gefördert vom
Bundesministerium
für Forschung
und Technologie

Verbmobil
Verbundvorhaben

Context Construction as Subtask of Dialogue Processing – the VERBMOBIL Case

Elisabeth Maier

DFKI GmbH

Report 134
Juni 1996

Juni 1996

Elisabeth Maier

Deutsches Forschungszentrum für Künstliche Intelligenz GMBH
Stuhlsatzenhausweg 3
D-66123 Saarbrücken

Tel.: (0681) 302 - 5347
e-mail: `maier@dfki.uni-sb.de`

Das diesem Bericht zugrundeliegende Forschungsvorhaben wurde mit Mitteln des Bundesministers für Forschung und Technologie unter dem Förderkennzeichen 01 IV 101 K/1 gefördert. Die Verantwortung für den Inhalt dieser Arbeit liegt bei dem Autor.

Context Construction as Subtask of Dialogue Processing – the VERBMOBIL Case

Elisabeth Maier *
DFKI GmbH
Stuhlsatzenhausweg 3
D-66123 Saarbrücken
Germany
maier@dfki.uni-sb.de

Abstract

This paper presents the dialogue component of the speech translation system VERBMOBIL. In particular, it describes the *Dialogue Memory* which has been developed to represent contextual information acquired during dialogue processing. Information is stored both chronologically, i.e. in the order of appearance (in the *Sequence Memory*) and conceptually (in the *Thematic Structure*). We show how the Dialogue Memory is used to draw contextual inferences, some of which serve as basis for the detection of indirect speech acts.

1 Introduction

This paper presents an approach for the representation of contextual information that was implemented as part of the face-to-face speech translation system VERBMOBIL. The content of the so-called *Dialogue Memory* in which contextual information about the ongoing dialogue is stored has been very much determined by the requirements of a number of system components, like e.g. speech recognition, semantic processing, transfer and generation (a more detailed discussion of this issue can be found in [11]). In VERBMOBIL, it is a task of the dialogue component to incrementally construct a representation of the context.

*This work was funded by the German Federal Ministry for Research and Technology (BMBF) in the framework of the VERBMOBIL Project under Grant 01IV101K/1. The responsibility for the contents of this study lies with the author. The author wishes to thank Jan Alexandersson, Norbert Reithinger, and Adelheit Stein for comments on earlier drafts of this paper, and Martin Klesen for his help in implementing the Dialogue Memory.

This paper is a reprint from the Proceedings of the Eleventh Twente Workshop on Language Technology, TWLT 11, Dialogue Management in Natural Language Systems, June 19-21, 1996

This paper first gives a brief presentation of the dialogue component and the subtasks it has to fulfill in VERBMOBIL (section 2). After a discussion of the two basic knowledge types that are represented in our context model – dialogue acts (section 3.1) and time information (section 3.2) – we show the two submodules of which the Dialogue Memory consists: *Sequence Memory* and *Thematic Structure* (section 3.3). To illustrate the incremental construction of the context representation in the Dialogue Memory we give an example for the processing of a dialogue fragment from our corpus of appointment scheduling dialogues (section 3.4). Finally we describe how contextual inferences are supported by our Dialogue Memory (section 4) and indicate how this approach can also be used to handle indirect speech acts (section 5).

2 Dialogue Processing in VERBMOBIL

It is a key issue of our project to establish robust processing methods that can cope with unreliable and incomplete input as it is typical for spoken language systems. One means to achieve this is the availability of contextual information. In VERBMOBIL this information is used e.g. to predict follow-up speech acts in dialogue processing, to disambiguate translational equivalents during transfer, to resolve anaphoric expressions in semantic evaluation and to control lexical variation in the generation of target language expressions. In our case, contextual information is provided by the dialogue component [2]. It is the task of this component to monitor the progress of the dialogue and to provide a representation of what has been said. On this basis the dialogue component is able to constrain decisions made by other system components and to predict follow-up dialogue states.

The dialogue component has been realized as a hybrid architecture: it contains statistical and knowledge-based methods. Both parts work with dialogue acts [3] as basic units of processing.

The statistics module is based on data automatically derived from a corpus of dialogues that have been manually labeled with dialogue acts. On the basis of this knowledge the statistics module determines possible follow-up dialogue acts for every utterance (see [14]). The plan recognizer as knowledge-based module of the dialogue component incorporates a dialogue model, which describes sequences of dialogue acts as occurring in appointment scheduling dialogues (see [1]).

3 Context Construction

As explained in [11] we took a rather "pragmatic" approach to the design of the Dialogue Memory: we decided to include only information types which are required by subcomponents of VERBMOBIL. Therefore we took the needs of speech recognition, syntactic-semantic processing, transfer and generation into account. In the following, we discuss the two most prominent information types included in our Dialogue Memory, namely dialogue act and temporal information. We explain how these information types are represented in two subcomponents of the Dialogue Memory, the Sequence Memory and the Thematic Structure. We conclude this section by providing an example for the construction of the context representation when processing a sample dialogue.

