A plan-based model of misunderstandings in cooperative dialogue

LILIANA ARDISSONO, GUIDO BOELLA AND ROSSANA DAMIANO
Dipartimento di Informatica, Università di Torino, Corso Svizzera n.185, 10149 Torino, Italy. email: {liliana, guido}@di.unito.it

This work is going to appear in the International Journal of Human-Computer Studies, Special Issue on *Detecting, Repairing, and Preventing Human-Machine Miscommunication.*

We describe a plan-based agent architecture that models misunderstandings in cooperative NL agent communication; it exploits a notion of coherence in dialogue based on the idea that the explicit and implicit goals which can be identified by interpreting a conversational turn can be related with the previous explicit / implicit goals of the interactants. Misunderstandings are hypothesized when the coherence of the interaction is lost (i.e. an unrelated utterance comes). The processes of analysis (and treatment) of a misunderstanding are modeled as rational behaviors caused by the acquisition of a supplementary goal, when an incoherent turn comes: the agent detecting the incoherence commits to restore the intersubjectivity in the dialogue; so, he restructures his own contextual interpretation, or he induces the partner to restructure his (according to who seems to have made the mistake). This commitment leads him to produce a repair turn, which initiates a subdialogue aimed at restoring the common interpretation ground. Since we model speech acts uniformly with respect to the other actions (the domain level actions), our model is general and covers misunderstandings occurring at the linguistic level as well as at the underlying domain activities of the interactants.

1 The phenomenon of misunderstandings

1.1 INTRODUCTION

In order for a dialogue to proceed successfully, it is important that the speakers maintain a consistent view of the interaction. Nevertheless, vagueness and ambiguities can intervene in a speaker's utterance and interfere with the process of interpretation of that utterance. Many researchers have recognized the role of various ambiguity types in generating misunderstandings, from different perspectives (Zaefferer 1977, Dascal 1989, Blum-Kulka & Weizman 1988, Vendler 1994).

In the linguistic and philosophical research, misunderstandings have been analyzed under different points of view: for example, Zaefferer (1977) and Dascal support the idea that the study of misunderstandings is related with comprehension in an indirect, external way, "much in the same way as pathological behavior is often said to be able to illuminate the nature of 'normal' behavior" (Dascal 1989). More recently, misunderstanding

has been recognized by Blum-Kulka & Weizman (1988) and Weigand (1997) as a normal phenomenon in communication, to introduce into an harmonical model of dialogue.

In the research on conversational analysis, Schegloff (1987), Schegloff (1992) has analyzed misunderstandings with respect to the sequences of turns in a dialogue, in order to identify which specific mechanisms are used by the interactants to defend the common set of beliefs necessary for the interaction to go on successfully: he calls *"intersubjectivity"* this common ground in dialogue and points out that, during an interaction, the speakers monitor their partners' reactions and interpret them as displays of understanding / misunderstanding of the previous turns. "Having registered the observation that, through their talk, speakers can display aspects of their understanding of prior talk, it remains to be noted that, in doing so, they can reveal understandings that the speakers of that prior talk find problematic - in other words - what they take to be *mis*understanding" (Schegloff 1992). Schegloff considers various cases where there are problems in maintaining the intersubjectivity among the interactants; he studies different types of repairs, which can be started by the misunderstood speaker, or by the agent who has misunderstood a previous turn, when they realize that something is wrong with the interpretation. He explains that repairs are an important instrument for reconstructing the mutuality in the interactants' beliefs.

In this paper, we describe a plan-based agent model that takes into account the problem of misunderstandings in (cooperative) agent interaction. This model supports NL communication and interprets dialogue as a rational, cooperative form of interaction among agents. The coherence of a dialogue is assessed by identifying the relation among each turn and the previous interaction context: we model a goal-based notion of coherence, which builds on the idea that the explicit and implicit goals identifyable by interpreting a conversational turn can be related with the previous explicit / implicit goals of the interactants (Allen 1983, Litman & Allen 1987, Carberry 1990). Under the assumption that the interactants cooperate and that every turn is performed to carry on some of their goals jointly, a misunderstanding is hypothesized when the coherence of the dialogue is lost, i.e. an utterance comes which is not related with the contextual goals. In fact, the presence of an utterance that does not contribute to the goals pursued in the receiver's interpretation of the dialogue is taken as a sign that his interpretation context is different from the speaker's one (who certainly considered his own turn coherent).

The processes of analysis and resolution of a misunderstanding are rational behaviors caused by the acquisition of a supplementary goal when an incoherent turn comes: when an agent A interprets a turn and finds it incoherent with respect to the previous context, he can adopt the intention of looking for an alternative interpretation of the interaction. This choice is the consequence of another goal which he decides to pursue: A wants that his view of dialogue converges with his interlocutor's one. If he considers it unlikely that the speaker B has performed a topic shift, or that a breakdown in cooperation has occurred, he looks for an action to modify the wrong interpretation which caused the misunderstanding. Such an action must lead one of the interactants, possibly himself, to change the interpretation context: so, A restructures his own interpretation of the previous dialogue, or he induces B to restructure his dialogue context, depending on who A believes to have made the mistake. This commitment leads A to produce a repair turn, which initiates a subdialogue aimed at restoring the common interpretation ground.

Many researchers (e.g. consider Pollack (1990), Perrault (1990), Hobbs *et al.* (1993),

Nagao (1993) and Hirst *et al.* (1994)) have used abductive frameworks for carrying on the interpretation of a dialogue and possibly restructuring the dialogue context, when a failure occurs in the integration of new utterances. In our plan-based model, the whole interpretation process is represented as the execution of interpretation (and reinterpretation) actions, as well as restructuring actions which modify the agent's dialogue interpretation; moreover, the maintenance of a correct interpretation context is a mutual goal of the interactants (Cohen & Levesque 1991). In this sense, the actions undertaken by agents to recover from a misunderstanding are not a separate activity with respect to their normal rational behavior: they are adopted as commitments by them, when necessary.† We believe that this is a major difference with respect to the other models of dialogue processing, where the treatment of misunderstandings (and general communication problems) is embedded in the normal interpretation process, or is managed by external recovery strategies (like the metarules used in Eller (1993) to restructure the dialogue context). Moreover, some of these models are mainly focused on misunderstandings in the interpretation of the illocutionary force of utterances and are constrained by a rigid notion of coherence. Instead, we claim that the misunderstanding phenomenon goes beyond dialogue and regards wider kinds of interactions; in our model, we analyze misunderstandings due to ambiguities both at the level of communicative and domain activity.

Incoherent turns are not always due to misinterpretations: also topic shifts and intentional breakdowns in cooperation should be considered. Currently, we don't model topic shifts due to the initiation of new dialogues; however, as pointed out in other works,‡ focus and topic shifts are usually marked by the presence of "cue" words. So, they can be distinguished from misunderstandings. In our model, this means that when an unexpected change in a dialogue occurs, the presence of a cue word should trigger the hypothesis that a topic shift has occurred, before hypothesizing that a misunderstanding has occurred. So, no misunderstanding analysis would start. On the contrary, since we model cooperative dialogues, we exclude the hypothesis that a breakdown in cooperation can occur.

The paper is organized as follows: after a brief description of the phenomenon of misunderstandings (section 1.2), we will present our notion of coherence in dialogue and our approach to the repair of a misunderstanding (section 2). Then, we will describe our computational model of dialogue (section 3): the agent model (section 3.1), the interpretation process of utterances (section 3.2) and the recognition and recovery from a misalignment in dialogue (section 3.3 and 3.4). We will then provide a detailed example to show how our model works (section 4). Section 5 describes the evidence about the occurrence of misunderstandings gathered by studying a number of dialogues. Finally, section 6 compares our model to other related works and section 7 concludes the paper.

1.2 BACKGROUND

Schegloff (1992) describes the behavior of agents when a misunderstanding occurs and identifies different types of repairs. He also implicitly supports the idea that misunderstandings can be distinguished according to which type of information is actually misinterpreted; for example, an incorrect identification of the referent of a description represents

† The interpretation and restructuring actions undergo the same planning and execution process as the other actions, like domain-level actions and speech acts.

‡ For example, consider Cohen (1984), Cohen (1987), Litman (1986), Litman & Allen (1987), Grosz & Sidner (1986) and Grosz & Sidner (1990).

a different level phenomenon with respect to an incorrect identification of the intended speech act, or of the plan underlying the interaction.

In Schegloff's description, third position repair corresponds to the cases where the misunderstood agent realizes that the partner has a wrong interpretation and urges him to restructure it. For example, consider turn T4 in the following interaction (Schegloff (1992), page 1317):

Example 1:

T1: Dan: *"... See Al tends, it seems, to pull in one or two individuals on his side [...]"*
T2: Al: *"WT"*
T3: Roger: *"Well so do I."*
T4: Dan: *"Yeah. I'm not criticizing, I mean we'll just ... "*
T5: Roger: *"Oh you wanna talk about him"*
T6: Dan: *"look, let's just talk"*

In the example, Dan realizes that he has been misunderstood by Roger and provides him with the intended interpretation of his own previous sentence by performing a third position repair in turn T4: in this way, he urges Roger to change his interpretation of the dialogue.†

In fourth position repair,‡ a speaker A produces a turn that is misunderstood by B, who replies with another turn, sequentially appropriate to his own (wrong) understanding of the first turn. A doesn't realize that the two interpretations are not aligned any more and responds with a further turn, coherent in A's view of dialogue. At this point, B understands his own mistake, restructures his own interpretation accordingly and informs the partner about the realignment. For example (Schegloff (1992), page 1321):

Example 2:

T1: Marty: *"Loes, do you have a calendar,"*
T2: Loes: *"Yeah"* ((reaches for her desk calendar))
T3: Marty: *"Do you have one that hangs on the wall?"*
T4: Loes: *"Oh, you want one."*
T5: Marty: *"Yeah"*

The recognition that a misunderstanding has occurred is triggered by the lack of coherence between the interpretation of the misunderstood turn and that of the subsequent turns. Having misinterpreted, in fact, the hearer gives his contribution to dialogue in a way that is coherent in his wrong interpretation, but not from the point of view of his partner's intended meaning.

† Turn T5 looks like a fourth position repair by Roger, but we will discuss this aspect in section 2.2.

‡ Third and fourth position repairs are not necessarily the third / fourth turns of the interaction, starting from the misunderstood turns: several turns (even long subdialogues) can come, before an agent realizes that an interpretation problem has occurred. Anyway, in third position repairs, the misunderstood speaker performs the repair turn after a "sequentially inappropriate" turn from his partner comes (so, the repair comes in third position, in an extended way); instead, in fourth position repair the misunderstander detects his own mistake after the partner speaks again; so, the repair comes in an extended fourth position. In Schegloff (1992), the expression *"sequentially appropriate"* has been used to indicate the coherence relation between a participant's understanding of a turn and his reply.

