PRESUPPOSITION-BASED AI NATURAL LANGUAGE SYSTEM

Victor Raskin, Hebrew University
Jerusalem, ISRAEL

The first Israeli AI natural language group is developing a system in which the purpose of the semantic analysis of a sentence is to explicate both its own meaning and the whole framework of relevant knowledge associated with it. The latter is covered by the set of presuppositions of the sentence. The emphasis in the research is on justification of the semantic instruments used in the system rather than on originality of its language or notation.

The system is based on a quantitatively unrestricted set of final love dialogues from the Mills and Boon series of romance novels, which, while varying a great deal in details, are all very similar schematically. (In a typical dialogue the two protagonists, the male and the female, discuss the history of their love relationship from the point they stopped being indifferent to each other and on to the present moment when they love each other passionately and are going to be married.) The material is thus thematically restricted but the system is designed to process any dialogue of this type, even it has to ignore meaningfully certain details it cannot and should not handle. At the present stage the lexicon includes over 500 lexical entries represented in tree-form frames and based on a dozen of primitives such as want, together, realize, carry out, can, and the standard variables for the male protagonist, the female protagonist and the object. The data base also contains a number of laws and regularities in the "world" of the dialogues such as that consummation (carrying out) of love means marriage since the dialogues exclude pre- or extramarital sex. All such laws are written in the terms of the lexical entries and, together with them, constitute the two sources for the calculation of presupposition, which is already feasible in the system.

The presupposition orientation is explained both by the fact that any AI system should "understand" more than the meaning proper of the sentence and, more importantly, by the fact that presupposition provides the system with a criterion of justification for the proposed format of the lexical entries and the primitives constituting them.

A related series of psycholinguistic experiments demonstrates that the semantic ability of the native speaker includes presupposition-awareness as its aspect in the sense that, when exposed to a sentence and asked a question of the type, "What else should you know in order to understand the sentence properly?", the collected answers tend to be reliably uniform. Presupposition is broadly thought of here in terms of enablement.

At the initial stage the system operates in cycles, the general goal being the perfection of the lexical entries. Postulated initially a priori, the primitives and the format of the lexical entries lead to a certain set of presuppositions associated with every sentence. The latter is selectively compared with the human data, then automatically corrected and fed back to the entries, thus modifying them accordingly.

At subsequent stages, the system, semantically tuned up in this way, is put to various uses such as participating in a dialogue, summarizing, lying and making jokes.

It is claimed that the fact of semantic justification of the instruments used in the system makes them extrapolable to other linguistic material, and possibilities of such extrapolation are briefly outlined.

The system is implemented in a LISP-based and ACTOR-semantics oriented language by Mr. Boris Klebansky.