Getting Meaning into the Machine

Yorick Wilks, *University of Sheffield*

My own life in AI does not go back, quite, to the Dartmouth Conference, but I do recall giving a BBC talk in the sixties, when that was still possible on public radio in Britain, with the title "Getting Meaning into the Machine." I dare not look back at the content, but I still like the title since I realize I took it as a lifetime's goal.

What intrigues me now is how that meaning might (or might not) have changed with time. What I was expressing then was the belief that some form of compositional semantics, distinct from logic and formal linguistic semantics, would capture enough of meaning to carry out "language understanding" tasks—from dialogue to translation. There was quite a group of such people thinking such things through the seventies and even eighties—of whom Roger Schank was the best known. It didn't work out; although like all AI strands, it never quite died either. It had absolutely no empirical foundation and its content was the direct product of intuition, just as core AI and much of linguistics have continued to be, though in a different style. All you could point to by way of justification was that counts over machine-readable dictionaries showed that the set of *primitive concepts* (in systems like my own, called Preference Semantics in the 1970s) corresponded pretty closely to the most frequently used defining words in big dictionaries. But that was no basis on which to found a theory of meaning.

The approach survived best at the core of a number of Japanese machine translation systems (probably still running) based on an interlingual representation between languages, such as Fujitsu's. But some notions in it survived into the general stream of consciousness:

- that some determinable notion of *preference*, or affinity between concepts or words, existed and
- that semantic theories had to have some procedural notion within them and not be mere static representations, because of the constant change and adaptation of word meanings over time.

> The Semantic Web is a development of great importance to AI as a whole—even though we still dispute what it means and how it can come into being.

False starts

Since that time, two waves of empiricism and machine learning have gone through language processing. The first, *connectionism*, looked promising but never scaled up. The second wave was set off by Frederick Jelinek's translation system at IBM in 1990. That system attempted to move successful computer speech techniques into language processing itself and didn't work above 50 percent correctness over sentences. But it set off a huge range of applications in every area and level of language processing—right up to the assignment of speech acts and determination of inference patterns—on the way catching up notions such as semantic preference, dear to my old radio broadcast. One way to look at this movement is to either say that it's *unsupervised learning*—that can give you concepts from nothing but can't tell you what they are—or say it's all supervised learning over annotations of text. The latter consists of humans annotating onto training text whatever they want, from semantic types to parts of speech, and then the system learning them from more corpora. However, this is a circular process and still rests entirely on intuition as to what, say, the annotation `company` actually means when attached to a word.

Both these views are wrong: unsupervised learning might not name concepts, but it's a remarkable demonstration of the power of huge computations over corpora to capture undeniably semantic notions. Look at http://labs.

google.com/sets to see inductions of semantic types from your own instances; that was high-class research only 15 years ago. Annotations are more complex, but the only way to justify their sense is to show (as we do pretty effectively now) that we can replace annotation by human intuition with machine-learning-based automated annotators working from seed examples. That fact alone helps justify the annotator term having meaning, because it's then a learnable notion. This is a vital idea for giving meaning progressively to the so-called "upper level" concepts in ontologies, as used in language processing, scientific knowledge structures, and in the future Semantic Web (SW). We still use intuition to write down these upper-level concepts, and intuition can have its own validity in scientific areas if done by experts—who else can write a map of biology? But if we don't ground it in something firmer, it's as much at risk as all AI knowledge structures.

The Semantic Web

What we're seeing in the SW is a growing together of these upper conceptual levels based on the name spaces and concept triples derived from texts by *information extraction*—a successful shallow technology for extracting items and facts. Information extraction now rests wholly on automated annotation's success and has now been successfully extended to the automatic induction of ontologies. I believe that the top and bottom levels will grow together and that interpretation or meaning will "trickle up" from the lower levels to the higher. This is the only way I can imagine the higher conceptual labels being empirically justified. This process reminds me of the concept of *semantic ascent,* which Richard Braithwaite pioneered in the 1950s as a description of the way in which interpretation trickled up scientific theories from observables (such as cloud-chamber tracks) to unobservables (such as neutrinos). I can't see any other route from the distributional analysis on which the revolution in language processing rests to the interpretation of serious concepts. This also reminds me of Immanuel Kant's great dictum synthesizing rationalism and empiricism: "Concepts without percepts are empty; percepts without concepts are blind."

The SW is a development of great importance to AI as a whole—even though we still dispute what it means and how it can come into being. Many seem to believe that it means "good old-fashioned AI" is back in a new form—a rebranding of the old tasks of logic, inference, agents, and knowledge representation. It's true that core AI tasks have come to something of an impasse: we haven't seen them marketed in products much in the last 50 years. But a key feature of the SW, I assume, is that its delivery must be gradual, coming into being at points on the World Wide Web, possibly starting with the modeling of biology and medicine. I can't imagine how the SW could start somewhere completely new, without being piggybacked on the WWW; however, its implementation will achieve much more than having the same texts "annotated with their meanings."

The key possibility the SW offers to traditional AI is delivering some of its value in a depleted form initially, by trading representational expressiveness for tractability, as some have put it. The model to follow here could be search technology and machine translation on the WWW (or even speech technology): each is available now in imperfect forms that we can't imagine living without. This might all seem obvious, but machine translation has only recently crossed the border from impossible (or failed) to commonplace. Isn't it far better for a field to be thought useful, if a little dim at times, than impossible or failed? We'll need to choose Web services using the SW that will merely be a nuisance, not critical, if they fail. My own current interests are in lifelong personal agents, or *companions,* as a new type of conversational interface to the Web. It shouldn't matter if our agents sometimes err or mislead—any more than when people do—as long as we have other ways of checking information.

This view of the future of the SW is personal and partial; many don't see the need to justify the meanings of logical predicates or ontological terms now any more than they did when they set out in AI and representation in the sixties. But the history of the Cyc project (www.cyc.com) is a good demonstration, if we needed one, of why this view can't be a foundation for AI in the long term. Similarly, the SW movement also promotes the idea that meanings will be saved or preserved by trusted databases of objects (URIs), referential items in the world, rather like the way digit strings ground personal phone numbers in a database. But this way out won't protect knowledge structures from the changes and vagueness of words human beings really use. Hilary Putnam considered this problem in the sixties and declared that scientists should therefore be the ultimate "guardians of meaning." As long as they knew what "heavy water" really meant, it didn't matter whether the public knew—and was perhaps better they didn't. But people call heavy water "water" because it's indistinguishable from water; otherwise, they would have just called it "deuterium dioxide." We, the people, are the guardians of meaning, and we should be getting meaning into the machine—probably via the SW—by doing it our way. ❐

Yorick Wilks is a professor of artificial intelligence and head of the Natural Language Processing Group at the University of Sheffield. Contact him at Univ. of Sheffield, Regent Ct., 211 Portobello St., Sheffield, S. Yorkshire S14DP, UK; yorick@dcs.sheffield.ac.uk; www.dcs.shef.ac.uk/~yorick.