COG 266 Michael Phillips
Final Paper 5/16/2014

On the Complexity of Language and Cognition

There are many different forms of life on Earth. If all of these life forms had a common source, then the diversity of life since the genesis of biology has diverged in unimaginable ways to form a plethora of species. If all life is fundamentally connected however, then we should also be able to point out all of the similarities that exist among all the living creatures in complement to their differences. In the animal kingdom, one example of diversity among animals is the differences in brain power that we see, that is, different levels of intelligence and different cognitive capabilities from species to species. So is there any way to talk about evolution, cognition, and the similarities and differences of living beings in one comprehensive model? In what follows, I argue that the aspect of “complexity” can be used to talk about all of these considerations. In Part I, I discuss what I mean by complexity, and introduce my model of complexity. In Part II, I discuss the topics of language and cognition in humans and other intelligent animals, and at the end of each discussion I briefly discuss how these topics are related to and can be explained by complexity. The goal is to show how complexity provides a means to talk intelligently about the similarities and differences between humans and other animals.

Part I – Complexity

In biology, evolutionary theory, linguistics, or other fields, the term ‘complexity’ can have a wide variety of meanings and applications. Therefore, it is important for me to identify what I mean by complexity. The term in this case is borrowed from the concept of Kolmogorov complexity, also known as descriptive complexity. In descriptive complexity, the measure of the complexity of an object is defined by how much information is in the object, and how compressible that object is. Consider the following two strings of alphanumeric characters:

String 1	a01a01a01a01a01a01a01a01a01a01a01
String 2	4c1j5b2p0cv4w1x8rx2y39umgw5q85s7

Despite being the same length, String 2 is said to have a higher descriptive complexity, because the information in the string cannot be broken down any simpler (it's not compressible), but in String 1 it would be easy to simply multiply the <a01> sequence over and over. I argue that this kind of complexity can apply to the brain (or other features) of organisms in evolution, but in biology, there are no "compressible strings", simply ones which contain more or less information. For example, we might able to describe the complexity of a "simpler" ape such as an australopithecine as: <a, b, c, d>, where each letter represents some cognitive ability (of a particular domain) of the animal ('a' could be memory, 'b' could be ability to hear, 'c' could represent some level of perception, etc…). If australopithecines have <a, b, c, d>, and humans evolved from them, then humans also have <a, b, c, d> (let's assume). However, humans have evolved to obtain many other cognitive functions and states, such as consciousness, high intelligence, and language, and so the "string" of humans might be much longer <a, b, c, d, e, f, g, h, i, j, k, l, m> but the original four variables are still intact. Therefore, we have demonstrated that the descriptive complexity of the human brain is much higher than the descriptive complexity of our ancestors from millions of years ago, so this shows the difference between us. However, the fact that both species contain similar features <a, b, c, d> also shows the thing we share, therefore showing us how we are similar, and how evolution connects us.

As we see, complexity is not a binary system of opposites (simple or complex) but rather a graded scale or spectrum, on which species lie relative to one another:

Species A — Species B — Species C

$\longleftrightarrow$

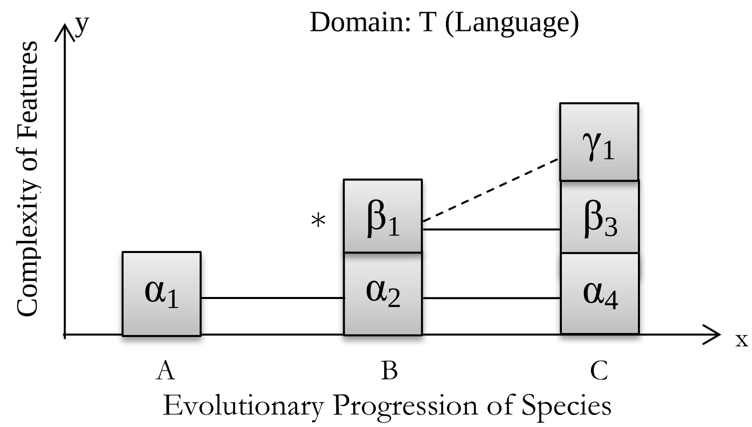
Less Complex — More complex

One thing to note is that in the biological/evolutionary term of "complexity", it is not always the case that a lineage of species through evolution aims to be more complex. Natural selection does not have any intrinsic directionality such as "more complex" that is implements over the time. It is just usually the case (including the case of our own species) that organisms become more complex, especially in environments where the organism needs to be versatile and adaptive.

