Probabilistic Strategy Selection for Flexible Cognition

Jennifer M. Roberts and Gregory Marton

CSAIL, Massachusetts Institute of Technology; jenmarie@mit.edu

While current artificial systems must be custom tailored to operate within a single domain of expertise, the human mind readily adapts to multiple domains. In moving toward artificial systems that display similar flexibility, we become interested in how efforts to learn about the mind can support the development of flexible artificial systems, and how the development of these systems can contribute to our understanding of how the mind works.

When contrasting natural and artificial systems, we find that people learn and maintain multiple competing strategies for performing the same task, while current software lacks this ability. When adding numbers, asking questions, referring to objects in the environment, or discerning spatial relationships, people develop and flexibly use a variety of approaches.

This paper describes a proposed research program that focuses on computationally modeling the ability to maintain and use multiple strategies for performing the same task. We first point to developmental evidence as motivation for this research program. We then describe a few computational efforts that seem to fall within this vein. Finally, we describe preliminary work within the field of lexical semantics based on a new proposed model of “word strategies”.

Developmental Evidence that People Rely on Multiple Strategies

Developmental studies provide evidence that people actively maintain and use more than one strategy for performing the same task (see Siegler 2006 for a review). As the first of three examples, Siegler studied how children learn to add and determined that children develop a set of different strategies and switch between those strategies when performing adjacent addition problems. For example, a child who is learning to add may go through a period when they are equally likely to calculate "2 + 4" by using either of the following two strategies. First, they may use their fingers to count out each addend (counting "1, 2" and "1, 2, 3, 4") and then count the number of extended fingers ("1, 2, 3, 4, 5, 6") to compute the sum. Alternatively, they may extend four fingers while saying "4" and then extend fingers while counting "5, 6". Surprisingly, children often go through a period where they know both efficient and inefficient strategies for performing the same task, but continue to use both strategies intermittently.

Similarly, Crain and Thornton observed that three year olds sometimes alternate between two strategies when asking questions about objects, such as questions that contain “who” or “what” (Yang 2006, p169). For example, a three year old might ask, “Who do you think who is in the box?” During the same period, the three year old might ask such questions using the correct English grammar as well.

Even when a child appears to produce the same output, they may be using different strategies that rely on different types of representation and levels of understanding. For example, French 4-year-olds correctly use the word “mes” to indicate ownership of a group of objects (“mes” is a plural form of “my” that roughly translates to “all of my”). Because the plural marker “s” is silent at the end of a noun, “mes” actually provides all of the information about whether a noun is plural or not. Around the age of 6, children start replacing “mes” with “toutes les miennes” (literally “all of my”). Shortly afterwards, they return to using “mes”, but only at this point can they explain that “mes” indicates both plurality and possession. Thus, while French children may produce the same output both before and after the “toutes les miennes” stage, the strategies they use involve differing levels of representation and understanding (Karmilloff-Smith 1992, p49).

Computational Models Using Multiple Strategies

The ability to flexibly develop and select between multiple strategies helps the human mind adapt to and respond to its environment. The following computational models provide steps toward modeling this capability, and they serve as examples of computational efforts that align with the proposed research program for modeling multiple human-like strategies for performing the same task.

First, Shrager and Siegler (1998) developed a computational model of how 4- to 5-year-olds learn to add that focuses on accurately modeling how the children’s strategies evolve over time. The SCADS model started with knowledge of the simplest addition strategy mentioned in the previous section and then modeled the development and adoption of new strategies. Strategy development relied on detecting and eliminating redundant behavior and on preferentially moving toward efficient sub-operation orders. During simulations, the model used a meta-cognition system to discover strategies in the same order that children do and probabilistically selected between current strategies to adopt strategies at a pace that was qualitatively similar to the children’s pace. Despite the model’s impressive agreement with human results, this appears to be the only model of its kind.

Charles Yang (2006) proposes a model for grammar learning in which kids first develop strategies that are consistent with the rules of universal grammar and then slowly weed out strategies that do not agree with their native language. Based on universal grammar, the strategies that children develop produce sentences with word orders that are correct in some existing language, but not necessarily correct in the language a child is trying to learn. For example, questions of the form "Who do you think who is in the box?", where the question word repeats twice, are grammatically correct in German and Hindi, but not in English. As children explore strategies consistent with universal grammar, they search through the space of possible grammars to converge upon the grammar of their native language.

