

# Dialog Control in a Natural Language System<sup>1</sup>

Michael Gerlach      Helmut Horacek

Universität Hamburg  
Jungiusstraße 6

Fachbereich Informatik    Projektgruppe WISBER  
D-2000 Hamburg 36      F.R.G.

## ABSTRACT

In this paper a method for controlling the dialog in a natural language (NL) system is presented. It provides a deep modeling of information processing based on time dependent propositional attitudes of the interacting agents. Knowledge about the state of the dialog is represented in a dedicated language and changes of this state are described by a compact set of rules. An appropriate organization of rule application is introduced including the initiation of an adequate system reaction. Finally the application of the method in an NL consultation system is outlined.

## INTRODUCTION

The solution of complex problems frequently requires cooperation of multiple agents. A great deal of interaction is needed to identify suitable tasks whose completion contributes to attaining a common goal and to organize those tasks appropriately. In particular, this involves carrying out communicative subtasks including the transfer of knowledge, the adjustment of beliefs, expressing wants and pursuing their satisfaction, all of which is motivated by the intentions of the interacting agents [Werner 88]. An ambitious dialog system (be it an interface, a manipulation system, or a consultation system) which is intended to exhibit (some of) these capabilities should therefore consider these intentions in processing

the dialog, at least to the extent that is required for the particular type of the dialog and the domain of application.

A considerable amount of work in current AI research is concerned with inferring intentions from utterances (e.g., [Allen 83], [Carberry 83], [Grosz, Sidner 86]) or planning speech acts serving certain goals (e.g., [Appelt 85]), but only a few uniform approaches to both aspects have been presented.

Most approaches to dialog control described in the literature offer either rigid action schemata that enable the simulation of the desired behavior on the surface (but lack the necessary degree of flexibility, e. g., [Metzing 79]), or descriptive methods which may also include possible alternatives for the continuation of the dialog, but without expressing criteria to guide an adequate choice among them (e. g., [Meßing et al. 87], [Bergmann, Gerlach 87]).

Modeling of beliefs and intentions (i.e., of propositional attitudes) of the agents involved is found only in the ARGOT system [Litman, Allen 84]. This approach behaves sufficiently well in several isolated situations, but it fails to demonstrate a continuously adequate behavior in the course of a complete dialog. An elaborated theoretical framework is provided in [Cohen, Perrault 79] but they explicitly exclude the deletion of propositional attitudes. Hence, they cannot explain what happens when a want has been satisfied.

In our approach we have enhanced the propositional attitudes by associating them with time intervals expressing their time of validity. This enables us to represent the knowledge about the actual state of the dialog (and also about past states) seen from the point of view of a certain agent and to express changes in the propositional attitudes

---

<sup>1</sup> The work described in this paper is part of the joint project WISBER, which is supported by the German Federal Ministry for Research and Technology under grant ITW-8502. The partners in the project are: Nixdorf Computer AG, SCS Orbit GmbH, Siemens AG, the University of Hamburg, and the University of Saarbrücken.

occurring in the course of the dialog and to calculate their effect. This deep modeling is the essential resource for controlling the progress of the conversation in approaching its overall goal, and, in particular, for determining the next subgoal in the conversation which manifests itself in a system utterance.

We have applied our method in the NL consultation system WISBER ([Horacek et al. 88], [Sprenger, Gerlach 88]) which is able to participate in a dialog in the domain of financial investment.

## REPRESENTING PROPOSITIONAL ATTITUDES

Knowledge about the state of the dialog is represented as a set of propositional attitudes. The following three types of propositional attitudes of an agent towards a proposition *p* form a basic repertoire :

**KNOW :** The agent is sure that *p* is true. This does not imply that *p* is really true since the system has no means to find out the real state of the world. Assuming that the user of a dialog system obeys the sincerity condition (i.e., always telling the truth, c.f. [Grice 75]) an assertion uttered by the user implies that the user *knows* the content of that assertion.

**BELIEVE :** The agent believes, but is not sure, that *p* is true, or he/she assumes *p* without sufficient evidence.

**WANT :** The agent wants *p* to be true.

Propositional attitudes are represented in our semantic representation language IRS, which is used by all system components involved in semantic-pragmatic processing. IRS is based on predicate calculus, and contains a rich collection of additional features required by NL processing (see [Bergmann et al. 87] for detailed information). A propositional attitude is written as

(*<type>* *<agent>* *<prop>* *<time>*):

- *<type>* is an element of the set: KNOW, BELIEVE, and WANT.
- The two agents relevant in a dialog system are the USER and the SYSTEM. In addition, we use the notion 'mutual knowledge'. Informally, this means that both the user and

the system know that *<prop>* is true, and that each knows that the other knows, recursively. We will use the notation (KNOW MUTUAL *<prop>* ...) to express that the proposition *<prop>* is mutually known by the user and the system.

- *<prop>* is an IRS formula denoting the proposition the attitude is about. It may again be a propositional attitude, as in (WANT USER (KNOW USER *x* ...) ...) which means that the user wants to know *x*. The proposition may also contain the meta-predicates *RELATED* and *AUGMENT*: (*RELATED x*) means 'something which is related to the individual *x*', i.e., it must be possible to establish a chain of links connecting the individual and the proposition. In this general form *RELATED* is only used to determine assumptions about the user's competence. For a more intensive application, however, further conditions must be put on the connecting links. (*AUGMENT f*) means 'something more specific than the formula *f*', i.e., at least one of the variables must be quantified or categorized more precisely or additional propositions must be associated. These meta-predicates are used by the dialog control rules as a very compact way of expressing general properties of propositions.
- Propositional attitudes as any other states hold during a period of time. In WISBER we use Allen's time logic [Allen 84] to represent such temporal information [Poesio 88]. *<time>* must be an individual of type TIME-INTERVAL. In this paper, however, for the sake of brevity we will use almost exclusively the special constants NOW, PAST and FUTURE, denoting time intervals which are asserted to be during, before or after the current time.

## INFERENCE RULES

As new information is provided by the user and inferences are made by the system, the set of propositional attitudes to be represented in the system will evolve. While the semantic-pragmatic analysis of user utterances exploits linguistic features to derive the

attitudes expressed by the utterances (c.f. [Gerlach, Sprenger 88]), the dialog control component interprets rules which embody knowledge about knowing and wanting as well as about the domain of discourse. These rules describe communicative as well as non-communicative actions, and specify how new propositional attitudes can be derived. Rules about the domain of discourse express state changes including the involved action. The related states and the triggering action are associated with time-intervals so that the correct temporal sequence can be derived.

Both classes of rules are represented in a uniform formalism based on the schema *precondition - action - effect*:

- The *precondition* consists of patterns of propositional attitudes or states in the domain of discourse. The patterns may contain temporal restrictions as well as the meta-predicates mentioned above. A precondition may also contain a rule description, e.g., to express that an agent knows a rule.
- The *action* may be either on the level of communication (in the case of speech act triggering rules) or on the level of the domain (actions the dialog is about). However, there are also pure inference rules in the dialog control module; their action part is void.
- The *effect* of a rule is a set of descriptions of states of the world and propositional attitudes which are instantiated when applying the rule yielding new entries in the system's knowledge base. We do not delete propositional attitudes or other propositions, i.e., the system will not forget them, but we can mark the time interval associated with an entry as being 'finished'. Thus we can express that the entry is no longer valid, and it will no longer match a pattern with the time of validity restricted to NOW.

