is strategic-intentionality. It might be characterized as: “THIS is the nature of the world in which the WE exists and that nature shapes our SHOULDs; WE should want to exist in harmony with nature.”

In other words our thinking needs to see the objective reality of the world as it truly is and not as we would ideologically wish it to be. We already have a collective knowledge acquisition method - science - with which to see the world objectively. But we as a species have not fully developed the cognitive machinery to make it a normal thinking process. Science is currently dependent on the cognitive capabilities of just a few members of the species, and even then there are lapses in judgment along the way. Scientific thinking must be a norm for every member of the species (or I should say ‘a’ species of *Homo*) in a way that is not the case now. Every individual needs to be able to model reality without letting ideological and often false models subvert understanding. If a cultural construct is found to be in error then it should be changed or expunged in favor of understanding reality better than before. When ideology guides thinking it is more often the case of throwing out the facts in favor of the beliefs.

Strategic intentionality means that every individual would consider every thought from the perspective of where does it lead in the long run. In other words, judgments about what would be best to do in the present should be biased in favor of a long-term greater good outcome. The future should not be discounted in the way it currently is in our 2 ½-order consciousness.

Third-order consciousness will require significant improvements in several areas of cognition and group interactions already in evidence, but seemingly weak in nature. First an individual would need to have a more refined ability to construct more veridical models of the phenomena in the world. In other words human beings will have to be able to learn about the world with much more resolution and capacity to represent complexity. This will possibly require more memory capacity than we currently have, but it will also require more refined cognitive “tools” to utilize that capacity. As I will discuss below, one such tool involves a much better ability to see the world from a systemic perspective. In part this will depend on the internal language of the mind (see below for discussion of the language of thought) containing more of the semantics of systems. As I assert below, humans have a certain capacity to see systemness in the world now. But it is often vague or fuzzy and only covers limited scales of time, space, and complexity.

A more complete systems perspective, based on a more explicit systems language of thought would, in turn, lead to a much better explicit or public language for the purposes of sharing information among individuals. Much ambiguity in spoken languages now comes from under specified 3rd party descriptions of external phenomena and similarly under specified descriptions of 1st party interior mental and emotional states. There are conditions that are extremely hard to express in everyday language. Written language works slightly better in that more effort is put into expression completeness when writing. Thus reading about phenomena and mental states is less ambiguous for the information receiver.

My own suspicion is that a stronger systems-based semantics will emerge from a stronger systems language of thought. Expressing the world and one's own internal states in systems terms would provide much less ambiguous descriptions, thus increasing the efficiency of information transfer between individual. I am not talking about an "invented" language here. Rather I imagine a public language of systems will emerge naturally from deeper systems thinking just as our current languages emerge from our capacity to construct sentences based on a limited cognition of systemness.

Such a systems language should give rise to even higher capacity for groups to share perceptions. Coupled with a better ability to model the world with greater veracity, humans with 3rd-order consciousness should be able to dampen the tendencies to rely on ideologies over objective reality. In the last chapter I will explore the possible evolution of the brain that could conceivably give rise to this 3rd-order consciousness.

Sapience and the Level of Consciousness

The new capacity for cognition that makes modern human beings so different from all other animals is realized in our level of conscious awareness, of the world, of ourselves, of others, and of the future. It is embodied in the brain mechanisms that collectively I have called sapience. It is now time to dig deeper into what those mechanism are and what they provide us in our cognitions. As I presented in chapter 2, the major components of sapience can be identified as higher-order judgment, moral sentiments that modulate decisions, systems perspective that allows us to perceive the world as an organized whole with parts that are themselves organized wholes, and strategic perspective that lets us consider how things are working and what is likely to happen in the future given certain actions are taken in the present.

Judgment

Making a judgment is not the same thing as 'taking' a decision. A judgment may be used to guide a decision but the latter is perceived as a conscious act, whereas the former is perceived as a feeling or intuition. Judgments come automatically from the subconscious mind as part of a 'knowledge milieu', as in chapter 2, that operates in the intelligence processor to affect the option selected. But we should recognize that there are a range of judgments in terms of complexity, scope, and time scale that affect our decisions. There are also differences in judgments with respect to the strength or salience (or conviction) based on the depth of experiential knowledge held in tacit memories. Finally, there is an issue of efficacy of the model being applied. The inherent strength of sapience conditions this factor. Taken together, all of these factors determine how well our judgment succeeds in guiding our decisions toward efficacious outcomes.

Efficacious Models

Starting with the last point above about the efficaciousness of the models of the world and the self, which are used to generate a judgment or intuition, the recent research in human failings at making good, or even reasonable judgments in certain instances has brought to light a possible weakness in our mind's (brain's) capacity to make good decisions. It turns out that the brain is imbued with a number of heuristic-based mechanisms that serve as templates or baseline models for a number of different judgment tasks. There is now a rich literature in what is called the "Heuristics and Biases" program in psychology (Gilovitch, et al., 2002). Investigators have been systematically testing human judgments under a variety of circumstances to determine if there are a set of consistent mistakes made where judgments tend to be biased for some systemic reason. A number of mental heuristics have been proposed that lead to those biases that have been found.

For an example, the *representativeness heuristic* tries to match a sample (say a percept) with a canonical representation (a generalized concept) by an associative recall process. Features from the sample help recall representative concepts when they are the same or similar as those held in the concept (also called an exemplar). In many early evolutionary real world instances (meaning situations such as seeing a large cat-like creature and recognizing it as potentially dangerous) this is actually a pretty good way to go. It has the advantage of being extremely fast, which in many cases could have been the difference between life and death (Kahneman, 2011). There are, however several problems with reliance on representativeness in our complex modern world. First, what if you don't have an efficacious canonical representation that matches the percept? This could be a simple case of complexity; the representation you have is incomplete or does not incorporate aspects of the real thing itself because you never experienced this particular configuration before. Or, worse yet, you have an overly generalized representation that is missing subtle features that are actually important in distinguishing the true character of the percept (as when there are multiple different possible realities). For example suppose you, as a child, have this simple very general concept of a dog resulting from having only encountered a small sample, and all of them were friendly. You would be in possible danger if you were unable to detect that a specific dog was unfriendly. The brain will do the best it can, but what is introduced is a tendency to judge the percept on the best you have, and that may not be good enough.

Another example is the *availability heuristic*. This one seems to be related to the ease of recall from memory (either explicit or tacit). People have often noted what is called the 'recency' effect where someone will make a judgment that favors the last experience in a sequence of similar experiences, even if it happened a while ago. For example, clever real estate agents can show prospective buyers a series of houses in which they show the house they really want to sell last. As often as not the buyers have better thoughts about that house. Neurologically, the features of the most recent house may have damped down those of prior seen houses, which might have been even more attractive to the buyers if they could remember them. So there appears to be a consistent tendency for people to base their judgments (when decisions need to be made) on those factors that were most recently activated in memory rather than do a re-analysis of all of their memories. Again, this is actually probably evolutionarily sound. Foragers will frequently pass up smaller game or patches of food when they are not very hungry but will return to the last remembered patch when they are since it is also likely the closest in spatial terms; this even though it might not have been the largest of the ones visited. Here there is an energy saving issue. When you are hungry get to any amount of food that will sustain you in the quickest fashion. You can always go back to a prior visited patch if you need more.

As a last example of a heuristic that psychologists focus on we return to the affective influence on decisions already discussed. The *affect heuristic* is what I have claimed is the basic decision guide from our limbic systems. This is also what Damasio (1994) called the “somatic marker theory” or tagging an option with valence based on emotion-laden past experiences with the same option. Most of us know quite well what kind of trouble we can get into when we choose based on emotional feelings versus reasoning.

There are many more candidates for heuristics that are pre-programmed, so to speak, into the brain. These heuristics are all part of the ‘system 1’ (Kahneman, 2011 and chapter 1), fast, subconscious judgment process. They probably worked most of the time in early humans, or else we wouldn't be here or they wouldn't still be with us. But in our modern world they too often cause us to make serious errors in critical judgments. They, by themselves, were never able to handle complex social or cultural problem solving. That is why we developed sapience to the degree we did. Sapience strengthens the mind's capacity to override system 1 (system 2 is the rational thinking, but not necessarily just conscious thinking — chapter 1). Sapience allows us to learn much more in the way of tacit models of the world and its subsystems. Our representative concepts can be much more detailed and through sapience become ever more efficacious as we age and gain experience. Our control over availability of memories is stricter, allowing us to reflect more on not-so-available memories for comparisons. It makes us consider comparisons, seeing samenesses and differences in more subtle features. It also dampens down our affective tendencies to succumb to our wants and desires (or fears).

I believe a very useful measure of the strength of sapience may prove to be the degree to which our heuristics and biases rule our decisions in complex situations. I strongly suspect that observations of a sample of the general population will show that the heuristics are relied upon much more often than appropriate. Whereas we should find that more highly sapient individuals rely far less on these mechanisms. But that will take some significant advances in the kinds of tests and probes that psychologists have developed so far.

From Mechanical Judgments to Value Judgments to Strategic Judgments

Judgment processing covers a wide range of scope depending on the nature of the decision problem in focus. Something as simple as judging where to put ones foot in stepping forward in rough terrain involves low-level models of ‘difficulties encountered in walking’. In fact the basis for the evolution of higher level judgment capabilities probably started with this kind of background processing in early quadrupeds. Primitive cortices (paleocortex) are found in amphibians and reptiles that may be involved in such low level judgment processing.

The use of tacit knowledge and judgment in guiding current decisions is one of the characteristics of mammalian and avian life forms. What mostly differentiates creatures in the phylogenetic tree is the scope, degree of salience, and time scale over which tacit knowledge spans. For example a duckbill platypus (*Ornithorhynchus anatinus*, order *Monotremata*) is an example of an early mammal whose environment is relatively simple and whose modus operandi is fairly straightforward, even if a bit bizarre. Its brain, specifically its cerebral cortex (forebrain), is scarcely more than a shallow covering of the mid-brain (limbic brain centers). This cortex is presumed to code memories for specific places, mate, pups, good hunting places, etc. whatever is important in the life of a platypus. The frontal part of the cortex organizes attention and decisions and the rest of the cortex processes sensory data (much from the sensitive bill) and formulates learned motor responses. Judgments of this sort most likely amount to not more than moment by moment frames for guiding real-time activities.

In mammals and birds, living more complex life styles, the neocortex (in mammals) and the nidopallium[135] (the equivalent of the mammalian prefrontal cortex in birds) are much expanded and thickened owing to the vastly greater amount of knowledge these animals need to learn in order to succeed. Most of the knowledge has to be learned (vs. genetically endowed such as affect) because the complex environments are also subject to non-stationarity, meaning that individuals need to be able to adapt behavior over their entire lives.

Additional low level judgments include what we call values. Values are a set of attributes tagged with valences (good or bad) that we use to judge a wide variety of situations. The origin and cognitive processing of values is still an area begging for more research. Some values can be seen to be innate, so are probably part of the genetically mediated moral sentiment processor. Such values tend to elicit strong emotions when they have been violated. The emotional response of disgust, for example, seems to be elicited by a number of situations that can be interpreted as harmful if pursued. Evolution equipped us with an automatic revulsion of such situations (one commonly cited example is the near universal revulsion of incest). Other values appear to have been learned in the context of a particular culture. Our moral codes include locally evolved guidelines for what is right and wrong behavior. These learned guidelines generate judgments on on-going situations and even evoke emotional responses that can overcome the rational decision making system. Racial biases are often cited in this vein.

A great deal more research is needed to tease out the innate from the learned values models but the fact remains that these models, and their generated judgments, are powerful forces in guiding thinking and decision making. Higher sapience may dampen down these forces when and if it is exercised.

At an intermediate level we make judgments about a wide variety of situations in daily living. We decide to drive to work rather than take the bus when there is a chance we will have to work

---

[135] See: http://en.wikipedia.org/wiki/Nidopallium for background.

late and not be able to catch the bus home. We decide to wear that new shirt to the office because we want to look our best for the new employee that looks attractive. These are examples of judgments that involve a mix of knowledge and affect. These are the most common kinds of judgments we make routinely. And they do wonders to get us through daily life. We do not have to think long-term. We do not have to consider who all might be affected by our decisions. We simply consider the situation at hand and bring to bear a limited and restricted set of models that seem to apply.

At a somewhat higher level, however, we might think twice about taking the car to work because in the back of our minds is the higher price of gasoline and the fact that we want to save for some special reason. The judgment then comes down to simply make sure we leave the office in time to catch the bus. Another situation might involve that new attractive co-worker. Assuming the shirt trick paid off and you got a date, you might start considering the prospects for a longer-term relationship. Now not only emotional drives, but consideration of many other attributes about the person start to enter the mix. And a sufficiently sapient person might be starting to consider what life ten years hence might be like with this other person as a mate.

Still higher in the judgment hierarchy is the kind of considerations needed to think about the good of others. Coming up with worthy judgments about what is good for others requires far more complex models than just coming up with judgments about what is good for one’s self. The models have to include those of the persons involved as well as their environments, the context of their lives. Most people formulate opinions about what is good or bad for others. Most of the time these opinions are grounded in values of the person having them rather than being considered in the context of the other person’s life. They may or may not be good for the other (and too often are not). Wisdom involves arriving at judgments about and for others that are grounded in the others' circumstances and not just their own.

At a very high level of the hierarchy judgments are based on extensive models of the world, not just the local environment. These models are built with extensive knowledge, both explicit and implicit (tacit) and are constantly being modified, improved as new experiences are had. One must be motivated to grasp the ‘bigger picture’ even as they learn more. Persons possessing this level of sapience will naturally tend to think about the future, how the world will evolve, how people will be affected in that future, and what can be done now to improve the chances of good outcomes in that future. This kind of judgment formation involves both systemic and strategic subconscious thinking (see below), and this is the meaning of the overlapping ovals in figure 1.5 in chapter 1.

This rough hierarchy of judgments reflects the way in which functionality is accreted, that is added in layers of increasing complexity, in the process of evolution. I will say more about this in chapter 4, but the basic idea is that existing structures, that are responsible for certain functions, are not dropped as evolution produces more complex organisms. Rather these functions/structures remain but a new layer of function/structure is accreted to, or built on top of the existing one(s). The new functions might serve to modulate the lower level functions, i.e. suppress them when the higher level function produces a contrary result. We've seen this kind of down modulation with respect to the rational thinking system damping down the affective system in certain decision situations. Figure 3.4 shows the organization of judgment processing just described. As in the hierarchy of mind itself, we see judgment as a pyramidal structure with the low-level functions occupying the broad base and the highest level social judgment processing occupying the narrow and shallow upper peak. This again reflects the relative strengths of sapience. We might expect that in much higher sapient individuals the upper areas of judgment would be broader that this diagram implies. That is for future research to determine.

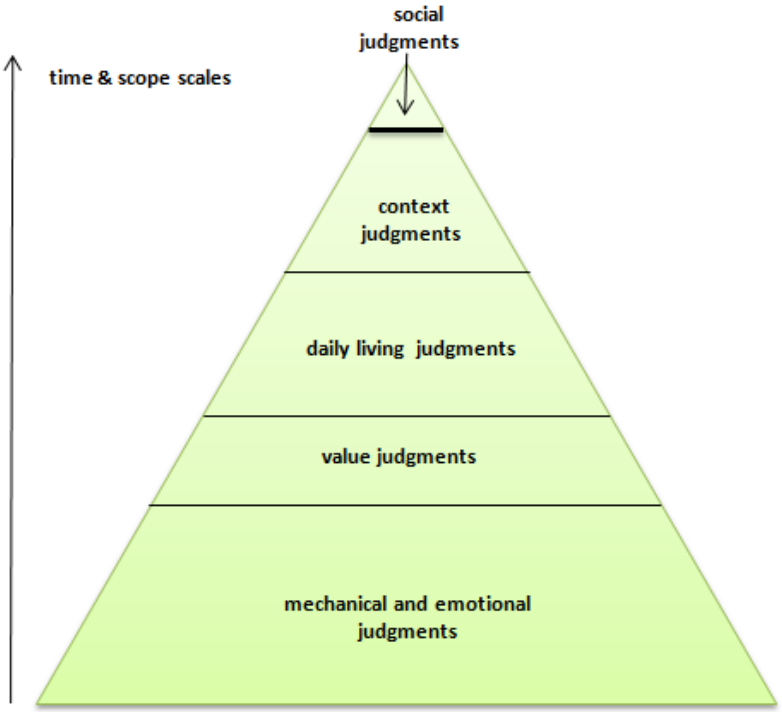


Figure 3.6. Judgment processing can be organized (categorized) in a hierarchical structure reflecting the amount of processing power given to various ‘depths’. The vast majority of judgments come from models of mechanical, value, and everyday levels of situations. Only a smaller fraction of judgments are directed at higher needs such as long-term thinking and social judgments.

Expanded Dimensions of Interaction

In the bigger, more advanced brains, judgment becomes more important in its role of guiding decisions. Decisions themselves have become much more complex along several dimensions[136]. First consider the role of time. While the duckbill might only be concerned about what is

---

[136] Horgarth (1980)

happening in real-time, more advanced mammals have to make judgments that may have an impact on their lives, days, perhaps weeks from the moment. Humans make judgments that can affect them their whole lives, and the lives of their offspring well past their own deaths.

Another dimension is physical space. Human judgments can extend to the whole world today. Indeed, by making the decision to visit the planets we may make judgments affecting the solar system. When modern humans evolved they were already capable of migrating without a specific destination in mind. They spread from Africa rapidly. But even with this ability, their capacity for building a model of much more than a few hundred kilometers radius around an individual was probably limited. The more area a tribe occupied as effective territory (the known world to any one tribe) the more other tribes and physical conditions they encountered and would have to have models for.

As personal worlds expand the number of distinct objects of all kinds will increase, as will the potential for interactions between these objects. The raw complexity of a larger world probably increases, at least as the square of the radius out from an individual's location.

Stationarity is a fourth dimension. The longer in time scale, the larger in space entailed, and the more complex an individual's world is, the more likely it is that unpredictable changes are going to happen. Murphy's Law comes into effect. No matter how large one's world might seem there is always a larger, more complex world surrounding it and interacting with things within it. And those interactions can lead to a cascade of changes in the individual's world. Think of the example of the invasion of a foreign species into an ecosystem. Change and difference is inevitable in the world if your world is complex. Species that have existed for so long, like the platypus, have survived in a relatively buffered world or a world where the kinds of changes that did occur had no direct impact on their fitness.

The situation for humans is seemingly at the extremes of time, space, complexity, and non-stationarity. We are omnivores, meaning that changes in species of plants and animals will cause changes in behavior as we attempt to adapt to the new varieties. Our memories, particularly our tacit memory system, must have a huge capacity in order to deal with all relevant knowledge needed to operate successfully in this expanded world. Moreover, it must be capable of constant and life-long learning as the accumulation of experiences act to refine concepts and their interconnections, or even revise older concepts as new evidence is encountered. As we now understand from the research, this latter is very hard for most people to do.

Once certain beliefs are encoded, especially in adults, it is extremely hard for most people to accommodate countervailing evidence and change their minds[137]. The more affective attachment to beliefs there is, such as religious doctrines and stories or nationalistic sentiments, the more difficult it is to question them and revise our thinking. This is an aspect of common sapience that keeps humans from attaining wisdom that is effective. Indeed, there is a kind of pseudo-wisdom

---

[137] For example see (Gardner, 2004), for insights into what it takes to change opinions and beliefs.

prevalent in humans who cling to old beliefs and use their tenants to guide decisions. This level of sapience was actually fine for primitive humans living in small bands and roaming over small territories. The wisdom of the elders was based on accumulated traditions (experiences) plus imagined explanatory beliefs. As long as the environment remained reasonably stable, these beliefs could serve a purpose in stabilizing the social framework of the tribe. But it is when the scope of the situation expands along the above dimensions that things go awry. Ancient belief-based wisdom starts to fail and decisions do not produce best outcomes.

One of the difficulties of complex, non-stationary environments is that causal relations are sometimes very confusing. Numerous causes can lead to the same effect. Causal chains can become causal webs. A single causal event can, due to non-linear, chaotic interactions, lead to multiple (stochastic) effects. As a result the veracity of our models of the world, as well as of ourselves, can become degraded too easily. We make poor observations of reality and any errors get encoded into our tacit storehouse. Then owing to the inherent biases of judgment[138] our judgments are further distorted from reality.

In short, our minds and our average level of sapience as a species is failing to handle the world of complexity and non-stationarity that we have helped to create. We've made wonderful discoveries in science and engineering thanks to our superior cleverness. But we have generally failed to make wise decisions regarding their exploitation in inventions and their uses. My singular paradigmatic example: nuclear weapons. Indeed the very need for weapons at all points to the massive failures of sapience at the scales beyond the tribal. Of course human evolution included the xenophobic tendencies — the “Us vs. Them” affective influence (Berreby, 2005) — that give rise to feelings of hostility toward others, especially if competition for resources is prevalent. But again, a higher level of sapience means having the ability to override the limbic impulses long enough to think things through.

Reflective Judgment — Meta-knowledge

Potter (1971) defined wisdom as "...knowledge of how to use knowledge." In other words, wisdom, as he saw it was a kind of meta-knowledge that transcended mere ordinary facts and episodes. He argued that tacit knowledge included an embedded moral aspect (see below) but also involved an ability to reflect on one's own knowledge, what one knew, what one suspected one didn't know, and what one should know to be considered knowledgeable. His emphasis was on explicit self-reflection, which must certainly be considered part of wisdom. But as I have argued, deep wisdom also involves non-conscious abilities to judge what tacit knowledge should be learned and attend to experiences that will further that mandate.

This view of meta-knowledge, whether tacit or explicit, helps to differentiate ordinary intelligence from sapience. One might be tempted to call sapience meta-intelligence, or meta-

---

[138] Marcus, (2008), gives a delightful and sometimes dismaying treatment of built-in biases in ordinary human judgment in *Kludge*.

cleverness to include the role of creativity. I prefer sapience because there are aspects of the latter that do not behave as just meta-decision making. But I will return to this subject when I cover the evolution of sapience and the brain structures involved.

The term 'reflective judgment' can be broadly interpreted as the mind (either conscious or not) reflecting on its own judgments and making judgments about those judgments! I call this 'second-order judgment'. That is, the sapient mind examines the results of judgments made, large or small, and presumably guides learning to modify or refine the tacit knowledge that gave rise to the original judgment. Judgments must be judged.

Somewhat more narrowly reflective judgment entails epistemic cognition or thinking about knowledge[139]. Wisdom is often described as including an ability to make complex decisions in the face of uncertainty. Uncertainty is inherent in incomplete knowledge so a wise person understands that the risks associated with any complex, especially socially pertinent, decision is due to an incompleteness in their own knowledge. How gracefully the person deals with such uncertainty, that they continue to apply their best judgment without claiming that it is based on absolute truth, for example, is a property of sapience. The sapient brain knows when it is dealing with incomplete or even inconsistent models of the world and adjusts its judgments accordingly.

This too is a guide to what needs yet to be learned. Uncertainty might be reduced in future such situations by learning more knowledge about the characteristics of the current situation. Models can be improved so that better decisions are made in the future. Or, and this is the harder problem, if outcomes from judgments do not meet expectations, then models may need to be altered. This requires determining the failure and taking steps to correct the model for future use. Many ordinary people find this very difficult or never really realize that it is essential in order to become wiser.

So, sapient judgment involves a self-monitoring and meta-judgment capacity that is essentially automatically invoked when dealing with uncertainty. Its actions should generally result in improving tacit knowledge models over time and multiple experiences. People who never reflect on their own progress toward wisdom (or evaluate the efficacy of their judgments) and never seek improvement are doomed to foolishness. An old (I think Chinese) paradox says that if you think you are wise, then you are not, but if you seek wisdom, then you are!

Moral Sentiment and Guidance to Reasoning

There has been a tremendous spurt of research into the basis of moral and ethical behavior in the last several decades. Neuroscientists, psychologists, and anthropologists have begun to unravel the neural substrates and universal behaviors associated with moral sentiments. And the study of

---

[139] Kitchener & Brenner (1990)

the evolution of moral behavior has demonstrated that the human capacity for moral reasoning is innate and has developed extraordinarily in the genus Homo.

The literature on the subject is extensive and I cannot begin to do the subject justice in this short work, though I have included a large set of reading resources in the bibliography. So I will focus on some of the main points that tie in with sapience. The keys to understanding the underlying motivations in sapient minds are the evolution and benefits of true altruism and cooperative attitudes that have allowed humans to form trust-based alliances even with non-kin and strangers.

At the core of the more sapient dynamics of social interactions is the sense of fairness and justice that helps maintain a generally well-functioning social network. Fairness sentiments have recently been demonstrated at work in lower primates. Details of how humans experience justice and fairness have lately emerged from behavioral economics where many of the classical economics assumptions about rational agents have been found baseless or called into serious question. Models of humans operating under innate sentiments of fairness and justice have helped explain a good deal of human behavior that is otherwise puzzling under theories of human nature that depend on rational thinking[140].

But sentiments of justice and fairness have a dark side that, if not controlled, can lead to a breakdown of social structure. That is retribution, the fairness sentiment that demands punishment for cheating and immorality. The invention of rule of law has been one of mankind's greatest achievements whereby people feeling cheated are restrained from wonton retribution. The governing jurisdiction (state) takes responsibility for punishment and when things work right, the accused are afforded due process to determine guilt or innocence lest retribution be wrongly taken and lead to further conflict. Of course even the rule of law is no guarantee that things work properly. There are always tyrants and cheaters among the judges who subvert the process for their own gain. The proper execution of the law, in the end, depends on the judgments of judges! This is a critical aspect of societies. Judges who have the power to condemn or free must have good judgment. Unfortunately, if my conjecture about the rarity of higher sapience is correct, then our court systems, dealing with unprecedented numbers of cases and requiring a greater volume of judges on the bench, is doomed to poor judgments more often than not.

And that is the difference between strong sapience and the lesser kind that seems to be our lot. Truly sapient individuals seem to have the ability to down-modulate their own desires for revenge and thus have greater control over their more affective (limbic) reactions to cheats and

---

[140] The "Ultimatum game" is used in experimental economics work to test ideas about what humans (and other primates for that matter) consider fair and just. It has revealed that many people would just as soon lose any possible gain by refusing to take a lesser offer from a partner player who has been given a sum of money and told they may offer the receiver partner whatever amount they want. Many people will offer 50% on average, but those that offer less than around 10-15% are often refused, which then means neither player gets a reward. See the Wikipedia article http://en.wikipedia.org/wiki/Ultimatum_game , for a good background on the game and how it is used. It also has many good links and references for those who want to dig deeper.

criminals. This doesn't mean that sapient individuals do not experience anger and desire for revenge or retribution. It simply says that stronger sapience somehow controls those limbic-based urges and keeps them in check so as to make decisions more wisely. Recent imaging studies of brain functions have shown that some individuals do seem to have more prefrontal activity during episodes of exposure to cheating or perceived unfairness, correlated with more restrained decisions. We also have evidence from neuroanatomy that inhibitory efferent fibers from the frontal cortex to various limbic areas, including the amygdala, act to dampen emotional responses to events in order for the prefrontal cortex to 'consider' the situation before acting.

Altruism, Empathy, and Caring/Sharing

A starting point for understanding moral sentiment as an underpinning of sapience is to see the role that altruistic motives play in human life. Our whole sense of wanting to be good, and help others starts with this basic biological mechanism for ensuring the success of tightly bound social groups in out-competing other groups. It is an evolutionary argument but it takes an interesting twist at the human level when altruism turns into empathy and caring with intentional altruistic behavior.

A fundamental premise of altruistic behavior is that one individual is willing (or compelled) to sacrifice its self-interest (or, more correctly, its fitness) for the benefit of one or more conspecifics. Sober and Wilson (1998) describe altruistic-like behavior in a nematode parasite (page 18). Other researchers have observed altruistic behavior in numerous species at all stages of evolutionary complexity.

For some time evolutionists had wondered about altruism and how it could have come about. Darwin himself had misgivings. It seems obvious, on the surface, that altruistic behavior would reduce the fitness of an individual (by exposing them more frequently to life-ending situations) and so selection would have weeded it out. But it is so clearly engrained in so many species that it must have an evolutionary purpose. Researchers have described various degrees of altruism and have generally provided satisfactory reasons why they would be favored. At the lowest level is the theory of kin selection, which basically posits that an individual's gene's chances of showing up in the next generation are improved if that individual ensures that close relatives are taken care of or protected, even if the individual itself does not reproduce. This mechanism is used to explain why female worker bees don't bother to reproduce. The principle may extend to tribes or colonies in some species where distant related individuals sound alarms when a predator is spotted, thus increasing their risk of calling attention of the predator to themselves but ensuring that their cousins take cover.

Somewhat more inclusive (beyond kin) is the theory of reciprocal altruism in which members of the same community are willing to sacrifice themselves or take non-reproductive roles even when there is a weak genetic connection between members. Presumably there are strong benefits of other kinds in tight social networks such that it still increases the fitness of the group to have this kind of behavior even in the absence of direct genetic benefit (see Sober and Wilson, 1998 for a model of group selection providing a solution).

Many social psychologists have not been able to admit that the above mechanisms apply to human altruism, or what some have labeled 'true' altruism. While cases of people jumping into a frozen lake to save a stranger may have some basis in the above described, seemingly automatic reactions, what appears to most of us as genuine conscious caring for others goes far beyond mere kin or group fitness improvement. There must be something more that produces these feelings in humans. And that may be the role of sapience. The key being that true empathy emerged as a function of social-biological coevolution.

Recently neuroscientists have discovered a remarkable kind of neuron in the brains of primates and possibly a few other non-primate mammals (and even birds). These neurons, called mirror neurons, have the interesting characteristic of firing both during the performance of an action by an individual and during that same act being performed by another individual when observed. These neurons, and in fact, systems of these neurons have been identified in human brains embedded in several higher order perceptual and integrative processing areas. This suggests that humans are capable of entailment with respect to much more subtle behavior by others, such as facial expressions conveying emotional state information of the other. There have been several studies that show that human subjects experience emotional mirroring wherein they not only grasp the emotion being expressed by others but actually experience a mood change in the direction of that emotion. The general phenomenon is what we would call empathy[141].

Altruism based in deep evolutionary roots of the brain along with this new mechanism of empathetic coupling may go a long way to explain so-called 'true' altruism. Other mental factors may still be active in some forms of altruism. It is suggested that a main motive for why people give to charitable causes is that they get a mental reward for doing so, thus suggesting there is no such thing as 'true' altruism. However, I do not see how these various mechanisms are mutually exclusive. Both empathetic-based behavior and subsequent reward are perfectly compatible means of reinforcing altruistic sharing and care.

Regardless of the details, it is recognized that deep caring for others is a core trait of wise people. Wisdom involves understanding that sharing and caring are at the heart of viable social groups. And, I suspect that even low levels of sapience involve a basic tendency toward more of this than being uncaring and selfish[142].

---

[141] See Goleman (2006, pp 40-43) for an introduction to mirror neurons and their possible role in empathy.

[142] It may seem hard to reconcile this idea with the major sentiments we see prevalent in our modern world. The neoliberal, *laissez-faire* market, capitalistic model of an economic system praises selfishness and self-interest in the name of economic growth using Adam Smith's metaphor of the 'invisible hand' as a justification. However if we keep in mind that the modern view is based on a world over-full of people who are competing for diminishing resources (in other words under stress) then we might understand why this is so. Taken out of the rat race of modern economic life, I assert, many people would revert to the biological norm of empathetic caring and sharing.

Yet selfishness is a problem in our societies today. The basic human propensity for sharing and caring might be muted when the social domain exceeds a limit based on sheer numbers and kinds of people encountered. Mankind evolved in a world where tribes rarely exceed several hundred individuals. It isn't unlikely that our subsequent evolution selected for those able to accommodate larger groups and strangers. After all, we have been living in such groups, villages, towns, states, etc. for five to eight thousand years. And we have adapted culturally, if not biologically, somewhat to those conditions. But the rate of change wrought in the information age has surely exceeded our abilities to accommodate the myriad strangers we encounter today. Indeed issues of xenophobia, ethnic conflict, etc. may have their roots in the fact that we have exceeded the number of others that we can extend our sharing and caring to.

One test of sapience might very well involve determining the extent to which a person can feel empathy toward strangers, and how much caring can be extended to different kinds of people.

Justice and Fairness

What is right behavior? What is wrong behavior? Are there universal rights and wrongs? And what makes it right or wrong in the first place?

Not that long ago the general belief was that different cultures around the world had different beliefs about what constituted right and wrong behaviors. The argument went so far as to claim that there were no universally held beliefs about right and wrong and therefore cultures should not be judged on the basis of their mores. It was all relative. And the evidence seemed solid. Even though, for example, most people throughout the world viewed incest as wrong, a few cultures, including some western ones, have practiced ritualistic incest (e.g. Hawaiian royalty marrying - brothers and sisters - to maintain the royal line). Similarly cannibalism is repugnant to most societies, yet some tribes have practiced ritualistic cannibalism for religious reasons.

However as the science has progressed it is becoming clear that there is something like a universal semantics of moral/ethical behavior in a manner not dissimilar to the universal disposition to language, e.g. speaking, hearing, and signing. All humans have a sense of right and wrong, even if the specific instances of what counts as right or wrong vary from one culture to the next. Moral sense is innate.

One of the clearest pieces of evidence for this innateness comes from experimental work with monkeys and apes which demonstrate a built-in sense of fairness[143]. Fairness involves a relational observation between the subject and others in the group. For example, when one individual perceives another getting an unearned reward he/she will tend to feel resentment toward the receiver if there is no apparent agent giving the reward (perceiving the recipient as a cheater). Or the observer may feel anger toward the agent that provided the unearned reward. It also works on the punishment end. If an individual observes a perceived cheater being punished,

---

[143] De Waal (2005, chapter. 5)

then he/she feels satisfied that this is an appropriate outcome — in other words justice has been served.

The fairness sentiment can also lead to jealousy. When one individual perceives another winning a reward he/she can feel jealous but not feel anger since the reward was earned. In a good way jealousy might lead the individual to efforts to seek a similar reward to even out the balance, to get one's fair share. Of course that can lead to frustration if the reward was actually a result of chance and not truly earned.

The main point is that our sense of what is right and wrong starts with an innate sense of fairness. Behaviors that help the members of the group achieve a fair balance of resource sharing, for example, are associated with right actions and lead observers of those actions to have favorable memories of the actor. Similarly, behaviors that unbalance the resources or harm others are perceived as wrong and lead observers to have negative memories of the perpetrator. Cheating is defined by this criterion, when an actor derives benefit unfairly. And feelings of retribution follow when the cheater is caught. All of this sense of righteous and moral sentiment plus many more related sentiments arise from innate mechanisms in the brains of social animals. And that includes humans. Our moral sentiments are grounded in innate senses of fairness and right and wrong actions (relative to the individual).

But another question that should be asked is: If fairness is innately based, how can there be cheaters and sinners in the first place? And that is a critical question to ask. The answer is likely grounded in evolutionary theory and the inherent variation in gene alleles in a population. At a more transcendent level of social life occasional cheating may be advantageous in terms of the exploitation/exploration trade-off that the evolutionary algorithm is always manipulating. Cheaters are in a sense a kind of exploration of the space of possible behaviors while conformers are exploiters of good behaviors as defined historically. Every so often the environment may change in a way that some form of cheating behavior leads to a survival advantage that assures at least some members of the species show up in the aftermath. Or it could lead to a behavior that actually helps the group.

So, innate cheaters are a consequence of normal variation in the population. Under ordinary circumstances the cheater's behavior is not helpful to the group and so mechanisms for detecting and punishing cheaters are necessary to maintain group cohesion and functioning. Hence justice. The sense that a cheater has been punished is part of the package. In humans we find a spectrum of the concept of justice with regards to the protection of non-cheaters for whom some evidence suggests they are cheaters. This problem probably doesn't arise much in non-human apes and other social mammals. But human social structures are complex and, as I have asserted, the causal chains are often obscure so that it is hard to abduce the cause of an innocent being caught looking like a cheater. Once again the human propensity to create laws that protect the innocent until evidence can be examined fairly provides a way to mitigate injustices. Of course the various laws and institutions for applying them are as imperfect as their creators and so they are no guarantee that justice will prevail. And cheaters can be found within complex institutions using laws and procedures for personal gain (picture the district attorney who is anxious to get a conviction so as to promote getting re-elected).

Cooperativity

Social animals cooperate with one another to get work done that benefits the whole group (recall this from the section above, "Second and a Half Consciousness"). Far back in evolutionary history, when organisms reached a significant level of complexity, the variability in alleles for genes involved in various aspects of physical structure and behaviors created situations in which specialization in "talents" made it possible for different individuals to become extra efficient in some tasks that contributed to the whole enterprise. Such specialists were selected for in the sense that the groups containing them were more fit. This is another form of group selection mentioned above.

From a systems perspective, improved talents are emergent at the level of the individual but quickly give rise to auto-organizing forces within the group[144]. The specialists tend to occupy the task performance roles and come to dominate them such that a new organization tends to emerge at the level of the group as a whole.

It would appear that early in evolutionary history a tendency for social animals to auto-organize into cooperatives so as to maximize their fitness by letting specialists be responsible for specific tasks. One can imagine in the early tribal context of Pleistocene humans some individuals having particular talent for making spears and other tools spending more of their time doing that and letting those more adept at hunting and gathering taking care of the food-getting tasks. As humans invented more varied technologies and tribal behaviors, especially after the invention of agriculture specialization and task efficiency became a dominating theme in social organization. Farmers farmed and craftspeople crafted. Farmers started specializing, some grew grains others husbanded animals for food and work. Craftspeople specialized. Some built houses, others worked metals. Self-sufficiency, except at the most primitive level of existence, became increasingly non-viable.

But none of this divergence in specialization could be possible unless people were able to cognitively (even if subconsciously) recognize the need to let specialists do their things and cooperate, particularly by trading products of those different skills. Commerce as we know it is only possible because humans possess a fundamental drive to cooperate with others whom they recognize as being skilled at doing something better and needing that something to sustain life. Once started down the track of technological invention the complex interplay between genetic variation, technological specialization, and moral sentiments became a self-fulfilling cycle, a

---

[144] See chapter 10 of *Principles of Systems Science* (Mobus and Kalton, 2014) for an explanation of the cycle of emergence, auto-organization, and subsequent emergence at a higher level of organization.

feedback loop that accelerated the pace of development. From the time of Paleolithic humans to our present situation, civilization essentially exploded into existence.

Cooperativity, much more so than empathy or altruism, is the basis of human eusociality. Just as with the latter two tendencies, cooperation is built into the brain[145]. It is a sub-component of sapience and in more strongly sapient individuals the desire to cooperate is particularly strong. Of course part of wisdom is realizing when someone with whom you would wish to cooperate is not so similarly inclined toward you. The better part of wisdom then is to go about your own business and let it go.

Emotional Control

The final piece I want to cover here involves the capacity of a sapient individual to dampen innate emotional responses to cheating and immorality in others. As mentioned above, this facility is what allows man to formulate laws and procedures to protect innocents. Handling the capture and punishment of cheaters and sinners requires dispassionate observation of the recoverable facts of the matter before finalizing a judgment. It is no accident that we look for some level of wisdom in those we elect or appoint as judges in our judicial system[146].

This control of the limbic responses to external events by the frontal cortex is found in all primate brains. But in humans it has reached its greatest effectiveness. There are many more efferent and afferent fibers connecting various areas in the frontal, and especially the prefrontal, cortices with numerous limbic nuclei in humans than in other apes[147]. The prefrontal cortex monitors limbic activities and acts to dampen the motor responses until the executive functions in the frontal cortex have time to evaluate the correctness of the limbic response. I will be providing some more detail in the chapter 4 regarding the neural basis for this ability in sapient brains.

---

[145] The term "hyper-sociality" is being used more frequently to differentiate the form of human sociality from that of other eusocial species such as ants, bees, and naked mole rats. I think the idea is to imply that humans are even more social than the 'true' social animals. However I'm not in agreement with such an implication. Humans are indeed highly social and meet most of E. O. Wilson's (2013) criteria for eusociality. But they are definitely not eusocial in the way that ants and bees are. The latter are often described as essentially automata, following rules blindly. Humans actively think about their cooperation and have the feeling of freely deciding how and when they will do so. More on this in chapter 5.

[146] It is becoming increasingly hard to find such people. Even the highest level of the courts in the United States, the Supreme Court, is now peopled with numerous ideologically driven individuals. I have no doubt they are well-intentioned people. But they are now observed to largely make decisions on the basis of their political and economic, if not religious beliefs, in spite of the evidence. The case of Citizens United (http://en.wikipedia.org/wiki/Citizens_United_v._FEC) has been cited as an egregious decision promoting a political agenda rather than a dispassionate interpretation of the US Constitution's First Amendment protection of free speech.

[147] LeDoux (1996)

Systems Perspective

The main hypothesis here is that all brains, even the most primitive, are evolved to perceive some level of systemness. A worm brain with its zeroth-order and only a hint of what would become first-order consciousness is able to perceive inputs to itself and regulate its own body state relative to those inputs. A fish brain (zeroth- and first-order without much in the way of learning beyond basic habituation or sensitization) is able to form images of external objects and perceive their main features insofar as they have relevance to the fish. It can process inputs from those objects, regulate the fish's internal states, and take relevant actions (eat, reproduce, or run) based on the features of those objects. It has a very primitive sense of self and other. A reptile brain does all of that and more first-order processing in the form of learning some new but still primitive images and relations (a hint of second-order consciousness emerging). The mammal brain, with its newly acquired neocortex, has reached the second-order consciousness. It has a model of itself and is capable of constructing models of others of its kind as well as non-kind. It can learn to adapt to varying environments to greater or lesser degrees, but certainly more so than the reptiles.

All of these levels are using neural substrates to process system models, either hardwired into the brains or learned from experience. In all cases the brain deals with inputs and outputs to/from the self and with representations of external conditions to which it responds. Animals perceive systemness from a primitive form up to the human capacity to perceive nuances. Full second-order consciousness has the ability to perceive external objects as systems, i.e. to see the objects receive inputs and produce outputs. Human, 2½-order consciousness is able to model what goes on inside those objects in order that they process their inputs to produce their outputs.

The modeling of systems is accomplished by virtue of a mental language of system that the brain "speaks." That is, the brain (of humans in particular) is evolved to construct (learn) percepts and concepts that represent models of system components like boundaries, networks of relations and the like. In Mobus & Kalton (2014) we describe a modeling language that can be used to construct computer-based simulations of systems[148]. The language of system that is used in the brain uses neural representations of elements such as given in chapter 12 of Mobus & Kalton (2014).

The Language of System

Philosophers of mind and linguists have proposed that the mind uses an internal (private) language that is more primitive than spoken (public) language, but must have structures related to public language, namely symbols (lexicon), syntax (grammar), and semantics (meaning)[149]. There has been a great deal of speculation about what this language of thought might be, what it

---

[148] Chapter 12 describes the lexicon for a system decomposition description, i.e. a language for describing a system as it is being analyzed. Chapter 13 discusses various system modeling language approaches.

[149] Fodor (1975); Pinker (1997), p69; Schneider (2011). Also see Deacon (1997), chapter 1 for a discussion of the various theories for language competence in humans vs. the evolution of language.

must be like, and how would it have evolved in humans to produce our kind of language and competence therein. There have also been many criticisms about the existence of such a language[150]. This language, if it exists, has been somewhat whimsically termed "mentalese." In this section I propose a hypothesis about what this mentalese might be that resolves some long-standing questions about its role in thinking, public language production and understanding, and how it evolved in brains as they became more complex. In the next chapter I will demonstrate how concepts as *symbols* for objects (nouns), actions (verbs), and relations (sentences and thought streams) are constructed in neocortical neural networks. And I will demonstrate how the activations of these networks due to bottom-up sensory stimuli (perception) or top-down recurrent activation (imagining) works in the brain to produce thoughts that translate from massively parallel production to sequential entry into working memory and consciousness. That is, I will tackle the problem of how natural or public language arises from mentalese. But first I will establish where mentalese itself comes from.

The current proposal is that indeed the brain speaks a primitive language, but the underlying language is that of *system*, may I call it "systemese?[151]" The brain contains mechanisms for representing system symbols, syntax, and semantics (explained in the next chapter). In this section I want to explain the psychological aspects of mentalese as the language of system. In chapter 5 I will return to the subject from the standpoint of the evolution of mentalese, as I contend that the language of system is inherent in all brains starting with the simplest (i.e., from zeroth-order consciousness). There are two senses of the use of the word "primitive" in describing the language. The first sense is that of being minimal in capacity to describe the world. A zeroth-order consciousness language speaks only input-self-output (self is the system). A first order consciousness language adds to this some additional terms for external sources and sinks as well as a few additional internal regulations. Second-order consciousness language adds to this a more elaborate (hierarchical) control set of predicates. Finally, in humans a much more elaborated set of predicates and objects (including proper nouns) gives rise to the public language we speak and hear (and write and read).

The second sense of primitive is based on the notion that mentalese is private (even to consciousness ordinarily). Public language is a combinatorial expansion of the primitive symbols of mentalese. That expansion is able to combine spatial, temporal, relational, quantitative, affective, and, recursively, combinations of those combinations. Presumably, this is only possible because our form of mentalese is nevertheless capturing enough of what is relevant in the real world to express all aspects of it. I infer that to mean that our mentalese is a complete systemese

The systems perspective that is a component of sapience is based on system mentalese, which means the brain naturally and automatically "speaks" systemness. In humans this primitive

---

[150] Strong connectionists, for example, deny that the brain can represent "symbols," argued by language of thought proponents (or those who insist on a computational theory of mind – Schneider, 2011). Symbols and syntax are considered necessary elements of an "language."

[151] See also, Mobus & Anderson (2016).

language, in the second sense, has achieved a substantial capacity to represent all aspects of systems of very high order (complexity and emergent properties). This should not really be surprising. Systemness is the main organizing principle of the Universe (principle 1 in chapter 1). That the brain has a natural language that reflects the way the Universe is organized, while not obvious, should not be surprising. Evolution has produced brains that are able to describe the rest of the "relevant" world. Our brains are evolved to be able to recognize and reason about reality as it is in the world. By relevant I mean the immediate environment of a brain of a given complexity. A worm's brain need not be aware of too much more of the world than the dirt and similar other organisms in it that it crawls through (primitive in the first sense). Human brains need to be able to represent much more of the Universe and manipulate those representations to produce descriptions of the past state of the world, the present state, and possible future states. The human brain has evolved a greater capacity to describe more aspects of the Universe (primitive in the second sense).

Concepts

In the next chapter I will explicate more thoroughly the *concept of concepts*, especially as they are represented in mental tissue[152]. Here I want to somewhat anticipate this but in the realm of psychological phenomena. A concept is a complex system in its own right. It is an encoding of systems in the world, and in humans, encodings of systems that are not necessarily in the world. A concept is composed of an assemblage of correlated percepts that are, in turn, assemblages of correlated features. Concepts are represented in patterned *networks* of neurons (principle 3 in chapter 1) that are connected by strong memory traces and co-activate. They are composed of sub-concepts (subsystems, principle 1). There may be a deep hierarchy (principle 2) of concepts that are simpler toward the base and more complex toward the apex of this hierarchy. For example, the abstract concept of a dog is a very high level one that is composed of a large number of sub-concepts such as characteristic behaviors (barking, baring teeth, sniffing indiscriminately), and body forms (highly variable in dogs compared with other mammals but with enough commonality across the species that they can be recognized, e.g. overall size – St. Bernard vs. Chihuahua – does not mask the main features that make a dog a dog). The concept of a dog includes relation to a more abstract concept, that of mammal. Mammals have general characteristics that are found in dogs such as fur. Dog fur is different from sheep wool but both share features that allow fur to be recognized (as opposed to feathers in birds).

The base of the hierarchy of concepts ultimately arises from aggregates of features. These are fundamental sensory inputs such as color, texture, motion, and so on. All objects have consistent sets of features that make up basic concepts. Features are perceptible but not directly accessible to conscious processing except through their participation in forming primitive concepts. In the next chapter I will show how this works in the brain, showing how low-level feature detection

---

[152] Mobus & Kalton (2014), chapters 5, 6, and 7, "Think Boxes" provide descriptions of concepts and how they are formed. Chapter 4 of this book provides essentially the same information.

gives rise to basic percepts, which in turn give rise to primitive concepts and those, in turn build up to more complex concepts, both specific (like my dog Fido) and abstract (like dog-ness).

Thus far I have been describing object concepts. It turns out that the brain also encodes actions and relations, e.g. objects doing something, objects doing something to other objects, and objects being in relation to other objects such as being on top of or below, etc. Figure 3.7 shows a schematic representation of this. Features of different objects and relations between objects contribute to percepts that, in turn, contribute to the encoding of concepts. In the figure an internal representation of object 1 (concept of object 1), a representation of an asymmetric action-relation, and a representation of object 2 mirror the actual objects and their relation in the world. The concepts as actual physical embodiments in neural tissue are essentially tokens of the objects and the encoding of the relation embodies a subject-verb-object structure. In other words, the brain is encoding sentences using natural tokens. This is going on in animals that have no public language (non-humans) with brains able to form representations of things and relations, generally speaking those with cortical structures able to learn, but which only have a more primitive (in the first sense) mentalese with which to think. But in even more primitive animals the tokens of objects and relations are encoded in neural structures but not by individual learning. Rather evolution of the brain produced instinctively coded representations of those things and relations that are important to the fitness of the beast. For example, we know that frog brains do not detect a flying insect as an insect per se. Rather it has an automatic recognition of particular features, such as an object of a certain size range flying by at a certain speed range. The frog will then automatically react by shooting out its tongue to capture the object. The frog's brain evolved to recognize these features and react accordingly. Internally the brain regions that do the recognizing speak to one another in this extremely primitive mentalese that effectively describes a system; the external object, the act of moving, and the "meaning" that this represents food.

Figure 3.7 depicts the encoding of representatives of very concrete objects and a concrete action. The pinkish ovals represent neuronal networks that have encoded three different aspects of what has been observed in the world. The relation (a predicate – "acts upon") is encoded by virtue of the causal perception (Mobus, 1994). The networks are activated in a temporal order, concept 1 first acts, that action then activates concept 2, which, in turn changes its behavior in the world (not shown). This causal direction establishes the actor and the patient relation. The neural representation of the action initiated by concept 1 is, itself, a concept.