In what follows, I propose a much more concise model for my definition of complexity in Figure 1, followed by a description of the model (the example demonstrates the linguistic complexity of hominids).

Figure 1

Linguistic Complexity in Hominids (Simplified)



How this model works:

The species: A, B, and C represent different species in the primate order. Primate A may be the earliest hominid, primate B is a hominid that preceded humans and that humans evolved from, and primate C is homo sapiens, or humans.

The features: The most important part of these models is to define the cognitive features (represented using the Greek alphabet). Even though we may use a similar alphabet in every model (α, β, γ…) we may use them to represent different features every time. These variables can also represent non-cognitive features (such as physical/biological features like vocal cords) as well as feature "groups", and this is especially the case in simplified models such as these. I define the features in this particular model as follows:

Variables of the model and the features they represent:
α – **the α-group:** a collection of simple cognitive abilities and features needed for everyday survival (some extent of memory, attention, ability to perceive sound, etc…). All primates have this "group" of brain features. These basic features are the ingredients needed for language (they provide the foundation/building blocks). β – Mimetic intelligence (more or less). A pre-linguistic system of motor-based representations, which enables the ability to communicate one's intentions. γ – **the γ-group:** the final piece needed for full human language. Contains features such as highly developed language faculty (brain module), complex vocal cords, etc…

In this model, it is proposed that a kind of "mimetic intelligence" (β) in an earlier primate eventually lead the acquisition of a mental language faculty (γ) in humans. This is only a speculative proposition on my part, however, so this is represented with a dotted/dashed line in the graph. ($\beta \rightarrow \gamma$). Mimetic intelligence was retained, however, from primate B to C (humans), showing the fundamental relationship between the two species in this domain. Note that primate A did not have a functional mimetic intelligence, but did have the α-group $<\alpha_1>$. Non-dotted or "full" lines are used to represent non-speculative relationships between features in various evolutionarily related species (we know that $\alpha \rightarrow \alpha$), and asterisks are used to indicate features which came about on their own (not necessarily because of another feature, but perhaps because of environmental factors). The subscript numbers of each feature represents a kind of **internal complexity** overtime (primate A and primate C both have memory (α), but the memory of C might be much greater than A). The higher the number is, the higher this level of complexity in some sense for that feature.

The Domain: The domain we specify is essential in determining what aspect we want to refer to, and therefore what features should be relevant in our model. For example, Domain T is the "linguistic" domain, that is, we only care about features that pertain to the linguistic capabilities of an organism. Without defining our domain, we'd have no way to know what features should be relevant (species A could have a lot of other features we could list, but none of them would contribute the complexity of their linguistic capabilities).

