Curtis M. Kularski
Dr. Mark Faust
Introduction to Cognitive Science
10 May 2012

DIFFERENCES IN CATEGORIZATION BETWEEN BIOLOGICAL AND ARTIFICIAL COGNITIVE SYSTEMS

Biological cognitive systems and artificial cognitive systems are different in the methods in which they form categories and the assignment of meaning to a category. Biological systems form categories out of necessity for survival or based on logical groupings formed out of experience. For artificial cognitive systems categories for survival are not a consideration as their existence is in the service of their human creators. Artificial systems, unlike biological systems, have no motivation or sense of free-will, they act upon programming alone. An artificial system may act as a situated agent and form categories and concepts based on its own experience with the world, which differs from that of biological systems. Alternatively, an artificial cognitive system can be instructed in the creation of categories based on a human understanding of the world. The former case is not likely to render a meaning that is intelligible to humans. The latter case does not demonstrate a genuine attempt at learning and establishing meaning, only emulating existing systems of categorization.

Categories are an essential component of cognition. From social cognition (Johnson, Freeman, & Pauker, 2012) to casual object recognition (Cohen & Lefebvre, 2005), categories are an integral component for the success of the cognitive system. Categorization is a process that is innate in biological cognitive systems, but is often difficult to implement accurately in artificial cognitive systems. Categories are involved at various levels of cognition in humans ranging from innate categories to learned categories. In the absence of the structure of categories the physical world would be unintelligible to the human cognitive system and inaccessible to an artificial intelligence (Cohen & Lefebvre, 2005).

Categorization is a function of learning in organisms with a sensorimotor system. Categorization must also be a function of artificial cognitive systems that are situated in real-world environments. It is through categorization that concepts are formed and templates are matched. In humans categories are often thought of as linguistic associations between a general type of object or concept and a word or phrase. This is most often the case with physical objects and constructs that are of conscious importance. Some categories do not meet conscious awareness and do not need to be coded into language to be useful, such as categories of phenomena that are a threat to personal safety. Categories that are defined in language are attached to meaning. The category cup has the meaning of a vessel that it is possible to drink liquid from, whereas bowl may be defined as a vessel that is designed to be used in conjunction with a spoon for the consumption of food. Meanings are culturally and temporally specific. To state it most simply, categorization is required for the creation of meaning in organic systems and for the illusion of meaning in artificial systems (Harnad, 2005).

"Artificial intelligence is the study of how computer systems can simulate intelligent processes such as learning, reasoning, and understanding symbolic information in context." (Liu & Ren, 2010) In the modern state of artificial intelligence technology there is adequate processing power and storage capacity to facilitate large-scale intelligent systems, but there are limitations in the acquisition of knowledge due to mechanisms utilized in the forming of concepts and categories. Key problems existing in the selection of features used for categorization and the formation of coherent concepts based on the features that are selected (Liu & Ren, 2010). Part of the difficulty in selecting features is the lack of a mechanism for weighting various attributes in making categorical decisions (Harnad, 2005). A proposed solution to the problem is a theory of Axiomatic Fuzzy Sets (AFS), which membership in categories are determined based on a consistent algorithm and the selection of a distribution of features that is optimized by fuzzy logic. Utilizing fuzzy logic allows the cognitive system to form categories based on approximate data rather than absolute criteria. Axiomatic Fuzzy Sets theory was created based on perceptions of the human recognition process for the purpose of allowing artificial intelligence agents to generate concepts in a more human way. One important mechanism that has been replicated from human recognition is the ability to describe objects through only a small subset of their actual member categories instead of enumerating the entire array of characteristics to form a description of what is being identified. This occurs primarily through the use of sorting algorithms applied to fuzzy sets which eliminate attributes which are shared by many members of the set and emphasizing attributes which are similar between items. Attributes that are determined to be unique are reported based on the weighted likelihood that they are notable attributes for identification. Through the development of the theory and various improvements on its implementation it is now possible to develop a categorization system that is more efficient than those used by humans, relying on fewer and less detailed features for more accurate classification. The use of consistent algorithms ensures that the categorization results received are reproducible, while the fuzzy logic allows for dynamic adaptation as additional data is collected. Algorithms developed by Liu & Ren based on AFS theory are designed to be used in situated and embodied cognitive systems that interact with the real world. Their algorithms utilize multiple phases in the decision process to determine if input matches a category or is part of a new category (Liu & Ren, 2010). Utilizing adaptive algorithms and techniques such as AFS it is possible for artificial intelligence systems to emulate human categorization relatively well, but also to move beyond the confines of human ability to fully exploit the potential of a technology-driven cognitive system.

