Expanding the linguistic coverage of a spoken dialogue system by mining human-human dialogue for new sentences with familiar meanings

Gregory S. Aist (gaist@cs.rochester.edu)[1]
James Allen (james@cs.rochester.edu)[1, 2]
[1]Computer Science Department, University of Rochester
P.O. Box 270226, Rochester, NY 14627 USA

Lucian Galescu (lgalescu@ihmc.us)[2]
[2]Institute for Human and Machine Cognition
40 South Alcaniz Street, Pensacola, FL 32502 USA

How can a system grasp linguistic variety?

Computer systems that interact with people using natural spoken dialogue offer many possibilities for effective and efficient interaction. But, a dialogue system for a new domain requires a great deal of expert attention to either collecting data for a new domain or designing a model of the language that people use when solving problems in that domain. Consider the extension of a checkers-playing dialogue system to other games such as Go: game-specific phrases and terms would need to be added – both "official" and colloquial versions. Prior work here includes identifying new lexical items, new word-sequence correlations (Galescu, Ringger, & Allen 1998), or new phrase patterns using existing words or word classes such as *color* or *animal* (Bulyko, Ostendorf, & Stolcke 2003). In this paper, we take a slightly different approach by focusing on semantics. Given a specification (as a sample dialogue), we want to identify alternate ways of talking about the same things. These alternates may not necessarily use any of the words that their counterparts in the script use. In fact, the more dissimilar such new phrases are in terms of surface features, the more helpful they would be if added to the vocabulary and syntax that the system can understand. Such a technique should prove useful when expanding the linguistic coverage of a dialogue system to cover what people say in practice. We describe these techniques in the context of an equipment purchasing system – part of the multisite CALO project, an intelligent personal assistant.

Start with a script

Any dialogue system typically starts with some data collection or user interviews or some other technique designed to give a basic idea of what the users of the system will eventually say. The requirements for the system can then be easily expressed in terms of a dialogue script. The initial script for our domain of computer purchasing began:

System: Hello, this is CALO. User: Hello CALO.
S: Hello. U: I would like you to buy a computer for me.
S: Ok; what kind of computer would you like?

Add human-human dialogues

As part of the dialogue system effort, Rochester collected a set of approximately 40 human-human dialogues carried out by ~20 people playing the role of "buyer" and ~4 people playing the role of "agent". In order to collect dialogue aimed at the computer purchasing domain, we provided a short scenario to role-play when buying the first computer - "You are a small business owner looking to buy a computer", etc. The second scenario was self-directed, that is, "Now get a computer for yourself." These dialogues were recorded and transcribed. Those utterances that closely matched lines from the script we used directly as additional training data. Those utterances about capabilities beyond the scope of the initial system – such as warranty purchases – we reserved for later use. That left a wide range of utterances concerning concepts that are present in the script, but using different words and syntax. We wanted those.

Mine them for semantic matches to script lines

We used Latent Semantic Analysis (http://lsa.colorado.edu) to extract utterances from the human-human dialogues that were similar to each line in the script. For each utterance in the dialogues, we calculated its similarity to each line in the script, and assigned it to the line with the highest similarity. We then hand-filtered the resulting data to yield new ways of saying lines in the script, such as:

On a similar topic, but system initiative rather than user:
Script/User: I would like you to buy a computer for me.
Dialogue/*System: Hello would you like to buy a computer*

With similar meaning, but very different words and syntax:
Script/*S: Ok, I'll start looking.*
Dialogue/S: I'll go ahead and uh save this.

Less similar in meaning, but still reasonable alternatives.
For example, after "*Ok; what kind of computer would you like?"*:
Script/U: A lightweight laptop computer with 500 mb of ram.
Dialogue/U: Pentium processor.

From the utterances paired with the 28 lines in the script, 41 were identified to be useful matches. CALO developers at Rochester and IHMC used these to extend the language understanding of the CALO dialogue system.

References

Bulyko, I., Ostendorf, M., & Stolcke, A. (2003). Getting more mileage from Web text sources for conversational speech language modeling using class-dependent mixtures. *Proceedings of the North American Association for Computational Linguistics*. Edmonton, May 2003.

Galescu, L., Ringger, E., & Allen, J. (1998). Rapid language model development for new task domains. *Proceedings of the First International Conference on Language Resources and Evaluation*. Granada, Spain, May 1998.