Semantic Interpretation of Pragmatic Clues: Connectives, Modal Verbs, and Indirect Speech Acts

Michael GERLACH, Michael SPRENGER

University of Hamburg
Department of Computer Science, Project WISBER, P.O. Box 302762
Jungiusstrasse 6, D-2000 Hamburg 36, West Germany

Abstract

Much work in current research in the field of semantic - pragmatic analysis has been concerned with the interpretation of natural language utterances in the context of dialogs. In this paper, however, we will present methods for a primary pragmatic analysis of single utterances. Our investigations involve problems which are not currently well understood, for example how to infer the speaker's intentions by using interpretation of connectives and modal verbs.

This work is part of the joint project WISBER which is supported by the German Federal Ministery for Research and Technology. The partners in the project are: Nixdorf Computer AG, SCS GmbH, Siemens AG, the University of Hamburg and the University of Saarbrücken.

Introduction

Much work in current research in the field of semantic - pragmatic analysis has been concerned with the interpretation of natural language utterances in the context of dialogs, e.g., determining the speaker's goals [Allen 83], deriving beliefs of one agent about another [Wilks/ Bien 83], and planning speech acts [Appelt 85]. In this paper, however, we will present methods for a primary pragmatic analysis of single utterances to construct user model entries which are the starting point for the higher level inference processes just mentioned. Our investigations involve problems which are not currently well understood, for example, how to infer the speaker's intentions by using interpretation of connectives and modal verbs.

Our work is a part of the natural language consultation system WISBER [Bergmann/Gerlach 87]. Consultation dialogs require a much wider class of utterances to be understood than other applications (e.g., for data base interface). In advisory dialogs wants and beliefs play a central role. Although a consultation system must be capable of handling the linguistic means which are used for expressing those attitudes, problems of how to treat modal verbs have received little attention in artificial intelligence and computational linguistics.

The interpretation processes described in this paper work with our semantic representation language IRS [Bergmann et. al. 87] and generate entries for the user model. Representations of utterances in IRS still contain uninterpreted linguistic features such as modal verbs, modal hedges, connectives, and tense information. We are presenting methods for deriving the meaning of these features as they occur in utterances: transforming idiomatically-used indirect speech acts, interpreting connectives in compound sentences, and resolving ambiguities in the meaning of modal verbs by using, i.a., temporal restrictions. The last chapter sketches the technical means used by these processes, i.e., the semantic representation language, the way rules are encoded, and the assertional knowledge base containing the user model.

Fig. 1 shows the different stages of the interpretation process. First, if a connective is found, the analysis process breaks up the sentence into separate propositions. In the next step idomatically-used indirect speech acts are transformed into a direct question. The propositions are then interpreted independently during the modal verb analysis which creates one or more propositional attitudes for each proposition. These interpretations are then related, depending on the natural language connective. Finally, after inferring the appropriate time intervals from verb tense, the sentence type is used to derive the propositional attitudes which are entered into the user model.

Transformation of Idiomatically-Used Indirect Speech Acts

Speakers often use indirect speech acts because they want to express politeness or uncertainty. Examples are : "Could you please tell me which bonds have the highest interest rate?" , "I'd like to know which...", "I do not know which...." We believe that for appropriately handling such an idiomatic use of indirect speech acts in a consultation system it is admissible to transform such utterances into a simplified form - the corresponding direct question. Therefore the first step in our semantic-pragmatic interpretation is mapping the different ways of asking questions onto one standard form which is the formal representation of the equivalent direct question.

Fig. 2 shows the rule which applies to the idiom "I do not know whether X." and transforms it into the representation of the direct speech act "X ?" The rule formalism will be described in detail later.

