

What Would a Human-Like Dialogue Agent Need to Know?

Lenhart K. Schubert

Department of Computer Science
University of Rochester
Rochester, New York 14627, USA

Abstract

Dialogue systems are still narrowly constrained in their ability to make inferences from utterances. These limitations stem in part from restricted expressivity of the representations used, lack of generality of the inference mechanisms, and inadequate quantities of linguistic and world knowledge. Recent work in Natural Logic (NLog) has highlighted the advantages of working with a representation close to language, for example enabling an immediate inference from a complex attitudinal sentence such as “I know that you won’t forget to give me a call” to “You will contact me”. The knowledge involved in such inferences is typically lexical, for example concerning the “factivity” of “knowing (that)”, the antifactivity of “forgetting (to)”, and the lexical entailment from “giving X a call” to “contacting X”. However, dialogue inferences (and discourse inferences more generally) also require the use of world knowledge. For example, the utterance of “I am happy to tell you that you have won a million dollars” not only commits the speaker to the proposition that the addressee has won a million dollars (an NLog inference), but also allows the inference that the addressee will soon receive a million dollars, will as a result be significantly wealthier, and is very likely delighted at this prospect. Such inferences will strongly affect the further course of the dialogue, but they could not be obtained by NLog methods. They can, however, be obtained by methods such as are employed in our EPILOG inference engine at the University of Rochester, the latest version of which was built by Fabrizio Morbini. We are currently able to show examples of such inferences, including ones involving metareasoning over classes of verbs or other syntactic entities, but not yet on a large scale. Thus we are working on various methods of knowledge accumulation, and the talk will include a brief progress report on extraction of many millions of items of general knowledge from text using our KNEXT system and other methods; and lexical knowledge engineering, with some semiautomatic help, building to some extent on top of WordNet, VerbNet and other resources.