## CONTROL STRUCTURE

So far, we have only discussed how the actual state of the dialog (from the point of view of a certain agent) can be represented and how changes in this

state can be described. We still need a method to determine and carry out the relevant changes, given a certain state of the dialog, after interpreting a user utterance (i.e., to decide which dialog rules may be tried and in which order). For reasons of simplicity we have divided the set of rules into three subsets each of them being responsible for accomplishing a specific subtask, namely:

- gaining additional information inferable from the interrelation between recent information coming from the last user utterance and the actual dialog context. The combination of new and old information may, e. g., change the degree of certainty of some proposition, i. e., terminate an (uncertain) BELIEVE state and create a (certain) KNOW state with identical propositional content (the consistency maintenance rule package).
- pursuing a global (cognitive or manipulative) goal; this may be done either by trying to satisfy this goal directly, or indirectly by substituting a more adequate goal for it and pursuing this new goal. In particular, a goal substitution is urgently needed in case the original goal is unsatisfiable (for the system), but a promising alternative is available (the goal pursuit rule package).
- pursuing a communicative subgoal. If a goal can not (yet) be accomplished due to lack of information, this leads to the creation of a WANT concerning knowledge about the missing information. When a goal has been accomplished or a significant difference in the beliefs of the user and the system has been discovered, the system WANTS the user to be informed about that. All this is done in the phase concerned with cognitive goals. Once such a WANT is created, it can be associated with an appropriate speech act, provided the competent dialog partner (be it the user or an external expert) is determined (the speech act triggering rule package).

There is a certain linear dependency between these subtasks. Therefore the respective rule packages are applied in a suitable (sequential) order, whereas those rules belonging to the same pack-

age may be applied in any order (there exist no interrelations within a single rule package). This simple forward inferencing works correctly and with an acceptable performance for the actual coverage and degree of complexity of the system.

A sequence consisting of these three subtasks forms a (cognitive) processing cycle of the system from receiving a user message to initiating an adequate reply. This procedure is repeated until there is evidence that the goal of the conversation has been accomplished (as indicated by knowledge and assumptions about the user's WANTS) or that the user wants to finish the dialog. In either case the system closes the dialog.

## APPLICATION IN A CONSULTATION SYSTEM

In this section we present the application of our method in the NL consultation system WISBER involving rather complex interaction with subdialogs, requests for explanation, recommendations, and adjustment of proposals. However, it is possible to introduce some simplifications typical for consultation dialogs. These are urgently needed in order to reduce the otherwise excessive amount of complexity. In particular, we assume that the user does not lie and take his/her assertions about real world events as true (the sincerity condition). Moreover, we take it for granted that the user is highly interested in a consultation dialog and, therefore, will pay attention to the conversation on the screen so that it can be reasonably assumed that he/she is fully aware of all utterances occurring in the course of the dialog.

Based on these (implicit) expectations, the following (simplified) assumptions (1) and (2) represent the starting point for a consultation dialog:

- (1) (BELIEVE SYSTEM  
(WANT USER  
((EXIST X (STATE X))  
(HAS-EXPERIENCER X USER))  
NOW) NOW)
- (2) (BELIEVE SYSTEM  
(KNOW USER  
(RELATED ((EXIST Y (STATE Y))  
(HAS-EXPERIENCER Y USER)))

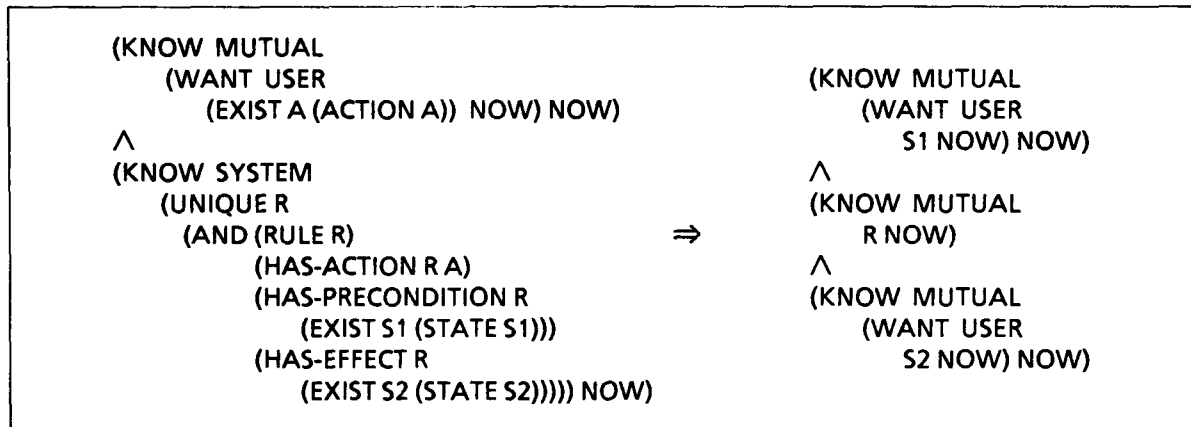
NOW) NOW)