3.1 Dialogue Acts for Appointment Scheduling Dialogues

On the basis of an extensive corpus of appointment scheduling dialogues (currently about 1000 dialogues are available in transliterated form) we determined a set of 42 dialogue acts [8]. These dialogue acts form the leaves of a dialogue act hierarchy, where the more abstract levels are rather independent from the domain at hand. In figure 1 we show only the abstract dialogue act categories; domain-dependent dialogue acts concern e.g. dates, locations and durations. The hierarchical organization of the dialogue acts has a number of advantages, one of them is the applicability of acts to other domains and applications. We expect, for example, that the set of abstract dialogue acts is easily portable to other negotiation dialogues; first experiments with travel planning dialogues confirmed this assumption.

The hierarchy given in figure 1 shows that three phases can be distinguished in appointment scheduling dialogues: an *initialization* phase (e.g. GREET, INTRODUCE), a *negotiation* phase (e.g. SUGGEST, ACCEPT) and a *closing* phase (e.g. CONFIRM, THANK). While the initialization and the closing phase fulfill rather social functions, the negotiation phase is mainly task-oriented: dialogue acts belonging to this phase are used to advance the negotiation and to achieve the goal of the interaction. Another class of acts, the *deviations*, cannot be properly attributed to one dialogue phase: they can occur at any point of a dialogue. They also do not contribute to the task as such; they are used to describe actions like e.g. giving feedback, thinking aloud, etc.

In this paper we focus only on the dialogue acts of the negotiation phase; they either co-occur with propositional material relevant to the task, i.e. information about proposed dates and times, or they present an evaluation of proposals (i.e. acceptance or rejection).

3.2 Time Information in Appointment Scheduling Dialogues

In addition to dialogue acts another information type that is relevant for the advancement of appointment scheduling dialogues is temporal information. To represent temporal information in our Dialogue Memory we use an intuitive hierarchical model of temporal categories. Our categories are *date, year, month, week*, etc. down to *time* as most fine-grained type of temporal information. These categories are ordered hierarchically insofar as an instance of a temporal category is refined by an instance of a category of finer granularity, i.e. an instance of type year can be specified further by adding an instance of type

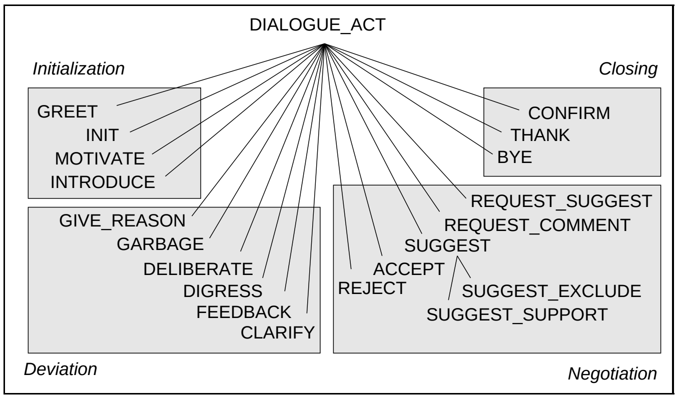

Figure 1: Domain-independent dialogue acts in VERBMOBIL

month. In this case we say that the instance of type year is *superordinated* to the instance of type month. For every temporal information mentioned during the conversation one or more instances of the appropriate categories are created and embedded into the temporal structure created so far.

3.3 Sequence Memory and Thematic Structure

In the Dialogue Memory of the VERBMOBIL system two subcomponents have been developed for the representation of context: the *sequence memory*, which mirrors the sequential order in which the utterances and the related dialogue acts occur, and the *thematic structure*[1], which consists of instances of temporal categories and their status in the dialogue, i.e. who they have been proposed by and whether they have already been accepted or rejected. Both components are closely intertwined so that for every utterance of the dialogue the available information can be easily accessed.

During dialogue processing contextual information is constructed as follows (see figure 2):

- Updating the Sequence Memory
  - *Determination of Dialogue Act*
    the dialogue act included in an utterance is determined either through shallow or deep processing; in the latter case the semantic evaluation component is responsible for the computation of the dialogue act. Depending on the quality of the translation either the results of shallow processing or of semantic evaluation are incorporated into the

[1]The term *thematic structure* is unrelated to the same term as introduced in the Prague School of Linguistics. The term rather refers to what the dialogue is about, i.e. its propositional content.

Dialogue Memory. While shallow processing computes dialogue acts mostly on the basis of key words the semantic evaluation component additionally uses sentence mood and contextual information. After dialogue act determination the Sequence Memory is updated by an object which includes the dialogue act together with the identifier of the utterance. This new object is then linked to the object representing the previous utterance.

