A Domain Independent Semantic Parser For Companion

Kathleen McCoy, Patrick Demasco, Mark Jones, Christopher Pennington & Charles Rowe

Applied Science and Engineering Laboratories
University of Delaware/Alfred I. duPont Institute
Wilmington, Delaware USA

Introduction

This work is part of an augmentative communication project being conducted at the Applied Science and Engineering Laboratories at the University of Delaware and the A.I. duPont Institute. The goal of this project is to increase the communication rate of physically disabled individuals via Natural Language Processing techniques.

We wish to take as input a compressed message (i.e., one containing mainly the content words of the desired utterance) from the disabled individual and yet pass a syntactically and semantically well-formed utterance to a speech synthesizer or text preparation system. At the same time, we wish to do this by placing as little a burden on the user as possible (Demasco et al., 89). Thus, we are not interested in a simple coding system (cf. (Baker, 82)) where sentences have been stored and are simply indexed by their content words.

The system is broken into several phases. The first phase, the semantic parser (see also (Small&Rieger, 82)) is responsible for determining the semantic role being played by each input word. It must determine which word is the verb, what role each noun phrase plays with respect to the verb (e.g., actor, theme), and what modification relationships are present. The resulting semantic representation is then passed to the translation component which is responsible for replacing the semantic terms with their language-specific instantiations. The final phase of the processing is a sentence generator which forms a syntactically correct sentence. This paper is concerned primarily with the semantic parser.

Knowledge and Processing

In (McCoy et al., 89) we reported on a semantic parser which relied primarily upon the domain of discourse. As the lexicon grew, this solution seemed unsatisfying because it was difficult to constrain the domain of discourse, and the knowledge management for an arbitrary number of possible domains proved unwieldy. A new approach is being developed which identifies the roles that words can play in a domain independent manner.

The problem faced by the parser is the ambiguity resulting from the compressed input because it contains no syntactic information. Consider the following input: "John study red house". The first problem faced by the system is determining the function of each word in the sentence. Each individual word can have different semantic classifications and thus its function in the sentence may be ambiguous. For example, the word "study" has two meanings: an action, as in "John studies", or a location, as in "John is in his study". In order to recognize all possible word meanings (and to constrain further processing) we classified all words into five categories according to the different semantic functions that the word can play: Actions (verbs), Objects, Descriptive Lexicons (adjectives), Modifiers to the Descriptive Lexicon (adverbs), and Prepositions. Within each category, we further classified the words so that finer semantic distinctions can be made.

If each individual word is looked at in isolation, it (usually) has many interpretations. However, if the input words are taken together, then many of the interpretations can be eliminated. In the above example, while "study" may either be an object or a verb, no other word in the sentence occurs in the verbs hierarchy. Thus "study" will be taken as the verb.

Once the verb of the sentence has been determined[1], additional knowledge sources can be applied to further reduce ambiguity. The main verb predicts much of the structure of the sentence. It dictates, in a top-down fashion, which semantic roles are mandatory and which roles should never appear, as well as type information concerning possible fillers of each role. For example, the verb "go" cannot have a THEME case in the semantic structure. Furthermore, it cannot have a FROM-LOC case unless it also has a TO-LOC. This information is captured with frame predictions which are associated with each verb in the verbs knowledge base.

Based on the particular verb, a set of semantic structures (called frames) is created. Each frame encodes one possible interpretation of the current sentence and contains typed variables which may be filled with words from the input. Each variable type corresponds

1. In the case that the system identifies more than one word that can fill the role of the main verb, it will spawn an independent process for each verb candidate.

to a type of possible input word (i.e., they are types taken from one of the system knowledge bases).

For example, the frame associated with the word "study" indicates that the AGENT must be human and that it can take a THEME (some abstract or physical object) and a LOCATION (a physical place). The AGENT is required while THEME and LOCATION are optional.

Once a frame has been chosen, bottom-up processing attempts to fill out the frame by fitting the input words into the appropriate slots. A frame is considered satisfactory if all of its obligatory roles can be filled.

A final source of ambiguity must be dealt with. So far we have considered words of the input modifying the verb by planning a role with respect to it. We must also consider the possibility of one word of the input serving to modify the meaning of another input word. Because a certain word can only modify certain types of other words, we attach words which can be modifiers onto the words of types that they can possibly modify. For example, the word "red" is attached to the type PHY-OBJ (PHYsical-OBJect) in the object knowledge base. Therefore, the word "red" can modify the word "house" since "house" is a PHY-OBJ, but cannot modify the word "idea" since "idea" is not. Using the modifier information the system attempts to associate any modifier words with the words of the input that they modify. In this way, the "red" from the input will be attached as a modifier to the word "house".

Using the above processing and knowledge sources the parser will come up with two possible interpretations of the input: "John studies at the red house" (here the red house is taken as a location) and "John studies the red house" (here the red house is taken as the thing that John is studying). Notice both interpretations account for all input words and fill all obligatory roles associated with the chosen verb.

Current Functionality

The current system is able to recognize different uses of a verb depending on the modifiers present in the input. Consider "John look Boston" and "John look tired". Respectively, the final outputs are "John looks at Boston" and "John looks tired". Notice that "look" can be used in two very different ways. The system determines which use is intended by considering the rest of the input. If there is an OBJect or LOCation to look at, then the first use is recognized. If this is not the case, but a modifier that may modify the agent occurs in the input, then the second use is recognized.

As an example of how the entire input is used to constrain possible interpretations, consider "John study weather university". Notice that in isolation "university" could either be a LOCATION or a thing being studied (a THEME). However, since "weather" can only play the THEME role with respect to "study" the system correctly generates: "John studies weather at the university".

In addition, the present system has the capacity to infer the verb intended by the user. The present system chooses between have and be as verb candidates depending on the given input elements. Given the input "John tired", the final product is "John is tired". Given the input "John paper", the final output is "John has the paper".

This system will also infer the actor (subject) of the intended sentence. In particular, if no agent is given, the system will infer the user to be the agent (and thus generate a first person pronoun for that slot). Given the single-word input of "hungry," the system will determine the intended verb and the intended agent, and generate the sentence "I am hungry".

Conclusion

A prototype implementation of the system has been completed and is currently being evaluated. The lexicon currently contains several hundred words of various types. The implementation is in Common Lisp and runs on several different hardware platforms. The evaluation currently includes a paper experiment to determine whether or not the current system mirrors the action of a human. Future plans for evaluation include testing the prototype system with potential users.

Acknowledgments

This work is supported by Grant Number H133E80015 from the National Institute on Disability and Rehabilitation Research. Additional support has been provided by the Nemours Foundation.

References

B. Baker. Minspeak. Byte, 186ff, September 1982.

P. Demasco, K. McCoy, Y. Gong, C. Pennington, and C. Rowe. Towards More Intelligent AAC Interfaces: The Use of Natural Language Processing. In Proceedings of the 12th Annual conference, pages 141-142, RESNA, New Orleans, Louisiana, June 1989.

K. McCoy, P. Demasco, Y. Gong, C. Pennington, and C. Rowe. A semantic parser for understanding ill-formed input. In Proceedings of the 12th Annual conference, pages 145-146, RESNA, New Orleans, Louisiana, June 1989.

A Domain Independent Semantic Parser

S. Small and C. Rieger. Parsing and comprehending with word experts (a theory and its realization). In Wendy G. Lehnert and Martin H. Ringle, Editors, Strategies for Natural Language Processing, 1982.

Contact

Kathleen McCoy
Dept. of Computer and Information Sciences
University of Delaware
Newark, DE. 19716
Email: mccoy@udel.edu