2 Coherence and misunderstandings in dialogue

2.1 COHERENCE AS A RATIONAL PHENOMENON

We analyze dialogue from the intention recognition point of view (Cohen *et al.* 1981, Allen 1983, Allen & Perrault 1980, Cohen & Levesque 1991): when an agent acts, a relation of his action with the interaction context is looked for, to see whether the action represents an attempt to satisfy any intention expressed explicitly by the partner, or an implicit goal which can only be inferred by reasoning on the partner's plans, or if it is a further step in a plan that the agent has already started. A different approach to the analysis of coherence was adopted by the conversational analysts, who introduced *adjacency pairs* (Sacks *et al.* 1974) to model the expected continuations of an interaction: adjacency pairs are sequences of speech acts (e.g. question-answer pairs) such that, after the first element occurs, the second one is expected. However, as pointed out in Levinson (1981), agent behavior can not directly be explained by means of such type of strict interactional rules. On the contrary, the intentional approach to dialogue interpretation makes it possible to adopt a flexible notion of coherence: a new contribution is considered coherent as long as a relation can be identified among the intentions underlying an action and the previous pending intentions of the interactants. Following the ideas of Castelfranchi & Parisi (1980), we consider an utterance coherent with the previous context if and only if its receiver can interpret it as a means of the speaker to achieve an unsatisfied goal g which realizes one of the following coherence relations:

1. *Goal adherence:* g is one of the goals addressed explicitly by the partner in a speech act (not necessarily in the last turn of the interaction).
   Goal adherence covers the phenomena treated by adjacency pairs, but it is a more general approach than that. We represent speech acts as actions (Austin 1962) having the effect that the interactants share the belief that the speaker intends the partner to execute a certain (domain-level or linguistic) action. An illocutionary act and the action occurring in its effect clearly correspond to the first and second components of an adjacency pair (e.g. a question is performed to induce the hearer to answer it). In the simplest case, dialogues are composed of sequences of adjacency pairs because the participants recognize each other's intentions and react to them as expected. However, other behaviors are modeled in our intentional approach: e.g. insertion sequences (Schegloff 1972), like "question-question-answer-answer" sequences (where further questions are asked before the answer to the initial question is provided) can be modeled as investigations on the actions addressed in the initial question. Asking questions about the preconditions of an action is explained as a coherent continuation, where the speaker tries to execute the action and, to do that, moves to the satisfaction of the subgoal of checking whether the action is executable or not. E.g. consider the following excerpt, taken from Merritt (1976):

   T1: A: *"Sell me a bottle of whisky, please."*
   T2: B: *"Are you 18?"*

   Also notifications of (un)successful performance of actions (e.g. acknowledgements) follow from the satisfaction of a joint goal (see Cohen & Levesque (1991)):

   T1: A: *Could you register me for the Artificial Intelligence exam ?*
   T2: B: *Ok, you are registered.*

2. *Goal adoption:* g is one of the goals that the speaker B has inferred the hearer A is aiming at; this goal has not been expressed explicitly by A, but B has inferred it by reasoning on A's plans. For example, in:

   T1: A: *"I need to borrow a book. Where is the library?"*
   T2: B: *"The library is over there, but it is closed for the whole week."*

   B replies to A's utterances by providing him with the information that his plan can not be executed, because one of its constraints is false. B's second utterance follows from the fact the he has adopted A's goal to know whether he may enter the library (Allen 1983). A didn't express such a goal, but it can be identified from the fact that A is checking another condition (a knowledge precondition (Morgenstern 1987)) of the plan of using the library.

3. *Plan continuation:* g is a subgoal deriving from the course of actions already undertaken by the speaker; so, the speaker is carrying on the execution of his plan. For example, in:

   T1: B: *"Where is the Computer Science library?"*
   T2: A: *"It is in the underground floor."*
   T3: B: *"Do you know if it is open today?"*

   T3 is not related with A's previous turn T2, but it can be related with T1, as another step in B's preparation for going to the library. Also turns aimed at checking whether actions have succeeded represent continuations of an agent's acting; for example, consider: *"Can you give me your book? .... Did you hear me?"*.

In summary, the coherence of the turns of a dialogue is assessed by identifying a structured set of goals which relate the linguistic and domain actions of the interactants. During the interpretation of a turn, a relation among the goals pursued by the turn and the contextual pending goals is searched for in a "bottom-up" way (as we will explain in Section 3.2): the turn is matched on the "nearest" pending goals and, possibly, on the other, higher-level ones. In this way, the existence of a goal adherence relation is identified before a goal adoption relation, which in turn precedes a possible plan continuation relation. This corresponds to hypothesizing that, before carrying on one's own plan, an agent tries to satisfy his partner's open goals.

Note that this notion of coherence is not necessarily associated with cooperation among the interactants. In fact, it is possible to respect the first two conditions above, while thwarting the satisfaction of a goal g; as pointed out in Castelfranchi (1992), when this happens, an exchange between two debating agents can be perceived as fluent.

In our model, the "display of understanding" described by Schegloff is mirrored on the ability of producing goal-coherent turns. For example, acknowledgements, as well as the production of a coherent turn, display understanding (e.g. consider a proper answer to a question). When an agent tries to interpret a new turn from his interlocutor, he might not be able to relate it with the previous context. In that case, he should reconsider such context to see whether there is an alternative interpretation of some turns which restores the coherence of the whole context.

2.2 REPAIRING TO A MISUNDERSTANDING

When an agent A receives an incoherent turn, he commits to restoring the intersubjectivity in the dialogue; in principle, A has two possibilities: the first one is to change his own view of the interaction; the second one is to persuade his partner B to change his own. The appropriate strategy depends upon the alternative (coherent) interpretation(s) of the interaction that A is able to find out; the existence of such an alternative interpretation is an applicability condition of the action of restoring the intersubjectivity:

1) If the last (apparently incoherent) turn becomes coherent by changing the interpretation of some previous turn T_i uttered by himself, then A can inform B he should correct his interpretation of T_i in order to adjust his view of the interaction.

2) If the last turn becomes coherent by changing the interpretation of some previous turn T_j uttered by B, then A will restructure his own interpretation of the dialogue.

When the hearer A chooses an alternative interpretation of a turn, he has to reconstruct the interpretation of the whole sequence of turns, in order to identify the partner's dialogue context. In fact, the speaker B who has uttered the problematic turn, should have an interpretation that is coherent from his point of view, although different from A's. Empirical data (see section 5) show that agents succeed in identifying the alternative interpretations of the misunderstood turns, both in cases they have been produced or recognized by them. In fact, repairs are often composed of a part where the recognized wrong interpretation is made explicit (e.g. in turn T4 of Example 1, *"I'm not criticizing ..."*). Referring to the classification of repairs in Schegloff (1992), we see that:

1) In third position repair, the speaker of the misunderstood turn stops the diverging dialogue and formulates a repair (a request for a reconstruction of the interpretation), to induce his partner to replace the interpretation of the earlier turns. Schegloff identifies a well defined recurrent schema for formulating a repair, with a few types of realizations and a clear rationality.

2) Instead, fourth position repair looks like a notification that a change of state happened: typically, it is structured as in Example 2 (section 1.2), with a marker like *"Oh"* (Heritage 1984), and a sort of grounding sentence (Clark & Schaefer 1989), where the agent informs the partner that he has finally understood what the other meant initially.

In Example 1 (Section 1.2), it can be seen that turn $T5$ (*"Oh you wanna talk about him"*), which follows the request to reconstruct the interpretation of the dialogue, is similar to an acknowledgement of action execution, just as it happens in a fourth turn repair. This similarity suggests that in fourth turn repair, the agent who restructures his own interpretation performs the same action triggered on another agent by a third turn repair. As a fact, both phenomena are characterized by the same underlying goal: the reconstruction of the intersubjectivity. However, the reason for adopting such a goal is different: when a third position repair is performed, the hearer is solicited by the partner's request, which typically also suggests the intended interpretation of the turn (*"Yeah. I'm not criticizing, I mean we'll just ..."*); instead, in fourth position repair, the hearer adopts the goal of recovering from the context misalignment and commits to changing his own interpretation autonomously. In both cases, the speaker notifies the partner that he has restructured his own dialogue context. If, instead, there is no way of realigning the subjective views of the interaction, it is not a problem: unfortunately, intersubjectivity sometimes breaks down; typically, in this case, the failure is notified to the partner.

3 A computational model of misunderstandings

3.1 REPRESENTATION OF THE AGENT KNOWLEDGE ABOUT ACTING

Our analysis of misunderstandings is performed in a plan based architecture for generating and recognizing agent behavior, based on a two-level representation of the knowledge about acting. At the metalevel, the Agent Modeling (AM) library (Ardissono *et al.* 1996) describes the recipes for planning and executing actions; at the object level, there are the Domain and the Speech Act plan libraries, and some internal actions like those describing the interpretation task; the Domain library (Ardissono *et al.* 1993) describes the recipes for obtaining the domain goals in a restricted domain; the Speech Act Library (Ardissono *et al.* 1995a, Ardissono *et al.* 1995b) keeps the linguistic knowledge. The three plan libraries are based on a Generalization and a Decomposition Hierarchies (Kautz 1991) and share the same representation formalism, so that the same procedures can be used on them. Moreover, plans support a declarative representation style, that can be used both for interpreting and generating the behavior of the agent. The idea of using the same structures to model interpretation as well as generation is basic to our notion of coherence, that presupposes that agents have the same Agent Modeling plans (they have the same rational behavior, although their beliefs and goals may differ substantially). The agent's knowledge also includes some prior information about the world state; such knowledge is strictly domain dependent and regards beliefs about other agents, about conditions which hold in the world, etc. Although the presence of this type of knowledge poses a limit in the applicability of the model to general domains, it is a basic component of any system which aims at performing a deep interpretation of the interactions among agents.

3.1.1 The object level actions

At the object level there are the Domain plan library and the Speech Act Library.
The *Domain plan library* contains pre-compiled recipes which describe typical well formed plans to obtain domain goals. We chose the University environment as an experimental domain; so, actions regard borrowing books from the library, taking exams, and so on. The structure of this library recalls other well known solutions, so we will not describe it here.
The *Speech Act library* contains the definition of speech acts: they are represented, similarly to domain actions, as acts that an agent performs to try to change the world state (in this case, the hearer's beliefs). The reason for representing speech acts at the object level is that, in this way, their planning and execution can be reduced to that of any other actions, by means of the Agent Modeling plans, and the treatment of dialogue is unified for any type of interaction. In particular, we model communication by introducing multiple (object level) actions, each of which has a different role:
- at the highest level, speech acts are performed to induce the hearer to act in a certain way (Smith & Cohen 1996); action "Get-to-do" represents this level and has the effect that the hearer intends to act in a certain way;†
- "Get-to-do" is realized by means of one of a set of alternative illocutionary acts (e.g.

† We model the action of altering one's own beliefs as an "Update" action, which is treated as any other object level action.

a request, an order, or other), which have a weaker effect: that it is shared among the interactants that the speaker has the communicative intention that he intends that the hearer intends to act in a certain way. "Get-to-do" can also be realized by means of a complex communicative act, composed of multiple speech acts, which are related by rhetorical relations (Moore 1995, Barboni & Sestero 1997);
- the illocutionary act can be performed, in turn, by means of different surface speech acts (e.g. a request may be performed as an imperative sentence, a question, or other). We use direct and indirect speech acts to describe some of the politeness techniques used in dialogue, following the taxonomy of politeness strategies in Brown & Levinson (1987);
- the surface acts are all performed by means of a "Locutionary-act", which has as an argument the syntactic and semantic representation of the speech act; in turn, this is realized by an "Utterance-act" which contains a suitable text string.

As an example, we report the "Ask-if" illocutionary act (asking information about the truth value of a condition):†

Ask-if :=

name:	Ask-if
roles:	((speaker x)(hearer y)(propositional-content *cond*))
var-types:	((person x y)(condition *cond*))
effect:	SH(x, y, Cint(x, y, Goal(x, Goal(y, Inform-if(y, x, *cond*)))))
constraint:	¬Bel(x, ¬Knowif(y, *cond*))
More-specific-actions:	Direct-ask-if(x, y, *cond*)
	Indirect-ask-if(x, y, *cond*)

3.1.2 The metalevel actions

The Agent Modeling (AM) actions take domain actions and speech acts as objects: the idea is that an agent who is performing a problem solving activity to obtain his goals can plan both domain and linguistic actions, according to what is better in the specific situation (Ardissono *et al.* 1996).

Figure 1 shows a portion of the metalevel library: as it can be seen, most of the AM actions have a *"source"* parameter (denoted as s) that represents the agent for whom the action is performed (i.e. its beneficiary). When an agent acts on his own, this parameter is bound to the agent himself; in general, *source* is used to model team cooperation. The actions are represented in the figure with the following graphical notation:
- specific actions are related to more general ones by means of thick arrows, labeled with the restriction on their executability which distinguishes a specific action from its alternatives (e.g. "agent(*action*)=a");
- the links in the Decomposition hierarchy are represented by boxed multiple arrows (the oval below "Cycle-do-action" represents a cyclic decomposition that is executed while the associated condition is true). Dashed lines denote conditional steps which are executed only when their "if" condition is true; the "wh" restrictions specify when an action can be a step of an higher-level one;

† In the action, SH is the mutual belief operator; Cint is the "communicative intention" operator introduced in Airenti *et al.* (1993): Cint(x, y, p) means that x has the intention that it is common belief among x and y that p holds, and that x had that communicative intention towards y. So, x wants that it is mutually believed p, and that x wanted to communicate it.

