THE IMPACT OF NATURAL LANGUAGE ACCESS ON DATABASE DESIGN AND IMPLEMENTATION

CHAIRMAN:
Dr. S. Jerrold Kaplan
Teknowledge Inc
Palo Alto, CA

PARTIAL LIST OF PARTICIPANTS:

Mr. Jeffrey Hill
Artificial Intelligence Corp.
Waltham, MA

Mr. Francisco Corella
Symantec
Sunnyvale, CA

Dr. Jim Davidson
Stanford University
Stanford, CA

Dr. Steven Shwartz
Cognitive Systems, Inc.
New Haven, CT

Prof. Robert Wilensky
University of California, Berkeley
Berkeley, CA

The central question to be addressed by this panel is how the provision of natural language access influences the design and implementation of database systems. Within this, there are (at least) the following four issues to discuss:

1. The augmentation of data models to support the specific requirements of natural language access.

This is perhaps the broadest category. Among the particular demands of natural language access are: treatment of lexicon and lexical ambiguities; use of parenthetical expressions which may require simultaneous update and query; support for the generation of natural language responses; storage of discourse information; representation of text structure models; accurate processing of natural language updates; and other interactions that, though not unique to natural language access, have special problems in that context, such as natural language queries to dynamic databases requiring natural language statements of alerting conditions, or answering "meta-questions" - questions about structure. Current work in all these areas suggests the need to augment data models in particular ways.

2. Ideas for data models that arise from the study of natural language semantics.

This category is related to the first, but the emphasis here is that researchers in natural language are developing ideas about modelling that can be very helpful in database modelling. One example of this is the uniform use of first-order logic as the "logical form" of natural language sentences as the representation of knowledge and data, and as the medium for deductive retrieval. Another example is the development of concepts of intensionality that support the notion of an "historical" database system.

3. The organization of information in database systems to enhance the portability of natural language interfaces.

Many researchers are investigating how to reduce the difficulty of moving a natural language interface from one database system to another. The problems in doing this include what information is needed, how the information should be divided into modules, and what algorithms are needed to acquire the necessary information for a new database. There are obvious ramifications on database structure for achieving various levels of natural language interface portability.

4. The requirements imposed on conventional database systems that function as the back end of natural language processing systems.

Looking at natural language research itself, most researchers develop their own information storage and retrieval mechanisms specifically tailored to the kinds of data structures that arise in natural language processing. However, there is some recent work in investigating how conventional database systems could be used in this role. This work is bound to have important effects in the future design of databases that deal in less regular structures than we are used to.

ACKNOWLEDGMENT:

The concept for this panel and the list of issues was developed originally by Dr. Jonathan King of Hewlett-Packard Laboratories, Palo Alto, California.