They express that the user knows something that 'has to do' (expressed by the meta-predicate *RELATED*) with states (STATE Y) concerning him/herself and that he/she wants to achieve a state (STATE X). In assumption 1, (STATE X) is in fact specialized for a consultation system as a real world state (instead of a mental state which is the general assumption in any dialog system). This state can still be made more concrete when the domain of application is taken into account:

In WISBER, we assume that the user wants his/her money 'to be invested.' The second assumption expresses (a part of) the competence of the user. This is not of particular importance for many other types of dialog systems. In a consultation system, however, this is the basis for addressing the user in order to ask him/her to make his/her intentions more precise. In the course of the dialog these assumptions are supposed to be confirmed and, moreover, their content is expected to become more precise.

In the subsequent paragraphs we outline the processing behavior of the system by explaining the application and the effect of some of the most important dialog rules (at least one of each of the three packages introduced in the previous section), thus giving an impression of the system's coverage. In the rules presented below, variables are suitably quantified as they appear for the first time in the precondition. In subsequent appearances they are referred to like constants. The interpretation of the special constants denoting time-intervals depends on whether they occur on the left or on the right side of a rule: in the precondition the associated state/event must hold/occur during PAST, FUTURE or overlaps NOW; in the effect the state/event is associated with a time-interval that starts at the reference time-interval.

In a consultation dialog, the user's wants may not always express a direct request for information, but rather refer to events and states in the real world. From such user wants the system must derive requests for knowledge useful when attempting to satisfy them.<sup>2</sup> Consequently the task of infer-



**Rule 1:** Inference drawn from a user want referring to an action with unambiguous consequences (pursuing a global goal)

ring communicative goals is of central importance for the functionality of the system.

There is, however, a fundamental distinction whether the content of a want refers to a state or to an event (to be more precise, to an action, mostly). In the latter case some important inferences can be drawn depending on the domain knowledge about the envisioned action and the degree of precision expressed in its specification. If, according to the system's domain model, the effect of the specified action is unambiguous, the user can be expected to be familiar with this relation, so he/she can be assumed to envision the resulting state and, possibly, the precondition as well, if it is not yet fulfilled. Thus, in principle, a plan consisting of a sequence of actions could be created by application of skillful rule chaining.

This is exactly what Rule 1 asserts: Given the mutual knowledge that the user wants a certain action to occur, and the system's knowledge (in form of a unique rule) about the associated precondition and effect, the system concludes that the user envisions the resulting state and he/she is familiar with the connecting causal relation. If the uniqueness of the rule cannot be

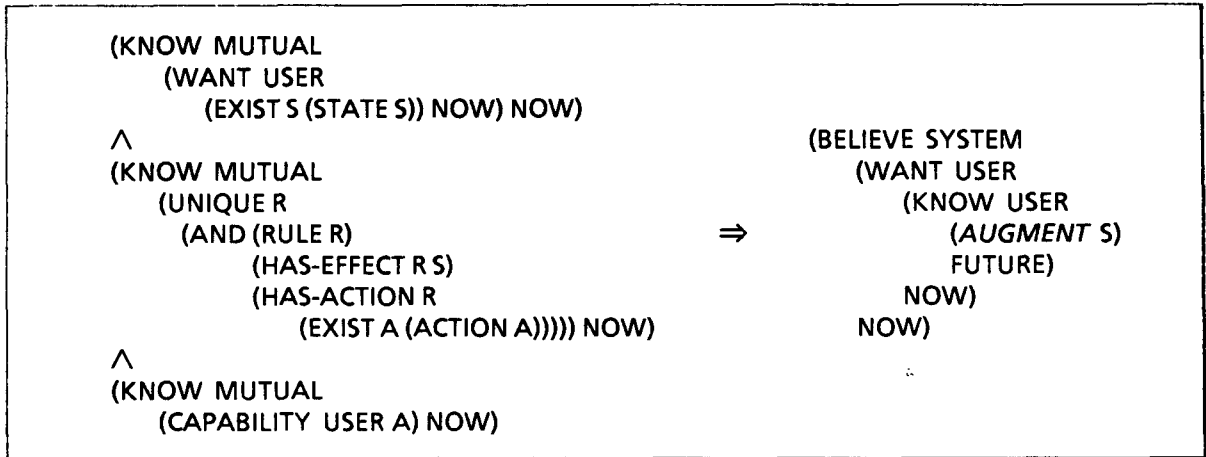
established, sufficient evidence derived from the partner model might be an alternative basis to obtain a sufficient categorization of the desired event so that a unique rule is found. Otherwise the user has to be asked to precise his/her intention.

Let us suppose, to give an example, that the user has expressed a want to invest his/her money. According to WISBER's domain model, there is only one matching domain rule expressing that the user has to possess the money before but not after investing his/her money, and obtains, in exchange, an asset of an equivalent value. Hence Rule 1 fires. The want expressed by the second part of the conclusion can be immediately satisfied as a consequence of the user utterance 'I have inherited 40 000 DM' by applying Rule 5 (which will be explained later). The remainder part of the conclusion matches almost completely the precondition of Rule 2.

This rule states: If the user wants to achieve a goal state (G) and is informed about the way this can be done (he/she knows the specific RULE R and is capable of performing the relevant action), the system is right to assume that the user is lacking some information which inhibits him/her from actually doing it. Therefore, a want of the user indicating the intention to know *more* about this transaction is created (expressed by the meta-predicate *AUGMENT*). If the necessary capability cannot be attributed to the user a consultation is impossible.

If, to discuss another example, the user has expressed a want aiming at a cer-

<sup>2</sup> Unlike other systems, e.g., UC [Wilensky et al. 84], which can directly perform some kinds of actions required by the user, WISBER is unable to affect any part of the real world in the domain of application.



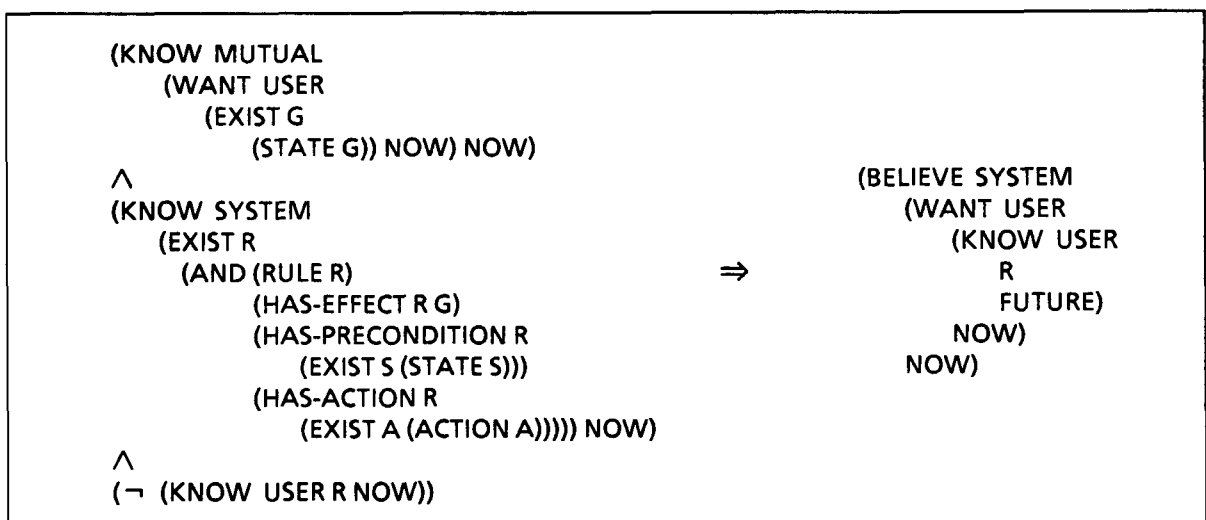
**Rule 2:** Inference drawn from a user want referring to a state, given his/her acquaintance with the associated causal relation (pursuing a global goal)

tain state (e.g., 'I want to have my money back'), the application of another rule almost identical to Rule 1 is attempted. When its successful application yields the association of a unique event, the required causal relation is established. Moreover, the user's familiarity with this relation must be derivable in order to follow the path indicated by Rule 2. Otherwise, a want of the user would be created whose content is to find out about suitable means to achieve the desired state (as expressed by Rule 3, leading to a system reaction like, e.g., 'you must dissolve your savings account').

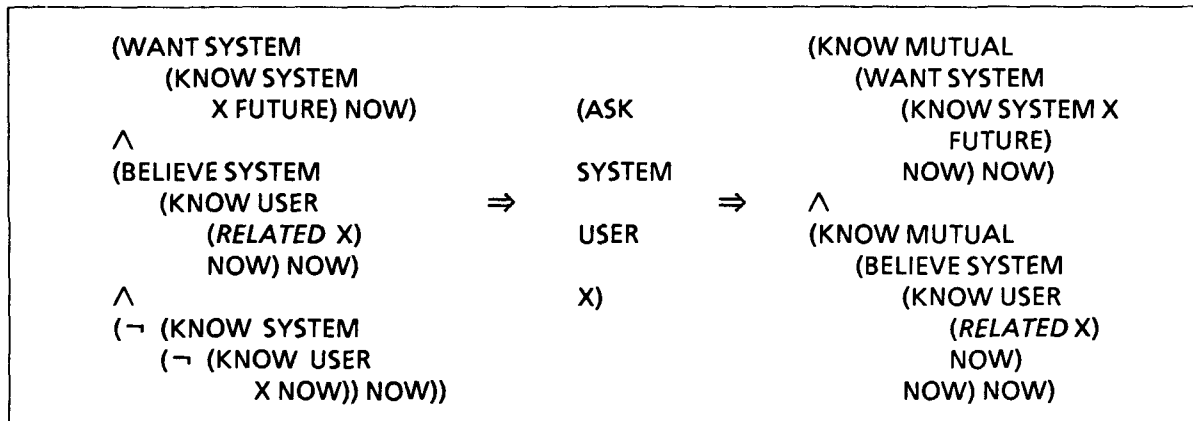
It is very frequently the case that the satisfaction of a want cannot immediately be achieved because the precision

of its specification is insufficient. When the domain-specific problem solving component indicates a clue about what information would be helpful in this respect this triggers the creation of a system want to get acquainted with it. Whenever the user's uninformedness in a particular case is not yet proved, and this information falls into his/her competence area, control is passed to the generation component to address a suitable question to the user (as expressed in Rule 4).

Provided with new information hopefully obtained by the user's reply the system tries again to satisfy the (more precisely specified) user want. This process is repeated until an adequate degree of specification is achieved at some stage.



**Rule 3:** Inference drawn from a user want referring to a state, missing his/her acquaintance with the associated causal relation (pursuing a global goal)



**Rule 4:** Inference drawn from the user's (assumed) competence and a system want in this area (triggering a speech act)

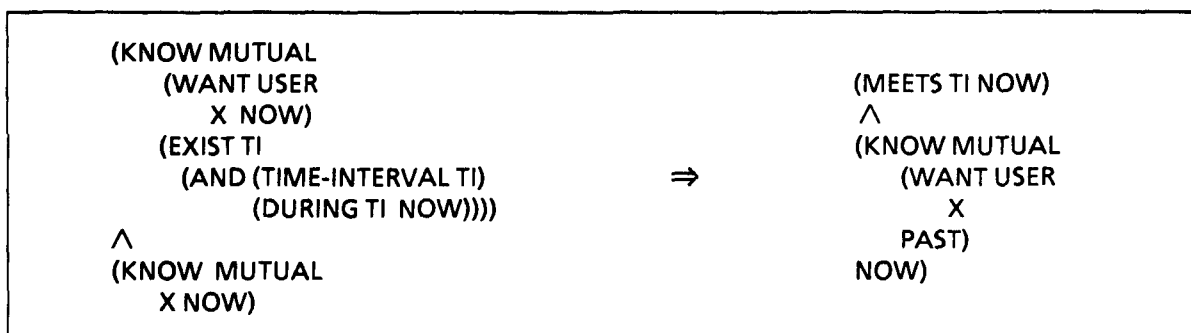
In the course of the dialog each utterance effects parts of the system's current model of the user (concerning assumptions or temporarily established knowledge). Therefore, these effects are checked in order to keep the data base consistent. Consider, for instance, a user want aiming at investing some money which, after a phase of parameter assembling, has led to the system proposal 'I recommend you to buy bonds' apparently accomplishing the (substituted) goal of obtaining enough information to perform the envisioned action. Consequently, the state of the associated user want is subject to change which is expressed by Rule 5.

Therefore, the mutual knowledge about the user want is modified (by closing the associated time-interval) and the user's want is marked as being 'finished' and added to the (new) mutual knowledge.

However, this simplified treatment of the satisfaction of a want includes the

restrictive assumptions that the acceptance of the proposal is (implicitly) anticipated, and that modifications of a want or of a proposal are not manageable. In a more elaborated version, the goal accomplishment has to be marked as provisory. If the user expresses his/her acceptance either explicitly or changes the topic (thus implicitly agreeing to the proposal), the application of Rule 5 is fully justified.

Apart from the problem of the increasing complexity and the amount of necessary additional rules, the preliminary status of our solution has much to do with problems of interpreting the *AUGMENT*-predicate which appears in the representation of a communicative goal according to the derivation by Rule 2: The system is satisfied by finding any additional information augmenting the user's knowledge, but it is not aware of the requirement that the information must be a suitable supplement (which is recognizable by the user's confirmation only).



**Rule 5:** Inference drawn from a (mutually known) user want which the user knows to be accomplished (pursuing consistency maintenance)

## FUTURE RESEARCH

The method described in this paper is fully implemented and integrated in the complete NL system WISBER. A relatively small set of rules has proved sufficient to guide basic consultation dialogs. Currently we are extending the set of dialog control rules to perform more complex dialogs. Our special interest lies on clarification dialogs to handle misconceptions and inconsistencies. The first steps towards handling inconsistent user goals will be an explicit representation of the interrelationships holding between propositional attitudes, e.g., goals being simultaneous or competing, or one goal being a refinement of another goal. A major question will be specifying the operations necessary to recognize those interrelationships working on the semantic representation of the propositional contents. As our set of rules grows, a more sophisticated control mechanism will become necessary, structuring the derivation process and employing both forward and backward reasoning.