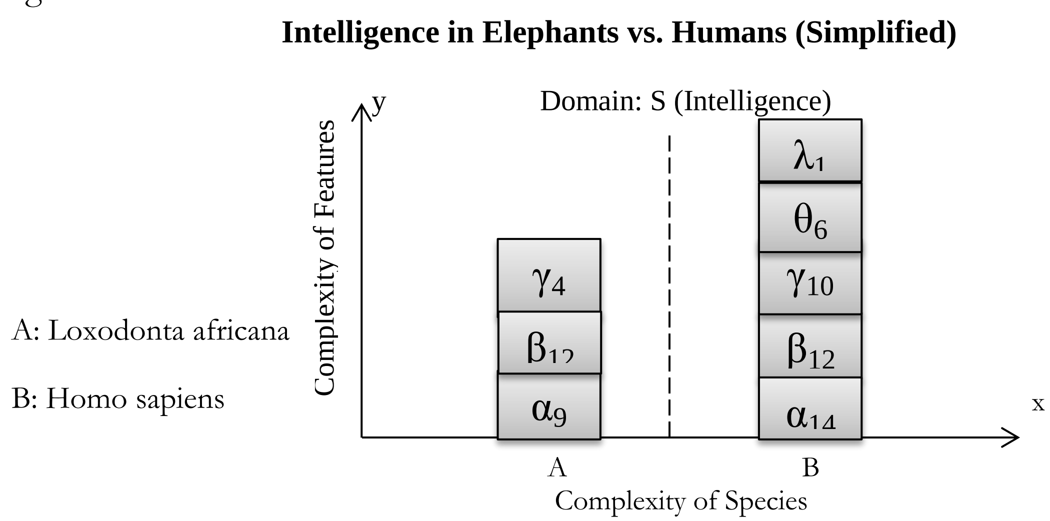
Therefore, the complexity of the language of organism C (*homo sapiens*) can be stated as $<\alpha_4, \beta_3, \gamma_1>$ in this (extremely) simplified model (this is called a **string of features**). Reverting back to the idea of descriptive complexity, this string can be seen as the "description". Note that the string $<\alpha_4, \beta_3, \gamma_1>$ is more complex than the string $<\alpha_1>$ (it has more information in it, all the while being uncompressible). Therefore, the descriptive complexity from the string of features in the linguistic domain T is higher for primate C (humans) than from primate A (some early ancestor to humans). This is one way to look at the complexity factor.

As an aside, note that one can be very specific when addressing a particular feature in the model, such as 'cognitive feature $<\alpha_4>$ of Organism C in Domain T ($T(C<\alpha_4>)$).

The main success here is to show how previous species are **similar** to later species because of how simpler features in the domain (such as α) are retained throughout these related species (the next species isn't going to lose the α-group, it's an essential building block), and how these species are **different** because the number of features and their complexity keeps growing (α, β → γ, etc…). Using this model, we are also able to successfully talk about the evolution of a particular domain (such as language) in a particular species (such as primates). This also helps show the fundamental similarities and differences between humans and primates today such as chimpanzees even though we didn't evolve from them (although we had a common ancestor).

But what about a complexity model for species of different families or orders? For example, what if I want to compare humans with elephants? Elephants are an example of one of Earth's most intelligent animals, but we did not evolve from them. Figure 2 shows a "comparative" model of complexity.

Figure 2



In this model, there are no lines connecting the features, because elephants do not "evolve" into humans. This time, we are simply observing and comparing the different features in the domain of intelligence. When we look at the features, we notice that even though these animals are separated by millions of years of evolution, certain similarities remain. In this case, the α-group is much bigger, and represents a lot of similar cognitive features needed for intelligence on the level for both elephants and humans (memory, attention, ability to respond to environment, self-recognition, etc…). In contrast to these similarities seen here, there are also differences, because only species B (humans) is shown to have <λ> (where <λ> represents some identifiable feature contained in human intelligence, but not elephants, so this could be a group of features), and the internal complexity of the features (represented in the subscript numbers) is constantly higher in humans, although this is an assumption. Notice the convergent evolution that takes place since both humans and animals evolved to have similar features of intelligence.

Let's revert back one final time to the original idea of descriptive complexity to reinforce the argument being made. Essentially, as a species evolves, certain new cognitive features and non-cognitive features get added over time to equal a higher level of complexity. Let's represent this in uncompressible strings:

Species A: [az1w27]
Species B: [az1w27] + [fg12d6]   Evolution
Species C: [az1w27] + [fg12d6] + [jkl4d1]
…

Of course instead of alphanumeric strings, the species are developing new biological features overtime (more information), but most importantly they are building off of old ones. For example, if an ancestral species had some extent of memory, the next species doesn't lose its ability to remember things, in fact, it will use memory to build newer and more complex cognitive features ($\alpha \rightarrow \alpha, \beta$). This is not always the case, but depending on our domain, certain things like intelligence or linguistic capability certainly follow this increase of complexity over time. As was the case in the comparative complexity model in Figure 2, one can also use complexity to compare to non-related species of animal, based on features that they share. Therefore, this definition and model of complexity is successfully able to compare species, discuss evolution, and describe cognitive features and abilities in a representational way.