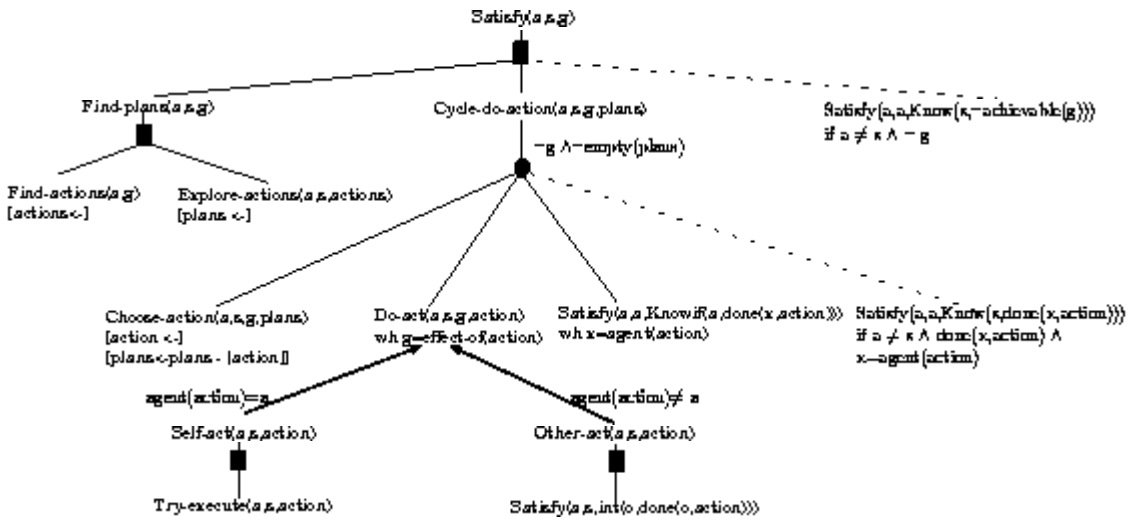

FIGURE 1. A portion of the AM plan library (the "Satisfy" plan).

- the notation "[x <-]" means that variable x is assigned a value as a result of the execution of the action to which the notation is associated.

The main AM action ("Satisfy(a, s, g)") describes the problem solving behavior of an agent a who wants to satisfy a goal g which he has committed to. Briefly, the agent looks for a plan which leads to g (this involves finding some actions - "Find-actions" - and exploring them hierarchically - "Explore-actions", if they are complex). Then, he chooses between two alternative behaviors: executing it by himself ("Self-act"), or inducing another agent to perform it ("Other-act"). After that, he checks whether the action has been executed successfully by its agent (who can be himself, or his partner): step "Satisfy(a, a, Knowif(a, done(x, *action*)))" describes this behavior (the "done" operator maps an action on the world state after a successful execution of the action). If the action was performed successfully and the agent is cooperating with another agent, he notifies the partner about the success (see the conditional step in the recipe). Instead, given that the execution can fail, the agent could loop in this activity ("Cycle-do-action"), by choosing another alternative for g and trying it. Finally, if none of the alternatives could be completed successfully and the agent is cooperating with another agent, he informs his partner that the main goal g is impossible.

The other major Agent Modeling action is "Try-execute(a, s, *action*)" (not shown in the figure), which describes the process of executing an object level action: when an agent is committed to performing an action, he first checks whether its constraints and preconditions are satisfied (this is described by an occurrence of the "Satisfy" action in the body of "Try-execute"); if the preconditions of the action are not satisfied, the agent can adopt them as subgoals. Then, if the action can be executed, he moves to its performance, or to the execution of its steps, if it is complex. At end, he has to check if the effects of the action hold, in order to know if he has succeeded; on the contrary, if the action can not be executed (because its constraints are false), the agent can inform the partner of the problem (see the notion of Joint Intention in agent cooperation introduced

in Cohen & Levesque (1991)).

Our agent model is a (*Belief, Desire, Intention*) BDI agent (Rao & Georgeff 1991). The notion of commitment adopted in our model follows the notion of intention and Joint Intention described in Cohen & Levesque (1990a), Cohen & Levesque (1990b), Cohen & Levesque (1991), as derived from the original description of intention given in Bratman (1990). The agent's behavior is ruled by a core interpreter which loops on three phases: the interpretation of the input, the decision of which high-level goal to commit to (on the basis of which is the internal state of the agent), and the reaction, where the agent starts acting to reach the chosen goal (by means of the execution of a "Satisfy" action on the goal). In our current implementation, the agent works only in the interpretation mode, so the system plays the role of the hearer for each turn of a conversation, by switching his role among the two interactants and letting the human user play the role of the "speaker" and type sentences on the keyboard.

3.2 INTERPRETATION OF UTTERANCES

During an interaction, the agent performs an incremental interpretation of the dialogue, by maintaining as a context the model of his and his partner's activities. Such dialogue context (*ctx*) is represented by the sequence of local interpretations of the (interleaved) turns of both speakers; in the context, the interpretation of each turn which has been considered coherent is explicitly related with some previous turn by reporting the coherence relation (adherence, adoption and plan continuation) identified with a contextual goal. Instead, incoherent turns are unrelated with respect to the previous context.

The local interpretation of a turn is a complex structure which represents its Agent Modeling, linguistic, internal and domain level actions (Figure 3 shows one of these structures; there, the three action types are contained in boxes labeled suitably, with the "AM", "SAM", "IM" and "DM" labels). In this interpretation structure, the AM plans relate the linguistic and domain activity, since they not only specify the steps necessary for (linguistic and domain) actions execution, but they also explain how a given speech act contributes to the problem solving activity of an agent who is working to satisfy a domain goal.

The process of local interpretation of a turn takes in input an interpretation structure which only contains the observation that the agent has performed an action (e.g. "Exec(A, Utterance-act(A, B, *"Do you have a calendar?"*)"); at this stage, the input utterance has not yet been interpreted syntactically and semantically. The interpretation process expands this structure by traveling along the plan libraries in order to find which high level goals explain the observation. Since the interpretation structures are composed of two levels, the AM actions and the object-level ones are expanded in parallel. During the expansion of a SAM (Speech Act Model), this process builds the syntactic (Lesmo & Torasso 1985) and semantic (Di Eugenio & Lesmo 1987) representation of the input sentence (the "Utterance-act" action generates a "Locutionary-act" if a semantic interpretation can be found for its object text string), and identifies the performed illocutionary act.

In such a complex interpretation structure, the possible ambiguities that lead to alternative interpretations (the main cause of the phenomenon of misunderstanding) may concern several aspects: the syntactic and semantic interpretation (including references)

of the sentence can cause the activation of more than one "Locutionary-act" hypothesis. The ambiguity in the illocutionary force leads to identify different actions representing speech-acts. A domain action can be interpreted as a means for more than one goal (e.g. going to the library can be a step of both borrowing book and studying in the library). Even the AM plans allow to construct different hypotheses on what the agent is doing (e.g. the applicability conditions of a domain action are checked in two places: when an agent is exploring the action and when he is executing it). Finally, the same local interpretation of a turn can be related to the previous context in different ways.†

The interpretation and coherence-seeking processes are represented by the "Build-interpretation" action, which is at the object level, and takes as arguments the interpreting agent x, the agent y to which x attributes the interpretation, the dialogue context for turns $t_1, \ldots, t_{n-1}$, and the last turn t_n. The effect of "Build-Interpretation" is to build agent x's view of y's interpretation of the dialogue, for the context $ctx_{1,n}$: $ctx_{1,n}$ is obtained from $ctx_{1,n-1}$ by incrementing it with the interpretation of turn t_n and, possibly, specifying the relation which links t_n with some previous pending goal of $ctx_{1,n-1}$:‡

Build-interpretation :=

name:	Build-interpretation
roles:	((agent x)(partner y)(old-inter $ctx_{1,n-1}$)(new-turn t_n))
var-types:	((person x y)(turn $t_1 \ldots t_n$)(context $ctx_{1,n-1}$ $ctx_{1,n}$))
effect:	Bel(x, inter(y, $[t_1, \ldots, t_n]$, $ctx_{1,n}$)
constraint:	inter(y, $[t_1, \ldots, t_{n-1}]$, $ctx_{1,n-1}$)

The Agent Modeling plans in a local interpretation provide the basic information to decide whether the new utterance is coherent with the previous context or not. In fact, the top AM action of a turn interpretation represents a speaker's goal G which led him to produce the turn: if the turn is coherent, G must appear also in the context. As described in section 2, a new utterance from a speaker A can be related to the dialogue context in different ways: it can aim at the satisfaction of some goal of the partner B (goal adherence / adoption), or at the continuation of the speaker's plan (plan continuation). In the first case, the goal which explains A's new turn appears in some previous turn interpretation of his partner B, having been expressed explicitly by him (by means of a speech act), or inferred by A from B's overt actions. Instead, in the plan continuation case, the top Agent Modeling action of A's turn interpretation can be linked to A's previous metalevel activity (the continuation of an AM plan likely corresponds to the continuation of an object level plan, either a domain or a communicative one, or to the satisfaction of some subsidiary goal due to the cooperation with the partner).

It must be noted that the process performing the local interpretation of a turn operates in a goal directed fashion, in cooperation with the coherence-seeking process: in principle, when a new action of uttering a sentence is recognized, one could try to interpret it locally, and take the context into account only after having recognized the speaker's goal (by looking for it in the context itself). On the contrary, we use some heuristics

† In this paper, we don't deal with the problem of how one specific interpretation is chosen for a turn. In Ardissono & Sestero (1996), we have discussed how maintaining a model of the speaker can help in discarding some alternatives, and how it can be useful to identify the information common to the alternative interpretations, as a minimal conveyed content.

‡ The meaning of predicate "inter(y, $[t_1, \ldots, t_n]$, $ctx_{1,n}$)" is that agent y interprets turns $t_1, \ldots, t_n$ as a context $ctx_{1,n}$.

which start from the (speaker's and hearer's) pending contextual goals to guide the interpretation of the new contribution. This is done by looking for the possible paths which relate an object level action with the plans occurring in the interpretation of the previous context. The search for those paths makes it possible to avoid trying every expansion of the AM local interpretations towards any higher-level goals. Anyway, if the top goal of a local interpretation can not be related with some goal pending from a previous turn interpretation, there is still the chance to expand these interpretations further, in order to find some higher-level goals that explain them. So, our heuristics also search for some higher-level goal, for which there are both a path leading to the new turn interpretation, and a path from some previous turn interpretation to the high-level goal.

Although the last turn of a dialogue may be incoherent with respect to the previous context, it is possible that the interpretation function produces a local interpretation, which remains unrelated. A real failure in the interpretation of a turn happens only when also the local interpretation fails, maybe because the sentence is syntactically / semantically ill formed, or it is impossible to identify the performed speech act, or the underlying domain goal.

3.3 RECOGNITION OF A MISUNDERSTANDING

In the linguistic research, the term *repair* has been used in a wide sense; in fact, phenomena due to problems in the understanding process (e.g. consider questions like *"Are you asking me or do you want to know?"*, *"What do you mean?"* or problems in the identification of a referent as described in Heeman & Hirst (1995)) have been classified as repairs (e.g. see Schegloff (1992)). In our model these questions are related with subgoals which derive from the execution of a "Build-interpretation" action: they are not associated with the resolution of misunderstandings because the speaker has not yet committed to one interpretation of his partner's turn, but he is trying to build one.