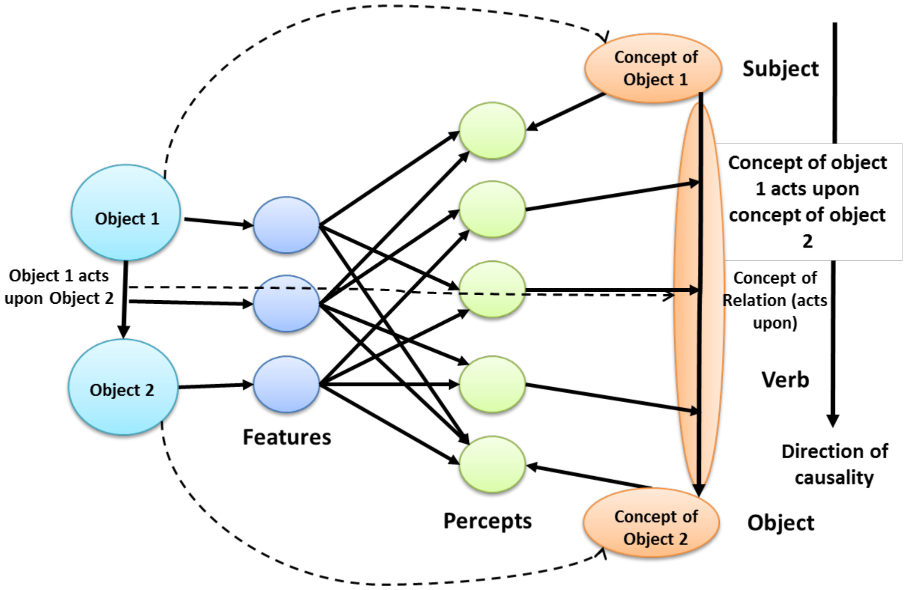


Fig. 3.7. Encoding of concepts in neural structures provides the tokens upon which the mentalese of system operate. Here is represented concepts of concrete (real) objects and an asymmetric action-relation detected by the brain structures described in the next chapter. This encoding scheme is the basis for forming yet higher order concepts.

By observing a large number of instances of objects that are the same as or sufficiently similar to the two objects in this instance to form a new higher concept of, for example, Object 1-ness – the set of percepts that are common to all encountered instances of objects like object 1 – which then represent the possibility of other instances of the relation depicted[153]. Figure 3.8 depicts the situation where a number of instances of similar objects give rise to a concept of the *category* of those objects. The same is the case for actions and relations. This is what we mean by generalization and categorization. The categories formed are necessarily fuzzy, i.e. there is a membership function that is employed to allow different objects to have variations in the features/percepts or sub-concepts that pertain to a specific instance. This is why we can have an instance concept of our pet dog (Fido) that is a member of the concept of dog-ness that includes other instances of dog with different features. In a similar fashion dog-ness is a member of a category we call mammal and there is a set of fuzzy features that make up mammal-ness.

All of this representation takes place in brains capable of forming these concepts. But note that the systems mentalese gives rise to descriptions of the world experienced. Concepts are actually models of the systems encountered in the world and the kinds of relations they can have with one another. The language of system is what is used to construct (or inherit) these models in the brain and have them available to help interpret the world in the future. They can also be used to simulate the future states of the world in animals with very advanced neocortices.

---

[153] This is the result of inductive learning, i.e. learning to represent a set of objects (and relations and actions) as a category. The higher order concept is a fuzzy category.

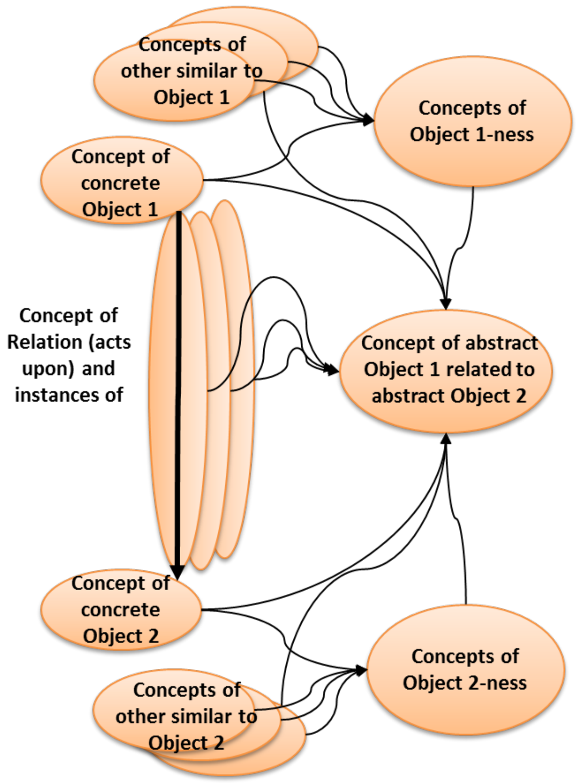


Fig. 3.8. Constructing higher order (more abstract) concepts that capture the basic structure of systems comprised of Object 1-like objects, Object 2-like objects and multiple instances of them in the action-relation (acts upon). At a higher level in a larger hierarchy the brains of more evolved animals (those with larger and more developed neocortices) can form concepts of the general form. The consistency of the acts-on relation on multiple observations reinforces the encoding of the generic (more abstract) relation concept.

Symbols and Symbol Grounding

Concepts as encoded in neuronal networks are effectively symbols representing all lower level concepts that go to make them up. In other words, the neuronal network encoding the concept of dog symbolically represents all characteristics of dog-ness. As I will show in the next chapter, the abstract concepts (believed to be encoded in neuronal circuits in the prefrontal cortex) can be activated by excitation from a motor sequencer-like circuit. Its activation corresponds to bringing the concept into working memory where it can be "processed" in the sense of a sequential computation (see below re: syntax). In this sense, the higher order concepts act as symbolic representations subject to language processing.

Higher level concepts, if strongly activated in this way, can send activation signals to their component lower-level concepts, in fact this recurrent activation can proceed all the way to secondary sensory cortical neuronal networks bringing the "thought" of the concept into greater vividness in consciousness.

The fact that concepts are based on lower level concepts all the way down to sensory features answers the symbol grounding problem (Harnad, 1990)[154]. All symbols (concepts) are grounded in sensory experience. Since written symbols are also sensory perceptions that feed higher order concepts in the human brain (e.g. the symbol 'Σ' appearing on paper is linked to a predicate concept, summation, in a discrete math context when it has been learned to so represent the operation) they too fit this model. The role of symbols in a particular sentence construction is based on the syntax of the language. However, when symbols such as concepts in the brain are previously grounded by inductive learning then their use is constrained by their meaning (semantics) and not merely the syntax (see footnote 36).

Syntax and Semantics of Mentalese

This is where systemese plays its role. The objects, actions, and relations described by systemese restrict or control what can be "said" in mentalese. Lexical elements of systemese include objects – bounded processes seen from the outside – and flows – actions that influence recipients. The syntax of systemese prevents us from saying that a flow (an arrow in the diagram) goes from a process to nowhere. There has to be another process or sink recipient (actor-patient relation). Thus there are construction rules for how the lexical elements can be combined (e.g. "Product flows from process to customer", or in predicate form, "Flows-from-to, process, customer"). The brain representations of these symbols can be processed in several orders but end up in the same relation. One could observe an object (the process) outputting some substance, then see it 'flow' to another object arriving at a slightly later time, and then seeing the second object 'accept' the flow. It is also possible to observe the flow of something, follow its channel backward to find out where it is coming from and then follow it forward to see where it is going. Or, the reverse. Either way, the brain observes cause and effect, actor and patient.

What remains is to connect this system mentalese to the human capacity for public language – the capacity to communicate concepts between individuals. As far as we know a dog cannot tell another dog about the humans he lives with or the nice comfortable bed he gets to sleep in such that the other dog forms similar concepts in his brain and can imagine the first dog's life (and possibly feel some jealousy for it). One of the most prominent differences between human beings and all other animals is the capacity to transmit concepts through a symbolic generative language. The ability to replicate concepts in other brains requires a common background for both sender and receiver since concepts are built from lower-level concepts or percepts. Lower down in the hierarchy of concepts/percepts and especially features, there are a huge number of common experiences that can be brought into the construction of the higher-order concepts. For example, I can tell you about an extraordinary sunset that I witnessed by describing the colors of

---

[154] The symbol grounding problem arises when one takes a strict computationalist view of cognition. Symbols in a formal language are basically arbitrary signs and their role is determined entirely by the syntax of the language. For example, the rules of simple syllogism, X is Y, A is X, and therefore A is Y, in which A, X, and Y might be substituted by any symbols without changing the inference relation, divorces the symbols from any intrinsic meaning. The meaning, semantics, must be supplied by the language users.

the clouds and their shapes and you can construct an imaginary image of such a sunset based on your experience with similar (but possibly less intense) sunsets. You share with me the notion of the colors red, yellow, orange, blue, etc. even if each of us has had many different instance experiences of these colors.

For language to work you first need these common constructions and an ability to recombine concepts to form new systems configurations (e.g. an object with which you are familiar that is taking in inputs different from what you have experienced in the past, or a system whose behavior you have not previously encountered). These capabilities are there in system mentalese but what is needed is a way to communicate the concepts in an efficient and compact form. That is accomplished by words and sentences.

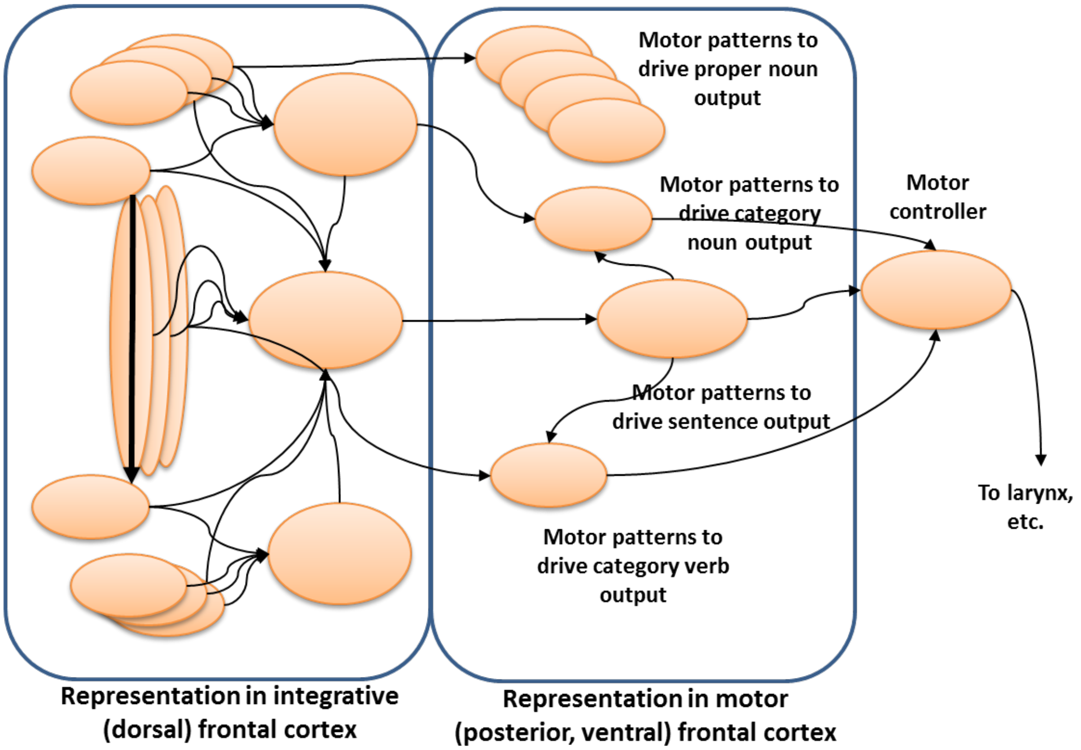


Fig. 3.9. The language production (not generation) part of the brain, Broca's area (see next chapter) forms concept-like structures that are used to encode concepts into motor patterns. These patterns are learned by association with auditory circuits and concepts (high order) in the integrative part of the cortex. These are essentially the inverse of concept patterns of neural circuit in that they are formed with exactly the same learning mechanisms. They control the vocalization machinery to produce sequences of modulated sounds we call "words".

Figure 3.9, above suggests how this is accomplished in human being. The posterior frontal cortex specializes in learning motor patterns for behavior control. A special portion of this brain area, called Broca's area, has evolved to learn patterns for controlling all of the (also) evolved machinery involved with vocalizations. These patterns are learned (constructed) using the same pattern learning that operates to form concepts in the integrative portion of the frontal cortex (dorsal, anterior). The learning is driven by associations with auditory inputs correlated with the formation of concepts. This takes place most dramatically in children who are acquiring their native language. In the next chapter I will bring this down to the level of brain circuits and modules where language is actually processed.

Mentalese becomes the mechanism for developing the capacity to encode complex concepts into short, small motor patterns that result in the emission of words and sentences. Of course there is a yet higher-order control on what words to actually produce at any given time. This is provided by areas of the prefrontal cortex – deciding what to say and how to say it. I will say more about this in the next chapter. What I wanted to accomplish here is to make the case for how a system mentalese can support the development of a public language. This is part of the emergence of sapience. It became possible when the frontal lobe became sufficiently large such that some portion of it could be dedicated to vocalization patterns.

In a fairly similar way, the auditory processing of the anterior parietal portion of the brain, part of a large percept integration region, could begin to encode sound patterns and sequences that constitute human speech sounds. In both speech production and understanding the key element is to be able to encode such patterns as a result of associated important perceptions such as the presence of other human beings. In particular the presence of a mother (or equivalent care-giver) for a young child creates a huge salience factor which, in essence, says what you are hearing now is important. There is a phonological neural loop (see the next chapter) which then causes a child to attempt production of those sounds and hearing them through the same neural machinery as hearing the sounds produced by other humans. The child is able to compare the sounds uttered by others with those uttered by themselves, and engage in making minor corrections as needed so that the two sounds match.

This salience-driven learning process organizes the image patterns being learned with the sounds being learned such that the sounds, which are effectively compact symbols, come to associate meaning with the images. The things and actions we see become coupled with names of objects and actions.

Grammar, the proper construction of sentences with the words we learn, comes about in a similar way but is learned by yet higher-order concepts probably located in the prefrontal cortex, though less is known about this presently. Grammar, or syntax, is just another pattern, but one that has special properties with respect to generating recursive structures through clauses. “Jane knows that Paul knows that she likes him.” The capacity to describe complex relations like this is needed to be able to communicate about hierarchical systems. “The company contains a department called sales whose employees must know a lot about the company.” Or: “The organism contains an organ that produces a necessary chemical that is used by other parts of the organism.” This ability permits constructing descriptions such as: “Department A sends a report to Department B, which in turn sends a response back to Department A.” In other words, the language produced by manipulable symbols using a recursively generative language allow human beings to perceive and describe (i.e. communicate about) systems having internal causal feedback loops.

Three Perspectives Processed by the Brain

In the next chapter I will go into some details about how the brain is evolved to process systemness at the neural circuit level. Recalling the mental architecture for human consciousness described above I will now relate that view with the notion of a systems construct that constitutes the contents of those four basic layers (as in figure 3.2).

The mind must consider three relations with respect to its perceptions of systems. The first perspective, relating to what I called, above, zeroth-order consciousness, is the basic self-awareness, or more correctly, processing the self as the system of interest. This involves the moment to moment construction of the body state map – the body is the system – from the state in the prior moment to the current given the current inputs from the environment and those from the body itself. In essence we can say that the brain (the most primitive parts of it) is observing the body and produces actions that are preprogrammed according to certain states. This level of awareness is far below that of conscious access except when we feel pain or pleasure.

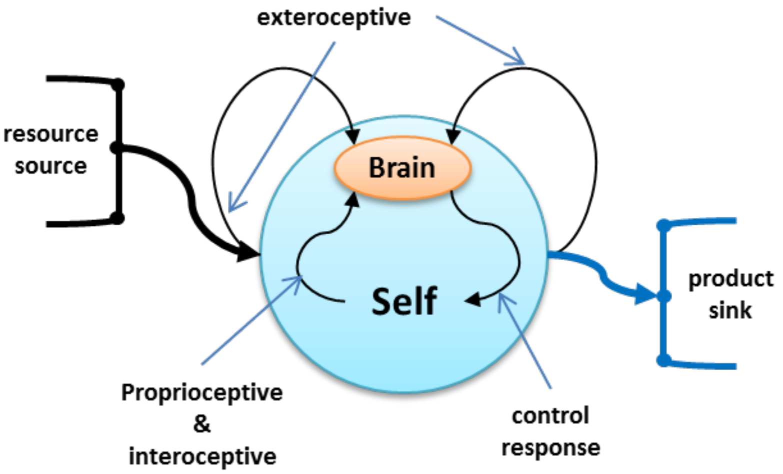


Fig. 3.10. The first perspective of a zeroth-order conscious entity is the receipt of exteroceptive inputs from environmental stimuli and proprioceptive inputs to construct a basic map of other vs. self.

The second perspective comes when the brain is capable of encoding images of correlated sensory inputs, i.e. forming images of things in the environment – what I called first-order consciousness above. In this situation the brain is capable of perceiving external entities as "other" systems. The mind now has simple models of other things that are the sources of physical and sensory inputs.

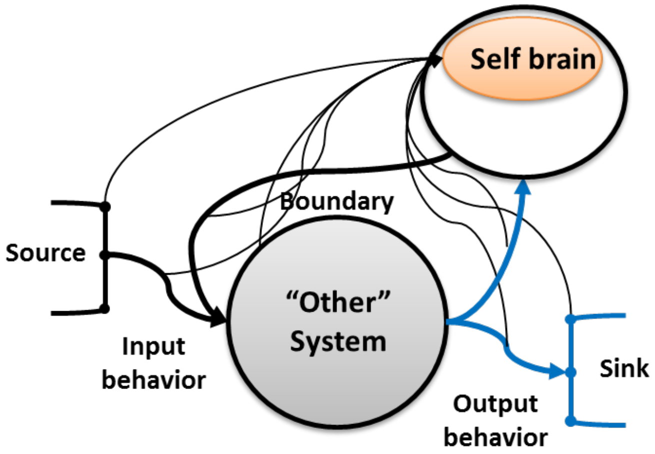


Fig. 3.11. With correlated information from the environment and an ability to detect boundaries and behaviors of "other" entities, the brain has the ability to construct a perspective of the other as a system of interest.

As brains evolved to become more capable of mapping more external entities and environmental sources and sinks (figure 3.12) the self came to a point of being able to situate mentally within a larger system.

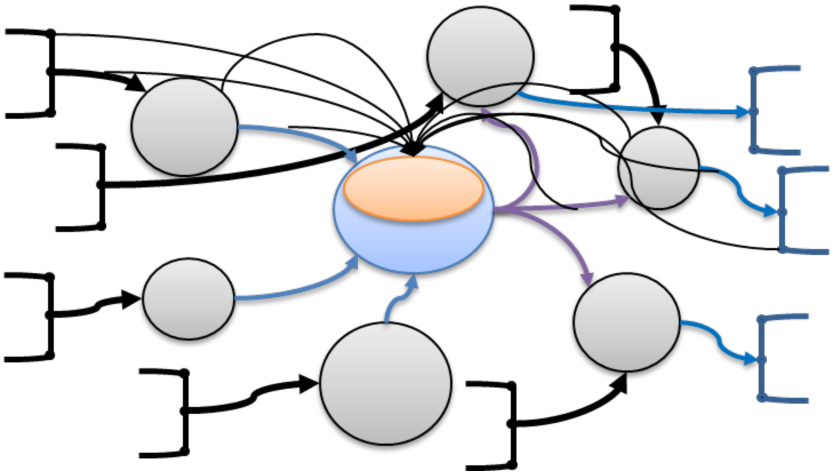


Fig. 3.12. The evolution of more complex self-others systems modeled in more complex brains such as mammals and birds gave individuals a greater capacity to model the world in which they were situated.

The third perspective is that of looking down upon a system that is external to one's self or a god's-eye-view. The mind finds causal relations between many of the modeled entities in the second perspective, not just with respect to the self but independent of the self. It is a view that provides objectification of larger complex systems[155]. For example we readily see and

---

[155] Tomasello (2014), page 113.

comprehend various kinds of organizations of human participants along with their artefacts, such as a corporation, or a church. Whether we are engaged with such an entity or not we can construct a mental model of its behaviors, structures, and components. More importantly, this perspective allows one to view one's own meta-system as if from outside (figure 3.13). The individual can do this because they have the capacity to construct a model of themselves within their own brains. Tomasello (2014) has developed a model of human perspective in which he explains that having evolved to have collective intentionality leads to group conventional thinking, norms and shared beliefs. Such thinking, in turn is necessary in order to support the notion of objectivity or a separation of the self from the system in order to observe without, for example, emotional involvement. Such an "objective" perspective does not eliminate or even necessarily minimize personal biases from emotions or ideological beliefs. It merely allows one to observe the system *as if* being outside it. This too is an important aspect of sapience. Principle 10, in chapter 1, claimed that sufficiently complex systems can contain models of themselves. That is a CAES can self-reflect and do what-if simulations about how they would react to hypothetical situations (the set of inputs they as a system might experience and the set of outputs they would produce). Individual human beings have this capacity but so do organizations comprise of human beings. Enterprises often engage in strategic planning exercises.

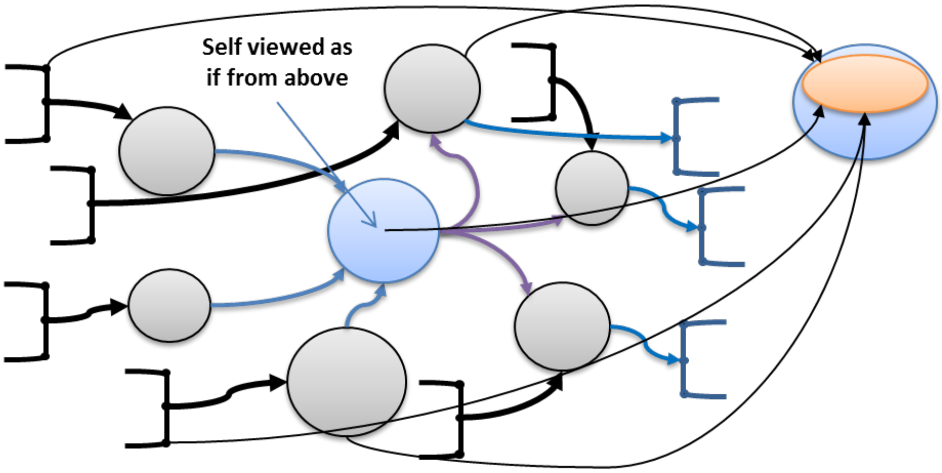


Fig. 3.13. Second and one-half order consciousness gave an entity the ability to perceive the meta-system as if from outside. An individual CAES can pretend to be observing the whole system within which it is, in reality, embedded with itself as one of the component subsystems. To do this it must contain a sufficiently realistic model of itself.

The success of this perspective depends critically on the individual having a highly veridical self-model. This, unfortunately, is where sapience appears to be weak. Most people, it seems, have very dim or inaccurate self-models that invariably produce erroneous predictions of the future[156]. But then this is not surprising. Most people have very dim or inaccurate models of other people

---

[156] An excellent treatment of this phenomenon is given in Gilbert (2006) where he describes peoples' inability to reasonably predict their future happiness (or sadness) under various hypothetical circumstances. Presumably this is because they possess weak models of themselves.

whom they observe directly. In fact, given the average human beings' inability to overcome built-in heuristics and biases[157] to give reasonably accurate predictions of future conditions it is understandable that our so-called "objective" view of the world is not nearly as objective as we would like to think. The reason is that our affect subsystem has a much higher influence on our thinking than we would like to admit. Sapience provides us with the "beginnings" of a systems perspective but until or unless the sapience regions of our minds (and brains) have a higher level of control over our limbic system we will suffer from what is almost an illusion of objectivity while working primarily from non-veridical beliefs, i.e. models of the world and ourselves.

Analogical Reasoning

"Why is a raven like a writing desk?" So asked the Mad Hatter of Alice. Lewis Carroll's intent seemed to be to demonstrate that there were such significant differences between some kinds of objects that there weren't logical answers that could be made. But in fact there are many reasonable answers that come to mind precisely because our brains are wired to explore them and we have the ability to abstract objects to ideas. Both objects are objects! Both are kinds of systems, where that latter term is used to label objects that have certain organizing characteristics.

Everything in the Universe is a system, part of a system, and composed of sub-systems[158]. This fact makes it possible to describe a small number of principles that encapsulate the qualities of being a system[159]. And it is thus that analogies between diverse systems can be made. All systems share a basic set of analogic qualities[160]. And then, on top of this, many systems share specific, higher-order qualities that make them particularly comparable. Horses and donkeys and zebras can all be recognized as being related by virtue of their body plans being so similar.

This is a powerful thing. The brain, by building good models of some things, can quickly assess other things and build analogic models that allow reasoning about the second object that stands a chance of being valid. It isn't necessary to observe every aspect of the new thing as long as its behavior (or form) strongly resembles something already known. And, wonderfully, it works for just about anything. It works for nouns (ravens and writing desks) and for verbs (flying high and being happy).

Analogic thinking probably arose early in vertebrate, certainly the mammalian and avian, evolution. Complex cortical structures are involved in what could be described as making copies

---

[157] See Gilovich, Griffin, & Kahneman (2002) and Kahneman (2011) for examples of how poor we humans are at reasoning rationally!

[158] Mobus & Kalton (2014), chapters 1 & 3 explain this using the term "systemness" to designate the qualities of being a system that take every object/process into account. There are, of course more and less complex systems. There are active and passive systems, etc. But everything that our senses can detect fits into the qualification of systemness.

[159] ibid

[160] See: http://en.wikipedia.org/wiki/Analogy for more on this powerful form of cognition.

of concepts and then “playing” with those copies to see other possibilities. More of this idea will be covered in the next chapter. But the point is that the brain has the capacity to compare two different concepts judging similarities and differences and using those for further reasoning. This is possible because all things are systems and systemness evaluation is built into the brain.

Seed Knowledge — the Systems Scaffold

Systemness is a form of seed knowledge that is genetically coded into brain structures for processing. There is a prototype model that every brain possesses that can be used to construct particular models of the things and processes we encounter[161]. Additionally, the way the brain is organized for learning constitutes a scaffold by which information is organized and incorporated into knowledge - by the construction of dynamic mental models of the things in our experience and how they work.

Our models of how the world, other people, and even ourselves (who we think we are) work are based on having a built-in intuition about how systems work in general. Indeed, our entire knowledge base is organized around systemness. And when we learn, we are incorporating our perceptions into a framework of systemness because that is how our brains are wired.

Our minds naturally look for things like boundaries, wholeness (Gestalt), cause-effect relations, and a myriad of characteristics of systemness. We automatically attempt to find patterns in noisy data, and categorize patterns in hierarchical structures. Our brains process incoming perceptions so as to see the systemic nature of nature. We can't help it. A system mentalese is the language of thought.

This is not surprising since through science, which is supposed to be objective, we have discovered that the world, the universe, is indeed comprised of systems and systems of systems. We find causal relations among system components everywhere we look. In fact, the drive behind the scientific approach to knowing is that when we find phenomena that are not previously categorized, for which a pattern of organization and causal relations have not been identified, then we are essentially forced to look for these things. It is as if evolution predisposes us to see systemness because everywhere there are systems. We are systems. And we are subsystems of larger meta-systems.

This propensity to see systemness, or discover it if we don't immediately see it, is a fundamental organizing principle which our brains are constrained to use to learn about the world. The generic system is a kind of seed structure upon which we map percepts in order to have a means of organizing our knowledge.

---

[161] Indeed a single neuron is a model system. It receives material, energy, and message inputs from other systems in its environment and produces outputs, especially messages (products), electro-chemical and chemical, that other systems use.

Every knowledge construction requires some kind of template upon which to organize new knowledge. The mind is not a blank slate (Pinker, 2002). The brain itself is organized in such a way that we begin our construction of knowledge with the aid of built-in biases for key perceptions and organization of those into early conceptual structures, like categorization and hierarchies of types. Thus as we grow and develop our models of the world and ourselves, we start with a foundation of generic systemness and a scaffolding that provides a basic shape to how we understand the world. Literally, we can't see it any other way. To that structure we start fitting our experiences into place. It is probably more a matter of jostling the bits and pieces around until they 'fit' into the scaffolding and among other bits and pieces already integrated. It is a stochastic process. Some bits won't fit anywhere in the edifice and so get dropped even if they should legitimately be part of the knowledge base. Fortunately, these bits are likely to be encountered later again so they have more than one opportunity to get incorporated.

The point is that knowledge is built upon prior existing knowledge and the ultimate seed knowledge is provided by evolution in the form of the ability to model systems.

In the next chapter I will provide some ideas about how the brain actually accomplishes this feat. For now all you need recognize is that the generic system can be represented as a network or, in mathematics, a flow graph. Figure 3.1, above, is such a network and it represents the system I have been describing in words. The dashed line circumscribing sapience demarcates the system of interest and the other entities provide inputs and take outputs from that system. Figure 3.14 shows a generic system with the expected kinds of components[162]. The system of interest has a boundary of some kind, it has component subsystems between which flows and associations occur in an internal network (not shown). It receives inputs of energy, material, and messages from environmental sources and it produces outputs of similar kinds that flow to environmental sinks. The arrows from and to environmental entities may also be reciprocal linkages with entities rather than explicit flows. This representation is kept simple for demonstration purposes.

---

[162] For a more detailed description of model components see Mobus & Kalton (2014), chapter 12, *Systems Analysis*. For more on the construction of models using those components also see Mobus & Kalton (2014), chapter 13, *Systems Modeling*.

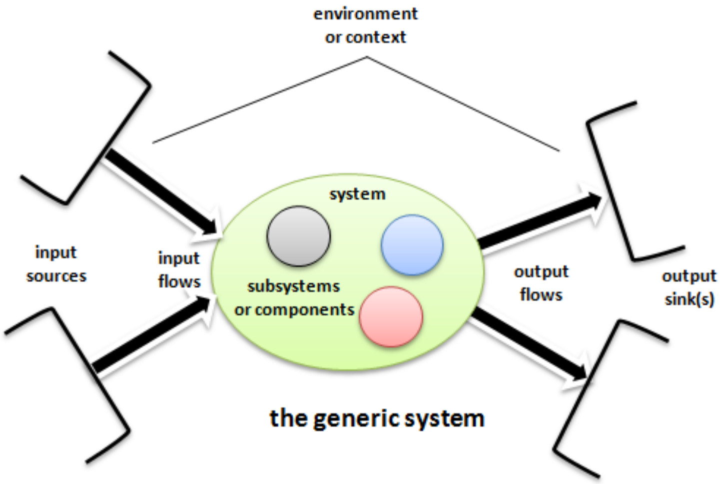


Figure 3.14. A generic system has all of the features/attributes of a basic system in generalized form. Neural networks can encode the various elements and their generic interactions. The human brain has the ability to make copies of this generic model and then learn the particular features of each kind of component.

A generic system concept is encoded into the brain as a template for the learning of all real systems/objects that the brain will encounter in the future. 'Systems learning' entails making a copy of the generic template somewhere in the cortex (probably in the frontal-parietal areas) and then beginning to link up specific perceptual and other conceptual features to the copy as it becomes particularized to the real system being learned. In chapter 4 I will revisit this in terms of plausible neural circuits. The point here is that our brains are wired to look for subsystems and boundaries and connections, etc. as we construct a larger network of particulars. Figure 3.15 is meant to capture some of this idea graphically. Starting with a fixed template copy, the brain learns the particulars of a system by identifying the features and attributes that should be attached to the model of the real system while also expanding and modifying some of the details. In many ways this template-attachment approach resembles the notion of "code reuse" in object-oriented programming of computer languages. Detailed objects are constructed by adding links of specific code to existing "abstract" objects already defined[163]. For example, the real system being modeled will have many more component subsystems with particular linkages back to the template model. Characteristics, such as the nature of the boundary, may be modified as well.

---

[163] When I talk about making a copy of a model, it should be understood that I am not really talking about a completely new copy being constructed in neural tissue, like making a new completely separate copy of a document on a copy machine. There would not be enough room in the brain for completely new copies of existing circuits. The copying is logical in the sense that new circuits are encoded with particulars and linked to existing model circuits with connections that are excited only when the "new" model is appropriate. See chapter 4 for more details on this neural method of encoding new concepts (models) as extensions of existing concepts.

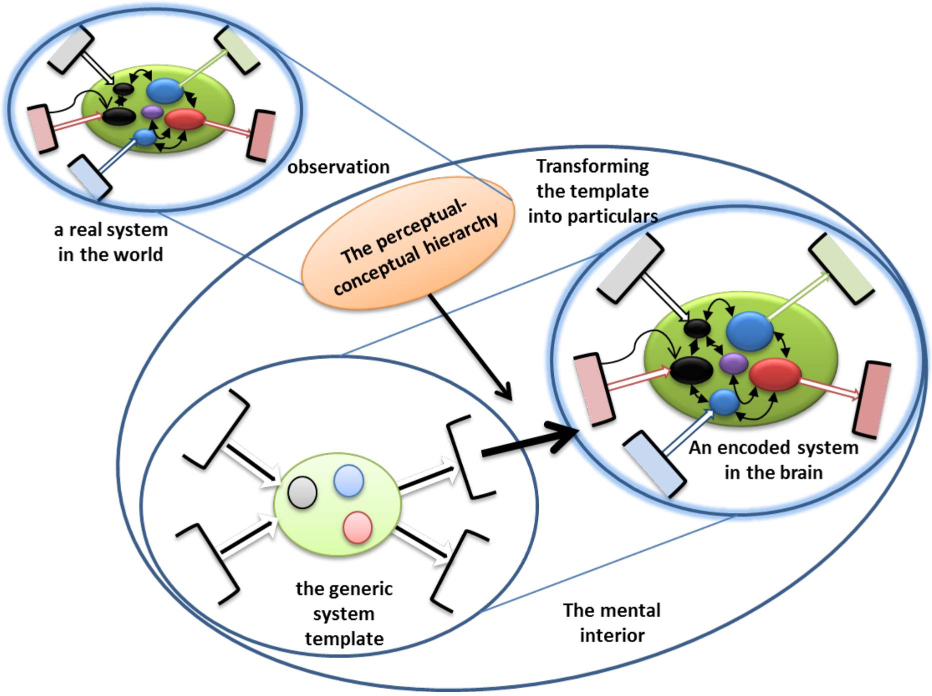


Figure 3.15. A particular (real) system is learned by attaching perceptual and conceptual features to a copy of the system template and expanding where needed, e.g. in the number of subsystems and their linkages. This is the basis for humans learning what is in the world and how things, and the world, work.

Since systems are subsystems of larger meta-systems, and are, themselves composed of subsystems, this copying-modifying procedure works in both the direction of the larger and the smaller. The brain can build a model of the meta-system by starting with an already built subsystem (now treated as a component) and situating it within the larger system. Note that the entities identified as sources and sinks can now be modeled in their own rights and their linkages constitute the more complete model of the meta-system.

Working from smaller systems to larger meta-systems is a synthesis/integration process. Working from a system inward to model the component subsystems as systems in their own right is analytical reduction. The brain automatically works at doing both of these. The former is driven by a need to understand the context of a particular system and leads to a grasp of a larger world. The latter is driven by the need to understand how a particular system works. Both of these processes are aimed at providing the brain with a basis for anticipating the future behavior of the systems it observes (see below).

Sapient Systems Perspective

In the section above regarding three systems perspectives processed by the human brain I tried to make clear that 2½-order consciousness, which is produced by sapience, is a more complete ability to see systems in the world from a more “transcendent” point of view. This included seeing one’s self embedded in the system and interacting with the other component subsystems. Much of this brain processing is done in the subconscious. That is, we see things from these perspectives without being aware that we are doing it or how it is done. But there is more to the systems perspective. There is also our conscious capacity to visualize and reason about systems. This is referred to in the literature as "systems thinking"[164].

Systems Thinking

Seeing objects as systems and connections between the behaviors of those objects as part of a larger system is commonly called *systems thinking*. This style of thinking is in contrast with simple localized and linear causal thinking (figure 3.16 A). The latter can be visualized as billiard balls striking, one after the other. Object 1 does something that causes object 2 to do something, which causes object 3 to do something, and so on. Very primitive brains such as amphibians and reptiles probably have an ability to represent this kind of thinking. More advanced brains have the ability to learn such representations.

Linear causal thinking only requires an ability to recognize distinct objects and their behaviors. It requires being able to hold a limited number of representations of the interactions between causing and responding objects. It is the simplest kind of thinking about the world and what happens in it. It is also the easiest kind of thinking to do and thus most often relied upon even by human beings[165]. While it might be called a very primitive form of systems thinking – that is it represents recognition of causal interactions between component subsystems it fails to qualify as adequate systems thinking because it does not really see the objects doing the interactions as part of a larger system.

In more evolved brains the ability to recognize more complex interactions between components produces the first level of systems thinking, but a very shallow form. Figure 3.16 B shows a depiction of the necessary conditions for claiming a shallow form of systems thinking. Here there are a larger number of components that can interact in various ways with one another in a network of relations. Interactions are probabilistic rather than simple deterministic (the probabilities indicated in the figure suggest this). And interaction strengths may vary depending on many conditions. Being able to think about such a set of interactions and relations is what most people can do when they take some time and effort to grasp complex behaviors.

They may or may not realize that these interactions are persistent over time, which would lead to a grasp of some kind of binding and boundary. Rather most people seem to rely on intuitions about those matters. Many people have the ability to grasp systems of this kind as long as they are limited in number of components and interactions. But they very often do not know much about more distal causes (the question marks in the figure represent a lack of knowledge about these distal causes).

---

[164] See Meadows (2008) for one view of what systems thinking means.

[165] This is another aspect of what Kahneman (2011) calls "fast thinking," calling upon those reptilian hard-coded causal relations that evolution found useful.

Shallow systems thinking is shallow not just in terms of scaling outward but also downward into the details of each of the components. One who thinks in this fashion can often do limited prediction (or anticipation) of system behaviors but they do not usually understand WHY these behaviors obtain. That requires understanding each subsystem, itself, as a system. In other words, the thinker needs to have insights into how each component accomplishes its transformations of inputs into outputs. Typically the shallow systems thinker treats the components as black boxes. One does not need to know how a computer works in order to use it to do word processing, for example.

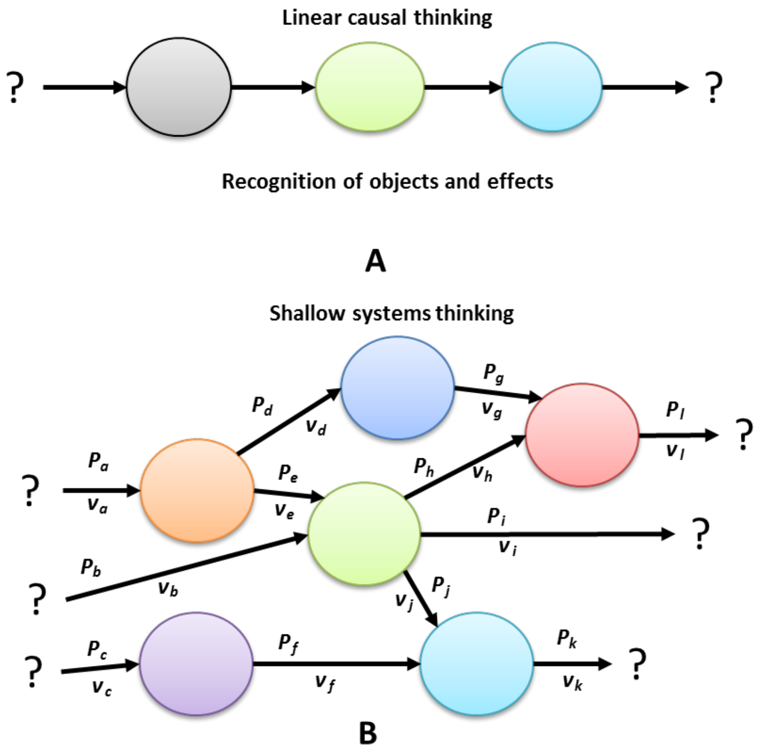


Fig. 3.16. (**A**) Linear causal thinking is simply the ability to associate the actions of one object on other objects in the vicinity. Objects are recognized by surface characteristics only (different colors). (**B**) Shallow systems thinking recognizes various objects have varying effects on other objects. The P_x refer to different probabilities, or more generally, likelihoods. The v_x refer to variable "forces" or actions. Shallow systems thinking recognizes more complex interactions between multiple component subsystems but may not involve deeper understanding of why these effects are the case. The question marks indicate that the thinker is unknowledgeable about the sources or external effects.

A transition that may happen as people learn more explicitly about systemness is what I will call "insightful systems thinking" (figure 3.17). This level of systems thinking is accompanied by a greater understanding of the whole system by virtue of grasping the boundary and boundary conditions as well as some preliminary understanding of the inner workings of at least some of the subsystems.

I think it is fair to say that the average human being is capable of reaching this level of systems thinking, at least in limited areas of expertise. For example an auto mechanic can not only drive a vehicle but can also repair one because s/he knows how the internal subsystems work. S/he might not have a similar understanding of cell biology even though both domains involve the hierarchical networked organization of subsystems. Living systems, however, behave very differently from mechanical systems and this disjunction most often obscures the deeper systemness perspective.

In figure 3.17 I have left the question marks at the boundary representing the idea that even when people develop a certain domain expertise they are still often incapable of extending their knowledge outward to encompass phenomena (other systems) that can still have an impact on the behavior of the system of interest. The ability to expand one's awareness of the elements of the environment of the system and their impacts on the system brings us to the next level of systems thinking.

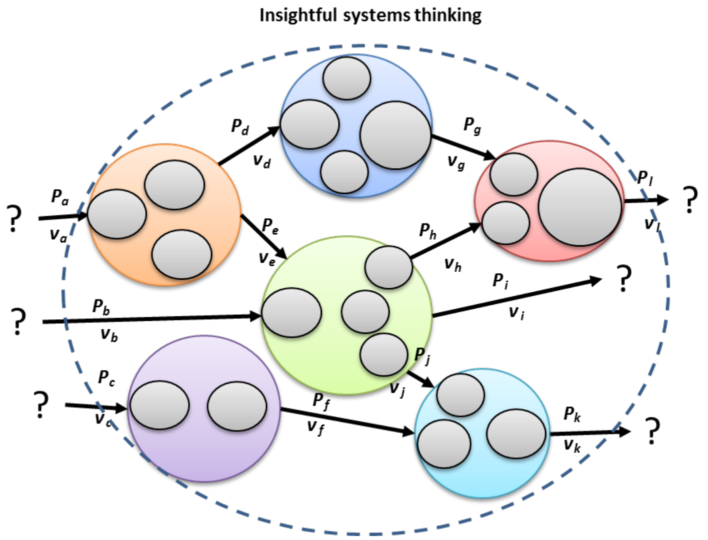


Fig. 3.17. More insightful systems thinking includes having more detailed knowledge of how each subsystem in the system processes its inputs to produce outputs. The internal ovals represent sub-subsystems. Internal arrows have been omitted, but the deeper knowledge would include the same kind of network structures inside each subsystem as is true for the whole system. The thinker also is capable of recognizing a boundary that makes this aggregate of subsystems into a larger system.

Deep systems thinking involves all of the principles of systems science from chapter 1. At a very minimum, however, it involves being mindful not only of the nature of the subsystems within the system of interest, and the interactions between them, but outwardly, it requires an awareness of the external environment and how it interacts with the system. Figure 3.18 depicts this kind of awareness. The actual sources and sinks in the environment may not necessarily be modeled within one's understanding of the system but they are identified and, if necessary they can be modelled in order to better understand their long-term influences on the system. This capability to see the larger picture follows from principle 1 – systemness.

The scope of understanding extends outwardly to encompass as much of the environment as is needed in order to have a thorough grasp of the system of interest. The scope of understanding also extend to deeper levels of the hierarchy of subsystems and sub-subsystems as needed to grasp how the system processes inputs to produce outputs. It takes into account the ambiguities and non-deterministic dynamics of the systems and subsystems as well.

This type of systems thinking goes on both at the subconscious and conscious levels of cognition. The more sapient brain is able to process models of the world in the subconscious as described in chapter 1 (figure 1.5). It can bring these models, or parts of them, to conscious awareness on an as-needed basis.

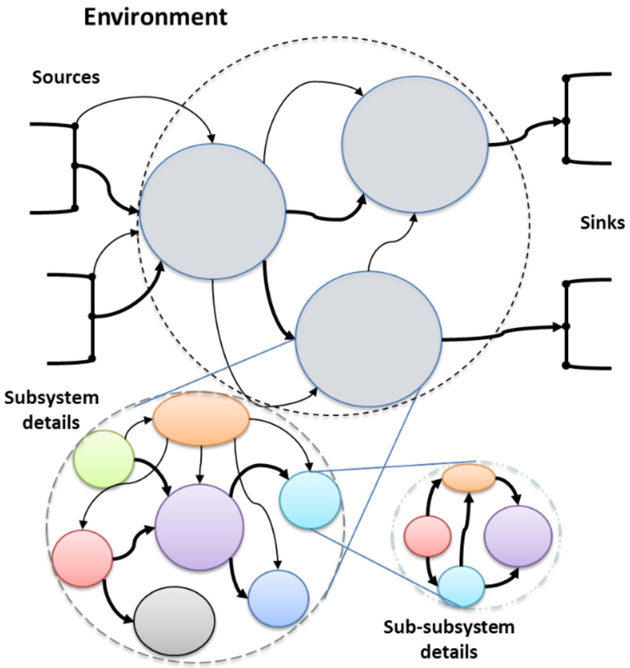


Fig. 3.18. Truly deep systems thinking means that one is able to recognize the system of interest in its environment, have an understanding of the sources, sinks, and flows. It means having deeper knowledge of subsystems and their internal sub-subsystems down to a detailed enough level to understand why transformation are what they are. All variables (probabilities, etc.) are known, but not shown in the figure.

Interactions between Systems Perspective and Other Components

As indicated above, one of the characteristics of judgment is in guiding what should be learned. We can now see that the systems bias is part of the basis for this. As our internal models of world systems improve over time and experience, our judgment derived from them can better guide the intelligence machinery in attending to perceptions that help improve the systems models. This is low-level judgment at work, the kind our biological ancestors had evolved. What makes for sapient systems thinking, and judgment so informed, is the role of strategic (long-term planning, see below) thinking, conscious reflection on knowledge being constructed and editing knowledge as needed (including editing plans for acquiring knowledge in the future). Such judgments provide guidance as to which systems need to be learned.

One succinct way of looking at this is that sapience expands the role of judgment in guiding future learning and refines the systemic nature of what is attended to in that future time. As noted above, the drives that produce the learning of particular systems causes us to explore both inward (reductionist analysis) and outward (synthesis and integration). The more sapient mind is equally interested in both directions. But all too often most humans run into limitations on what they are able to do in terms of expanding their models and understandings both inward and outward. This is a scope issue relating to the same problem as mentioned above for judgment. Most humans have limited curiosity and that appears to decline with age for many. They are not driven to explore past a certain point based on the degree of novelty involved, reached about middle age, I suspect. The old saying, "you can't teach an old dog new tricks," has some basis in psychology. As children, while the brain is still in rapid development, curiosity directed at learning the smallest details and the largest relationships is at a maximum. It is hard to say for certain when in a person's life the drive to curiosity starts to diminish. It is hard to say why it does. One can imagine a storage limit, but as I have argued, this seems less likely given the way the brain encodes systems by reusing features that are common to many systems and simply organizing appropriate linkages (see chapter 4). As an aside, I do posit that our modern education system may have a great deal to do with damping down children's enthusiasm as it attempts to force-feed knowledge, which is generally not systemically organized (think subject silos), into the minds of young people. By the time they graduate from high school (if they graduate) they have been told, in so many words, that the world contains many different disparate bodies of knowledge and they must choose one such body to learn well so that they can do a good job in the marketplace. It is hard to imagine how this message can promote curiosity and a love for learning.

But I also suspect that a continuing life-long drive to curiosity depends on the level of sapience in the individual, that is, their innate capacity for building wisdom. With lower sapience comes a limited scope and time scale for thinking. People learn just what they need to know to get by in the world to which they are accustomed. They do not, in general, expect that world to change very much. They expect whatever trends exist to continue on into the future. So at some point they are no longer concerned with expanding their scope (learning the yet larger meta-system in which they are embedded) and they feel competent knowing 'enough' about the daily systems they deal with that they do not need to know how they work inside. Lower sapience goes along with a limited world view. This is not surprising given the conditions of life for early humans where changes in the environment were generally slow in advancing, barely perceptible, and seemingly inconsequential over their lifetimes. They had no need to continue to be greatly curious as a rule as long as they stayed in one locale. However, some variations in the mental drives to explore, in some individuals, contributed to increasing the curiosity of some, and drove, for example, exploratory migrations out of Africa. Such inquisitiveness contributed to higher sapience and greater fitness for the species. See chapter 5 for more on this subject.

Sapience involves intentional model building such that one becomes more effective in problem solving in an ever wider scope as experience grows. One attribute of a wise person is grasping the interconnections between elements of a complex system, especially a social organization. Applying systems thinking to such organizations increases the probability of finding solutions that will work. And wise people seem to continue learning their whole lives.

Insofar as systems perspective interaction with moral sentiment the basic idea is that having a capacity to construct more veridical models of the world with larger scopes of time and space provides one with a much better ability to grasp consequences of actions and a capacity to see that those consequences might have negative impacts on one's self and those that one loves. Consider the concept of karma in the Vedic tradition. Karma means action that results in feedbacks from the environment affecting the actor. From a moral point of view, good actions in the present result in good feedback in the future. Conversely wrong actions, even those that bring some kind of immediate reward, will result in bad feedback consequences in the future. With systems perspective one can grasp this feedback principle as well as the nature of causality of distal consequences. Sapience includes a strong sense of being good for the good of the group, of performing cooperative behaviors that will benefit all as opposed to only the self. Systems perspective strengthens this sense by bringing the possibility of negative consequences being fed back from selfish actions.

Strategic Perspective

Strategic Thinking and Wisdom

The systems perspective allows the human mind to contemplate the future. When one begins to understand what we could call the 'bigger picture', one recognizes that the larger meta-system in which they are embedded is an on-going process operating on many time scales. Humans can think about a more distant future than other creatures. They can think about tomorrow, next week, next year, and even the next century. They can imagine the future state of things in the world. And their imaginings may be more or less veridical based on how well they know how the world works. The better their mental models of the world, the more efficacious their suppositions about what may transpire become. That knowledge depends of knowledge gained from historical experience and close observation of current conditions.

