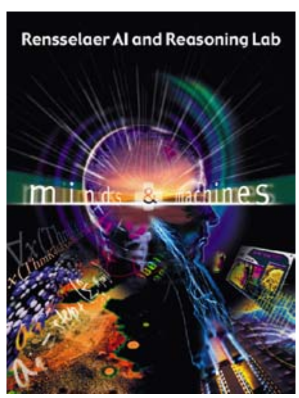
Learning and “Poised-For” Knowledge

⋆ white paper ⋆

Selmer Bringsjord & Kostas Arkoudas
Department of Cognitive Science (SB, Chair)
Department of Computer Science
Rensselaer AI & Reasoning Laboratory (SB, Director)
Rensselaer Polytechnic Institute (RPI)
Troy NY 12180 USA
{selmer,arkouk}@rpi.edu
SB’s web site: http://www.rpi.edu/∼brings

4.28.04



Rensselaer

Suppose we ask you to learn about a topic with which you are unfamiliar — astronomy, say.[1] How would you go about acceding to our request? Perhaps it's safe to say that if something is at stake (and to ensure that, let's suppose you must incorporate some substantive discussion of astronomy within a larger team project of pivotal importance to your career, and the careers of your associates), your first move will no doubt be to *read.* Perhaps you search the Web and read documents about our solar system, black holes, and so on; or perhaps you purchase a book or two, or more; at any rate, you will almost certainly seek to learn by reading.

In turning to written text and diagrams to learn, you will have done something quite interesting from the standpoint of learning as it has been rigorously studied in computer science, cognitive science, and AI: you will have done something that, traditionally, *isn't* considered learning in these fields.[2] In them, to learn is almost invariably to produce an underlying function f on the basis of a restricted set of pairs $(a_1, f(a_1))$, $(a_2, f(a_2))$, $\ldots$, $(a_n, f(a_n))$. For example, consider receiving inputs consisting of 1, 2, 3, 4, and 5, and corresponding range values of 1, 4, 3, 16, and 25; the goal is to "learn" the underlying mapping from natural numbers to natural numbers. In this case, assume that the underlying function is n^2, and that you do "learn" it. While this narrow model of learning can be productively applied to a number of processes, the process of reading isn't one of them. Learning by reading cannot (at least for the foreseeable future) be modeled as divining a function that produces argument-value pairs. Instead, your reading about astronomy can pay dividends only if your knowledge has increased in the right way, *and* if that knowledge leaves you poised to be able to produce behavior taken to confirm sufficient mastery of the subject area in question. This behavior can range from correctly answering and justifying test questions regarding astronomy, to producing a robust, compelling presentation or paper that marks success on the larger project we have assumed you and your team to have been tasked with.

How would we go about ensuring that, thanks to your reading, you

[1]The topic, as will soon become clear, is immaterial. The moral of the story we are about to tell could be derived from attempts to learn about horse-racing, golf, quantum mechanics, hurricanes — whatever. The essential ingredient is that when it comes to the topic to be investigated, you are a non-expert. If by coincidence astronomy happens to be your avocation, please imagine that we couched this white paper in terms of a topic suitably foreign to you.

[2]Of course, we don't deny that cognitive science and education have long been intertwined, but their relationship is not based on formal, mechanized accounts of high-level learning — the kind of accounts that are required for progress in computer science and engineering.

have learned about astronomy, in a deep and lasting way? One way is to test you. We could give you what was just mentioned: multiple-choice items, or short-answer ones, or perhaps essay questions that force you to integrate what you have learned with other knowledge, and other topics. Of course, tests like this are nothing new. They are effective in a limited way, but the fact of the matter is that they *are* limited — chiefly because, at best, they merely *reflect* the learning that has taken place. This explains why the field of psychometrics has tended not to concern itself with what happens *inside* the mind/brain. It's worth noting that though it isn't a likely occurrence, multiple-choice items can in principle be answered correctly in the complete absence of learning having taken place. Even in the case of essays, the problem remains: automatic essay graders can be given nonsense text that nonetheless contains the sought-for features, causing high grades to be returned by these systems; what should be an F is given an A. And it is certainly possible in principle to generate a coherent essay that receives high marks, where the generation is based on a shallow bag of programming tricks, instead of suitable underlying knowledge. The same possibility exists no matter *what* the output is. At any rate, even if we can trust that the output is genuine, surely it seems rather inefficient to check that learning has taken place by waiting for this output to be crafted. Finally, even if we agree that the output infallibly indicates that deep learning has occurred, we will inevitably find ourselves studying that output in order to determine if it reflects genuine understanding — if it indicates that the right sort of cognitive structures are in place.