Sajit Rao (1998) applies the idea of multiple strategies to image processing, using "visual routines" to extract spatial relationships. Shimon Ullman (1996) reviews ideas for multiple strategies that can perform the same visual task.

While not yet a computational model, Karmiloff-Smith (1992) suggests a theory of Representational Redescription that provides insight into how the mind's strategy-development mechanisms might relate to the development of multiple representations. According to this theory, the mind iteratively redescribes representations into more conscious, malleable pieces. As a motivating example, she discusses learning the Rubik's cube. First, she learned how to solve the problem without having any conscious awareness of how to regenerate the solution. After that, she had to watch herself repeatedly solving the cube in order to start slowly recognizing intermediate stages within her solution. Once she was able to consciously recognize a few stages, she built on this knowledge to eventually explain to herself how to solve the cube. Likewise, she was eventually able to index her solution so that she could reproduce the solution regardless of what state the cube started in. Karmiloff-Smith describes this process of Representational Redescription as a process in which unconscious representations become redescribed into more conscious representations, which later become redescribed into more flexible representations that can be consciously manipulated to create more complex behavior. Computationally, Representational Redescription can be thought of as iteratively rewriting the procedures and knowledge structures used to perform a particular task into successively more flexible and more efficient forms.

As a final example, McShane (2004) outlines semantic procedures for interpreting approximation words like "around" and "nearly". She provides a procedure for interpreting "around" that takes the numerical value being approximated and adds an uncertainty of 7%. She notes, however, that while this strategy of assuming a 7% error works in a remarkable number of contexts, people use very different approximation strategies when talking about young age, clock-times, or people's heights. This provides evidence that people also use multiple strategies when interpreting approximations.

In each case described above, sets of computer programs or procedures can in principle fulfill the same goals as the sets of strategies that people employ. From a computational point of view, learning new strategies translates into searching the space of programs. However, the hypothesis space of programs, even that of correct programs, is infinite. The fitness landscape contains sharp peaks, so most programs will not work at all, small changes in a program may result in large changes in outputs, and small changes in outputs may require large changes to a program. Nonetheless, we claim that human learners successfully solve this problem and spend significant amounts of time creating new programs in hopes of finding ones that are fitter than those they currently have. Thus, we advocate a research program centered around developing artificial systems that flexibly employ and ultimately discover multiple strategies, i.e., multiple computer programs, for performing a particular task.

Probabilistic Strategy Selection for Semantic Interpretation

In this section, we describe preliminary efforts for a new line of research that falls within the proposed research program. As mentioned above, the ability to develop and probabilistically select between multiple strategies appears to be an important component of flexible cognition. As a step toward creating an artificial system with this capability, we will explore probabilistic selection between strategies within the domain of lexical semantics. In this preliminary work, we will use multiple strategies to provide computer-interpretable semantics for time words using an existing semantic parser called Sepia (Marton 2003).

Temporal words provide a test bed for exploring reasoning in the face of ambiguity and testing the use of multiple human-like strategies, because we have a coherent understanding of reasonable semantics for time. Some attention has been paid to interpreting temporal expressions in ACE (Doddington 2004) and in TempEval (Verhagen 2007), providing a consensus semantic target for correct interpretations.

Traditionally, computational work on interpreting words, a task known as word sense disambiguation, focuses on selecting the best definition for a word from amongst a set of static dictionary definitions (Agirre & Edmonds 2006). In contrast, we wish to focus on identifying the best ways to interpret definitions.

McShane (2004) concretely expresses word senses in terms of procedures and notes that using a single procedure, even for a fixed word sense, will not work well in all contexts. However, she defers the task of specifying more comprehensive procedures to the future, not wanting to become a "lexical semanticist … [pursuing] an endless path of potentially unneeded research about approximation." As an alternative to the manual work she alludes to, we propose an approach for learning these lexical semantics.