If, during the interpretation of an input utterance, the process searching for the coherence of the last turn fails but a local interpretation of the turn is still possible, then the interpretation action produces a context in which this local interpretation is unrelated with respect to the previous context. Since our model does not currently manage topic shifts and breakdowns in partners' collaboration, the agent (A) can only hypothesize a misunderstanding. Hence, in the second phase of the main agent loop (the one where the agent decides which goal to adopt next) A assumes that this interpretation hypothesis is not admissible and does not reflect what his partner has in mind. Under assumption that the partner (B) is actually collaborating, this implies that the intersubjectivity in the interaction has been lost. In this case, A adopts the goal of realigning the subjective views of the dialogue; in our model, this is done by the execution of a "Satisfy" action on the goal that the A's (private) interpretation of the dialogue is the same as that of B, up to the misinterpreted turn t_j:

$\text{Satisfy}(A, B, \text{inter}(A, [t_1, \ldots, t_j], ctx') \wedge \text{inter}(B, [t_1, \ldots, t_j], ctx'') \wedge \text{equal}(ctx', ctx''))$†

Given this goal, there is a planning phase, in which the agent looks for the actions having the goal among their effects. The two alternative ways to achieve the goal are restruc-

† The "Satisfy" action does not manage complex goals in a general way. It can however satisfy separately the subgoals of a conjunctive goal. In this particular formula, this is enough because, during the diagnosis of the misunderstanding, one of the context variables is bound to one of the speakers' real interpretation subcontext; so, it is not necessary to satisfy both the subgoals.

turing his own context, or inducing the partner to change his context. These alternatives correspond to two different instantiations of the same "Restructure" action.
The "Restructure" action must be executed by the agent who is misunderstanding his partner and consists of modifying his own interpretation of the turns of the dialogue to obtain the correct interpretation (i.e. the one intended by the partner).

Restructure :=

name:	Restructure
roles:	((agent x)(partner y)(turns $[t_1, \ldots, t_n]$) (old-inter ctx)(intended-inter ctx_j))
var-types:	((person x y)(turn $t_1 \ldots t_n$)(context ctx ctx' ctx_j))
effect:	inter(x, $[t_1, \ldots, t_j]$, ctx_j)
constraint:	inter(x, $[t_1, \ldots, t_n]$, ctx) $\wedge$ inter(y, $[t_1, \ldots, t_n]$, ctx') $\wedge$ correct-subcontext(ctx_j,ctx,ctx') $\wedge$ agent(y,$ctx_j[j]$)

The execution of the "Restructure" action is ruled by the same Agent Modeling actions that describe the performance of the domain-level and linguistic actions. As we said in section 3.1, before committing to an action (and before executing it) an agent checks whether its constraints are true; this checking process corresponds to the diagnosis of the misunderstanding, since the constraints of "Restructure" are the following ones:
- the agent x (who has to perform the action in order to correct his wrong interpretation) has an interpretation ctx of the whole dialogue, up to the last turn t_n;
- an alternative interpretation ctx' of the dialogue can be attributed to x's partner y;
- ctx_j (determined by comparing ctx and ctx') is the coherent interpretation of the misunderstood speaker, up to the misinterpreted turn t_j ($j < n$);
- the agent y corresponds to the speaker of the first turn that has been interpreted differently in ctx and ctx' (i.e. the last turn of ctx_j, denoted as $ctx_j[j]$);

To check whether the constraints of an action hold, the agent must perform a further "Satisfy" action on the goal of knowing if they are true. The truth value of a partially instantiated condition depends on whether a value for its unbound parameters is found that satisfies the condition; in this specific case, the value is the interpretation of the dialogue (i.e. the goal is "Knowref(A, ctx', inter(B, [...], ctx'))").

The "Knowref(...)" goal is obtained by executing an action corresponding to the "Reinterpret" algorithm described in Figure 2. This action, that the agent undertakes to identify the alternative interpretations of the turns of the dialogue, is very similar to the one previously performed in the standard interpretation task. The main difference is that while in the basic interpretation process the agent considers the interpretation of his own turns as fixed, in this phase, he must consider the alternative interpretations of such turns as well: by doing this, he tries to understand why the partner believes that the interaction is coherent.

When a hearer A chooses an alternative interpretation of a turn T_i, he has not only to check its coherence with the last problematic turn, but he also has to reconstruct the interpretation of the whole sequence of turns. In fact, if the speaker B has uttered the last turn, he must have a wrong, but coherent view of the whole interaction, otherwise he would have started a repair.

The reinterpretation process traces back the dialogue, turn after turn and looks for a relation between the problematic turn and some previous (probably misinterpreted) turn. In principle, it should be possible to look for a global alternative interpretation of the dialogue, starting from the last contribution, like in a standard backtracking procedure. But time would be wasted in producing many candidate alternatives which fail to relate with the last turn. Instead, we exploit the newest information for pruning the inadequate hypotheses: first, we look for an alternative interpretation of a single turn T_i, which explains the last contribution; then, we propagate the change, to see if the rest of dialogue becomes coherent: if the interpretations of the other turns are not adequate any more, an alternative for them is looked for.

```
Reinterpret(agt, [t1, ..., tn]) ≡
begin i := n - 1; ip := ∅; restC := ∅;
      while ( i > 0 ∧ empty(ip)) do
      begin ip := Build-interpretation(agt, ti, tn);
            if empty(ip) then i := i - 1
            else begin restC := Reinterpret(agt, [t1, ..., ip[1]], ctx);
                       if restC then
                         begin j := i + 1;
                               while ( j < n ) do
                               begin
                                 restC := Build-interpretation(agt, restC, tj);
                                  j = j + 1;
                               end;
                               if empty(restC) then ip := ∅ /* Try another interpretation pair */
                         end
                       else ip := ∅ /* Try another interpretation pair */
                 end
      end;
   return(restC ▫ ip[2]) /* Return the whole reinterpreted context */
end.
```

FIGURE 2. The algorithm for finding alternative interpretations of a dialogue.

"Reinterpret" has two arguments: the agent (*agt*) and a sequence of turns ([t_1, ..., t_n]). When it is called on a list of uninterpreted turns [t_1, ..., t_n], it goes backward from the problematic last turn t_n, towards the beginning of the dialogue, and stops when it finds the most recent turn t_i for which "Build-interpretation" has found a new interpretation, coherent with that of t_n.† Then, "Reinterpret" propagates the new interpretation *ip* (*interpretation pair*) to the whole context, using the interpretation of t_i in *ip* (the first component of this context, denoted by *ip[1]*) as a "pivot", backward by calling itself recursively, and forward by means of a loop of "Build-interpretation" on the remaining turns. When the loop stops, the new interpretation context is returned; this context is composed of the (sub)context for turns $t_1, \ldots t_{n-1}$ (*restC* in the figure) with that of turn t_n (denoted as *ip*[2]).

The algorithm stops changing the previous interpretation when an alternative interpretation of a turn t_i is found that is coherent with the interpretation of the turns t_1 ..., t_{i-1}. The speaker of the last changed turn (t_i) is the agent who has been misunderstood.

3.4 RECOVERY FROM A MISUNDERSTANDING

After the agent has identified the trouble source turn of the dialogue (i.e. the first misunderstood turn), he can react in two ways, according to who is the speaker of the turn: 1) If he has been misunderstood, he can try to persuade his partner to restructure his

† Although "Build-interpretation" takes in input a context and a turn, there is no contradiction here, because a single turn is by itself an elementary context.

interpretation by performing the "Restructure" action.† As described in section 3.1.1, the (object level) action "Get-to-do", is used to induce other agents to perform actions for one's sake; so, in this case, a request to restructure the context is performed (see T4 in Example 1, section 1.2).

2) If, instead, he has misunderstood his partner, he has to execute the "Restructure" action himself. As a result of this execution, his model of the previous dialogue is changed and he can go on with the interaction, having reestablished the interpretation context. In this new context, both agents share the same interpretation. Note that, before continuing the dialogue, the agent has still something to do: he notifies the partner that he has succeeded in realigning the interpretation.

In both cases, the recovery goal is shared among the speakers; when everything ends up well, the agent informs the partner about the success of their aims, otherwise, if no repair is feasible, the agent warns his partner that the intersubjectivity is unrecoverable.

Our model directly supports the (positive and negative) notifications to the partner.‡ In fact, the hearer is committed to the goal that the two dialogue interpretations meet, and this goal is naturally shared with the partner, in that speakers involved in a cooperative interaction want that their intersubjectivity is maintained.

Moreover, when, after the occurrence of a repair, the interactants have finally restored the intersubjectivity, they can find out that some utterances expressed between the trouble source turn and the repair turn may be no longer interesting. So, the pending intentions that were created by them should be considered irrelevant, too, and no reply to them would be expected.

Some more words must be spent about the recognition of the turns expressing a request for a repair: differently from the recognition of a misunderstanding, which is triggered by a failure in the interpretation of the last turn, this type of recognition takes place in the standard interpretation process. The interpreting agent B accepts this topic shift since the dialogue could not go on anymore and the partner is still collaborating with him in some way.

4 Example

We will show how our model works on Example 2 of section 1.2: in the interaction, Loes, who is denoted as B, is the receptionist and the keeper of supplies. Marty is denoted as A.

† In some cases, agents recognize that their interlocutors have misunderstood them but, for politeness reasons, they let the interaction go on, without making any repairs. In general, any planned action (in this case "Restructure") can be discarded if it is in conflict with other goals of the agents. In fact, their behavior is influenced by several factors, like the relationship among the interactants, and how much the misunderstanding can influence the subsequent talk. Here, we don't deal with these side behaviors.

‡ In principle, a notification is necessary only when the agent has produced some turn which could mislead his partner. It is a limitation of our model the fact that it always prescribes an explicit notification. The notification goals are managed by conditional steps of the "Satisfy" action (see the rightmost steps in Figure 1): "Satisfy(A, A, Know(B, done(Restructure(...))))" and
"Satisfy(A, A, Know(B, ¬achievable(inter(A, ..., ctx) ∧ inter(B, ..., ctx') ∧ equal(ctx, ctx'))))".

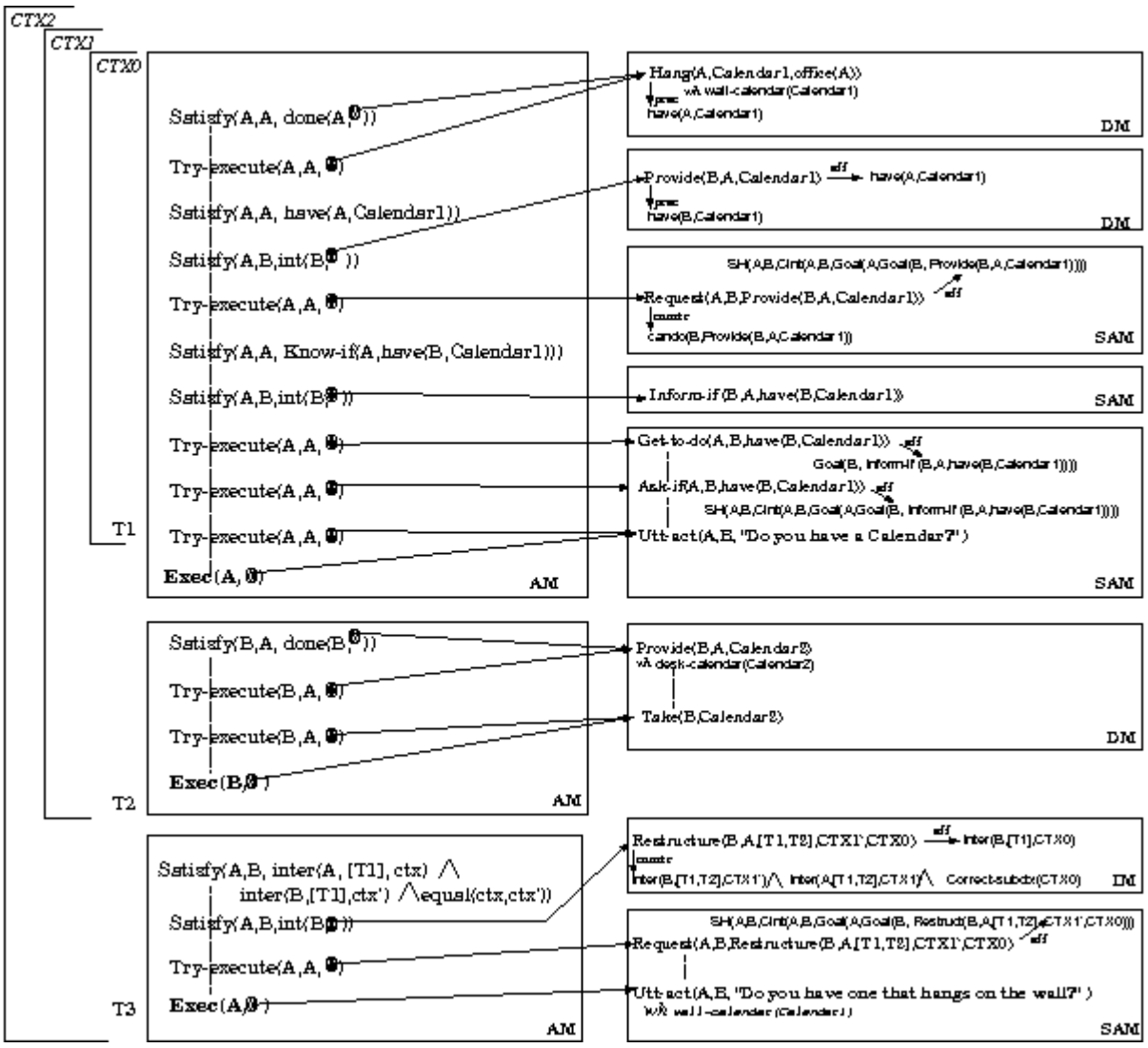

FIGURE 3. Interpretation of Example 2 from Martin's (A's) point of view.

T1: A: *"Loes, do you have a calendar,"*
T2: B: *"Yeah"* ((reaches for her desk calendar))
T3: A: *"Do you have one that hangs on the wall?"*
T4: B: *"Oh, you want one."*
T5: A: *"Yeah"*

In this interaction, there are two repairs directed towards two distinct instances of misunderstandings; the first one is a third position repair in T3 and the second one is a (different) self-repair in T4, due to the fact that B recognizes a further misunderstanding from the request for repair in T3.

Figures 3, 4 and 5 will be used in the following to describe the different interpretations of this interaction by the two speakers. In the figures, we only show the most important actions of the rightmost path in the recognized Agent Modeling plans. As a fact, the structures would contain other previous AM actions, so we only report the interpretation

branch which contains the pending goals. We will now analyze the dialogue from the different points of view of the two speakers.

Figure 3 shows A's point of view: in CTX0, A has the goal of hanging a calendar in his office ("Satisfy(A, A, done(A, Hang(A, Calendar1), office(A)))"). In order to do that, he must satisfy the precondition of action "Hang", that is "have(A, Calendar1)" (see the DM containing "Hang(A, Calendar1, office(A))" in the figure); so he starts another "Satisfy" action on the goal of having a calendar. The execution of this AM action leads A to plan an adequate object level action to induce B to provide him with the calendar (the "Find-plan" AM action is not shown in Figure 3). A request is a good action for that, but it is not very polite; so, A refers indirectly to his intentions by checking whether B can execute the action (i.e. whether she has any calendars; see the action "Satisfy(A, B, Knowif(A, have(B, Calendar1)))"). In other words, A hides his intention to execute the request by performing a question ("Ask-if(A, B, have(B, Calendar1))"), from which B can infer A's intention (this way of acting is a "pre-request", an indirect politeness strategy described in Levinson (1983)).

We now analyze B's (Loes') view of the dialogue, shown in Figure 4. While B does not have any problem in interpreting A's question as a request for a calendar, she misinterprets her partner's high level intentions underlying the underspecified sentence (see CTX0″): she thinks that A wants to borrow the calendar to have a look at it (the underlying domain action is "Read"). As a consequence of her misinterpretation, she chooses a calendar of the wrong type, a desk one: Figure 4 shows her action by associating the ("wh") restriction "desk-calendar(Calendar2)" to action "Provide(B, A, Calendar2)" in context CTX1″.
Since B is collaborative, she adopts A's still unexpressed intention that she gives him a calendar: this fact has been denoted in the figure by means of the dotted arc that relates the effect of the request ("SH(A, B, Cint(A, B, Goal(A, Goal(B, Provide(B, A, Calendar2)))))") with action "Satisfy(B, A, done(B, Provide(B, A, Calendar2)))".

Going back to Figure 3, we can examine how A interprets the observation that B is taking a calendar ("Exec(B, Take(B, Calendar2))"). The interpretation heuristics mentioned in section 3.2 are used to perform the interpretation task: instead of identifying all the possible reasons for taking a calendar, the interpretation process tries to find a relation between A's goals and the new contribution. Starting from the low-level goals, taking something has nothing to do with A's explicitly expressed goal of being informed about something; however, the heuristics relate A's implicit request that B provides him with a calendar with the fact that B has just taken one: to provide an agent with something, you have to take the thing and pass it to the receiver.
Unfortunately, B's action can not be related with A's dialogue context CTX0 because B is handling the wrong type of calendar; so, the coherence-seeking procedure fails and, in CTX1, the new turn remains unrelated from subcontext CTX0.

At this point, A tries to restore the intersubjectivity by planning a "Satisfy" action on the goal that A and B have the same interpretation context on a subpart of the dialogue.†
This results in the instantiation of an action
"Satisfy(A, B, inter(A, [T1], ctx) ∧ inter(B, [T1], ctx′) ∧ equal(ctx, ctx′))".
A plans a "Restructure" action and evaluates its constraints (in doing so, he applies the

† The subpart is initially unspecified, but, in this case, the only possible dialogue context that A and B can share is the one composed of the interpretation of T1: in fact, T2 has been produced from an incorrect understanding of T1.

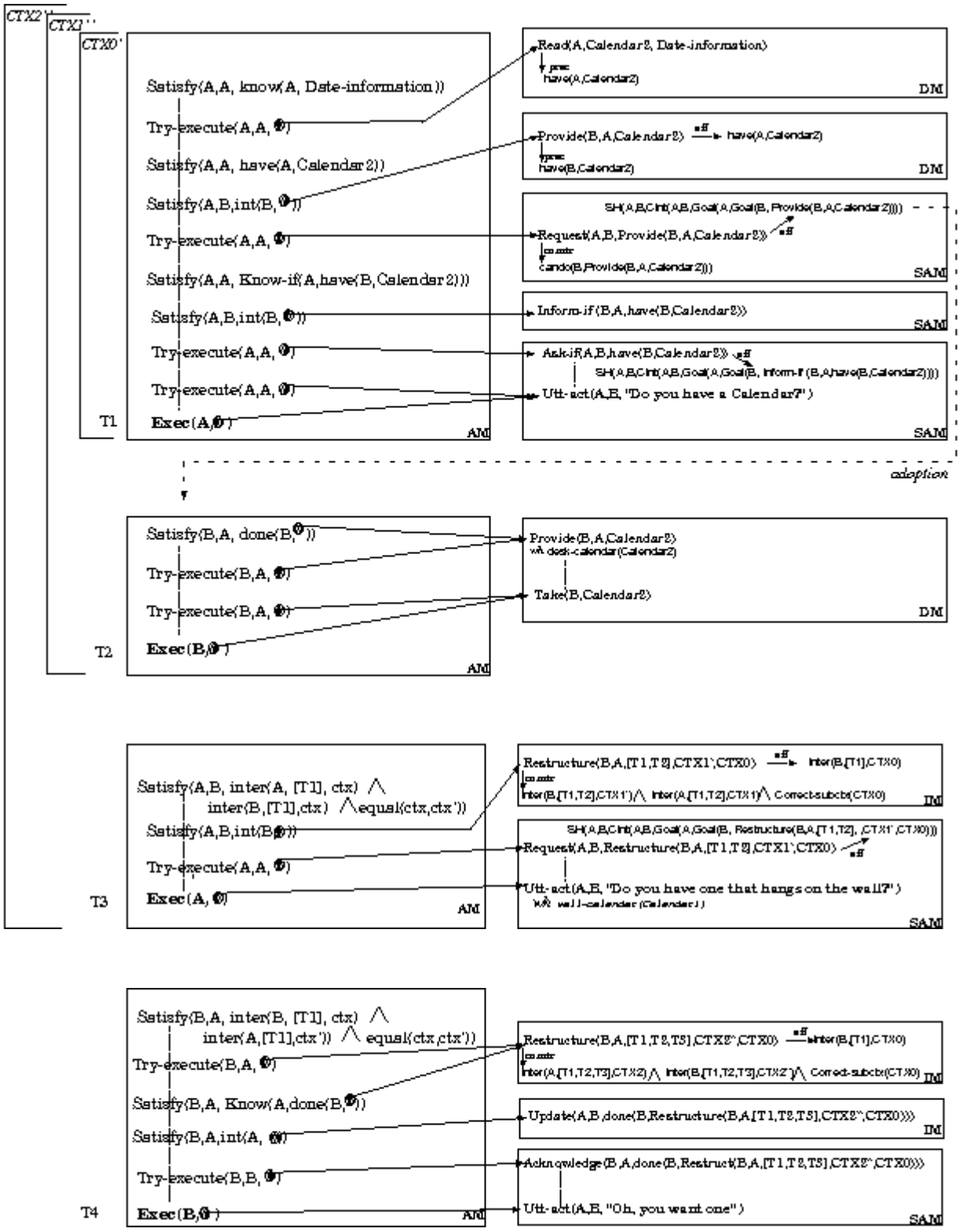

FIGURE 4. Interpretation of Example 2 from Loes' (B's) point of view.

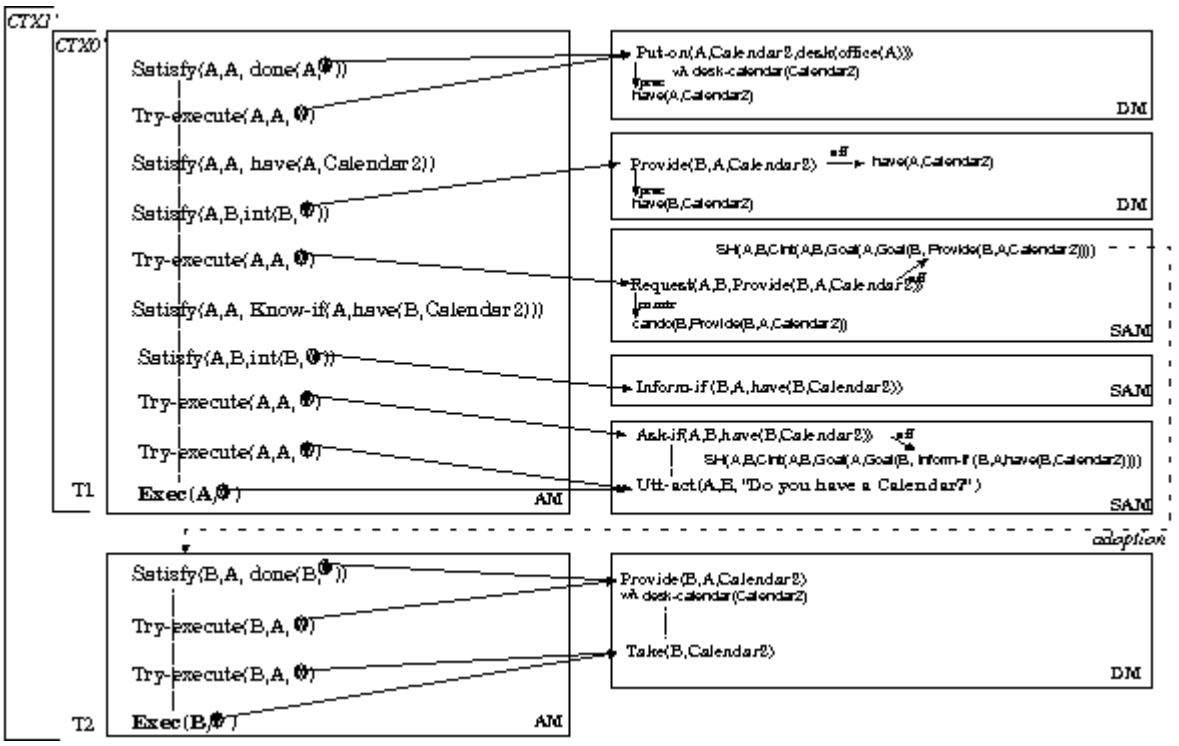

FIGURE 5. The reconstruction of turns T1 and T2 from Martin's point of view.

algorithm described in section 3.3, which, in this case, only makes him reinterpret locally T1 and relate it with T2).
A finds an alternative interpretation of his own turn T1 which explains T2 by means of an adoption relation (see CTX1′ in Figure 5). So, he assigns the responsibility of the misalignment to B. In the interpretation of CTX1′, which is still different from what B really has in mind, A attributes to B the belief that he wants to put a calendar on his desk. However, A has chosen the wrong hypothesis; since he is unaware of his error, he consequently formulates a request for repair ("Restructure(B, A, [T1,T2], CTX1′, CTX0)"): he exploits the strategy of repeating the first turn with the addition of some more specific information about the type of calendar he wants. The new information serves to enable the partner to disambiguate between the two alternative interpretations (CTX0 and CTX0′).