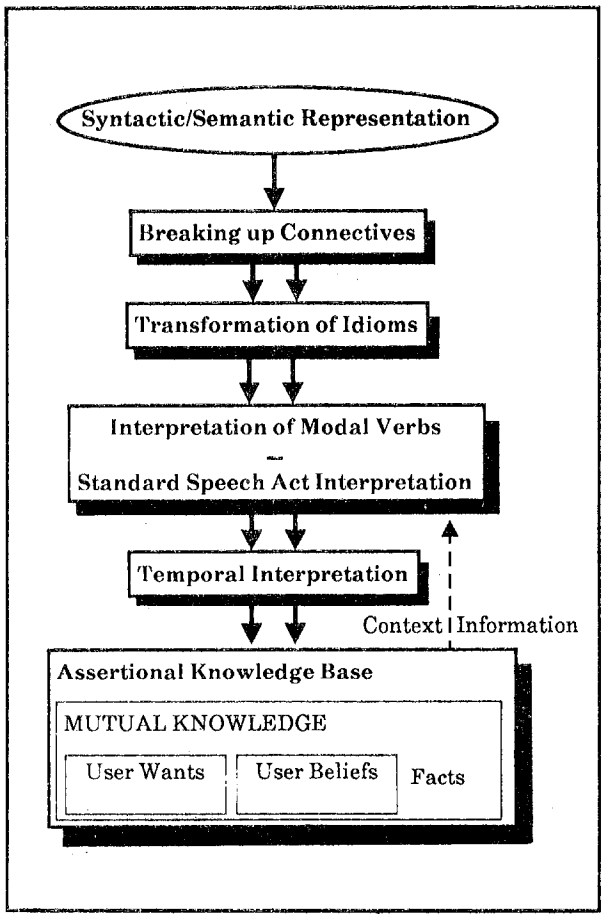

Fig. 1: The stages of the interpretation process

During that transformation process we do not loose any information which might be relevant to the dialog control component of the system (not described in this paper). Before answering any question - direct or indirect - the system has to check whether it is able to answer that question. If this is not the case the user must be informed about the limitations of the system's competence, anyway. This argumentation is similar to that of [Ellman 83], who argues that it is not relevant whether an utterance is a request or an inform as long as the hearer can detect the speaker's superordinate goals.

```
(AND (ASSERTION ?A)
     (HAS-AGENT ?A USER)
     (HAS-PROP ?A ?P)
     (PROP ?P
          (NOT
               (AND (KNOW ?K)
                    (HAS-EXPERIENCER ?K USER)
                    (HAS-OBJECT ?K ?X)))))

  ⇒

(AND (QUESTION ?Q)
     (HAS-AGENT ?Q USER)
     (HAS-OBJECT ?Q ?X))
```

Fig. 2: A rule for transforming the idiom:
"I do not know X."

The transformation of indirect speech acts works on the semantic level by applying rules which specify formal transformations of semantic representations of sentences. In this our approach differs from that taken in UC [Wilensky et. al. 84 and Zernik/Dyer 85] where a *phrasal lexicon* is used and the semantic interpretation of idioms is done during the parsing process.

Interpretation of Modal Verbs

An adequate treatment of modal verbs is necessary for determining the attitudes of the speaker concerning the state of affairs expressed by the proposition he is asserting.[1] The main problem in interpreting modal verbs is their typical ambiguity, e.g.,

(1) Mein Sohn soll viel Geld haben.
In English the two readings are:
'My son is supposed to have a lot of money.'
vs.
'I want my son to have a lot of money.'

Our rules for disambiguating the different readings are based on information which is stored in the semantic representation of the utterance: information about semantic categories of the subject of the modal verb (e.g., ANIMATE, GENERIC, DEFINITE), the relation between the time expressed by the modal verb and the time of the proposition and whether the proposition denotes a state or an event.

(2) Ich habe 10000 Mark geerbt und möchte das Geld in Wertpapieren anlegen. Sie *sollen* eine Laufzeit von vier Jahren haben.
'I have inherited 10000 Marks and would like to invest the money in securities.'
Two readings of the second sentence:
'They are supposed to have a term of four years.'
vs.
'They should have a term of four years.'

In the first reading of the second sentence the entry for the user model must contain the proposition embedded in a belief context, while the second reading must lead to an entry under speaker's wants. In order to resolve this ambiguity, the rules compare the time of the proposition with the tense of the modal verb. For example, if the tense of the modal verb is present and the time of the proposition is sometime in the future, the system decides that the "want" reading is appropriate. The problem in our example is to determine the time of the proposition: We have only the information of tense *haben* (to have) which is a present infinitive and might also denote a future state. Hence the system tries to find out whether the object of the proposition appears in a want context of the speaker. This is the case as is clear from the previous utterance ... *and I want to invest the money in securities* and therefore the system decides to put the proposition of the second sentence into the user's want context as well. (Even if the second utterance is taken to be a belief of the speaker, the fact that it is cited in this context is sufficient to infer that it is also a want, why else should the speaker cite this fact in connection with his decision to invest in securities?)

1) For the semantics of English modal verbs, which is quite different from the German, see [Boyd/Thorne 69]. For German modal verbs see [Brünner/Redder 83], [Reinwein 77], [Sprenger 88].

Usually the user's questions are interpreted as *user wants to know p* (or more formally: (WANT USER (KNOW USER P))), where p denotes the propositional content of the question. For example,

(3) Können Pfandbriefe mehr als 7% Rendite haben?

'Can bonds have an interest rate of more then 7%?'

is interpreted as: the user wants to know whether the proposition is true, which means in our example, taking into account the modal verb *können*, whether it is possible for bonds to have an interest rate of more then 7 %.

One problem arises when the modal verb *sollen* occurs in a question. Normally it is interpreted as indicating a want, e.g.,

(4) Soll ich das Fenster schließen?

'Should I close the window?'

Here the speaker wants to know, whether there is some other person (probably the hearer), who wants the proposition to be true. But this interpretation doesn't make any sense in a *consulting* dialog. In a consultation the speaker is not interested in the wants of the advisor, e.g.,

(5) Soll ich Pfandbriefe mit 5% Rendite kaufen?

'Should I buy bonds which have an interest rate of 5 %?'

Rather than inquiring about someone else's wants, as in (4), the speaker is interested in a recommendation:
(WANT USER (KNOW USER (RECOMMEND SYSTEM P)))

The interpretation of modal verbs is further influenced by connectives which may occur in complements. Consider the following sentence:

(6) Meine Schwester muß viel Geld haben.
'My sister must have a lot of money.'

In this case one can only infer that the speaker believes that the proposition is true, namely that his sister has a lot of money. The interpretation completely changes when we have:

(7) Meine Schwester muß viel Geld haben, *um ihr Haus zu bauen.*
'My sister needs to have a lot of money *in order to build her house.*'

It is possible that the speaker believes as in (6) that his sister has a lot of money, but this cannot be inferred from the statement. Here we can only infer that the speaker believes that the second proposition (his sister's building her house) *implies* the first one (his sister's having a lot of money).

Connectives

Connectives are a means of expressing the argumentative and logical structure of the speaker's opinions by linking propositions. Such relations between propositions are classified into several categories such as *inferential, temporal, causal* linkages [Cohen 84 and Brée/Smit 86]. The system interprets underlying beliefs and wants and enters them into the user model in accordance with the different classes of connectives.

As an example, take the class of connectives which express inferences of the speaker, e.g.,

(8) Ich will eine Anlage mit kurzer Laufzeit, *damit* ich schnell an mein Geld herankommen kann.

'I want a short term investment *so that* I can get my money back quickly.'

Because of the connective *damit* the system concludes that the proposition of the second part of the sentence is the superordinate goal rather than the first proposition although this is the want which is expressed directly. The user supposes that the first proposition is a necessary condition for the second, which expresses his goal. When further processing this logical structure, the system can recognize the underlying misconception, namely that it is not the term of an investment which is important for getting the money back quickly, but the liquidity.

The interpretation of connectives depends on the occurrence of modal verbs, as the following examples demonstrate:

(9) *Soll* ich meine Wertpapiere verkaufen, *um* meine Hypothek *zu* bezahlen?

'*Should* I sell my securities *to* pay off my mortgage?'

(10) *Muß* ich Gebühren bezahlen, *um* mein Sparbuch aufzulösen?

'Do I *have to* pay a fee *to* desolve my savings account?'

In (9) the modal verb *sollen* inside the question indicates that the user wants a recommendation. It indicates further that the connective *um-zu* has to be interpreted as a user's want. The correct interpretation is that the user wants to know whether the system would recommend that the user attempts to attain a certain goal (paying off his mortgage) *by* selling his securities.

Such a want is not inferrable from (10). It may be that the user wants to desolve his savings account at some time in the future, but the modal verb müssen (must) inside the question does not indicate a current want. Therefore only the relation between the two propositions is the focus of attention. Hence we can paraphrase the user's want as 'Do I have to pay a fee if I want to desolve my savings account?', or, again more formally,

(WANT USER (KNOW USER (IMPLIES P2 P1))),

where P2 denotes the desolving event and P1 the fee paying.

The Computational Model

The processes described in this paper work on a formal representation of utterances which reflects their semantic structure but also contains lexical and syntactic information (hedges, connectives, modal verbs, tense, and mood) which has not yet been interpreted. Our formal representation language is called IRS (Interne RepräsentationsSprache, [Bergmann et. al. 87]). It contains all the standard operators of predicate calculus, formalisms for expressing propositional attitudes, modalities, and speech acts, natural language connectives (*and, or, however, therefore,* etc.), a rich collection of natural language quantifiers (e.g., articles, wh-particles), and modal operators (*maybe, necessarily*).

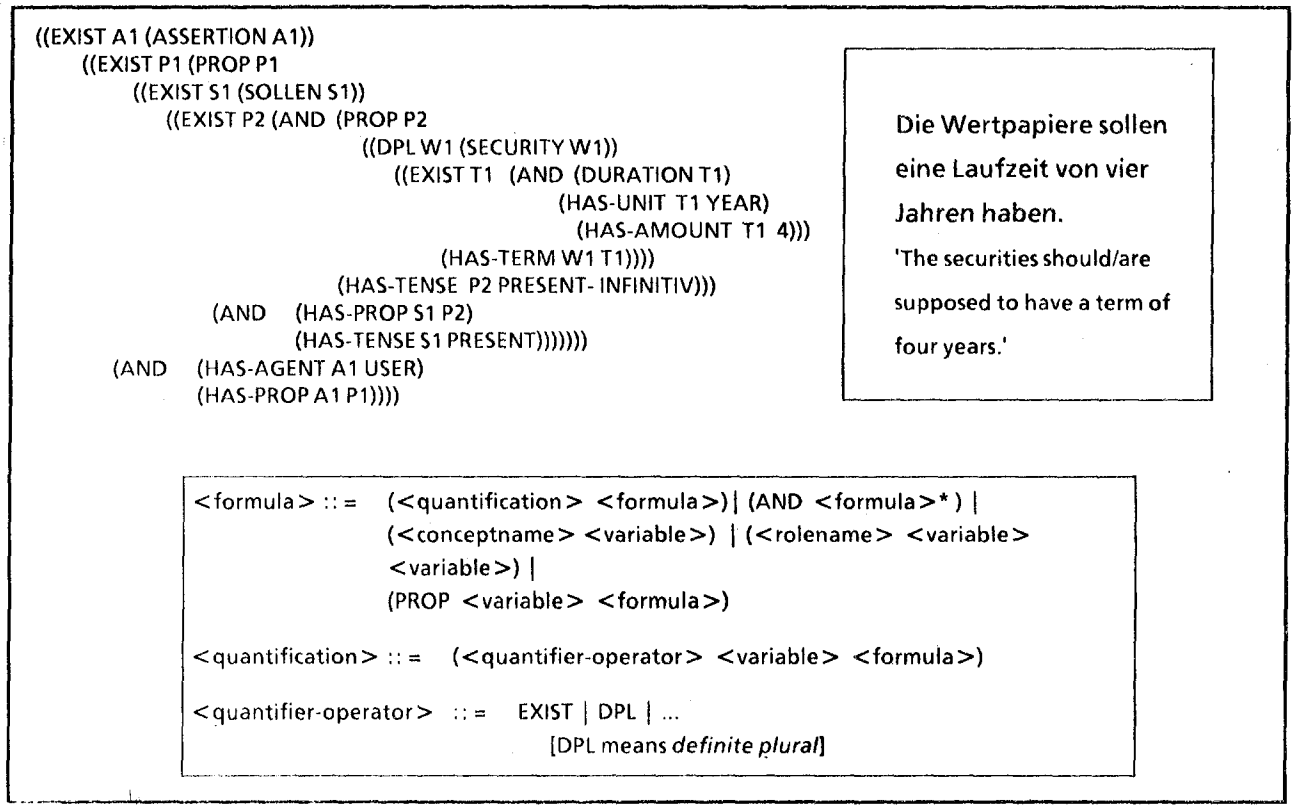

Fig. 3: An example of IRS and the corresponding part of the syntax of IRS

Fig. 3 shows a part of the syntax definition of IRS and the representation of the sentence

(6) Die Wertpapiere sollen eine Laufzeit von vier Jahren haben
'The securities should/are supposed to have a term of four years.'

This example contains some important features of IRS:

Only one- and two-place predicates are allowed. They correspond to the concepts and roles defined in our terminological knowledge base QUIRK [Bergmann/ Gerlach 86] except for SOLLEN and HAS-TENSE which still need to be semantically interpreted.

Quantifications are always restricted to a range which may be described by an arbitrary formula.

The operator PROP allows for associating a variable to a formula. In subsequent terms the variable may be used as a denotation of the proposition expressed by that formula .

In the formula given in Fig. 3 the variable A1 denotes the assertion as an action with agent USER and propositional content P1. S1 reflects the occurrence of the modal verb *sollen* which is represented like a predicate, but has not yet been semantically interpreted. The "propositional content" of S1 is P2 which denotes the proposition *the securities have a term of four years.*

For characterizing sets of structures to which one specific interpretation may apply, we use IRS patterns[Gerlach 87], i.e., highly parameterized semantic structures which specify an arbitrary combination of features relevant to the interpretation process: The surface speech act, tense information, modal hedges, and restrictions on the propositional content.

A quite simple example for an IRS pattern is given in Fig. 4. Its elements are

- variables (symbols starting with '?'),
- constants (all other symbols),
- a concept pattern (matching any one-place predication),
- role patterns (matching two-place predications).

```
(AND (?INFO-TRANS-TYPE ?INFO-TRANS)
     (HAT-SOURCE ?INFO-TRANS USER)
     (HAT-GOAL ?INFO-TRANS SYS)
     (HAT-OBJECT ?INFO-TRANS ?OBJECT))
```

Fig. 4: An IRS pattern

This pattern is used for matching the top level of the representation of an utterance of the user, directed to the system. When matching the variable ?OBJECT is bound to the whole propositional content of the utterance and is used by the subsequent steps of analysis.

As described above, we do not only infer new user model information directly, but also perform transformations on IRS structures, e.g., to reduce idioms to more primitive speech acts. This kind of processing involves applying a set of transformational rules to an IRS formula where a rule is a pair of IRS patterns as described above (for an example, see Fig. 2). When instantiating the right hand side of the rule the interpreter will create new variables for unbound pattern variables and quantify them in the appropriate way (in Fig. 2 this is the case with the pattern variable ?Q).

In WISBER the user model is a section of the central assertional knowledge base (A-Box, [Poesio 88]) which allows for storing and retrieving assertional knowledge in different contexts which denote the content of propositional attitudes of agents . Hence a new entry is added to the user model by storing the propositional content in the A-Box context which contains the user's wants.