This particular model is in a very developmental/fragile stage, and as such still has problems. For example:

- The features at this point are relatively arbitrary. It's too difficult or near impossible to correctly identify each and every feature that should be in a domain, especially if we want to talk about extinct organisms.
- Difficult to correctly identify the connections between different features. For example, did α really assist in the development of β, or was this independent and/or based on something else?
- Only shows that there are features of a particular domain involved, does not fundamentally show how they "work together" in the brain.

Part II – Language and Cognition

Now, let's use two concrete concepts to compare humans and other intelligent animals: language and cognition. As mentioned, I relate each subject to our model of complexity at the end of the discussions.

1. Language

Since we are comparing the similarities and differences of human features with other animals, then the following question must be considered for language: What is the difference between systems of animal communication and natural language as possessed by humans? In order to answer this question, I first talk about human language, followed by an example of animal language (ape language).

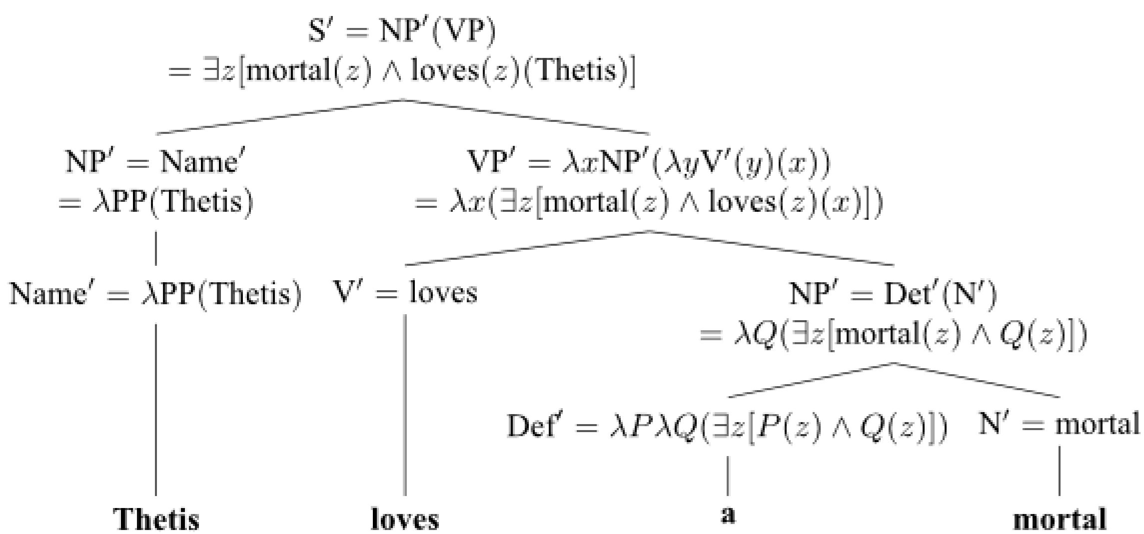
The origins and evolution of human language are not as well understood as we'd hope. The ability to communicate verbally with complex collections of words and phrases has helped to make our civilizations and cultures what they are today. So little is known about the origin and evolution of language because researchers/linguists have only started studying this topic for the past few decades, and the research remains to increase. In a recent article entitled "The mystery of language evolution", Hauser, Yang, Chomsky, and others claim that we know much less even than we think we do, stating "all modeling attempts have made unfounded assumptions, and have provided no empirical tests, thus leaving any insights into language's origins unverifiable". Therefore, there is not as of yet any one particular theory that has been accepted to explain the true detail of language origins, however many hypotheses have immerged. To list a few:

The Bow-wow hypothesis:	The earliest words were simply imitations of cries and noises of animals in the nearby environment.
The Mother Tongue Theory (scientific monogenesis):	One original primitive language gave rise to all human languages on Earth.
Gestural Theory:	Human language developed out of the use of gestures for simple communication.
Ritual/Speech coevolution:	Holds that no theory of language origin is valid, because language is not a separate adaptation, but an internal aspect of something much wider

These hypotheses and theories barely scratch the surface, but it shows how difficult understanding early human language truly is.