  - *Computation of Predictions*
    on the basis of the dialogue act the statistical component computes the most likely dialogue acts for the following utterance; this information is then entered into the representation of the current Sequence Memory object;
  - *Plan Recognition*
    using the dialogue act the plan recognition component determines how the current utterance fits into the expected course of an appointment negotiation dialogue. It also determines which dialogue phase the utterance belongs to; this information is then added to the Sequence Memory.

- Updating the Thematic Structure
  - *Parsing of Time Expressions*
    using the information supplied by the parser the semantic evaluation component retrieves the temporal information included in an utterance[2] and maps it into an expression of a time description language which was developed specifically for the purposes of VERBMOBIL. This language (for details see [9]) is designed to model temporal knowledge that can be directly entered into the Thematic Structure (*absolute* temporal information), and temporal information that requires additional inferences in order to be added to the Thematic Structure (*relative* temporal information)[3]. Some of the relative expressions, in particular those expressions representing temporal references (e.g. *then, that day*), are resolved by the semantic evaluation component on the basis of focus information provided by the Thematic Structure. The majority of the relative expressions, though, requires additional processing in order to be expressible in terms of the Thematic Structure.
  - *Computation of Absolute Time Expressions*
    relative expressions are, for instance, referring expressions (e.g. *last Sunday, next week*), temporal modifiers (e.g. *early June, late afternoon*), public holidays (e.g. *Christmas, Good Friday*) and combinations of these categories (e.g. *Thursday in three weeks, the day after*

[2] For the time being shallow methods for the extraction of time processing are not integrated in our prototype. It is foressen for later stages of the project, though.

[3] A similar distinction for temporal expressions can be found in [12].

Easter); for an outline of this process see [7]). Some of these computations, in particular the resolution of referring expressions, are made on the basis of a so-called *reference point* or *focus* which is the time point under consideration at the current stage of the dialogue. The reference point is updated as the dialogue proceeds and the participants' attention moves on to other time points.

- *Embedding in Thematic Structure*
  with the temporal information once mapped onto the appropriate categories the Thematic Structure is either updated with newly created instances in case the time frames have not been mentioned yet or the focus is shifted back to objects already mentioned and therefore available in the Thematic Structure.
- *Inferencing over Thematic Structure*
  under conditions that will be discussed in section 4 the entry of a new temporal instance into the Thematic Structure induces a number of follow-up actions, like e.g. the propagation of dialogue act or evaluation information to subordinated or superordinated instances.

3.4 A Running Example

In this subsection we show how a part of a sample dialogue is processed and subsequently represented in the Dialogue Memory. The dialogue together with the dialogue acts and time expressions for the individual utterances, i.e. the output of Semantic Evaluation, are given in figure 3.

While processing the dialogue the system creates an object for every turn. i.e. for every contribution made by one speaker and adds it to the Sequence Memory. The information concerning the individual utterances is then linked to this turn, mirroring their order of occurrence (see figure 4 for a snapshot of the Sequence Memory after processing the sample dialogue).

In the Thematic Structure the first proposal made by dialogue partner A leads to the creation of time objects for the month, day, weekday and the period-of-day. These objects also represent information concerning the speaker who made the suggestion. In the next utterance the proposal is rejected by speaker B. For doing so, an anaphoric expression is used. Since anaphoric expressions by default refer to the temporal objects which are currently in focus, i.e. which have been mentioned in the previous utterance, the dialogue act of the current utterance refers to those objects. Because for anaphoric expressions rejections only refer to the most specific object in focus (rejecting a Thursday morning as time for a meeting does not necessarily mean that all of Thursday is rejected) only the object concerning the period of day is updated with the corresponding information. The full Thematic Structure for the dialogue segment is given in figure 5.

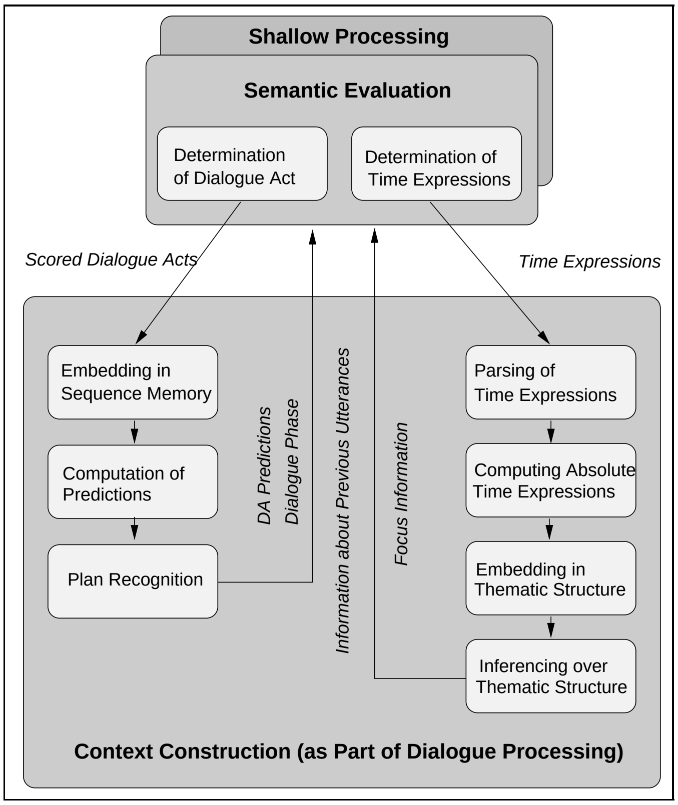

Figure 2: The process of context construction.

