Proceedings of VEXTAL'99 - Venezia, San Servolo, V.I.U. - 22-24 November 1999

A Computational Approach to the Semantics of Function Words in Dialogue Context

B. Ludwig, G. GÖrz, H. Niemann
Bavarian Research Centre for Knowledge Based Systems
Am Weichselgarten 7
D-91058 Erlangen
Phone: +49 9131 691451
Fax: +49 9131 691185
email: fludwig,goerz,niemanng@forwiss.de

Abstract

The present paper discusses the importance of function words for understanding dialogues. In a spoken language dialogue system (SLDS), they serve for reconstructing assumptions of the user about the domain. The paper gives examples providing evidence for that claim, presents an approach to parsing function words, and shows how their semantics can be integrated into a DRT-based semantic formalism. On that basis, the paper outlines how function words influence the assignment of speech acts to utterances, and in this way reflect the user's mental state. Finally, implications and consequences of the approach to the theory and development of dialogue systems and their applications are discussed briefly.

1 Introduction

In spoken language dialogue there exists a variety of ungrammatical phenomena like repairs, interruptions, or syntactically ill-formed sentences caused on the one hand by the speaker himself or on the other hand by the speech recognizer due to a "misunderstanding" of the speech signal. Therefore, any dialogue management system has to cope with the problem of reconstructing as precisely as possible the utterance represented by the speech signal. For this purpose, a set of hypotheses coded in a word lattice is generated.

A successful and promising approach is that of extracting parts of the utterance that represent information relevant to concepts defined in the domain of the dialogue manager's application. This information then has to be incorporated into the current dialogue context ([Fis98], [Nag98]).

While extracting relevant information from word lattices produced by the speech recognizer, systems make—up to different degree—intensive use of the linguistic knowledge about the syntax of the (natural language) to be processed. As shown e.g. by [Nag98] exploiting this knowledge is an important cornerstone for understanding natural language. Additionally, linguists and philosophers of language ([Eng88], [Eco93]) explain how natural language produces semantic content using syntactic form. Therefore, analyzing syntax is a major key to the interpretation of meaning.

Besides mentioning problems that from the viewpoint of computational linguistics are strongly related to parsing, it should be noticed that the interpretation of utterances in a given discourse should be based on a multi-level discourse model as various researchers have shown (e.g. [Hee98], [Tra94], [TraHin91]). This means that the semantics of different parts of an utterance is to be defined by various elements or processes of the discourse model. On the other hand, as noted e.g. by [StrJön98] or [LudGörNie98], a dialogue model combines notions defined in the knowledge base of the application of the system and other notions defined in the application-independent knowledge. Consequently some parts of an utterance have to be interpreted with respect to the application-dependent knowledge, whereas other parts can be understood only with respect to the application-independent notions of the dialogue model.

In this context, function words do not describe content relevant in any application, but independently from a specific domain declare (logical) relations between utterances (not only of the same speaker) made in the discourse. From this point of view, they express explicitly assumptions of the speaker about the domain. Based on this observation, the aim of the present paper is to show how the analysis of function words and their incorporation into the processing of utterances in a dialogue model serves to reconstruct the user's assumptions. We show that function words help in reasoning about propositions stated by the dialogue participants and about the rules the speaker considers to be true and therefore uses in generating his utterance. The overall aim of this analysis is to make the dialogue model able to explain the user's communicative behaviour and in this way to strengthen its ability to be cooperative. Finally, we present our approach to integrate the analysis of function words into our existing dialogue model and system.

2 Phrase Structures in Natural Language

As this paper does not develop a theory about language understanding in general, we restrict the presentation here to a small number of syntactic constructs found frequently in (corpora of) information retrieval dialogues.

The two main phrase structure types are:

- verb phrase (the verb introduces an action or relation into the current discourse)
- noun phrase (the noun introduces an object, i.e. a discourse referent)

Phrase structures of different types (e.g. prepositional phrase, adverbial phrase, adjectival phrase, sub clauses) can modify the two basic structures mentioned above.

In addition to these types of phrase structures, function words are very common in spoken language. Nevertheless, in current implementations of spoken language dialogue systems (SLDS) function words usually are not taken into consideration for two reasons: One the one hand, they do not carry content that is relevant for the application. Therefore they cannot be mapped onto values for parameters serving as necessary input for the pragmatic component of the SLDS (e.g. a database). On the other hand, their recognition quality is poor because they are very short. This situation can be improved by defining and training specialized language models for the speech recognizer.

In the following we will show how function words are processed during parsing and composition of corresponding semantic descriptions in the framework of processing different phrases types containing application relevant information.

3 Constructing Semantic Representations from Phrase Structures

3.1 Parsing Phrase Structures in a DRT-framework

Our approach to parsing is guided by [Jac77]'s $\overline{\text{X}}$- and [Abn95]'s chunk-theories.

The most prominent and important structures are noun phrases and verb phrases. They introduce new referents for objects and actions, respectively. According to grammatical rules, minimal structures can be extended or modified by phrases of other types. Complete sentences are headed by a verb phrase extended at least by a noun phrase functioning as the subject of the sentence.

For example, a noun phrase can be analyzed in the following way (simplified):

$$\mathrm{NP} \longrightarrow \mathrm{DET}\,\mathrm{ADJP}^*\,\mathrm{N}\,\mathrm{PP}^*$$

According to composition rules describing the semantic dependency of the sub phrases, during the syntactic analysis, a representation based on Discourse Representation Theory (DRT) as developed by [KamRey93] is generated for the entire NP (out of the representations for the sub phrases). It contains a discourse referent for the noun and a number of assertions about this referent. We use the usual notation for Discourse Representation Structures (DRS):

$$\left[\frac{\text{set of referents}}{\text{set of assertions}}\right]$$

Variables for referents to be introduced later by different sub phrases and for DRS of sub phrases are notated as λ-variables. DRS composed by substituting λ-variables by the DRS of the appropriate sub phrase (see the example below).

The relation of a sequence of phrases matching the pattern above to a given domain model is the following: It is seen as a valid or plausible NP only if the resulting semantic representation is defined in the domain model. This condition combines the syntactic expression (or term of the natural language) with a notion (concept definition) of the domain model. In this way, a meaning is assigned to the phrase structure. In our approach, the domain model is represented using a Description Logic (DL). In order to prove whether a semantic representation is defined in the domain model, we check its validity utilising DL reasoning mechanisms (subsumption and consistency check).

By imposing this constraint we can make use of pragmatic information while parsing. In this way we avoid constructing readings of the input that are syntactically correct, but otherwise implausible. Empirically we have found that this approach also speeds up parsing as the number of edges constructed by the chart parser is rather small compared to a purely syntactic approach to parsing.

To give an example, the sequence

the fast red car

composed of

$$\lambda\, H.\left[\frac{\mathrm{n}}{\emptyset}\right] + H(\mathrm{n}),$$

$$\lambda\, x, H.\left[\frac{\emptyset}{\text{has-velocity}(x, \text{fast})}\right] + H(x),$$

$$\lambda\, x, H.\left[\frac{\emptyset}{\text{has-color}(x, \text{red})}\right] + H(x),$$

$$\lambda\, x.\left[\frac{\emptyset}{\text{car}(x)}\right]$$

is parsed as a possible hypothesis in the word lattice (in our example it is the only hypothesis, because the lattice consists of one single path) only if its semantic representation

$$\left[\begin{array}{l}\mathrm{n}\\ \hline \text{has-velocity}(\mathrm{n}, \text{fast})\\ \text{has-color}(\mathrm{n}, \text{red})\\ \text{car}(\mathrm{n})\end{array}\right]$$

is valid in the domain model. This requires the concept

$$\exists\text{has-velocity.fast} \sqcap \exists\text{has-color.red} \sqcap \text{car}$$

to be derivable in the domain model.

In an analogous way, DRT representations for actions formulated as verb phrases are constructed during the parsing process.

So, in our view, attaching phrases is not only a syntactic, but mainly a domain pragmatic problem. Therefore, our parser combines phrases whose combination is correct from a syntactic point of view only if the resulting semantic representation is consistent with the represented knowledge about the domain and the beliefs of the dialogue participants. From this perspective, parsing of word lattices is a process that should not be considered in isolation from reasoning about the content of the appropriate dialogue structures as sketched above.