The trick I introduced in the last section, regarding the copying and modifying of neural representations of systems, allows another wondrous capability[166]. Since our model of the world and how it works is our highest meta-model it seems to have been an easy though incredibly significant step in human evolution to apply the copy-modify and run-the-model trick to that model, or subsystems within it. To imagine the future it is only necessary to copy our model of the world into another neural circuit and then, under the auspices of our prefrontal cortex, adjust the model in certain ways to essentially play, 'What If' games. We can, under conscious or

---

[166] A reminder to read footnote 20 above!

subconscious sapient control, try new things. We can construct hypotheses in mental space and then let the internal dynamics of the model play out in fast time to see what that hypothetical world would be like under those ‘assumptions’

We can, in fact, generate any number of alternative scenarios. Sometimes we do this in a serious vein meaning to try to anticipate what would happen in the real world if we took such-and-such actions right now. At other times we do it for entertainment or affective reward, creating fantasy worlds that we know (if we are sapient enough) are just play things for our amusement. Either way we can generate a seemingly endless number of futures by considering a variation in behavior of ourselves or others, constrained by real world rules (like the laws of nature), and then see where the dynamics lead. Not only can we set up initial conditions and run the model (somewhat analogous to running a computer model) but we can actually interact with the model as it runs changing this or that parameter to ‘guide’ the process if we think it is going astray.

This ability is in every human brain. What makes a difference in terms of its value is how we use the capability. Daydreaming and wishful thinking are necessary refreshments, I imagine, so should not be excluded. But failure to use the facility for thinking seriously about the future is too often the case with lower sapience. Sapience seeks to model the world sufficiently accurately to be useful as a tool in projecting possible futures. It also seeks to manipulate the model for the purpose of finding actions (decisions) that can produce good future outcomes. The world the sapient mind models must be at the largest possible scale and over the longest possible time frame in order to include all that the sapient being cares about (values). The higher the sapience the larger that model.

This is, of course, what we mean by strategic thinking and planning. Our models allow us to play 'what-if' games with the future. And if our future projections include a favorable outcome given we take a particular action now, then it makes sense to do so. All that is necessary is that we have knowledge of causal relations that provide leverage over the way the future unwinds. Sapience involves such a power.

Wise individuals have always concerned themselves with long-term and wide-scale outcomes when making judgments about actions to be taken in the present. Indeed this actually applies to follow-on actions in the future that will need to be taken in order to keep the world moving in the desired direction. Alternatively, the wise also contend with the issue of worlds that are not changeable by any action that can be taken. Sometimes the future is inevitable because the forces in action are overwhelming to human actions. A truly wise person will have a sufficiently efficacious model of the world that they can tell when action is appropriate and when it is not.

Strategic thinking, like moral sentiments, can have a dark side when sapience is unbalanced among its components. Thinking about the future and what might happen if such-and-such is done now is not always done with good intentions. A more Machiavellian person might set plans to establish or keep power over others. And those plans might include coercion and threats of violence as the actions to be taken. So by itself, strategic thinking is not sufficient to constitute a more sapient mind.

We can see now why strong positive moral sentiments are a necessary element in sapience. We don't tend to think of those who plot against others for personal gain of wealth or power as particularly wise people. Indeed one can argue that such plotting really isn't strategic thinking because in the long run evil purposes will result in failure (well we like to believe that good always triumphs in the end). However it is hard to assess a time horizon that matters in human affairs. Is a year a strategic time horizon? Is ten years? What about several generations? The answer is probably all of the above. The major issue in strategic thinking is that it is oriented toward the future, that it involves multiple possible outcomes each with associated likelihoods, and that it focuses on actions to take in the present to improve the odds of a favorable outcome in that future time. But to be sapient the favorableness has to be defined for more than the single individual. The wise elders look to the future for their people, not just themselves. They can see beyond their own deaths to a time when their grandchildren will face the environment on their own. What should we do today to assure those grandchildren have favorable options?

What Are the Functions of Strategic Thinking?

It turns out that many aspects of strategic thinking have been formalized much as aspects of judgment have been formalized under the rubric of decision science. Strategic planning and management have been developed into a model that has been used in the commercial and non-profit organizational world for many years now. Much of the formal model was worked out ages ago for military management. These formal aspects, I assert, reflect what goes on in each of our heads now to some degree or another. In our formal approaches we have simply succeeded in codifying what goes on in strategic planning/management activities in our own thinking and built systems within organizations to carry these functions out.

Strategic control is at the epitome of a system designed to keep an active agent (a person, a tribe, a company, or an army) effectively operating in the context of a dynamic, non-stationary environment. Psychologist David Geary (2005) has written an extremely lucid book on the evolution of the mind, integrating neurobiological aspects with behavioral (I will be returning to this reference in the next chapter). In it he expands on the theory that the human brain evolved as it did to increase the animal's control over its environment. Indeed, he argues that much of human social interaction involves subtle innate strategic interactions between members of the society as they collectively cooperate to tame the biophysical environment, and compete with one another for social influence without it becoming a violent process. The human brain represents the growth in importance of strategic thinking on top of logistic, tactical, and operational thinking.

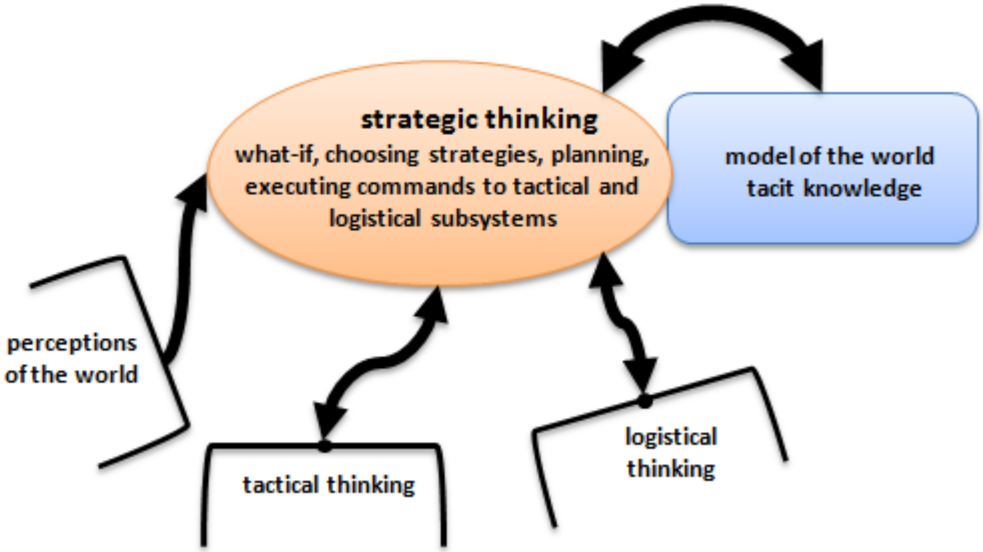


Figure 3.16. Strategic thinking involves monitoring the external world as well as the activities of tactical and logistical thinking/planning/control subsystems. It uses and adapts the world model (the world system from above) to play what-if and generate scenarios. It plans and chooses tactical and logistical commands that will be carried out by those systems. And it monitors the unfolding of events to see if the long-term desired outcomes (achieving goals) are being reached. Not shown, intelligence + creativity (cleverness) is also directed in terms of refining the model as experience is gained.

Planning

There are some well-known pieces of strategic planning that you can map to your own approach to your future plans. For example you have to think about your overarching mission in life! Most of us don't have a conscious mission (some zealots think they do) but we all have our biological mandates to survive and procreate. Some people are driven by a need to accumulate material wealth, so they see their mission in terms of making a lot of money. Others just want to muddle through without too much hassle. So many people don't actually make mission-oriented plans even though they are subconsciously driven toward certain end goals. They behave as if they have a mission.

Each of us is constantly rolling out various scenarios about our social lives. We consider questions like: “What do I need to do to win Mary Jane's (or Billy Bob's) heart?” That is usually high on the priority list when we are young and/or single. Once the mission is in place you then have to consider a number of subsequent questions. Who are my allies? Who are my competitors? What talents and charms do I have? What flaws might I have? What will my competitors do if I exercise this charm, how will they counter? What might my competitor do to exploit my flaws (most of us have difficulty thinking about our own flaws, which is probably a good thing since if we did we might never try to get a date)? What goals should I set to take actions on? Coming up with answers to those questions is strategic planning. And for most people a lot of this goes on subconsciously with actions also being chosen subconsciously.

However sapience permits us conscious access to these plans. Indeed it allows us to bring more critical thinking to bear on formulating and re-formulating them.

These are the kinds of questions each person asks themselves when they are in the courting mode. Similar kinds of questions suit other modes as well. Each of us is forever assessing our situation and our long-term prospects within the social networks to which we belong. Later, when we have a mate and offspring our strategic questions and approaches turn to how we might best position the family for the future. This growth of strategic thinking extends when we have grandchildren. Indeed, a few individuals demonstrate an ability to think strategically for multiple families, the wise elders of years gone by are an example.

And that is what sapient strategic thinking is about. Sapience involves the ability to include many others, including non-relatives, in the circle of others for whom strategic advice might be needed. In sum, sapient strategic thinking poses a future that is the best situation for the most people possible in accordance with good moral sentiments. It uses the veridical models of how the world works (systems thinking) to generate plans and judgments on what actions to take now and in the future to obtain that best situation.

Conclusion (So Far)

Sapience is built upon formerly evolved functions such as simple judgment, systems perspective and moral sentiments, primitive versions of which are seen in many mammals and some birds. Strategic perspective and thinking, at least in terms of thinking beyond the immediate future, seems to be a unique function in the genus *Homo*; I will return to the evolutionary perspective in the final chapter.

Wisdom is described as a complex set of psychological functions. Sternberg and others have determined that it is not just intelligence, though it is correlated with intelligence in several important ways[167]. Wisdom is viewed as a kind of meta-knowledge with a meta-intelligence and meta-creativity to process the meta-knowledge. I am suggesting that the model of sapience presented here helps to organize our understanding of the psychological underpinnings of wisdom.

Sapience, like intelligence, is a behavioral and mental construct with identifiable neurological underpinnings. Wisdom, by itself as a psychological construct, is a window into how humans make hard or “wicked-problem” decisions based on a wealth of tacit knowledge. But wisdom seems to require aging — one has to live a long time and aggregate the needed tacit knowledge. Sapience is a native capacity, like intelligence, that makes wisdom possible, just as intelligence makes rational thought possible. Whether or not wisdom obtains in an individual will depend at least in part on what kind of world the individual lives in, just as how smart someone becomes depends on how their intelligence is exercised in life. If one lives among similarly sapient and

---

[167] Sternberg (2003)

wise elders, then it is possible to see a wise person emerging from a life of learning. On the other hand, just as a potentially smart individual, trapped in a non-stimulating environment, might turn out dull in later life, so too a more sapient person might still turn out foolish if trapped in a foolish environment.

The really big question that this model raises has to do with how much sapience do ordinary people have relative to what is needed to live successfully in this complex, overcrowded world. That is, how powerful is the computational competence of the brain with respect to sapience? General fluid intelligence is characterized in terms of speed of memory acquisition and recall, working memory capacity and other psychometric measures. Collectively these are attributes that determine how intelligently a person is in problem solving and learning. In a similar fashion I expect there are measures of attributes associated with judgment, moral sentiments, systems modeling, and strategic thinking that collectively constitute sapience level. This model provides a way to generate testable hypotheses with regards to overall decision making competency with respect to complex, uncertain problem domains.

The question of competency level is an important one. In the case of intelligence the definition of the norm is a statistical property of the population. We assign the value of 100 as the intelligence quotient of the average person (the peak of a bell curve). And for the issues in life that intelligence, or cleverness, is good for addressing, this system seems to work pretty well to attribute relative intelligence levels. Since the curve is Gaussian the bulk of people are near the norm and there are jobs for everyone. But with sapience the situation may be different. If it is a newly emerged capability in *Homo*, as I suspect, then the distribution curve may have a more skewed shape. It may be that the majority of people fall in the lower end of the curve. In the last chapter I will explore this possibility more fully. Consider, for now, that such a distribution might well explain the seeming paucity of wisdom in our current societies. That we as a species are in the mess we are in because our cleverness exceeds our wisdom would be a reasonable conjecture.

Chapter 4 - The Neuroscience of Sapience

Brains and Sapience

Sapience is a relatively new brain function in the evolution of animal cognition. As we have established only humans seem to possess it. Its basis in brain structures is likely to be found in new or recently modified brain regions. The prefrontal cortex (PFC) is implicated in the paleoanthropological records. Cranial endocasts from late Pleistocene skulls indicate that there was a rapid expansion of the frontopolar region of the PFC. In particular the patch of tissue identified as Brodmann area 10 (BA10), right behind the eyebrows appears to have increased in proportions relative to the rest of the prefrontal cortex roughly one and one half to twice its size in previous species of *Homo* as recently as 150,000 to 200,000 years ago. The prefrontal cortex is known to be the seat of "executive functions," the organizing functions that orchestrate all other memory and processing functions in the brain. In order to understand why BA10 and the prefrontal cortices are the keys to sapience we first need to understand what is being orchestrated. And from a systems perspective that will most certainly start with the nature of brain complexity at different scales. We need to understand how each level of organization contributes to the ultimate cognitive capacity of human beings[168]. Then with some very low level details regarding neural representation of concepts in neuronal networks in the neocortex, the basic building components of thoughts, both conscious and subconscious, I will make a case for how the expansion of the BA10 patch (and the effect of that on other brain structures) brought into the sphere of cognition the capacity for strategic thinking. I have asserted that what the brain does is build models of what is in the world and how the world works, I will include some details about how these models are instantiated in neural tissues. With this explanation I will then show how the function of sapience controls the use of these models for a variety of purposes that affect decision processing.

Some Necessary Caution

A first caveat: While in the prior chapters I constructed a very different framework for thinking about the basis of what we call wisdom than has previously been presented in psychology, I managed to keep it constrained by actual psychological work on that subject. I may have flirted with speculation in the area of systems and strategic perspective, but I don't think it strayed too far from observation of human thinking and behavior. In what follows there is necessarily a great deal more speculation involved simply due to the nature of the underlying science — the science of the human brain. In recent years a tremendous amount of information about the frontal lobes and especially the prefrontal cortex has come to light and the pace seems to be accelerating. So, while I will attempt to stick close to what is known about brain function with respect to the

---

[168] Complexity and levels of organization are the subjects of chapters 5 and 3 respectively in the Principles book (Mobus & Kalton, 2014). As will soon become apparent the tie between these perspectives will be network theory, the subject of chapter 4 in that book.

functional components of sapience you should recognize that this, at times, could be getting toward extreme conjecture!

My second caveat is: The material presented here is assembled from book sources more than from primary literature (journal articles) and so is going to be, to a degree, dated. In other words the field is much farther advanced than represented in these sources. I have endeavored to keep track of some of the important latest work and use those findings in guiding the integration of what is in the books. Nevertheless, the rate of finding new research results these days is staggering, as is the volume of material. On the plus side, books tend to present a more integrated view of knowledge resulting from the authors being familiar with research from multiple sources. Thus book sources can tend to represent something of a consensus view, at least within sub-disciplines, of the state of knowledge at the time of writing. Understand that new research should always be considered tentative until replicated and the results confirm the earlier findings. Otherwise tentative results raise some uncertainty with respect to interpretation of those results. Thus reliance on books by highly credited authors seems a fair basis for anything that is, itself, speculative.

What We Will Cover

The intent of this chapter is to map the components of sapience, explicated by the prior chapter, to brain functions. This ranges from neural circuits, starting with how neural nets might be representing percepts, concepts, and models, and going to how specific brain regions might be organizing and processing tacit knowledge (organized as systems perspectives) for strategically controlled and morally-motivated judgments.

For those less familiar with brain anatomy or neurobiology I will provide as many Wikipedia references, for easy and quick tutorials, and general reading suggestions that can be used for greater background than can be contained in this work. I'll try to be gentle in expectations of what the general reader knows about neurobiology, but at some point it will be necessary for those with no background but interested in understanding these concepts better to dig deeper on their own. An excellent resource for those who want to become more acquainted with cognitive neuroscience is the comprehensive book by Barrs and Gage (2000). And if you are looking for a very readable treatment of the subject with wonderful graphics look at Rita Carter's *Mapping the Mind* (1999).

Neurons are the active computing elements that form strongly linked networks representing perceptual models at a low level of processing and complexity, and conceptual models at a higher level. A single neuron may participate in multiple networks at different times, which explains the massive capacity of the brain with a finite number of neurons to work with. I will provide a basic explanation of neural encoding and how these networks are formed (learning) and used in thinking (recall and reasoning).

There is a hierarchy of complexity in the brain organization that accounts for the functionalities that contribute to the brain's work in "controlling" the organism and its behavior[169]. This hierarchy corresponds with the hierarchical cybernetic model. In particular the human brain has structures that perform the functions of the highest level in the hierarchy, strategic management, which, as argued in the last chapter, is by far the capacity that distinguishes humans from all other animals.

The point of this chapter is to integrate the science of the brain (neurobiology) with the psychology of wisdom as covered in the prior chapters by applying the principles of systems science. It is my contention that this will provide some new insights into the overall phenomenon.

Brain Complexity at Multiple Scales

It must seem to many people to be the height of audaciousness for us humans to actually believe we can understand how the brain works! And yet it often comes as a surprise to many non-neuroscientists just how much we do understand about the brain's workings. Any complete understanding is still far off, but there are glimmers of hope. For one thing we have a reasonably good handle on the scope of complexity involved in brain functioning. By this I mean the way in which the brain is a complex system comprised of complex subsystems, which, in turn are comprised of complex sub-subsystems, and so on down to the molecular level. The state of neuroscience, viewed from a macro-level, shows a hierarchy of complexity from the molecular functions within synaptic junctions (and associated glial cells) through the workings of whole neurons, through circuits of neuronal assemblies, all the way up to the functioning of whole brains comprised of complexly intercommunicating modular regions.

There are still many details not yet understood, such as exact wiring diagrams for circuits and regions. We still have a long way to go to delineate the intricacies of all of the neuromodulator molecules and their effects on various kinds of neuron types. Even the zoo of neuron types is probably not complete. But nevertheless the broad outlines of some very powerful principles appear to be taking shape that allows us to produce better hypotheses leading to better experiments and yielding more useful information over time.

Brain Complexity Overview

Figure 4.1, below, shows a schematic diagram of brain organizational complexity so far as it is understood today. There are multiple different dimensions of complexity and here only four are shown. The first dimension of complexity in the figure is called "levels of organization," from the synapse through to functional neuronal clusters such as nuclei (globular-like clusters) and cortices (sheet-like clusters). These levels are dependent on those below as indicated in the concentric ovals. Synaptic complexity is itself dependent on chemical processing complexity at a

---

[169] Hierarchies of networks in the brain and how that architecture functions can be found in Seung (2013) and Sporns (2011).

lower level than shown. The second dimension could be called functional identity (modularity); various brain regions process different kinds of information even if the underlying mechanisms for doing so are the same as delineated in the first kind of complexity. The third dimension of complexity is indicated as the inter-module wiring organization or schema that produces the whole brain activity and the animal's behavior. All of this is shown as against a fourth dimension, that of genetic and developmental complexity, which is beyond the scope of this book. I will, however, touch on the subject as part of the last chapter on evolution.

Other dimensions of complexity, to some degree, map onto these dimensions rather naturally. For example the temporal dimension roughly maps onto the levels of complexity with the fastest time scale represented by events in the synaptic junctions (chemical reactions). However, it turns out that longer-term processes also impact the synapses as is the case for the development of long term memory traces (see below). Longer temporal scales are involved in neuronal activities, and longer still scales in circuit activities.

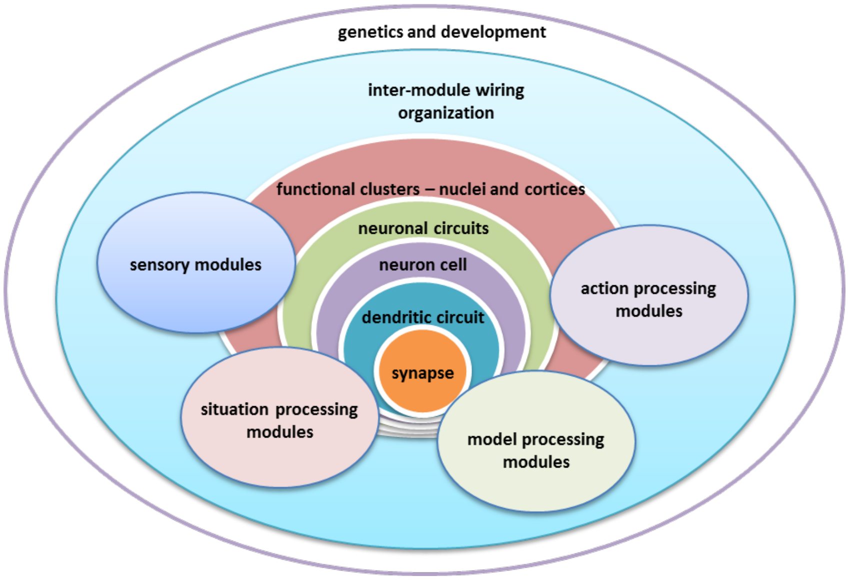


Fig. 4.1. A schematic representation of four dimensions of complexity in the functioning brain provides a sense of how deep the levels of organization go. Several important module types are shown at the levels of "functional clusters" and extending into the level of inter-module wiring that gives rise to ultimate behavioral functioning.

Another important dimension of complexity addressed in the schematic is genetics and particularly the role of evolution and development on the shaping of the phenotypic form. To some degree these can be seen in the various functional modules, their sizes and relative processing powers being due to both genes (evolution) and development influences (e.g. epigenetic factors).

There has been a spate of predictions from the artificial intelligence crowd over the years about when the computational processing power of computers would overtake the capacity of the brain[170]. Relying on the so-called Moore's Law phenomenon[171] as an indicator of processing power, and a very overly simplistic estimate of neuronal processing power (in megaflops per second — a wholly unjustified comparison), these futurists estimated that we would soon be able to build computers that exceeded the human brain's abilities. Aside from the fact that they started with almost no real understanding of just what sorts of algorithms they might need to emulate to get that kind of ability, they seriously and grossly underestimated the informational complexity of just what is going on in the brain in the first place. If one were to hazard a guess at the processing capacity of the brain, they should start by looking at the complexity of just a synapse all by itself. Neural network modelers are infamous for trying to simplify the synapse down to a simple scalar variable (the synaptic weight) and a linear learning equation that would change the weight according to simplistic interpretations of Hebb's rule[172]. In fact the complex chemical processing taking place at the synapse is extremely difficult to emulate in our very fastest computers. The best we can do is approximate some functional form of what it is doing to form and maintain memory traces. The real thing is wildly complex. Just ask any neurobiologist.

Systems Analysis and Bottom-up Integration

Ordinarily when we are faced with a very complex system and want to understand it we start with a top-down systems analysis through structural/functional decomposition. The brain and nervous system might, in theory, be analyzed through this procedure. As it turns out, in biology as in most sciences, the analysis of form and function has been carried out in a somewhat fractious top-down-like fashion through the normal reductionist procedures. That is we started by dissecting organisms, then their organs, then the tissues in the organs, then the fine structures of cell groups within the tissues, and then the cells themselves. Biology has reached the stage where “dissection” of the tiniest aspects, for example the genetic code and the machinery for converting it into living tissues, proceeds apace, bolstered from below by biochemistry and that by organic and physical chemistry[173]. Though, strictly speaking, the main organizing principle of systems, the hierarchical levels of organization, were not followed per se, one cannot help but recognize this feature of biological systems in practice. Therefore we can work as if structural and functional decomposition has been used to do a top-down analysis. To a large degree, that is.

---

[170] See Moravec (1990).

[171] Gordon Moore observed that the power of computers and memory devices seemed to double every 18 to 24 months. This is due to the nature of digital components being made smaller by technological advances. More transistors could be fit on a single chip. Owing to the step-function like investment in new processing equipment to make the chips, in each time period the power of the devices magically seemed to jump. As of this writing Moore’s (sort of) law is still in effect but the physical limit of miniaturization will be reached some day. Of course, by the time that happens many computer engineers expect a completely new medium rather than silicon will come into use, allowing Moore’s law to seem to be a true physical law of nature! See the Wikipedia article: http://en.wikipedia.org/wiki/Moore%27s_law

[172] Hebb’s rule essentially says that “neurons that fire together wire together.” See also: Rumelhart & McClelland (1989).

[173] An excellent description/review of hierarchical structure in living systems starting from the molecular level can be found in Smith & Morowitz (2016)

To accomplish the objectives of systems science we need to now synthesize our findings from a bottom-up approach of integration. The below discussions of brain complexity (and function) is based on do just this. It will start with the lowest essential level of complexity and work its way up to whole brain complexity (and function) to show how the brain, as a whole processes cognition – encodes memories, recalls memories on an as needed basis, and manipulates those memories in various and sometimes new ways to produce new thoughts.

However, I want to take just a bit of time to explain the top-down analysis results in order to prepare you for the bottom-up integration. The brain is an organ found in the heads of most animals. In worms and insects there is little more than a thickened set of ganglia (neuronal clusters with specific processing jobs) representing the fact that most of the sensory modalities had evolved in the cephalic (head) portion of the bilaterally symmetrical bodies at an early point in biological history. By the time we get to vertebrate body plans we have the beginnings of what we would call a true brain, an organ in which all of the previous ganglia have been organized into a single, though still bilaterally symmetrical, aggregate of tissues – a brain subsystem that controls most of behavior.

Brains evolved with the complexity of the animals in response to the complexity of the environments in which their behaviors provided improved fitness (see below). The first major innovation was the division of the single large globular node in an anterior direction. That was followed rather quickly by another such division to form three major nodes, the hind, middle, and fore brain modules. The hind brain remained the major interface between the more frontal lobes and the body insofar as general body state and sensing external pressures (touch) as well as sending control signals to the viscera and muscles (recalling Figure 3.2). The mid brain, at first more of a replicated hind brain providing some redundancy and backup, was free to evolve in new directions as long as the basic functions of sensing and controlling were being handled. In fish we start to see the results of this in both the mid and forebrain modules. Eventually, by the time we get to amphibians and early reptiles, the three module architecture is established and is conserved in later evolutionary transitions. However, the neat trick of accreting new tissues to the front end of the brain did not end there. The middle brain module would evolve more complex response mechanisms that we find in our modern limbic core. The forebrain would itself evolve tissues that were more advanced at learning by association, primitive cortical sheet structures that had the ability to encode and represent objects – the precursors of second order consciousness.

The modern mammalian brain, including the hominid brain, retains the basic three-module plan but the various modules have been extremely expanded and complexified. The prefrontal area of the forebrain has been the site of expansion of consciousness and executive functions such as planning and creative thinking.

Neurobiologists have long been teasing apart the numerous sub-modules in every part of the brain. Just as the main brain modules of the early animals replicated and differentiated in functions, it seems that all brain sub-modules have done effectively the same thing. Areas responsible for one kind of processing were duplicated due to some developmental control gene being mutated (one possible mechanism). Duplicates are redundant, but if not necessarily detrimental to the organism they can persist. Because they are redundant they are free to differentiate evolutionarily. If such differentiation contributes to the fitness of the possessor then the brain has become more complex and has additional functionality.

Various regions of the brain appear to have similar ways of processing information, but involve slightly different cell types (neurons) and connectivity – what is called the microarchitecture of the tissues in that structure. They also receive inputs from different other parts of the brain which tends to make them specialists in processing particular kinds of information (i.e. having particular meaning). They also send their results to different other parts of the brain to elicit modified responses, and so on. The brain has seemed to particularly adept at responding to evolutionary demands from changing environments. And its changes have, in turn, caused changes for other entities that share that environment, e.g. their predators or prey.

All of the work of reductionist analysis has given us a set of basic structures at different levels of complexity from the whole brain (and its operations in the whole animal) down to the level of neurons and their interconnections. Actually it goes deeper in dissecting the workings of synaptic junctions down to the molecular level with respect to encoding memory traces at the circuit level. In what follows below I will provide a brief tour of these mechanisms starting at the lowest level of complexity and attempt then to build up the higher levels of complexity operating at the level of the functioning whole brain.

Complexity in the Brain Reflects Complexity in the World

The concept of complexity is deeply related to the dual aspects of information and knowledge[174]. The former is conveyed in a message being received by an entity when that message contents are not completely expected by the entity. Gregory Bateson (1972) famously described information (the technical concept advanced by Claude Shannon[175]) as “news of difference that makes a difference.” By that he meant a message, the contents of which varied from that expected by the receiving entity. The entity is “informed” by something that differs from its *a priori* “knowledge” (expectation). The latter, then, is just the internal condition of the entity that forms that *a priori* expectation. In the terms that I have been developing here, that is the “model” of some aspect of the world about which the message conveys content. Put succinctly, the less one knows about aspects of the real world, by virtue of having incomplete models of phenomena, the more “surprised” one is by messages received. Conversely, the more you know the less surprised you are by messages. Thus the amount of information you receive from any kind of message depends on how much you already know.

---

[174] Chapter 9 in Mobus & Kalton (2014) covers the concepts of information and knowledge.

[175] Shannon & Weaver, 1949

Spatial Complexities

The complexity of the brain, in terms of the levels of organization shown above, is a mirror to the complexity of the real world to which that brain is attuned. And the complexity of the real world to which human brains are attuned is literally astronomical. From the precision of our senses at minute scales to the vastness of our solar system, and through the extensions of our senses that we have developed, far further down into the quantum world and far further out into the Cosmos, we have the ability to be exposed to, so far as we know, every reachable level of organization. The human brain is, in other words, a structural and functional encoding of how the world works. Or, at least it is capable of capturing a significant model of how the world works. Whether its owner actualizes that capability or not depends on how much of the world a person's brain exposes itself to.

What can a snail know? The size and complexity of the snail brain reflects the limited amount of the real world with which it must interact to survive and reproduce. That isn't much so the snail's brain is not, by comparison to the human, very complex. The snail's brain only needs to encode a model of the world in which it crawls around looking for food and mates while avoiding becoming food. Even so, the snail brain is sufficiently complex in its own right to cause neuroscientists pause[176] in saying they know its complete workings.

The brain of any animal is designed to process information about the world in which they must live. It brings in messages from arrays of sensors for light, odor (or chemicals), perhaps sound, and touch, to name a few. Those arrays are mapped into specific modules in the brains where meaning about the spatial arrangements of things in the world can be determined and appropriate actions can be taken. Thus, the spatial complexity of the world that must be attended to is reflected directly in the spatial distribution of processing tasks in whatever brain structures exist. But spatial complexity is only the beginning of the problem that brains have to solve. You not only need to know where things are in your environment, you also need to know when they are impacting you and to what degree. These are issues of temporal complexity.

Temporal Complexities

Below, in the section called "Representing Causal Relations in Neurons" I will explain how it is that even the most primitive brains can deal with the temporal complexities of the world. Here I want to introduce the reader to a concept that they may not have heard about unless they have studied "stochastic processes." Stochastic means that there is some randomness involved in determining the exact value of a process variable. For example, the temperature of a specific location can be measured, say, every hour during the day. You could easily grasp that the temperature goes up during the day, reaching a peak in the mid to late afternoon, and then falls to a low that valleys in the wee hours of the morning. But it turns out that the curve of this time

---

[176] An even simpler animal, a primitive roundworm called *Caenorhabditis elegans* or *C. elegans* for short has a substantially simpler brain than gastropods. Indeed, it has only 302 neurons in its entire neural system! See the Wikipedia article: http://en.wikipedia.org/wiki/Caenorhabditis_elegans

series of measurements is not actually a smooth curve. Were you to take the temperature more frequently, say every minute, you would see shorter term deviations from the smoothed curve. These are due to many factors that cause fluctuations in the local temperature, partly due to local spatial complexities, like terrain that causes turbulence in air flows. The resulting time series curve is now more jagged with seeming random up and down deviations randomly distributed in time. That is basically what we mean by a stochastic process. In order to work with such stochasticity we use statistical methods such as computing the time average of the temperature so as to have a definite, if maybe not totally accurate number with which to work.

Such stochastic processes are everywhere in nature. Nothing happens with the smoothness of, say, a pendulum swing. Everything is to some degree or another, jerky. It turns out that sensory circuits in the brain evolved to deal directly with this problem and our senses use time-averaged values of physical attributes (like temperature, or chemical concentrations). That is all well and good for short-term time scales.

However, over longer time scales even the statistical properties of a stochastic process can change, and in really unpredictable ways. One very important example of this, right now, is the mean global temperature (along with the daily variance of temperatures). Due to anthropogenic warming, the mean temperature is climbing such that each passing year it is just a bit warmer (on average) than prior years[177]. Of course due to the complexities of the Earth's atmosphere and hydrosphere, the process is very stochastic and so the change in mean temperature may be masked by random noise over, for example, half a decade spans.

Such a process is called "non-stationary" in stochastic process language. There are two types, basically. The first type, similar to the example of global warming, is called homogeneous non-stationary. Essentially it describes a trend. The trend itself might be somewhat predictable within limits of time horizons. That is how scientists are able to "predict" the mean global temperature in the future as a function of the released CO^2. The idea of a trend may seem to be predictable but the reality is even more complex. For one thing the system being measured might actually be subject to internal nonlinearities that could cause the measure to suddenly jump unexpectedly when it reaches critical values. That is a real surprise.

The other form of non-stationarity is called, not surprisingly, non-homogeneous non-stationary. Essentially this is nothing like a steady trend. Rather the measurements can appear stationary on short time scales but if made over much longer time scales would produce completely different statistical values. These are truly unpredictable processes. And the fact is they probably are the norm, in the sense that they are more ubiquitous than other kinds of processes.

---

[177] Talk about complexities! The global average for both atmosphere and hydrosphere are increasing "on average." That means that the annual average may only be detectable on longer time scales, such as decades rather than each year. But I think you get the general picture.

Non-homogeneous non-stationary processes represent the hardest problem to solve when it comes to learning causal relations. The latter is absolutely necessary for purposes of prediction over the long run. *C. elegans* might not need to worry much about non-stationarity in its normal environment, but *Homo sapiens* is exposed to the whole Universe. Generally speaking just about everything of any importance is essentially non-homogeneous non-stationary. And that means that no one person can have adequate knowledge encoded in their brains to be able to "understand" the world and not be surprised on occasion.

The marvel of the human brain is that it is evolved to handle all of these kinds of spatial and temporal complexities, at least within limits imposed by biological constraints. The reason is that its levels of organization reflect those of the world with the dynamics of each level tuned to those of the world. The environment of the Earth has gotten more complex over geological time. Once, four billion years ago, the Earth surface was basically rocks, water, and atmosphere, all with very different compositions than what we find today. Once life emerged (cellular entities) the biosphere was born and the chemical complexity of all other spheres began to complexify. Living systems synthesize more complex chemicals such as proteins and lipids and let them loose in the world. Over time the levels of organization, mediated primarily by life processes, increased in a hierarchical layer-like fashion, and so too did the complexity of the whole. Biological evolution was responsible for acceleration of this complexification. At each emergence[178] information flux increased and that drove the evolution of information processing power, especially in the brains of mobile entities – animals. The brains of all animals reflect this 'layered' structure. Simpler animals operate in simpler environments and have lower level layers of processing. The human brain is constructed with the maximum of layered capacity (recall figure 3.4). And those layers reflect the levels of complexity we find in the world.

The levels of organization/layers of complexity start with the chemical processes that take place at the sub-cellular level. The first clear example of a mechanism that encodes knowledge based on information received is the synaptic junction of neurons. That is where I will start.

The Synapse

Synapses are the junctions between neurons. First note that there are actually two major categories of synapses, electrical and chemical. Actually both have chemical and electrical phenomena going on so that distinction is a little confusing. Chemical synapses probably ought to be called semi-discrete or pulsed synapses since they operate on the basis of pulses of electrochemical discharges of neurotransmitters. Electrical synapses might better be called continuous or gradient synapses since they communicate information more like an analog device. The problem with either of these descriptors is that they don't quite own up to what the phenomenon actually entails. For example, in the case of the pulsed synapse, the release of neurotransmitter chemicals into the synaptic gap leads to an analog build up. If the pulses come

---

[178] As in Morowitz (2004).

in rapid succession, there is something like a gradient response in the post-synaptic membrane. Nature, it seems, could not make up her mind in terms of the discrete signal vs. analog and so chose to create two versions of a mixture between the two. It could be that discrete-like pulsed signals are more reliable over long distance communications channels (the axons of the neurons - all of the electrical synapses are between near neighbor cells). But analog (continuously fluctuating) signals are needed for reliable storage and integration of state information.

Figure 4.2 shows a highly simplified diagram of a typical chemical synapse and its multi-time scale processing. An axon or 'cable' from another neuron, or a sensory end organ, conducts a pulsed signal called an action potential. This is an electrochemical wave that travels down the axon (always away from the origin) and terminates in a synaptic bud. The latter contains packets of neurotransmitters of various kinds depending on the type of neurons involved. There are many different kinds of neurotransmitters and many more kinds of neuro-active molecules (also called neuromodulators). Some neurotransmitters stimulate the receiving neuron to fire, others inhibit it. Right here we find many sub-dimensions of complexity. The dynamics of post-synaptic membrane behaviors are highly variable and dependent first on the kinds of neurotransmitters and modulators found in the vicinity of the post-synaptic membrane.

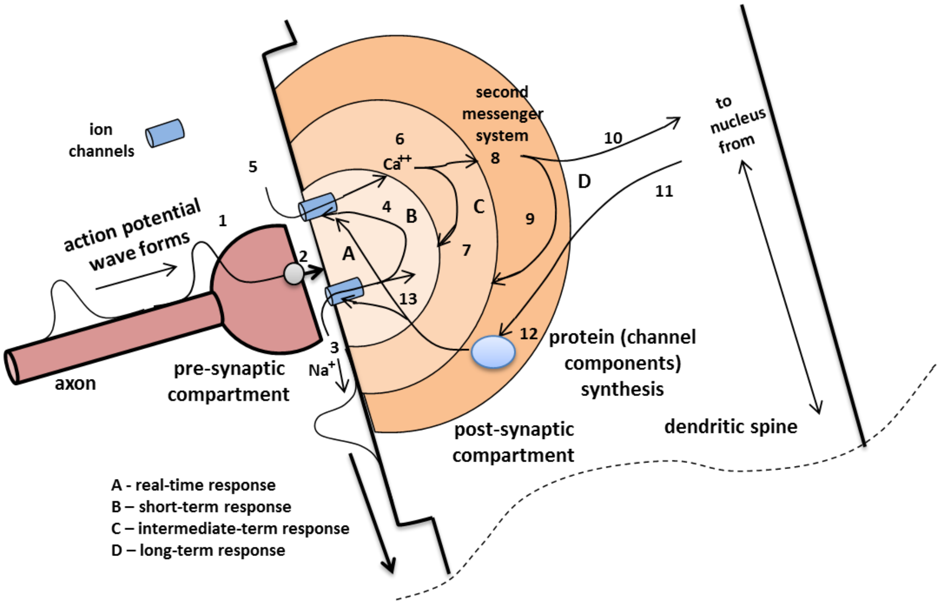


Fig. 4.2. The synapse is an incredibly complex, dynamic actor in the whole panoply of neural processing. Here chemical and electrical forces interact to produce cell to cell signaling with signal processing, such as noise filtering, and even noise introduction for some purposes. The synapse is the site of memory trace encoding in that the receptivity of the post-synaptic membrane increases under certain signal and associated conditions. See text for details.

Synapses can form just about anywhere on the receiving neuron's body, but in most instances form where an axon comes in contact with a dendritic spine (as in the figure). The events summarized here were derived from Alkon (1987), LeDoux (2002), and Squire & Kandel (2009).On the arrival of an action potential (1) the pre-synaptic boutton (bulge) releases the packet contents into the gap between pre- and post-synaptic membranes, where they rapidly diffuse across and couple with receptor sites on the post-synaptic membrane (2). These receptors are tied to protein channels that open (or close depending on which neurotransmitters) allowing the influx of sodium ions (3) into the interior of the post-synaptic compartment. This sudden influx of positive ions changes the electro-potential across the membrane which can cause various voltage-sensitive channels to open as well (4). There are actually a number of other ions that may cross the membrane in response to a signal, but the important aspect here is that the electro-potential is changed dramatically. This depolarization wave can travel down the dendritic spine toward the cell body (shown as an arrow from #3 to a wave form traveling toward the receiving neuron's main body). When I get to the neuron level of complexity I will further elucidate what happens then. For now it is only important to recognize this and certain chemical changes in the immediate post-synaptic compartment as real-time responses to the incoming action potential signal.

One major real-time effect is for calcium ions (Ca^{++}) to enter the post-synaptic compartment (5). This starts a series or cascade of chemical events that operate on a slower time scale (6) than the real-time events and depends on the accumulation of calcium ions over time. The ‘intensity’ of a pulsed neural signal is encoded in the frequency of action potential arrivals (pulses per unit of time). The more frequently that they arrive at the synapse, the more calcium ions build up (and are slow to be removed) and thus stimulate several different second messenger effects (7 and 8). One effect (7) reinforces the sensitivity of the post-synaptic compartment as a form of short-term trace of the preceding action potential stream called short-term potentiation. This simply means that the synapse becomes more likely to fire a depolarization event even with weaker (less frequent) subsequent inputs until the cell manages to sequester or remove the accumulated calcium. Effectively, the calcium buildup acts as a kind of leaky integrator or capacitor storing a short-term memory of recent events[179]. At the end of a long chain of second messenger events and in correlation with other chemical conditions brought on by either the activities of nearby synapses or by neuromodulators in the extra cellular matrix, chemical signals are sent to the protein construction mechanisms (called ribosomes - 9) and back to the cell body and to the nucleus (10). The former would appear to reinforce instructions to keep producing proteins needed to keep the channel concentration up to snuff. The latter is thought to be activating genes in the nucleus to increase the production of messenger RNA (somehow tagged to be delivered to the synaptic compartments that sent for it!) that will up the concentration of channels in the post-synaptic membrane (11, 12, and 13). These membranes undergo structural changes that are very long lasting and constitute longer-term strengthening of the memory trace. Synapses, so strengthened, are capable to generating strong depolarization events even with weak incoming signals, and even after long periods of quiescence.

---

[179] See this Wikipedia article on the leaky integrator: http://en.wikipedia.org/wiki/Leaky_integrator

Synaptic dynamics and the encoding of memory traces in synaptic strength in several different time domains is the basis for memory phenomena at the circuit level as I will discuss below. But as you can see, already we are dealing with immense complexity (chemical and temporal) and we haven't even considered what is going on in the rest of the neuron.

The Neuron

The first thing to understand is that there are many kinds of neurons in the brain (Figure 4.3). I can barely do justice to the variety and what is known of their different functions. Here I will just consider a single type of neuron, one that is common in the cortical structures discussed later. That is the pyramidal cell (see upper left corner of the figure below). This kind of neuron, of which there are sub-types depending on what cortical area one is looking at, appears to act as a major integrator of diverse convergent signals from both local and distant neurons. But its axons are long and branch considerably so that it also acts to send signals to divergent other cells.

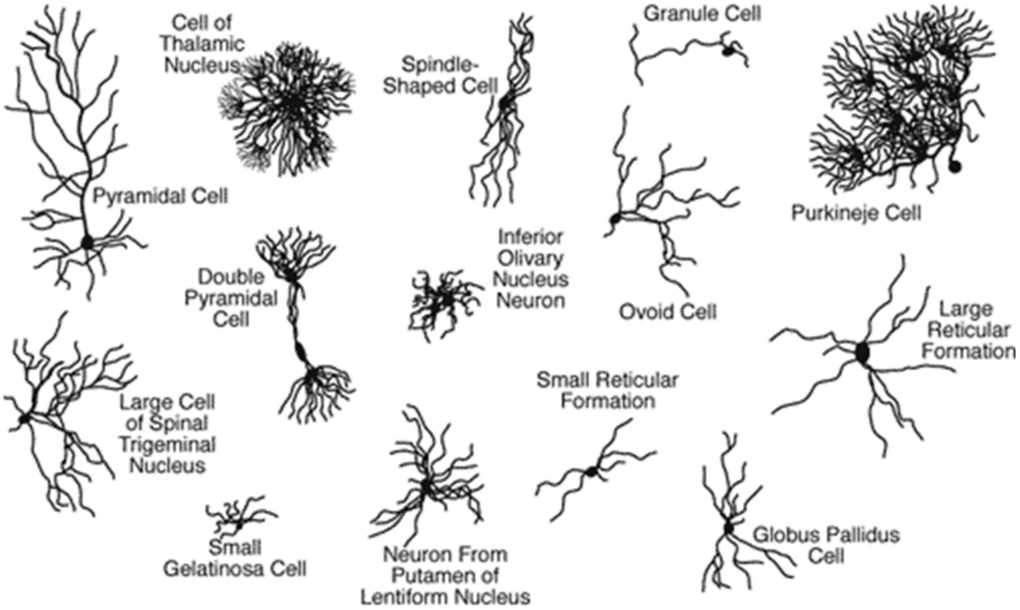


Fig. 4.3. Various neuron types found in the brain. Image from Consortium on Cognitive Science Instruction (image: http://www.mind.ilstu.edu/curriculum/neurons_intro/imgs/neuron_types.gif , accessed 2/6/2019)

Almost all of these types of neurons derive from a base-type and they all have many features in common insofar as their functioning for receiving and sending signals. Figure 4.4 shows what we might call a typical arrangement for neurons in their role as communications devices.

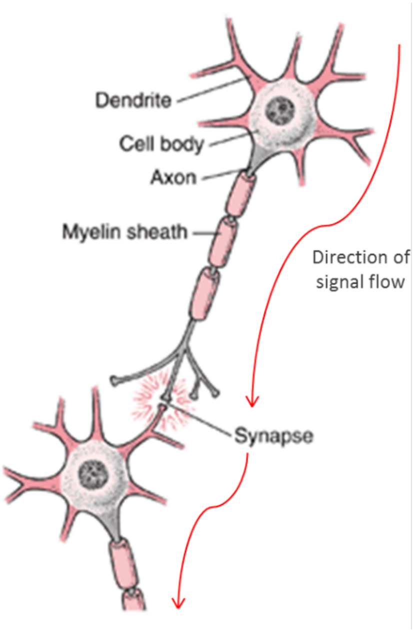


Fig. 4.4. Typical neurons receive input signals (action potentials) in the dendrites or on the cell body and send signals down the axon toward other neurons or end effectors (like muscles). A fatty sheath (myelin) insulates the axon much as a plastic coating insulates an electric wire but also helps to speed the signal in those axons that are so equipped. Not all axons have sheaths, however. Generally only long distance axons (from one brain region to another) are sheathed. The bundles of long-distance axons form what is known as the white matter in the brain. Neuron bodies form the grey matter. Image courtesy Tutorbene,. http://www.tutorbene.com/ accessed 2/6/2019.

Neurons are said to 'fire' each time an action potential is sent down the axon. Whether or not a neuron fires depends on a complex integration of all of the incoming signals received in the dendrites and cell body synaptic junctions. At the root of the axon (see slightly darkened region of the cell body above where the axon emerges) a 'threshold' function determines if the sum excitation from the inputs should, at a given instance in time, generate an output action potential. If the inputs are weak, or out of synchrony, the neuron can fail to fire or fire very sparsely (low frequency). On the other hand, if the inputs are strong and closely aligned in the temporal domain, the neuron can fire vigorously (high frequency). Thus, the neuron is a relay, a filter, and an amplifier depending on the total input activity. Basically the dendritic processes and synapses provide a temporal integration over all incoming signals. The neuron's dendritic tree structure and its body membrane provide a spatial integration over all incoming signals. This is how it is possible for even a single neuron to act as a pattern recognizer.

The dynamics of neuronal firing is based on all of the previously mentioned complexities of synaptic dynamics plus those resulting from the complexities of the cell itself. Synapses that strengthen due to high levels of activity at critical moments when other inputs have excited a particular cell (or even just a particular branch of a dendrite!) produce a memory trace through the neuron. If those synapses have a tendency to excite the cell body they can produce an output signal in the cell whenever their source cells are excited. It is the case that a single neuron may participate in many different memory traces, acting as a convergence zone for different memory features that constitute a cause for activating whatever other cells the axon runs to.

Neurons participate in networks in which each neuron acts as a little spatio-temporal integrating computer. Many neurons can innervate a single neuron, which, in turn, can activate many other neurons. The dendrites pull in signals from all over, both local (nearby) neurons and distant ones. The axons generally branch out and run to many other neurons, again, both local and distant. The possibilities are literally infinite, especially since we now know that neurons are forever forming new connections even while breaking off old unused ones[180]. The brain is continually reforming as new memory traces are built from new experiences. At the same time it reinforces memory traces that have proved to be useful in living experience. But each kind of neuron is a complex processing unit in itself. We are just beginning to unravel some of the mysteries of various sub-functions that the different types of neurons perform, their dynamics, and their interactions with one another.

Interested readers may find more information about synaptic and neuronal processing in my paper on simulating these (Mobus, 1994).

The Local Network

Neurons work together in complex networks to process even more complex spatiotemporal patterns of inputs (say from the senses), associate these patterns with context (other patterns) as well as with the internal state of the organism (drives, emotions, ideas, etc.) to produce meaningful output (e.g. behaviors). In the basic sense, meaningful output means producing behaviors that support the fitness of the beast in question. It turns out that many researchers have demonstrated how neural networks and their activities produce meaningful behavior in various animal models. While this is still preliminary there are many really convincing demonstrations of how neurons (including primitive brains), working together in specifically organized networks, produce fit behaviors.

One of the most primitive networks that evolved for the purposes of movement control is called a ‘central pattern generator’ (CPG). This is an arrangement of neurons that mutually excite and inhibit dynamically in such a way as to produce an undulating (not-quite sinusoidal, but cyclical) wave when innervating opposed muscle groups[181]. I, along with my colleague Paul Fisher, while we were at the University of North Texas, explored this phenomenon when developing a search control strategy for our robot MAVRIC. Our paper, “Foraging Search at the Edge of Chaos” provides a detailed description of simple CPGs as found in nature and the one we simulated (using the simulated neuronal and synaptic dynamics described above) to produce the ‘drunken sailor walk’ search pattern described in that paper[182]. CPGs of various kinds are responsible for most of the kinds of dynamic muscle coordination needed, for example, to generate varying gaits

---

[180] Sebastian Seung (2012) provides an excellent description of the networks formed from neuronal connections as well as of neuronal plasticity.

[181] For example see: Alford & Alpert (2014) for a description of a lamprey swimming CPG. http://journal.frontiersin.org/Journal/10.3389/fncel.2014.00290/full

[182] See Mobus & Fisher (1999). Also see the Wikipedia article: http://en.wikipedia.org/wiki/Central_pattern_generator

in running. The same circuit can respond to different input signals by changing average frequency and amplitude of the wave pattern output. This is an example of nature's phenomenal way to provide multiple functionality from single components.

