APPROXIMATE RESPONSES FROM A DATA BASE QUERY SYSTEM: AN APPLICATION OF INFERENCING IN NATURAL, LANGUAGE

Aravind K. Joshi and S. Jerrold Kaplan
University of Pennsylvania
Department of Computer and Information Science
The Moore School of Electrical Engineering
Philadelphia, Pennsylvania 19104

Ronald M. Lee
Department of Decision Sciences
The Wharton School
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Our' goal is the development and application of various techniques for generating approximate responses to data base queries. An "approximate response" is a response other than a direct answer to the question. Approximate responses are frequently referred to by linguists as "indirect answers" or "replies" (e.g. in BS76). What is approximate is not so much the response as the relationship between the response and the initial query. Our approach is to regard an interaction between a user and a data base as a discourse, having the properties and constraints normally associated with human dialog. (Conversational Postulates of Grice (G67) are examples of such constraints.) Many of the conventions of human dialog can be implemented through approximate responses which, for instance, 1) aid a user in formulating a suitable alternative query when the precise response to the initial query would be uninteresting or useless; 2) inform a user about the structure or content of the data base when the user is unfamiliar with its complexities; and 3) summarize at an appropriate level, eliminating unnecessary detail.

Natural language (NL) query systems *are* of benefit to users who are only partially familiar with the structure and/or content of the underlying data base. Such "naive" users are typically hampered by their lack of knowledge in formulating a query which will retrieve the. desired information. We believe that NL can do more than simply provide the user with a convenient, higher-level replacement for a formalized query syntax. NL questions frequently embed information about the user's understanding of the structure of the data. This information can be exploited to inform and guide the user in the use of the data base.

Of particular interest to us is the key role that shared knowledge between conversants plays in the effectiveness of human dialog. As observed in (CH75), dialog tends to proceed with statements which offer a specific piece of 'new' information to the conversation which is differentiated from information considered as 'given' or already known to the other party. Breaches of this 'Given-New Contract' can point to the need for additional background information to be supplied in order for communication to be effective. We believe that this observation can be effectively utilized within the context of queries to a data base system. Our approach here is to pay special attention to the 'given' information contained in the user's questions in the form of presuppositions. If these turn out to be false, we interpret this as a signal that the user misunderstands some aspect of the data base's structure or content and is in need of additional clarification. An approximate response explicitly contradicting the failed presupposition and perhaps suggesting an alternative is appropriate, as it is in human dialog. Such a response serves to correct the users mis impressions and provide suggestions for alternatives, hopefully relevant and useful ones.

A presupposition of a sentence S can be broadly defined as any assertion that must be true in order for S to be meaningful. In the case of questions, the presupposition must be true for a direct answer to be meaningful.

Presuppositions come in many forms. There are presuppositions which are primarily syntactic (JW77). Others deal with implied restrictions on the size, or a claim about the completeness of the answer set (BS76). Of particular interest in a data base context *are* those presuppositions of an NL question which are implied by a corresponding formal query to a given data-base structure. We have observed that each stage in the execution of a formal query, except for the final one, has an interpretation as a presupposition of the NL question. If a particular stage of execution returns a null set, the corresponding presupposition has failed and can be explicitly contradicted, rather than returning an obviously uninformative or misleading null response.

Consider the query
"WHICH LINGUISTICS MAJORS GOT A GRADE OF B OR BETTER IN CS500?"
Assuming a suitable structure for the data (see Figure one), a corresponding formal query might perform the following operations: 1) Find the set of students and restrict it to linguistics majors; 2) Find the set of courses and restrict it to CS500; 3) Find the class list (set of students) associated with the result of 2; M) Restrict the class list of 3 to those with grades ^ B; and 5) Intersect 4 with 1 to produce the response. An empty set at each stage could be used to produce the following approximate responses contradicting the failed presuppositions: 1) There are no linguistics majors; 2) There is no course "CS500"; 3) No students were enrolled in CS500; and 4) No students received a grade of B or better in CS500. A failure in the final stage leads to the direct answer NONE. It is worth noting that different data structures will reveal different presuppositions. For instance, a different data base might produce the response "No linguistics majors took CS500."