Going back to B's view of the interaction (see CTX2″ in Figure 4), she successfully recognizes T3 as an unrelated turn, which is a repair following an interactional breakdown. However, in this case, it is a justified topic shift: in fact, this turn satisfies the joint goal of maintaining the intersubjectivity in the dialogue.

After B interprets T3, she finds out that the beliefs underlying her interpretation context CTX2″ are inconsistent with A's request for restructuring. While B reconstructs T3 with the newly conveyed information about the desired calendar type, she understands what A thinks that she has in mind. In fact, she can recognize the "Restructure" action in A's plan only by hypothesizing that A believes that the constraints of such an action are

true: in particular, she has to reconstruct A's alternative context interpretation CTX1′ (Figure 5); she can do so thanks to the fact that A has specified that the calendar should be hung on the wall, in order to contrast with the "Put-on" interpretation (A believes that constraint "inter(B, [T1,T2], CTX1′)" is true). At the same time, B succeeds in identifying A's original intention, displayed in T1: hanging the calendar on the wall (A believes that the other constraint "inter(A, [T1,T2], CTX1)" is true).
On the basis of all this information, B adopts the goal of resolving the complex misunderstanding: she restructures her interpretation context by changing it from CTX0″ to CTX0, as a consequence of the effect of "Restructure(B, A, [T1,T2,T3], CTX2″, CTX0)" and then she notifies A that she has understood what he meant initially (turn T4). The notification is performed by means of a "Satisfy" action on the goal of letting A know that a "Restructure" action has been successfully executed.

It must be noted that B could not accept A's repair of turn T3; in fact, this would have mislead A by inducing him to believe that B's dialogue context really corresponded to CTX1′. Instead, from her acknowledgment, A understands that B's "Restructure" action is not the one requested by him in T3: T4 (*"Oh, you want one"*) means that B has realized only at this point that A wanted to keep the calendar for himself, instead of borrowing it. Now, the dialogue can go on correctly and B can provide A with the calendar to be hung on the wall.

For Schegloff (1992), Marty's controversial turn T3 is a follow-up question that specifies his intentions, rather than a repair for correcting Loes' interpretation. However, Schegloff does not exclude the other interpretation, since he admits the existence of repairs which do not have the recurrent linguistic features characterizing their canonical form: turn T3 could be a polite form of repair where, instead of correcting Loes explicitly, Marty responds to her with another question that addresses the interpretation problem. Anyway, Schegloff also points out the fact that, if Marty's turn T3 were a repair, it would be addressed to a different interpretation problem (calendar type) from the one which is addressed by Loes' fourth turn (keeping or borrowing). In our model, the double misunderstanding interpretation, that Schegloff partially admits, is not problematic: in fact, thanks to the planning context in which it is embedded, the role of T3 in triggering T4 is recognized.

5 Evidence from a corpus of conversational data

In order to fully understand the phenomenon to be modeled, we based our work on the analysis of a corpus of data collected from natural spoken conversations. The corpus contains 63 instances of misunderstandings; all the excerpts are in Italian, except two instances taken from an English transcription.†

The analysis of the data well evidenced the complexity of the phenomenon and highlighted some important features of natural occurring misunderstandings, such as the flex-

† The transcriptions used for the research consist of two published corpora of spoken Italian, and two unpublished minor corpora, plus a few examples transcripted for the specific purpose. de Mauro *et al.* (1993) contains various types of interactions, while Gavioli & Mansfield (1990) is a collection of book shop encounters (with a portion in English). The unpublished corpora were collected in 1994 by C. Ferrus and G. Fornara for their theses: the first one contains only instances of misunderstandings of proper nouns, the second one contains transcriptions of job interviews.

ibility of repair mechanisms in human conversation, the need to take different levels of misunderstanding into account and the nature of resources exploited by human conversants in accomplishing the repair task. All these features served as guidelines during the development of the model, hopefully contributing to make its treatment closer to reality.

First of all, the empirical data show the impact of misunderstanding on dialogue, and confirm, apart from the theoretical expectations, the need for a treatment of this phenomenon in a computational model of dialogue. In fact, the episodes of misunderstanding found in 59 hours and 12 minutes of conversation amount to 52, with a quite high rate of occurrence of one every 68.3 minutes approximately. This datum justifies the intuition that, even if they can be considered a sporadic event, misunderstandings do still play a role in conversation and consequently need to be treated by an appropriate mechanism (see also Perlis & Purang (1996)).

By observing the misunderstood turns, for example, it is possible to attempt a distinction concerning *what* is misunderstood. With reference to the process of interpretation of an utterance, five levels can be identified as possible object of misunderstanding, namely the phonetic level, the syntactic level, the lexical level, the semantic level and the pragmatic level.† If we consider the origin level of misunderstandings, we have the following distribution: the semantic level gathers the highest amount of instances, with 54% of the total, followed by the pragmatic level (19%), the phonetic level (14%) and the lexical level (13%), while no instances were found belonging to the syntactic level. More precisely, looking at the internal composition of the semantic level, divided into propositional content and reference (including wrong reference to the allocutor, as evidenced in Schegloff's data) the rate importance of the whole category can be attributed to the misunderstandings of the referent intended by the speaker. As far as the pragmatic level is concerned, it includes several types of misunderstandings that hadn't been highlighted before. At this level, misunderstandings concern not only the comprehension of the illocutionary force, but also the speaker's plan, the role of participants, the grasping of the indirect meaning, the choice of the right interpretation context and the discourse topic. Interestingly, only a few instances (3 instances) were found in which the hearer's reaction was misunderstood, generating a wrong repair, i.e. a repair that it was not necessary.

Regarding the causes of misunderstanding the analysis of the real examples underlines the substantial role played by ambiguity in allowing for wrong interpretations through all the levels (see Zaefferer (1977), Vendler (1994)). In detail, the responsibility for misunderstanding can be attributed to the presence of a relevant ambiguity at one of the identified levels in approximately two thirds of the instances, while only one third takes place in complete absence of ambiguity (e.g. in the presence of mishearings and other phenomena like misconceptions).

Finally, it is interesting to have a look at the position of repair with respect to the misunderstood turn. The evidence from the corpus suggests that, while most repairs are executed within two or three turns of distance from the misunderstood turn, a significant set of examples shows that repairs can be delayed at up to several turns of distance from the point where the misunderstanding has occurred. There are examples of repair at the 12th, 15th, 19th and even later, at 31st turn, confirming the data collected in Hansen *et al.* (1996) about the possibility for the speaker to delay the repair. Moreover, most of the analyzed repairs were performed by the misunderstood speaker on the under-

† The presence of different levels of misunderstanding has been pointed out by many researchers, like Zaefferer (1977), Dascal (1989), Schegloff (1992), Vendler (1994) and Bazzanella & Damiano (1997).

standing shown by the other participant (i.e. third turn repair), while only three instances can be classified as repairs initiated by one participant on his own interpretation of a previous turn. So, although this datum confirms the preference for third position repair, it underlines the need for a flexible diagnosis of misunderstanding, which doesn't rely only on a pre-ordered sequence of turns, but which is able to go backward in the dialogue in search for a misunderstood element that may reside in any previous positions and at any possible level of meaning.

6 Related work

First of all, some comments must be made about our representation of the context of a dialogue: our Context Models (indicated as CTXi in the figures) are rather different from the dialogue model described in Lambert & Carberry (1991), Carberry *et al.* (1992) and Lambert (1993). As discussed in Ardissono *et al.* (1996), although both frameworks contain domain, communicative and metalevel plans, we make a different use of our (AM) metalevel and communicative plans. In Lambert's model, the problem solving plans relate communicative actions to the domain level plans pursued by the speakers; instead, in our approach, both domain level and linguistic actions are the objects of the metalevel actions and can be performed interchangeably. Moreover, our structures can be extended for answer production, adding them some supplementary information about how to execute actions (Ardissono & Cohen 1996).

As far as the management of dialogue is concerned, other frameworks have exploited the view of dialogue as a collaborative activity among agents. For example, Chu-Carroll & Carberry (1995a), Chu-Carroll & Carberry (1995b) describe a framework where a system and its user collaborate to build a (correct) plan for obtaining the user's goals: in an interaction, the system and the user negotiate their beliefs, until they agree on a solution. In that work, however, no negotiation of the *meaning* of sentences is performed.

Dialogue has been modeled as a collaborative activity also in Heeman & Hirst (1995), where a planning framework has been used to interpret and recognize referential expressions. In this model, the conversational turns are seen as actions that contribute to the interaction by establishing some communicative goals, which become shared between the interactants. Although our work is based on the same idea, it differs in the way how it models negotiation among agents. In particular, Heeman and Hirst introduce specific actions to plan the form of a referential expression, and a number of actions which describe the reactions which an agent can adopt when he receives a referential expression: e.g. the agent can accept it, refuse it, correct it, and so on, depending on the fact that he has been able to resolve the expression, or he has encountered any troubles in the interpretation process. In any case, he starts collaborating with his partner at the resolution of the problem, by means of a negotiation subdialogue.

In our work, the interpretation of referential expressions is not separated with respect to the other linguistic phenomena: we describe a generic framework of interaction, where the various interpretation levels (syntactic, semantic, etc.) are modeled in a unique way, by means of the execution of object level interpretation actions, under the control of the Agent Modeling actions which manage the reaction to a possible failure. The idea is that an acceptance, a refusal, or other, may happen at different levels, not only at the referential one: for example, an ambiguous word may cause the hearer to build more than

one semantic interpretation of a sentence, so that he would have to start a clarification subdialogue with the speaker to establish the intended meaning of the utterance; in fact, this is quite similar to the case of an underspecified referential expression. For this reason, we adopt a unique (metalevel) model for handling action execution and failure: this model rules the execution of domain actions, as well as speech acts and interpretation actions. During the interpretation of an utterance, it supports the possible reactions to a failure (in terms of notification to the partner) at the utterance level, as well as at the locutionary and illocutionary ones.

Our representation of the utterance interpretation process as multiple object level actions does not only offer a framework for treating interpretation failures, but also for managing the phenomenon of *grounding* in dialogue (Traum & Hinkelman 1992): we treat the acknowledgments that the receiver of an utterance has interpreted it successfully and accepted it (e.g. *"Okay."*) exactly in the same way as the acknowledgements of successful execution of a domain action, as prescribed by our Agent Modeling plans; similarly, the occurrence of requests for an acknowledgement by the speaker of an utterance (e.g. *"Right?"*, *"Ok?"*) are collapsed into the process by which an agent monitors that his partner has executed a requested action successfully: in the case of linguistic actions, this corresponds to checking whether the partner has been able to perform the interpretation of the utterance and the updating of his own beliefs with the communicative effects of the utterance itself.

