Web chatbots
The next generation of speech systems?

Core speech recognition technology hasn’t changed much in the past 20 years; the advances have been in how speech is embedded in applications, rather than the core pattern-recognition algorithm. However, a step-change in this pattern-matching may be coming, leading to a new market for web-based chatbots who can chat naturally, without trying to sell you something, at least, not obviously. Over the past 20 years, speech technology has moved from research lab prototypes to viable commercial products, thanks to faster processors, better integration into applications, and simpler, more user-friendly training and adaptation. However, the underlying speech recognition algorithms have not changed much in the same period: The standard core speech recognition engine still uses a cascade of probabilistic Markov models. This sort of language model is best suited to applications where the system needs to understand exactly what the user says. For office dictation, command and control systems, telephone banking services, and so on, it is important to capture the user’s spoken language verbatim, to be passed onto to word processor, database or other back-end system.

However, in real life, people don’t usually use language so precisely. Much of our natural language use is in conversation: chatting without precise goals, where the hearer does not expect or need to capture and process our exact words and sentences. When having a friendly chat, or networking with current and prospective clients and contacts, we don’t have to understand exactly, to be understanding. Alternative language models are better suited to latching on to key cues and phrases, and filtering out the rest. Online chatbot engines such a Pandorabot.com use a different language model called AIML, which does not look for exact matches.

A Markov model can work out the probability of a whole word or sentence from its parts. But to do this, it needs to analyse all of the parts: it can’t process Eric Atwell is a genius without processing each of the four word-pairs making up this sentence.

About a decade ago, an American computer scientist called Wallace developed AIML, the artificial intelligence mark-up language. He used AIML to build language models for online chatbots: programs we can chat to informally, in a natural free-flowing conversation. Wallace realised that a language model for conversation analysis need not process all parts of each input sentence: it is better to focus on key cues and phrases.

Online chatbot engines such a Pandorabot.com use a different language model called AIML, which does not look for exact matches.

For example,
\[ p(\text{atwell}) = p(a) * p(t) * p(w) * p(e) * p(l) * p(l) \]

To get a better estimate, use the probability of each letter-PAIR, and multiply these:
\[ p(\text{atwell}) = p(at) * p(tw) * p(we) * p(el) * p(ll) \]

Speech recognition researchers realised that a Markov model works not just for letters, but also for phonemes (spoken letter-sounds), and for word-sequences, for example:
\[ p(\text{Eric Atwell is a genius}) = p(\text{Eric Atwell}) * p(\text{Atwell is}) * p(\text{is a}) * p(\text{a genius}) \]
which indicate the general direction of the conversation. An AIML model can use these key phrases in generating replies which go with the flow of the conversation, rather than treating user input as an information-request to be matched rigidly.

The AIML model is gaining popularity, not least because Wallace was able to demonstrate its merits by winning the Loebner prize international chatbot contest. Not once, but three times. An AIML-chatbot website, Pandorabots.com, hosts thousands of chatbots built by Wallace and his followers. The most popular Pandorabots for the last 24 hours web-page regularly lists chatbots developed by researchers and hobbyists, and also some commercial systems. For example, Cyber-Sandy and Nickie act as portals to adult-entertainment websites; Jenny introduces the English2Go website, and lets English language learners practice their chatting technique. The first Pandorabot chatbots were text-only: the user typed a sentence via keyboard, then the chatbot reply appeared onscreen as text too. Now, some Pandorabot chatbots incorporate speech synthesis; for example, Jenny talks with an educated British accent, via a speech synthesis engine. However, Pandorabot chatbots cannot recognise speech: the user still has to type their input via a keyboard. This is because existing Markov-model-based speech recognition is too cumbersome, and does not fit the AIML key-phrase model. Existing speech recognition systems would take a lot of time and memory trying to recognize everything in the input, even though little of this is subsequently needed by the AIML language model.

Researchers at Leeds university are developing techniques to automate the training of AIML language models from dialogue corpuses, large dataset of recorded conversation. A way forward is to apply the AIML model to speech recognition as well as conversation modeling. The next generation of Pandorabot chatbots can then include real-time speech recognition, allowing a fully natural spoken conversation. This will inevitably make AIML-based Pandorabots even easier to use, and hence more popular.

But will web chatbots make money? Certainly, at the moment most are built by academics and hobbyists. But the minority of chatbots fronting commercial websites, like Cyber-Sandy, Nickie, and Jenny, illustrate a market which could mushroom once chatbots can hear as well as speak. The talking chatbot can be a natural, friendly front-end to any website. Budgets currently set aside for layout and graphics on a commercial website homepage will be better spent on training the company chatbot portal, a friendly front-desk clerk who can chat to visitors. If you know what you want, of course you can still go direct to the page; but if you want to explore a company’s website without pressure or hard-sell, then the chatbot is a friendly companion.

BT adverts made good use of its catchphrase. It’s good to talk, to boost their telecoms sales. The mobile phone industry has also tapped into people’s need to chat: the telephone is not just for communicating information.

I predict that the next generation of speech systems could also focus on conversation, rather than dictation and command-and-control applications. Maybe it’s time your company website got a chatbot.

About the author
Eric Atwell is a Senior Lecturer in the School of Computing at Leeds University. The British Council commissioned his market review of “The Language Machine”, online: