TOWARDS A COMPUTATIONAL MODEL FOR THE SEMANTICS OF WHY-QUESTIONS

W. Wahlster

Germanisches Seminar Universitaet Hamburg Von-Melle-Park 6 D-2000 Hamburg 13 Federal Republic of Germany

Summary. This paper discusses aspects of a computational model for the semantics of why-questions which are relevant to the implementation of an explanation component in a natural language dialogue system. After a brief survey of all of the explanation components which have been implemented to date, some of the distinguishing features of the explanation component designed and implemented by the author are listed. In the first part of the paper the major types of signals which, like the word why, can be used to set the explanation component into action are listed, and some ways of recognizing them automatically are considered. In addition to these linguistic signals, communicative and cognitive conditions which can have the same effect are discussed. In the second part the various schemata for argumentative dialoque sequences which can be handled by the explanation component in question are examined. Particular attention is paid to problems arising in connection with the iteration of why-questions and the verbalization of multiple justifications. Finally schemata for metacommunicative why-questions and for why-questions asked by the user are investigated.

Introduction

The explanation component of a natural language AI system is that component whose job it is to generate, in response to a why-question an explanation which is both understandable to the user and appropriate to the current state of the dialogue.

Although there has been relatively little research into the semantics and pragmatics of whyquestions^{1,5,9,17} and the cognitive processes underlying the answering of them, several Al systems do exist which are capable of handling certain types of why-questions. The practical value of the incorporation of an explanation component lies essentially in the fact that, as Stallman and Sussman have put it, "such programs are more convincing when right and easier to debug when wrong".15

Figure 1 provides an overview and comparison of the explanation components which have been implemented to date: BLAH²², DIGITALIS ADVISOR¹⁶, EL¹⁵, EXPOUND², HAM-RPM^{19,21}, LUIGI¹³, MYCIN^{12,14}, NOAH¹¹, PROSPECTOR⁷, SHRDLU²⁴, TKP2¹⁰ (The symbol ''-'' signifies that the attribute in question is not applicable to the given system). This paper presents some results of my experience in designing and implementing an explanation component²¹; together, they represent a step toward a computational model for the semantics of whyquestions. The explanation component was designed as a module which could in principle be incorporated into any natural language AI system. It has been tested within the natural language dialogue system HAM-RPM⁶, which converses with a human partner in colloquial German about limited but interchangeable scenes.

In implementing HAM-RPM we have taken into account the human ability to deduce useful information even in the case of fuzzy knowledge by approximate reasoning. The model of fuzzy reasoning used in HAM-RPM can be characterized by the following four properties²⁰:

- (a) A fuzzy inference rule represents a weak implication; a particular 'implication strength' must thus be associated with each such rule.
- (b) The premises of a fuzzy inference rule are often fulfilled only to a certain degree.
- (c) The applicability of a fuzzy inference rule in the derivation of a particular conclusion is likewise a matter of degree.
- (d) Several mutually independent fuzzy inference rules can corroborate each other in the derivation of a particular conclusion.

The explanation component which I have developed differs from BLAH²², one of the most advanced explanation components which have similar goals, in that on the one hand fuzzy inference rules and facts can be modified by appropriate hedges (in accordance with (a) through (c) above), and on the other hand the system is able in the course of a dialogue to generate multiple justifications for an explanandum (in accordance with (d) above). A further important difference between this explanation component and the other systems included in Figure 1 is that the system is equipped with a fairly sophisticated natural language generator, which is ATN-based and includes algorithms for generating pronouns and definite descriptions¹⁹.

Only two aspects of this explanation component will be discussed in this paper: The signals on the basis of which the explanation component generates an argumentative answer to a question asked by the user and the speech act schemata for the argumentative dialogue sequences which can be

SYSTEM'S GENERAL CHARACTERISTICS					EXPLANATION COMPONENT						
					LINGUISTIC CAPABILITIES			COMMUNICATIVE CAPABILITIES		COGNITIVE CAPABILITIES	
Name of system	lmplementation language	Domain ofdiscourse/ application	Formalism for inference rules	Natural language input of why-questions?	Basis for gener- ation of natural language explana- tions	Explanations in technical (TECH)/ colloquial (COLL) language	Selection based on user model (MDD)/ dialogue strate- gies (D1A)	Variable level of detail (DET)/ trans- parent structuring (STR)	Explanation of fuzy (FU2)/ hypothetical (HYP) reasoning	Explananda	
BLAH	AMORD	U.S. in- come tax laws	AMORD rules	No	schemata	COLL	MOD, DIA	STR, DET	нүр	assertions, suggested al- ternatives, decisions	
DIGITALIS ADVISOR	OWL	medicine: digitalis therapy	0WL procedures	No	schemata, canned text	тесн	DIA	DET	-	system's questions, reasoning chain	
EL	ARS	electrica circuit analysis	ARS rules	No	-	-	-	-	нүр	system's conclusions	
EXPOUND	LISP	logic: formal proofs	predicate calculus formulas	No	simple case grammar	TECH	-	STR	-	theorems	
HAM~RPM	FUZZY	traffic scene, room- booking	DEDUCE procedures	Yes	ATN-based generator, schemata	COLL	MOD, DIA	DET, STR	FUZ	system's ques- tions and con- clusions, rea- soning, incl. multiple der- ivations	
LUIGI	SOL	kitchen world	SOL procedures	Yes	specific generation procedures	COLL	-	-	-	simulated actions	
MYCIN	LISP	medicine: bacterial infections	production rules	Yes	schemata	TECH	DIA	DET	FUZ	system's ques- tions and con- clusions, rea- soning chain, meta-infer- ences	
NOAH	SOUP	repair of electro- mechanical equipment	SOUP procedures	No	-	-	-	-	-	system's instructions to user	
PROSPECTOR	LISP	geology: mineral explo- ration	rules in inference net	No	schemata	TECH	-	-	FUZ	system's questions	
SHRDLU	MICRO- PLANNER	blocks world	consequent theorems	Yes	specific generation procedures	COLL	-	-	-	system's simulated actions	
ТКР2	LISP	logic: formal proofs	predicate calculus formulas	No	schemata	TECH	-	STR	-	theorems	

Figure 1: Comparison of all explanation components implemented to date

realized in the system.

A Formal Description of the Signals Suggesting an Argumentative Answer

The purpose of the present section is to list the major types of signals which are capable of setting an explanation component into action. The resulting classification of linguistic expressions does not, of course, imply that all of the expressions in a given category are completely synonymous.

Signals for Argumentative Answers in the User's Utterances

From the point of view of algorithmic recognition, the simplest case is that in which the user elicits an argumentative answer from the system by asking a direct question. The word *why* can often be interpreted as a signal for an argumentative answer. On the other hand, its exact meaning depends on the dialogue context and it can be used within speech acts which have nothing to do with explanation, such as making a suggestion or a comment⁵. In spite of its ambiguity, the word *why* represents the only means of eliciting an argumentative answer in most AI systems which have an explanation component.

Special idiomatic expressions such as those listed in (L1) can have the same function as the word why. In the system HAM-RPM expressions like

(L1) How come, what ... for, how do you know

these are recognized through pattern matching during lexical analysis⁶.

Indirect questions such as those in (L2) require that the system be able to divide the utterance into matrix sentence and embedded sentence syntactically; only then can it process the latter using the same means as in the case of direct questions containing why or the questions in (L1).

(L2) Please tell me why A, I'd like to know why A

Further types of signals include direct (see L3) and indirect (see L4) requests. The problem of

(L3) Please explain why A, prove that A

(L4) I'd be interested in hearing why you think that A, Are you prepared to justify your conclusion that A?

how indirect speech acts such as the requests in (L4) can be recognized automatically is one which has recently been attracting much attention from natural language Al researchers³¹⁸.

The word why and the expressions in (L1) needn't accompany the proposition to be explained within a single utterance, as they do in the example (E1); they can also be used alone after the system has answered a question to elicit an explanation of the answer (cf. E2).

(E1) USER (U): Why is Glenbrook Drive closed?

(E2.1) USER (U): Is Glenbrook Drive closed?

(E2.2) SYSTEM (S): Yes.

(E2.3) USER (U): How do you explain that?

The expressions in (L3) and (L4) can also be used to achieve just the opposite: An argumentative answer is requested in advance, before the corresponding question has been asked of the system.

(E3) Please explain your answer: Do you think that A?

As the continuation of (E2.1) and (E2.2) represented by (E2.4) and (E2.5) illustrates, a speaker often explains a previously given answer when the listener - perhaps using an expression such as the ones in (L5) - shows signs of doubt as to (L5) Really? Are you sure? That's strange.

(E2.4) U: Really?

(E2.5) S: Yeah, they're repaving it.

the truth of the answer.

A kind of signal which suggests an argumentative answer in a still more obvious manner is the category of utterances by the user which indicate an opinion contrary to that expressed by the system (cf. L6). The idiomatic expressions in (L5)

(L6) I doubt that, That doesn't follow, I can't believe that..., Since when?

and (L6) which always express doubt or a contrary opinion no matter what the current dialogue context may be, can be handled adequately if information concerning their implications is stored in the system's 'idiom lexicon'⁶.

A further way in which the user can indirectly ask a why-question is by himself suggesting an explanation of what the system has just asserted, while at the same time indicating a desire to have the explanation confirmed by the system. For example, after the system has given the answer (E2.2), the user should be able, by asking the question (E2.6), to elicit an explanation like (E2.7) from the system. If this kind of behavior

(E2.6) U: Because of an accident? (E2.7) S: No, because they're repaying it.

is to be realized in a dialogue system, the program must be able to recognize (E2.6) as a proposed explanation. Algorithms which recognize explanations in certain contexts have been developed, e.g., for the ICAI system ACE¹⁴ and the text-understanding system PAM²³.

Leading and rhetorical questions which suggest an affirmative answer may be seen as containing an implicit request to justify the answer if it is negative. If the system's answer to (E3.1)

(E3.1) U: You aren't going to restrict me to 40k of core today again, are you?
(E3.2) S: Yes, in fact I am. I've got 47 jobs

logged-in in the moment.

is not something like (E3.2), but rather simply Yes, in fact I am, the system isn't exhibiting the sort of cooperative behavior which we would like to have in a natural language dialogue system.

These last two types of speech acts cannot at present be handled adequately by AI systems. The same is true of explanations within the schema reproach-justification (cf. E4.1 and E4.2). (E4.1) U: You erased my file COLING.TMP!

(E4.2) S: Yeah, your log-out quota was exceeded.

Communicative and Cognitive Conditions as Signals for Argumentative Answers

Two further kinds of signals which suggest argumentative answers deserve mention in this section. In contrast to the preceding types they can be incorporated without difficulty into existing AI systems, e.g. HAM-RPM²¹.

Both kinds of signal lead to the question's being over-answered in that they suggest an argumentative answer in the absence of any explicit or implicit request for such an answer in the user's question.

On the one hand, the system may offer an unsolicited explanation for reasons of partner tactics if it has already noticed that the user seems to have a tendency to ask for explanations of answers⁶.

On the other hand, over-answering may even be reasonably expected of the system in the case where the answer is based on uncertain beliefs and approximate or hypothetical reasoning. This kind of behavior can be modelled to a limited extent if the system is programmed so as to attempt to generate an explanation as soon as its confidence in its own answer sinks below a certain threshold, e.g., because the implication strength (see (a) above) of one of the inference rules it has used is low (cf. E5.1, E5.2). The

(E5.1) U: I wonder if the Mercedes is cheap. (E5.2) S: I imagine so -- it's pretty old and rusty.

generation of an argumentative answer in such a context falls outside the usual scope of linguistic analysis; it is a good example of an application of the AI paradigm in that the condition which gives rise to the generation of an argumentative answer is a certain property of a cognitive process, namely the inference process by which the answer is derived.

Figure 2 summarizes the various signals for argumentative answers which have been discussed in this section (types of signals which have been implemented in HAM-RPM's explanation component are indicated by a *).

- question _____ question word * idiomatic expression *
- request _____ direct *
- evidence of doubt in user *
- evidence of a contrary opinion in user *
- inadequate explanation suggested by user
- unexpected answer to a leading or rhetorical question
- evidence of reproach in user
- over-answering _____ partner-tactics * uncertainty about own answer *
- Figure 2: Signals which can elicit an argumentative answer

Speech Act Schemata for Argumentative Dialogue Sequences

This section deals with argumentative dialogue sequences and their reconstruction in Al systems. The speech act sequence depicted in schema 1 will serve as a starting point.

- (S1.1) U: <yes-no-question>
- (S1.2) S: <affirmative answer> (with restricting hedge)
- (S1.3) U: Why?

(\$1.4) S: <argumentative answer>

Interpretation of S1.3 by S: What is the basis for the assertion (belief) in S1.2 that A?

Schema 1: A simple argumentative dialogue sequence

In schema 1, as in the schemata to follow, the word why represents the entire class of signals in the user's utterances for argumentative answers which were discussed in the previous section.

Here is an example of a simple argumentative dialogue sequence:

(E6.1) U: Do you know if the Mercedes is cheap? (E6.2) S: I think so.

- (E6.3) U: What makes you think so?
- (E6.4) S: It's in need of repairs.

Iterated Why-questions and Ultimate Explanations

A sequence such as (E6.1) through (E6.4) may be continued by one or more repititions of schema 2, in which the user requests that the system's argumentative answer itself be explained.

(S2.1) U: Why?

(S2.2) S: <argumentative answer>

Schema 2: Iteration of a why-question

The dialogue sequence (E6.5) through (E6.8) is a continuation of (E6) in which two further whyquestions occur. The answer (E6.8) is an example

- (E6.5) U: Why?
- (E6.6) S: It's in need of repairs because its rear axle is bent.
- (E6.7) U: How come?
- (E6.8) S: That's just the way it is.

of an ultimate explanation. Though it is debatable whether ultimate explanations in a philosophical sense are in fact possible, it is clear that participants in everyday dialogues frequently offer explanations which they are not in a position to explain further. Some typical formulations which are used in such cases are listed in (L7).

The Ambiguity of Iterated Why-questions

A further problem in connection with iterated why-questions is the ambiguity which they regularly involve. Each of the why-questions after the first one can refer either to (a) the assertion which constituted the explanans, or (b) the

inferential relationship between the explanans and the explanandum.

(\$3.1) U: Why Q?

(S3.2) S: Because P. -(a) Why P? (\$3.3) U: Why?= -(b) Why $(P \Rightarrow Q)$?

Schema 3: The ambiguity of an iterated why-question

If the second sort of interpretation is applied to the question (E6.7), an answer such as (E6.9)becomes appropriate.

(E6.9) S: A machine is in need of repairs when one of its parts is in need of repairs.

It is of course possible to eliminate this ambiguity with a more precise formulation of the whyquestion, as when, for example, (S2.1) is replaced with (S2.1).

(S2.1') U: I know that. But why does that make you think that Q?

Although interpretation (a) is far more common than (b) in nontechnical dialogues, the occurrence of questions such as (\$2.1) shows that it is nonetheless worthwhile to provide an Al system with the ability to answer in accordance with either of the possible interpretations. For interpretation (b), this means that the system must be able, like HAM-RPM²¹, to verbalize the inference rules it uses.

If the system is requested, via a further why-question, to explain an inference rule that it has verbalized in this way, the existence of a third type of argument in addition to the presentation of factual evidence and the verbalisation of inference rules becomes evident: The system may supply a backing¹⁸ for its inference rule. A backing usually refers to a convention, a theory, or observations.

An explanation component which uses backings must have access to the corresponding metaknowledge about its inference rules.

The Elicitation of a Multiple Justification

A further variant of schema 2 can be used to exhibit the step-by-step elicitation of a multiple justification. Instead of simply asking another why-question, the user specifically requests further corroborating evidence for the explanandum. Some typical expressions are listed in (L8).

- (L8) Is that all? Any other reason? Just because of that?
- (\$4.1) U: <request for further evidence>
- (\$4.2) S: <corroborating evidence for \$1.2>
- Schema 4: The elicitation of a multiple justification

As the example (E6.10) through (E6.13) shows, schema 4 can be instantiated several times in succession.

(E6.10) U: Is that the only reason?

(E6.11) S: Well, it's pretty old and beat-up. (E6.12) U: Anything else?

(E6.13) S: It's a bit rusty.

Dialogue Schemata with Metacommunicative Why-questions

In all of the dialogue schemata we have examined so far, a why-question asked by the user followed an answer by the system to a previous question. In this section we shall discuss dialogue sequences in which why-questions refer to questions or requests. In fact, of course, any kind of speech act, e.g. a threat or an insult, can give rise to a metacommunicative why-question; the two types to be discussed here are those most relevant to foreseeable applications of natural language Al systems.

Schema 5 will serve as a starting point. In clarification dialogues schema 6, a variant of schema 5, can be instantiated.

- (S5.1) S: <question>,<request>
 (S5.2) U: Why?
- (\$5.3) S: <argumentative answer>

(\$5.4) U: <response to \$5.1>

Interpretation of S5.2 by S: What was the intention underlying the speech act in \$5.1?

Schema 5: A dialogue sequence with a metacommunicative why-question

- (S6.1) U: <question>
- (S6.2) S: <clarification question concerning S6.1>,<request for a paraphrase of \$6.1>
- (S6.3) U: Whu?
- (S6.4) S: <argumentative answer>
- (S6.5) U: <response to S6.2>
- (S6.6) S: <response to S6.1>

Schema 6: A metacommunicative why-question within a clarification dialogue

Here is a dialogue sequence containing a metacommunicative why-question asked by the user:

(E7.1) U: Please list all articles since 1978 on the subject of 'presposition'. (E7.2) S: Do you really mean 'presposition'? (E7.3) U: Why do you ask?

- (E7.4) S: I don't know this word.
- (E7.5) U: I meant 'presupposition'.
- (E7.6) S: I have the following entries: ...

Why-questions Asked by the System

Although all of the why-questions considered so far have been asked by the user, the system can also ask why the user has made a particular input. This situation is described by schema 5 except that the roles of USER (U) and SYSTEM (S) are reversed.

Providing an application-oriented AI system with the ability to ask such why-questions is worthwhile because there are many situations in which the system requires further information about the user's intention to guide its search for an answer or to help to formulate its answer in a communicatively adequate manner. Of course,

the system can only make use of the user's answer to such a why-question if it is equipped with the ability to analyse argumentative answers. The example (E8) might occur in one of HAM-RPM's dialogue situations, in which the system simulates a hotel manager who is anxious to rent a particular room to a caller who is inquiring about it. It illustrates the way information about the dialogue partner's intentions can influence the way a particular state of affairs is described.

(E8.1) U: Has the room got a big desk? (E8.2) S: Why do you ask? (E8.3) U: Because I've got a lot of work to do. (E8.4) S: Yes, the desk is fairly large. (E8.3')U: I hate big desks. (E8.4')S: It isn't particularly big.

The schemata we have investigated in this and the previous sections can also be embedded in one another, as can be seen from schema 7. In this schema, (S7.4), but not (S7.3), is a metacommunicative why-question.

(S7.1) U: <yes-no-question> (S7.2) S: <affirmative answer> (with restricting hedge) (S7.3) U: Why? (S7.4) S: Why do you ask? (S7.5) U: <argumentative answer to S7.4> (S7.6) S: <argumentative answer to S7.3>

Schema 7: Successive why-questions of different types

In mixed-initiative systems, in which either of the partners can initiate a dialogue sequence, the system must be able both to ask and to answer why-questions, including those of a metacommunicative nature. Summary and Integration of All Argumentative Dialogue Schemata Relevant to Al Systems

Figure 3 summarizes and integrates the schemata for argumentative dialogue sequences discussed in the preceding sections. The arrows joining the rectangles indicate that one speech act follows another in time. If arrows join two rectangles in both directions, loops such as those discussed in connection with iterated why-questions are possible. Double vertical lines on the left- or right-hand side of a rectangle indicate that the speech act in question can be the first or the last speech act in a sequence, respectively. The system's criteria for recognizing at each point which of the possible speech acts the user has performed and for selecting its own speech acts are not included in the diagram.

If one extends Figure 3 by permitting the reversal of the roles of system and user, all schemata for argumentative dialogue sequences²¹ are included which are relevant for foreseeable applications in dialogue systems with mixed-ini-tiative.

Technical Data

A non-compiled version of HAM-RPM is running on the DECsystem 1070 (PDP-10) of the Fachbereich fuer Informatik of the University of Hamburg under the TOPS10 operating system. Comprising approximately 600 LISP/FUZZY procedures, the current version occupies 150K of 36-bit words and requires from one to fifteen seconds for a response.