In comparison to the complicated algorithms and various theoretical approaches to categorization utilized in artificial intelligence it may seem as though human categorization is simple. On the contrary, human categorization is governed by numerous complicated systems. The key ways in which categories are formed for use in human cognition are cultural, institutional and individual categorization. In cognitive science the primary method of categorization that is given attention is cultural categorization. This includes areas such as language, common objects and social classifications. Ignoring individual and institutional categorization leaves a large gap in the space of what is categorized and how categories are formed. Categorization on the individual level begins early in life as an individual constructs categories of likes and dislikes or categorizes experiences based on culturally defined attributes. In recent times the emphasis of cultural constructs on individual categorization has become less notable as technology has become more focused on the experience of the individual. Through moving a majority of media and communication to electronic mediums it has become possible for arbitrary categories to be formed, destroyed and created again at the whim of the individual. In digital systems of organization there is no confinement to a single category for any item. Through the use of a system of multiple categories, often called “tagging”, it is possible to assign an item to many categories and then retrieve the item again later by retrieving one of the categories to which it belongs. On the surface this concept may seem like simply an approach to organizational systems which has nothing to do with cognitive categories, but in fact the concept of multiple categories indicates quite accurately how human categorization works on the individual level. When an individual develops categories for personal use there is not a single category in which each item belongs, it is associated with many categories (Glushko, Maglio, Matlock, & Barsalou, 2008). An example of this would be the way in which the individual categorizes music. It is possible to categorize music based on the artist, genre and tempo at a cultural level. At the individual level it is possible to preserve those categories, but new categories emerge that have no relevance to broader culture. An individual may organize their music based on the mood that is sets for them personally or have music separated into categories for different types of events. Recent research has revealed that cognitive categorization for different types of categories may be maintained by separate systems in distinct parts of the neural net. Theories relating to a multiple-systems model also posit the potential of each system having its own mechanism for learning categories (Ashby & Maddox, 2011). Individual categories exist inside of institutional categories and institutional categories exist inside cultural categories. The categorical systems build upon each other and inform each other (Glushko, Maglio, Matlock, & Barsalou, 2008).

The development of complex systems for teaching an artificial intelligence to categorize and the existence of individual categories are not completely dissimilar topics. An artificial cognitive system that has not been taught cultural concepts of categories must develop categories based upon environmental experience. Humans on the other hand spend their lives learning categories and learning how to map items into those categories. Artificial intelligence systems may be given some cultural training, depending upon their intended goal. Artificial intelligence systems that are designed to interact with humans typically are given basic categories that are important to interacting gracefully with humans. This basic level of knowledge helps shape the way that other categories are formed as it serves as a base for a system of categories that is weighted toward categories that are of interest to humans. The adaptation that occurs in humans through learning also occurs in artificial intelligence agents, but after an initial training period adaptation is typically prevented on artificial systems to prevent the development of an overly complicated system of categorization that would ultimately become unintelligible. Human cognitive systems have the benefit of life-long learning to adapt and refine categories with an organic understanding of when categories are and are not in need of adaptation or correction (Kirstein, Wersing, Gross, & Körner, 2012). Individual categories are adaptations made by the individual based on their own environment and their own needs or perceived needs. An artificial intelligence adapts category usage based on feedback and optimizes categories through learning algorithms (Liu & Ren, 2010). Individual human categories and fuzzy artificial intelligence categories rely on an individualized construction of categories with minimal consideration for social or cultural influences, but yet must respond to cultural categories when they are presented.