CPGs generally do not involve a great deal of learning. They are basically multi-phasic oscillators that switch modes depending on variations in input (through synapses from other circuits) signal. Those inputs, however, can be subject to long-term modification due to learning, i.e. memory traces encoded in the synaptic chain in other circuits. The above referenced paper on “Causal Inference” also explains the basic neuronal unit of networks that learn associations between signals that are hard wired to convey semantic information (e.g. the smell of food) and signals that come from incidentally firing circuits. If the latter consistently fire a short time prior to the firing of the semantic signal, then a longer-term association is encoded between them, such that the incidental signal (called a conditioned stimulus in the psychology literature) may be sufficient, by itself, to cause the receiving neuron to fire, leading eventually to a motor response (the conditioned response).

Another wonderful example of a local circuit network that preprocesses sensory information is found in the retina of the eye. Here a variety of neuron types receive signals from the rods and cones (light detectors) and process information about intensity changes between neighbors (including the timing) such that they generate signals conveying information about direction of motion of objects in the visual field. Indeed the networks in the retina supply a fair amount of information that is then sent to the brain via many fewer axonal processes than one might have thought necessary. The eye produces something like primitive meaning extraction from the visual world that saves communications costs in getting signals to the brain.

Local circuits perform specific processing tasks concerned with things like feature detection, association processing, and motor synchronization. Various processing module (next section) contain many such local circuits that then need to be coordinated. Often these circuits compete with one another through mutual inhibition and a “winner-take-all” emergence of activation.

Nuclei and Cortical Modules

There are two basic kinds of structures into which neural circuits are built. One is essentially a globular-like cluster called a nucleus[183]. The other, discussed below, is a more extended structure called a cortex.

The nuclei are not just aggregates of undifferentiated cells or homogeneously distributed neurons. They have internal structure and may be comprised of many kinds of neurons. But they are generally more primitive in that they usually perform fairly specific functions, such as acting as distribution relays for incoming sensory data. Or they may process the data to extract

---

[183] In actuality there are many different “shapes” for these circuits but as a rule they are globular-like as opposed to the sheet-like structures of cortices.

meaningful spatiotemporal patterns, and if they find such patterns to signal the presence and location of objects in the environment that match those patterns. The patterns to be recognized are hard coded into the neural circuits with very little if any learning taking place. These nuclei simply respond to the presence of a specified pattern in the data stream from the sensors and activate or operate to select an evolutionarily determined response. For many of these nuclei the main output is to secrete neuromodulators that may have both a neural signaling and endocrine signaling purpose. The latter have various impacts on the physiology of the body. For example the perception of a threat may trigger a fight or flight response that includes both neuromuscular priming and prepping the body for elevated metabolism to support fighting or fleeing.

Brains and Hierarchical Cybernetics

Figure 4.5 shows a schematic of this primitive brain architecture. This brain is designed to be a purely stimulus-response processor (as described in the previous chapter), having evolved from even more primitive neural networks in worms and such. It is the first brain that provides a primitive form of tactical control. That is, it is driven as much by sensory inputs from the environment as internally generated signals (e.g. hunger). It recognizes basic patterns of inputs that represent basic stimuli in its environment, such as the presence of food, danger, or a mate. Affective-driven action selection is its basic job. It maps input patterns to output actions (behaviors) that coordinate the body of the animal with what is happening in the environment for the benefit of the animal.

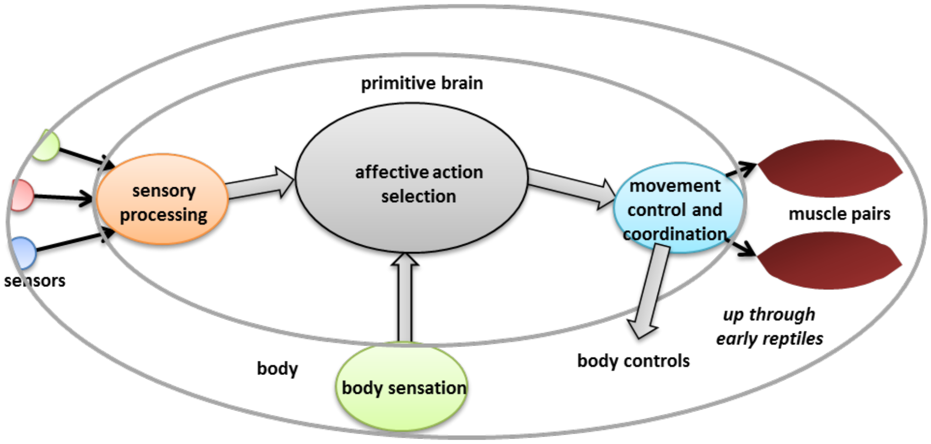


Fig. 4.5. The earliest brains developed a consistent architecture with sensory processing feeding into several general pattern recognizing modules that were hard coded by evolution to 'perceive' the environment. This primitive central nervous system matched sensory data with pre-programmed reactions (affective action selection) that were then processed into motor programs for output response. Sensory inputs included body sensations that could modify the response under certain conditions. The motor programs sent signals down the spinal cord to the muscle pair groups that moved the limbs and tail (e.g. swimming fish).

Even though this kind of brain does not learn in the conventional sense, it still possesses short-term and even intermediate-term synaptic potentiation for encoding temporal traces in the circuits.

It has a very primitive capacity for short-term memory traces so that responses that need to play out over a little longer period can do so. It would do no good to detect a threat, start to flee, and then quickly 'forget' the presence of danger. The beast would stop running as soon as the threat pattern was out of perception. Thus some form of memory is available in even these primitive brains, through the mechanisms described for synapses above.

Over the course of biological evolution the continuous emergence of new species created more complex environments for all. The world became increasingly complex requiring increasing information processing capability in those genera that were exposed to these increases[184].

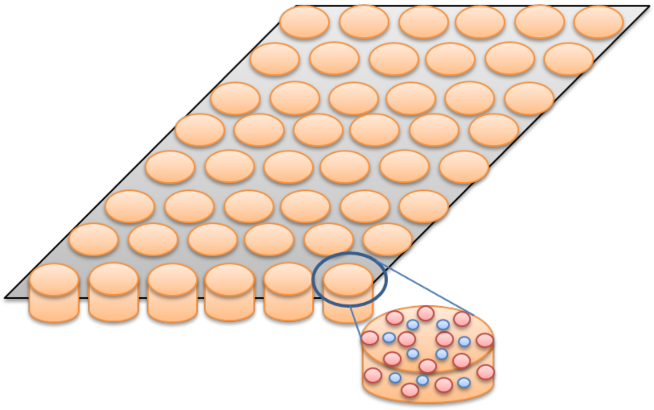


Fig. 4.6. Cortical "modules" or clusters of cells are arranged on a larger sheet-like structure. Each of the ovals represents a local cluster of cells arranged in a columnar fashion (shown at the bottom of the sheet). One cortical column is enlarged a bit to see that it is composed of many neurons that communication locally, but also some of them communicate over greater distances. The columns are actually organized in functional layers (not shown). See Baars & Gage (2002) for more detailed descriptions.

Around 250 million years ago there was a major revolution in brain architecture with the development of cortical structures. These are multi-layered sheets of neurons of many different kinds that have small-scale local networks (near neighbors talking to one another) and some long-distance communications between small clusters located far away in the sheet. Figure 4.6, above, shows a schematic representation of such a sheet with cortical columns arrayed. This figure is just meant to convey the geometrical sense of a cortical structure. The details are much more complex, even in simple vertebrate brains.

Locally, that is within a columnar cluster, neurons interact with one another, both excitatory and inhibitory connections are found. Much of this interaction is regulatory. For example some helper neurons among the pyramidal cells, can help maintain excitation of the module while others provide inhibitory feedback to keep the cluster from overdriving. Some of the pyramidal neurons are responsible for sending signals longer distances; their axons are long. Some of these

---

[184] Only some animals were exposed to increasing information loads and in those species the selection pressures that favored increasing capabilities in brains resulted in larger brains with more memory capabilities. The other species (the majority) did not need to undergo expansion of brains because their niches did not suffer increases in information loads.

long distance communications will be with other columns not too far away (for example columns have mutually inhibitory connections with next door neighbors) while others may extend to distant columns and even columns on the opposite hemisphere cortex! Figure 4.7 shows a schematic representation of these communications Note that within-column connections are dense while long-distant connections tend to be sparse.

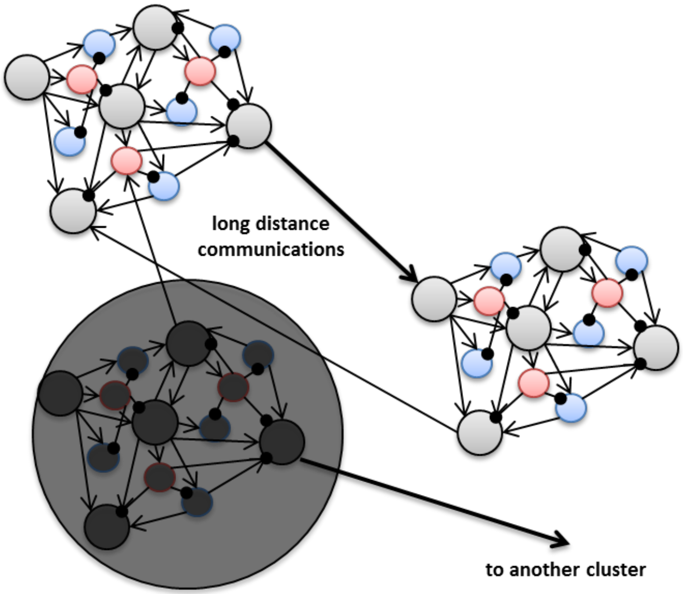


Fig. 4.7. In cortical structures one can find local clusters of tightly communicating neurons of different kinds. In this figure the red neurons send inhibitory feedback to the grey (e.g. pyramidal) cells and the blue (helper) cells. Pointed arrows are excitatory synapses and circle termini represent inhibitory synapses. Local clusters are wired to regulate the activity of the cluster. Pyramidal cells are most likely doing the bulk of long-term learning (engram encoding) whereas the other neurons are providing dynamic control, e.g. to prevent runaway excitation in closed loop feedback. Long distant communications is generally excitatory and allows the development of associations between clusters.

The main revolutionary development had to do with the cortices being organized in layers and small local clusters. The sheet arrangement enables the construction of elaborate maps, where clusters can act like positional locations of discrete percepts and concepts. Each cluster receives input from a different layer in the cortex that is receiving signals from a sensory or post-sensory processing module, such as a nucleus in the lower brain. Between the horizontal layers and the vertical clusters (e.g. cortical columns) cortical tissue appears to be well suited to encode more complex patterns than was possible in the more primitive brain. Moreover, the patterns can be learned from on-going experience rather than be hard coded into the tissues. This means that late reptiles, early mammals and birds, were able to learn important patterns in their lifetimes, meaning that they were more adaptable to changes in the environment that could never have been anticipated through evolution.

Cortices probably evolved from nuclei, some of which have layered architectures (like concentric layers in an onion). Figure 4.8 shows a schematic representation of an early cortical brain. Note that the old brain is still there, with some minor modifications, it is still doing its job. But now the newer cortex is added atop the old brain and receives neural as well as endocrine (neuromodulator) signals from the older. The newer brain adds additional motor control (esp. learned behavior patterns) to the whole organism, which is relayed down through the older brain with its established interface to the musculature.

Cortical structures capable of longer-term adaptation (learning) enabled creatures to build more elaborate mental representations of the entities and processes they experienced. Thus, animals that possessed these innovations were able to use learned models of their world to tactically interact with it. Their brains tended to specialize in tactical control development.

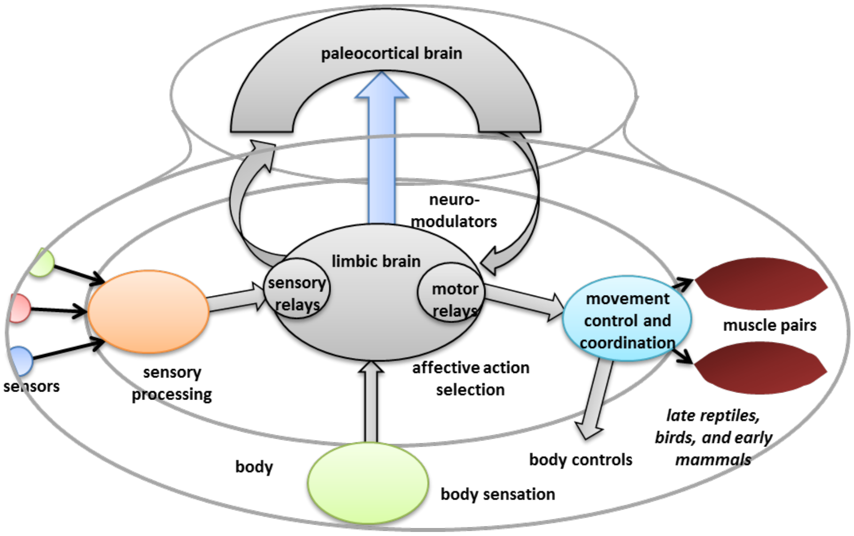


Fig. 4.8. A major leap in the evolution of brains involved the development of cortical structures that could encode long-term memory traces of patterns and relations. The paleocortex evolved in later reptiles, birds, and early mammals. Cortices are sheets of cells in which local circuits (e.g. cortical columns) can encode object patterns (learned objects) and long-distant communications between these circuits provide ways to encode relations between objects. The cortical structures provided the first real environment learning capabilities.

The new brain is comprised of the old brain and a new cortical structure (called the paleocortex since it was evolved before the newer neocortex). The old brain still does its jobs and influences the newer cortex in doing its job. Initially this is just learning patterns and associations that help modulate the older affective pattern-response matching. To some degree the newer abilities had to regulate or modulate the older capacities in order to produce more nuanced responses to patterns that the old brain might have misclassified. But the newer brain did not yet have the more elaborate model building capabilities that would come with the evolution of the neocortex in latter mammals.

The Neocortical Brain, the Prefrontal Cortex and Executive Control

The final major revolution in brain architecture comes with the advent of a yet newer layer of cortical tissue called the neocortex. Like its predecessor it is comprised of additional layers and is internally organized to represent many more perceptual and/or conceptual details (Fig. 4.9 below). It is also capable of forming transitory connections between clusters (now representing wholly formed concepts) to do ‘what-if’ analysis. That is, it can form ad hoc connections between dynamically represented concepts just to see what happens. This is the origin of creativity and invention. It is also the basis for the development of rational decision processing. Maps of situation concepts, decision nodes, and learned outcomes from experience can now be represented and used to guide future decisions under similar sets of conditions. This applies to learned decision paths (e.g. in expertise) and possible decision paths (models).

Figure 4.9 provides a summary schematic of the neocortex. The layout is roughly what we find in the brain where the left side of the schematic represents the posterior primary sensory and early perceptual processing regions (the occipital lobe for vision, etc.) The posterior to central regions of the parietal and temporal lobes are convergence zones[185] where early sensory modality associations are encoded as simple concepts (e.g. ‘fur’ composed of hairs, colors, and numerous other features). In the section below, entitled, “Representing Concepts” I will provide finer details on how neurons and neuron clusters form these convergence zones and do the encoding of engrams that form neural assemblies representing concepts[186]. Further forward (toward the anterior portions of these two lobes) more complex concepts are constructed from many of these simpler concepts through more complex convergences (e.g. ‘animal with teeth, tail, four legs, and having fur’).

The posterior frontal lobe cortex is yet another convergence ‘region’ with convergence zones for meta-concepts or abstractions (the above described animal is a ‘mammal’) and relations between concepts (‘the mammal runs’) and between the concepts and the self (‘the mammal is running toward me with mouth open!’).

Roughly speaking the complexity of concept construction and abstraction increases from left to right (posterior to anterior) in the schematic. Low level percepts and simple concepts send projections (straight black arrows) into higher cortex regions. Inputs come from many lower nodes converging on the target concepts. In turn these higher nodes project back onto the lower nodes that innervated them (red arrows from right to left on the upper half). The details will be covered later. Thus these higher nodes are divergence zones as well as convergence zones.

---

[185] Damasio (2010) describes the notion of convergence-divergence regions and zones. See chapter 6, ‘An Architecture for Memory’ and especially starting on page 151, the section on the subject.

[186] Dehaene (2014) describes these phenomena from a high level, conceptual point of view in chapter 5, “Theorizing Consciousness.” The term, “neural assemblies,” is accredited to Donald Hebb (1949).

Tactical Planning – Instantiating Motor Sequences to Achieve a Goal

The more anterior parts of the frontal cortex are responsible for assessing the current situation of the self in relation to the world and the state of the self with regard to affective states (drives and such). Inputs from the limbic system enter this region (not shown in this figure) and are part of the on-going second-order conscious assessment of the immediate future and what sort of tactical plans should be invoked. The frontal cortex is involved, then, in low-level *planning* of tactical 'moves' that the agent should take to optimize its situation in the near future. Tactical plans are a temporarily linked set of learned behaviors (light red squares). They are also meta-concepts, using the exact same encoding machinery as was used by the object and relation concept formations in the top half of the schematic. This machinery is explained below in the next section on "Neural Basis of Systems Representation and Models." Later still in the section on "Systems Intelligence" I cover the construction of heuristic programs in neural tissue to show how just such a linked set can be created and executed. As with all other areas of the cortex, temporary linkages can become more permanent (incorporated in long-term memory) with sufficient positive reinforcement after being used.

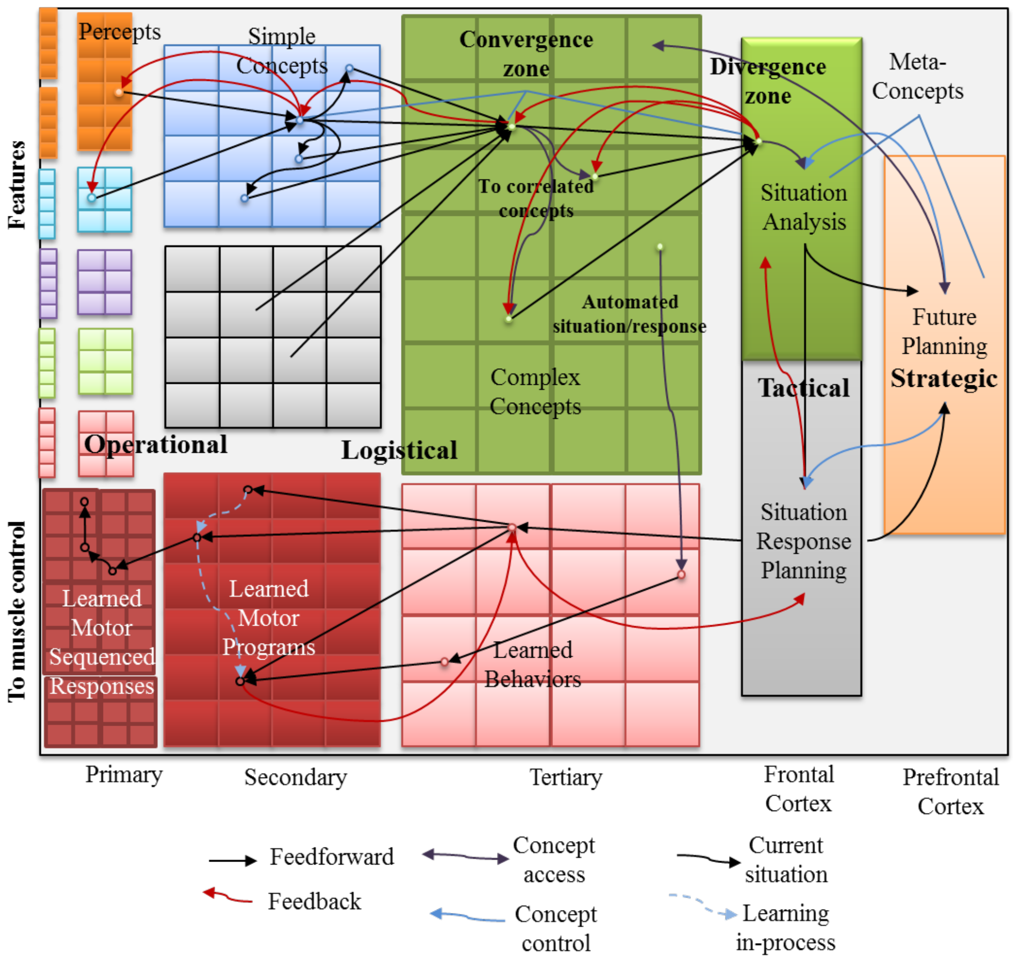


Fig. 4.9. The neocortex provides a much higher level computational capability for all of the hierarchical cybernetic system. The neocortex adds a capability to encode more complex relations between concepts. The basic mammalian brain includes everything shown up through the tactical level systems (on the right). The advanced mammalian brain extends the frontal cortex with the prefrontal cortex in which executive controls allow access directly to concepts for longer term planning purposes. The prefrontal cortex has direct access to much more of the conceptual encoding regions of the neocortex (in anterior parietal lobes, for example). In humans the prefrontal cortex is highly expanded with the Brodmann area 10 (discussed below) substantially expanded relative to the rest of the prefrontal region. See text for detailed explanations.

In the figure, in the Complex Concepts (green) area to the right, I show a single concept that constitutes a previously learned situation that has a feedforward link (labeled "Automated Situation/Response") to a Learned Behavior node (red square). Whenever that concept is activated it can automatically activate the learned behavior. Tactical plans (meta-concepts) can be saved for future use in similar circumstances, saving the time needed to organize and link a sequence of behaviors on the fly. This is how routine actions come into being. When you drive to work every day using the same route and encountering the same traffic conditions, the tactics eventually become rote. You no longer need to 'think' about what to do next; the meta-concept of getting to work becomes automatic.

In the Learned Motor Programs region there are three nodes in the process of being linked in sequence. There is a new Learned Behavior node being activated that is activating the three nodes in the Program region in sequence and the dashed blue arrows represent the fact that these are not permanent links, just yet. If this cascade of behaviors is repeated often with appropriate feedback then the links will become permanent or at least more so, so that faster response times are achieved when the initiating situation activates the higher level behavior node. See the section below, Representing Concepts in Neural Networks, for more details regarding the mechanisms for creating these links.

Conscious thinking is the process where new temporary linkages are being made under control of areas toward the anterior regions of the prefrontal cortex (the blue arrows from the Future Planning – far right – back to the frontal lobes in the figure). This process involves search through the existing learned behaviors for a pattern match between the current goal state and the desired outcome (figure 4.10 below). In my journal article (Mobus, 1999), "Foraging Search: Prototypical Intelligence" I developed the thesis that CPG like circuits in the cortex or closely linked sub-cortical structures could be responsible for driving an analogous search through concept space in the cortex (as opposed to searching for food in the environment).

> An argument is advanced that searching for resources in the physical world, as per the foraging model, is a prototype for generalized search for conceptual resources as when we think. A problem represents a conceptual disturbance in a homeostatic sense. The finding of a solution restores the homeostatic balance. The establishment of links between conceptual cues and solutions (resources) and the later use of those cues to think through to

> solutions of quasi-isomorphic problems is, essentially, foraging for ideas. It is a quite natural extension of the fundamental foraging model[187].

Since that time evidence for a process of search through concept space (abstract, object, relations, and behaviors) to find appropriate matches has increased[188]. The search is based on pattern matching (as discussed in chapter 1), but also running simulations (of behaviors) to see if the outcomes match the goals. When candidate tactical behaviors are found the next step is to consider the inputs to the behavior that result in the desired outputs and use those to develop the sub-goals needed to reach in order to make progress toward the ultimate tactical goal.

The process proceeds from final state back through time and sub-goal states to the current state and the first step tactical move (behavior) to be generated. Once a sequence of behaviors (along with alternate paths through contingent behaviors) is linked the execution of a tactical plan can be initiated.

There are a number of important constraints on the thinking process that keep it from being 'perfect.' For one thing the process of search and testing each candidate behavior takes time, even if implemented in a massively parallel form. Thus there is a need to pare the search candidates. In chapter 2, the section titled, "Making Decisions: Putting the Constructs Back Together," I covered this basic notion of how decision nodes in a search tree are pruned away in order to save time. This is the same basic idea. Concepts are the nodes in such a tree (here generalized to a network) and the prefrontal cortex may be using cue learning to keep the search manageable in time and space.

Another major constraint is space – neural circuit space in the cortex. Here I am referring specifically to the space called 'working memory' which is suspected to be managed by the prefrontal cortex. As in figure 1.5, where the amount of space given to conscious thought is represented as a very small percentage of total mind space, this is essentially what we mean by working memory. Much research has shown that this memory is limited to a few 'chunks' of memory, what I have called meta-concepts, which can be held in conscious memory at one time. The amount of space accessible by conscious thinking is limited and relatively small. But, as it turns out, most near-term tactical plans need not be built or recognized in conscious memory. Indeed, as figure 1.5 suggests (and the neurological research confirms) most of our thinking goes on in the subconscious mind. What may be the process is that we consciously think about the ultimate goal we want to achieve (and the story of how motivations drive selection of goals is another whole story!) We then may do a conscious search for the final behavior that we have

---

[187] Mobus (1999).

[188] Stanislas Dehaene (2014) and Jean-Pierre Changeux developed a computer model of the cortex, a simulation of what they call the "global workspace" where an "ignition" of excitation in higher (anterior) cortex represents the consciousness of thoughts. In Dehaene's book (chapter 5) he describes this very same idea of a central pattern generator kind of search through the global workspace (page 189). It was quite gratifying to see this idea being proposed by neuropsychologists using both imaging data and computer modeling evidence after my hypothesizing based on systems science principles!

learned produces that outcome. And then we submit that to subconscious thinking that takes over a massively parallel search through sub-goal behavior space, linking candidates in contingent chains, and continuing until the current situation analysis matches the inputs to a starting behavior, or in other words, the next behavior that should be undertaken to start the execution of the most likely chain.

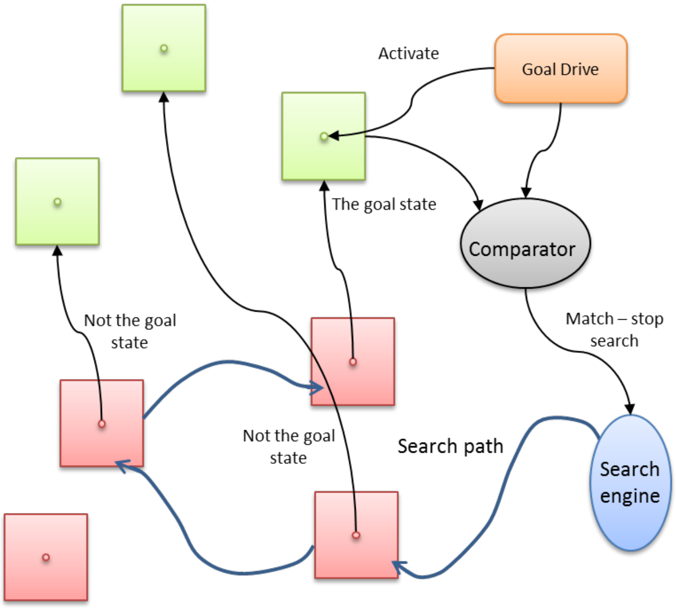


Fig. 4.10. Foraging search is a likely candidate for finding potential goal states and then determining if it compares favorably with the goal drive. Here a goal drive activates its learned associative goal state concept. The search engine (a hypothetical circuit in the prefrontal cortex) conducts a foraging search (blue wavy arrows) among behavior or motor programs (red boxes). The search is not exhaustive, skipping around in a quasi-random fashion (see Mobus, 1999). When it finds one that is associated with an activated concept a comparator circuit (possibly in the anterior cingulate cortex) determines the degree of efficacy of the found concept. If the goal drive and represented goal state are similar, the search is halted and the behavior concept is linked into the growing tactical plan list.

Strategic Planning – Instantiating Long-term Behaviors for Big Goals

The final level of management in the higher mammalian and human brain involves the future planning capabilities of the prefrontal cortex areas. Higher mammals such as the primates, and particularly the great apes, are able to plan whole combinations of tactical plans in advance of executing them. However the ability is limited to fairly short time horizons (on the order of only a day in the case of our nearest cousins).

In order for this new capability to work, however, it needed a much more elaborate form of executive control over the forming and testing of new circuits. The motor system (or rather the premotor cortex that had evolved to do a primitive form of planning for multiple behavioral options that could be chosen as needed in highly volatile environments) gave rise to the prefrontal cortical regions in the primate brains. There we see a final convergence zone for every kind of neural signal from the most primitive parts of the brain as well as the later evolved parts.

This region is responsible for thinking, but also for melding affective signals and sapient signals (judgment) with on-going rational decision processing. The latter is probably under the direct control of the premotor area (dorso-frontal area) and acts not too differently from a computer churning away through a graph search algorithm. The analogy is not really that good, however. Effectively the brain progressively activates concentric rings of clusters representing possible decision points. Those clusters that have strong damping either from the limbic system or from the sapience system are pruned from the decision tree. Those that have strong positive (excitation) from either system will be preferred such that the nodes they are subsequently connected to will be activated. Mutual cross inhibition of nodes further reduces the number of parallel paths as decision options are eliminated. Thinking through a problem requires both conscious directional control (such as always checking the direction against *a priori* desired results) and subconscious affective/judgment steering control.

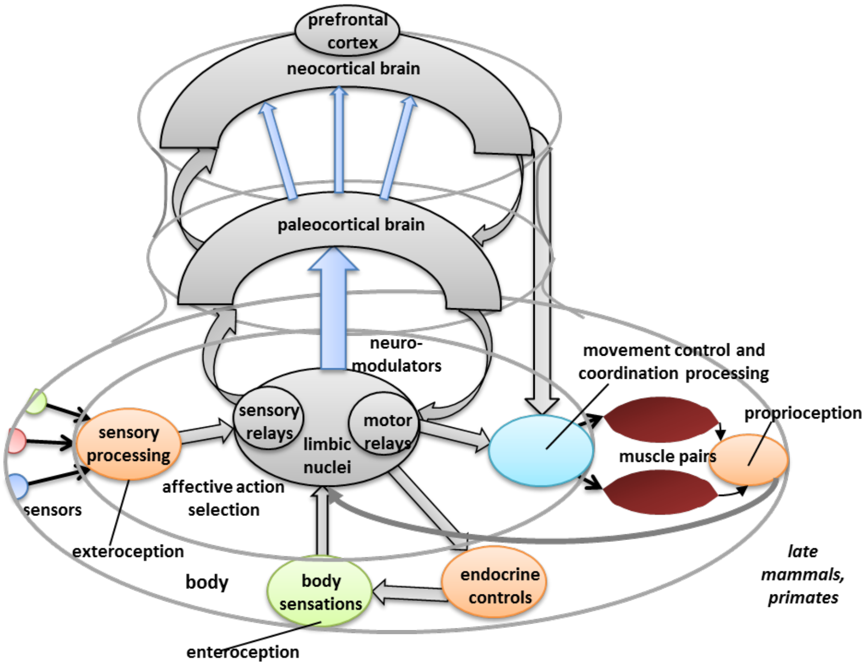


Figure 4.11. The neocortex adds another level of complexity to the established paleocortical brain. The neocortex can encode much more detail and many more patterns than could the paleocortex. Moreover, the neocortex allowed the development of circuits that could be used to build models of the dynamics of the world as learned through experience. Much of this is encoded tacitly and can be brought to recall/usage by special executive control circuits evolved in the prefrontal cortex.

As fantastical as all of this sounds it actually represents a straightforward development of hierarchical levels of complexity based on the lower levels described above. The capacity to rapidly and tentatively form excitatory connections between cell clusters ultimately depends on the same factors that control neural wiring development (say in early embryonic development) and the dynamics of synapses able to locally modify their efficacy and maintain it for long periods of time relative to the life of the animal. These are inherent in the most primitive animal neurons. It is just minor modifications here and there in the genetic program that produces neurons and circuit forms that have been successfully selected throughout the course of animal evolution that result in the high degree of complexity we see in the human brain today.

The last step in hominin evolution involved the enlargement and capabilities of the prefrontal patch of tissue known as Brodmann area 10 (BA10 discussed below). It is the patch right behind the eyebrows. It appears to have made a rapid expansion about the same time that we think humans developed language and other social cognitive skills (see the next chapter for more explanation). It is known to be involved in many judgment functions. It is also known to have a more advanced form of communication with other parts of the brain through von Economo (spindle) cells[189]. Finally it is highly connected (recurrently) with every other part of the prefrontal cortex, suggesting that it is truly an ultimate convergence zone for control signals.

I posit BA10 is the seat of both consciousness (self-, other-, and of awareness) and sapience. Our brains (and here I refer to animals in general) first evolved to monitor our bodies and our environment, to match the needs of the body and its safety with the situations found in the environment. That monitoring and control function still exists and is still carried out by the limbic and lower brain areas. This is part of Kahneman's (2011) *fast system 1*. What we think happened in the evolution of sapience is that the high speed connections back to the limbic system (see the discussion in the section on Emotions below) that help to regulate or dampen the limbic responses increased substantially. Perhaps in most humans, however, these connections or the control competencies of BA10 are weak enough that the limbic system still has undue sway over our actions. Our urges, our desires, and our reactions to what is happening carry more weight in our decisions than they should. But there is *some* control, even so. The model of a *level* of sapience is that some humans have exceptional capacity in down modulating emotions so that sapient judgment can provide the stronger influence over intelligence and creativity. Such a capability would necessarily involve much longer range and longer time scale thinking since it would not be motivated by immediate gratification needs associated with limbic-driven decisions.

As with all but a few (r)evolutionary developments in the brain the changes have been based on simple modifications at one scale of complexity that gave rise to new competencies and even new scales of complexity. I suspect we will find the evolution of BA10 was based on some pretty minor tweaking at the genetic level having to do with the developmental control over the prefrontal cortex (you know the old story about how little difference there is between chimpanzee DNA and our own!). But because it was built on an incredible hierarchy of complexity the resulting uber-complexification was nothing short of spectacular.

In embryological and fetal development of the brain we know that various parts of the brain develop (grow in size and complexity) in response to internal stimulations. The early primary visual cortex, for example, experiences self-excitation which stimulates the growth of neurons

---

[189] See the Wikipedia article: http://en.wikipedia.org/wiki/Spindle_neuron

and their axons/synapses. It is possible that even a seemingly minor change in the size of BA10 could stimulate significant increases in the other areas of the brain that it innervates. It wouldn't have been necessary for multiple genetic changes to have coincidentally happened at the same time to produce the expansion of the large human neocortex. A simple increase in BA10 could have influenced increased growth in the rest of the prefrontal cortex, and that, in turn influenced other areas of the neocortex. This developmental process continues long into adolescence and even young adult life as the prefrontal cortex is still developing into the second decade of life.

Now that we have surveyed the architecture of the cortical brain I want to go into some details of how this structure can encode the components of representations of the world of experience – the concepts that are accessible for conscious thinking and also for subconscious modeling.

Neural Basis of Systems Representation and Models

As I claimed in the prior chapter our brains are literally pre-wired to encode the perception and conception of systemness – the properties and dynamics of systems. In this section I will build the argument for neural circuits that encode those properties. Ultimately the very mental models that we have in our brains for how things work and how the world in general works are systems models, not too dissimilar from the kinds of computer models that system dynamics researchers build in attempts to make predictions about how the systems they study will behave in the future under varying environmental conditions[190].

The brain is a modelling engine. It perceives the world in a way that allows it to construct models of objects, agents, and their interactions in the world. The mind uses those models to reconstruct the world (memory recall) and to make predictions or anticipations about the future states of the world should various contingent conditions obtain. All of this must be grounded in the way in which neural circuits are formed during both prenatal (mostly inherent) and postnatal (largely experientially driven) development. The inherent ability of the brain to perceive and conceive systems is in our biology. The ability to build modifiable and adaptive models is in our ability to learn as we live. Both together make it possible for humans to understand the workings and meanings of the world.

Representing Causal Relations in Neurons

Amazing as it may sound, the capacity to represent the world out there with networks of neurons inside the head begins with the tiniest bit of neural machinery, the synapse, the connecting point in communications from one neuron to another. During a particularly fruitful time in the 1980s a number of artificial neural network (ANN) models were developed[191]. These were computer

---

[190] Chapter 13 of Mobus & Kalton (2014) describes several different approaches to modeling real phenomena including the approach generally called system dynamics. All dynamic models embody the structural and functional aspects of the system being modeled and attempt to compute the future states of the system based on a stream of inputs and starting conditions. The brain, that is the cerebral cortex, does all of this as well.

[191] See the Wikipedia article: http://en.wikipedia.org/wiki/Artificial_neural_network for more details on the subject of ANNs.

simulations of what many researchers believed to be a semblance of what goes on in the brain. They were necessarily quite simple and treated synapses as scalar weights. This never seemed quite right to me and I pursued a different route, which I will describe below.

Most of the work on artificial neural networks centered on a concept called "distributed representation"[192]. This term was used to describe a scheme in which pattern encoding was distributed among all synaptic junctions in a fully connected network (see the diagram in the Wikipedia article referenced in end note 20 on ANN above.) The main idea was that every synapse in such a network participated, in some small way, in the encoding of every pattern that the network was trained to recognize. In fact this actually worked for relatively small sets of non-heterogeneous patterns (e.g. recognizing individual faces from a large library of faces). But the concept ran into trouble for very large sets or for dealing with heterogeneous patterns. One of the main problems had to do with the amount of time it would take to train the networks. At least one computer scientist determined that as the network grew in size to accommodate larger problem sets the time it took to train the network grew exponentially large[193]. This means that such networks are going to be limited severely in terms of what they can learn to represent.

The early researchers pushing the distributed representation paradigm were convinced that their first successes meant that the brain actually stored information in the way their model networks did. There was a long and deep debate regarding the differences between distributed representation versus what is called *local representation* — the idea that a limited number of neurons encode specific patterns. That debate has been largely settled of late by recognition that specific clusters of neurons do fire differentially in response to specific patterns presented to the sensory system. For example we now know that there are neurons that fire whenever a generic face (even the 'happy face') is presented (work done in monkeys)[194]. Another set of neurons fire whenever the face of a generic monkey is presented. Another set, still, fires when the face of a conspecific is presented, and another set when a specific member of the colony is viewed. When the latter example is the case, all of the prior clusters are also firing, suggesting that 1) there is a hierarchy of representation from generic down to specifics, and 2) that patterns are indeed represented locally, in a hierarchy of features, rather than distributed throughout the brain. As I will show below, this local representation is actually just a focal point for specific concepts (encoded patterns). However, it turns out, too, that the total feature representation of a specific pattern is distributed, but only amongst local clusters at lower levels in the hierarchy.

While the ANN work was receiving so much attention in the 1980s and 90s (and actually continues to dominate some neurological thinking even today) I felt dissatisfied with the lack of biological realism being modeled. Synaptic weights, as represented in the artificial neurons, did not seem to me to adequately represent the dynamics of what was then known to occur in real

---

[192] Rumelhart, et al (1986); Hanson & Olson (1990)

[193] Judd (1991) – warning! Heavy duty math.

[194] Scalaidhe, et al (1999).

synapses[195]. So I set about trying to formulate a computer model that did a better job at emulating biological synapses.

In work that I did in the early 90s I built a computer simulation model of a more biological-like synapse, which I dubbed the Adaptrode (Mobus, 1994; and see below). The main feature of the Adaptrode as a mechanism for learning is its ability to capture multi-time scale associative information through a reinforcement mechanism. What this means is that the Adaptrode could record a memory trace in the short run, from incoming action potentials, and, if that recording were reinforced by a signal coming through a second channel shortly after, it would record another intermediate-term memory trace at a somewhat weaker level. The second trace, while weaker, nevertheless kept the memory trace for a longer period of time. Then after a longer time had passed, if yet another signal arrived via a third (or even fourth) channel, the memory trace ended up in a long-term form. I was able to show that this approach went a long way toward solving the non-homogeneous, non-stationary relation learning problem.

In other words, the Adaptrode mechanism, when incorporated into a simulated neuron, allowed that neuron to have short-, intermediate-, and long-term memory traces recording the association between two or more external sources. In Mobus (1994) and in my PhD dissertation I showed that the Adaptrode could emulate Pavlovian conditioning[196]. And I further showed that such conditioning is a necessary part of any representation of causal relations.

A causal relation is of the general form:

$$A \Rightarrow_C B, \text{ if } T_A <_{int} T_B$$

where event A and event B are observed in near proximity and the time of event A, T_A, comes before the time of event B, T_B by a small interval, *int*.

There are more specific forms of such relations, for example a stochastic form allows that events A and B are probabilistic but that the above situation must be true more often than not. There can also be restrictions placed on the relation, such as that B must never precede A within a certain time interval. However, these are just ways to formally capture what we all readily perceive when we say that A causes B.

One of the more interesting versions of causal relations is that of circular causation. Most of us are satisfied with a simple A causes B kind of explanation for things, like hitting the cue ball with the stick causes it to roll and hit the target ball in billiards. In such situations we are happy with a unidirectional flow of causality. But we might just ask, what caused one to hit the cue ball in the first place? Most would imagine a chain of causal relations going back to something like the

---

[195] See Alkon, (1987) and Mobus (1994) for details of synaptic dynamics.

[196] Pavlovian conditioning is a form of predictive modeling in that the animal learns to predict a meaningful situation (like the availability of food) based on cue events. This is an anticipatory computation that is learned by experience.

person wanted to get the target ball into a corner pocket. Our explanation can go even further back and suggest that the individual wanted to play the game, and so on. But few will ever say something like: the memory of the target ball going into a pocket, if the angle it is hit by the cue ball is just right, causes the person to hit the ball! In other words, something about the ball and pocket (an effect) actually is part of the cause of hitting the cue ball.

This, possibly strained, example is an instance of circular causality, wherein an effect loops back somehow to be a component cause of the event that caused it. The notion of circular causality is an absolute no-no in most logics, but it goes on all around us all the time. *A* causes *B*, which causes *C*, which feeds back to affect *A*, and so on. In the billiard example, if the person misses her shot that fact feeds back into her memory and her brain tries to build a more correct aiming strategy for the next time she shoots. If she finds herself repeating the same kind of shot, she might do a better job of it.

It turns out that an ability to represent causal relations is essential to building models of how things work. Causal relations are also captured in simple algebraic functions such as: $y = f(x)$. This is interpreted as the value of y changes in proportion to some change in the value of x. The function associates two variables, x and y where the latter is considered the dependent variable. Hence, a change in x 'causes' a change in y. What the Adaptrode allowed me to do was to build neural circuits (networks of neurons) that represented such relations but also learned those relations and their strength of association over very long time frames. For example, my MAVRIC robotic experiments showed that a wandering robot could learn to associate one kind of sound with a light, that meant reward (well, robot reward) and another combination that meant harm and then always approach the reward-combo while avoiding the harm-combo. Thus the robot learned to represent cause and effect, a sensed phenomenon with a reinforcing signal, either reward or punishment[197].

Representing causality is something neurons can do quite well. Figure 4.12 shows a stylized neuron (Fig. 4.12A) and a schematic of what happens inside the neuron to cause associative (and causal) adaptation to occur (Fig. 4.12B).

---

[197] See Mobus, (1994 and 2000).

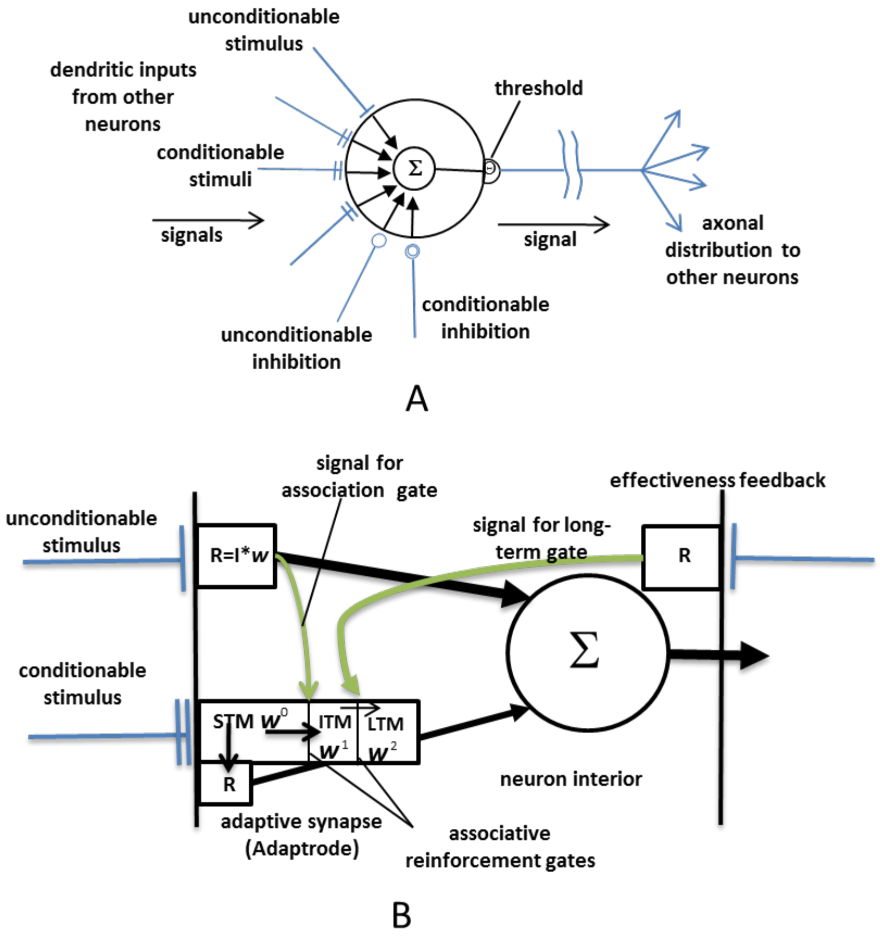


Fig. 4.12. An adaptable (learning) neuron (A) and how it associates a conditionable stimulus with an unconditionable stimulus (B). See text for details.

The neuron in Fig. 4.12A receives multiple excitatory and inhibitory inputs from many other neurons in the network. The single barred terminus, labeled "unconditionable stimulus" brings a non-adaptable stimulus signal to the neuron. This signal is generally highly effective in contributing to the overall excitation level of the neuron. The circle with a Σ is a spatiotemporal integrator (actually the cell membrane) that adds all of the incoming signals and sends the current level of excitation to the Θ threshold in the axonal hillock. If the summation exceeds the threshold the cell fires output action potentials. Their frequency is proportional to the fluctuating excitation of the cell. Both excitatory and inhibitory inputs can contribute. The double barred termini represent adaptable (meaning plastic) synapses (inputs) that can become stronger in their influence under conditions of sustained inputs as well as temporally correlated excitation of the unconditionable stimulus. If any of these synapse become stronger they can begin to cause the neuron to fire even in the absence of the unconditionable stimulus. Thus the pattern of learned inputs can cause the neuron to come to represent that pattern and fire the neuron whenever that pattern is present in the input array.

Figure 4.12B details how the learning occurs and gives a hint at the role of time in the adaptation process. Here we see an unconditionable and a conditionable stimulus together. This is representative of signaling circuits (actually internal flows of chemical concentrations!) inside the neuron. Just inside the synapse is a compartment that responds to the input signals. In the figure the response of a synaptic signal, R, is determined by the synaptic efficacy or weight (w) times the real-time input signal (I, frequency of action potentials). That response, in turn, signals the integrator. In the case of the unconditionable input, the weight, w, is a high and fixed value, meaning that the synapse is competent to cause the neuron to fire by itself, what Daniel Alkon (1987) calls a "flow-through" synapse. Such synapses carry important semantic information. In the Pavlovian dog-salivation conditioning experiment, this signal was the smell of food, while the conditionable signal was the ringing of the bell which had to precede the food.

The conditionable synapse, immediately beneath the unconditionable one, is inherently weak to begin with; its weight value is near zero, and so it cannot fire the neuron. Note that the compartment behind this synapse has three sub-compartments representing three stages in the evolution (dynamics) of a memory encoding. Initially w^0 is low. A steady (strong) input signal will tend to elevate the w^0 value as a short-term memory (STM), but as soon as the signal falls off, w^0 will decay again toward its initial low level. However it takes a bit of time for this decay to occur. During that time, should the unconditionable stimulus cause the cell to fire, it also opens a 'gate' that allows whatever the value of w is at that instant to change a similar weighting value in the intermediate-term memory area (ITM w^1). This means that a short-term memory trace is saved for a longer time (the decay in the ITM area is much, much slower than in the STM). What has happened is that an association between the conditionable and unconditionable signals has been established. Moreover, the way these synapses work, the conditionable signal had to occur a short time interval prior to the unconditionable signal otherwise the gate is locked shut. What this means is that the conditionable signal represents something that is "causally" associated with the unconditionable signal (or whatever generated it). The conditionable signal becomes, in a very real sense, a predictor of the unconditionable signal.

Over an even longer time scale, if the firing of the neuron sets into motion some downstream activity then that will provide excitatory (or possibly inhibitory) feedback to the cell, either at another synapse or through neuromodulators. That feedback can then open a second gate that will allow the STM to raise the w^2 value in the long-term memory area (LTM). Thus the memory trace becomes associated not just with an original semantic signal, but also with longer-term rewards (or possible punishments). Because of the time lags involved, these associations are strongly causal in inferential nature. The conditionable synapses are locked out from encoding associations if either the unconditionable signal or the feedback signal comes before the conditionable signal in real time.

Over many reoccurrences of these temporal associations of signals, the conditionable synapse will develop a much stronger weighting (efficacy) that will allow it to contribute significantly to the cell firing on its own. Typically, in large pyramidal neurons with thousands of inputs, it will be a pattern of many synapses (coming from multiple sources) that build up enough strength such that collectively they can cause the neuron to fire in the absence of semantic inputs. But the principle of temporal encoding shown above is operative in these cases (see, Mobus, 1992 for more details).

It might not be immediately obvious, but the multiple time domains involved in Adaptrode (and real synapse) encoding is the solution to learning non-homogeneous non-stationary relations. The synapse not only encodes traces in multiple time scales, it also “forgets” traces, or trace strength, in those same time domains. The trace will decay over time if not reinforced relatively frequently. The synapse retains a weak but long-term efficacy that cannot necessarily excite the neuron, but does not disappear entirely. This kind of weakening is necessary in order for the neurons to be able to participate in multiple traces and different times.