Since the origins and evolution of language are not (yet) well understood, let's focus on aspects of natural language of which we can be certain. Language is so complex and so autonomous from other features of the human brain that it merits its own field of study: linguistics. Theoretical linguistics encompasses a variety of areas of study, such as phonology, morphology, syntax, semantics, and pragmatics. These fields are used to look at the complex properties of language, from the physical articulation of phonemes, to the morphosyntactic properties of grammar, to the way words and phrases apply an interactive meaning in social contexts. On a certain level of complexity, the linguistic capabilities of humans is unparalleled by any species alive today. In order to see this complexity, consider the following parse tree from the *Stanford Encyclopedia of Philosophy*:

Figure 3 **Natural Language with Formal Components**



This is a model showing a parse tree of a simple English sentence "Thetis loves a mortal". There are a few things worth observing from this figure: different areas of linguistics come together in order to form the final picture (semantics + syntax), the use of formal logic (e.g. lambda calculus) grants us a sense of the complexity of natural language, and there seems to be an underlying structure to language seen here that reveals how language is not purely linguistic but is reliant on (as well as an intrinsic part of) other cognitive features (consider the ritual/speech coevolution hypothesis).

Now that we have briefly discussed the origins and evolution of language, areas of linguistics, and models which demonstrate the complexity of natural language, let's look at the systems of communications in other intelligent animals, with the example of nonhuman primates. Much research has been done to test the language capabilities of the great apes, including chimpanzees, gorillas, and bonobos. Most notably, attempts are used to communicate with these species uses some form of sign language, or artificial languages based on lexigrams such as Yerkish. A number of great apes are quite successful at learning some form of language (like American Sign Language – ASL) such as the gorilla Koko who knows about one thousand unique signs, and the chimpanzee Washoe who knew roughly 350 words of ASL. This has been one way for humans to see the true intelligence of their ancestors. However, it's worth noting that these apes don't use their own form of sign language in the wild – instead apes tend to communicate much more simplistically with a system of vocalizations such as grunts, screams, and some social gestures.

But why wouldn't they use something more efficient like sign language if it appears they're capable of it? This implies for these species that a certain cognitive level has been reached in order to express some level of complex thought, but the communication of these thoughts have not yet been formalized or actualized among the species themselves – that is, they have not yet found a way to express themselves naturally on this level, but the capability is there in some form. This shows us how the non-linguistic factors of concepts such as mimetic intelligence must precede the actualization of a natural language system. Therefore, the extent to which a research animal (such as Koko) is able to communicate is revealing of their cognitive development as a whole. Some great limitations for ape language include: apes can only learn a finite amount of words (a thousand or so maximum), apes fail to ask questions using language (reveals cognitive limitations), and apes lack complex vocal cords capable of human-level communication (a non-cognitive feature).

It should be noted that by "animal language" I mean the implementation of human (natural) language in animal systems. Only a finite amount of animals are even capable of this, which is why apes provide such a great example due to their intelligence. Natural systems of communication in the animal kingdom also exist – such as bird songs, whale songs, communication in elephants, and even dance behaviors in bees. Communication among species of animals is certainly a demonstration of intelligence and cognitive development.

- How does language relate to the model of complexity?