On the other hand, it should be noted that by defining semantics in the domain model and therefore keeping it separate from syntax, we obtain independency from the application as far as the grammar rules are concerned. This is due to the fact that phrase structures and syntactic phenomena like agreement occur in the same form in any domain. To generate semantic representations of word sequences, a domain model is a necessary prerequisite as it declares (by using a domain independent inference procedure) which syntactically correct sequences have a well-defined meaning in the domain model.

3.2 Function Words

According to [Kno96], speakers use function words to express coherence between a set of preconditions and the currently uttered proposition. Coherence is based on a defeasible implication rule describing the speaker's assumption about how the arguments of the function words cohere logically. A defeasible implication is an implication that allows exceptions to be valid. A function word defines whether the two propositions in its scope are consistent with the implication or not for the special case of the given utterance. The function word is chosen according to the intended state of consistency. In our dialogue model, we use FIL (see [Abd95]) to formulate and reason about the implications:

$$*(\forall x : \alpha(x) \rightarrow \beta(x), \forall x : \alpha(x) \rightarrow \beta(x))$$

says that $\alpha(x) \rightarrow \beta(x)$ in general holds for any x, but there can be some discourse referents which violate the assumption.

To illustrate this point more in detail, we will discuss a simple example:

I have got a lot of money. So I will buy a Ferrari.
I have got a lot of money. But I will not buy a Ferrari.

In the first sentence, *so* indicates that the speaker assumes *I will buy a Ferrari* to follow from *I have a lot of money.* To express this relationship he uses *so* thereby signalling a material consequence.

Obviously, the speaker makes the same assumption in the second sentence as in the first one. In contrast to the first sentence however, he does not intend the consequence *I will buy a Ferrari* to be true. This intention is expressed by stating *I will not buy a Ferrari.* From this point of view, the second sentence is a contradiction to the default assumption the speaker states to be valid in the general case. Nevertheless he wants to make an exception for the individuals involved in the sentence. To reach this aim, he must achieve to make the default assumption fail. This effect can be obtained by using *but* which means: the consequence of the default assumption is intended to fail.

Formulating the default assumption in FIL, one would write:

$$[\forall x\,\mathrm{have}(x,\mathrm{y}) \wedge \mathrm{money}(\mathrm{y}) \longrightarrow \mathrm{buy}(x,\mathrm{f}) \wedge \mathrm{Ferrari}(\mathrm{f})]$$

where $[\phi]$ is used as an abbreviation for $*(\phi,\phi)$.

Using this FIL construct, we have obtained a simple possibility to express the defeasible character of the function word's semantics as proposed by Knott. Additionally, the verification whether the assumption holds or not in the current dialogue can be delayed until the utterance is integrated into the dialogue context (see the following section).

[Kno96] also discusses a number of additional semantic properties for function words which in particular serve for describing the relation of an utterance to the discourse (i.e. to what has been said so far) and to the assumptions the speaker makes about the domain. For a detailed discussion of these properties, called features by Knott, the reader is referred to the literature.

The point to be discussed here is how to integrate Knott's approach into the DRT framework sketched so far. To study this in detail it is worth revisiting what we have outlined above. As we have seen during the discussion of the two examples, a function word of the type conjunction (in the examples *so* and *but*) has got two so called *text spans* (i.e. the text in the scope of the function word) as arguments. These text spans contain propositions stated by the speaker as well as an implicitly stated assumption in the form of an implication (i.e. another proposition). Consequently, we can devise the following DRS scheme for *but*:

$$\lambda P, Q.\left[\begin{array}{l}\emptyset\\ \hline *(P \rightarrow \neg Q, P \rightarrow \neg Q)\end{array}\right] + P + Q$$

And the scheme for *so* is in analogy to *but*:

$$\lambda P, Q.\left[\begin{array}{l}\emptyset\\ \hline *(P \rightarrow Q, P \rightarrow Q)\end{array}\right] + P + Q$$

In the examples above, we obtain in the case for *so*:

$$P = \left[\begin{array}{l}\mathrm{s\,y}\\ \hline \mathrm{speaker(s)}\\ \mathrm{have(s,y)}\\ \mathrm{money(y)}\end{array}\right]$$

$$Q = \left[\begin{array}{l}\mathrm{s\,f}\\ \hline \mathrm{speaker(s)}\\ \mathrm{buy(s,f)}\\ \mathrm{Ferrari(f)}\end{array}\right]$$

After substituting the appropriate DRS for P and Q respectively, we obtain the following result (in the implication, the substitution is not performed due to lack of space):

$$\left[\begin{array}{l}\mathrm{s\,f\,y}\\ \hline *(P \rightarrow Q, P \rightarrow Q)\\ \mathrm{speaker(s)}\\ \mathrm{buy(s,f)}\\ \mathrm{Ferrari(f)}\\ \mathrm{have(s,y)}\\ \mathrm{money(y)}\end{array}\right]$$

In the case for *but* on the contrary, while P remains as before, Q looks as follows:

$$Q = \neg \left[\begin{array}{l} \mathrm{s\,f} \\ \hline \mathrm{speaker(s)} \\ \mathrm{buy(s,f)} \\ \mathrm{Ferrari(f)} \end{array} \right]$$

Substituting P and Q into the DRS for *but*, we obtain the same assumption as for *so* while the semantic representation for the second text span is modified as shown above.

The remaining question is: When is the substitution into the DRS for a function word performed during the parsing process? Extending the approach for parsing phrase structures with a standard chart parser, we require the implication and the propositions representing the text spans to be valid given the domain model and to be consistent with what has been uttered before. If this verification fails, the selected P and Q are not coherent with respect to the function word under consideration, and therefore no new edge can be introduced into the chart. Otherwise a newly constructed edge will represent the found interpretation hypothesis.

4 Integrating an Utterance into the Discourse

To show how the analysis of function words facilitates the processing of the whole utterance in a given discourse we will discuss an example (a fragment of a train information dialogue, see [Die87]):

> **AAA**: *Auskunft Hauptbahnhof*
>
> **AAA**: *Information Service Central Station. How can I help you?*
>
> **BBB**: *ja, guten Tag! Ich hätt' gern eine Zugverbindung nach Venedig, und zwar für Montagnachmittag oder Montagvormittag,* **dass** *ich so um halb vier in Venedig bin.*
>
> **BBB**: *Good morning! I would like to ask for a connection to Venice on Monday morning or Monday afternoon* **in order to** *arrive in Venice at half past three.*

Using Knott's terminology the function word **dass** (translated as **in order to**) is cause-driven and has positive polarity (see [Kno96]). Consequently, it means that the desire for a connection to Venice on Monday morning or afternoon is a precondition for arriving there half past three the same day.

To see the importance of correctly interpreting **dass** in the given dialogue situation, it has to be noticed that "Monday morning or afternoon" as well as "half past three" are temporal expressions. When assuming for the second expression "today" as a default for the missing day, the second expression could function as "overwriting" the first one if one does not consider **dass** correctly. Using the terminology of [Tra94], one could therefore see the second expression as a *repair* of the first one whereas due to the semantics of **dass** it is intended to be a *continue* act.

Consequently, if one uses an algorithm for extracting meaning from word lattices that simply spots domain relevant concepts without considering discourse relevant function words, the word lattice can be misinterpreted easily leading to a wrong and confusing processing of the utterance.

Other than only constructing the correct semantic representation for an utterance using Traum's model of grounding as shown above, the description of the semantics for **dass** in the way [Kno96] would do it, one can take advantage of the implication stated explicitly in the semantics of the function word: The domain reasoner can try to validate the implication in the domain model representing the system's knowledge about the domain. If it can be verified, the utterance can be acknowledged and an appropriate answer can be given. If not, grounding is impossible and the system must indicate via a *reqrepair* act that it cannot process the utterance or propose an alternative solution by generating a *repair* act.

Grounding of utterances is supported by function words in still another way: By checking whether the implication introduced by a function word is consistent with what the user has uttered before in the dialogue, the dialogue manager is able to detect inconsistencies in the user belief modality. In this case, there are two possibilities. First, the system could try to find another path through the word lattice resulting in a different interpretation of the lattice consistent with the content of the belief modality. Otherwise, the system can refuse grounding of the utterance again by generating a *reqrepair* or *repair* act.

This last point leads us directly to the discussion of further consequences of the semantics for function words: They give a possibility to reconstruct some assumptions of the user about propositions valid in the domain model.

5 Modifying the User Model

Despite the fact that existing generic dialogue managers do not maintain an explicit user model, there are numerous applications for SLDS having user models and mental images as a necessary prerequisite. In [StoBlo97] it is demonstrated how listener models holding a representation of a visual scenario described by a sequence of utterances serve for verifying the plausibility and correctness of an utterance and for resolving references of relative expressions like "left of" or "near" used by the speaker in his description of a scene.

In [TR633] the dialogue manager must be able to maintain a model of how the user views the state of affairs in a complex transportation optimization problem.

As explained in [ArdCoh96] and [ArdBoeDam], user models in dialogue systems have the role to provide a basis for describing the dialogue participants as autonomous agents and—on the basis of this observation—deliver the necessary facts for reasoning about their rational behaviour in communication. This reflects the fact that dialogue participants try to execute their own plans during conversation in order to reach their pragmatically founded goals. For doing so, they use their own domain knowledge. This should be known as precisely as possible to the dialogue manager in order to be cooperative.

In order to study how the analysis of function words supports the reconstruction of the user's assumptions about the domain, we have to look carefully at the semantics of function words: The important point to notice is that a function word (implicitly) introduces an implication into the current DRS that defeasibly expresses an general (default) rule possibly violated by the discourse referents in the current utterance.

This rule is part of the user's domain model. Therefore, by stating it explicitly, one can extend the content of the user's belief modality. Using the content of this modality, the system is capable of understanding and explaining the user's way of thinking about the domain.

There are at least two important advantages which demonstrate why it is worth advocating approaches to dialogue systems that are able to handle complex user models:

1. It is easier for the dialogue system to be cooperative. In a mixed-initiative dialogue system that does not restrict user utterances to be responses to very

precisely stated questions, but allows the user to speak in a quite free manner, it is a very important point to detect the topic of the user's interest in his utterances. Knowing the user's intentions more precisely, the system can react adequately and cooperatively which is highly relevant for dialogue success rates and average transaction lengths. These two factors are the most important ones for large-scale "real time" applications of dialogue systems.

2. The dialogue system can permanently test the user's domain model against its own. Doing this, in the case of inconsistency between the two domain models, the system is able to explain immediately wrong assumptions to the user—a communicative behaviour augmenting the transparency of the system responses. This is the only way for making the system behaviour plausible to the user, particularly when speech recognition errors cause the user's expectations not to be met.

As [ArdBoeDam] show, keeping an extensive user model is necessary for tracking the coherence of the user's plans and actions formulated in his utterances and for describing and solving misunderstandings due to failed coherence between the user's plans and the actions as described in a domain model.

6 Summary and Outlook

The present paper outlines an approach to integrate the analysis of function words into an existing dialogue model. It shows how function words are parsed and how their semantics can be incorporated into a DRT-framework. On that basis it is shown how such an extended semantic representation of an utterance supports the process of grounding and provides a foundation for user models that can be updated and modified by the content of the user's utterances.

The combination of grounding utterances and user models is a very powerful approach to dialogue understanding because it has a number of very far reaching consequences for the study of cooperative dialogues.

- Many very interesting applications that rely heavily on information about the user can be realized in the framework of a dialogue system. Among the areas of interest are medical care (e.g. help systems for disabled persons), teaching (e.g. interactive software for language learning), or instruction (e.g. online help facilities for software, telecommunication services, home electronics). They all have in common that in order to function properly and effectively they have to perform an evaluation of the user's behaviour. The aim of such an evaluation is to employ a dialogue strategy adaptive to the user's capabilities, his level of experience, and to several resource limitations (as studied e.g. in [Wei97]).

- Another interesting aspect is that the integration of utterances into a dialogue context can be used to verify hypotheses about the interpretation of a word lattice by tracking whether the user accepts a certain hypothesis or not (grounding of utterances). Such a verification can be used to learn the (possible) meaning of results produced by the syntactical analysis in a example-based and interactive way.

For the realization of these two ideas function words play a key role because the provide the necessary data as has been sketched above. Interestingly enough, Knott's approach shows to be a powerful tool for improving the analysis of discourse in dialogue systems and to enhance the capabilities of such systems.

References

[Abd95] **N. Abdallah**, *The Logic of Partial Information*, Springer, New York 1995

[Abn95] **S. Abney**, *Chunks and Dependencies: Bringing Processing Evidence to Bear on Syntax*, in: *Computational Linguistics and the Foundations of Linguistic Theory*, CSLI 1995

[ArdCoh96] **L. Ardissono** and **R. Cohen**, *Extending the Role of User Feedback in Plan Recognition and Response Generation for Advice-giving Systems: an Initial Report*, in: LNAI 1081, 1996

[ArdBoeDam] **L. Ardissono**, **G. Boella** and **C. Damiano**, *A plan-based model of misunderstandings in cooperative dialogue*, in: *International Journal of Human-Computer Science*, special issue on *Detecting, Repairing and Preventing Human-Machine Miscommunication*, to appear

[Die87] **G. Diewald**, *Telefondialog: Beschreibungsversuche zwischen Gesprächsanalyse und Textlinguistik*, Zulassungsarbeit, Universität Erlangen 1987 (in German)

[Eco93] **U. Eco**, *La ricerca della lingua perfetta nella cultura europea*, Laterza, Rom, Bari 1993

[Eng88] **U. Engel**, *Deutsche Grammatik*, Groos, Heidelberg 1988 (in German)

[Fis98] **J. Fischer** et al, *Empowering Knowledge Based Speech Understanding through Statistics*, in: *Proceedings of ICSLP 98*, Sydney, Australia 1998

[Hee98] **P. Heeman** et al, *Beyond Structured Dialogues: Factoring out Grounding*, in: *Proceedings of ICSLP 98*, Sydney, Australia 1998

[Jac77] **R. Jackendoff**, $\overline{X}$ *Syntax: A Study of Phrase Structure*, MIT Press 1977

[KamRey93] **H. Kamp** and **U. Reyle**, *From Discourse to Logic*, Kluwer, Dordrecht 1993

[Kno96] **A. Knott**, *A Data-Driven Methodology for Motivating a Set of Coherence Relations*, PhD thesis, University of Edinburgh 1996

[LudGörNie98] **B. Ludwig**, **G. Görz**, and **H. Niemann**, *Modelling Users, Intentions, and Structure in Spoken Dialogue*, in: *Proceedings of Konvens 98*, Bonn, Germany 1998

[Nag98] **A. Nagai** and **Y. Ishikawa**, *Concept-Driven Speech Understanding Incorporated with a Statistical Language Model*, in: *Proceedings of ICSLP 98*, Sydney, Australia 1998

[StoBlo97] **Stopp** and **Blocher**, *Construction of Mental Images and their Use in a Listener Model*, in: *Proceedings of the Symposium on Conceptual and Semantic Knowledge in Language Generation*, Heidelberg, Germany, 1997, pp. 270–280

[StrJön98] **Strömbeck** and **Jönsson**, *Robust Interpretation for Spoken Dialogue Systems*, in: *Proceedings of ICSLP 98*, Sydney, Australia 1998

[TraHin91] **D. Traum, E. Hinkelman**, *Conversation Acts in Task-Oriented Spoken Dialogue*, in: *Computational Intelligence*, 8(3):575–599, 1992

[Tra94] **D. Traum**, *A Computational Theory of Grounding in Natural Language Conversation*, Ph.D. Thesis, Computer Science Dept., University of Rochester 1994

[TR633] **D. Traum et al.**, *Knowledge Representation in the TRAINS-93 Conversation System*, Technical Report 663, University of Rochester, 1996

[Wei97] **S. Weis**, *Resource-Adaptive Action Planning in a Dialogue System for Repair Support*, in: *Proceedings of KI-97*, Freiburg, Germany 1997