Some computational models of dialogue use a notion of coherence based on an analysis of the expected behavior of agents in conversation. In those approaches (for instance, consider McRoy & Hirst (1995), Traum & Hinkelman (1992), Traum & Allen (1994), Danieli (1996)), the speech acts occurring in the last conversational turn, together with the existing dialogue context, are used to predict which speech acts the interlocutor should perform if the interaction goes well; a deviance from the expected behavior is taken as a sign that some interaction problem is occurring and the presence of a misunderstanding is hypothesized.

Our work uses a deeper notion of context, where different types of intentions (related with the interaction, as well as with the agent's domain activity and the goals deriving from his participating to a conversation) are represented explicitly and maintained as a context; they contribute to the identification of more general relations that can exist with the new input. In particular, we model linguistic expectations and underlying intentions uniformly as goals, although they fall at different levels. So, we provide a unified model of the dialogue context, where the pending goals which happen to correspond to the linguistic expectations fall at a lower level with respect to those pertaining the underlying activity of the agents. An immediate consequence of this is that, when a turn satisfies some low-level pending goals, there is no need to inspect the higher level ones (so, the effort spent in the interpretation task is limited). At the same time, however, when the low-level goals are not matched, the dialogue context is rich enough to be analyzed, searching for the possible relations between the turn and the previous part of the interaction.

Our model also differs from the above mentioned systems because of the two-level organization of our plan libraries, where the agent modeling plans rule all other actions, including interpretation and repair actions. In particular, in McRoy & Hirst (1995), metalevel plans model the expected continuations of an interaction: following the ideas developed in Litman & Allen (1987), the strongest expectation is that the receiver of a turn accepts

the turn and reacts by contributing to it, but he might also start another (sub)dialogue, by introducing a new communicative goal. McRoy and Hirst extend Litman's approach and introduce metaplans to diagnose misunderstandings and formulate repairs, when the expected behavior is violated. However, their metaplans only analyze the surface expectations introduced by performing of a speech act; the absence of a deeper intentional analysis limits their approach to the treatment of misunderstandings on speech acts, while our model also treats misunderstandings on domain level actions (see the analysis of the example in section 4).

7 Conclusions

In this paper, we have described an agent model that provides a computational treatment of misunderstandings in dialogue. We have focused on third and fourth position repairs, and we have described how they can be detected when a turn incoherent with the dialogue context occurs. On the contrary, we have left apart first and second turn repairs, because they are rather different phenomena: third and forth position repairs are caused by the recognition that one of the interactants has committed to a wrong interpretation of the previous part of the dialogue (they are concerned with interpretation mistakes which occurred before the repair turn); instead, first turn repairs are performed by a speaker in order to avoid that the hearer has interpretation problems; finally, second turn repairs are typically performed by the receiver of a turn, when he experiences some problems in the interpretation of his partner's turn, so he has not yet committed to a specific interpretation (e.g. they are usually related with misspellings, and other similar phenomena).

Our dialogue model consists of a plan-based representation of the knowledge about the way to reach one's own goals. This knowledge takes into account the normal planning behavior of an agent, as well as his capability to recover from problems in the interaction with other agents. Misunderstandings are treated as one of the specific problems that can occur when interacting with other agents and are dealt with by adopting the goal of maintaining the safety of the intersubjectivity in dialogue. In summary, the main ideas underlying our work are:

1) Misunderstandings in dialogue are explained in the same way as problems occurring in non-linguistic interaction. In fact, in our agent model, linguistic actions are means to obtain one's goals exactly as the other (domain level) actions. The maintenance of such a general interaction context allows us to model misunderstandings without posing any limitations to the object of the misunderstanding and to the distance between the misunderstood turns and the repair turns performed by the agents.

2) Different levels of interpretation have been identified as the possible objects of a misunderstanding, from the utterance level (which regards the syntactic and semantic interpretation), to the pragmatic level, which covers misunderstandings on the illocutionary force of speech acts, and on the domain level activity underlying the linguistic behavior of agents.†

3) The recognition of misunderstandings follows from an agent's attempt to resume the intersubjectivity, after he has not been able to interpret the last turn coherently. How-

† Although our model is compatible with the presence of different levels of misunderstanding, we don't consider mishearings: they are important for speech recognition, as discussed in Smith & Hipp (1994) and Danieli (1996), but they are not a problem for our system, because it takes its input in textual form.

ever, the identification of the mistake and the reinterpretation of the dialogue are fully embedded into his goal-directed behavior, in that the agents participating to an interaction share the common goal of maintaining their intersubjectivity and act to restore it, when threatened. Restoring intersubjectivity is a goal shared among the interactants because the intersubjectivity is the basis for any *social action* (Schegloff 1992).

4) There is a clear distinction among repairs to misunderstandings (i.e. requests of repair) and notifications that one of the interactants has just adjusted his dialogue context to recover from a misunderstanding.

5) Repairs can be the object of a misunderstanding, too; moreover, an agent can believe that he has been misunderstood by another one, so generating a repair, but he might be wrong (so, the other one could make another repair). Our model takes into account these phenomena, as it can be seen in the example described in section 4.

Our agent model is implemented in Common Lisp and runs on workstations. The interpretation of utterances is fully implemented, starting from the NL (Italian) form, to the construction of the dialogue context of the interaction; the generation of agent behavior is under development.

Some aspects of our model have to be fully developed and represent an avenue for future work:

1) A deeper study of the structure of dialogue, as mentioned in section 1.1, in order to analyze subdialogues and topic shifts. The cases where a topic shift is an acceptable interpretation hypothesis, even in a cooperative environment, can be determined by the satisfaction of the previous goals or by the urgency of the newly introduced goals (e.g. consider the problem of managing interactions in highly dynamic environments, as studied in the reactive planning research).

2) The development of the linguistic strategies to perform a repair (in terms of sequences of speech acts (Schegloff 1992)). An interesting related problem is to find the minimal information necessary to disambiguate the meaning of the trouble source turn when a misunderstanding occurs. The idea of finding some discriminating information for disambiguation purposes recalls the strategies used in plan recognition to deal with ambiguous hypotheses on the observed agent's plans (e.g. see the initiation of clarification dialogues to disambiguate the partner contribution (van Beek & Cohen 1991, Cohen *et al.* 1994)).

3) The ability of the speakers to talk about what is happening in dialogue, by means of anaphoras, verbs and nouns referring to speech acts (such as "to criticize", "to order", etc.); Goy & Lesmo (1997) provide a first analysis of the lexical semantics of communication verbs.

4) The development of a method to analyze the "weaknesses" of one's own utterances (in terms of ambiguities and underspecification). Empirical data show that humans identify the causes of the misunderstanding very easily, and that they repair them in an effective way by "filling the gaps" of their previous turns. We believe that the type of failure arising when an agent tries to relate the incoherent turn to the previous context could be useful for an efficient search of the alternative interpretation for the misunderstood turn.

5) The ability to prevent misunderstandings by performing early repairs, like first-turn repairs. While we have not worked at this problem, there are some dialogue models which try to predict how the hearer will interpret an utterance, in his mental state. For example, Ndiaye & Jameson (1996) use the concept of *anticipation feedback loops*

(introduced in Wahlster & Kobsa (1989)) to choose the "most promising" utterance that can be expressed to the user, in order to convey an information. Although this does not model the occurrence of utterances immediately followed by repairs (it is a way to plan what to say), the idea underlying anticipation feedback loops could be exploited to recognize first turn repairs in the interpretation of a turn.

6) The ability of a speaker to identify a misconception of the partner. Usually misconceptions make it impossible to interpret a turn locally (e.g. consider the problems studied in McCoy (1986) and McCoy (1988)); however, if they aren't promptly recognized, they can lead to misunderstandings. When an agent looks for an alternative interpretation of the previous dialogue, he should take into account also the possibility that a misconception of the partner led him to a different understanding of what has been said.†

7) It also must be noticed that our analysis of coherence does not take into account the role of cognitive load in dialogue interpretation. As evident from the analyzed corpora, speakers often misinterpret complex utterances, or they forget what has been previously said (Jordan & Thomason 1996). However, this problem is currently outside our interests.

Acknowledgements

We are very grateful to Leonardo Lesmo, Carla Bazzanella and Morena Danieli for the fruitful discussions and good advice that they provided us with. We also thank them and the anonymous reviewers for having carefully helped us to improve our paper with their comments on the first version. Finally, we thank Sandra Carberry for her comments to the agent modeling architecture which represents the conceptual framework of this work. This work was partially supported by MURST 60% and by the Italian National Research Council (CNR), project "Pianificazione e riconoscimento di piani nella comunicazione". In particular, the analysis of the corpora which guided us in the construction of our model of misunderstanding was carried out within this project by Carla Bazzanella and her students, and we thank them a lot for their collaboration.

References

G. Airenti, B. Bara, & M. Colombetti (1993). Conversational and behavior games in the pragmatics of discourse. *Cognitive Science*, 17:197–256.

J.F. Allen & C.R. Perrault (1980). Analyzing intention in utterances. *Artificial Intelligence*, 15:143–178.

J.F. Allen (1983). Recognizing intentions from natural language utterances. In M. Brady & R.C. Berwick, editors, *Computational models of discourse*, pages 107–166. MIT Press.

L. Ardissono & R. Cohen (1996). Extending the role of user feedback in plan recognition and response generation for advice-giving systems: an initial report. In *Lecture Notes in Artificial Intelligence n. 1081: Advances in Artificial Intelligence*, pages 109–120. Springer Verlag, Berlin.

† For an analysis of misconceptions, see Goodman (1985), Pollack (1987), Pollack (1990), Calistri-Yeh (1991), Carberry (1986), Carberry (1987), Eller & Carberry (1992), Eller (1993), Donaldson & Cohen (1996). Note that plan misconceptions have been implicitly reduced to the presence of different interpretation contexts in all the approaches based on the presence of *buggy plan libraries*. This suggests that a basic model of misunderstandings can be extended to the treatment of misconceptions, by considering also the buggy plans as alternatives.

L. Ardissono & D. Sestero (1996). Using dynamic user models in the recognition of the plans of the user. *User Modeling and User-Adapted Interaction*, 5(2):157–190.

L. Ardissono, A. Lombardo, & D. Sestero (1993). A flexible approach to cooperative response generation in information-seeking dialogues. In *Proc. 31st Annual Meeting ACL*, pages 274–276, Columbus.

L. Ardissono, G. Boella, & L. Lesmo (1995a). A computational approach to speech acts recognition. In *Proc. 17th Cognitive Science Conference*, pages 316–321, Pittsburgh.

L. Ardissono, G. Boella, & D. Sestero (1995b). Recognition of preliminary sentences in dialogue interpretation. In *Lecture Notes in Artificial Intelligence n. 992: Topics in Artificial Intelligence*, pages 139–144. Springer Verlag, Berlin.

L. Ardissono, G. Boella, & L. Lesmo (1996). Recognition of problem-solving plans in dialogue interpretation. In *Proc. 5th Int. Conf. on User Modeling*, pages 195–197, Kailua-Kona, Hawaii.

J.L. Austin (1962). *How to Do Things with Words*. Harvard University Press, Cambridge, Mass.

P. Barboni & D. Sestero (1997). Flexible response choice using problem-solving plans and rethorical relations. In *Proc. 5th Congresso della Associazione Italiana per l'Intelligenza Artificiale*, Roma, Italy. to appear .

C. Bazzanella & R. Damiano (1997). Il fraintendimento linguistico nelle interazioni quotidiane: proposte di classificazione. *Lingua e Stile*, 32(3).