Acknowledgements

I wish to thank Anthony Jameson for careful reading of an earlier draft of this paper.



Figure 3: Schemata for argumentative dialogue sequences in Al systems

References

- Bromberger, S. (1966): Why-questions. In: Colodny, R. (ed.): Mind and Cosmos. Pittsburgh: Univ. Press, p. 86-111
- [2] Chester, D. (1976): The translation of formal proofs into English. In: Artificial Intelligence, 7, 3, p. 261-278
- [3] Cohen, P.R. (1978): On knowing what to say: Planning speech acts. Univ. of Toronto, Dept. of Computer Science, Technical Report No. 118
- [4] Davis, R. (1976): Applications of meta level knowledge to the construction, maintenance and use of large knowledge bases. Stanford Univ., Technical Report STAN-CS-76-562
- [5] Freeman, C. (1976): A pragmatic analysis of tenseless why-questions. In: Mufwene, S.S., Walker, C.A., Steeuer, S.B. (eds.): Papers of the twelfth regional meeting of the Chicago Linguistic Society. Chicago: Chicago Linguistic Society, p. 208-219
- [6] v. Hahn, W., Hoeppner, W., Jameson, A., Wahlster, W. (1980): The anatomy of the natural language dialogue system HAM-RPM. In: Bolc, L. (ed.): Natural language based computer systems. Munich: Hanser/Macmillan
- [7] Hart, P.E., Duda, R.O. (1977): PROSPECTOR -A computer-based consultation system for mineral exploration. Stanford Research International, AI Center, Technical Note 155
- [8] Hayes, P., Reddy, R. (1979): An anatomy of graceful interaction in spoken and written man-machine communication. Carnegie-Mellon-Univ., Dept. of Computer Science, Technical Report CMU-CS-79-144
- [9] Heringer, H.J. (1974): Praktische Semantik. Stuttgart: Klett
- [10] Nakamishi, M., Nagata, M., Ueda, K. (1979): An automatic theorem prover generating a proof in natural language. In: IJCAI-79, Tokyo, p. 636-638
- [11] Sacerdoti, E.D. (1977): A structure for plans and behavior. N.Y.: Elsevier
- [12] Scott, C.A., Clancey, A., Davis, R., Shortliffe, E.H. (1977): Explanation capabilities of production-based consultation systems. In: American Journal of Computational Linguistics, Microfiche 62
- [13] Scragg, G.W. (1974): Answering questions about processes. Univ. of California, San Diego, Ph.D. Thesis

- [14] Sleeman, D.H., Hendley, R.J. (1979): ACE: A system which analyses complex explanations. In: International Journal of Man-Machine Studies, 11, p. 125-144
- [15] Stallman, R.M., Sussman, G.J. (1977): Forward reasoning and dependency-directed backtracking in a system for computer-aided circuit analysis. In: Artificial Intelligence 9, p. 135-196
- [16] Swartout, W.R. (1977): A Digitalis therapy advisor with explanations. MIT Lab. for Computer Science, Technical Report TR-176
- [17] Tondi, L. (1969): Semantics of the question in a problem-solving situation. Czech Academy of Science, Prague
- [18] Toulmin, S. (1969): The uses of argument. Cambridge: Univ. Press (1st ed. 1958)
- [19] Wahlster, W., Jameson, A., Hoeppner, W. (1978): Glancing, referring and explaning in the dialogue system HAM-RPM. In: American Journal of Computational Linguistics, Microfiche 77, p. 53-67
- [20] Wahlster, W. (1980): Implementing fuzziness in dialogue systems. In: Rieger, B. (ed.): Empirical Semantics. Bochum: Brockmeyer
- [21] Wahlster, W. (1980): Automatic generation of natural language explanations for conclusions based on fuzzy inferences. (in preparation)
- [22] Weiner, J.L. (1979): The structure of natural explanation: Theory and application. System Development Corporation, Santa Monica, Technical Report SP-4035
- [23] Wilensky, R. (1978): Why John maried Mary: Understanding stories involving recurring goals. In: Cognitive Science, 2, p. 235-266
- [24] Winograd, T. (1972): Understanding natural language. N.Y.: Academic