An area where differences between artificial and human cognitive categorization abilities have created much interest is in facial recognition. Humans handle facial recognition as an innate function with the ability to recognize and sort faces into categories with minimal mental effort, but their artificial counterparts struggle with. It is believed by some researchers in the area of artificial vision systems that the difficulty with establishing a reliable artificial facial recognition system is the reliance upon human intervention in the training of the artificial agent. It is believed that there are too many factors involved in facial recognition for a system to develop the capability to categorize facial features based on the limited input and experience that can be acquired through human training. Instead a proposed alternative is to allow the system to teach itself by utilizing an online environment as a perpetual training mechanism with constantly updating input (Raducanu & Vitri, 2008). Generally humans are social beings and interact with each other on a regular basis in situations which are optimized by instantly recognizing familiar faces. Perhaps the specialized procedures that are used for this effective process developed as an evolutionary adaptation, the ability to categorically differentiate between friend and foe. Currently due to the limitations of artificial vision system recognition of human faces research in the area has been restricted to recognition and classification of facial expressions. The difference between the two areas of study is that they require a different type of recognition. Facial recognition requires an artificial system to recognize the same face based on differences in angle of image, differences in facial expression and other variables that change in a face over time. To recognize emotion or facial expression only requires that the face be able to be oriented based on general features of a face and then compare the expression displayed to common templates and then render a decision of what expression is displayed based on the closest match (Ilbeygi & Shah-Hosseini, 2012). There are billions of faces, but only a few dozen facial expressions. This combined with the variability in individual faces over time makes reading expressions easier than reading individual faces for an artificial system. For humans both processes are relatively simple and perhaps even an entirely innate occurrence.

Remarkably, artificial intelligences have much less trouble with determining emotion in human speech. The capability of categorizing emotion in speech is important because of the role that emotional information plays in human communication. For an artificial system to effectively interact with a human it is important for the system to be able to extract more information from speech than the words that are being spoken. An experiment conducted in 2010 found that when the same emotional categorization test was administered to an artificial intelligence, trained from sample speech to categorize emotion, and human subjects that the performance was almost identical. The human subjects and the artificial intelligence were given identical recorded speech clips and a set of possible categories into which the speech could be placed. The researchers claim that they utilized a “psychologically inspired strategy” to conduct their experiment. The researchers used universal emotional states as the categories from which the experiment participants could select. The experiment did not attempt to have the artificial intelligence interpret meaning of the speech, only to place it into an emotional category (Shaukat & Chen, 2010). Artificial intelligences use communication as a mechanism for obtaining information. Humans utilize communication for more than transmitting information. Components of human speech such as humor and sarcasm are essentially useless input to an artificial system and the capacity to categorize emotion in human speech may be an important step in creating a mechanism to filter out such communication and perhaps to eventually understand it or participate in it.

Another area of categorization that carries distinct properties for both types of cognitive systems is the use of labels. An artificial system can be given the language label for a specific collection of features that compose a category, but the label itself has no meaning to the artificial system beyond the scope of a name for a group of features. On the other hand, a human cognitive system often will develop an opinion or predisposition regarding a category based on the label. A recent study on this phenomenon revealed that human perceptions of members of a category are often shaped by the interpretation of the labels used. Humans sometimes react to certain labels with an emotional reaction based on previous experience with the label (Foroni & Rothbart, 2011). A feature of human cognitive ability is spreading activation. This feature activates nodes that are associated with a node that is retrieved, in a task called priming. For humans this improves the speed of retrieval for items that are related. This is one of the ways in which a label can carry additional meaning for a human cognitive system (Reitter, Keller, & Moore, 2011). Artificial cognitive systems that function based on a connectionist model or that integrate connectionist and situated cognition concepts are often given the ability to associate multiple categories with each other in a meaningful way. In this form artificial cognitive systems share a capability with their biological counterparts.