What happens is that if a relation (between a cue and a consequence) that is learned somehow becomes no longer true in the real world, then the brain does not need to retain the trace of the relation, at least not at the level it did when the relation was predictive. The trace decays over time but never really totally disappears, however, so that if the relation were to again become true, then the trace could quickly be brought back up to strength and thus operative in making predictions.

Armed with the notion that small neuronal circuits can capture and represent causal relations as well as general associations (pattern recognition) it is just a small step to develop a theory of construction of dynamic systems models using neural circuits. System dynamics modeling gives us a clue. Such models are, in fact, networks of components (or 'stocks') and connected by dynamic directional links (or 'flows') that implement feedback loops and, yes, causal relationships. Networks of living neurons, or rather small neural circuits called cortical columns are able to learn to represent perceptual patterns (like faces) and learn to associate patterns to form meta-patterns, or concepts. The latter are not just static representations, however. The causal dynamic described above allows these concepts to interact with one another as models of how the real-world counterparts interact. The invocation of one concept (or percept, for that matter) can invoke related concepts. We experience this as things coming to mind, or being reminded. When we think of that beautiful sunset we also think about our lover with whom we shared the experience once.

Representing Concepts in Neural Networks

The Concept of a System

In the last chapter I demonstrated how concepts of things and of relations can be formed in both specific (episodic) and general (tacit) memory forms (figures 3.7 – 3.9). Below I will provide a general scheme for how all concepts get encoded. Coupled with the causal relation encoding covered above, we have the necessary ingredients for representing system models in neural networks. Remember the main lexical elements of system language name processes (transformations of inputs to outputs), boundaries, input and output flows (matter, energy, and messages), sources and sinks, and stocks (buffers, reservoirs, etc.). Mobus & Kalton (2014, chapter 12) contains a more complete lexicon along with icons used to represent the terms in visual form. The key is to grasp how neural networks using the causal encoding mechanism described above can come to represent each of these lexical elements, the syntactic structures using those elements, and the semantic connections to sensory modalities and their images.

Here I provide a basic description of such encoding for the basis of a template system, that is a system model that can readily give rise to detailed mental models of actual systems as found in the world. All systems share a fundamental architecture that is described by the system language. The representation in neural tissues (the neocortex and its cortical columns that form mutually exciting clusters) is based on a simple, yet non-obvious fact. Neurons and neural circuits, like the columns, are themselves systems and so can represent systems in the world according to principle 3. Systems are networks of relations between components at any level of organization. So it should not be surprising that the brain, which represents everything in this fashion, is also a natural framework for representing systemness itself.

Linguists and semioticians have long recognized the primacy of three basic kinds of verbal representation of "things" in the world. Signs (semiotics) come in three basic forms: icons, visual resemblance of a sign to the thing in the world, indexes, pointers to the thing in the world, and symbols, arbitrary and abstract representations of things in the world that are manipulatable in ways the others are not. We should be able to identify neural encodings of all of these aspects if we are to understand how humans have risen to 2½-order consciousness that, in turn, leads to the emergence of public symbolic, recursive representative language.

Icons are visual representations of the thing itself in the world. An icon of an object would be a two-dimensional (usually) graphic that is shaped in a way to suggest the object itself. In the book, so far, I have represented generic objects as ovals. The oval shape is a stand-in for the actual shape of an object. It merely represents that there is a real "thing" that has an identifiable boundary and suggests that there is internal structure (the black box view). The oval is also indicative of an object whose internal structure and functions are knowable following decomposition procedures. Another iconic representation of an object, the internals of which are not known, is the open rectangular form (see figures 3.10 - .15 for examples). These objects are known to exist and are sources or sinks of flows to/from the central object.

Mentally the icons or visual representations of objects in neural networks are the results of a process of learning (constructing) concepts as described below. The actual visual representation of a real object is encoded in a vast associative network in primary and secondary visual cortex. The details of the visual representation are stored in that cortex and are only activated by sensory input (see the mechanism below) or by recall activated by abstract representations stored in frontal/prefrontal cortex as compact neural clusters (actually clusters of cortical columns) that activate the lower level visual percepts and, from them, the features.

Iconicity is the representation capacity that allows one to recognize a smiley as a face, even when the features are extremely simplistic. It is the arrangement of, themselves, iconic features (mouth, nose, eyes) that form the basic visual pattern of a face. There are “face” neurons in the inferior temporal gyrus (where integration of visual sensory input is accomplished) that “fire” when any kind of pattern that has two orbs, a single central protrusion (or simile), and a half-moon or half-circle at the bottom is observed. Other, higher up, neuron clusters fire when more detail is provided. Still higher neural clusters fire when sufficient details are in the vision such that it is recognized as a particular face (e.g. grandmother’s face). The whole hierarchy of clusters from icon (smiley) to grandmother collectively fire (synchronously) whenever grandma is in the field of view (or is remembered visually).

Icons of visual objects is fairly straightforward owing to the nature of the visual encoding system in the brain.

Similarly, in my system language lexicon, I have used arrows of varying thicknesses and colors to represent flows of materials, energies, and messages (communications), starting in the Preface. The arrow is iconic of something going from one location in space to another location. In this case arrows go from one object to another, for example from sources to our object of interest. In the brain, however, representations of icons for flow are not necessarily visual, but are inferential. Some flows are visually representable, such as the flow of water in a stream. Flows of messages or forces are, however, not directly seen and so must be inferred from causal effects. The source object has to “act” first, followed by a reaction by the recipient (sink). The causal representation mechanism described above can capture this relation in neural tissues, thus encoding a flow relation between two objects. Thus, as depicted in figures 3.7 and 3.8 in the last chapter, concepts of flows can still be modeled. It turns out that such models are more like indexical relations than iconic ones. That is, the concept of a causal flow relation points to the idea of a flow rather than “looks” like a flow per se. Thus neural networks can represent flows but not necessarily as icons (unless one can see the flow of a visible substance from source to sink as in the case of a tube that conveys the substance.)

The combination of icons for objects and inferential concepts of flows provides us with a concept of system (figures 3.14 and 3.15). The circuits of the neocortex can embody the encoding of a system by virtue of encoding the elements of boundaries, objects, and flows with causal consequences. Second-order conscious brains have the capacity to leverage a template encoding of a system such as in figure 3.14 by adding details in the form of links to additional concepts (e.g. flow controls and stocks as described in Mobus & Kalton, 2014).

What 2½ order consciousness adds to this language modelling is a new kind of sign called a “symbol.” Symbols are almost arbitrary concepts in that can be used in a language as stand-ins for the objects, components, relations, and other elements of a system, but also represent abstractions of systems elements and their interactions. Symbol encoding takes place in the same way as iconic and indexical concept encoding except that it is based on linkage to sound productions (see figure 3.9 in prior chapter). In addition to the coupling of symbol concepts to sounds being produced it is also coupled to sounds being heard. These sounds are, of course words. Below I describe what is called the phonological loop which is responsible for our ability to tie what we hear (e.g. our native language being spoken) with what we say (the motor programs we learn to emulate the sounds we hear) and the connecting concepts held in the prefrontal cortex that are effectively the seat of symbols (in language). We now know that this loop is so flexible that it can even be built using visual signs, which is sign language where arbitrary symbols of hand movements, for example, can be used to make the connections between words signed to words viewed.

Moreover, this exact same flexibility allows the loop to be built for signs affixed to a stable medium (as icons had been done before). These signs are writing and their input interpretation through reading. Both signing and reading/writing are skills that require more work to encode as concepts because the areas responsible for production and interpretation are the auditory and motor planning areas (see figure 4.17 below). The sensory area for signing and reading (and writing) is the visual cortex which is further back in the occipital lobe. It is a testament to the malleability of the neocortex that allows the language organizing areas to recruit other sensory and motor areas when the primary senses (audition) are compromised. The same thing is true for touch replacing vision when a blind person learns to read braille.

Symbols are concepts that have all the right features for language. They are themselves compact, since they are just abstract representations. This allows the prefrontal cortical areas in which they are encoded to handle millions of lexical elements. All higher order concepts are recursively interrelated, so symbols lend themselves to encoding sentences that are syntactically and semantically based on the internal system language, i.e. subject-verb-object (or actor-action-patient) relations that can be nested wherein the outer relation's participants can invoke inner relations, e.g. "The **man** (actor) [*who (actor) threw (action) the brick (patient)*] **broke** (action) the **window** (patient)." The actor in that main sentence invokes a phrase that has the same structure. Such sentences are easily represented by concept maps, which in turn can be translated to neural circuits in the neocortex. Thus, while system language is the mentalese that I claimed in chapter 3, its internal semantics and syntax are what guides the learning of public language as the relation of symbols representing systems interacting with their environments. Any set of lexical, syntactic, and semantic symbols can be learned. But they are going to, in the end, describe systems and what they do.

Neural Encoding of Concepts: From Sensory Features to Complex Models

Now that I have covered the overall notion of concept encoding as producing models, of which language is the main element of representing models, it is time to demonstrate the basic mechanisms by which concepts of all kinds are actually encoded in memory traces in neural tissues. Here I am mainly considering what goes on in the neocortex of the mammalian brain as represented in figure 4.9 above.

The representation of concepts is accomplished in a hierarchical fashion (see figure 4.13 below). That is small bits of representation, called features, when they form a consistent spatiotemporal pattern, generate a percept either as sensory driven or as mentally activated (the arrows in figure 4.9 above). The links between the feature (detectors) and the percept representation neuron cluster are learned in precisely the causal association manner given above. A small cluster of neurons might be activated when that set of features is activated and the cluster learns (as a unified group) that the particular set of features 'means' that percept. Pavlovian conditioning actually provides an example of the attachment of meaning to arbitrary causal associations. Pavlov's dogs learned to associate a bell with the availability of food in their near future (seconds later). The bell had no intrinsic meaning to the dogs. But it came to have meaning when paired in this manner with food, which does have meaning. The bell caused the dogs to salivate as if food were present. We can readily model this association of arbitrary events/patterns with meaningful stimuli or previously learned meaningful concepts. Indeed I suspect that this is at the heart of what Damasio (1994) called 'somatic marking' and I have referred to as semantic tagging.

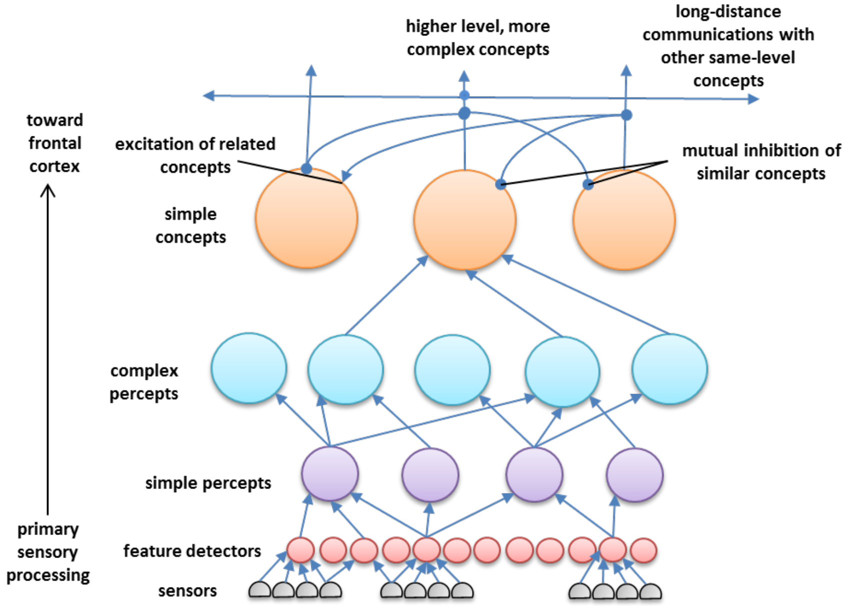


Fig. 4.13. A hierarchical representation of the world in features, percepts, and concepts is built up in layers across the cortex. This is a "side view" of the same mapping as in figure 4.9 above.

In the figure 4.13, above, sensory inputs come in from below to activate feature recognizers. Those features of the world that are present in the sensory field (e.g. in the visual field) are activated and then they, in turn, activate the perceptual field above[198]. Percept clusters

---

[198] Sensors are generally arrayed in a planar map. For example the retina is an array of rod and cone light detectors backed by local circuits that detect movement and a few other features. These feature detectors as well as those in the primary sensory cortices are laid out in a map that contains local 'fields'. Each field represents a topographical region of the sensory array and contains the same feature detectors in every field. Thus the feature being detected can be detected in multiple fields across the map. This is how the brain locates the position of clusters

(represented by a single circle but not to be taken as a single neuron) are activated when the set of features maps to that cluster. The connecting lines in the figure are actually complex channels that provide two-way activation, i.e. from above or from below (see Fig. 4.14 below). There are numerous adjunct neurons in these channels that prevent run-away activation, but are not shown here. Also note that some features are shared between several or even many percepts. The lines shown are channels subsequent to any learning that has taken place to form the mappings. Only a few maps are shown in the figure.

I should make a quick note on memories and their location in the brain. In the above diagram and the last paragraph I indicate that the connections between clusters are recurrent (also referred to as reentrant). That is, the lower clusters can be activated from higher clusters and vice versa. During sensory stimulus the flow of information is from bottom, more 'primitive' clusters to higher more integrated ones (features to percepts to concepts, or primary sensory cortices to associative cortices further forward in the cortex). However, during thinking or imagery recall the flow can go in the other direction, from higher level concepts to lower level features. We now know from imaging studies (e.g. fMRI) that memories are not formed from simple clusters at one point in the cortex. Rather, the recall of memories excites the same sensory areas of the brain involved in perceptual tasks. Perceptual memories are learned from repeated or reinforced bottom up activations and perceptual recognition proceed from bottom up mappings. Once learned, however, when a higher level concept is activated during cognition, it can send downward a wave of (presumably milder) excitation that recursively spreads over the mapping from bottom upward. So memories are found to be diffuse across the sensory and associative cortices. Higher level concepts can be constructed from the reuse of lower level concepts, and those from yet lower level percepts. Thus the brain does not have to construct memories from all features recorded anew with each new experience. The fur on your dog’s body is similar to the fur on other mammals and the details of features and constructions need only be derived inductively from a few instances. Once done they can be used in all perceptions of fur. This is very similar to how certain generally useful computational algorithms can be developed once, encoded in a generally accessible form (e.g. a shared library function in Unix) and used by any higher-order program from anywhere in the computer’s memory. It is not necessary to have unique copies of the same algorithm available to every single program that is running[199].

---

of features that are being stimulated simultaneously, or in other words, the position in the whole sensory field of the objects of interest.

[199] For those interested in the more technical details the Wikipedia article gives a reasonable introduction. See Wikipedia: http://en.wikipedia.org/wiki/Library_%28computing%29#Shared_libraries

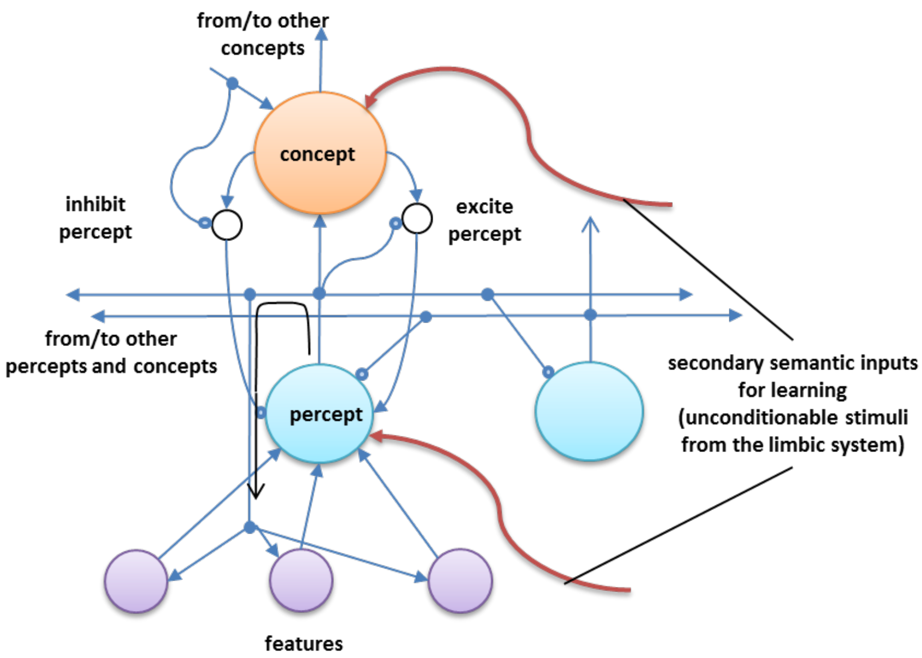


Fig. 4.14. Some greater detail than provided in Fig. 4.13 shows that the excitation of higher level clusters (e.g. concepts) can feed back to lower level clusters that comprise the components of patterns that activate the cluster in bottom-up (perceptual) processing. However, the downward activation can be driven from yet higher level clusters. For example, a meta-concept, or other associated concepts (not shown) can activate a concept that then sends recurrent signals down to the percepts that are part of the bottom-up activators of that concept. In turn, percepts can activate the feature set that would activate them from the bottom-up. Thus perceptual experiences activate concepts from the bottom up, while higher or associated concepts activate lower level percepts/features from above. This mechanism enables single cell clusters to encode complex objects and relations by reusing the lower level clusters. This recurrent wiring helps to explain why areas in the perceptual cortex are activated even when someone is just thinking about an object rather than perceiving it. The two smaller clusters represent associated control neurons that either drive downward activation or inhibit positive feedback from upward activation. The open circular termini are inhibitory synapses. The solid dots on lines (axons) are used to denote connection of multiple lines. Also note the red inputs from the limbic system. These are affective-based unconditionable stimuli to the cortical neurons that, in essence, tell the cells when there is something important that needs to be encoded as per the description of synaptic learning above.

In a recursive fashion, sets of percepts that have been activated from their various feature maps activate concepts to which they are mapped. Concepts are complex versions of percepts with other inputs considered, for example, other concepts. The channel arrows above the concept level represent communications between various concepts. These can be excitatory, as when concepts have been learned in association. Or they can be inhibitory, as when concepts are mutually exclusive or clash.

Concepts that intercommunicate can form meta-concepts. For example, a 'dog' is a meta-concept. It groups all animals having certain perceptual characteristics (aggregates of features) in common into a category of things. A specific dog, say your pet Fido, has some unique instances of those characteristics which you recognize. Yet the uniqueness of your pet does not preclude your understanding that it is, after all, a dog in the more generic sense. This ability to categorize and generalize while maintaining specificity of instances is probably a general mammalian (possibly some avian cousins can do it as well) mentation feature.

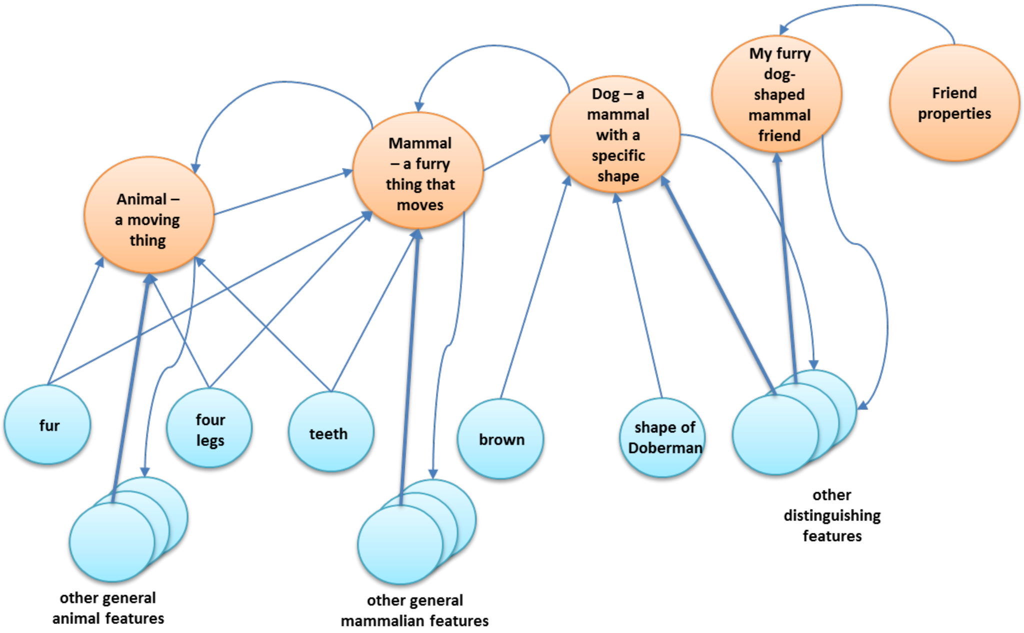


Fig. 4.15. Features and concepts form super-clusters or meta-concepts. Feedback signals help reinforce the relations between such clusters over time using the leaning mechanism covered above. Thinking about my dog (a hypothetical since I don't currently own a dog) activates all of the associated concepts, e.g. mammal, animal, etc., as well as numerous features that contribute to the properties of my friend dog. If I were to physically see (perceive) the dog, these features would be activated driving activation up the chain to my concept of him cluster. These kinds of associations allow us to answer such questions as: Is my dog a mammal? Note that the concepts referred to here are images of things, not names of things (see below for discussion of symbols).

All of these structures are composed of myriad neural clusters with many neurons participating in forming dynamic cluster activations when a concept is activated in the mind (even if subconsciously). Neurons and new sub-clusters can join and leave these structures over time as learning takes place. New associations can be made at all levels and old associations can fade if not reinforced. Moreover, old associations can be inhibited in the case where new learned associations provide contradictory or dampening weight to the various activations. The structure of neural representation is in constant flux as new associations are learned. Some are so often reinforced with new experiences that they become essentially permanent in long-term encoding (changes in synaptic morphology suggest that some connections develop long-term stability).

As life goes on we form larger scale conceptual networks (networks of networks of networks...) to represent more abstract concepts like cities and corporations. We don't really have a good idea about the capacity of the brain to form these fractal structures, that is, what is the largest scale. But we do know that better brains, those more intelligent brains, can form more complex structures at more abstract levels than more common ones. This is part of general intelligence.

Creativity plays into this in providing ways in which brains can form novel, if temporary, meta-concepts at abstract levels and explore the applicability of these structures.

Constructing Concepts and Percepts in Specific Brain Regions

Learning associative linkages and forming ever more elaborate networks of sub-networks appears to go on throughout the entire cerebral cortex. However, the functions of different regions put this capacity to different uses. The basic concept formation process that I have just described appears to take place primarily in the forward (anterior) portion of the parietal lobes and the temporal lobes — the integrative processing parts of the cerebral cortex. These regions are responsible for integrating sensory features to form percepts and at least the first layers of concepts. It isn't completely clear where more abstract concepts are formed, but a very likely candidate is regions within the frontal lobes, in particular posterior prefrontal cortex.

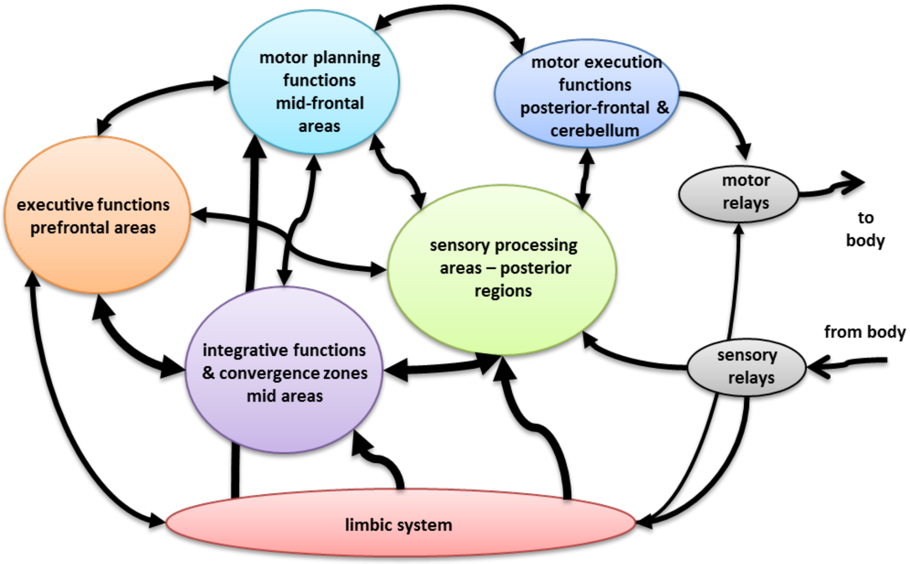


Fig. 4.16. Brain functions are distributed in a rough conceptual map of information flows. See explanation in text. This is a modular view of the brain similar to the architectural diagram in figure 4.10 above.

Superimposed on this conceptual structure formation is the sensory, thinking, motor activation processing that constitutes the major activity of the cortex (as in figure 4.9). Sensory information comes in through more posterior structures. Primary visual processing, for example, is handled in the occipital lobes. In general the more posterior portions of the lobes (except for frontal) process sensory input patterns, building percepts. The more central regions (except for occipital and frontal) integrate multi-sensory percepts forming or activating early concepts. These, in turn are the stuff of scene recognition and understanding — determining what is immediately in one's environment from sight, sound, touch, smell, etc. All of this is forwarded to the frontal lobes where decisions about what to do are computed. The inputs include the current situation in the environment along with the current state of the self and already present drives, motives, and relevant memories. These decisions then loop back to the posterior frontal areas, sometimes called the pre-motor cortex, where the decisions activate motor intentions that eventually culminate in some kind of action, even if only to generate new thoughts without overt activity.

Knowledge, concepts and their relations are encoded in neural clusters that start at the very lowest feature level in the primary sensory cortices, are integrated into perceptual structures in the secondary and early integrative cortices, and are further integrated in “abstract” concept clusters in the frontal and prefrontal lobes. Appropriate concepts are activated (excited) from either sensory (bottom-up) or thinking (top-down) or both. Those entities that are most strongly activated in the most abstract concept regions are said to be in “working memory” where they are the subjects of conscious attention. This brings us back to the nature of consciousness presented in the last chapter. What we have been reviewing is the neural and brain architectural bases for consciousness, a phenomenal experience that all healthy human beings experience. And that brings us to the question of sapience itself. The prefrontal cortex is the seat of “executive functions” in mammals. The expanded prefrontal cortex in primates and especially in apes provides us with explanations for the flexible behaviors and learning capabilities of these animals. But it is the tremendous expansion of the prefrontal cortex and one particular patch of tissue that gave rise to what I have been calling sapience. I now turn my focus to this part of the brain.

Symbols – A New Kind of Concept

Up to this point I have been describing essentially the formation and recall of ‘images’ that take place in mammalian (and possibly avian) brains. Concepts are little more than associations between percepts that are directly manipulatable in the cortex. For example, when we dream, we “see” the images of people, places, and things. They are behaving, perhaps incongruently with strict rules of reality, but identifiably. The situations and behaviors are, in some sense, acceptable to our consciousness. We might also hear people speak, but the main point is that our brains can reconstruct complex images that act out play-like scenarios from images stored in our memories.

It is a good bet that dogs and cats dream. Even rats seem to go through REM phases in their sleep time suggesting they too have dreams. This is because the major contents of dreams are the images (sights, sounds, and feelings). Indeed we suspect that our pets and other mammals think using these images. For example an ape could conjure the image of a particular kind of fruit they know to grow on a tree not too far away as the thought of desiring said fruit. The image of the thing is the only representation of the thing that the animal needs to think about and motivate action. The action itself is an image of body motion – thinking about walking over to the tree.

Human beings have evolved a new trick with the advent of a sapient brain. They have an area of frontal cortex (and most likely coordinated by a module in prefrontal cortex) that encodes a new kind of concept. It uses all the same mechanisms that we have seen here for doing the encoding but the nature of this kind of encoding is extremely compacted. This new kind of concept is a *symbol*. Specifically the symbol is a quasi-unique concept that involves a restricted set of temporally sequenced auditory features – a sound that we call a word. Over the course of child development, especially in the first two years of life, this area of the brain is busy creating image-like maps of these specific sounds, motivated by extremely strong limbic inputs that tell the encoding clusters to pay careful attention to certain sound patterns because the older members of the family, and in particular the mother, are routinely making these patterns with their voices. The child is primed to learn these patterns and, around age two, starts trying to imitate them. In this special brain region, there is the beginning of an associative mapping from auditory patterns to motor patterns as the child learns to produce the same sounds in the same sequences.

But that isn't the entire story. The sounds (words) themselves have meaning. So the association task is to couple the sound concepts with the image concepts forming in the new brain. Thus we find a three-way coupling of images of people, places, and things being represented by concepts of sounds that can either come through auditory channels or be produced by vocalization (motor programs). Any one of these three major concepts can activate the others in the exact same way I described above for how concepts can activate lower level percepts and other related concepts. There is nothing particularly new in terms of mechanics of association, only that this particular kind of association is very specific in forming the triplet relation of sound, meaning (image), and vocalization. We have the beginning of an abstract language.

Symbols are abstract concepts in the same way that categories are abstractions, as covered above. What has changed is now we have words to associate with each of the kinds of concepts, for example as shown in figure 4.15. For each of the pinkish ovals we can create new ovals in a different region of frontal/prefrontal cortex where these new concepts are encoded. They are the names of the things, e.g. "animal", "mammal", dog, my dog "Max."

Not only things have names (nouns). We also form concepts of actions that we also attach names to (verbs). Similarly we form concepts of relations such as "thing A is on top of thing B." They too get names.

The advent of symbol encoding provided an extraordinary benefit to human beings. While in the waking state it turns out to be extremely effortful to manipulate images in the same way we do in dreams. In part this is probably because in the waking state we are distracted by the activities and things around us taking so much processing power. When we daydream we effectively tune out the activities around us, go into a quasi-dream mode and can envision people, places, and things to some degree. To think quickly and well in the waking state, to conjure thoughts about things or scenarios that might be, and as I will argue below, to communicate states of affairs to others our brains need a more compact representation that can be manipulated with much less effort. A symbol concept is just the thing. Language, including or especially the language you silently speak to yourself, allows rapid manipulation of ideas. It allows relatively efficient communications between individuals.

Humans learned one more trick that extended the use of symbols. They invented signs to be yet another category of symbol, a visual and very terse one, which could be engraved or drawn on a suitable medium. *Homo sapiens* had already learned how to draw images from memory of animals and other humans in ritualistic poses as seen in numerous cave paintings dating back to 40,000 years ago. There are even older carvings that suggest humans were transferring images to external media, mostly rock. This is not surprising given that humans before sapiens had learned to carve (or chip) rocks to shapes that they could use for hunting and butchering. It is possible that early *Homo sapiens* were carving wooden objects to represent things in the world but such articles are not normally preserved so we might not ever know for sure. What we do know is that by about 10,000 years ago, probably with the spur of developing agricultural trade, people were making marks on clay tablets and into clay urns that signified amounts of various products. These markings appear to have been generally agreed upon as symbols meaning various kinds of products and counts of units of those products (an accounting system!)

Thus humans appropriated the hearing and speaking system based on word symbols to the reading and writing system based on visual, but very compact, symbols.

Words alone are not language. They are the lexicon that must be tied together with syntax to produce higher order meaning (semantics). In the last chapter I argued for the idea that the language of systems, derived from the properties of systemness, is the basic mentalese that provides for the manipulation of thoughts, and now we can see this means concepts. With the advent of associated name concepts or symbols there would seem to be a natural pathway for using that mentalese as a substrate for syntax. That is, the description of what systems do, what they do to each other, and how they relate to one another in a meta-system does, in fact, constitute a mental model for public language. The use of symbols naming things, actions, and relations would seem to be a natural follow on to the use of system mentalese by the brain. However, the brain consists of thousands of processing modules talking mentalese to thousands of other modules concurrently. The channel of communications and the limits of working memory restrict the use of a public language to a serial mode of transmission. You can only say one word at a time so some additional mechanism is needed to handle that.

There are in the world some several thousand languages, many with unique properties (uses of sounds like clicks or intonations to change meanings of words, etc.) So it is not surprising that there appear to be many different syntactical forms in use. This is a result of needing to sequence words and to nest phrases to approximate the more natural concurrent ideas that were the basis of generating sentences in a language. So long as the specific word ordering does not violate the basic structures of system mentalese, any surface syntax will perform its task. That syntax (i.e. word order and nesting rules) is a culturally derived feature. And surface syntaxes can morph over time; that is the language can evolve in many dimensions.

Below I will revisit language in terms of the way Brodmann area 10 has come to organize other brain regions for sapient purposes.

The Seat of Sapience: The Prefrontal Cortex and Brodmann Area 10

The prefrontal cortex is recognized as the seat of executive control that gives rise to consciousness (chapter 3). I have suggested (in the section above, The Neocortical Brain) that Brodmann area 10 is something like an executive-executive that gives rise to the higher order consciousness experienced by human beings. Figure 3.3 showed the “reflection” map that I asserted gives rise to what I referred to as $2^{½}$ order consciousness. This map is, I believe, the result of the expansion of BA10 and that which gave rise to the suite of human capacities that make us unique among animals - sapience.

Among other effects of the operations of BA10 is the induction of a more complex and effective communications capability that allows humans to share their ideas and thoughts, the same ones produced by the actions of BA10. The brain’s ability to process speech sounds and to generate speech acts is tied to the way in which BA10 interacts with the various preexisting areas of brains in lower primates. These areas were already developed for precursor functions. The advent of BA10’s executive-executive administration (the basis for strategic thinking, etc.) along with the increases in group interactions, empathy (from moral sentiments), and the need to cooperate with others in order to support strategic enterprises, is what helped shape specific modules in the brain for language capabilities.

Human language is unlike anything in the rest of the animal kingdom. Characterized as recursively generative, language allows humans to create an infinite number of sound combinations to which they attach semantic purpose (words represent things, actions, situations, etc.) They can recombine those combinations in higher-order combinations (sentences), which can, in turn, embed smaller units of sentence-like structures (phrases)[200]. Sentences are combined into larger scale units that attend to situational context. Language is a necessary outcome of breaking through to the next level of consciousness, of becoming sapient. This is part of the ½ aspect of human consciousness at level 2!

Language Production and Understanding

The posterior regions of the frontal lobes are designated as the premotor cortex. This is where intentional behaviors and movement planning is coordinated. When you reach for a glass of wine, this region organizes the sequence of movements and sets in motion the motor control programs (possibly 'run' in the cerebellum) that perform the behavior (recall figure 4.9). The anterior portion of the frontal lobes is called the prefrontal cortex and has been designated as organizing the 'executive' functions[201]. These are the functions associated with higher cognition and consciousness. This is the region of the brain where plans are formulated, memory managed, and mentalese is translated into interior narrative in the language of speech. When you 'listen' to

---

[200] Suddendorf (2013), see especially chapter 4.

[201] Goldberg, (2001); Goldberg & Bougkov, (2007)

your interior monologue (thinking to yourself) your prefrontal cortex is driving the formulated thoughts from working memory into the speech area of the (usually left hemisphere) frontal lobe — Broca's area — where the preformation of voiced sounds is initiated and syntactical structures (concepts) are formed (figure 4.17 below). The posterior portion of Broca's area lies in or near the temporal lobe where integration of the sensory inputs occurs. Thus a posterior portion of that lobe in conjunction with the parietal lobe feeds the auditory area of the brain — Wernicke's area — where the reverberations of what would otherwise be sounds (voice) are 'heard' in the head. The brain forms this phonological loop because those reverberations are then subsequently interpreted and end up back in the frontal lobe. The prefrontal cortex, particularly the reflective map that was introduced in chapter 3, is having an on-going conversation with itself in the same language as the individual uses to communicate with other individuals. But this conversation is merely the tip of the proverbial cognitive iceberg. The underlying, that is sub- and pre-conscious thinking activity, not summarized in the speaking language, comprises a much greater volume of conceptual organization. It involves the generation of temporary concept hierarchies and sequences (some novel) in working memory, and then the filtering and reorganizing of those temporary thoughts guided by the mental models of the world already established[202]. Presumably much more of this pre-conscious shuffling and construction takes place than ever surfaces to the level of internal voice generation.

It is likely that the prefrontal cortex, which is greatly expanded in humans relative to other apes, provides the main organizing circuitry for building and adjusting our models of how the world works. Once called the 'silent' lobes, because their functions were not obvious, the frontal lobes are now known to be involved in taking in the situation in the world from our sensory integration cortices, using our mental models to interpret the situation as well as integrate our inner drives and goals, and formulating behavioral responses based on anticipated results, or a change in the situation. It is the prefrontal region which is primarily responsible for controlling this activity. Then the output from the behavioral plan is pushed back into the premotor regions for action to be initiated. This, at least, accounts for voluntary behavior. As discussed below, involuntary, or reflexive behaviors have a different origin and triggers.

---

[202] See Daniel Dennett’s theory of multiple drafts in Dennett (1991).

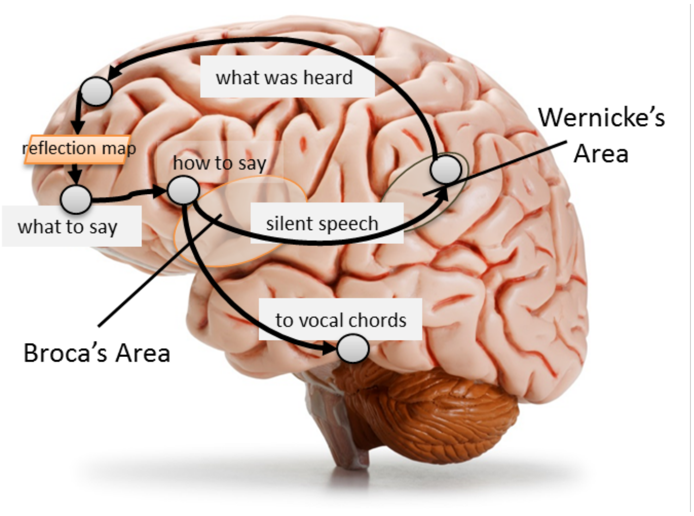


Fig. 4.17. The internal conversation loop (phonological loop) involves the listening (hearing language processing) in Wernicke's area receiving echoes from the speech production area of the frontal lobe, Broca's area. The same channels from Broca's area, along with control signals from the cerebellum, go to the vocal chords, tongue, diaphragm and other voicing muscles, but are presumably blocked when you are thinking to yourself. When you are talking to yourself you probably get both actual sound inputs along with internal (non-voiced) inputs in Wernicke's area[203].

The Uniqueness of Brodmann Area 10

The cortex has been mapped into smaller regions based on cytoarchitectonic (organization of cellular structures and cell type distributions), and to some degree on functional, considerations[204]. Many executive functions have been derived from psychological testing and have subsequently been correlated with the activities in specific areas in the prefrontal cortex. One area that has remained tantalizingly elusive is BA10, the most frontal polar area of the brain (see Fig. 4.18). It is also the case that this area is the most recent to expand in extent, and may have been most recently modified, cytoarchitectonically, in the evolution of humans. The most recent evidence from cognitive neuroscience regarding the activities in this area strongly suggests that it represents the highest level of control of all other areas of the brain, what I referred to above as 'executive-executive administration'.

---

[203] Brain image courtesy C^{3}NL The Computational, Cognitive & Clinical Neuroimaging Laboratory, Imperial College http://www.c3nl.com/research-traumatic-brain-injury/

[204] See the Wikipedia article on Brodmann areas: http://en.wikipedia.org/wiki/Brodmann_area

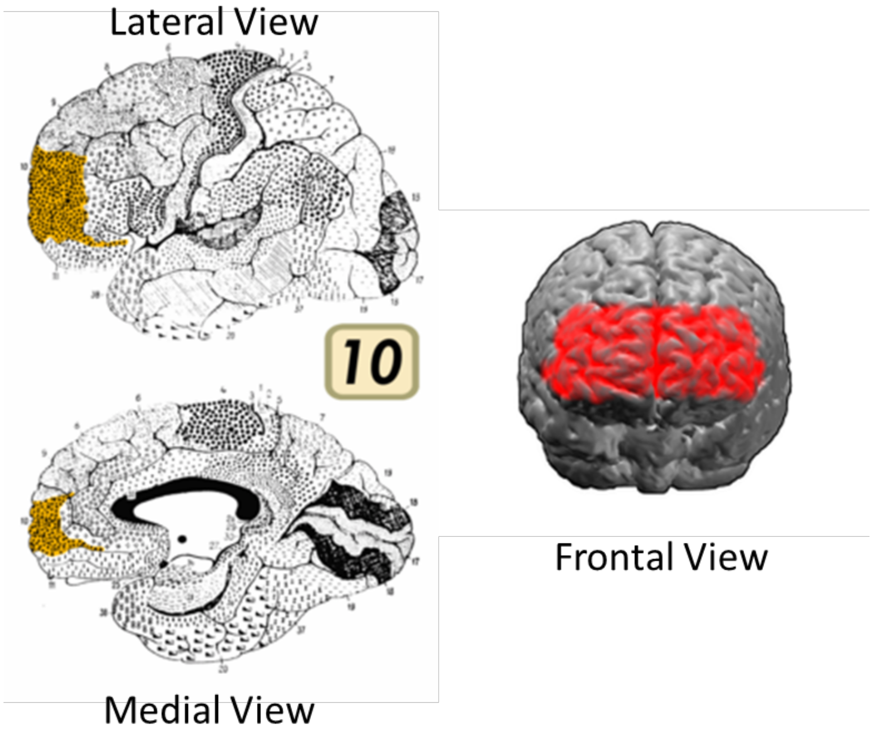


Fig. 4.18. Brodmann area 10 (BA10, colored regions) is the foremost region of the prefrontal cortex (frontal-polar). It is strongly implicated in high order judgment and moral reasoning executive tasks. It appears to be the final convergence zone for all other prefrontal areas as well as many other cortical and limbic brain areas. [Source: Wikipedia.[205]]

The prefrontal cortex is re-entrantly wired (either directly or indirectly in a few steps) to every other part of the brain, including the brain stem nuclei where the most primitive (operational level) controls are located. The prefrontal lobe areas collect information from all areas of the brain and provide recurrent afferents to those areas. Much of the information is summary, i.e. conceptual rather than perceptual. Even so, because of the way information flows from feature detection, through perceptual levels, to conceptual levels, the prefrontal cortex, and hence BA10 is informed of everything that is going on in the organism and the sensory fields outside of the body. Again, most of the information is subliminal or subconscious, not necessarily accessible to conscious attention. Nevertheless, the high-level models executed by the prefrontal cortex and BA10 receive everything that is relevant to management of thought. Thus to conclude that the prefrontal cortex is the seat of strategic and higher order tactical control is not unwarranted. Hence the strong suspicion that this patch of tissue is responsible for the highest level of strategic control of the organism is not without merit.

I should point out that I am not saying that BA10 is solely responsible for sapience, in the sense that all of the processing I have been talking about (conscious and subconscious) goes on just in BA10. Rather, I think of BA10 as having executive coordination control over all of the other

---

[205] See: http://en.wikipedia.org/wiki/Brodmann_area_10

parts of the prefrontal cortex (executive-executive) and that these, in turn, are involved in controlling all of the processing activities that collectively produce an integrated sapience function. BA10 is like the board of regents of a university. It is responsible for the long-term fate of the whole organization (of knowledge acquisition). The other areas of the prefrontal cortex are the management team, from the president down to tactical managers like the deans of colleges. This vision of the functional aspects of prefrontal cortex architecture is in keeping with the hierarchical representation scheme presented above[206]. In a real sense, concepts encoded into BA10 are the ultimate meta-concepts that tie everything together. The more competent this region is, the more comprehensive these meta-concepts will be. Think of this meta-concept region as a person's ultimate model of the world. The president of a university (or CEO of a corporation) has to be focused on near-term activities that support the mission of the university (teaching and research or quarterly profits), while the board of regents (board of directors) have to be concerned with the long-term health and success of the organization in a forever changing world.

All of the other brain regions encode and process their own kind of knowledge and would continue to do so if BA10 were eliminated (as has happened in certain brain lesion accidents[207]). Indeed we have reason to believe that elimination or reduction in effectiveness of BA10 does not seem to diminish intelligence as we normally think of it. The university can continue to operate quite well without the board of regents, at least for a while or until the environment changes radically. But the organization will have lost its ability to plan for the long-term because it lost its knowledge of the larger world outside. Or, rather, it will have lost its access to knowledge that might very well be found in independent departments, but cannot be brought to bear on long-term thinking for the whole institution.

My thesis, specifically, is that the expansion and differentiation of BA10 in *Homo sapiens* greatly increased the level of strategic perspective which, in turn, increased the existing moral sentiments and systems perspective in support of high-order judgments and intuitions. It promoted greater functionality in various other parts of the brain associated with memory mechanisms and communications. I will reopen that discussion in chapter 5 on the evolution of sapience as resulting from this expansion and why it occurred.

Ongoing substantial research on the cognitive functions of the prefrontal cortex is producing elegant results regarding the nature of consciousness and elucidating the 'I' in phrases like: "I think, therefore I am." My hope is that neuroscientists will see value in a framework for thinking about human psychology in terms of sapience as the epitome of strategic (executive) control and

---

[206] Also see the description of the hierarchical cybernetic model as it pertains to management theory.

[207] For example see Damasio (1994) for a description of Phineas Gage's frontal lobe damage and changes to his personality as a result. Gage retained most of his cognitive abilities. Also see Wikipedia: http://en.wikipedia.org/wiki/Phineas_Gage

its role in integrating moral sentiment with systems-oriented perceptual/conceptual model construction. Below I argue for how this might be accomplished.

Affect and Moral Sentiments

There has been a fair amount of work in neuroscience regarding the brain regions involved in processing emotions, innate drives, and moral judgments. In broad brush strokes, the basic emotions appear to involve primary or triggering processing in the deep limbic areas. Fear, for example, has been studied extensively by neurologist Joseph LeDoux (1996) and has been found to be triggered by areas in the amygdala, a nuclear module near the center of the brain. Other triggering areas in the limbic system are known. All of these seem to trigger automatic body responses that are thought to have been evolutionarily helpful in preserving our early vertebrate distant cousins. The general responses seemed to have followed the “shoot first and ask questions later” kind of philosophy. Except for vertebrates below the reptiles there probably were never any questions asked. These were instinctive, hardwired, responses to perceptual events that had, over evolutionary time, proven to be behaviorally sound.

Emotions, according to several neurobiologists, are not the kinds of feelings that, say, fish experience. They are more nuanced. It takes a cortex that can do additional processing and uses the same sensory information to produce more evaluative responses. But that cortical system also provides a new kind of map that recognizes the images of these triggering primitive brain centers. It is possible that these images are what instigate the additional processing. In second-order conscious animals, this is the experience of emotional response that can possibly moderate any innate reactive response. In the more complex worlds that more complex animals lived in it does not make sense to react to every shadow as a threat or every green thing as if it were food. More discrimination is needed. In humans the prefrontal cortex gets into the act of evaluating what the body is feeling. The experience of fear is considered more carefully because there can be many false positives that would produce inappropriate responses. And human emotions are more subtle still. Instead of outright fear a person may feel a sense of dread or just uneasiness in some situations that could anticipate danger.

The limbic system[208] is the more central region of the brain, often called the primitive mammalian (paleomammalian) brain. It preceded the neocortex evolutionarily and overlies the most primitive parts of the sub-cortical brain (or “reptilian” brain). In short the limbic system is responsible for a wide array of functions that involve distribution of sensory inputs to the various sensory cortices (thalamus), early warning of environmental contingencies that have emotional content (amygdala), laying down long-term memory traces in the cortex (hippocampus), and numerous other logistic and tactical control functions. In particular, the limbic system is responsible for reporting affective state to the prefrontal cortex. In other words, we feel our

---

[208] The Wikipedia article at: http://en.wikipedia.org/wiki/Limbic_system has some issue flags as of this writing, but it provides a reasonable background description of the parts of the paleomammalian brain I am referring to here.

emotional state by virtue of the limbic monitors reporting that state to our consciousness. Note that our emotional state is usually determined by our environmental situation prior to our becoming aware of it. First we lose our temper and then we realize we are angry and need to assess what caused us to be so. The same is basically true for all the other primary emotions.

Emotions

The fundamental basis of all behavioral decisions is an animal's attraction toward or repulsion from an external stimulus. This is ancient in neural terms. The simplest animals, indeed the simplest bacteria, have a fundamental taxis (movement toward or away from) stimuli that are good for them or bad for them. Everything else is neutral.

Except that some of those other "things" may provide cues to anticipate the good and the bad. In which case, the creature evolves sensory mechanisms for perceiving those otherwise neutral cue stimuli (as demonstrated in Representing Causal Relations above).

We humans emote consciously but it all begins with our most primitive taxes. We are drawn to those situations that provide food and sex. We are averse to situations that threaten to prevent us from finding food and sex or threaten our existence. Everything we feel (consciously recognize as emotional propensity) ultimately is based on these fundamental biological needs (basic drives).

Starting with Darwin, psychologists and sociologists have recognized that we humans have a particular ability to recognize and react to emotional states that are brought on by our physiological reactions to environmental cues. Our purely animalistic emotions are inherited from our evolutionary predecessors. Our primary emotions, anger, disgust, fear, sense of wellbeing, sense of poor being, and surprise are reminiscent of responses we see in much more primitive vertebrates such as fish and reptiles. In mammals and especially humans these emotions are experienced at the conscious level as temper, revulsion, anxiety, happiness, sadness, and startle. We humans can elicit memories of these feelings and form impressions of how events in the world trigger such feelings. Ultimately, however these emotions have their roots in those most basic reactive systems of attraction and avoidance (repulsion).

The basic emotions are triggered in the lower limbic system, holdovers from the most primitive reactions to environmental situations. The conscious awareness of the emotions is the responsibility of the prefrontal cortex. The latter is alerted to analyze the details of the situation and try to make decisions that produce more appropriate actions, not just reactions. The latter might not be appropriate for the nuances of the actual situation. This is certainly true in far more complex environments where triggering conditions can be illusions.

In the human brain a particular region, the anterior cingulate cortex (the ACC already mentioned in the section on strategic planning above), a fairly ancient region that lies underneath the prefrontal cortex, contains unique cell types called von Economo neurons or spindle neurons.

These neurons appear to transfer signals from the limbic areas, such as the amygdala (which mediates aversion responses) to Brodmann area 10 (see below). These neurons seem especially adept at more rapid transmission of action potentials. Thus the prefrontal cortex, responsible for conscious awareness, is informed that the animal has experienced something that might be dangerous. The prefrontal cortex can then get to work analyzing the situation and make decisions about appropriate behavior. It may decide that that thing that at first looked like a snake is just a piece of discarded garden hose. Or if it was a snake, perhaps it is a kind known to the observer to not be dangerous and therefore need not require escape. The conscious brain dampens down the signals coming from the limbic brain's automatic responses. A great deal more research is needed to work out the mechanisms for the prefrontal cortex's ability to modulate or inhibit the purely primitive animal responses to perceived danger (and likely other of the basic emotions).