## REFERENCES

- Allen 83**  
Allen, J.F.: *Recognizing Intentions from Natural Language Utterances*. In: Brady, M., Berwick, R.C. (eds.): *Computational Models of Discourse*, MIT Press, 1983, pp. 107-166.
- Allen 84**  
Allen, J.F.: *Towards a General Theory of Action and Time*. In: *Artificial Intelligence* 23 (2), 1984, pp. 123-154.
- Appelt 85**  
Appelt, D.E.: *Planning English Sentences*. Cambridge University Press, 1985.
- Bergmann, Gerlach 88**  
Bergmann, H., Gerlach, M.: *Semantisch-pragmatische Verarbeitung von Äußerungen im natürlichsprachlichen Beratungssystem WISBER*. In: Brauer, W., Wahlster, W. (eds.): *Wissensbasierte Systeme - GI-Kongress 1987*. Springer Verlag, Berlin, 1987, pp. 318-327.
- Bergmann et al. 87**  
Bergmann, H., Fliegner, M., Gerlach, M., Marburger, H., Poesio, M.: *IRS - The Internal Representation Language*. WISBER Bericht Nr. 14, Universität Hamburg, 1987.
- Carberry 83**  
Carberry, S.: *Tracking User Goals in an Information-Seeking Environment*. In: *Proceedings of the AAAI-83*, Washington, D.C., 1983, pp. 59-63.
- Cohen, Perrault 79**  
Cohen, P.R., Perrault, C.R.: *Elements of a Plan-*
- Based Theory of Speech Acts*. In: *Cognitive Science* 3, 1979, pp. 177-212.
- Gerlach, Sprenger 88**  
Gerlach, M., Sprenger, M.: *Semantic Interpretation of Pragmatic Clues: Connectives, Modal Verbs, and Indirect Speech Acts*. In: *Proc. of COLING-88*, Budapest, 1988, pp. 191-195.
- Grice 75**  
Grice, H.P.: *Logic and Conversation*. in: Cole, Morgan (ed.): *Syntax and Semantics, Vol. 3: Speech Acts*, Academic Press, New York, 1975, pp. 41-58.
- Grosz, Sidner 86**  
Grosz, B.J., Sidner, C.L.: *Attention, Intentions, and the Structure of Discourse*. In: *Computational Linguistics* 12 (3), 1986, pp. 175-204.
- Horacek et al. 88**  
Horacek, H., Bergmann, H., Block, R., Fliegner, M., Gerlach, M., Poesio, M., Sprenger, M.: *From Meaning to Meaning - a Walk through WISBER*. In: Hoepfner, W. (ed.): *Künstliche Intelligenz - GWAI-88*, Springer Verlag, Berlin, 1988, pp. 118-129.
- Litman, Allen 84**  
Litman, D.J., Allen, J.F.: *A Plan Recognition Model for Clarification Subdialogues*. In: *Proc. COLING'84*, Stanford, pp. 302-311.
- Meßing, et al. 87**  
Meßing, J., Liermann I., Schachter-Radig M.-J.: *Handlungsschemata in Beratungsdialogen - Am Gesprächsgegenstand orientierte Dialoganalyse*. Bericht Nr. 18, WISBER-Verbundprojekt, Dezember 1987, SCS Organisationsberatung und Informationstechnik GmbH, Hamburg.
- Metzing 79**  
Metzing, D.: *Zur Entwicklung prozeduraler Dialogmodelle*. In: Metzing, D. (Ed.): *Dialogmuster und Dialogprozesse*. Helmut Buske Verlag, Hamburg, 1979.
- Poesio 88**  
Poesio, M.: *Toward a Hybrid Representation of Time*. In: *Proc. of the ECAI-88*, München, 1988, pp. 247-252.
- Sprenger, Gerlach 88**  
Sprenger, M., Gerlach, M.: *Expectations and Propositional Attitudes - Pragmatic Issues in WISBER*. In: *Proc. of the International Computer Science Conference '88*, Hong Kong, 1988.
- Werner 88**  
Werner, E.: *Toward a Theory of Communication and Cooperation for Multiagent Planning*. In: *Theoretical Aspects of Reasoning about Knowledge, Proceedings of the 1988 Conference*, Morgan Kaufman Publishers, Los Altos, 1988, pp. 129-143.
- Wilensky et al. 84**  
Wilensky, R., Arens, Y., Chin, D.: *Talking to UNIX in English: An Overview of UC*. In: *Communications of the ACM*, Vol. 27, No. 6, pp. 574-593.