The difference between biological cognitive systems and their artificial counterparts that makes artificially intelligences “not quite human” is the fact that all artificial intelligences are driven by algorithms and trained by theories that are designed to simulate human cognition and humans are essentially cognitive machines. Artificial intelligences can conduct categorization by being driven by neural networks (Barsalou, 2008), utilizing fuzzy logic algorithms to sort features (Liu & Ren, 2010), simulating cell assemblies (Huyck, 2007) or simple pattern matching based on features that are significant to humans (Shaukat & Chen, 2010; Ilbeygi & Shah-Hosseini, 2012). Weighting of features is an attribute that all cognitive systems use in categorization. Humans lower the cognitive weight on unused categories and concepts that are determined to be unimportant to the point where they are forgotten and no longer impact categorization, except in cases where the category is explicitly retrieved through priming links (Reitter, Keller, & Moore, 2011). Artificial cognitive systems lack the ability to forget. This can be seen as an advantage, as the artificial system should never be unable to categorize an item, no matter how many cycles have occurred since it was last recognized. It can also be a disadvantage as it may leave data in place longer than it is useful and ultimately have a negative impact on categorization efficiency (Kirstein, Wersing, Gross, & Körner, 2012). While human cognition has been studied thoroughly there is still no exact understanding of how exactly human cognition functions. Categorization is a basic component of cognition which is the easiest to simulate and perhaps the most essential to further improvements in the ability of artificial cognitive systems to mimic human cognitive behavior (Huyck, 2007). As Stevan Harnad states in *To Cognize is to Categorize*, “We organisms are sensorimotor systems” which interact with the world using sensory surfaces and the affordances therein (Harnad, 2005). Essentially he argues that categorization is the basis of all cognition and categorization is based on sensorimotor interactions. While artificial intelligence systems are being given electronic equivalents to all of the sensors that humans have, few if any cognitive systems have been given all of them simultaneously. To do so introduces too many pieces of data for the current level of artificial intelligence technology to utilize in a productive way. Humans integrate all of their sensory input gracefully through the use of attention and short term sensory memory without necessary consciously acknowledging each piece of information. Artificial systems, or their subsystems, must process this data to extract what is useful. Artificial systems are not as graceful at attention processes as their human counterparts. Attention is one of the most basic category systems in humans as it allows for the separation of relevant information from irrelevant information. Artificial systems do not have a concept of meaning for their input, only the existence of new data, and therefore have difficulty with determining what is relevant to their task or not unless specifically primed or trained for a specific type of attending (Carotaa, Indiverib, & Dantec, 2004). The differences between artificial and biological cognitive systems is not as wide as it once was and the only boundary that really still exists between them for categorization is the difference between eons of biologically evolving hardware and decades of software algorithms.

Despite the differences in the way that biological and artificial cognitive systems approach categorization there are some common features between their approaches. The most essential similarity is the use of features. Features are identifying pieces of information about an item or a concept that makes it distinct from anything else. All cognitive systems must decide what features are important and place a certain weight on them, whether it is a computational weight or being weighted in number of neural links to the feature (Harnad, 2005). An example of a specific artificial cognitive system acting very similarly to a biological system is DARPA's Big Dog, which adapts to situations of losing balance or sliding with its adaptive legs. It is reasonable to assume that Big Dog's cognitive system can learn from experience much like an organic cognitive system would. If Big Dog encounters a particular set of variables and is able to successfully overcome the situation using a particular strategy, then those variables may become a distinct category in which the successful solution may become the recommended first solution if those variables are encountered again (Hornyak, 2012). Another common element between biological and artificial cognitive systems is the reliance upon sensory data. While an artificial system can categorize text, it is only categorizing the text itself, no other meaning is being recognized or encoded. All of the meaning extracted from categorizing text by an artificial system is based upon meaning that is supplied to that text by the context in which it exists, which it is given by human cognition (Guoa, Shao, & Hua, 2010). To obtain meaning for a reliable categorization does not necessarily require direct sensory contact, as information may be transmitted through "hearsay" (Harnad, 2005). This method may work well for human cognitive systems as there is a standard method of transmission, in the form of language, which the cognitive system already functions in. Most artificial intelligence systems are not designed to communicate in a generalized way so that they would be able to communicate information directly to other artificial systems, except through a direct data transfer which is not analogous to any function that biological systems are capable of performing. In the same vein though, direct copying of categorical data is not something that humans are capable of and as such artificial systems have an advantage in retaining cohesion of meaning between each other. However, theoretically that cohesion would be broken as soon as one of the artificial systems gained new experience with that category. A system still learning would continue to adapt the category to be compatible with new experiences (Kirstein, Wersing, Gross, & Körner, 2012). From a technological standpoint, because most artificial systems use different methods for creating and storing category data, two dissimilar artificial systems would be unlikely to be able to share information.