S. Blum-Kulka & E. Weizman (1988). The inevitability of misunderstanding: discourse ambiguities. *Text*, 8(3):219–241.

M.E. Bratman (1990). What is intention? In P.R. Cohen, J. Morgan, & M.E. Pollack, editors, *Intentions in communication*, pages 15–32. MIT Press.

P. Brown & S. C. Levinson (1987). *Politeness: some universals on language usage*. Cambridge University Press, Cambridge.

R.J. Calistri-Yeh (1991). Utilizing user models to handle ambiguity and misconceptions in robust plan recognition. *User Modeling and User-Adapted Interaction*, 1:289–322.

S. Carberry, Z. Kazi, & L. Lambert (1992). Modeling discourse, problem-solving, and domain goals incrementally in task-oriented dialogue. In *Proc 3rd Int. Workshop on User Modeling*, pages 192–201, Wadern.

S. Carberry (1986). User models: the problem of disparity. In *Proc. 11th COLING*, pages 29–34, Bonn.

S. Carberry (1987). Pragmatic modeling: toward a robust natural language interface. *Computational Intelligence*, 3:117–136.

S. Carberry (1990). *Plan Recognition in Natural Language Dialogue*. MIT Press.

C. Castelfranchi & D. Parisi (1980). *Linguaggio, conoscenze e scopi*. Il Mulino, Bologna.

C. Castelfranchi (1992). No more cooperation, please! In search of the social structure of verbal interaction. In J. Slack. A. Ortony & O. Stock, editors, *Communication from an Artificial Intelligence perspective. Theoretical and Applied Issues*. Springer Verlag, Berlin.

J. Chu-Carroll & S. Carberry (1995a). Communication for conflict resolution in multi-agent collaborative planning. In *Proc. First International Conference on Multiagent Systems*, pages 49–56.

J. Chu-Carroll & S. Carberry (1995b). Response generation in collaborative negotiation. In *Proc. 33rd Annual Meeting of the ACL*, pages 136–143.

H.H. Clark & E.F. Schaefer (1989). Contributing to discourse. *Cognitive Science*, 13:259–294.

P.R. Cohen & H.J. Levesque (1990a). Intention is choice with commitment. *Artificial Intelligence*, 42:213–261.

P.R. Cohen & H.J. Levesque (1990b). Rational interaction as the basis for communication. In P.R. Cohen, J. Morgan, & M.E. Pollack, editors, *Intentions in communication*, pages 221–255. MIT Press.

P.R. Cohen & H.J. Levesque (1991). Confirmation and joint action. In *Proc. 12th IJCAI*, pages 951–957, Sydney.

P.R. Cohen, C.R. Perrault, & J.F. Allen (1981). Beyond question answering. In W. Lehnert & M. Ringle, editors, *Strategies for Natural Language Processing*, pages 245–274. Lawrence Erlbaum, Hillsdale. NJ.

R. Cohen, K. Schmidt, & P. van Beek (1994). A framework for soliciting clarification from users during plan recognition. In *Proc. 4th Int. Conf. on User Modeling*, pages 11–17, Hyannis, MA.

R. Cohen (1984). A computational theory of the function of clue words in argument understanding. In *Proc. 10th International Conference on Computational Linguistics*, pages 251–258, Stanford, CA.

R. Cohen (1987). Analyzing the structure of argumentative discourse. *Computational Linguistics*, 13:11–24.

M. Danieli (1996). On the use of expectations for detecting and repairing human-machine misomunicaiton. In *AAAI 1996 Workshop: Detecting, Repairing and Preventing Human-Machine Miscommunication*, pages 87–93, Portland.

M. Dascal (1989). The relevance of misunderstanding. In M. Dascal, editor, *Dialogue: an interdisciplinary approach*, pages 441–459. Benjamin, Amsterdam - Philadelphia.

T. de Mauro, F. Mancini, M. Vedovelli, & M. Voghera (1993). *Lessico di Frequenza dell'Italiano Parlato*. ETASLIBRI.

B. Di Eugenio & L. Lesmo (1987). Representation and interpretation of determiners in natural language. In *Proc. 10th IJCAI*, pages 648–653, Milano.

T. Donaldson & R. Cohen (1996). Addressing user misconceptions within a goal-oriented, turn-taking framework in dialogue. In *AAAI 1996 Workshop: Detecting, Repairing and Preventing Human-Machine Miscommunication*, pages 98–101, Portland.

R.M. Eller & S. Carberry (1992). A meta-rule approach to flexible plan recognitionin dialogue. *User Modeling and User-Adapted Interaction*, 2:27–54.

R.M. Eller (1993). *A Plan Recognition Architecture for Ill-Formed Dialogue*. Ph.D. thesis, University of Delaware.

Gavioli & Mansfield (1990). *The PIXI corpora: bookshop encounters in English and Italian*. CLUEB, Bologna, Italy.

B. Goodman (1985). Repairing reference identification failures by relaxation. In *Proc. 23th Annual Meeting ACL*, pages 204–217, Chicago.

A. Goy & L. Lesmo (1997). Integrating lexical semantics and pragmatics: The case of italian communication verbs. In *Proc. 2nd International Workshop on Computational Semantics*, Tillburg, The Netherlands.

B.J. Grosz & C.L. Sidner (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12:175–204.

B.J. Grosz & C.L. Sidner (1990). Plans for discourse. In P.R. Cohen, J. Morgan, & M.E. Pollack, editors, *Intentions in communication*, pages 417–443. MIT Press.

B. Hansen, D. Novick, & S. Sutton (1996). Prevention and repair of breackdowns in a simple task domain. In *AAAI 1996 Workshop: Detecting, Repairing and Preventing Human-Machine Miscommunication*, pages 5–12, Portland.

P.A. Heeman & G. Hirst (1995). Collaborating on referring expressions. *Computational Linguistics*, 21(3):353–382.

J.C. Heritage (1984). A change-of-state token and aspects of its sequential placement. In J.M.Atkinson & J.C.Heritage, editors, *Structures of Social Action*. Cambridge University Press.

G. Hirst, S. McRoy, P. Heeman, P. Edmonds, & D. Horton (1994). Repairing conversational misunderstandings and non-understandings. *Speech Communication*, 15:213–229.

J.R. Hobbs, M. Stickel, D.E. Appelt, & P. Martin (1993). Interpretation as abduction. *Artificial Intelligence*, 63:69–142.

P.W. Jordan & R.H. Thomason (1996). Refining the categories of miscommunication. In *AAAI 1996 Workshop: Detecting, Repairing and Preventing Human-Machine Miscommunication*, pages 13–20, Portland.

H. Kautz (1991). A formal theory of plan recognition and its implementation. In R.J. Brachman, editor, *Reasoning About Plans*, pages 69–125. Morgan Kaufmann Publishers.

L. Lambert & S. Carberry (1991). A tripartite plan-based model of dialogue. In *Proc. 29th Annual Meeting of ACL*, pages 47–54, Berkeley,CA.

L. Lambert (1993). *Recognizing Complex Discourse Acts: A Tripartite Plan-Based Model of Dialogue*. Ph.D. thesis, University of Delaware.

L. Lesmo & P. Torasso (1985). Analysis of conjunctions in a rule based parser. In *Proc. 23rd Annual Meeting ACL*, pages 180–187, Chicago.

S.C. Levinson (1981). The essential inadequacies of speech act models of dialogue. In M. Parret, M. Sbisa, & J. Verschueren, editors, *Possibilities and Limitations of Pragmatics*, pages 473–492. Benjamins, Amsterdam.

S.C. Levinson (1983). *Pragmatics*. Cambridge University Press, Cambridge.

D. Litman & J. Allen (1987). A plan recognition model for subdialogues in conversation. *Cognitive Science*, 11:163–200.

D. Litman (1986). Linguistic coherence: a plan-based alternative. In *Proc. 24th Annual Meeting of the ACL*, pages 215–223, New York, NY.

K.F. McCoy (1986). The ROMPER system: responding to object-related misconceptions using perspective. In *Proc. 24th Annual Meeting of the ACL*, pages 97–105, New York.

K.F. McCoy (1988). Reasoning on a highlighted user model to respond to misconceptions. *Computational Linguistics*, 14(3):52–63.

S.W. McRoy & G. Hirst (1995). The repair of speech act misunderstandings by abductive inference. *Computational Linguistics*, 21(4):433–478.

M. Merritt (1976). On question following question (in service encounters). *Language in Society*, V:315–357.

J.D. Moore (1995). The role of plans in discourse generation. In D. Everett & S.G. Thomason, editors, *Discourse: Linguistic, Computational and Philosophical Perspectives*.

L. Morgenstern (1987). Knowledge preconditions for actions and plans. In *Proc. 10th IJCAI*, Milano.

K. Nagao (1993). Abduction and dynamic preference in plan-based dialogue understanding. In *Proc. 13th IJCAI*, pages 1186–1192, Chambery.

A. Ndiaye & A. Jameson (1996). Predictive role taking in dialog: global anticipation feedback based on trasmutability. In *Proc. 5th Int. Conf. on User Modeling*, pages 137–144, Kailua-Kona, Hawaii.

D. Perlis & K. Purang (1996). Conversational adequacy: Mistakes are the essence. In *AAAI 1996 Workshop: Detecting, Repairing and Preventing Human-Machine Miscommunication*, pages 47–56, Portland.

C.R. Perrault (1990). An application of default logic to speech act theory. In P.R. Cohen, J. Morgan, & M.E. Pollack, editors, *Intentions in communication*, pages 160–185. MIT Press.

M.E. Pollack (1987). Some requirements for a model of the plan-inference process in conversation. In Ronan Reilly, editor, *Communication Failure in Dialogue*, pages 245–256. North-Holland, Amsterdam.

M.E. Pollack (1990). Plans as complex mental attitudes. In P.R. Cohen, J. Morgan, & M.E. Pollack, editors, *Intentions in communication*, pages 77–103. MIT Press.

A. Rao & M.P. Georgeff (1991). Modeling rational agents within a bdi-architecture. In *Proc. 2th Int. Conf. Principles of Knowledge Representation and Reasoning (KR:91)*, pages 473–484, Cambridge, MA.

H. Sacks, E.A. Schegloff, & G. Jefferson (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50:696–735.

E.A. Schegloff (1972). Sequencing in conversational opening. In J.J. Gumperz & D. Hymes, editors, *Directions in Sociolinguistics*. Rinehart and Winston, New York.

E.A. Schegloff (1987). Some sources of misunderstanding in talk-in-interaction. *Linguistics*, 25(1):201–218.

E.A. Schegloff (1992). Repair after the next turn: The last structurally provided defense of intersubjectivity in conversation. *American Journal of Sociology*, 7(5):1295–1345.

I.A. Smith & P.R. Cohen (1996). Toward a semantics for an agent communications language based on speech-acts. In *Proc. 14th Conf. AAAI*, pages 24–31, Portland.

R. Smith & D. Hipp (1994). *Spoken Natural Language Dialogue Systems: a Practical Approach*. Oxford University Press, New York.

D.R. Traum & J.F. Allen (1994). Discourse obligations in dialogue processing. In *Proc. 32nd Annual Meeting of ACL*, pages 1–8, Las Cruces, New Mexico.

D.R. Traum & E.A. Hinkelman (1992). Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3):575–599.

P. van Beek & R. Cohen (1991). Resolving plan ambiguity for cooperative response generation. In *Proc. 12th IJCAI*, pages 938–944, Sydney.

Z. Vendler (1994). Understanding misunderstanding. In D. Jamieson, editor, *Language, Mind and Art*, pages 9–21. Kluwer, Dodrecht-Boston-London.

W. Wahlster & A. Kobsa (1989). User models in dialog systems. In A. Kobsa & W. Wahlster, editors, *User Models in Dialog Systems*, pages 4–34. Springer Verlag, Berlin.

E. Weigand (1997). Misunderstanding: the standard case. *Journal of Pragmatics*. to appear .

D. Zaefferer (1977). Understanding misunderstanding: a proposal for an explanation of reading choices. *Journal of Pragmatics*, 329-346.