In humans combinations of primary emotions are possible. For example one can feel anger and disgust simultaneously. We can feel happy and surprised at the same time. The modulation of basic emotions by the prefrontal cortex is what makes the emotions seem graded rather than all-or-none. Some emotions, like anger, may be triggered and result in immediate and strong reactions, getting away from the control of the prefrontal cortex, at least for a while. I suspect that in more sapient brains, however, the capacity for the prefrontal cortex to regulate the expression of emotions to fit the actual situation is much greater, leading to a calmer demeanor and measured interactions with the world.

Social emotions: Examples are embarrassment, guilt, shame, jealousy, envy, empathy, and pride

On top of basic emotions human beings express and feel a range of social emotions related to their being eusocial creatures. These emotions are innate; they are inherited along with a number of innate behaviors. So they must be rooted in areas of the brain deeper than the general purpose circuits of the neocortex. The emotions are processed by the prefrontal cortex and experienced consciously, though at times the source of the cause of the emotional experience may not be immediately perceived. For example, a married man who unconsciously flirts with a pretty young woman may feel a twinge of guilt later.

The social emotions provide us with a secondary set of controls on our behavior in groups where members are not necessarily related. Examples of social emotions include guilt, revenge, dominance, and subordination. These play a role in regulating humans cooperating but also in human competition and violence[209].

---

[209] Sapolsky (2018) has produced a compendium of human behaviors and their regulation,

Higher-order Drives

The basis of positive moral sentiment seems to be the twin affective drives of altruism and empathy[210]. These are higher-order drives. Altruism[211], or self-sacrifice for the good of others, is thought to be the basis of behaviors in a large number of social species, particularly among social insects, birds and mammals. Its presence is difficult to explain through ordinary fitness selection arguments since the animal doing the sacrificing is thought to experience a reduction in the transmission of its genes to the next generation. Kin selection theory is generally invoked to explain how, through inclusive fitness (the family or kin overall are more fit) altruism is actually a favored evolutionary strategy[212].

Empathy is something more difficult to explain. First it appears to be a phenomenon observed only in humans and possibly in some other great apes[213]. Fundamentally it requires that a mind contains a model of the 'other' sufficient to pick up on emotional cues and to mirror those emotions to some extent in one's own model of the self (principles 9 and 10 covered in chapter 1). There are a few, currently unresolved, neurological hypotheses explaining empathy in terms of the actions of what are called "mirror" neurons. These are neurons that appear to fire actively when the subject observes another being performing an action that they themselves can do. Those same neurons fire when the subject performs that same action. This is enticing (especially since the phenomenon is observed in primates!) as a possible mechanism for empathy, but it is not by any means sufficient to explain what causes one person to, for example, feel sorry for someone they observe suffering, or feel joy when someone else is expressing joy.

It is not clear, at present, what basic layout in the limbic system gives rise to these drives or the resulting behavioral programs we witness. However, they seem to be undergirded by more primitive drives, indeed the most primitive drives, of *seeking* for resources and mates, and *avoidance* of dangers[214]. These two most primitive drives (found in all motile creatures including bacteria!) are augmented in more complex animals with additional drives such as rage, panic, and sex drive. In yet more advanced organisms care giving and play (in mammals and some birds) round out the set of drives[215].

Seeking behaviors are driven by associations of physical sensations with rewards. For example, the taste of a food generates a reward loop, with dopamine delivered to associator neurons to

---

[210] De Waal (2010 & 2014).

[211] A reasonably good review of the concept of altruism can be found in the Wikipedia article: http://en.wikipedia.org/wiki/Altruism.

[212] Kin selection theory has been fairly popular in "explaining" the phenomena of one animal (or person) giving up their reproductive investment in favor of those related. In recent years the theory has come under increasing criticism for attempting to explain all instances of cooperation and sacrifice, especially among human beings. See, for example, Wilson & Wilson (1994, 2008) and Wilson (2013).

[213] De Waal tackles this issue with respect to the apes, especially the Bonobos (2014). See also Suddendorf (2013) for explanations of how hard it is to experimentally confirm empathy in non-humans.

[214] This is the basis of my robotics work in Mobus, (1999).

[215] See McGovern, K. in Barrs and Gage, (2007), Chapter 13, pp 369-388.

'bless' the association. That is the follow-up with dopamine validates the association of the substance as a food so that it will be sought after in the future[216].

These basic drives underlie the biasing of perceptual and conceptual systems in the neocortex. We saw this in chapter 2, Relationships. We will see that the neocortex provides an important set of new facilities to match these basic drives to social behavior. But there seems to remain a puzzling middle piece to the story of how basic drives translate into social behaviors, especially in animals with simpler brains. The origin and mechanisms underlying eusocial insects, for example, cannot rely on mirror neurons as found in neocortical tissue. Pre-mammalian phyla contain many examples of social behavior, such as schools of fish, so sociality must have a deep mechanism in the primitive brain.

By the time we get to mammals we find many social species that do not depend on, say, instinctive behaviors per se, or division of labor to the extent seen in the eusocial insects. What we do see is altruism and mimicking behaviors emerge from the primitive cortex and later expanded in the neocortex. So we are not in a position to identify the brain structures directly involved in going from primitive drives to social behavior yet, although a link between the automatic reaction to bodily excreta and rotten food — disgust — has been suggested. It may be that soon we will have more insights into how primitive drives are linked to our more evolved emotional centers in the limbic system.

Sentiments and Morality

There has been much progress in connecting more advanced affective modules in the limbic system to the basis of moral reasoning in the neocortex. These 'more advanced' structures include the hippocampi and the amygdalae (from above) as well as the cingulate cortex (see below), especially with its strong connectivity with the frontal lobes[217]. A great deal is now known about the iterative processing through re-entrant circuits that takes place between the limbic centers and the frontal lobes and most especially the prefrontal cortex. It is clear that these two widely separated circuits (in evolutionary terms) are intimately cooperating to produce final behavior.

There is reason to believe that a ‘moral instinct’ in humans is similar to a language instinct[218], or the drive to acquire specific interpersonal capabilities. That is, just as all normal humans have the capacity and inclination to learn language at a young age, to learn the vocabulary and grammar of their native tongue with a minimum of explicit instruction, so too, all normal humans have a built-in moral instinct that becomes particularized to a given culture. All humans, by this theory, have a built-in sense that there are right and wrong social behaviors, to experience empathy with other minds, and to instinctively seek to cooperate with familiar conspecifics in social networks.

---

[216] Mobus (1999). This was the basis of how my MAVRIC robot learned to seek out specific cue stimuli. It wasn’t actually looking for ‘food’ as such, but rather it learned to associate various ‘clue’ stimuli with rewards (like good tastes) and thereafter seek those clues as cues.

[217] Damasio (1994); LeDoux (1996)

[218] Pinker (2007a)

Since it is fairly certain that the motivation and mechanisms for language acquisition is built into our brains, and if there is an analogous situation with respect to the acquisition of morals (e.g. there is a moral sentiment that moves us to acquire specific moral codes) then morality is essentially built in[219].

By built in, of course, I mean that genetic propensities exist which guide the early development of the brain to hard-wire these instinctive tendencies into the limbic areas and ensure communications with the appropriate cortical regions. The latter areas are where learning complex concepts allow one to learn the specific rules of a given society; what constitutes specific good and bad behaviors within the context of that society. The impetus to want to belong to the group tends to bias our behavioral decisions to what that group counts as good and avoid that which it counts as bad. Of course, sometimes it is tempting to do something that would otherwise be counted as bad (cheating) because, as biological creatures seeking self-gratification, it can be advantageous if the behavior can be pulled off without getting caught. Our brains have evolved an elaborate set of filters for self-inhibition (regulation) as well as to detect cheating in other members of the group.

The basis of our moral (and ethical) reasoning, processed in the neocortex, is grounded in evolved instinctive behaviors that allow us to form strong social bonds and functional structures. The essence of morality is how we tend to treat others within our group and those outside our group[220]. The adaptive capacity of the human neocortex is substantial, as evidenced by our ability to belong to different 'groups' and to expand our sense of group-ness beyond mere tribal levels (150 - 200 people in primitive tribes according to some estimates[221]) to include large institutions (e.g. our work or religious affiliations), states and nations, and perceived racial affinities[222]. Our brains have allowed us to behave socially with strangers with whom we perceive a common membership in some conceptual group[223]. Nevertheless there are limits. It is just as easy to perceive an exclusionary set of attributes that make the 'other(s)' seem inferior, even non-human, allowing for more aggressive tendencies to emerge. Such perceptions, if driven by emotional forces from the limbic system, readily lead to all kinds of horrors (as judged from the outside by uninvolved observers). This is very likely an evolutionary hold over from our Pleistocene ancestors, their tribal organizations (groups) and the effects of group selection that favors in-group altruism and out-group suspicion and hostility[224].

So far as sapience is concerned, it is my hypothesis that the prefrontal cortex, especially BA10, has responsibility for bringing to bear a large body of tacit knowledge about good and evil, right and wrong, costs and benefits and many other dichotomies, on the decisions that must be made in

---

[219] Also see Suddendorf (2013);

[220] Wright (1994)

[221] Dunbar's number ~ 150 stable cognitive relations between people. See Wikipedia: http://en.wikipedia.org/wiki/Dunbar%27s_number

[222] Berreby (2005)

[223] Seabright (2004)

[224] Sober & Wilson (1998)

maintaining the social order when problems emerge. That groups are composed of many highly variable individuals, each with a high sense of autonomy and a generally unique array of desires and motivations, as well as personality type inventories, ensures that problems (conflicts) will arise. To what degree can a strong sapient individual bring their wisdom to the judgment process, motivated by strong, positive, moral values may often mean the difference between success and failure of the group? I will explore this from an evolutionary perspective in the next chapter.

Primitive to Higher-order Judgment

Every time you make a decision (conscious or unconscious!), no matter how trivial or 'local' it may be, some portion of your brain is applying judgment while another part is applying affect to bias that decision. Most of us go through daily life making mostly trivial decisions. What to eat, what news story to read, what to wear, etc. are the stuff of daily life. Most of the time, these decisions are made subconsciously without thinking too much about it. Even when the number of choices is larger (do I want Mexican, Chinese, Thai, ...?) and we spend a little time actively engaged in analysis (what do I feel like eating?) we don't go at it with any kind of rigor. We just decide based on what we feel.

Low-level or primitive judgment refers to the evolutionarily early application of learned preferences to guide decision processing. Some brain studies suggest that the anterior cingulate cortex (ACC) plays a role in conflict resolution and mediating between the affect centers and the neocortex[225]. It is conceivable that this region of the primitive mammalian brain actually was responsible for early mammalian judgment processing. Now it is still involved in applying biases to lower-level decision making.

The ACC lies just beneath the frontal lobe neocortex and near the prefrontal cortex specifically. The latter is richly connected with the former, suggesting an integration of functions. The ACC has a rich concentration of specialized neurons called spindle cells (see the section) which are implicated in high-speed communications between the limbic system and the prefrontal cortex. But in addition the ACC is in communication with the parietal cortex as well as the frontal eye fields of the frontal cortex (implicated in eye movement control). It appears that the ACC is situated in such a way that all of the information needed to evaluate a person's situation and make moment-to-moment judgments is available. Simple judgments may include directing the eyes (and possibly the auditory system focus) in deciding what to look at next.

I suspect that the prefrontal cortex extends the judgment processing role of the ACC wherein the former has expanded the scope and complexity of learned tacit knowledge application to decision processing. In other words, the prefrontal cortex has become responsible for higher-order judgments. The lowest sorts of such judgments are those involving simple matters but requiring conscious analysis and consideration. Decisions like 'what paint color would look best in this room?' require more thought as well as awareness of emotional responses and thus, I

---

[225] See the Wikipedia article: http://en.wikipedia.org/wiki/Anterior_cingulate_cortex

suspect, are processed by the prefrontal cortex with 'help' from the judgment application (to decision processing) circuitry already available in the ACC.

Higher order judgment requires reflective examination of the 'problem', the factors involved, the beliefs and feelings of others (wicked problems are invariably social in nature), one's own feelings and desires, and, most importantly, how choices may play out in the future. Thus the prefrontal cortex is the orchestrator for bringing all of the intelligence, creativity, and affective resources to bear on our models of how the world works and what are the possible outcomes of different choices. Only some of this reflection need be done in conscious awareness. More likely a larger portion of judgment processing takes place in subconscious mind but sometime later comes into conscious awareness more or less fully formed. In the former case the decision is taken without conscious reflection. In the latter case, the ACC may be responsible for promoting comparative processing in the prefrontal areas to resolve differences.

This model has the neocortical regions of the prefrontal cortex essentially accreted onto the more primitive paleocortex and adding substantially more processing and representational power to handle substantially more complex problems.

I cannot leave this discussion without pointing out that while high-order judgments improve the efficacy of decisions in general, they are still subject to many biases. In chapter 3, under the heading, “Efficacious Models” I discussed several systemic biases that obtain in most people's judgments due to the fact that the brain relies heavily on heuristic processes (see next section) to come to quick (and sometimes dirty) conclusions. The quickness can be a benefit in most ordinary circumstances where speed is more valuable than precision. This would certainly apply to many kinds of survival decisions where it is often alright to error on the side of caution and react quickly. These heuristics, however, still dominate much of human judgment. Ordinary (that is average) sapience does not have the ability to force every decision to be made contemplatively (nor should it). But it also seems to still be weak when it comes to judgments about what judgments should be contemplated before affecting decisions. This is likely the domain of the $2^{1/2}$-order consciousness that I introduced in chapter 3 (figure 3.3). This ‘ultimate’ reflective map has the job of *judging judgments* and, if sufficiently developed, determines when heuristic-based biases are potentially damaging to good decisions. It would then direct judgment processing to kick into second-system (rational processing) mode to filter out the effects of biases to a greater degree. It is not likely that all biases can be filtered. And it is likely that biases still persist, even if at a reduced effectiveness level. But the judgments should be greatly improved. As an example, consider the effect that mood has on judgments. It is well known that mood (happy or sad) can have a bias effect on judgments[226]. But it is also known that by pointing out to someone the attribution of their current mood, which is not related to the target subject of the judgment, that many people can discount their moods and arrive at better judgments than they would have otherwise. In other words, it is possible to trigger more rational judgment processing in the

---

[226] Schwarz (2002)

particular case of mood biases. Hence we know the mental processing power is available. But this also shows us that it is not always the case that people self-monitor their judgment processing and recognize that their mood: 1) is not attributable to the current problem; 2) will influence judgments unless contemplation is initiated. This is evidence of the still relative weakness of sapience in most people where the standard cognition could be described as "just go with what you feel."

Mood-based biasing is just one of many forms that result from heuristic processing that attempts to speed the decision process along. It has been shown that mood-based biases can be mitigated to some extent with conscious attribution. But this may not be the case for all such biases. Indeed it has been shown that some forms of biases cannot be mitigated simply by calling attention to the possibility of biases entering the judgment process. What we are left with, for the present, is an understanding that heuristic processes leave the sapient (higher-order) judgment process vulnerable to mistakes. There is some evidence that this need not always be the case, but generally is the rule rather than the exception (except in exceptional cases, which we recognize as wisdom!) This inherent weakness in sapience will become a topic of discussion in chapter 5 where I discuss the evolution of sapience and its future.

Systems Intelligence

In the section above on "Neural Basis of Systems Representation and Models" I laid out a scheme for how networks of neurons might reasonably wire up (through learning) to form representations of things and processes, including a mechanism for representing causal relations. This was left as a basis for building percepts and concepts. The relationship to 'models' was not further developed. In this section I want to further explore how dynamic models can be built and 'run' in brain tissue since such models are the basis for understanding how sapience models systems in the world outside the head, and, as it turns out, inside the head as well.

Heuristic Programming in the Brain

In computing we build models of real-world systems by writing programs. A program is, at base, just a set of steps, like in a recipe, that collectively do three basic things: they change a memory variable's value using arithmetic operations, they perform input/output operations, and a few computer instructions cause a program to branch, that is take one of two different tracks of steps, based on the condition of a value resulting from an arithmetic operation (e.g. a value is positive or negative)[227]. Typically, computer programs are algorithmic. That means the steps are a) unambiguously defined; b) a finite sequence of steps; and c) produce a clear answer, usually within a reasonable amount of time. Algorithms are very definite and specific. They can be implemented in computers because the nature of logic elements is matched to the need for

---

[227] See especially chapter 8 in Mobus & Kalton for an extensive treatment of computation (and programming) as a more general process, not restricted to human-designed computers.

definiteness and specificity. Algorithms, and their implementation in computer programs, very clearly solve well-defined (mathematically and logically) problems.

There is another kind of programming that can be simulated in a computer but is a better description of what can go on in living neural tissue. And that is heuristic programming. Heuristics are generally described as 'rules-of-thumb' in that they are not guaranteed to produce a definite result but generally do. It turns out that vastly more kinds of real-world problems cannot be sufficiently specified such that some algorithm might be found that would solve the problem. In this case it may still be possible to design a program involving heuristics to find a possible, even probable, (sometimes approximate) solution to such an under specified problem. It turns out that motor (muscle) control is of this nature. Muscles are not like solid gears and machine motors. They are deformable under stress, yet resilient and pliant. They do not respond exactly to electrical stimulation, although a bundle of muscle cells will perform reasonably well as an averaged response.

As brains evolved, and animal behaviors required more and more flexible motor control in spite of the uncertainties in response for muscle fibers, the neural circuitry of the cerebellum developed several interesting capabilities[228]. Behaviors involve sequences of activation of different muscle groups, often balancing the responses of opponent groups (e.g. biceps vs. triceps) in order to generate coordinated movements, say of appendages. Thus motor control is essentially a program of different motor responses evolved to carry out a motion primitive (like lifting a leg). More complex programs, like walking, are built out of meta-sequences of these primitives. The higher-level control, or decisions about the goals of making motions, what behavior to elicit, etc. is accomplished in the motor region of the cerebral cortex (frontal lobes), but the execution control and feedback control is handled by linear circuits of neurons in the cerebellum. In other words, cerebellar circuits instantiate motor programs. More complex programs (like riding a bicycle) are actually learned and represented in the cerebral frontal cortex, and like the motor cortex, the cerebellar cortex is also capable of adaptation with repetition.

Heuristic programs can be built in neural tissue in a relatively straightforward manner. Figure 4.19, below, shows a highly simplified, but biologically plausible model of how neurons can sequence a previously learned set of steps as well as branch to an alternate step plan should circumstances dictate (through body feedback).

---

[228] See the Wikipedia article: http://en.wikipedia.org/wiki/Cerebellum for an overview.

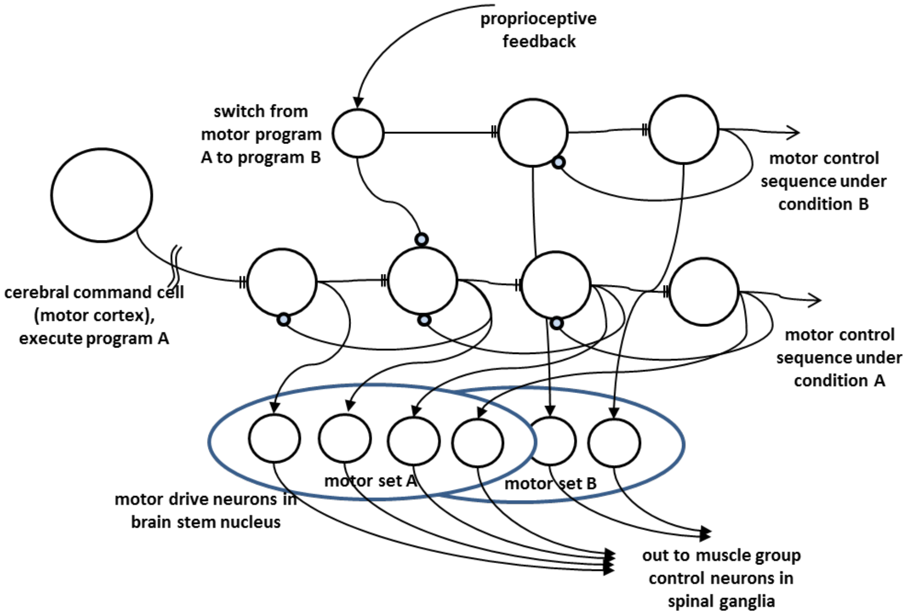


Fig. 4.19. This is a hypothetical, but plausible, circuit of neurons for implementing a heuristic (motor) program. See the description in the text.

Note that the program is initiated by a command cell in the cerebral cortex (frontal lobe motor region). The program is simply a sequentially activated set of neurons. In this simple version, each neuron activates the next one in line, but is subsequently inhibited by the neuron it just activated (negative feedback). In the event some body sensor (proprioceptive) determines a change in conditions that require a new sequence, a branching to a set of cells representing that different sequence can be initiated by inhibiting the old sequence (in step 2 in the figure) and starting an alternate sequence. All of these program step neurons activate actual motor drive neurons in the brain stem. Note that these sequences are learned (the double flat termini in keeping with figure 4.12. In figure 4.9 we saw the general layout for the cerebral cortex in which concepts were built from sensory inputs in a hierarchy of increasing abstraction. Then in the frontal cortex behavior responses were “planed” and actual motion was initiated in the posterior frontal lobes. These, in turn activated learned motor responses leading to the activation of the cerebral command cell in figure 4.19. The sequence of neurons that form the motor sequence control chains in the figure activate the clusters of cells that, in turn, activate the muscle groups themselves. The sequencing circuits are unique and are believed to be found in the cerebellum, which has received scant interest in the cognitive neuroscience literature until recently.

Thus the brain includes the capacity to build dynamic programs capable of the same sorts of processing we see in computer programs, but done through heuristic rules rather than strict logic.

Mental Simulation Models in Neural Circuits

I have discussed this aspect of motor programs because it turns out that the brain can use the same basic architecture to build mental models for simulating real-world systems. These are not motor programs per se; they do not involve directing muscles to activate. Rather they can be used to control sequencing of concept activation back in the cerebral cortex. Recent research has shown that the cerebellum is implicated not just in motor control and coordination, but also in some cognitive processes, including the generation of speech. It isn't hard to see that motor sequencing is needed for actual phonation, but it is thought now that the cerebellar circuits are involved in more complex structure formation, such as words and possibly even sentences or phrases. Other afferents from the same kind of circuits are now known to innervate regions of the cortex other than Broca's area or even other motor regions so speech production is not what these fibers are for. It is quite possible that branching program segments in the cerebellum are used to control the timing and sequence of activation of various concepts that represent real-world objects and actions. In other words, cerebellar programs may actually be at the heart of what we experience as ‘thinking’, or the sequencing of thoughts in natural (but learned) progressions, with branching capabilities when the context of thought is varied.

Once one can specify a method for producing branch-able sequences everything is in place to produce more complex programs that simulate real-world dynamics. In other words, the brain has the inherent capacity, between concept representation in the cerebral cortices and timed and conditional program structures in the cerebellum to construct and run models of the world! And running a model in fast forward is the essence of thinking about the future, of planning, and of feeding the judgment process with options.

Constructing Models — Guidance from the Prefrontal Cortex

What remains is to say how models of real-world systems are constructed in the first place. What guides the general intelligence system in deciding what to learn, how to organize it, etc.? These questions are yet to be answered in terms of evidence from neuroscience, unfortunately. If I were asked to guess, and generate hypotheses that might be tested, I would start with an assumption that the guidance is carried out by the prefrontal cortex, but using templates of systemness from deep in the limbic areas. The perception of systemness is evolutionarily ancient, I suspect. At the heart of native knowledge of how the world works must come from basic senses of boundaries of coherent things, dynamics, and cause and effect assignment. This is an area of interest I wanted to pursue with my robotic models, some day.

Clearly lower animals (at least birds and mammals) show an ability to recognize 'things' in their environments. They show a capacity to recognize process and dynamic relationships. These abilities suggest that there is something very basic about systems representation and model building. In what follows I provide a review of what we know about the brain that provides clues as to how it builds dynamic models of systems composed of subsystems and themselves composing meta-systems.

Object-Oriented Programming: A Close Analogy to System Programming in the Brain

Though computers and brains do not work in the same exact way with respect to logic or data storage (memory) or representation, there is one way in which a relatively new approach to designing software has captured an important aspect of how the brain encodes template 'thingness.' This is called object-oriented programming (OOP) and is functionally analogous to how brains represent objects/systems/entities, first as abstract 'things' and then as particular instances.

In OOP a programming language like Java or C++ defines a template "object" or the most abstract form of a unit of computation. This object has boundaries (in computer memory terms). It has internal basic attributes (internal variables). And it has basic behaviors that all and any derived object would have. For example it is generally convenient that every object in a program be able to send a textual representation of itself to an output device (screen). Thus in most of the OO languages the most abstract object (in Java called, not surprisingly, "Object") contains a behavior (called a method in Java) that does this. The behavior might be called "toString", meaning output a 'string' of characters that contain representations of the object's current state. In the most abstract such object this state is limited to just those variables which are possessed by all objects.

This abstract object is not very useful for programming. Instead the language allows for a way in which this object and its fundamental code can be 'inherited' by less abstract objects in which more particular internal state variables and more useful behaviors are defined. The less abstract objects are said to be 'derived' from the more abstract ones. For example, in graphics programming it is customary to start with the simplest and most fundamental graphic object, a *point*. Thus it is possible to define a point as a 'kind of' object in which one defines state variables like position as a set of coordinates (a tuple, x and y) that locate the point on the visual field (screen). Other state variables might include color and whether the point is 'on' or 'off.' Specific behaviors such as turnOn or turnOff are then defined (programmed) and added to the repertoire of behaviors already possessed by the point object.

The toString behavior would have to be modified so as to print out these new, added on, state variables, of course. But rather than starting from scratch, it is possible to simply add on the new code for the new variables and still use (reuse) the existing code.

The point is a fundamental object in a graphic system. A vector of points could be defined such that the result is a line. The starting point location and the ending point location in a Cartesian grid are sufficient to define a line. So these attributes are added to the definition of a point object to make a Line object. Again the attributes of the more abstract object are extended along with new behaviors (drawing a line as turning on all the points in the vector) to produce a new kind of object.

This kind of software architecture is actually emulating what the brain does in constructing concepts in the cerebral cortex. Objects defined in the software language can be used to derive more particular objects and resemble the categorization framework (e.g. animal → mammal → dog → Fido) which the brain uses to save on neuron resources as described above.

The brain doesn’t work the same way as an OOP, of course. But the logical construction of derived representations of concepts and the ability for the cerebellum/cerebrum programming process to build dynamic models of the world are strikingly similar. In Mobus & Kalton (2014, chapter 8) we go to lengths to point out how what the brain does is computation of this kind.

One More Trick the Brain Knows

I haven't said anything as yet about the fact that the two hemispheres of the brain are lateralized or functionally dual. This issue is terribly overplayed in popular psychology (left-brain/right-brain people!) but there are some obvious differences in functions performed on either side by mature brains. One of the more interesting findings is that the left hemisphere (or at least the frontal lobes and parts of both the parietal and temporal lobes) is the site of enduring patterns of processing. Most often noted is that ordinarily Wernicke's area and Broca's area work their speech processing magic in the left hemisphere. Other evidence suggests that other routine processing modules are instantiated in the left hemisphere cortex as well. This leads to questions about the popularly viewed 'heart' side of the brain — the right hemisphere. Goldberg (2001, 2006) has developed a very interesting model that suggests that the right cortex is largely involved in processing novelty or newly developed circuits — new models. It could be that the left hemisphere, in particular of the prefrontal and pre-motor areas of the frontal lobe, has the machinery in place to guide the construction of a model to be built in the right hemisphere where it can be 'tried out'. The model could also arise by copying circuit relations from an existing model (from the left hemisphere) into the right hemisphere and then guiding changes[229]. This would be essentially what we mean by *analogic* thinking. Once a model is constructed and

---

[229] Copying a neural circuit or network from one region of brain to another is based on the same learning mechanism in synapses as was discussed at the outset. Neural networks representing established concepts in the left hemisphere could be activated (possibly during dream sleep) while corresponding regions in the right hemisphere that have not been committed to long-term memory encoding are simultaneously activated by long-distance axons across the corpus callosum. Subsequently, using the causal encoding rules, the right hemisphere region just activated receives the secondary semantic signals (unconditionable stimuli) that cause the activations from the left side to become encoded in short or intermediate term memory traces. In this way the left side concept can be temporarily written into the right side. Modifications can be made to the new right side model by the prefrontal cortex selectively activating left side attributes and shortly after activating the semantic signals to the right side so that it appends those attributes to the newly constructed model in the right side. Thus the right hemisphere acts as a programmable memory where creative modifications can be applied without modifying the long-term memories in the left hemisphere. What would be required then is that: 1) if a new model proves to be more efficacious than the currently stored version in the left side, it must be copied back to the left side to reconstruct the older version; 2) the right hemisphere must prevent long-term memory traces from obtaining so that the neural circuits can be returned to scratchpad status waiting for the next new model to be built. All of this is, of course, highly speculative. But it should be a testable hypothesis.

'tested', perhaps validated by experience, it might be copied back into the left hemisphere for future use in routine thinking or as the basis for a new analogy.

This scheme requires a tremendous degree of plasticity in the wiring between neurons and cortical columns in the right hemisphere. If this is the case one test of the hypothesis would be to look for dynamic and possibly amorphous (that is, dense, but weakly activating) connectivity patterns in the right hemisphere. Indeed a great deal of work on working memory involving novel task learning implicates the right hemisphere frontal lobe. Barrs (2007) had developed a theory of working memory that is accessible to all relevant regions of the brain as a 'Global Workspace', though the idea here is related to consciousness and would not apply to subconscious processes of model building, strictly speaking[230]. Nevertheless, a general vision of the right hemisphere acting as a giant white board where images can be temporarily written and adjusted is appealing. The left hemisphere, frontal lobe acting as a controller, initiating the writing, guiding the adjusting, initiating testing, and finally encoding a permanent image of a dynamic model for later automatic use provides for a compelling model of how the brain can think new thoughts.

Thinking, whether it results in an inner voice, that is, in consciously registered thoughts, or goes on subconsciously, has an appealing connection with premotor processing. Unlike actual motor output (movement) the command signals direct a program of sequencing concepts and their modifications. What started out in mammals as a system for constructing sophisticated muscular choreographies spawned a system for choreographing conceptual dances — models of the world — that could then be tested on the stage of mentation. This version of ‘thinking’ is clearly related to the idea of thinking as a search through concept space via a ‘drunken sailor walk’ generated by a central pattern generator as described above in the section on Tactical Planning and shown in figure 4.10.

Executive Functions and Strategic Judgment

In my upcoming book on applying systems science to governance, I will elaborate on the application of hierarchical cybernetic theory to the establishment of social governing mechanisms that might usefully replace our current *hodgepodge* of mechanisms[231]. I liken this to the control architecture of the brain which is divided into three basic layers (the triune brain) of operational, coordination (tactical and logistic), and strategic management. Now we can see that the brain is indeed a hierarchical cybernetic system with the highest level devoted to strategic management. In the prior chapter, “The Components of Sapience Explained,” I identified the strategic perspective as a central component in sapience psychology. Now it seems plausible that

---

[230] Baars’ theory is the source of inspiration for Dehaene’s (2014) global workspace model mentioned above.

[231] By this I mean that most current governments are comprised of many different and often conflicting concepts and departments. For example the US government has numerous ‘agencies’ responsible for security issues and all too often these agencies either are redundant or dispute each other’s authority and jurisdiction.

we will find the seat of strategic thinking in the prefrontal cortex and, as I have argued, largely to be found in Brodmann area 10.

Strategic thinking requires that we build models of how the external environment works. This includes making decisions and judgments about what should be learned and how to go about learning. The models are used to simulate how the world will evolve into the future under different starting conditions, particularly with respect to actions that we might take in the present. The models themselves are organized concept clusters probably residing in the right hemisphere anterior parietal lobe. Their dynamics (i.e. running the models) is likely orchestrated by the premotor regions of the frontal lobes in conjunction with the choreography directions mediated in the cerebellum. The outputs from these models are analyzed by regions of the prefrontal cortex (other Brodmann areas conjoined with area 10) and supplied to BA10 for disposition and ultimately for making strategic (e.g. long-term) decisions. Such decisions might easily include life changing decisions such as what do I want to do when I grow up (career), or who do I want to marry, etc. In other words, the brain has evolved a capacity to plan for the future. And the vast majority of this thinking takes place in subconscious mind, only to surface as judgments and intuitions.

Strategic thinking involves the following characteristics:

- How is the world changing in the future?
- What goals should I set? Are they reasonable? What evidence do I have to suggest these are attainable and worthy?
- What are my assets (strengths) that I can bring to bear to meet these goals?
- What weakness or shortcomings do I have that would prevent me from meeting these goals?
- What assets should I attempt to obtain to help me meet these goals (sub-goals)?
- How will others react to my actions in attempting to reach these goals?
- Can I enlist others aid or support? Will I harm anyone else?

The list can be more extensive than this, but I think you get the idea. The main concern at the strategic level is with how the world is going to change, both intrinsically and in response to things that the individual does. The strategic thinker has to anticipate how the world will react over the long haul. How will other people react? How will the future state of the world and the individual be better off as a result of current decisions? These are all strategic questions that need attention. Most of the time, these questions are being processed subconsciously. They only rise to consciousness when the context of one's situation demands. In all cases they are connected to our internal models of the world, people, and ourselves (e.g. our own self-image and our beliefs about what others think of us).

Wisdom is often associated with those that think about the future, the long-term future, whether consciously or subconsciously. A wise person often thinks about or intuitively grasps what will be good for the future of the tribe. They think about what would be the best actions to take today to ensure a good outcome for the majority. They think strategically, not just for themselves, their own benefit, but for their group. Early hominins likely had the capacities listed above. They could think strategically for themselves. Modern humans have evolved a capacity to think strategically for many others; replace 'I' with 'we' in the list above. Or, at least, some humans seem to have this ability. True sapience surely includes this capacity to care and think for the benefit of family, friends, tribe, and even strangers. There may be a good reason, however, why this level of sapience is not widespread in the population of humanity. As I will dig further into the evolution of sapience in the next chapter, and to somewhat anticipate that exploration, note that not everyone in a tribe needs to be a strategic thinker for the group's benefit. Indeed, it is best if there are few wise folk in a tribe so that there is a higher chance that they will reach agreements. In other words, it seems reasonable that high sapience is not to be found in the majority of humans. However, sufficient sapience in the average person is needed so that they can recognize higher sapience in those whose judgments they will trust.

Conclusion

As we learn more and more about the functions of the frontal lobes, in particular the prefrontal cortices, we find that much of our unique human capacity to deal with our complex social milieu, our ability to think about the future, and to plan our actions in the present with the future in mind is associated with this remarkable region. Specifically, when we look for those qualities of mind which truly differentiate humans from other primates and all other animals, we are led to focus on the role of the fronto-polar region of the brain. The patch designated as Brodmann area 10 is particularly intriguing with respect to those capacities we recognize as the basis of wisdom.

I hope that neuropsychologists, in the not-too-distant future will concentrate on teasing out capacities for wise decision making (judgments). I strongly suspect they will find the activity of the BA10 patch is highly correlated with making wise choices.

Chapter 5 - The Evolution of Sapience: Past and Future

Overview

By all available evidence *Homo sapiens* was the first species ever to develop the sort of consciousness that allows not only awareness of the world and an ability to construct mental models of components of that world, but to be aware of its own thinking. It is the first species to develop a capacity for recursively generative abstract-manipulating language for more refined inter-individual communications. It is the first species to be capable of imagining combinations of objects that produce tools that give it non-biological adaptivity to a wide variety of environments. And it is the first species to form strong social connections not just based on familial (genetic) ties and able to cooperate with others in the society to accomplish complex work that could not be done by one person alone. It is the first sapient species.

In some ways I think the logic of placing our kind into a genus, *Homo*, and then as we discover precursor species that share some features with us putting them into the same genus is not right. The various species of beings that preceded our kind were clearly ancestors but it seems to me that the radical step up in consciousness and capacities warrants a different logic. In effect, I would have us be recognized as a new genus, one that departed sufficiently from prior species that we are more than just another species of the same genus. In keeping with the binomial nomenclature practice, I would propose that our genus be named *Homo* as now, but the prior species be identified with a term that means ‘human predecessor’. The current phylogenetic tree represents *Homo* as one of four members of a ‘tribe’ called *Hominina*, which contains *Pan* (chimpanzees), *Australopithecine* (with a number of species), *Paranthropus*, and *Homo* (also with a number of species besides *sapiens*, which are now extinct). In my view the radical situation with respect to *sapiens* suggests that there needs to be a fourth categorization, one that includes all the species formerly designated as *Homo* and leave the latter term to exclusively identify modern human beings in which there is only one species – us. The ‘new’ phylogenetic tree would branch at *Hominina* as now giving rise to the *Pan*, *Australopithecine*, *Paranthropus*, and this new category. Subsequently, at roughly 200,000 years ago the new branch would produce *Homo*. Just before that branch is established a few sub-branches, *Neanderthalis*, for example would have been very close to modern humans but perhaps not sapient. Thus, the prior genera subsequent to the split between the pre-human branch and the Pan branch are now extinct leaving only the genus *Homo* today.

Of course this is a suggestion that would probably not be given serious consideration. Think of all the textbooks and journal articles that would have to be amended. Think of the anthropologists’ careers that would be seriously altered. No, I offer this suggestion simply to try to put this idea of modern humans representing a radical leap from our predecessors into perspective for the following discussion. Thus in all of my references herein, I maintain the conventional terminology and phylogenetic labels.

Even so, the evolutionary forces that shaped sapience in the primate line are coming into better focus as paleoanthropologists uncover more evidence for the historical precedents of biological and technological (cultural) innovations such as language. At least the outlines of how sapience came into being are emerging. But what is needed is a grasp of what sapience is, how it makes humans different from all other contemporary and predecessor species, how it offers an opportunity for a new major transition in evolution. The previous chapters have attempted to answer the question of what sapience is. This chapter will attempt to explore the last idea: How does sapience allow the emergence of a new level of organization in the world?

First we need to take a look at the evolution of sapience up to the present. The perspective, as with the previous chapters, is from the principles of systems science. I will develop arguments about what I see as a kind of non-teleological trajectory of human evolution that provides the basis for making some speculations about the future of sapience and the genus *Homo*.

The Emergence and Development of Sapience

The story of sapience and its conjoined twin, 2½-order consciousness, is told in the evolution of hierarchical cybernetics, both within the individual human brain and in the societal matrix in which humans operate. It is the emergence of elaborated strategic management of the self, coordination management (tactical and logistical) in the social group, and the beginning of strategic management for the group in which we see the nature of wisdom and its relevance to human success in evolutionary terms. However, humanity finds itself in an awkward adolescence of consciousness. We have emerged from ape-hood as beings with superior future-thinking, mental models of both ourselves and others, and the language facility to share abstractions among ourselves. But we are not fully mature sapient adults, on average. As a species, on average, we appear to be only *somewhat* sapient. The capacity of the brain to build what we recognize as wisdom over one’s life is newly emergent and underdeveloped. I suspect the story of our species' evolution has more chapters to open.

In this final chapter I want to consider that story in terms of the evolution of sapience in our genus, *Homo*, its current status with respect to the degree to which it is expressed in modern populations, and its future potential as the next stage in human evolution. This is a hard story to tell. I often wonder if we have the courage to face up to our weaknesses as implied — that we are not currently sufficiently sapient in light of the problems we need to solve. As I asked at the very beginning, why haven't we developed a more perfect world or at least a world that is not threatened by destruction coming from our own hands? Indeed why is our world undergoing degradation as a result of our foibles[232]? The answer should now be clear. As sapient as we are, it just isn't enough to allow the development of global-scale wisdom in a sufficient number of

---

[232] Consider just a small sample of works that attest to the fact that humans are in the process of destroying themselves and their world. Catton (1982, 2009); Meadows, et al (2004); Lovelock (2006); Leaky & Lewin (1995); Montgomery (2007); Flannery (2005); Speth (2004); Heinberg (2007).

people that are in positions of *effective authority* [233]so as to affect the needed changes. We have reached a fundamental threshold for testing sentience on our planet. Enrico Fermi asked a very relevant question, now called the Fermi Paradox[234]. He was puzzled by the accumulating evidence, from evolution theory, that life would actually be a common phenomenon in the Universe. Frank Drake proposed a theoretical model, the Drake equation, which indicated that, given certain not-unreasonable assumptions, intelligent life should be more abundant in the galaxy[235]. Fermi asked, if this were the case, why haven't we seen evidence of this? Our Search for Extraterrestrial Intelligence (SETI)[236] has yet to uncover any indications that there are other sentient beings with advanced communications or travel technologies out there in the galaxy. Why not?

One very obvious answer to Fermi's paradox is that extraterrestrials are hiding from us! If you were a super-intelligent being, capable of interstellar travel and presumably having overcome the kinds of aggressive tendencies that would lead to self-destruction on your own planet (see below), observing what was happening on Earth, would you be inclined to drop in and say “Hi!” Another answer is that Drake's approach was misleading and that there really isn't much if any life out there in the universe. But yet another, more scary, solution to Fermi's paradox is that intelligent life reaches a point at which it cannot sustain itself because it has created an overly complex society without adequate sapience to guide further development. Indeed, Drake's equation includes an estimate of how many civilizations might actually succumb to their own cleverness. What if the answer were ‘most’? Is this the possible fate of humanity on this planet? Put very simply, what if we are too clever for our own good?

Is this some kind of universal fitness test? Are we facing such a test? If we are, will we prove worthy of continuation into the future of our planet? These are very likely questions that we cannot ignore. One reason for hope follows the logic that if, indeed, the emergence of living systems is actually quite likely and the trajectory of biological evolution leads to sentience/sapience (Losos, 2017; Morowitz, 2004; Smith & Morowitz, 2016; Stewart, 2000), and if evolution is a universal continuing process, then it follows that the current species is but the beginning of a new major transition. More on that later.

---

[233] By ‘effective authority’ I mean not a power relation in the conventional sense, but rather an influence relation in which prominent people (what we might call ‘thought leaders’), by virtue of their wisdom, guide the rest.

[234] Fermi famously asked, “Where is everyone?” See Wikipedia: http://en.wikipedia.org/wiki/Fermi_paradox

[235] See the Wikipedia article: http://en.wikipedia.org/wiki/Drake_equation for more details.

[236] See this article in Wikipedia for a description: http://en.wikipedia.org/wiki/Search_for_extraterrestrial_intelligence

Evolution and Hierarchical Cybernetic Systems

Simple systems have a way of developing into complex ones as long as free energy flows and materials for construction of parts are available[237]. This has been true for life on Earth, it is true for human organizations, and it is true for human minds. As these systems age they gain more structure and function as well as more knowledge. Yet, at some point, the complexity threatens to overcome whatever benefits have accrued from gaining it. Complexity solves certain problems, but it also creates new problems. Specifically, when the number of active components exceeds some threshold, the communications and interactions become burdens on the components and the system's overall organization and function is compromised. The energetic overhead begins to dominate the dynamics. There is a law of diminishing returns that applies to increasing complexity and that law can be seen in action throughout nature's systems[238].

Living systems were once just single celled bacteria-like organisms, living happily in the primordial seas. Somewhere along the line some of these organisms linked up through a process Lynn Margulis called endosymbiosis to form more complex eukaryotic cells, single celled animals and plants[239]. After a very long time, aggregates of such cells began to form and certain cells in the aggregates took on specialized functions. In particular, some became germ cells, able to carry on reproduction, and the rest became somatic cells, responsible for feeding and nurturing the germ cells. The evolution of life is a story of on-going increases in complexity in response to environmental selection forces and opportunities to exploit the free energies available in the environment[240]. How living systems have managed to evolve more complexity without suffering the consequences of diminishing returns on that complexity is, itself an extremely interesting story. It is the subject of chapters 10 and 11 in Mobus & Kalton (2015).

The bottom line on how life has evolved without running into the limits of complexity is simply the tendency to organize its structures according to the principles of hierarchical cybernetics. The hierarchical part involves the formation of functional network modules within time-domain layers and the cybernetic part involves communication and control/regulation within that hierarchy. Once nervous systems in animals appeared, especially with actual brains instead of bunches of neuronal ganglia, evolution of behavior came under the same laws of organization. The brain is a hierarchical cybernetic system. And in humans it has become the epitome of that model. The operational level (i.e. the physiological management of the body) is largely handled by the brain stem and endocrine systems, the logistical and simpler tactical (coordination) level

---

[237] Mobus & Kalton (2015), chapter 10 provides a fairly detailed description of the process of complexification, or what I now refer to as ‘ontogenesis’ or the origin and development of systems and systems of systems. .

[238] Joseph Tainter (1988) developed a rather comprehensive theory about how increases in complexity led to collapses of historical societies. See also Homer-Dixon (2006) for a study in how imperial civilization expansionism (and increasing complexification) leads to diminishing free energy per capita and consequent collapse.

[239] Margulis & Sagan (1995)

[240] The theme of emergence of social organizations from molecules making cells, to humans making societies is found in a number of very good references such as, Bourke (2011); Calcott & Sterelny (2011); Morowitz (2004); Smith & Szathmáry (1995); and Volk, 2017 to list just a few.

is handled by the limbic system and paleocortex (in early mammals, augmented later by the neocortex), and the more sophisticated tactical (e.g. maneuvering in complex terrain while stalking prey) is handled by the frontal lobes (specifically the prefrontal cortex). In the human brain the capacity for strategic management arose with the expansion and specialization of the Brodmann area 10 patch of prefrontal cortex, as described in the last chapter. Strategic thinking involving longer time scales and wider scope may have emerged from this kind of shorter-term tactical ‘planning.’

And because humans have coevolved the ability to think and plan strategically along with their hyper-sociality the complexity of a social organization of people as an emergent phenomena demands some form of internal management. It turns out that the society of humans has its own form of hierarchical cybernetic organization. This system arises from the ways in which humans interact with one another in familial and neighbor relations where cooperation and competition act to provide coordination for the group. It depends on the use of language (symbolic-based communications) to allow individuals to share their inner thoughts and experiences. It depends on individuals’ abilities to construct models of the others in the group as well as models of themselves within their tacit knowledge base for use in interpreting behaviors and speech acts.

Just as individuals gained strategic thinking abilities as part of this dynamic, so too social evolution involves strategic decisions regarding the group's situation in the larger environment and especially with respect to other groups. The social structure requires a hierarchical governance structure similar in form to that in the brain of individuals.

In human groups (tribes) where individuals tend to specialize in skills the need for cooperation exceeds that of competition as a general rule. People doing specific jobs, like arrow making, or hunting, need to work in concert for the good of the group. But doing so successfully means that groups can get much larger and if their sizes exceeded a threshold (e.g. Dunbar’s number[241]) then some forms of coordination management were required. Thus operational level cooperation may be augmented with tactical and logistical specialists. The long-term survival and fitness of the group as a whole might easily require a “leader” who was wise in the ways of the world and could provide the strategic guidance needed. Figure 5.1 suggests a social hierarchical group system.

---

[241] Robin Dunbar, an anthropologist, theorized that based on the size of the human cortex (as relative to other primates) that social group sizes of around 150 people was the upper bound on stability. See the Wikipedia article: https://en.wikipedia.org/wiki/Dunbar%27s_number for background. Accessed 4/12/2019.

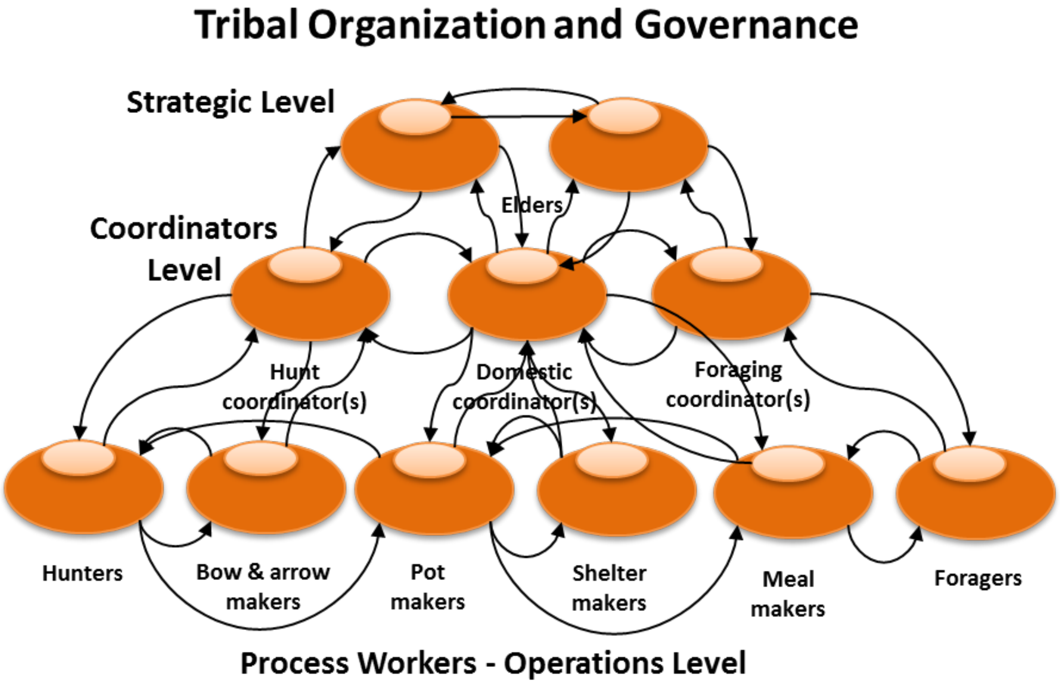


Fig. 5.1. Humans tend to have relatively high levels of specialization in skills and abilities, including abilities to help coordinate activities needed to sustain the group.

Every member of the group possesses a hierarchical cybernetic brain having operational, coordination, and strategic thinking capacities for their individual lives. Each "specialist worker" uses all of that thinking capacity to organize and manage his/her own work. They all use their strategic thinking ability to manage their interrelations with one another. Figure 5.2 shows the management hierarchy where each individual at each level still possesses the capacities to do strategic, coordination, and operational thinking needed to be a quasi-autonomous member of the group. Because each is quasi-autonomous and under the influence of inherent talent development, some develop better coordination competencies than others. That is they can provide help to the group as a whole by taking on tactical or logistical specializations. The hunter with the most experience and prowess might become the hunting party leader, planning and directing the hunting party activities. Another might be good at helping direct village work activities and distributing various resources, becoming a logistical manager of sorts. When these talents emerge the tribe, as a whole, is much better off. They become more efficient in providing the necessities of life. And, indeed, in times of plenty may make it possible for surpluses to accumulate.