As artificial cognitive systems evolve in both their background programing and their ability to adapt they are becoming closer to being able to mimic and surpass human cognitive abilities for categorization. Artificial and biological cognitive systems have some important distinctions in how they carry out tasks, but as research moves forward the gap between human cognition and artificial intelligence is growing very narrow. Categorization is one of the most basic ways in which a cognitive system forms understanding of items in its environment and is able to become a participant in its environment. The future of artificial cognitive intelligence may be one in which the artificial systems can interact with biological organisms as easily as we interact with each other. Artificial systems can categorize for themselves or emulate human categories, but no matter how much an artificial system resembles a biological cognitive system the question of whether or not they have a ‘true’ understanding of the meaning of those categories remains. Harnad states that ontology is beyond the scope of cognitive science (Harnad, 2005). Perhaps until artificial cognitive systems are theorizing the meaning of their own existence we should respect the differences and embrace the similarities.

BIBLIOGRAPHY

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, 147–161.

Barsalou, L. W. (2008). Cognitive and Neural Contributions to Understanding the Conceptual System. *Current Directions in Psychological Science*, 91-95.

Carotaa, L., Indiverib, G., & Dantec, V. (2004). A software–hardware selective attention system. *Science Direct*, 647 – 653.

Cohen, H., & Lefebvre, C. (2005). *Handbook of Categorization in Cognitive Science.* Maryland Heights, MO: Elsevier Science.

Foroni, F., & Rothbart, M. (2011). Category Labels and Category Boundaries: When Does A Category Influence Percieved Similarity of Category Members. *Social Cognition*, 547-576.

Glushko, R. J., Maglio, P. P., Matlock, T., & Barsalou, L. W. (2008). Categorization in the Wild. *Trends in Cognitive Sciences*, 129-135.

Guoa, Y., Shao, Z., & Hua, N. (2010). Automatic text categorization based on content analysis with cognitive situation models. *Information Sciences*, 613–630.

Harnad, S. (2005). To Cognize is to Categorize: Cognition is Categorization. In C. Lefebvre, & H. C. (Eds.), *Handbook of Categorization in Cognitive Science* (pp. 19-43). Elsevier.

Hornyak, T. (2012, 02 08). *DARPA takes bigger BigDog out for walkies*. Retrieved from CNET: http://news.cnet.com/8301-17938_105-57373477-1/darpa-takes-bigger-bigdog-out-for-walkies/

Huyck, C. R. (2007). Creating hierarchical categories using cell assemblies. *Connection Science*, 1-24.

Ilbeygi, M., & Shah-Hosseini, H. (2012). A Novel Fuzzy Facial Expression Recognition System Based on Facial Feature Extration from Color Face Images. *Engineering Applications of Artificial Intelligence*, 130–146.

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is Gendered: How Covarying Phenotypes and Stereotypes Bias. *Journal of Personality and Social Psychology*, 116–131.

Kirstein, S., Wersing, H., Gross, H.-M., & Körner, E. (2012). A life-long learning vector quantization approach for interactive learning of multiple categories. *Neural Networks*, 90–105.

Liu, X., & Ren, Y. (2010). Novel artificial intelligent techniques via AFS theory: Feature selection, concept categorization and characteristic description. *Applied Soft Computing*, 793–805.

Raducanu, B., & Vitri, J. (2008). Face Recognition of Artificial Visual Systems: A Cognitive Approach. *International Journal of Pattern Recognition and Artificial Intelligence*, 899-913.

Reitter, D., Keller, F., & Moore, J. D. (2011). A Computational Cognitive Model of Syntactic Priming. *Cognitive Science*, 587–637.

Shaukat, A., & Chen, K. (2010). *Emotional State Categorization from Speech: Machine Vs. Human.* Cornell University Library.

Yamauchi, T. (2009). Finding abstract commonalties of category members. *Journal of Experimental & Theoretical Artificial Intelligence*, 155–180.