SUGGEST_SUPPORT_DATE AB

`[month:may,day:26,day_of_week:thu,period_of_day:morning]`

A000: ja , das erste Treffen würd' ich gerne im Mai mit Ihnen machen und da explizit am, <äh> ja, am sechsundzwanzigsten Mai, Donnerstag vormittags.

yes, the first meeting I'd like to make with you in May, and then explicitly on the twentysixth Thursday morning.

REJECT_DATE BA

`[ana]`

B001: <äh> das paßt mir nicht gut.

that doesn't suit me.

REJECT_DATE BA

`[after([day:25])]`

ab sechsundzwanzigsten <hm> kann ich nicht.

I can't starting with the 26th.

SUGGEST_SUPPORT_DATE BA

`[before(1*week,ana)]`

vielleicht die Woche davor,

perhaps the week before (that),

REQUEST_COMMENT_DATE BA

`[ana]`

paßt Ihnen da irgend etwas?

does anything suit you then?

SUGGEST_SUPPORT_DATE AB

`[day_of_week:tue,day:17]`

A002: ja. da können wir am Dienstag, den siebzehnten, vielleicht ?

yes. Then we could on Tuesday the seventeenth, perhaps?

ACCEPT_DATE BA

`[ana]`

B003: das geht in Ordnung für mich.

that's ok with me.

SUGGEST_SUPPORT_DATE AB

`[day_of_week:tue,day:17,period_of_day:afternoon]`

A004: da machen wir's doch Dienstag, siebzehnten, am Nachmittag.

then let's make it Tuesday, seventeenth, in the afternoon.

Figure 3: A sample dialogue

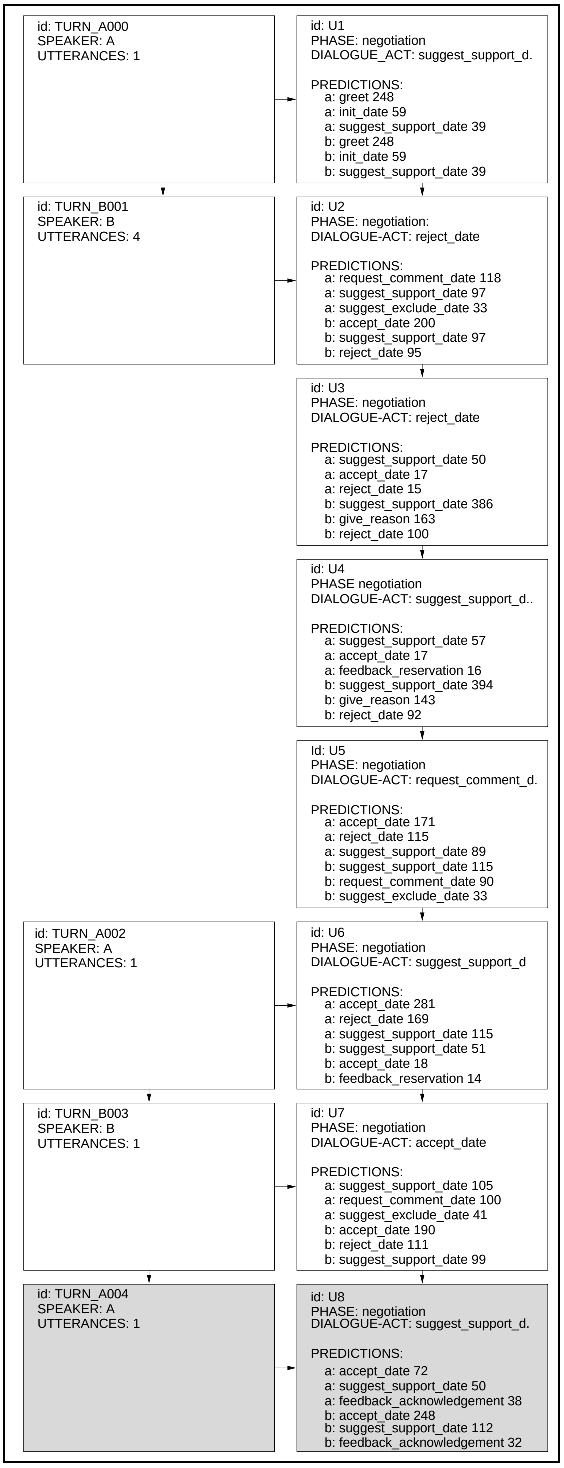

Figure 4: The Sequence Memory after processing the sample dialogue.

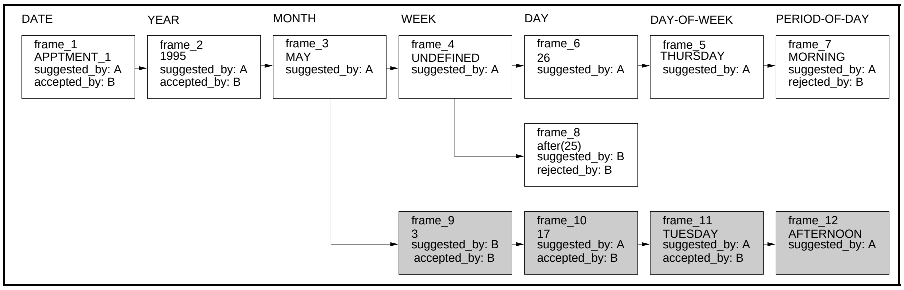

Figure 5: The Thematic Structure after processing the sample dialogue.

4 Contextual Inferences

In this section we have a closer look at the construction of the Thematic Structure: it can be observed that dialogue acts which concern a specific time object, i.e. the rejection or acceptance of such an object, can have consequences for other related time objects: they can be implicitly accepted or rejected. In the following subsections we discuss two different cases: inferences induced by *individual dialogue acts* (subsection 4.1) and inferences initialized by *chains of dialogue acts* (subsection 4.2).

To this end we introduce a more formal account for dialogue acts, for the Thematic Structure and for the operations manipulating it.

4.1 Inferences Induced by Processing Individual Dialogue Acts

When dialogue acts are related to individual time objects of the Thematic Structure, certain follow-up actions for the manipulation of related time objects are induced. An example for such inferences can be observed in the following example:

GBP002: I guess we could meet on the tenth
(SUGGEST_SUPPORT_DATE)
I am free in the afternoon
(SUGGEST_SUPPORT_DATE)
SKH003: I am <äh> taking the tenth off
(REJECT_DATE)
<äh> I have the <äh> afternoon of the eleventh available
(SUGGEST_SUPPORT_DATE) ...

After processing turn GBP002 the Thematic Structure contains two objects: one of type day, representing the tenth of the month under consideration, and one of type period-of-day, which stands for the afternoon of that day. In the first utterance of turn SKH003 the tenth is rejected. This information leads to the information rejected_by: SHK being added to that object. Since rejecting a day implies the rejection of all parts of this day, i.e. of all subordinated time objects, this information has to be passed down the time object hierarchy. In the given case this means that the information is inherited to the afternoon object.

While a rejection *inherits* its information *downward* to subordinated time objects the information that a time object is accepted is *inherited upward* to superordinated time objects. This becomes clear by looking at the following example:

AKK001: well <ähm> the day that I am free the most is probably Friday the twentythird
(SUGGEST_SUPPORT_DATE)
I am free all day
(SUGGEST_SUPPORT_DATE)
is <äh> that okay with you
(REQUEST_COMMENT_DATE)
JEB002: I have a meeting that afternoon
(SUGGEST_EXCLUDE_DATE)
but <äh> the morning would be great <äh> around nine or ten o'clock
(SUGGEST_SUPPORT_DATE)
we can go 'till around two
(SUGGEST_SUPPORT_DATE)
AKK003: nine o'clock sounds pretty good
(ACCEPT_DATE)

The acceptance of the time object representing nine 'o clock implies the acceptance of all superordinated time objects, like the objects for morning, Friday, twentythird, etc.

Downward and upward inheritance can be captured formally as follows:

Downward Inheritance
(REJECT_DATE x A) $\Rightarrow$
$\forall y$ (superordinated x y) | (REJECT_DATE y A)

Upwards Inheritance
(ACCEPT_DATE x A) $\Rightarrow$
$\forall y$ (superordinated y x) | (ACCEPT_DATE y A)

where (REJECT_DATE x A) and (ACCEPT_DATE x A) stand for the time object x being either rejected or accepted by speaker A, and (superordinated v w) stands for the fact that in the Thematic Structure the time object v is superordinated to the time object w.

Other dialogue acts of our set that induce inheritance in the same way are SUGGEST_EXCLUDE_DATE which is used for mentioning time frames that are **not** available for an appointment (downward inheritance), SUGGEST_SUPPORT_DATE, that is employed when a time is proposed for a meeting (upward inheritance), and CONFIRM which is used to wrap up an agreed upon date (upward inheritance).

4.2 Inferences Induced by Processing Dialogue Act Sequences

In appointment scheduling dialogues we very often find that the evaluation of the content, i.e. of certain time frames, remains implicit and has to be inferred from the surrounding context. Such inferences occur frequently when a speaker change takes place. In this section we show how such inferences can be made exploiting the dialogue act history and the information represented in the Thematic Structure. In the following sections we focus on the discussion of inferences made on the basis of utterance pairs.

Counterproposal with Implicit Rejection

In our corpus of appointment scheduling dialogues we can find cases, where a dialogue partner makes a new proposal thereby implicitly rejecting a proposal made by the respective other dialogue partner. As can be seen from the following example, this information can be inferred from the propositional content of the utterance, i.e. from the information stored in the Thematic Structure.

JAK010: 'n Termin würde mir ganz gut passen, Montag, der achte, bis Freitag, der zwölfte, da hätt' ich dann irgendwann Zeit
there is one date that would suit me well, Monday, the eigth, until Friday, the twelfth, around that time I would be available.
SUGGEST_SUPPORT_DATE

REK011: ja ich denke an einem Wochenende,
yes I think on a weekend
SUGGEST_SUPPORT_DATE
weil 's ja wohl abends 'n bißchen länger wird.
because in the evenings it will get a bit late.
GIVE_REASON
ähm zwanzigster, einundzwanzigster.
twentieth, twentyfirst.
SUGGEST_SUPPORT_DATE

The proposal to meet around the twelfth which has been made in the last utterance of turn JAK010 is followed by a proposal made by speaker REK to meet around the twentieth. Since the two proposals are not compatible, it can be inferred that the latter proposal implicitly serves as rejection of the former. In general, it can be observed that the suggestion of a time object followed by the suggestion of an incompatible time object of the same type serves as implicit rejection. This principle can be described as follows:

Chained Proposals = Implicit Rejection

$(\text{SUGGEST\_SUPPORT\_DATE}_n\ x\ A) \wedge$
$(\text{SUGGEST\_SUPPORT\_DATE}_{n+1}\ y\ B) \wedge$
$(\text{type } x) = (\text{type } y) \wedge$
$(\ x \neq\ y) \Rightarrow$
$\quad(\text{REJECT\_DATE}_{n+1}\ x\ B)$

with the indices after the speech acts indicating the number of the utterance and with type standing for the category of the time object. In short this formula means that the time object proposed by A in the nth utterance can be considered rejected by B in the (n+1)th utterance if speaker B proposed an incompatible time object in utterance (n+1).

New Proposal with Implicit Acceptance of Previous Proposal

Suggestions followed by a different proposal made by the other dialogue partner do not necessarily imply a rejection of the initial proposal – the contrary can be the case:

SRH003: ...how 'bout any time in the afternoon
(SUGGEST_SUPPORT_DATE)
DTL004: how does three o'clock look
(SUGGEST_SUPPORT_DATE)

The time object presented in turn DTL004 is a refinement of the proposal made in the preceding turn SRH003, thereby implicitly accepting the proposal. This principle can be described as:

Chained Proposals = Implicit Acceptance

$(\text{SUGGEST\_SUPPORT\_DATE}_n\ x\ A) \wedge$
$(\text{SUGGEST\_SUPPORT\_DATE}_{n+1}\ y\ B) \wedge$
$(\text{superordinate } x\ y) \Rightarrow$
$\quad(\text{ACCEPT\_DATE}_{n+1}\ x\ B)$

Acceptance by Downward Inheritance

As shown in the previous subsection, the acceptance of an *individual* time object induces the acceptance of all superordinate time objects by means of upward inheritance. Here we show that also downward inheritance may be triggered by the acceptance of a time object, in case subordinate time objects are available in the Thematic Structure. The following dialogue fragment presents an example for such a behavior:

SKH003: ... <äh> I have the <äh> afternoon of
the eleventh available <äh> fourteenth
<äh> maybe morning of the sixteenth
(SUGGEST_SUPPORT_DATE)
GBP004: I am free all day the eleventh
(ACCEPT_DATE)

Among time objects introduced and proposed in turn SKH003 are the time objects for the afternoon of the 11th. This turn is followed by an utterance where the other speaker accepts the day. Because being available the whole day also implies having the (proposed) afternoon free, the acceptance information can be inherited down to the object of type period-of-day. In this case, therefore both upwards inheritance as introduced in section 4.1 and downward inheritance take place. The downward inheritance of acceptance information is captured by the following rule:

Acceptance by Downward Inheritance
(SUGGEST_SUPPORT_DATE$_n$ x A) $\wedge$ (ACCEPT_DATE$_{n+1}$ y B) $\wedge$ (superordinate y x) $\Rightarrow$ (ACCEPT_DATE$_{n+1}$ x B)

Implicit Acceptance through Phase Change

In some cases a date proposed by one speaker can be considered as implicitly accepted if the respective other dialogue participant moves on to another dialogue phase. An example for such a phenomenon can be observed in the following dialogue fragment:

JDH004: ... well I get out of class at one,
DELIBERATE_EXPLICIT
so give me about a half hour or so and maybe we can get together
for one thirty in the afternoon then
SUGGEST_SUPPORT_DATE

I will see you then
SUGGEST_SUPPORT_DATE
SMA005: do you want to meet in my office or mine <äh> yours
SUGGEST_SUPPORT_LOCATION

In this example the proposed date is implicitly accepted by speaker SMA since he switches to the phase where the location for a meeting is being negotiated[4]. Similarly, progression to the closing phase implies an acceptance of everything the dialogue partners proposed before, assuming that the two interactants show cooperative behavior.