* This work is partially supported by NSr Grant MCS 76-19466.

We wish to thank Peter Buneman, Rob Gerritsen, and Ivan Sag for many fruitful discussions.

Another type of approximate response deals with the generation of a response to a substitute query. For instance, "Is Venus the fourth planet?" may be responded to by "No, it is the second planet."(see (L77) for similar examples). A determination of the focus and topic of the question can be used to generate an appropriate alternative, as opposed to (say) "No, Mars is the fourth planet." Syntactic and contextual cues are under investigation to determine the topic and focus in the face of partial information. Careful construction of the formal query can provide a relevant piece of alternative information forfree by selecting the most appropriate access path to the desired information.

An important convention of human conversation is that no participant monopolize the discourse, so that control can be shared. One implication of this is that all responses given in a conversational mode must be short. Thus where the system would otherwise respond with a lengthy list, we would prefer to be able to return a non-enumerative, or "intensional" response". Lengthy response sets could be summarized, or defined by a characteristic or attribute. For instance, the question "Which employees engage in profit sharing?" may be answered by listing the extension of a set containing (perhaps) 10,000 names, or by the intensional response "All vice-presidents." The summary might be computed from the data or infered from the data base schema, and can be used to avoid unnecessary and distracting detail. In these cases, the response may implicitly incorporate the restrictions of the question. For instance, a response to "Which students were invited to the party?" of "The girls living in West Philadelphia." clearly implies that only those girls who are students were invited (KH 73).

Conversations also allow hypothetical questions, or questions about the structure of the world (in our case, the data base). Questions such as "Can supervisors profit share?" may be answered affirmatively by the contents of the data base (finding an instance), or negatively by noting that the data base structure precludes such a possibility. If neither of these alternatives are successful, an approximate response of "maybe", or "I don't know" may be returned, (since many constraints to the data base may be imposed by the logic of the updating programs or organizational procedures).

Finally, conversations admit answers of a statistically approximate nature. "What percentage of welfare recipients are single mothers?" may be sufficiently answered by "About 80%". This concept is of use in the execution of queries on very large data bases, when precise responses are both unnecessary and expensive. If the user is willing to accept an approximate response which is within a given confidence level, this can frequently be computed for a fraction of the cost of a complete one.

Existing data base systems could be described as "stonewalling", giving only limited, precise answers, which inhibited browsing and query formulation, Approximate responses, as they are used in human dialog, can significantly increase the usefulness and convenience of data base query systems.

References:

(BS76) Belnap, N.D. and Steel, T.B., The Logic of Question and Answers, Yale University of Press, New Haven, 1976.

(CH75) Clark, H.H. and Maviland, S.E., "Comprehension and the Given-New Contract," in Discourse Production and Comprehension (ed. R. Freedle), Lawrence Erlbaum, Hillside, N.J., 1975.

(G67) Grice, H.P., "The Logic and Conversations", in The Logic of Grammar (eds. D. Davidson and G. Harman), Dickinson, Encino, Callf., 1975.

(KH73) Keenan, E.L. and Hull, R.D., "The logical presuppositions of questions and answers," in Prasuppositionen in Philosophic Und Linguistik"(eds. J.S. Petofi and D. Franck), Athenaum Verlag, Franfurt, 1973.

(JW77) Joshi, A.K. and Weischedel, R., "Computation oi a subclass of inferences: presupposition and entailment," American Journal of Computational Linguistics, January 1977.

(L77) Lehnert, W., "Human .and computational question answering", Cognitive Science, vol. .1, no. 1, 1977.

In the relational formalism:

STUDENTS(STUDENTS,MAJ0R)
OFFERINGS*(COURSES,SEQUENCER)
ENROLLMENTS (SEQUENCED , STUDENT# ,GRADE)

Figure 1

* NOTE: SEQUENCE# uniquely identifies an offering of a course.