People plausibly maintain multiple strategies for processing and interpreting time words. For example, if today is Tuesday, "next Thursday" could alternately be interpreted by identifying the next Thursday to occur (two days from today) or by identifying the Thursday during "next week" (9 days from today). We model these interpretation strategies as two procedures, one of which increments today's date until it finds "Thursday", and the other which identifies the Thursday within "next week". These two interpretations for "next" can be viewed as two different word senses. Different ways of implementing each word sense can be considered different strategies. Likewise, one strategy can encompass two traditional word senses.

Multiple strategies can also distinguish between different types of conceptual understanding. People learn both intuitive perception-based concepts, where one day might correspond to a cycle of daytime and nighttime, and definition-based concepts, where one day equals 24 hours (Reigeluth 2007). Intuitive interpretation strategies persist throughout life, as a four-year-old might describe tomorrow as "the time after my next long sleep" and a college student who is pulling an all-nighter might maintain that tomorrow does not start until after they sleep or until after the sun comes up. We will model such competing strategies by creating multiple procedures for each word and probabilistically selecting among them based on how well the final interpretations of phrases fit human judgments.

We will propose a mechanism for automatically generating novel semantics for each word, semantics that the computational learner will then evaluate and perhaps eventually "adopt". This model differs from usual word sense disambiguation in that the individual strategy for each word is less important than how it combines with the words nearby. As such, we propose a procedural notion of word sense – a word strategy – that is, to our knowledge, new.

Expected Contributions and Future Work

In this paper, we propose a research program centered on computationally modeling the ability to maintain and use multiple strategies for performing a given task. Our work with time semantics provides a test bed for exploring whether probabilistic selection between multiple strategies can provide more flexible performance. This test bed will also allow us to explore how multiple strategies may interface with multiple types of representation, and it will provide a language for thinking about strategies and how they can adapt over time.

Our approach centers around a new model of procedures as word strategies that makes several questions concrete: How do strategies build on one another? Can we identify parameters across which strategies can develop (as in the universal grammar case)? How does our ability to cluster and identify frequency information influence or interact with these capabilities? How do strategies relate to the construction and utilization of representations?

Acknowledgements

J.M.R. thanks the Fannie and John Hertz Foundation and the National Science Foundation for their financial support.

References

Agirre, E., and Edmonds, P. 2006. *Word Sense Disambiguation: Algorithms and Applications (Text Speech, and Language Technology)*. Kluwer.

Doddington G., Mitchell A., Przybocki, M., Ramshaw L., Strassel, S., and Weischedel R. 2004. The Automatic Content Extraction (ACE) Program — Tasks, Data, and Evaluation. In *proceedings of LREC*, 837-840.

Karmiloff-Smith, A. 1992. *Beyond Modularity: A Developmental Perspective on Cognitive Science.* Cambridge, MA: MIT Press.

Marton, G. 2003. *Sepia: Semantic Parsing for Named Entities*. Master's Thesis, Dept. of EECS, MIT.

McShane M., Beale S., and Nirenburg S. 2004. Some meaning procedures of ontological semantics. In *LREC 2004: Fourth international conference on language resources and evaluation*, 1885–1888. Lisbon, Portugal.

Rao, S. 1998. *Visual Routines and Attention*. Ph.D. diss., Dept. of EECS, MIT.

Reigeluth, C. M. 2007. Order, First Step to Mastery: An Introduction to Sequencing in Instructional Design. In *In Order to Learn: How the Sequence of Topics Influences Learning*, Ritter, F., Nerb, J., Lehtinen, E., and O'Shea, T. eds, 19-40. New York: Oxford Univ. Press.

Siegler, R. S. 2006. Microgenetic Analyses of Learning. In *Handbook of Child Psychology 6th Ed, Vol 2: Cognition, perception, and language*, Damon, W., Lerner, R.M., Kuhn, D., and Siegler, R. S. eds, 464-510. Hoboken, NJ: Wiley.

Ullman, S. 1996. *High-level Vision: Object Recognition and Visual Cognition.* Cambridge, MA: MIT Press.

Verhagen, M., Gaizauskas, R., Schilder, F., Hepple, M., Katz, G., and Pustejovsky, J. 2007. *Semeval-2007 task 15: Tempeval temporal relation identification.* 4th International Workshop on Semantic Evaluations.

Yang, C. 2006. *The Infinite Gift: How Children Learn and Unlearn the Languages of the World.* NYC: Scribner.