A formal account for the implicit acceptance of a proposal through phase change looks as follows:

Phase Change = Implicit Acceptance

$(\text{SUGGEST\_SUPPORT\_DATE}_n\ x\ A) \wedge$
$(<\text{dialogue\_act}>_{n+1}\ y\ B) \in$
$\{\text{NEGOTIATE\_LOCATION}, \text{CLOSING}\} \Rightarrow$
$(\text{ACCEPT\_DATE}_{n+1}\ x\ B)$

Of course, the principles discussed above can be freely combined in order to process also highly complex examples from our corpus of appointment scheduling dialogues. Whether there is a specific order in which these rules have to be applied and whether it is necessary to develop resolution strategies for conflicting rule applications is subject to future research.

5 The Treatment of Indirect Speech Acts - Discussion

A difficult problem in the processing of natural language utterances is the determination of *indirect speech acts* (for a detailed account of indirect speech acts see [15]). These speech acts occur in cases where the surface form of an utterance does not – or at least not fully – coincide with the intention that stands behind that utterance. An utterance can fulfill *one or more* purposes, one being clearly expressed in the surface form and the other being left to the hearers' inferences.

Systems that rely only on surface cues for the determination of dialogue acts are not able to capture indirect speech or dialogue acts. Keyword spotting techniques, for example, cannot be used to find indirect speech acts unless they are combined with mechanisms like focus tracking. Therefore, *deep* methods are

[4] For the treatment of some of our appointment scheduling dialogues it is necessary to divide the negotiation phase into two subphases: NEGOTIATE_DATE where a date for a meeting is being negotiated, and NEGOTIATE_LOCATION where a location for a meeting has to be determined. The negotiation of a location is not part of the official VERBMOBIL scenario, though.

required that abstract away from the surface form and that determine indirect speech acts on the basis of semantic and pragmatic evidence.

The approach proposed in the previous chapter, i.e. the incremental construction of the Thematic Structure and the derivation of additional information can be seen as an approach for the detection of a subset of indirect dialogue acts in appointment scheduling dialogues. This is the case in particular for implicit acceptance and rejection in the case of chained proposals and for the implicit acceptance in the case of phase change. These rules introduce new dialogue acts that have not been explicitly attributed to the previous utterances. For the rules of downward and upward inheritance, instead, no new dialogue acts can be inferred. It is rather the case that a dialogue act already available for the previous utterances is extended to new time objects. These rules therefore do not contribute to the discovery of indirect speech acts.

6 Related Research

Although there is a growing interest in the field of discourse processing and therefore in the area of contextual reasoning, not much work is reported about the use of contextual information in spoken dialogue systems.

Among the few approaches described the treatment of context in the P2 system (see [13]) bears most resemblance to our approach: the elements of which context is composed in P2 have been motivated from the specific requirements of a spoken dialogue system, like e.g. robustness and efficiency; from all the possible aspects of which context can be theoretically composed (for an overview see, e.g. [4]) only those aspects are used that are of immediate use by other system subcomponents. The context representation in P2, is composed of a *task record* that captures task-relevant information being exchanged during the dialogue, and a *dialogue contents history* which records the order in which the subtasks have been executed together with their propositional content. These two aspects of context can best be related to Thematic Structure and Sequential Memory in our system, respectively.

Relevant for our work is also the approach taken in [5]: this system processes dialogues with the input being analyzed thematically, intentionally, and interactionally. The results are stored in a short term memory, a working memory and a long term memory, the latter providing a frame-like representation of syntactic, semantic and pragmatic information.

Another approach that influenced the design of our Dialogue Memory is the three-tiered model proposed in, e.g. [10]: *Linguistic Tier Information* captures linguistic realizations of the concepts under discussion, *Discourse Tier Information* captures the speakers' model of the dialogue, i.e. of *what* has been said, and *Belief Tier Information*, finally, captures the cognitive state of the dialogue participants. While the information captured in the Discourse Tier can be related to the time objects of our Thematic Structure, the Belief Tier vaguely corresponds to the evaluation information attached to these objects.

Most of these approaches refer back to the three-level representation of discourse structure as proposed by Grosz and Sidner [6]; they distinguish (1) the intentional structure that describes the goals that are followed in a dialogue, (2) the attentional structure, that corresponds to the propositional content focused in the discourse, and (3) the linguistic structure which is related to the linguistic means used to convey this information.