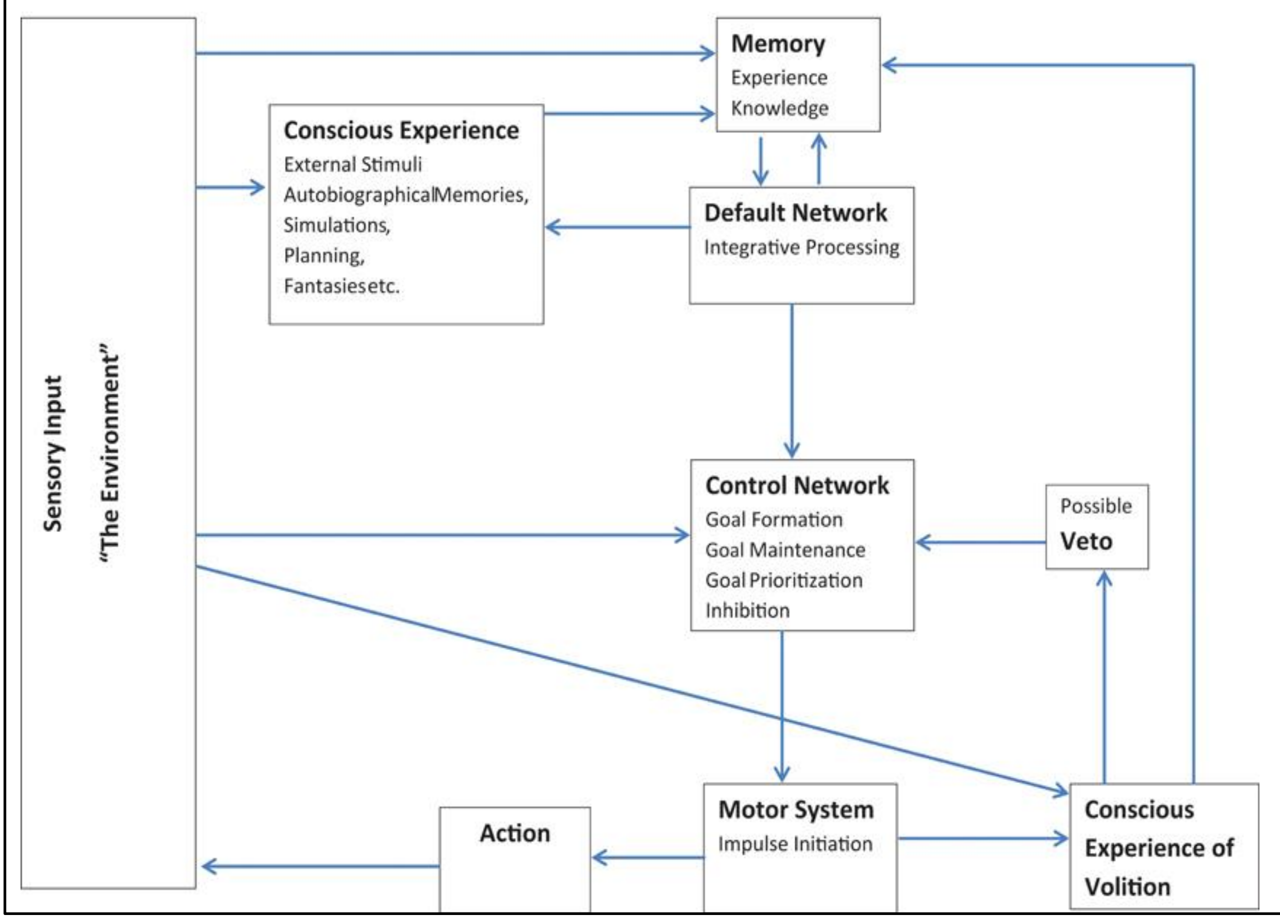
Language is a concrete example of a domain that can be described among different species of animals by complexity. If we were to build a comparative model of complexity for humans and apes that know sign language, the descriptive complexity for humans would be much higher, but we would also share some features of language with them and other animals, on a primitive level. Let's now readdress the original question: what's the difference between the systems of communication used by animals and natural language used by humans? Our complexity model would suggest that nothing more than differing levels of complexity based on specific and relevant features of the linguistic domain separates the two. In our discussion, we found that the ability to conceptualize the environment and to represent said environment using signs was a skill possessed by both humans and apes (just with differing internal complexity, a concept discussed earlier). This discussion of language has shown us how language, systems of communications, linguistic features, cognitive development, intelligence, and the general similarities and difference among humans and other animals can all be explained and quantified using the model of complexity from Part I.

2. Cognition

Again, let's start by talking about humans. Human cognition is a complex topic, with many involved processes such as calculation, reason, decision making, the ability to take in information about the environment (input) transform and elaborate that information (process) and be able to store, respond with, recover, and use information in order to perform a variety of higher functions (output). Any demonstration of this system with respect to thought processes is considered "cognition" for the intents and purposes here. Because cognition is so complex and intricate in human beings, there are a multitude of ways in which cognition can be demonstrated. Take for example the demonstration of consciousness in humans. Based on the brain architecture and neural networks of the human mind, humans are able to display a highly complex level of self-awareness, thinking capability, and other psychological factors involving the cognitive state of consciousness. Figure 4 shows a model of consciousness from an article entitled "Re-Conceptualizing Free Will for the

21st Century: Acting Independently with a Limited Role for Consciousness" that demonstrates the complexity of the conscious architecture (although this is only a proposal, as it is not the case that any one network or architecture for consciousness has been accepted by all cognitive scientists).