The capacity as well as the need for wisdom arose in this recursive, self-similar framework of individual expansion of judgment coupled with the dynamics and cohesion of groups in which individuals lived. Sapience and group management co-evolved in that one spurred the other and vice versa. Language, too, evolved in this framework, being essential for sophisticated communications between specialists needed for a distributed hierarchical system to work.

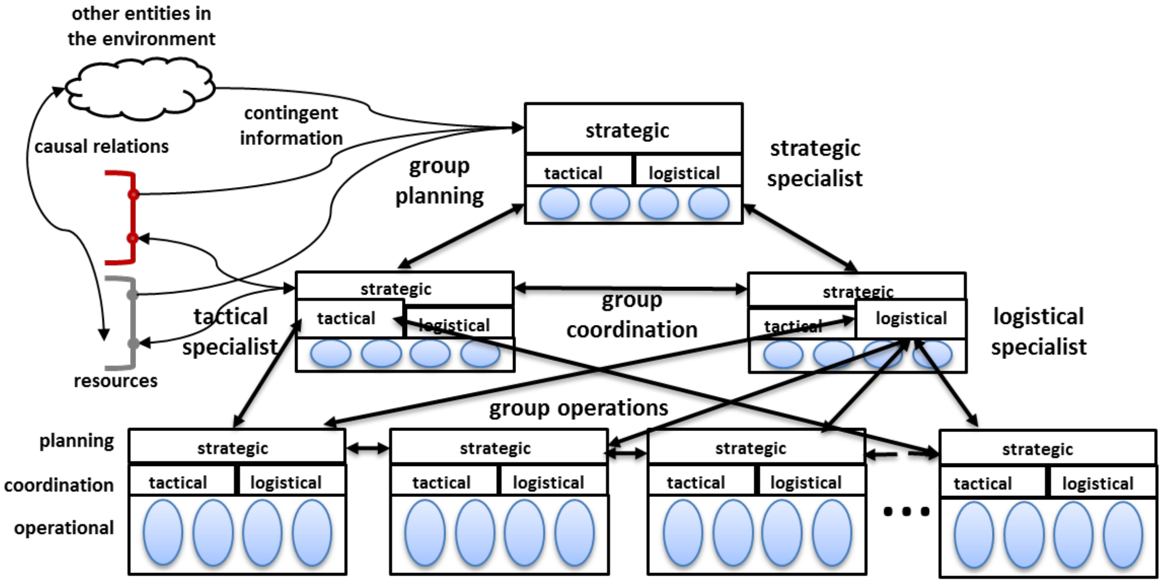


Fig. 5.2. The evolution of hierarchical cybernetics at both the individual level and that of the group involves the recursive application of the model from the individual level to the group level. As group sizes get larger, more coordination and long-term planning are involved in maintaining group cohesion and success in surviving. Individuals evolved to have greater strategic capacity with respect to their own lives within the community. But some individuals, representing the higher end of the sapience distribution, could function as strategic thinkers for the group as a whole. Likewise, some individuals may specialize in tactical management (lead hunting parties) and others specialize in logistical management (making sure there are enough arrows and bows, etc.). The strategic thinker needed to take in far more information regarding the environment than just the tactical messages to/from the resource sources (red and grey open rectangles). They needed to understand many more causal relations between the necessary resources and other entities (e.g. competing tribes) in order to plan future tactical activities.

Key individuals who showed particular strength in strategic thinking could assume the duties of doing such thinking (as in figure 5.1), or at least influencing the group thinking, for the group as a whole. The success of such individuals rested on their ability to understand the external environment, including other groups, and understanding one's own group members and their capabilities. Providing guidance for the group's long-range activities, such as where and when to plant seed or hunt, etc. based on a lifetime of accumulated tacit and explicit knowledge would set such individuals apart, with others in the group looking up to them as wise elders.

But this is the picture that probably held toward the end of the Pleistocene when humans were just starting to live in larger extended families and groups containing non-related individuals. What got human evolution going in this direction and led eventually to the emergence of hierarchical management of the group was something more fundamental. It was tied to the reproductive success of early hominins which took several interesting turns. And one of those turns probably got initiated much further back in hominid evolution.

The Evolution of the Genus *Homo*

The genus *Homo* is thought to have emerged from a prior genus, *Australopithecus*, about 2 to 2.4 million years before the present. The line of this genus is that which led to our species, *Homo sapiens* with numerous other species having emerged in various regions at different times along the history. Several species appear to have been contemporaneous at various times and may have shared ranges[242]. We know that members of our species migrated into Europe while the Neanderthals were still occupying that continent for example. There is archeological evidence that they co-occupied various sites and even inter-bred.

Since there are so many very good (though sometimes seemingly conflicting) accounts of human evolution I will not attempt to replicate that story in this chapter. Interested readers will find numerous references in the bibliography. There are many transitions from mere animal mentality and behaviors to human that are part of the story. The initial taming of fire, the origins of language, complex tool making, and, especially, the cohesion of tribes for successful survival of naked and vulnerable apes are all important parts of the story. The role of climate change in selecting for traits and cognitive capacities is also an important aspect[243]. All of these and more are incredibly fascinating parts of the story. If the reader truly wants to understand the wonder of the story they need to read some of the references. I simply cannot do justice to it.

Rather I am concerned with the specifics of the evolution of the human brain, and most particularly its evolution as the genus *Homo* gave rise to the species '*sapiens*.' This particular event, as I claimed in chapter 1, was a breakthrough in cognitive capabilities. Some bases for it may have existed in species prior to *sapiens*, for example the prior capacity for systems thinking and morally influenced judgments. But the real breakthrough came with the rapid expansion of the most fronto-polar part of the brain that eventually generated true strategic thinking and greatly increased the sociality of humans. Language as we know it emerged from the capacity of the brain to build models of others' minds as well as one's own mind and to share concepts through abstract representations (words and sentences). This appears to be a late development in the emergence of modern humans.

The evolution of new traits, both physical and behavioral, or the modification of existing ones, is the result of two basic situations. First there must be a modification to the genome that generates a change in the phenotypic form (anatomy or physiology for physical traits and influences on

---

[242] For a very complete and up-to-date review of human species evolution see Tattersall (2012). Also see Suddendorf (2013) and Tomasello (2014, 2016, and 2019) for very recent perspectives on human cognition, its differences from other great apes and prior hominins, and its evolution. For a "big picture/deep history" view of human evolution see Hariri (2015).

[243] A particularly interesting article by Marean (2010) describes how climate change in Africa, approximately 120,000 years ago led to the near extinction of *Homo sapiens*. All that survived may have been an extremely small population in southern Africa that learned to live off of shell fish. All modern humans may have descended from this population. The situation is referred to as an evolutionary bottleneck and may yet become important for modern humans to think about.

brain wiring that modify behaviors). Then the new/modified trait has to be subjected to environmental selection – does it make the individual possessor more fit than its conspecifics, i.e. can it reproduce more offspring over time. The evolution of human cognition is no different.

What was going on in the African continent with respect to the changes in the environment due to changes in the climate for the last two and a half million years is implicated in driving human evolution[244]. The fluctuations in climate associated with this time put heavy adaptation requirements on many species. These selection pressures were not targeting increasing cognitive capabilities per se, but coupled with the already impressive abilities of the hominids helped to steer the genus down the path of increasing intelligence and creativity. For example there is evidence that Australopithecines had learned how to handle fire for cooking and warmth. The hominins showed evolutionary adaptation to the stresses of relatively rapid changing environments possibly because they had a head start in terms of behavioral repertoires. Evolution usually proceeds by building on existing capabilities rather than inventing something new. So it seems to be the case with human brain and cognitive evolution[245].

What existed at the time these fluctuations started was a species that had already achieved great ape level consciousness, what Tomasello (2014) describes as "joint-intentionality". This is the mind that knows what it wants and knows (or has a pretty good idea) what other minds want and therefore knows how to maneuver a primitive social structure to get its own needs met[246]. Tomasello takes the stand that modern apes (which he has studied extensively) probably 'think' in similar ways to the last common ancestor between hominids and apes (6 million years ago). Apes today essentially think the same way as they have for all that time, while humans have diverged due to their unique environmental pressures.

The fact that apes (at least chimpanzees and bonobos) live in social groups is the key factor in what came next. Perhaps due to the increased survival pressures put on a (by that time) naked, almost toothless, upright standing ape pushed the genus toward a new form of eusociality. Put simply, humans needed to draw closer together, to cooperate with one another, in order to survive in a constantly changing but ever hostile world. Natural selection (working through group selection) imposed that *strategy* for humanity. Mere altruism based on kin selection was no longer enough. The human brain started evolving greater capacity to build much more elaborate models of the world but more importantly more concrete models of others' minds in the group. Working through the stage of joint intentionality, as Tomasello calls it, in which individuals acquired brain-based propensities to cooperate with other individuals on tasks that neither one could accomplish by themselves, and for which the rewards of success were much

---

[244] The Pliocene-Quaternary glaciation which started a little more than 2 $^{1/2}$ million years ago involved a sequence of ice expansions and retreats that reflected the fluctuating changes in climate every 40,000 to 100,000 years.

[245] Another source for getting a comprehensive picture of brain evolution is Striedter (2005).

[246] Tomasello (2014) describes the role of causal relations in thinking. Recall from chapter 4, I have provided a fundamental basis for how such relations are encoded at the level of neurons and circuits!

more than both individuals needed (a profit was made), humans evolved the capacity to cooperate. This success put pressure on the improvement in abstract and functional communications. Language evolved to meet the needs.

From a functional (systems) perspective what evolved in human thinking capabilities that made the biggest difference was the rise of strategic thinking linked with the other components of sapience and those with intelligence, creativity, and affect. The emergence of a strategic brain is, I assert, directly responsible for the emergence of the next level of consciousness, the 2½ order kind that gives us consciousness of our own thinking and reflection on the fact that we are conscious of our own consciousness. This capability is at the heart of our capacity to have recurrent models of other minds. I think about what you are thinking about what I am thinking! I think about what I should do if you think I am thinking I should do something and I either don't want you to think poorly of me, or I am trying to conceal something from you that would make you think poorly of me. And the chain doesn't need to stop there. Nor does it need to only involve two-way relations. I could be concerned about what Jane thinks about Dick because he thinks poorly of me!

The environment that dramatically selects for our strategic thinking is the collective, the social environment that includes the products of our own cleverness, our cultures. We have to cooperate with many different kinds of individuals with many different kinds of joint intentionality acts simultaneously processed. Humans developed what Tomasello calls "collective intentionality." The ability to construct complex, network-based models of causal relations involving multiple personality models in which we can propose work relations that will make us all better off is unique to human thinking.

Evolution of Strategic Thinking

One of the most cogent aspects of the differences between sapient beings and all previous animals is a capacity to think about possible future states of the world[247]. And not just in a passive way. The capacity to think about what one *wants* the future states of the world to be and plan tactical programs to achieve that future is the core of strategic thinking. Of course what makes the difference between strategic and tactical planning is, essentially, the time frame and the scope or complexity of the world in which the plan will be executed. Another aspect of the difference is that a strategic plan is actually a "temporally ordered" set of tactical plans in various levels of detail. The near-term tactical plan is usually sharp in detail, many steps and considerations are taken into account. What actions need to be taken now are usually clear in mind. Tactical plans further out (and contingent on the success of the near term plan execution) are generally fuzzier or less distinct (and more abstract). The site for organizing the strategic plan is the prefrontal cortex, as discussed in the prior chapter, specifically I think BA10. Recall in the section titled, "The Neocortical Brain, the Prefrontal Cortex and Executive Control" and figures

---

[247] Suddendorf (2013), especially chapter 5.

4.9 and 4.10, I showed the overall relation between strategic planning circuits in the prefrontal cortex and the linkages between outcome concepts and learned behaviors (behavior concepts) which constitute tactical plans.

Figure 5.3 gives a sense of how a sequence of (more or less discrete) tactical plans can be set up as a long-term strategic plan under the auspices of the prefrontal cortex. A similar, though probably less well formed, sequence of logistical plans could be formed, but the bulk of strategic planning involves tactical sequences so that is what I will focus on.

Each tactical plan is a learned behavior and has, for itself, a near-term goal or outcome that represents a considered necessary sub-step toward the long term goal (or sub-goals). The tactical plans come from a "library" of pre-defined tactics and are chosen for their appropriateness to achieve the long-term goal. However, the further out in time the tactic is placed in the plan the certainty of its need at the appropriate time or its outcome if it is used is increasingly low. Thus those sub-plans are more fuzzy and tentative.

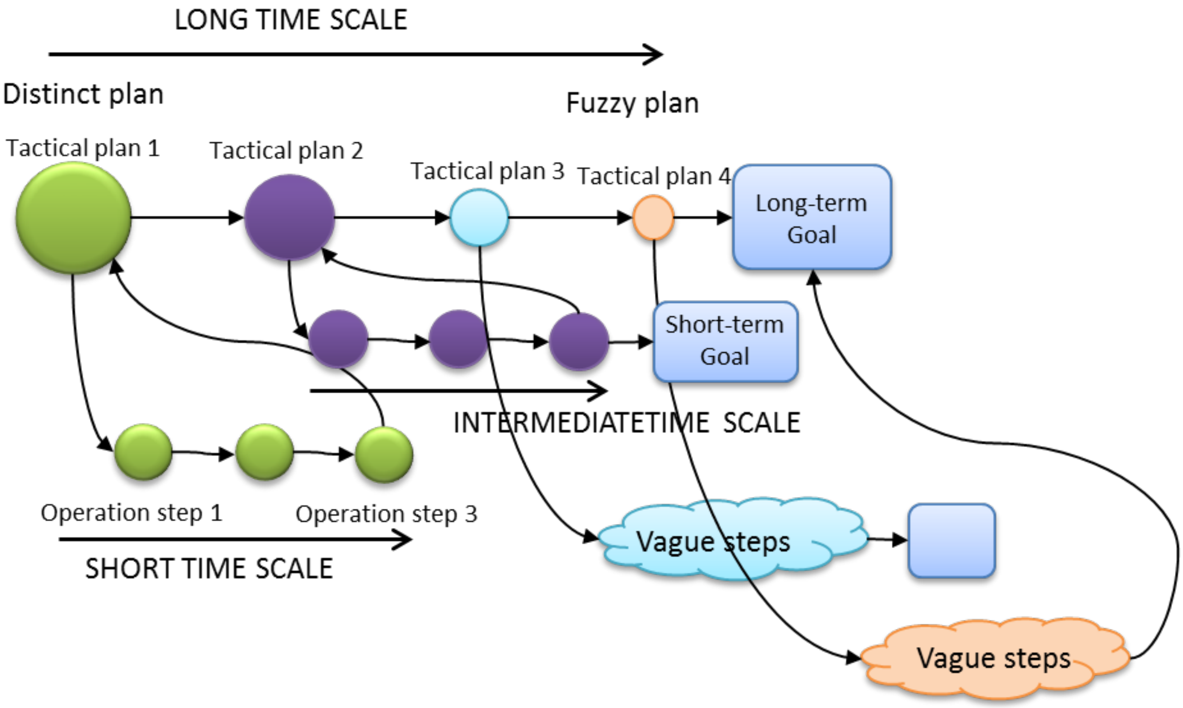


Fig. 5.3. A strategic plan is a series of tactical and logistical plans (each a near term sequence of operational steps needed to reach a sub-goal). Here I show the mechanism for forming a strategic plan as a sequence of tactical plans of varying time scales (and vagueness). The green series is what the agent will do next, to be followed at a later time by the purple series if the green one is successful. Similarly the light blue tactical plan is much less certain (it is fuzzy) and the steps may be vague when the plan is struck. Ultimately the plan is generated in order to achieve some long-term goal.

The circles in the figure represent the abstract representation of the plans that are held in memory and called to consciousness as the agent considers each step in the sequence. This logical structure is linked in the same manner as the motor program construction concept discussed in the last chapter (figure 4.19). It is not too much of a stretch to imagine it is implemented in neural circuitry the same way but at a much higher level in the cognitive hierarchy, i.e. by the networks in BA10 for the top sequence in figure 5.3 and in more posterior regions of the frontal cortex for the tactical plan templates (small green and purple circles).

Figure 5.3 might leave the impression that plans are linear structures only. However this is not the case. As we saw with heuristic programming using neural networks in figure 4.19 it is possible to have contingency programming providing alternative pathways in the event that conditions change or the chosen execution path plan doesn’t work out. That is, at any point in the sequence in figure 5.3, for both the strategic plan level and the tactical plan level, it is possible to have alternative templates or operation steps available. These are actually part of a branching tree with as many branches as the processing capacity of the planning brain can manage. They are technically ‘scenarios’ rather than fixed plans.

Each of the smaller circles in the figure represent dynamic models of what happens IF. The models provide outputs that pass to the next stage based on the inputs they receive from the prior stage and, most importantly, feedback from the environment in terms of the outcomes up to that point. Planning is not more than ‘running’ the sequence of models making choices about the inputs that ‘should’ be relevant in the real world at the real time. If the outputs from the models fit the needs of achieving the short-term goals, then that model or tactical plan template is inserted into the strategic plan. This is what is meant by *strategic thinking*.

The efficacy and completeness of strategic plans depend entirely on two things. The working capacity of the brain circuits given to doing the planning (i.e. model selection and running, checking, and heuristic programming) determines how many contingencies can be taken into account and thus how many scenarios can be included in the overall plan. The second thing is how complete and veridical the individual operational step and tactical plan templates are. And that is a function of the overall tacit memory capacity and the experience with real-world models of the agent. Once again we arrive at the issue of wisdom. Constructing complex strategic plans is not something that people can do strictly in their conscious minds (i.e. using intelligence alone). The mind doesn’t have the bandwidth, or what we call working-memory capacity to do so.

Most of our ability to think strategically is subconscious and dependent on our tacit knowledge stores, particularly the knowledge of models of actions that lead to goals in coordination with the rest of the world. When we consciously think about these models it is the abstract version of them that enters out thoughts, not the detailed step-by-step motions that we take. Like Shank & Abelson's (1997) scripts, we don't think about every detail of every step in going into a restaurant and ordering a meal. It is part of a flexible heuristic program that gets executed in the background. We only become conscious of any particular step when something doesn't go according to plan. Nevertheless, our subconscious minds that learned the many scripts and put the overall plan together has done a stupendous amount of work to plan for longer time horizons.

As far as we are able to discern no other contemporary species is able to do this. From the limited archeological record we only find evidence that early hominins planned. For example there is evidence that humans carried stones for long distances from their source to the sites at which they were worked into hand tools. This isn't exactly 'long-range' planning, but it does show that the mental machinery for thinking about a *possible* future was emergent in species prior to *sapiens*. In the record of the latter we see considerable evidence of planning and executing tactical plans, like constructing temporary shelters for seasons of hunting and wintering over.

In a tightly coupled social unit like a tribe it became increasingly important for an individual to consider what those around them would do given some behavior that was anticipated. For example, what would the other young males in the tribe do if one were to court a particularly attractive female? Would they react with negative actions? Jealousy? Would they accept the actions as good for the group? The individual needed to take these aspects into consideration before making a 'move.'

How would the female react? Would she reject his advances? So many complex possibilities needed to be taken into consideration. Should he be tentative to test the water? Should he make a bold move to prove his manhood? Real considerations give rise to extremely 'bushy' scenario trees, every branch conditioned by different likelihoods. We are back at the decision tree model, but now writ orders of magnitude larger and more complex.

The social environment created an incredibly complex milieu in which to operate. There are so many different consequences of actions taken in the near-term to be considered before acting.

What made this so difficult is the need for group cohesion. Each individual needed to take the feelings and thoughts of the others into consideration in making plans. The reason is that the group as a whole had become the 'unit of selection.' The success of the individuals ultimately depended on the success of the group. The more cooperative they were, the more they shared their cognitions with one another, the more they divided up the work load the more likely they would be to be around next year to produce more children. It mattered what others thought about what you did. And you cared very much about what others did.

Each individual needed to consider strategic consequences for themselves in the group context. But, as group cohesion meant more cooperation, the group as a whole had to consider strategic consequences for them as a unit.

This was the essence of the selection pressure to expand the prefrontal cortex and especially BA10. The brain needed more space to model others' mental states and personalities. So the whole cortex expanded substantially as the pressure increased. The more capabilities early humans had to cooperate and consider one another's thoughts and feelings the more joint success they realized and the more their populations expanded. This, in turn, increased the pressure to expand further the capacity of modeling since there were more individuals to model. We don't actually know when Dunbar's number reached the 150-200 person level but we do know that there is a correlation between that number and the expansion of the cortex that enabled more capacity to model others and the future. From a systems perspective this is not unusual. What was set up was a positive feedback between expanding cortex, fitness, expansion of group sizes, and back to selection for expanding cortex.

Of course this positive feedback would be countered, eventually, by some negative feedbacks. Perhaps most relevant was the problem of expanded cortex in the young creating problems in childbearing. A woman's body can only adapt so far (mechanically speaking) to allow a large headed baby to pass through the cervix without doing serious damage. In fact serious damage *is* endured by sapient women (as compared with their great ape cousins) along with increased risk of death from complications in childbirth. Human babies have to be born in a premature state, compared with other primate young, leading to complications in child rearing (see below). Surprisingly, however, this adaptation actually led to further positive feedback in the development of sapience.

The need to think strategically reinforced the propensity for the enlargement of the prefrontal cortex patch, BA10, which, in turn led to concomitant increase in the whole prefrontal cortex, which in turn led to the expansion of the whole cortex (possibly the whole brain). This led to the increased success of the group (and species) and the need for strategic group-think. Some groups were lucky enough to have individuals who had an extraordinary level of strategic thinking, thinking that covered the whole group for a very long time frame and for a very large environment. These were the wise persons that the group could turn to for advice for life and for success.

Evolution of Judgment

The expansion of the cortex meant something very important for *Homo*; it meant that there was more real estate available for the construction of far more complex, detailed, and accurate models of everything. These models (their templates particularly) were needed to do the planning in the last section. But there remained a problem in that there was not that much more real estate in the prefrontal cortex for conscious access to all models in their increasing levels of complexity, detail, and accuracy. Instead the evolution of the brain took a different, more compact route. Consciousness remained tuned to just the most abstract representations and then only when they were activated strongly, were most cogent[248]. Recall the claim I made in chapter 1 that most of our cognition is accomplished in the subconscious. We now know this is the case from neurobiology[249]. In essence, a far greater amount of thinking takes place in a 'preconscious' level and only becomes conscious when it needs the attention of the mind as a whole. More often, our thoughts enter consciousness not directly as whole-formed thoughts, but as subtle messages that affect our decisions that are currently being attended in consciousness. This is what we mean by judgment and intuition.

Judgment and intuition, at their best, provide our species with substantial problem solving abilities without conscious thought being involved. What we are conscious of is having found a solution, or an idea of a solution. We are conscious of a feeling of certainty that our judgment, or intuition, is correct (veridical). And we can use that feeling and judgment to plan actions. We can advise ourselves or others based on them.

As I have already considered, above, the increase in correctness of our judgments, and sense of certainty, due to being able to construct more complex, but more veridical models, due to the expansion of the cortex and the prefrontal capacity to build those models, led to greater group success in living. The very same argument that applies to the positive feedback that led to greater strategic thinking applies to the increase in capacity for judgment. Better judgment, better success and greater fitness.

On the level of the group the same dynamic applies. Any individual or subset of the group that gave better advice would likely be supported more by the group as a whole. Any offspring of such an individual would be cared for with somewhat greater effort. The wise elder(s) and their offspring would take a special place in the group dynamic simply because their contribution to group success would be noted.

Evolution of Eusociality and Language

Below I will consider some specific aspects of the evolution of sociality (eusociality and what some now call hyper-sociality). Here I want to review several important qualities that make the kind of human sociality possible as regards sapience. These are some of the more important attributes or processes that arose in humans in their breakthrough to sapient status. These are

---

[248] Dehaene (2014) describes the "ignition" effect in which the higher levels of the neocortex (prefrontal cortex) is sufficiently excited from lower levels (e.g. sensory inputs) to become self-reinforcing and dominate the cortical workspace. This, he claims, is the access to 'consciousness' that results in our subjective experience of the 'thing' we perceive.

[249] Dehaene (2014), chapter 2, provides the experimental evidence that our brains are capable of carrying on very complex thinking and problem solving in the unconscious. For example, we might sleep on a problem and wake up the next morning with a solution in mind. We do not need to be conscious of our problem solving in even most complex situations for it to work!

extremely important because their further development I see as crucial to humans achieving a much higher (and needed) level of sapience for future success of the genus *Homo*.

Theory of Mind (ToM)

I mentioned Tomasello's "joint intentionality" and "collective intentionality" above as well as the notion of a human mind able to construct a model of the mind of another individual. This is commonly referred to as the 'theory of mind' theory of interpersonal relations (see figure 5.4 below). In a 'weak' version of ToM the possessor of the theory may only have a notion that the other individual has a mind and is capable of perceiving the same things as the possessor sees. For example, Tomasello claims that great apes are capable of conceiving that other members of their group are seeing the same things, such as a piece of food, and have the same intentions, getting the food, as the possessor. But in humans the ToM concept is much stronger. Essentially, one human being constructs a model of another person, their way of thinking, their personality, their desires, capabilities, etc.[250] That is, one individual comes to 'know' another individual well enough so as to 'predict' what that individual will do in the future under various conditions. I can imagine that you will get mad at me if I imply that you are fat! That is because I know you to be sensitive to your weight, easily provoked, and likely to lose your temper if I breach a protocol[251]. So, not wanting to make you mad, I avoid any reference to your weight!

I argue that our brains have evolved the capacity to derive extended models of other people from a base model of people (or the individual intentionality possessed by apes). This includes a model of our own selves. We are a personality with intentions, desires, etc. just as any other person. We construct a model of our own minds or our so-called 'self-image' as a way to consider ourselves in future scenarios. This is not more than the mechanism by which we do strategic planning for our own future. But it also is used to evaluate how we actually preform in the real world. We might predict that we will act in a certain way based on our current model of ourselves, and then after we actually interact with the real situation, use the information to "adjust" our model. That is, if we are self-honest.

Sapience plays a huge role in deciding on veracity of our model of ourselves as well as those of others. If the affective inputs are strong, we might tend to construct models in which we are seen in positive light (the default condition), what Freudians might call "ego protection." More sapient individuals are more likely to see their own weaknesses and predict their own failings (under certain stressful conditions, for example). And they are likely to honestly assess their actual performance. Most of all, however, they are more likely to be able to use the information they get from experience to adjust their own thinking. That is, they can use their theory of their own mind to improve that mind. That is what I mean by a $2^{½}$-order consciousness. Only the

---

[250] See Mitchell, et al. (2006), page 65 for a discussion of neuroimaging studies showing the parts of the prefrontal cortex involved in processing social cognition.

[251] One of the more meaningful plot aspects in the movie "Avatar" (http://en.wikipedia.org/wiki/Avatar_%282009_film%29) was when the Na'vi female, Neytiri tells the human, Jake, "*I see you*," where she means I understand your mind/personality. She had constructed a deep model of Jake!

argument might be advanced that such a sapient mind deserves a full 3rd-order designation! I will present that argument later.

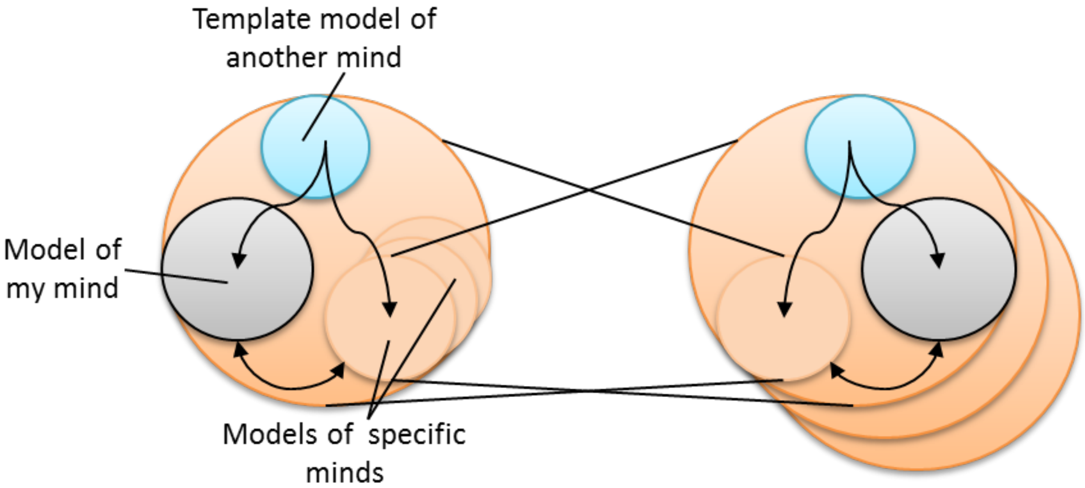


Fig. 5.4. Every human mind contains models of other minds (pale circles) along with a self-model (grey circle). Both models derive from an a priori systemic model (blue circle). The template model can be reused any number of times based on how many individuals the person meets and gets to 'know.'

Empathy

Having a theory of another person's mind leads to an extremely important facility. What if the other thinks the same way I do? Or, even more importantly, what if they feel the same way I do under similar circumstances? Can my theory of another mind include a way to connect to that mind by feeling what they feel?

Empathy is not the same as altruism. In the latter case, behavior can be hard-wired and automatic (as in warrior ants attacking an invader without thinking to protect the nest). Empathy involves actually experiencing a form of what the other is experiencing in order to increase the likelihood that one will cooperate with or help the other. As far as I know, no ant ever felt sorry for a nest mate or experienced the pain of seeing a nest mate attacked by an enemy.

To be able to actually get inside another person's thoughts and feelings, to experience even a shadow of what they are experiencing is a wholly new mechanism for reinforcing mutual intentionality. There is no compelling evidence that other animals feel another individual's feelings or put themselves in another's mind to consider what they must be thinking in quite the way that humans do. Pet dogs do seem to have some kind of empathy with their owners it seems but we cannot know that this is nothing more than a conditioned response to a human behaving in a certain way whenever they are feeling joy or sadness. The dog cannot tell us that it is feeling our pain. The case with some of the great apes, like bonobos, is somewhat ambiguous. It is clear that they possess a weak form of ToM, and sometimes demonstrate behaviors that sure look like they are sympathizing (as when another chimp rushes to groom a tribe-mate who has just been thumped by the dominant male, as if to soothe their hurt). The strong form of ToM likely took shape during the transition from early forms of *Australopithecines* to *Homo*. From some evidence in Paleolithic archeological record early hominins were clearly behaving as if they were possessed of a stronger form of ToM[252].

As with other cognitive activities much of empathy is processed subconsciously or preconsciously. We become aware, sometimes, of feeling sympathy or enjoying another's happiness. But most of our empathizing goes on beneath the surface and rises from sapience as our intuitions about the other person.

Moral Sentiments and Cooperativity

Our normal tendency is to not want to cause another person distress or pain. Indeed we seem to feel best when we are the cause of joy and happiness in the other. In essence we are strongly motivated to do good, being defined as that which uplifts the spirits of others. Empathy can reinforce our tendencies. We would not want to cause another distress only to feel that distress ourselves. And we would desire to cause another joy if we then feel their joy ourselves.

Moral sentiments, senses of right and wrong, fair and unfair, proper and improper, etc. appear to have arisen in early primates[253]. Capuchin monkeys are known to get mad when they think they have been slighted, exposed to a situation in which a cage mate receives a more tasty reward (a grape rather than a slice of cucumber), for no apparent reason.

In the context of collective intentionality (as described by Tomasello, 2014) it is not too hard to construct an argument for why moral sentiments as found in humans provided significant group fitness. The same is the case for cooperativity. In chapter 3 I argued that altruistic tendencies as seen in numerous species could lead to empathy once a brain could construct a model of another individual's thinking, as in shared intentionality. Taking that one step further the strong sense of cooperativity may have emerged in conjunction with moral sentiments. Cooperation is not the same thing as altruism, or even empathy. It stems from a desire to achieve a collective intention by working with others so that all will benefit. No one need sacrifice any fitness. Quite to the contrary, all participants come out ahead, achieving a goal that no one individual could have accomplished alone.

Coupled with the tendency toward skill specialization and an ability to trade products and services, suggested by figure 5.1, cooperativity started playing a major, *the major* role in organizing social systems. To be sure, the capacity to resort to competition when the conditions demanded it (scarcity for example) remained as a backup plan and could even be raised to the level of dominance in human affairs at times. Competition in modern markets might well be viewed in this way. It is a holdover from the era when human tribes needed to compete with one

---

[252] De Waal (2010).

[253] ibid

another to obtain needed resources. Individuals retain an ability to compete with other individuals (for example in the labor market) and companies compete with one another to sell products to a limited customer base.

Still, when individuals view themselves as members of the “In-group” they are more inclined to cooperate in working together.

Love

This review of the evolution of sapience could not be complete without at least a mention of something that provides extraordinarily strong bonds between individuals in more than just familial situations. In English there is only one word to encompass a range of positive emotional feelings that individuals can develop for one another. Romantic love and love for children, siblings, and parents, are relatively easy to explain evolutionarily. But love for friends extends beyond cooperation or empathy. It can be seen as a bonding energy not unlike gravity or electromagnetic forces. It mediates keeping people in contact with one another and that contact provides a milieu of well-being for all parties.

It may well be that momma chimps love their babies, though they also seem to get over losing one since mortality rates are relatively high and grieving isn’t exactly a luxury they can afford. Standard chimpanzees (not bonobos) do not seem to have any great affection among males and females in general. Mating is mediated by pheromones during a female’s estrous period. Afterward the pair doesn’t seem to stay a pair and mothers raise their young independently of fathers. For bonobos the situation is different, as far as mating is concerned. They are sometimes called the “make love not war” species because the engage in promiscuous sex that seems to augment the place of grooming activities in their cousins. It isn’t known for sure what a bonobo “feels” toward others in the group. Perhaps it is something like affection. But it is not likely the kind of love that humans have evolved. More research on both human forms of love and bonding in other animals is needed to be able to say much more.

One thing is certain. The feelings of deep affection one person can have for another (and hopefully mutually shared) is a source of strong sociality.

Communications

All animals seem to communicate with conspecifics in various ways. Chemical transmissions, sound, and coloration are used to send messages between individuals or groups. Some apes have been ‘trained’ to use symbols to convey their mental states to human researchers. Some cetaceans are thought to communicate more complex messages via their sonic vocalizations (e.g. whale songs). But, as far as we know, no other species communicates complex thoughts about abstract concepts like subject-action-object and with embedded phrases of the same form. No species has words representing real world things, let alone words that represent ideas or feelings.

The origin of human language is still somewhat mysterious. No language fossils exist to allow us to date when language became a regular part of human social life. Many guesses based on other behavioral correlates have been made. What is clear is that language was in full form by the time sapiens came to be. There is some speculation that other late species such as the Neanderthals had some basic language competence, but much more evidence, such as preserved voice boxes, would be needed to resolve the questions.

From a systems perspective of evolution[254], namely the roles of auto-organization and emergence, it is most likely that language emerged from the interactions that the expansion of BA10 had with the other prefrontal cortical areas and from there to the other associative cortical areas. In particular, of course, Broca's area in the posterior frontal (left side) lobe and Wernicke's area in the left temporal/parietal lobes, already poised to process conspecific semantic verbalizations responded to the increases in innervation from the prefrontal cortex (see figure 4.17) where abstract concepts, their causal relations, and projections of future states of the world were being generated.

In other words the language facility coevolved with the other facilities of sapience.

All of these unique characteristics, ToM, empathy, moral sentiments and strong cooperativity, love, and language-based communications are mutually reinforcing and contribute to the success of eusociality in humans. Once launched in their various nascent forms they helped human beings become increasingly successful in adapting to rapidly changing environments while remaining relatively weak and naked apes. Humans succeeded in competition with all other species because they cooperated and collectively solved problems. *Sapiens* succeeded best of all. Over the last 100 thousand years (or less) *Homo sapiens* emerged as supremely successful, supremely fit in every environment into which they migrated.

Of Grandmothers and Wisdom

Sometime long ago, perhaps before the human line and the chimpanzee/bonobo lines split, child rearing among apes underwent an interesting expansion such that other related females, in an extended family framework, began to assist mothers in raising infants and juveniles. Several related changes in family life among the apes that would eventually give rise to humans were under way as well. First the life span of the animals got longer such that several generations were living, and even breeding, at the same time. But time between pregnancies became extended as well, fewer offspring over time, and tribal living with social dynamics mediating interpersonal relationships emerged[255]. There is an intriguing hypothesis that grandmothers, who were living longer and had personal relations with their daughters and their daughters' young, began to assist directly in the day-to-day care of the young. The evolutionary advantage of this arrangement was that mothers were better able to scavenge for food than the older, presumably more sedentary,

---

[254] Chapters 10 and 11 in Mobus & Kalton (2014).

[255] See, for example, the Wikipedia article: http://en.wikipedia.org/wiki/Grandmother_hypothesis

females and could better support the group effort in finding resources, reflecting the increasing mutual intentionality covered in the last section.

Somewhat later in human evolution older females lost the ability to reproduce (menopause) yet were still able to care for the young children of their offspring. This may have been an entrenchment of the extended family view above, but something else was occurring in the reproductive habits in the human line that were not seen in the other apes. Namely fathers were investing in care for the young as well. The act of mating carried with it a commitment by fathers to support the family. Human children, because they are born in a much more immature state, need much longer periods of care than, for example, ape babies. A contribution from fathers would seem to argue for a weakening of the grandmother-care requirement. But, I suspect, yet another factor was taking shape in human evolution that provided a strong selective advantage for grandmother, and grandfather, involvement in child rearing.

Human children were requiring much longer development times in order to learn more about their world to be successful adults. They did not require the coddling care of an infant, but did need something much more important for long-term success. They needed to gain knowledge of what had become a much richer and more complex world of social affairs and the environment around them. Grandparents, I submit, provided youngsters with extended education in intricacies that they had learned over their lifetimes. They provided cultural transference and *wisdom*[256].

And such learning extended well into reproductive adulthood. Learning included not just fact knowledge, but a tremendous amount of tacit knowledge about how other people work and think. It required learning how to plan for longer spans of time and to consider many variations on what the future might bring. It required drawing on the stored wisdom of the elders until such time as the individual amassed sufficient tacit knowledge to become one of those elders.

Sometime within the last several hundred thousand years the tiny patch of cortical tissue, right behind the eyebrows on the prefrontal cortex, expanded and developed to provide the processing capabilities for enhanced strategic thinking and interpretation of moral sentiments. I submit that so strong a selective advantage was this innovation for humans that it became the most rapidly evolved aspect of human mentation. Sapience became a predominant factor in human success.

The Evolution of Spiritistic Thinking

In chapter 2 I introduced the notion of a surprising consequence of the rise of sapience in *Homo*, namely what I have called *spiritistic* thinking. In evolutionary terms this kind of thinking is a *spandrel*[257]. That is, it was not particularly adaptive in itself when it emerged, but rather a behavioral side effect of the other components of sapience that gave rise to 2½ order consciousness and were adaptive.

---

[256] c.f. Tomasello (2019), esp. chapter 5.

[257] See Gould (1997). Available at: http://www.pnas.org/content/94/20/10750.full

However, there are now good arguments supporting the notion that once in operation in the context of group selection, spiritistic thinking provided a way for members of a group to solidify their bonds by holding and maintaining spiritualistic notions (as described in chapter 2) that tended to reinforce the "specialness" of the group. Thus what was a behavioral result of the species just crossing the threshold of sapience, but not evolving a sufficiently strong form of sapience was co-opted by cultural evolution in a manner that probably reinforced its maintenance as a way of thinking in individuals. The results are what we call religions[258]. It may indeed be argued that the ability to hold spiritualistic notions made possible the creation of groups larger than Dunbar's number wherein members holding the same unsubstantiated beliefs saw themselves yet as members of this extended "tribe."

Spiritistic thinking is, I assert, provides more evidence that sapience is still weak in the brains of modern humans. If everything else I have argued about systems and strategic thinking is valid, then one has to believe that religious beliefs that are based on imaginative stories would not arise in the truly sapient mind. True wisdom relies on evidence and veridical models of the world. It is not faith-based in spite of arguments to the contrary.

Why Sapience Strength is Probably Not Normally Distributed

The idea of sapience strength hinges on a concept similar to intelligence strength. Intelligence researchers often categorize intelligence into general (g) and specific (e.g. logico-mathematical). General intelligence is further categorized into fluid intelligence, which includes psychometric characteristics such as speed and retention of memories and recall, memory capacities such as that of working memory, and so on. Crystallized intelligence is essentially the knowledge base itself, what one has learned and knows and how much one can demonstrate that one has learned? Psychologists attempt to assess the strength of several of these categories when measuring a generalized intelligence quotient (IQ). Sapience is using the same kind of brain machinery as intelligence so it isn't completely odd to suggest that there is a similar set of measures of strength of that facility. For example, the idea that there are general and special forms of sapience is quite justifiable.

And just as with intelligence, sapience is not uniformly distributed in the population. We now realize that intelligence, and this goes for creativity as well, takes on characteristic distributions that appear to fit normal or near normal curves (the famous, or infamous, bell curve). How people come to have a specific IQ rating is still hotly debated, though the evidence strongly suggests it is very much a heritable trait (at least 50%). The rest might be influenced by environment (nurture). Nevertheless, measures of IQ show remarkable consistency in terms of the normal distribution (where the peak represents the mean and is arbitrarily assigned a value of 100, meaning that someone with an IQ of 100 is performing on tests at their age level).

---

[258] See Bulbulia, et al (2008) and Atran (2002) for arguments regarding the evolution of religious thinking and the cultural benefits of religions.

Sapience could prove to have a similar mix of nature and nurture. But I also suspect that sapience might not share the quality of normal distribution, however. We have to explain an observation from the modern world that seems to hold. Wisdom is not common[259]. Sapience, in its capacity to produce wisdom, while clearly operative in some individuals, seems weak in most people in modern civilization. There are no data, of which I am aware, regarding relative sizes of BA10 from either anatomical (postmortem) or MRI studies but there seems to be a distinctly skewed aspect to the distribution of sapience in the population. The vast majority of people do not display strong capacity for good judgment and an ability to think far ahead (higher strategic level). Indeed, some psychologists have suggested that most people do not participate in self-reflection very often, the latter a result of 2½ -order consciousness that is one of the attributes given to wise people.

This is, of course, an anecdotal observation, not a scientific one. But it could be worked into a testable hypothesis. There are psychological tests of judgment and even some purported tests of wisdom that might be used along with correlational studies of fMRI images of the prefrontal cortex. It would be interesting to see if BA10 shows any particularly strong activity with some of these tests[260]. Then it would be interesting if there is sufficient image resolution to detect variations in the patch area activated. The hypothesis is that a majority of people will give poor judgments to particularly difficult moral judgment tests and that these will correlate with smaller patches of BA10.

---

[259] Sternberg (1990b)

[260] And to see what other parts of the brain are active during a mental act requiring wisdom!

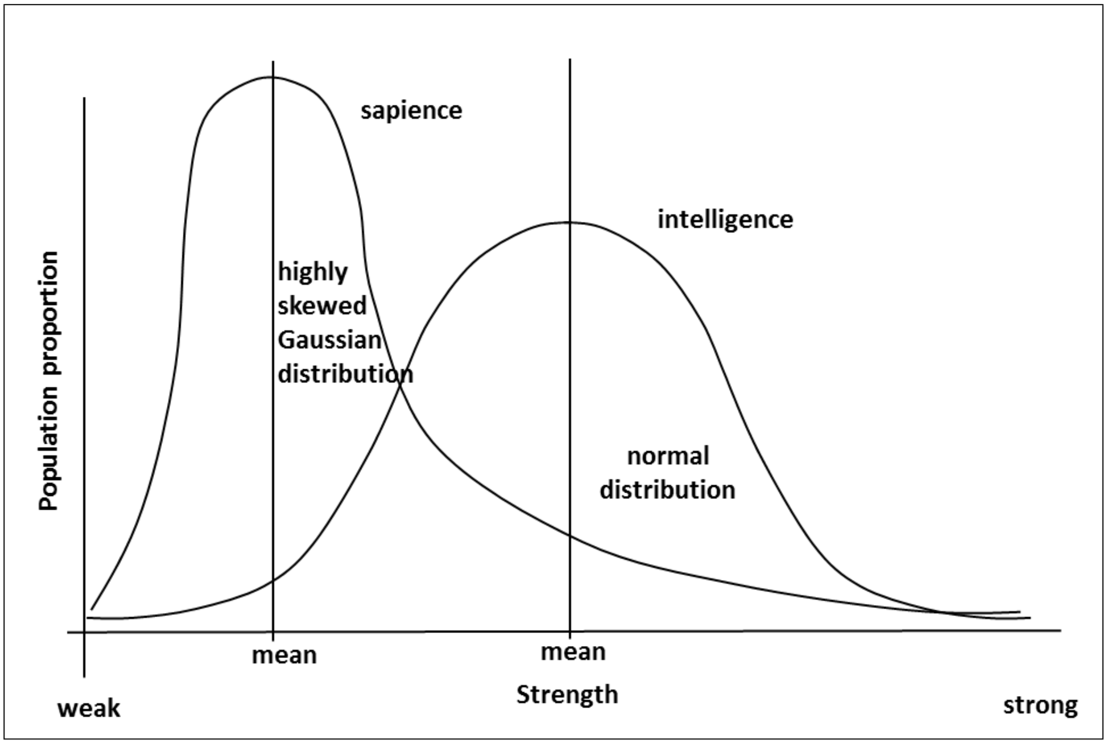


Figure 5.5. Intelligence is distributed in a normal Gaussian form where the majority of people are toward the center of the distribution (mean ~ mode). For newly emerged attributes in evolution, there is a greater tendency for the population to cluster toward the weaker end of the scale, producing a highly skewed Gaussian-like curve. This is a hypothetical view of the distribution of sapience (compared with intelligence) in the extant population. Over a long enough period of time, with selection favoring stronger sapience, the bulge would be expected to migrate to the right and eventually come to look like the normal distribution for intelligence.

The superior sapients were the wisest men and women in the Late Pleistocene tribes. They were the ones who had the proclivity to think about the long-term, and think about the larger scale of the world, particularly about the surrounding tribes and their ‘personalities’ (friendly, aggressive, etc.). Strong or superior sapience appears to be rare in our species today. Early humans, in whom sapience newly emerged, lived in small groups and there is now thought to be a basis for claiming that group selection was a strong force in shaping the social milieu of human culture. Groups that were lucky enough to have a few individuals with brain development genetics that favored greater expansion and development of the prefrontal area of the cortex and the other members of the group having just enough average sapience to be just wise enough to listen to those individuals (that is, wise enough to recognize superior wisdom when given) tended to fare better than other groups. Thus sapience need not have been selected for within groups, but would have had a strong selection advantage between groups. Such selection skew would result in there being fewer individuals with the "right" genetic traits (see below) in larger merged populations, say after the advent of agriculture.

Why the Evolution of Sapience May Have Come to a Halt

While the distribution of sapience strength may have always been skewed, nevertheless the early socio-cultural/physical environment might have still provided selection pressures that would have tended to bring the average sapience upward just as it did for higher intelligence and creativity in earlier hominin evolution. Then, about ten thousand years before the present that superior intelligence found a new way to acquire food, most importantly cultivated grains that were readily stored for longer periods. And that one ‘technological’ invention changed everything[261].

Throughout the Neolithic period humans had been increasingly subject to a new evolutionary pressure, not from the climate or other animals, but from their own developing technological developments and other cultural aspects. Cultural-biological coevolution began to play a larger role in the further development of biological and mental traits. With the advent of agriculture on a large scale, the changes in selection pressures due to cultural practices became the dominant forces shaping the human future.

The first major change involved the settling of tribes in fertile areas where crop reliability provided a certain amount of certitude in provisioning. But the settling also meant that tribes became responsible for claiming that land and protecting their claims against other tribes. ‘Property’, a greatly refined cultural concept derived from the biological need for ‘territory’ became necessary for people to consider. Protecting property required tactical thinking revolving around possible conflict, possibly heightening the sense of competition between groups[262].

Another major change, a result of having more food available over time, was an expansion of the populations in these settled regions. As populations expanded beyond the Dunbar optimum (150-200) new ways of thinking about other people became necessary. An ability to cooperate with strangers put significant emphasis on interpersonal relations and an individual’s capacity for building ToM models of other people. The individual’s tactical models for how to work with others (both cooperators and cheaters) became far more important than simple kin interactions. Politics was born. Religious thinking enabled the larger group sizes, which may be why up to the current day the link between politics and religion remains so pervasive. At the outset state governments and state religions went hand-in-hand to manage the larger populations.

Along with increasing populations the need to coordinate the work of the community increased considerably so there was an increased pressure to favor logistical talents at the group level.

---

[261] Scott (2017) provides an up-to-date description of the domestication of grains in Mesopotamia by humans and the domestication of humans by grains!

[262] Hodder (2012) describes how humans become ‘entangled’ in networks of relations upon which they come to depend. He traces the growing complexity of human entanglements through time from simple Neolithic villages to the modern world.

With the advent of agriculture a new group selection pressure emerged that favored more intelligence and creativity, along with more focus on power hierarchy and control. Planting and harvests needed to be managed and marauders fought off. The organization and management of agriculture demanded skills and thinking quite different from what had traditionally been considered wisdom. Once crops and animals had been domesticated, their care became more or less routinized. The need was to have a very large operational level control structure that would need more strict coordination, more emphasis on top-down command and control, and less strategic-level processing. The need for long-range thinking and planning diminished. And as it did so the favorable selection of greater sapience faded away.

As civilizations evolved great emphasis was placed on ingenuity and invention for the purpose of exploiting natural resources. Humans, like all animals have built-in drives to take advantage of the good times (gorge themselves in times of plenty) in order to survive the inevitable bad times. But with the development of human cleverness and social organization, humans began to diminish the frequency and amplitude of the bad times. With the growing of grains and storage in granaries for supplies during the non-growing seasons or in droughts, humans learned to hedge against uncertainty and improved their reproductive fitness tremendously. Under these conditions, the need for really long-term planning was, surprisingly, minimized. The only real planning horizon that counted was the seasons of the year. Humans invented clocks and calendars to help them plan the activities associated with planting and harvesting. The level of sapience in the most sapient members of society was more than adequate for this activity. Sedentary lifestyles and increasing specialization in tool making and work greatly diminished whatever selection pressures had been pushing humans toward greater sapience. Emphasis on producing more products than was actually needed — profits — and trading that for niceties, if anything, established a selection pressure on greediness and diminished the selection of stronger moral sentiments toward sharing. Humans learned to seek expansion rather than contentment with the status quo. As villages grew into cities, and cities grew into civilizations the size and complexity of societies grew exponentially and the modern human mentality emerged without much need for increasing the average wisdom of the average member of society. The coevolution of the human genome and the human culture favored humans with superior tactical and logistical skills in thinking. Wisdom took a back seat.