Conclusion

We have implemented our interpretation module in an Interlisp programming environment. It is a part of the natural language consultation system WISBER. The module's coverage includes all German modal verbs occuring in assertions and questions, some connectives (e.g., *and, so that, because*) and the most common indirect questions. On the one hand our future work will concentrate on extending the performance of the system inside the framework which is described in this paper. On the other hand we will integrate the concept of *expectations*, i.e. expectations the system has according to the users next utterance depending on the actual state of the dialog. This will enable us to resolve more kinds of ambiguities in user utterances.

References

Allen 83:
J. F. Allen: *Recognizing Intentions from Natural Language Utterances*, in: M. Brady and R. C. Berwick (Ed.): Computational Models of Discourse, MIT Press 1983, pp. 107-166

Appelt 85:
D. E. Appelt: *Planning English Sentences*, Cambridge University Press, 1985

Bergmann et.al. 87:
H. Bergmann/ M. Fliegner /M. Gerlach/ H. Marburger/ M. Poesio: *IRS - The Internal Representation Language*, WISBER Bericht Nr. 15, Universität Hamburg, Fachbereich Informatik, November 1987

Bergmann/Gerlach 87:
H. Bergmann/ M. Gerlach: *Semantisch-pragmatische Verarbeitung von Äußerungen im natürlichsprachlichen Beratungssystem WISBER*, in: W. Brauer, W. Wahlster (Eds.): Wissensbasierte Systeme - GI-Kongress 1987. Springer Verlag, Berlin 1987, pp. 318-327

Bergmann/Gerlach 86:
H. Bergmann/ M. Gerlach: *QUIRK - Implementierung einer TBox zur Repräsentation begrifflichen Wissens.* WISBER Memo, Universität Hamburg, Fachbereich Informatik, Dezember 1986

Boyd/Thorne 69:
J. Boyd / J. P. Thorne: *The Semantics of Modal Verbs*, in: Journal of Linguistics 5, 1969, pp. 57 - 74

Brée/Smit 86:
D. S. Brée / R. A. Smit: *Linking Propositions*, in: Proceedings of COLING-86, Bonn 1986, pp. 177 - 180

Brünner/Redder 83:
G. Brünner / A. Redder: *Studien zur Verwendung der Modalverben*, Tübingen 1983

Cohen 84:
R. Cohen: *A Computational Theory of the Function of Clue Words in Argument Understanding*, in: Proceedings of COLING-84, Stanford 1984, pp. 251 - 258

Ellman 83:
J. Ellman: *An Indirect Approach to Types of Speech Acts*, in: Proc. of the 8th IJCAI, Karlsruhe 1983, pp. 600-602

Gerlach 87:
M. Gerlach: *BNF - a Tool for Processing Formally Defined Syntactic Structures.* Universität Hamburg, Fachbereich Informatik, WISBER-Memo Nr. 17, Dezember 1987

Naito et. al. 85:
S. Naito / A Shimazu / H. Nomura: *Classification of Modality Function and its Application to Japanese Language Analysis*, in: Proceedings of the 23rd Annual Meeting of the ACL, Chicago 1985, pp. 27-34

Poesio 88:
M. Poesio: *Dialog-Oriented A-Boxing*, Universität Hamburg, Fachbereich Informatik, WISBER Bericht, to appear.

Reinwein 77:
J. Reinwein: *Modalverb-Syntax*, Tübingen 1977

Sprenger 88:
M. Sprenger: *Interpretation von Modalverben zur Konstruktion von Partnermodelleinträgen*, Universität Hamburg, Fachbereich Informatik, WISBER Memo Nr. 18, Januar 1988

Wilensky et. al. 84:
R. Wilensky, Y. Arens, D. Chin: *Talking to UNIX in English: an Overview of UC*, in: Communications of the ACM 27(6), pp. 574-593 (June 1984)

Wilks/Bien 83:
Y. Wilks, J. Bien: *Beliefs, Points of View, and Multiple Environments*, in: Cognitive Science 7, pp. 95-119 (1983)

Zernik/Dyer 85:
U. Zernik, M.G. Dyer: *Towards a Self-Extending Lexicon*, in: Proceedings of the 23rd Annual Meeting of the ACL, Chicago 1985, pp. 284-292