

Natural Language Systems: How are they meeting human needs?

Kathleen R. McKeown

Department of Computer Science
Columbia University
New York, N.Y. 10027

This paper appeared in *Proceedings of the Annual Conference - ACM '83*,
New York, N.Y., October 1983.

One goal of natural language research is to make systems more accessible to their users by allowing them to interact with machines in a language as close as possible to the language people use among themselves. Developing systems that answer natural language questions is one way to allow users to interact easily with computers. In the following sections, I briefly describe ways in which researchers have been successful in this attempt, limitations inherent in existing natural language systems, and current efforts to bring systems closer to meeting the needs of their users. To illustrate these issues, I focus on *natural language database systems* as one example of question-answering systems.

1. What Systems Can Do

Natural language database systems (NLDB) typify the success researchers have had in making systems accessible to users. These systems allow a user to ask questions in natural language about information stored in large databases instead of using formal query languages to retrieve information. Most systems support a reasonable range of questions and allow language features such as pronominalization¹ or ellipsis² so that a user can state a request concisely. Examples (1)-(3) are questions handled by LADDER [HENDRIX 78], an NLDB that answers questions about naval ships (1) is a typical question handled by the system, (2) shows the use of a pronoun (*it*) to refer to the ship mentioned in the previous question, and (3) illustrates the use of ellipsis to easily ask the question *What is its speed?* If the natural language interface were not available, the user would have been required to use the formal query shown in (4) to ask question (1).

1. What is the length of the Kennedy?
2. What is its home port?
3. Speed?
4. (ida '((nam eq John#F.Kennedy) (? length)))

What does it mean to *understand* and *respond* to questions in NLDB systems? An NLDB usually consists of a natural language interface and an existing database

¹The use of pronouns, such as *it*, *he*, *she*

²deleting repeated phrases in a follow-up question

with its database management system and formal query language. To answer a natural language question, the question is translated to the formal query language. This is done by *parsing* the question, using a grammar, into a canonical form in which the sentence constituents and relationships between them are identified. The canonical form can then be easily translated into the systems's formal query language. The resulting query is executed, causing a search of the database for the specified objects. The results of the search are formatted in an English sentence or a table.

2. What Systems Can't Do

Natural language question-answering systems are typically limited to handling questions about a narrow domain of discourse. For NLDB systems, this is the domain covered by the database. That is, they can only answer questions about knowledge contained in the database and that use vocabulary that pertains to the current database contents. While extending the system's knowledge and vocabulary is not usually a theoretically difficult task, it does require implementation effort. Some recent work [MARTIN 83] aims at automating the acquisition of this type of knowledge, thus making it easier to transport the natural language interface to a new database domain. Systems are also limited in English coverage to what their grammars have been designed to handle, although some systems (e.g. [HENDRIX 78]) allow their users to extend the system's vocabulary and coverage themselves.

But systems fail their users in more significant ways that cannot be corrected through implementation effort alone and it is these issues that will be emphasized in the talk. The use of natural language in a system implies that it is capable of the full range of interaction that conversation participants make use of, but a system which does no more than translate a question to a formal query is actually more restricted than it seems. Conversation participants take advantage of the flexibility that natural language provides to tailor their responses to different questioners, to provide detailed descriptions if necessary, and to determine the speaker's intended meaning and beliefs behind a question. Incorporating such capabilities in systems requires additional work in both the interpretation of questions and the generation

of responses. Some examples of recent and ongoing research which attempts to provide a fuller range of interaction are outlined below.

3. Meeting Users' Needs

- It has been shown [MALHOTRA 75] that users need to ask questions to familiarize themselves with the database before asking specific questions about its contents (e.g. "What data do you have?", "What is manufacturing cost?"). Natural language is particularly appropriate for answering such questions as they require definitions, descriptions, and longer textual sequences. To generate these kinds of responses, a formulation of strategies that can be used to organize and determine content of the response is required [MCKEOWN 82].
- Frequently, users reveal in their questions a presumption about the database which turns out to be incorrect. The encoding of presumptions in utterances is a formal feature of natural language which can be exploited to detect and correct misconceptions that the user has [KAPLAN 79; MAYS 78; MCCOY 83]. If such presumptions are not corrected, the user may be left with false beliefs about the database even if all his/her questions have been answered correctly.
- Flexibility in natural language means that the same question can be responded to differently in different discourse situations. A number of efforts in *user-modelling* maintain information about the user's beliefs, plans, and goals in order to provide more information than was literally requested in a question, but which any cooperative conversational participant would recognize was actually expected in a response (e.g. [ALLEN 80]). Developing the ability to tailor a response to a given user by varying level of detail, point of view, or textual strategy is another aspect of user-modelling [MCKEOWN 83; LEBOWITZ 83].

4. Summary

NLDB systems meet the needs of certain classes of users simply by allowing them to use a mode of interaction with the database that they are already familiar with. While these systems are not as flexible as human conversation participants, a number of ongoing research efforts are currently engaged in bringing them closer to meeting the needs of their users. Examples of these efforts will be presented in more detail in the talk.

5. References

- [ALLEN 80]. Allen, J.F. and Perrault, C.R. "Analyzing intention in utterances," *Artificial Intelligence* 15, 3, 1980. pp. 143-78.
- [HENDRIX 78]. Hendrix, G.G. "Developing a natural language interface to complex data," *ACM Transactions on Database Systems*, Vol. 3, No. 2, June 1978. pp. 105-47.
- [KAPLAN 79]. Kaplan, S.J., Cooperative responses from a portable natural language database query system. Ph.D. dissertaion, Univ. of Pennsylvania, Philadelphia, Pa 1979
- [LEBOWITZ 83]. Lebowitz, M., "RESEARCHER: An overview," In *Proceedings of the 3rd National Conference on Artificial Intelligence*, Washington D.C., 1983.
- [MALHOTRA 75]. Malhotra, A., Design criteria for a knowledge-based English language system for management: an experimental analysis. MAC TR-146, MIT, Cambridge, Mass., 1975.
- [MARTIN 83]. Martin, P. Appelt, D., and Pereira, F., "Transportability and generality in a natural-language interface system," in *Proceedings of the 8th International Joint Conference on Artificial Intelligence*, Karlsruhe, Germany, 1983.
- [MAYS 80]. Mays, E. "Correcting misconceptions about data base structure," In *Proceedings 3rd CSCSI Biennial Meeting*, Victoria, B.C., 1980.
- [MCCOY 83]. McCoy, K. "Correcting misconceptions: What to say when the user is mistaken," In *Proceedings to CHI 83*, Cambridge, Mass. 1983.
- [MCKEOWN 82]. McKeown, K.R., Generating natural language text in response to questions about database structure, Ph.D. dissertation, Univ. of Pennsylvania, Philadelphia, 1982
- [MCKEOWN 83]. McKeown, K.R., "User-oriented explanation for expert systems," Technical Report, Columbia Univ., New York, 1983