By our lights, therefore, a much better and more efficient test of learning in the case of the astronomy scenario is to look directly inside your brain after your reading has concluded. If we find the "right stuff" in there, we could confidently declare that you had indeed achieved deep and durable learning. Now, looking inside someone's brain like this isn't yet feasible from the standpoint of cognitive neuroscience, but if our aim is to build *machines* able to learn by reading (and also by other effective high-level techniques much more robust than function-based learning), then, given the research and development we propose, the "look inside" approach *is* feasible. If a machine is to have truly learned about astronomy by assimilating a book on the subject, then when we examine the "brain" of this machine, we will be able to find what we call "poised-for" knowledge, or just *p.f.-knowledge*. The basic idea is that: p.f.-knowledge is **knowledge poised for the semantically correct generation of output that would provide overwhelming psychometric evidence that deep and durable learning has taken place**. It's key that the generation be *semantically correct*. This condition ensures that, corresponding to output about, say, black holes, is knowledge that corresponds directly to the *meaning* of such entities.

There are six distinguishing features of p.f.-knowledge:

1. *Mixed Representation Mode.* Poised-for knowledge is not simply syntactic or traditionally "logicist" in nature. It must include *mental models*, in Johnson-Laird's sense. For example, your reading about astronomy will have produced knowledge in the form of mental images of our solar system and the orbits of our planets and their distance from our sun. This is of course to be expected, since the material you will read (as indicated above when we said 'written text and diagrams') will itself include not only straight text, but pictorial representation of galaxies and planetary orbits and so on.

2. *Tapestried.* Poised-for knowledge of a new topic is woven together with knowledge of previously learned topics. If you have really and truly learned about astronomy, then your new knowledge has been integrated with old knowledge. For example, you may now grasp the nature of a supernova by relating it by analogy to what you understand about ordinary combustion.

3. *Extreme Expressivity.* Poised-for knowledge is highly expressive knowledge. In the case at hand, it will include knowledge of who wrote the material you have read, knowledge about why I asked you to investigate the topic in the first place, knowledge about my knowledge of the topic, knowledge about my beliefs about your knowledge before you launched your investigation, and so on. In addition, because the texts in question are themselves rich and robust, p.f.-knowledge in this case will include not just doxastic operators (which pertain to belief and knowledge), but also operators reflective of historical facts about the progression of ever-better theories of the cosmos: the transition from geocentric to heliocentric views of our system, the transition from Newtonian mechanics to relativity theory and the empirical phenomena that confirmed this theory after Einstein proposed it, and so on.

4. *Mixed Inference Types.* The knowledge in question will be poised to allow for various types of justification to be articulated. Every now and then justification may come in the form of a well-reasoned deductive argument, but this is surely rare. P.f.-knowledge is structured to allow for educated guesses from the learner, and also inductive explanations, suggestive diagrams, and so on.

5. *Deep Connection to Natural Language.* The simple sort of learning that has tended to dominate computer science, cognitive science, and AI has tended to be divorced from natural languages, such as English. This kind of impoverished learning has often been based on purely numerical values, and when declarative information is involved, it tends to be disconnected from the power and subtlety of natural language. In the case of p.f.-knowledge, we know that it can give rise to explanations, presentations, answers, essays, justifications . . . *in English.* P.f.-knowledge is knowledge that is poised for communication expressed in English.

6. *Multi-Agent Structures.* In the scenario we have invoked, you were imagined to be part of a team working on a larger project. P.f.-knowledge takes account of multiple agents, their individual agendas, and their shared agendas. Poised-for knowledge of astronomy must include knowledge about how your knowledge is to be best employed by a team of agents working in concert. In addition, sometimes multi-agent structures will apply to a single agent, as when you (qua agent 1) consider whether what you (qua agent 2) now know about astronomy is sufficient to meet the demands at hand.

We seek to build an intelligent system, MARMML (pronounced "mar-mul"), designed from scratch to advance the cause of robust machine learning by formalizing and mechanizing (i) these six aspects of p.f.-knowledge, and (ii) the use of such knowledge to produce appropriate output.

A Note on The Technical Side

We have authored numerous technical papers, books, and systems that set the stage for an attack on sophisticated learning through the mechanization of both p.f.-knowledge and what it can generate in the form of the MARMML system. Some of this prior work has involved key collaborators, for example (related to the Mixed Representation Mode dimension), Johnson-Laird, the "grandfather" of mental models within cognitive psychology/science. Full references to and discussion of prior research and development goes beyond this brief white paper, which is intended to be both self-contained, and to be pitched at a level that allows specialists and non-specialists alike to learn the core of what we have in mind — by reading.