Figure 4 **A model for consciousness that demonstrates free will**



Whether or not humans have free will is a whole other debate, but Figure 4 at least shows just how complicated a model for consciousness is (external stimuli, action, memory, control network, knowledge, planning, possible volition, etc…). The model, despite being geared towards free will (and being rather simplistic) demonstrates the cognitive system: input → process → output.

In addition to free will, one really important aspect of consciousness is its ability to provide us with a cognitive state wherein which we are **aware of ourselves**. The **self** can be a complicated issue, with many different positions as well as psychological and philosophical considerations and implications. Concepts such as knowledge, memory, and other complex thought processes all add up to give us a sense of self, and one aspect or integral part of the self is self-awareness, otherwise known as the ability (or capacity) for a living being to recognize oneself as an individual who is separate from the environment. Philosopher René Descartes used the term "I think, therefore I am" in order to determine and prove that he is a self-aware being/entity. Not only do humans have self-awareness, but we can see self-awareness as only one small aspect of consciousness as a whole. Can we say the same for any other animals?

It is very difficult to imagine what goes on inside the brain of another animal. A branch of ethology known as **cognitive ethology** is generally concerned with this issue. Cognitive ethology tries to get at the heart of questions such as Does animal x have conscious awareness? What is the motivation for the behaviors displayed in animal x?

In order to address some of these questions, let's consider an animal: dolphins (*Tursiops truncatus*). Dolphins have large brains (relative to their body size), and physiologically they have a highly complicated neo-cortex (which of course is the part of the brain involved in higher functions such as conscious thought). Their brains seem to have the neurological capacity for self-awareness, and even for emotional processes. Some even think that the combination of whistles and clicks used by dolphins are signs of a complex language. But do dolphins demonstrate a high level of cognition, perhaps in the form of self-awareness? One experiment done to test the self-awareness in animals is the mirror test, where an animal is presented with a mirror and is tested for signs of self-awareness. Bottlenose dolphins are known passers of this test. In front of a mirror they tend to dance around, show off, and demonstrate "play" behaviors. Also, if a human marks one of their flippers, the dolphin will use the mirror to inspect their flipper, knowing that it's themselves that they see.

Some think, however, that the display of intelligence by dolphins is not a demonstration of complex cognitive abilities, but the ability to perform simple behaviors due to imitation (not much unlike a dog trick). However, one experiment by researchers at the Kewalo Basin Marine Mammal Laboratory (KBMML) was done to test the ability of dolphins to learn sign language, and syntax (word order) based on combinations of different signs. When the dolphin is asked to "go through the hoop" (expressed in sign language) even though the hoop is lying flat at the bottom of the pool, the dolphin recognizes the words, thinks about the situation, finds a way to prop the hoop up so that she can go through it, and then darts through the hoop. This experiment demonstrates that the dolphin not only had the concept of the simple command "go through a hoop", but she also had a concept of "throughness". These same dolphins were able to demonstrate the same understandings of a multitude of concepts, respond to questions, and manipulate the environment based on the commands given. More research needs to be done, and we'll never be able to full understand what it's like to be a dolphin, but I think experiments such as these demonstrate some degree of self-awareness and higher thought processing in these animals. It may be no comparison to the level of consciousness possessed by human beings, but the fact that there are identifiable signs of intelligence, cognition, and self-awareness are thought-provoking.

- How does cognition relate to the model of complexity?

In this section, we define cognition as any neurological system that invokes thought processes, with one of our examples being the cognitive state of consciousness. As we see (as was the case with language), the cognitive capability of humans seems to be far superior to the consciousness of nonhuman animals, however, comparing aspects of consciousness with dolphins showed us similarities which exist between the human species and cetaceans. Cetaceans have repeatedly shown to be one of the smartest orders of animals on the planet through their ability to pass the mirror test, ability to demonstrate an understanding for cognitive concepts, ability to communicate (among themselves and with humans), intricate hunting skills, etc… One of the amazing considerations here is that humans are separated from dolphins on the evolutionary tree by millions of years, and yet there are features of both species that we could compare. Convergent evolution has helped both species become intelligent with similar features through many years of adaptation.