7 Conclusion and Future Work

In this paper we presented a detailed account for the representation of context information in a spoken dialogue system. The Dialogue Memory which fulfills this task consists of a Sequence Memory to capture the chronological structure of the Dialogue and of a Thematic Structure that captures the task-relevant propositional information. We also showed, how the Thematic Structure supports inferences over time information and how these inferences can serve as the basis for the treatment of indirect speech acts in appointment scheduling dialogues.

Together with the dialogue component the Dialogue Memory is fully implemented and incorporated into the VERBMOBIL Research Prototype. As for the inferences induced by the processing of dialogue acts the system so far is able to capture downward and upward inheritance. The implementation of the other principles proposed in this paper will follow in a later stage of the project. Additionally, we will examine our corpus for the occurrence of more phenomena that can be explained by means of inferences over the Dialogue Memory, and over the Thematic Structure in particular.

Since we plan to extend the VERBMOBIL scenario to a different domain we also will examine how the principles that have been identified here carry across to different applications. Depending on our findings we intend to develop techniques that allow an easy adaption of the inference mechanisms for other domains.

References

[1] Jan Alexandersson. Plan recognition in VERBMOBIL. In Mathias Bauer, Sandra Carberry, and Diane Litman, editors, *Proceedings of the IJCAI-95 Workshop The Next Generation of Plan Recognition Systems: Challenges for and Insight from Related Areas of AI*, pages 2–7, Montreal, August 1995.

[2] Jan Alexandersson, Elisabeth Maier, and Norbert Reithinger. A Robust and Efficient Three-Layered Dialog Component for a Speech-to-Speech Translation System. In *Proceedings of the 7th Conference of the European Chapter of the ACL (EACL-95)*, pages 188–193, Dublin, Ireland, 1995.

also available as Verbmobil Report Nr. 50, DFKI GmbH, Dezember 1994. available in the cmp-lg electronic archive under no. cmp-lg-9502008.

[3] Harry C. Bunt. Rules for the Interpretation, Evaluation and Generation of Dialogue Acts. In *IPO Annual Progress Report 16*, pages 99–107, Tilburg University, 1981.

[4] Harry C. Bunt. Context and Dialogue Control. *Think*, 3:19–31, May 1994.

[5] Brigitte Grau, Gérard Sabah, and Anne Vilnat. Control in man-machine dialogue. *Think*, 3:32–55, May 1994.

[6] Barbara J. Grosz and Candace L. Sidner. Attention, intentions and the structure of discourse. *Journal of Computational Linguistics*, 12(3), 1986.

[7] Bernd Hildebrandt, Gernot A. Fink, Franz Kummert, and Gerhard Sagerer. Understanding of Time Constituents in Spoken Language Dialogues . In *Proceedings of International Conference on Spoken Language Processing (ICSLP'94)*, 1994.

[8] Susanne Jekat, Alexandra Klein, Elisabeth Maier, Ilona Maleck, Marion Mast, and J. Joachim Quantz. Dialogue Acts in VERBMOBIL. Verbmobil Report 65, Universität Hamburg, DFKI Saarbrücken, Universität Erlangen, TU Berlin, 1995.

[9] Uwe Küssner and Manfred Stede. Zeitliche Ausdrücke: Repräsentation und Inferenz. Technical Report Verbmobil Memo 100, Technische Universität Berlin, December 1995. In German.

[10] Susann LuperFoy. Implementing File Change Semantics for Spoken-Language Dialogue Managers. In *Proceedings of the ESCA Workshop on Spoken Dialogue Systems – Theories and Applications*, pages 181–184, Vigsø, Denmark, 1995.

[11] Elisabeth Maier. A Multi-Dimensional Representation of Context in a Speech Translation System. In Lucja Iwanska and Wlodek Zadrozny, editors, *Proceedings of the* IJCAI*-95 Workshop on Context in Natural Language Processing*, pages 78–85, Montreal, August 1995.

[12] K. Pepelnjak, J. Gros, F. Mihelič, and N. Pavešić. Linguistic Analysis in a Slovenian Information Retrieval System for Flight Services. In *Proceedings of the ESCA Workshop on Spoken Dialogue Systems – Theories and Applications*, pages 65–68, Vigsø, Denmark, 1995.

[13] Laila Dybkjær, Niels Ole Bernsen, Hans Dybkjær, and Dimitris Papazachariou. Contextual Elements in the Danish Dialogue System. In Laila Dybkjær, editor, *Proceedings of the Second Spoken Dialogue and Discourse Workshop, Topics in Cognitive Science and HCI 8*, Centre for Cognitive Informatics, Roskilde University, Denmark, 1995.

[14] Norbert Reithinger and Elisabeth Maier. Utilizing Statistical Speech Act Processing in VERBMOBIL. In *Proceedings of the 33rd Annual Meeting of the Association for Computational Linguistics (ACL-95)*, Cambridge, MA, June 1995.

[15] J.R. Searle. Indirect Speech Acts. In P. Cole and J.P. Morgan, editors, *Syntax and Semantics*, volume 3: Speech Acts. Academic Press, New York, 1975.