

Natural Language: It's All About Meaning

Scott E. Fahlman

Language Technologies Institute & Computer Science Department
Carnegie Mellon University

Human language is, fundamentally, a tool for identifying a piece of meaning – some symbolic knowledge structure – in my mind and building a good, effective approximation of that knowledge in your mind. The necessary information must be transferred across a narrow, noisy, linear channel: modulated sound waves or (just in the last few thousand years) a string of text.

Over time, human language has evolved into a beautifully effective and efficient tool for this task. We have developed all sorts of clever linguistic signals that indicate how a linear string of symbols should be folded back into a complex, tangled web of entities and statements about them. Prepositions probably rank right up there with fire as a great human invention (or perhaps, if you believe Chomsky, a great invention of biological evolution).

Because the channel is so narrow, I don't want to waste precious bandwidth by telling you things that you already know. My goal is to use a stream of language to paint a thin veneer of new knowledge-structure over what I assume you already have in your head. To my computer-science colleagues, I know that I can use a word like "hashtable" without further explanation; to convey the same meaning to the (presumably) non-technical person sitting next to me on the bus might take a while. To my wife, I know that I can say, "Morty was sick again today"; to a stranger, I would go on to explain that Morty is our cat and that he has a chronic problem with hairballs. When two people meet for the first time and want to communicate, a good fraction of their speech acts are devoted to exchanging information about what each person already knows, so that subsequent conversation can be reasonably efficient (starting, sometimes, with "Do you speak English?" and then maybe "Have you heard the news today?"). If I guess wrong about what the other guy knows, I'll have to backtrack and try again.

So in natural-language communication, we want to say as little as possible and rely as much as possible on shared background knowledge. Consider how much background knowledge is required to understand a short news item like this:

"In response to last Tuesday's café bombing in Tel Aviv, Israeli tanks today crossed into the Gaza strip and demolished the houses of several suspected Hamas leaders before retreating under a hail of rocks and Molotov cocktails."

If we wanted to create a headline for this story, we might go with *"Israeli Army Retaliates"* or *"Angry Palestinian Mob"*, but none of these words (except "Israeli") appears in the original story. The passage does not even contain synonyms for these words. So to understand and label this brief fragment of text requires some knowledge of Israel, Palestine, Gaza, Tel Aviv, Hamas, tanks, and Molotov cocktails. It requires some understanding of concepts like terrorism and revenge. It requires the knowledge that crushing a house is a hostile act, likely to anger the owner and his friends. And so on...

I think it is clear that most linguists, computational linguists, and AI researchers have accepted for a long time that meaning – both the knowledge conveyed by an utterance and a large amount of background knowledge – is central to the enterprise of producing and understanding natural language. And most of them accept that meaning is really the key to most of the other operations that we want to perform on chunks of language: categorization, summarization, question answering, search, translation, and so on. But having accepted that, they see that collecting and representing very large amounts of knowledge efficiently and effectively is a very hard problem – perhaps the hardest unsolved problem in AI. So most of them go off and work on something else instead: parsing, morphology, meaning-free search engines, statistical translation based on parallel corpora, and so on. These are all valuable efforts, and some of them have yielded valuable results, but I think that many of these researchers would admit (if they feel like being honest) that they're not really attacking the heart of the monster. They're dancing around the edges of the really hard problem.

The amazing thing, really, is that a meaning-free bag-of-words model can work at all as a surrogate for meaning. The truly astounding thing is that these models (perhaps augmented with some additional shallow features) work well enough to produce truly useful tools – even supporting some huge, successful companies such as Google. But almost all of these technologies deliver only partial success – maybe 80% – because they are not really working in the same feature space as the humans who wrote the text and the humans who form the intended audience. So that last 20% error (or whatever) is going to be very hard to eliminate by meaning-free approaches.

In many domains, that's OK. We all use Google (or some competitor) every day, and we've learned to live with the fact that a lot of the documents it retrieves for every search are not relevant – these are just documents that happened to have some of the right words. The result is still very useful, most of the time. But imagine having a search engine that actually reads all these documents, understands them to some degree, and that can answer questions about them. The reason that Google returns full pages and not just answers is that it depends on its human users to figure out whether a document really contains the answer to their question, and what that answer is.

The realization that, in an information-rich space like the Web, we can do so many useful things by statistical means – without really understanding anything – has led to a sort of gold rush. It has been an exciting time. There was (and still is) much low-hanging fruit, ready for harvesting. This gold rush has produced a lot of valuable technology and it has created a few great fortunes. But it has not brought us much closer to understanding the central problem of extracting and truly understanding the deep meaning of an utterance. And without that, I think we have no hope of eliminating that last tenacious component of the error.

As difficult as the task may be, we will not achieve real natural-language understanding until we confront the problem of representing meaning (including all the background knowledge) head on. The representation may be in some kind of logic, or a semantic network, or perhaps something more neural and fuzzy and distributed. Probably it will be all three at once: these may be just different viewpoints – different ways of characterizing (or implementing) the same underlying concepts. But whatever form the knowledge takes, we can't get much farther without it.

We can't go on indefinitely ignoring the elephant in the room. His name is Clyde, and he's not going away.