Again, if we were to make a string of variables representing features of cognition, one for humans and one for dolphins, the human string $<\alpha, \beta, \gamma,\ldots >$ would be bigger (therefore humans have a higher "descriptive complexity" for cognition), but both strings would have to include some of the same features, because as dolphins have shown, they are intelligent, self-aware (or at least self-recognizing) beings who, with their large brains, almost give us a run for our money! The fact that the human descriptive complexity string would be bigger highlights the differences between humans and these other animals, but a comparative model of complexity would reveal that we could also highlight similarities by seeing that humans and dolphins share features such as self-recognition <λ>, capacity for memory <α>, ability to respond to external stimuli with contemplation and decision making <β>, etc… Therefore, using descriptive complexity has again been able to show the differences between humans and other animals (what makes us unique, what features are unique to us) but also show the similarities in the form of shared features in the descriptions.

Conclusion

In this paper, I started by introducing a model of complexity based on Kolmogorov complexity, wherein which biological features are used to form variables in a string of a particular relevant domain. The amount of features making up information in that string is used to determine the descriptive complexity of that species, in that domain (language, cognition, intelligence, etc…). When we discussed language and cognition, and compared humans with other (intelligent) animals, we found that there were similarities and differences between the two in each case. There are features of our minds which make us uniquely human due to our intellect, and so we see the stark differences between humans and other animals, as the descriptive complexity derived for humans in the domains discussed here is much higher. However, as was seen in the models from Figure 1 and Figure 2, even if two species are different, that does not mean they do not contain the same features: often times the less complex species will at least have some essential building blocks relevant to the domain. This shows us how we aren't as different from other animals as we think, and that certain similarities can be highlighted in this way.

Therefore, using this description of complexity enabled us to make intelligent comparisons between humans and other intelligent animals, highlight the similarities and differences, enabled us to talk about aspects of evolution, cognitive development, and discuss the complexity of two specific domains: language and cognition. This model has therefore brought together concepts related to the components of the physical brain, the power of the mind, and even the mystery of consciousness.

References

"Animal Communication". Department of Linguistics, The Ohio State University. 1994. Retrieved 14 May 2014.

"The Language of Dolphins." Exploration Films, 13 May 2008. Web. 14 May 2014. <http://www.youtube.com/watch?v=jz3sQsTE5tA>.

Bonn, Gregory. "Re-Conceptualizing Free Will for the 21st Century: Acting Independently with a Limited Role for Consciousness" *NCBI*, Published 9 Dec 2013. Web. Retrieved May 14 2014 <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3856385/>

Fortnow, Lance. "Kolmogorov Complexity." University of Chicago (n.d.): n. pag. Web. 14 May 2014. <http://people.cs.uchicago.edu/~fortnow/papers/kaikoura.pdf>.

Hauser, Marc, Charles Yang, and Noam Chomsky. "The mystery of language evolution." Frontiers. Frontiers in Psychology, 7 May 2014. Web. 14 May 2014. <http://journal.frontiersin.org/Journal/10.3389/fpsyg.2014.00401/full>.

Parker, S.T., and R.M Mitchell. "Self-Awareness in Animals and Humans: Developmental Perspectives". New York, NY: *Cambridge University Press*, 1994. Print.

Patterson, Francine. "Ape Language" *Science* 211 (4477):86-88 1981. Web. Retrieved 14 May 2014. <http://www.sciencemag.org/content/211/4477/86.1>

Schubert, Lenhart. "Computational Linguistics." Stanford University. Stanford University, 6 Feb. 2014. Web. 14 May 2014. <http://plato.stanford.edu/entries/computational-linguistics/>.

Vajda, Edward. "The Origin of Language" Western Washington University. Web. 14 May 2014 <http://pandora.cii.wwu.edu/vajda/ling201/test1materials/origin_of_language.htm>