Rapid Evolution of the Prefrontal Cortex and Implications for the Future

Despite the diminishment of selection for greater sapience in the whole population due to the coevolution of culture and the mind, this has turned out to be a non-optimal situation. Man has been so clever and greedy and hooked on novelty seeking that he has created a world impossible for anyone to fully understand. Our culture and its institutions, inventions, and procedures, has taken on a life of its own. It has become an immensely complex system that contains individuals who are no longer able to grasp the meaning of the whole. And as a result we are making serious errors in judgments at all scopes. It seems to me, we have created a situation in which, once again, the selection for greater sapience may be coming to the fore.

But how is this to be worked out in time to avoid the extinction of our kind due to our poor judgments? Natural selection, even group selection, will be far too slow to produce results. Indeed, the problem is that the most sapient people on the planet have to outbreed the least. And there is no evidence that this is occurring. But there might be another option, which I outline below. What we have to recognize first is that a mechanism for rapid brain evolution is already in place.

One of the more interesting aspects of the evolution of the human brain is the rapidity with which the frontal cortex expanded relative to other cortical lobes, and the extremely rapid increase in size of the fronto-polar prefrontal cortex in evolutionary terms. Humans became anatomically modern a mere two hundred thousand years ago (roughly) and mentally modern around one hundred thousand years ago or less. The rapid expansion of the BA10 and its influence on the other parts of the brain with which it communicates suggests something besides ordinary genetic mutation and selection as the underlying mechanism for this development.

In recent years a new understanding of the genome has emerged. This new picture does not regard protein-coding genes as the only form of inheritance[263]. Indeed, what used to be called 'junk' DNA is now being shown to include large numbers of segments that code for short RNAs, some of which play a role in gene expression regulation through epigenetic mechanisms. Some segments of DNA were already known to play a role in regulation of gene expression (switching on and off), but it has turned out that there is an elaborate network of regulatory segments that act as a kind of developmental program that is activated by and affected by elements in the environment. The new synthesis of embryonic development and evolution is called Evo-Devo[264]. One of its principles is that various genes are turned on or off during development based on where in the embryo a particular cell sits in the overall body plan. This is a solution to the problem of cell differentiation during development. Every cell in the body has the same complement of genes, but only certain ones are active in cells from different tissue types. It turns out that the timing of turning genes on or off has the effect of controlling the overall development process and the resulting morphology. It helps explain why species with very different morphologies actually have genetic complements that are fundamentally the same. Turning the same genes on or off at different times in the development program lead to different body forms. And in terms of brain development it can lead to different competencies.

This model resolves two puzzling problems that we have had ever since the human genome was decoded. First, how is it that humans only have between twenty and twenty-five thousand protein coding genes (some simpler organisms actually have many more!)? This is considered a paucity of genetic material for specifying the complexities of the human brain. Second, why do the chimpanzees and humans share so much of their genomes in common, yet look and behave so

---

[263] Jablonka & Lamb (2005)

[264] Carroll (2005, 2006) walks us through the new vision of how evolution and development are linked and what it means in terms of fitness and speciation.

differently? What Evo-Devo provides us with is a possible solution to these seeming problems. First it suggests that it isn't the proteins per se that are important in cell differentiation, but rather the timing of when and which ones are turned on or off, and, likely, their combinations in the cells. In other words, it is the control program embedded in the non-coding DNA that can make all the difference in the world. A slight change in a control circuit can have a drastic impact on form and function of cell types. Indeed whole new cell types might easily derive from existing ones with just very minor changes in one little piece of control DNA. Actually it is probably more the case that one little change can impact numerous cell types since the cells are using the same complement of proteins, but simply using them in different contexts and at different times during development.

This is essentially an amplified effect. A small change can have many separate impacts on the form and function of the whole system. It could also explain the apparent increase in speed of evolutionary changes, especially in the brain. Recent research on control segments responsible for brain cell development has shown that these regions have been undergoing extremely rapid evolution[265].

A key question, then, is this: Is there a small number of DNA segments that control the difference between levels of sapience in humans? The reason I focus on genetics rather than environment is that we already know that, while wisdom development is correlated with intelligence, and the latter is somewhat impacted developmentally by the child's environment, still we have the puzzle of having brilliant people who are terribly foolish[266]. Ergo, while environment may play a larger role in the development of intelligence, I suspect it plays a much more minor role in the development of sapience. If it turns out that there is such a set of segments that regulate the timing of gene expression in, especially, Brodmann area 10 during development, then we may have found a leverage point for the future of evolution.

The Future of Sapience

Would Humans Be 'Better Off' With More Sapience?

Today there is much speculation about 'designer babies', where parents pre-select the genetic attributes of their to-be-conceived children. Talk revolves around improving attributes that parents would like their children to have, such as athletic build, and especially greater intelligence. The thinking driving the latter desire is based on believing that intelligence is the end-all and be-all of mental prowess. As I have been arguing throughout this book, intelligence and creativity are not the ultimate heights of human mentation. Sapience is what makes us uniquely human. It is true that we have greater intelligence and creativity as compared with our

---

[265] See "Brain Evolution: Neurogenomics Targets the Genes That Make Us Human", (2014) BrainFacts.org, http://www.brainfacts.org/brain-basics/evolution/articles/2014/brain-evolution-neurogenomics-targets-the-genes-that-make-us-human/ (Accessed: 1/11/2016).

[266] Sternberg (2002)

hominin cousins, but, as I have said, these are as much a part of the problem with the modern human condition as they are part of the potential solutions. I don't think we need more, smarter people to create cleverer but ultimately ill-advised stuff. What we need is better long-term, morally motivated judgment. What we need is better systemic understanding of the whole of the Ecos. What we need is more comprehensive strategic thinking for the world, not just our states or nations. In other words, if I could specify my designer baby, it would be to expand the capacity of the prefrontal cortex so that the child and later adult would develop wisdom over their lifetime.

Suppose sapience were distributed normally as is the case for intelligence. Further suppose that it had undergone a general strengthening under evolutionary selection pressures so that the average human being would develop what we today would easily recognize as real wisdom. Real wisdom is so rare today that when it is encountered it stands out starkly. But imagine that the average person, when older, exhibited such wisdom.

Furthermore, imagine that stronger sapience was active in developing youth and young adults. They would not have accumulated a lifetime's worth of tacit knowledge, though they might absorb by observation of the older adults' behaviors, so they might not exhibit what we normally think of as wisdom, per se. But I rather suspect their tendency to make better judgments, and hence better decisions, would result in a very different set of attitudes and behaviors early in life. For one thing, I suspect that as in the days of the small tribes, the young would be more prone to attend to what the wise elders had to say. The youth would still make mistakes, of course, because they would still be motivated by primitive limbic impulses. And those impulses might not be completely controllable until the frontal cortex undergoes its later maturation, just as is the case now. Nevertheless, I suspect that overall, or on average, the more sapient youth would tend to show earlier signs of wisdom than anything we are used to today.

I would argue that life would be much different for a more strongly sapient society. I don't think they would be any less clever or inventive. But I do suspect they would think long and hard about the judiciousness of exploring inventions that were two edged swords. Atomic energy might be a good thing if carefully developed and managed. But atomic weaponry has proven to be one of our greatest blunders and haunts us to this day. Might a very sapient human have learned over life that every new invention carries with it one or more side effects with costs, some of which might not be payable? Might they then counsel caution in developing a technology that could lead to future destruction?

Finally, might not a population of higher sapient beings be inclined to employ sapient governance in their economic and between-group affairs?

Will Nature Favor More Sapience?

We modern humans are actually not that far from some of what I have described in this book. We are weak when it comes to judgments that integrate moral sentiments with strategic thinking, but we do have the beginnings of those capabilities. And some members of our species have demonstrated a greater capacity for doing so, implying that the underlying sapience — the brain mechanisms necessary — is represented in the gene pool. That is to say, the genetic basis for greater sapience very likely already exists in a few individuals now. If such individuals were to interbreed and the resulting genome breeds true in subsequent generations we could have the onset of an incipient new species of humans that displayed all of the traits I have described. So, in effect, the raw material is already here.

The problem is that there has to be some way for such individuals to find and recognize one another with a mating preference strongly attracting them to one another. This might take the form of sexual selection and there is actually some preliminary evidence that such selection is taking place now, but at a very diffuse level. Too, the nature of our modern (western) societies and cultural complexities make recognition a hit or miss proposition. Our tendency toward nuclear, mobile families rather than extended families with grandparents in residence further weakens any of those former natural selection pressures described above.

It seems to me that the answer to the question is that nature, alone, will not provide the needed selection pressures in time to avert a total catastrophe to the human species. The race is just too fast.

Much as it pains me to write this, I think that the evidence that human beings are headed for collapse is substantial[267]. In a worst case scenario this leads to a complete collapse of not just civilization but a major population crash. Part of the problem stems from our running out of high grade fossil fuels and net energy to maintain our civilization[268]. Part of it stems from the complications arising from drastic climate changes. The future looks dim for us and for a large percentage of other species as well. Should such a collapse occur the question would then be what would nature favor in the survivors? After every major die-off, evolution produces an efflorescence of new species to fill empty eco niches. What niche would any human survivors, assuming there are some, fit into?

In an impoverished world such as is imagined an argument that our more bestial attributes will lead to more fitness than our higher mental capacities can be, and often is, made. Is this a valid assumption or just the result of too many Mad Max movies? The truth is we don't know and probably can't know what traits might be favored (or if any humans could survive at all). Biological evolution is a craps shoot. It can go any number of ways and chance plays a major role in determining who or what gets a chance to shoot again[269]. There is an argument, from systems science, that if there is sufficient flow of free energy then the whole system should be

---

[267] Catton (2009)

[268] See Hall & Klitgaard (2011) for an explanation of the role of energy in our economies and particularly the fate of fossil fuels in the looming future.

[269] Losos (2017), Smith & Morowitz (2016) provide enlightened approaches to understanding the roles of chance and necessity in biological evolution.

capable of rebuilding complexity (society in the case of humans) in some fashion. What we do not know is whether or not such a flow would be obtainable after a major crash of civilization.

But natural selection, sexual selection, etc. might not be the only routes to further evolution and some future speciation. Being the first species of animal to have understood evolution for what it is and how it works, as well have having begun to grasp the nature of genetics at a molecular level, we are in the unique position to engineer an intentional selection pressure that might just help bias future evolution in the direction of greater sapience.

Can We Humans Intervene To Force the Selection of Greater Sapience?

In a sense we already know how to breed plants and animals for qualities we favor. We act as the selection force and since we take measures to prevent breeding by non-conforming members of the population of interest, we accelerate the effects of selection (artificial selection). We have already set up the conditions for speciation in, say, dogs. If we were to eliminate all other breeds of dogs except Chihuahuas and Great Danes we would have the necessary separation (through the sheer physical impossibilities of mating across the two breeds) for allopatric speciation to occur, given a bit more time. It wouldn't be hard to imagine a day when the two would not even recognize one another as conspecifics and only be interested in breeding within their kind (it would probably require changes in the pheromones that each exudes as well as the difference in size).

Most people don't care for the idea of breeding people, since it raises the specter of coercive eugenics[270] and associations with Adolph Hitler's ideas of a master race. It would be good if people could remember that Hitler's and his minions' motives were strictly political and had nothing to do with science; they simply put their tactics under the veil of a scientific concept. Also when talk of breeding out unwanted traits was recognized as just another form of bigotry the whole idea of breeding human stocks took on a sinister aura. The real issues, of course, were the choices of what to breed for or against and who was making those choices. In suggesting that higher sapience is a positive, desirable quality, I am, naturally, making a similar subjective choice. But there is a huge difference. Everything that I have written here about sapience, its qualities, and how it is produced, is scientifically testable without any kind of coercive techniques. Plus I am not suggesting any kind of forced breeding program. Rather, what I am suggesting is an assist to assortative mating[271] that will help concentrate the prospective high sapients within a region and then let nature take its course. Breeding will be voluntary and mutually-selecting. But first must come the real science.

It does little good to only examine older individuals for traits of wisdom as a gauge of sapience strength. Women past menopause wouldn't be able to mate, and men sufficiently mature to

---

[270] See the Wikipedia article on eugenics: http://en.wikipedia.org/wiki/Eugenics for more background.

[271] See the Wikipedia article on assortative mating: http://en.wikipedia.org/wiki/Assortative_mating for more background.

exhibit wisdom would be less attractive to women in their prime (not to mention the possible need to stock the shelves with Viagra!) One could, of course, infer from wise behavior that the sage or crone so behaving possesses the right genes and so their offspring should as well. But there are several problems with this approach. For one, wise people, these days, may tend to have fewer offspring because they have realized what is happening in the world due to overpopulation and/or thinking their offspring will have to pay the price for the foolish excesses of their own generation. A second problem is that there is no guarantee that offspring will actually carry the exact or even a partial complement of the right DNA. Everybody has two parents and in our society there is no guarantee that both will have the DNA needed. It would take many, many generations of intensive breeding of the offspring of the offspring before we would even begin to concentrate the needed DNA into a pool. And we've already agreed no eugenics!

Fortunately, if my hypotheses are correct about the prefrontal cortex size and development owing to the 'right' DNA segments, then we have several methods at our disposal for working backwards from studies of wise behavior correlated with size, etc. of the prefrontal cortex and the genetic components that regulate the development of it (and particularly Brodmann area 10). Research on the timing and activation of genes during development of the prefrontal cortex is currently underway. It is not a difficult step to take to sequence the DNA segments associated with both the genes themselves and the control network snippets mentioned earlier. We should be able to find markers (possibly several) that correlate highly with wise behavior in adults in later life! In other words, just as we do genetic testing today for disease potentials or risk factors for diseases, we could test a subject's DNA for the potential for high sapience. Then tell the subjects the results. Let them decide what to do with the information.

Along with this testing we would set up a 'sapience matching service' similar to on-line meeting services now. Counseling high testing sapients that they would be doing us all a service by having as many children as they wanted (and please out-compete the low sapients if you can!) would help accelerate the accumulation of strong sapience DNA. Better still would be providing financial support for high sapient couples to raise really big families.

Unfortunately even this amount of assisted intentional selection might not be sufficient to make a difference. If there is a general population crash, as I suspect strongly will be the case, there would be no gain from assisting high sapients to meet and mate. But if there is a crash it also means that there will be an evolutionary or population bottleneck event as has apparently happened before. A prudent Plan B would be to not only assist high sapients in finding one another but also providing for a secure colony with all the necessary technology and energy supplies they would need to survive the event and emerge in the new world with some semblance of civilization (especially knowledge of essential artifacts and processes), what has been called a "lifeboat" approach. Such a breeding population, though small, would then be in a position to out-compete any other surviving lower sapients, with the hope that they would generate the next incipient species of the genus *Homo*. Call it an "ark for humanity."

Of course nothing, no amount of planning, no high minded intentions, or anything else, could guarantee the success of such a program. There are no guarantees in nature except that you will eventually die one day. So too, all species are guaranteed to go extinct eventually. Ours is no exception. But the extinction of one species does not need to mean the extinction of the whole genus. Our case is desperate. We are the only living species within our genus, and our genus is the only known one that has achieved second and a half-order consciousness, complete with abstract symbol representation and manipulation for communication of complex ideas.

What a waste of potential and a shame it would be to simply stand by and let our species go extinct or worse yet, devolve into something we would consider sub-human. We would be demonstrating at least one answer to Fermi's Paradox. What if, throughout the galaxy hundreds, even millions of sentient beings faced this same threshold and failed to act. It would explain why we have no evidence of intelligence in the universe besides ourselves. And whatever evidence we might have accounted for is about to be nullified. My personal preferred (admittedly science fiction-based) belief is that those sentient beings who made it to higher sapience are simply laying low and possibly watching to see what happens here on Earth!

I firmly think there is a real potential to achieve a higher level of sapience in our genus. We know there is because we have seen rare glimpses of individuals with the capacity to display real wisdom. The challenge facing mankind is to preserve that potential. Just as when each person realizes and accepts their own mortality and prepares for the wellbeing of their progeny survivors, we as a species probably need to recognize and accept our collective mortality and make preparations for our survivors if there are to be any at all. Are we wise enough to do this?

The Next Major Transition

Earlier I mentioned, in passing, the fact that evolution has produced new levels of organization through the emergence of new structures built from associations of existing entities in new relations. It is now recognized that rather than competition being the main driver of innovation in evolution producing such new levels, that cooperation between disparate entities leads to emergent properties and functions. The cooperation of some forms of early bacterial life led to the emergence of eukaryotic cells. The divergence and specialization of functions (e.g. between somatic and germ cell lines) among cells in early multicellular clumps, and the subsequent cooperation among them, led to multicellular life forms.

John Maynard Smith and Eörs Szathmáry (1995) wrote *The Major Transitions in Evolution* in which they traced these emergences of more complex forms over evolutionary history[272]. Each transition led to new structures, composites of previously existing lower complexity structures, in which the component parts achieved high degrees of inter-component communications that

---

[272] See the Wikipedia article: http://en.wikipedia.org/wiki/The_Major_Transitions_in_Evolution for a listing of the transitions they identified. Also see: Calcott & Sterelny eds. (2011) for some more recent views on the transitions theme.

facilitated cooperative inter-component activities. The communications were part of the hierarchical cybernetic subsystem discussed at the beginning of this chapter. These new structures performed new behaviors, new functions, and served new purposes, not predictable from the behaviors or functions of the independent components taken alone. And those behaviors, etc. were tested against the selection forces generated in this new higher level of organization[273]. They identified eight major transitions from the macromolecular level at the origin of life to the social communities of primates. Harold Morowitz (2002) started with the emergence of matter itself and enumerated twenty-eight "steps" or "emergences" of higher organization that has led to the current state of affairs. His scheme teases out more detail than Smith & Szathmáry but the pattern of auto-organization, emergence, and subsequent selection processes is the same. In a similar vein, John Stewart (2000) highlights the evolution of cooperativity and evolvability over the history of the Universe, leading to the stage we find ourselves in currently. He posits that the next transition is imminent. More recently, Tyler Volk (2017) describes a recurrent process of this kind of cooperativity leading to new levels or stages of evolution. He calls the process "combogenesis" and the stages form what he calls a "grand sequence" of increasing complexity (nestedness). In a forthcoming book I will be describing essentially the same process that I have termed "ontogenesis". Clearly, many authors are realizing that the path to increasing complexification is a real outcome of universal evolution.

The key consideration in all of these transitions is that the component parts that associate find ways to "communicate" with one another to form stable configurations and process collective inputs to outputs (functions) cooperatively. And those outputs are found by the environment, the other entities with which the new forms interact, to be of utility (purpose). If either the inputs become unavailable or the outputs are not acceptable by the environment (i.e. the outputs become inputs to other entities) then the new form will not be supported and will ultimately disintegrate. For example, if a system produces too much of a waste substance that is not absorbed by the environment, then that waste can accumulate and poison the system[274].

Communications or the transmission of messages between components that need to cooperate convey information of the type that allows components to coordinate their activities such that the overall function of the new form is maintained. However, at some level of complexity mere cooperation is insufficient, usually due to time delays in communications and responses. At a certain point in complexity (number and kinds of components interacting in the whole new form) there will need to emerge a coordinator function. Successful emergence of that coordinator means the form will continue to perform its function and fulfill its purpose in the environment. If

---

[273] Mobus & Kalton (2014), chapters 10 & 11 describe this emergence process, starting with what they call auto-organization (generally called self-organization), which the carefully define, leading to emergence of new behaviors (covered in chapter 6) and how the new structures/functions then become subject to classical selection pressures thereafter.

[274] I've written at length about the need for a hierarchical governance system to achieve long-term persistence, sustainability, and thrivability. See Mobus (2015, 2017) for details.

the coordinator does not emerge or is faulty, then the form will ultimately disintegrate from the same failure to adapt to the environment.

It is my feeling that we humans, with the level of sapience we possess are the beginnings of what could be the next major transition[275]. Our form of eusociality is unique in the magisterium of life and even in the hominid line. We have all of the attributes of components about to form collectives (societies) that operate as single entities. We have the motivations, the languages for extraordinary communications, and a collective ability to specialize and cooperate. Those are the gifts of sapience. But, as I have argued throughout the book, we are not the end product necessarily. We are just the beginning of a new level of organization, a new transition.

We still retain too much competitive urge, too much lack of wisdom, too little personal and inter-personal (shared) knowledge. We are still very much mostly biological. But we are on the brink of a transition to the supra-biological. At present we are mostly following our biological mandates to consume and expand. Hence the difficulties we are facing. But we are also capable of understanding that blind following of those mandates will be our demise. And, I strongly believe, with but a little more sapience, a little more wisdom of the collective, we will transcend our biological selves and establish a hierarchical cybernetic structure (governance of the society) that will take the place of natural selection and culling of the population due to predation and disease. That will be replaced by intentions to thrive within the limits of the Ecos. Mere biology does not work this way. It blindly absorbs as much negentropy as it can from the environment without regard for the consequences. That is the way of mere biology because the ecosystem itself contains such diversity of selection forces that no one population or species can capture all of the resources. That is, that was the case until humans and their incredibly clever minds came into being. We, as a species, broke out of the usual limiting factors that involved large-scale negative as well as positive feedback loops that keep other species in check (e.g. the famous Lotka–Volterra or predator-prey dynamic[276]). Through the seeming miracle of resource substitution we have managed to fend off, at least for a time, the resource limitations (e.g. Liebig's law of the minimum)[277] that have kept all other biological systems operating in cycles rather than continuous growth. But only for a time.

We are beginning to experience resource limitations because we have finally exhausted what the Ecos can provide to support a developed world standard of living. It is time to face the reality of the human condition as it is now and consider if our current capacity for sapience can lead us to establish our own intentionally imposed controls on growth and consumption.

My hope, and belief, is that humankind will enter into a new kind of evolution, beyond mere biological evolution, but evolution in which we become active participants – intentional

---

[275] See also this discussion of Posthumans on Wikipedia: https://en.wikipedia.org/wiki/Posthuman for another take on the idea of future human evolution. Accessed 4/22/2019.

[276] See the Wikipedia article: http://en.wikipedia.org/wiki/Lotka%E2%80%93Volterra_equation

[277] See the Wikipedia article: http://en.wikipedia.org/wiki/Liebig%27s_law_of_the_minimum

evolution. We will become both the source of variability with respect to sapience capacity, with efforts to expand that capacity, and the force of selection that will prevent us from overreaching. The proper combination of intelligence, creativity, affect, and sufficient sapience should help us become truly pan-social beings. As always in evolution there are no guarantees. There will always be random variations unpredicted that need to be accommodated. The environment of the Ecos will always change and require evolutionary adaptations. But the human genus will become evolvable and thus establish persistence and true sustainability.

Bibliography

References and Further Reading

Abdou, Kareem, Shehata, Mohammad, Choko, Kiriko, Nishisono, Hirofumi, Matsuo, Mina, Muramatsu, Shin-ichi, and Inokuchi, Kaoru (2018). “Synapse-specific representation of the identity of overlappinc memory engrams”, *Science*, **360**:6394, pp 1227-1231.

Alford, Simon T. and Alpert, Michael H. (2014). “A synaptic mechanism for network synchrony,” *Front. Cell. Neurosci.*, 18 September 2014 | doi: 10.3389/fncel.2014.00290.

Alkon, Daniel L., (1987). *Memory Traces in the Brain*, Cambridge University Press, Cambridge.

Andreasen, Nancy C. (2005). *The Creative Brain*, Penguin, New York.

Ariely, Dan (2008). *Predictably Irrational: The Hidden Forces that Shape Our Decisions*, HarperCollins, New York.

Ash, Robert B. (1965, 1990). *Information Theory*. Dover Publications, Inc.

Ashby, W. R. (1952). *Design for a Brain*, Chapman and Hall, London.

Atran, Scott (2002). *In Gods We Trust: The Evolutionary Landscape of Religion*, Oxford University Press, New York.

Avery, John, (2003). *Information Theory and Evolution*, World Scientific, Hackensack NJ.

Baars, Bernard J. & Gage, Nicole M. (eds.) (2007). *Cognition, Brain, and Consciousness: Introduction to Cognitive Science*, Elsevier, New York.

Baars, Bernard J. (2007). Chapter 10: Thinking and problem solving, in Baars & Gage (2007), pp 294 - 316.

Bateson, Gregory (1972). *Steps to an Ecology of Mind: Collected Essays in Anthropology, Psychiatry, Evolution, and Epistemology*. University Of Chicago Press, Chicago.

Beer, Stafford (1959). *Cybernetics and Management*, English University Press, London UK.

Beer, Stafford (1966). *Decision and Control: The Meaning of Operational Research and Management Cybernetics*, John Wiley & Sons, New York.

Beer, Stafford (1972). *Brain of the Firm: A Development in Management Cybernetics*, Herder and Herder, New York.

Berreby, David (2005). *Us and Them: Understanding Your Tribal Mind*, Little, Brown, and Company, New York.

Blackmore, Susan (2004). *Consciousness: An Introduction*, Oxford University Press, New York.

Block, Ned (1998). "On a confusion about a function of consciousness". In N. Block, O. Flanagan, G. Guzeldere. *The Nature of Consciousness: Philosophical Debates*. MIT Press. pp. 375–415, Cambridge MA.

Bourke, Andrew F.G. (2011). *Principles of Social Evolution*, Oxford University Press, Oxford UK.

Brodt, S., Gais, S., Beck, J., Erb, M., Scheffler, K., and Schonauer, M. (2018). "Fast track to the neocortex: A memory engram in the posterior parietal cortex", *Science*, **362**:6418, 30 November, 2018, pp 1045-1048.

Bulbuila, Joseph, Sosis, Richard, Harris, Erica, Genet, Russell, Genet, Cheryl, and Wyman, Karen (eds., 2008). *The Evolution of Religion: Studies, Theories, & Critiques*, Collins Foundation Press, Santa Margarita, CA USA.

Buller, David J. (2005). *Adapting Minds: Evolutionary Psychology and the Persistent Quest for Human Nature*, The MIT Press, Cambridge MA.

Burton, Robert A. (2008). *On Being Certain: Believing You Are Right Even When You're Not*, St. Martin's Press, New York.

Cacioppo, J.T., Visser, P.S., and Pickett, C.L., (2006). *Social Neuroscience: People Thinking About Thinking People*, The MIT Press, Cambridge MA.

Calcott, Brett & Sterelny, Kim (eds. 2011). *The Major Transitions in Evolution Revisited*, The MIT Press, Cambridge MA.

Calvin, William H. (1996). *How Brains Think: Evolving Intelligence*, *Then and Now*, Weidenfeld & Nicolson, London

Calvin, William H. & Ojemann, George A., (1994). *Conversations with Neil's Brain: The Neural Nature of Thought and Language*, Addison-Wesley Publishing Company, Reading, MA.

Carroll, Sean B., (2005). *Endless Forms Most Beautiful: The New Science of Evo Dvo*, W.W. Norton & Company, New York.

_____________, (2006). *The Making of the Fittest: DNA and the Ultimate Forensic Record of Evolution*, W.W. Norton & Company, New York.

Carter, Rita (1999). *Mapping the Mind*, University of California Press, Berkeley CA.

Catton, William R. Jr. (1982). *Overshoot: The Ecological Basis of Revolutionary Change*, University of Illinois Press, Urbana IL.

___________________ (2009). *Bottleneck: Humanity's Impending Impasse*, Xlibris Corporation, USA.

Chalmers, David (1996). The Conscious Mind: In Search of a Fundamental Theory, Oxford University Press, New York.

Coen, Enrico (2012). *Cells to Civilizations: The Principles of Change that Shape Life*, Princeton University Press, Princeton NJ.

Csikszentmihalyi, Mihaly (1997). *Creativity: Flow and the Psychology of Discovery and Invention*, Harper Perennial,

Damasio, Antonio R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*, G.P. Putnum's Sons, New York.

________________, (1999). *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*, Harcourt Brace & Company, New York.

________________, (2010). *Self Comes to Mind: Constructing the Conscious Brain*, Vintage Books, New York.

Deacon, Terrence W. (1997). *The Symbolic Species: The Co-evolution of Language and the Brain*, W. W. Norton & Co., New York.

Dennett, Daniel C., (1995). *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Simon & Schuster, New York.

_______________ (1991). *Consciousness Explained*, Little, Brown and Company, New York.

Dehaene, Stanislas (2014). *Consciousness and the Brain: Deciphering How the Brain Codes Our Thoughts*, Viking, New York.

DeSalle, Rob & Tattersall, Ian (2012). *The Brain: Big Bangs, Behaviors, and Beliefs*, Yale University Press, New Haven CT.

De Waal, Frans (2005). *Our Inner Ape*, The Berkley Publishing Group, New York.

_____________ (2010). *The Age of Empathy: Nature's Lessons for a Kinder Society*, Broadway Books, New York.

_____________ (2014). *The Bonobo and the Atheist: In Search of Humanism Among the Primates*, W.W. Norton & Company, New York.

_____________ (2016). *Are We Smart Enough to Know How Smart Animals Are?*, W.W. Norton & Company, New York

Dilworth, Craig (2009). *Too Smart for Our Own Good: The Ecological Predicament of Humankind*, Cambridge University Press, Cambridge UK.

Diamond, Jared (2005). *Collapse: How Societies Choose to Fail or Succeed*, Viking, New York.

Donald, Merlin, (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*, Harvard University Press, Cambridge, MA.

_____________ (2001). *A Mind So Rare: The Evolution of Human Consciousness*, W.W. Norton & Co., New York.

Edelman, Gerald (1987). *Neural Darwinism. The Theory of Neuronal Group Selection*, Basic Books, New York.

_____________ (2006). Second Nature: Brain Science and Human Nature, Yale University Press, New Haven CT.

Ehrlich, Paul R. & Ehrlich, Anne H. (2004). *One With Nineveh: Politics, Consumption, and the Human Future*, Island Press, Washington.

________________________________ (2002). *The Dominant Animal: Human Evolution and the Environment*, Island Press, Washington D.C..

Ehrlich, Paul R., (2002). *Human Natures: Genes, Cultures, and the Human Prospect*, Penguin Books, New York.

Findeisen, W., Bailey, F.N., Brdys, M., Malinowski, K., Tatjewski, P. & Wozniak, A. (1980). *Control and Coordination in Hierarchical Systems*, John Wiley & Sons, New York.

Flannery, Tim (2005). *The Weather Makers: How Man is Changing the Climate and What it Means for Life on Earth*, Atlantic Monthly Press, New York.

Fodor, Jerry (1975). *The Language of Thought*, Harvard University Press, Cambridge, MA.

Friedenberg, Jay & Silverman, Gordon (2006). *Cognitive Science*, Sage Publications, Thousand Oaks, CA.

Fuster, Joaquín (1999). *Memory in the Cerebral Cortex*, The MIT Press, Cambridge MA.

Gangestad, Steven W. & Simpson, Jeffry A. (2007). *The Evolution of Mind: Fundamental Questions and Controversies*, The Guilford Press, New York.

Gardner, Howard (1999). *Intelligence Reframed: Multiple Intelligences for the 21st Century*, Basic Books, New York.

______________ (2004). *Changing Minds*, Harvard Business School Press, Boston.

______________ (2007). *Five Minds for the Future*, Harvard Business School Press, Boston.

Gazzaniga, Michael S. (2005). *The Ethical Brain*, Dana Press, New York.

______________ (2005). *Human: The Science Behind What Makes Us Unique*, HarperCollins, New York.

Geary, David C., (2005). *The Origin of Mind: Evolution of Brain, Cognition, and General Intelligence*, American Psychological Association, Washington, DC.

Gilbert, Daniel (2006). *Stumbling on Happiness*, Alfred A. Knopf, New York.

Gilovich, Thomas, Griffin, Dale, & Kahneman, Daniel (eds) (2002). *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press, Cambridge.

Goldberg, Elknonon (2001). *The Executive Brain: Frontal Lobes and the Civilized Mind*, Oxford University Press, Oxford UK.

________________ (2005). *The Wisdom Paradox: How Your Mind Can Grow Stronger as Your Brain Grows Older*, Gotham Books, New York.

________________ (2009). *The New Executive Brain: Frontal Lobes in a Complex World*, Oxford University Press, Oxford UK.

Goldberg, Elkhonon & Bougkov, Dmitri (2007). "Goals, executive control, and action", in Barrs & Gage (eds.) *Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience*, Chapter 12, Elsevier, Amssterdam.

Goleman, Daniel (2006). *Social Intelligence*, Bantam Books, New York.

Gould, Stephen J. (1997), “The exaptive excellence of spandrels as a term and prototype”, PNAS vol. 94 no. 20, 10750–10755, doi: 10.1073/pnas.94.20.10750.

Griffin, Dale & Tversky, Amos (2002). “The Weighing of Evidence and the Determinants of Confidence”, in Gilovitch, et al (2002, pp 230-249).

Hall, Charles A.S. & Klitgaard, Kent (2011). *Energy and the Wealth of Nations: Understanding the Biophysical Economy*, Springer, New York.

Hanson, J. S., & Olson, C. R. (eds. 1990). *Connectionist Modeling and Brain Function: The Developing Interface*, The MIT Press, Cambridge, MA.

Harari, Yuval Noah (2015). *Sapiens: A Brief History of Humankind*, Harper, New York.

Hartley, R.V.L. (1928). "Transmission of Information", Bell System Technical Journal, July 1928.

Hawkins, Jeff (2004). *On Intelligence*, with Sandra Blakeslee, Times Books/Henry Holt and Company, New York.

Hebb, Donald (1949). *The Organization of Behavior*, John Wiley & Sons Inc. New York.

Heinberg, Richard (2007). *Peak Everything: Waking Up to the Century of Declines*, New Society Publishers, Gabriola Island, BC, Canada.

Hodder, Ian (2012). *Entangled: An Archeology of the Relationships between Humans and Things*, John Wiley & Sons Limited, New York.

Hogarth, Robin (1980). *Judgement and Choice*, John Wiley & Sons, New York.

Homer-Dixon, Thomas (2006). *The Upside of Down: Catastrophe, Creativity, and the Renewal of Civilization*, Knopf, Toronto CA.

Jablonka, Eva, and Lamb, Marion J. (2005). *Evolution in Four Dimensions*, The MIT Press, Cambridge MA.

Johnson-Laird, Philip (2006). *How We Reason*, Oxford University Press, Oxford UK.

Judd, J. Stephen (1991). *Neural Network Design and the Complexity of Learning,* Bradford Books, New York.

Kahneman, Daniel (1982) *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge UK.

_______________ (2011). *Thinking Fast and Slow*, Farrar, Straus, and Giroux, New York.

Kauffman, Stuart (1996). *At Home in the Universe*, Oxford University Press, New York.

Kitchener, Karen S. & Brenner, Helene G. (1990). “Wisdom and Reflective Judgment”, in Sternberg, 1990, Chapter 10, pp 212-229.

Klein, Naomi (2014). *This Changes Everything: Capitalism vs. the Climate*, Simon & Schuster, New York.

Koch, Christof (2012). *Consciousness: Confessions of a Romantic Reductionist*, The MIT Press, Cambridge MA.

Kramer, Deirdre A. "Conceptualizing wisdom: the primacy of affect-cognition relations", in Sternberg, Robert J. (ed.) (1990). *Wisdom: Its Nature, Origins, and Development*, (pp 279 - 313) Cambridge University Press, New York.

Labouvie-Vief, Gisela (1990). “Wisdom as integrated thought: historical and developmental perspectives”, in Sternberg (1990, pp 52 - 83).

Leaky, Richard & Lewin, Roger (1995). *The Sixth Extinction: Patterns of LIfe and the Future of Humankind*, Doubleday, New York.

LeDoux, Joseph (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, Simon & Schuster, Touchstone, New York.

_____________ (2002). *Synaptic Self: How Our Brains Become Who We Are*, Viking, New York.

Lehrer, Jonah (2009). *How We Decide*, Houghton Mifflin Harcourt, New York.

Losos, Jonathan B. (2017). *Improbable Destinies: Fate, Chance, and the Future of Evolution*, River Head Books, New York.

Lovelock, James (2006). *The Revenge of Gaia: Earth's Climate Crisis & the Fate of Humanity*, Basic Books, New York.

MacKay, David J. C. (2003). *Information Theory, Inference, and Learning Algorithms*, Cambridge: Cambridge University Press.

Marcus, Gary (2004). *The Birth of the Mind: How a Tiny Number of Genes Creates the Complexities of Human Thought*, Basic Books, New York.

_____________ (2008). *Kluge: The Haphazard Construction of the Human Mind*, Houghton Mifflin Co. New York.

Marean, Curtis W. (2010). “When the Sea Saved Humanity”, *Scientific American*, August 2010.

Margulis, Lynn & Sagan, Dorion (1995). *What is Life?*, Simon & Schuster, New York.

Maturana, Humberto R. & Varela, Fransico J. (1980). *Autopoiesis and Cognition: The Realization of the Living (Boston Studies in the Philosophy of Science, Vol. 42)*, D. Reidel Publishing Company, Dordrecht, Holland.

Meadows, Donella (2008). *Thinking in Systems: A Primer*, Chelsea Green Publishing Company, White River Junction, VT.

Meadows, Donella, Randers, Jorgen & Meadows, Dennis (2004). *Limits to Growth: The 30-Year Update*, Chelsea Green Publishing Company, White River Junction, VT.

Miller, Geoffrey (2000). *The Mating Mind*, Anchor Books, New York.

Miller, James Grier (1978). *Living Systems*, McGraw-Hill, New York.

Mitchell, Jason P., Mason, Malia F., Macrae, C. N., and Banaji, Mahzarin R. (2006). “Thinking about Others: The Neural Substrates of Social Cognition”, in Cacioppo, et al (eds, 2006), *Social Neuroscience*, pp 63-82, The MIT Press, Cambridge MA.

Mithen, Steven (1996). *The Prehistory of the Mind: The Cognitive Origins of Art, Religion, and Science*, Thames and Hudson, London.

Mlodinow, Leonard (2012). *Subliminal: How Your Unconscious Mind Rules Your Behavior*, Random House Inc., Vintage Books, New York.

Mobus, George. E. (1994). "Toward a theory of learning and representing causal inferences in neural networks", in Levine, D.S. and Aparicio, M. (Eds.), *Neural Networks for Knowledge Representation and Inference*, Lawrence Erlbaum Associates. Available on-line: http://faculty.washington.edu/gmobus/Adaptrode/causal_representation.html.

Mobus, George E. & Fisher, Paul (1994). "MAVRIC's Brain", Technical Paper, University of North Texas and presented at *Seventh International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems* IEA/AIE'94, San Diego, CA. Available on-line: http://faculty.washington.edu/gmobus/Mavric/MAVRICS_Brain_rev.html

___________________________ (1999). "Foraging Search at the Edge of Chaos", in D. Levine, V. R. Brown and V. T. Shirey (eds), Oscillations in Neural Networks, Lawrence Erlbaum Associates, Hillsdale, New Jersey. Available on-line: http://faculty.washington.edu/gmobus/Mavric/chaos.html

Mobus, George E. (1999),"Foraging Search: Prototypical Intelligence", The Third International Conference on Computing Anticipatory Systems, Liege, Belgium, August, 1999. Available on-line: http://faculty.washington.edu/gmobus/ForagingSearch/Foraging.html

_______________ (2000). "Adapting robot behavior to a nonstationary environment: a deeper biologically inspired model of neural processing". *Sensor Fusion and Decentralized Control in Robotic Systems III*, SPIE - International Society for Optical Engineering, Nov. 7, 2000, Boston, MA. Available on-line: http://faculty.washington.edu/gmobus/Mavric/Nonstationary/spie-paper.html

Mobus, George E. & Kalton, Michael C. (2014). *Principles of Systems Science*, Springer, New York.

Mobus, George E. (2015). "A Systems Science Framework for Understanding the Nature of Governance", *Proceedings of the 59th Annual Meeting of the ISSS* - 2015, Berlin, Germany.

________________ (2017). "A Framework for Understanding and Achieving Sustainability of Complex Systems", *Systems Research and Behavioral Science*, **34**:5, John Wiley & Sons, Ltd, doi: 10.1002/sres.2482.

Mobus, G.E. & Anderson, K. (2016). "System Language: Understanding Systems", *Proceedings of the 60th Annual Meeting of the ISSS* - 2016, Boulder, Colorado.

Montague, Read (2006). *Why Choose the Book? How We Make Decisions*, Dutton, New York.

Montgomery, David R. (2007). *Dirt: The Erosion of Civilizations*, University of California Press, Berkeley CA.

Moravec, Hans (1990). *Mind Children: The Future of Robot and Human Intelligence*, Harvard University Press, Cambridge MA.

Morowitz, Harold J. (1968). *Energy Flow in Biology*, Academic Press, New York.

_________________ (2004). *The Emergence of Everything: How the World Became Complex*, Oxford University Press, Oxford UK.

Odum, Howard T. (1983). *Systems Ecology: An Introduction*, Wiley, New York.

________________ (2007). *Environment, Power, and Society for the Twenty-First Century: The Hierarchy of Energy*, Columbia University Press, New York.

Pinker, Steven (1997). *How the Mind Works*, W. W. Norton & Company, New York.

_____________ (2002). *The Blank Slate: The Modern Denial of Human Nature*, Penguin Books (Paperback Ed.), New York.

_____________ (2007a). *The Language Instinct: How the Mind Creates Language*, Harper Perennial, New York.

____________ (2007b). *The Stuff of Thought: Language as a Window into Human Nature*, Viking, New York.

____________ (2011). *The Better Angles of Our Nature: Why Violence Has Declined*, Viking, New York.

Potter, Van Ransselear (1971). *Bioethics: Bridge to the Future*, Prentice-Hall, New York.

Rees, Martin (2003). *Our Final Hour — A Scientist's Warning: How Terror, Error, and Environmental Disaster Threaten Humankind's Future in This Century—On Earth and Beyond*, Basic Books, New York.

Rosen, Robert (1985) *Anticipatory Systems: Philosophical, Mathematical, and Methodological Foundations*, Pergamon Press, New York.

Robinson, Daniel N. (1990). "Wisdom through the ages", in Sternberg (1990, pp 13 - 24).

Rumelhart, D. E., McClelland, J. L. & the PDP Research Group (eds. 1986). *Parallel Distributed Processing: Explorations in the Microstructures of Cognition*, Volumes 1 & 2, The MIT Press, Cambridge, MA.

Ruse, Michael (2003). *Darwin and Design: Does Evolution Have A Purpose?*, Harvard University Press, Cambridge, MA.

Sapolsky, Robert M. (2018). Behave: The Biology of Humans at Our Best and Worst, Penguin Books, New York.

Scalaidhe, Séamas P. Ó, Wilson, Fraser A.W., & Goldman-Rakic, Patricia S. (1999). "Face-selective Neurons During Passive Viewing and Working Memory Performance of Rhesus Monkeys: Evidence for Intrinsic Specialization of Neuronal Coding", *Cerebral Cortex,* Volume 9, Issue 5 Pp. 459-475,

Scott, James C. (2017). *Against the Grain: A Deep History of the Earliest States*, Yale University Press, New Haven CT.

Searle, John R. (2004). *Mind: A Brief Introduction*, Oxford Uninversity Press, New York.

Seung, Sebastian (2013). *Connectome: How the Brain's Wiring Makes Us Who We Are*, Mariner Books, New York.

Schank, R.C. & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding*, Earlbaum Assoc., Hillsdale, NJ

Schneider, Sandra L. & Shanteau, James (eds. 2003). *Emerging Perspectives on Judgment and Decision Research*, Cambridge Series on Judgment and Decision Making, Cambridge University Press, Cambridge UK.

Schneider, Eric D. & Sagan, Dorian (2005). *Into the Cool: Energy Flow, Thermodynamics, and Life*, University of Chicago Press, Chicago IL.

Schwartz, Barry (2004). *The Paradox of Choice: Why More is Less*, Harper, New York.

Seabright, Paul (2004). *The Company of Strangers: A Natural History of Economic Life*, Princeton University Press, Princeton, New Jersey.

Shannon, Claude E. & Weaver, Warren (1949). *The Mathematical Theory of Communication*. Univ of Illinois Press, Chicago IL.

Simon, Herbert A. (1998). *The Sciences of the Artificial*, Third Edition, The MIT Press, Cambridge MA.

Sloman, Steven A. (2002). "Two Systems of Reasoning", in Gilovich, et al (2002, pp 379 - 396)

Smith, Eric & Morowitz, Harold J. (2016). *The Origin and Nature of Life on Earth*, Cambridge University Press, Cambridge UK.

Smith, John Maynard (1964). "Group selection and kin selection". *Nature* 201 (4924): 1145–1147.

Smith, John Maynard & Szathmáry, Eörs (1995). *The Major Transitions in Evolution*, Oxford University Press, Oxford UK.

Sober, Elliott & Wilson, David Sloan (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior*, Harvard University Press, Cambridge MA.

Speth, James Gustave (2004). *Red Sky at Morning: America and the Crisis of the Global Environment*, Yale University Press, New Haven, CT.

Sporns, Olaf (2011). *Networks of the Brain*, The MIT Press, Cambridge.

Squire, Larry R. & Kandel, Eric R. (2009). *Memory: From Minds to Molecules*, Roberts & Company, Greenwood Village, CO.

Sternberg, Robert J. (ed.) (1990a). *Wisdom: Its Nature, Origins, and Development*, Cambridge University Press, New York.

_______________ (1990b). "Wisdom and its relations to intelligence and creativity", in Sternberg, Robert J. (ed.) (1990, pp 142 - 159).

_______________ (1997). *Thinking Styles*, Cambridge University Press, Cambridge.

_______________ (ed.) (2002). *Why Smart People Can Be So Stupid*, Yale University Press, New Haven.

________________ (2003). *Wisdom, Intelligence, and Creativity Synthesized*, Cambridge University Press, New York.

Stewart, John (2000). *Evolution's Arrow*, The Chapman Press, Canberra AU.

Striedter, Georg F. (2005). *Principles of Brain Evolution*, Sinauer Associates, Inc., Sunderland, MA.

Suddendorf, Thomas (2013). *The Gap: The Science of What Separates Us from Other Animals*, Basic Books, New York

Surowiecki, James (2004). *The Wisdom of Crowds*, Doubleday, New York.

Swanson, Larry W. (2003). *Brain Architecture: Understanding the Basic Plan*, Oxford University Press, Oxford UK.

Tainter, Joseph A. (1988). *The Collapse of Complex Societies*, Cambridge University Press, Cambridge MA

Tattersall, Ian (2012). *Masters of the Planet: The Search for Our Human Origins*, Palgrave Macmillan Trade, New York.

Tomasello, Michael (2014). *A Natural History of Human Thinking*, Harvard University Press, Cambridge MA.

_______________ (2016), *A Natural History of Human Morality*, Harvard University Press, Cambridge MA.

_______________ (2019), *Becoming Human: A Theory of Ontology*, Harvard University Press, Cambridge MA.

Tversky, Amos & Kahneman, Daniel (1986). "Rational Choice and the Framing of Decisions", *The Journal of Business*, **59**:4, Part 2: The Behavioral Foundations of Economic Theory, pp. S251-S278.

Volk, Tyler (2017). *Quarks to Culture: How We Came to Be*, Columbia University Press, New York.

von Neumann, John & Morgenstern, Oskar (1944), *Theory of games and economic behavior*, Princeton University Press, Princeton NJ.

Wackernagel, Mathis & Rees, William (1996). *Our Ecological Footprint: Reducing Human Impact on the Earth*, New Society Press, Gabriola Island, BC Canada.

Wallerstein, Immanuel (2007). *World-Systems Analysis: An Introduction*, Duke University Press, Durham NC.

Ward, Peter (1994). *The End of Evolution: A Journey in Search of Clues to the Third Mass Extinction Facing Planet Earth*, Bantam Books, New York.

Williams, George C. (1957). "Pleiotropy, natural selection, and the evolution of senescence". *Evolution* **11** (4): 398–411

Wilson, D.S. & Sober, E. (1994). "Reintroducing group selection to the human behavioral sciences". *Behavioral and Brain Sciences* 17 (4): 585–654.

Wilson, D. S. & Wilson, E. O. (2008). "Evolution 'for the good of the group'". *American Scientist* 96 (5): 380–389.

Wilson, Edward. O. (2013). *The Social Conquest of Earth*, Liveright, New York.

Wiener, Norbert (1950). *The Human Use of Human Beings: Cybernetics and Society*, Avon Books, New York.

Wright, Robert (1994). *The Moral Animal, Why We Are the Way We Are: The New Science of Evolutionary Psychology*, Vintage Books, New York.

Index

About the Author

George Mobus is Associate Professor Emeritus of Computer Science & Systems at the University of Washington Tacoma, School of Engineering & Technology.

George received his Ph.D. in Computer Science in 1994 from the University of North Texas after years working in industry as a software engineer, engineering and operations manager, and CEO of an embedded systems design and manufacturing company in Southern California. Prior to joining SET at UWT (formerly the Institute of Technology) he taught computer science at Western Washington University where he did research on autonomous robotics. At UWT he taught courses in systems software, robotics, embedded systems programming, and the Global Challenges course in the UWT Global Honors Program. For the last decade Dr. Mobus has been turning his attention to global issues, such as global warming and peak oil, energy return on energy invested (EROI) and population dynamics where computational problem solving might be applied.

Dr. Mobus' overarching interest has been with systems science and how it might be applied to the analysis of these global and significant problems. He works with groups from the International Society for the Systems Sciences (ISSS) and the International Council on Systems Engineering (INCOSE) to develop a more consistent definition of systems science and systems engineering. He is the lead author of the well-received textbook, "Principles of Systems Science", published by Springer (2015) with co-author Michael Kalton. He is currently working on a new book that shows how to use the principles to guide deep systems analysis of both natural and engineered systems and has been a participant in two of the last International Federation for Systems Research Conversations (http://www.ifsr.org/index.php/ifsr-conversations/) working on this subject. He was appointed, in 2017, as Editor-in-Chief of the Springer book series on Systems Science and Systems Engineering, https://www.springer.com/series/6104). Currently, Dr. Mobus is concentrating on the systems approach to engineering a more naturalistic human socio-economic system than has evolved in societies thus far.