Representing Conversation Acts in a Unified Semantic/Pragmatic Framework

Massimo Poesio
University of Edinburgh
Centre for Cognitive Science
2 Buccleuch Place
Edinburgh EH8 9LW, Scotland, UK
poesio@cogsci.ed.ac.uk

David Traum
UMIACS, University of Maryland
A. V. Williams Building
College Park, MD 20742 USA
traum@cs.umd.edu

Introduction

We would like to argue for a unified formal theory to represent and allow automated reasoning about both the semantic and the pragmatic aspects of conversational context, particularly the common ground between conversational participants. Central to this theory is the hypothesis that speech act occurrences are recorded in the common ground and can serve as the basis for a model of language processing in context that reconciles in a general and systematic fashion the differences between the theories of the common ground adopted in current theories of discourse processing which focus more narrowly on sub-issues such as reference resolution, intention recognition, and dialogue management.

One example of poor fit between existing theories of context is the contrast between, on the one hand, linguistically motivated theories developed to account for the semantics of anaphora (e.g., (Kamp & Reyle 1993)); and on the other hand, the models of context proposed for intention recognition and dialogue management, whose emphasis is on capturing the effects of speech acts on the beliefs, intentions, and obligations of the participating agents (Allen 1983; Carberry 1990; Cohen & Levesque 1990; Perrault 1990; Traum & Allen 1994). These traditions resulted in very detailed proposals about context and context update;[1] but the resulting models of context differ significantly. It is not possible to simply adopt one or the other model. While the linguistically motivated theories of context integrate well with current theories of semantic interpretation, their relation with current work on planning and plan recognition is less clear; the opposite is true of theories of context based on actions and their effects.

We desire a unified model for several reasons. For one thing, we wish to develop theories of the *interpretation* of, discourse phenomena, such as definite descriptions, which requires information of both a semantic and pragmatic nature be brought to bear (semantic information about which discourse referents are accessible; pragmatic information about the current attentional state and how the current utterance relates to the intentional structure of the conversation) (Poesio 1993). Having all of this information within a single framework allows one to more easily combine different information sources and evaluate competing proposals. Likewise, semantic details of the contents of utterances and background contexts in which and about which they are uttered will also be important for speech act interpretation and disambiguation. Moreover, our model is meant to be one that can be used by an artificial agent engaging in conversations as an internal, on-line representation of context. While much has been written about individual contextual problems, many of the proposed representations are mutually incompatible, not usable by an agent involved in a conversation, or both.

Elsewhere, (Poesio & Traum 1997), we have written in more detail about the semantics of this model, and how we reconcile conflicting views on common ground into a unified model that can be used for interpreting definite referring expressions as well as tracking grounding by interpreting acknowledgments and repairs. Here we just present some of the flavor of that approach, concentrating on issues most directly related to communicative action.

Building Blocks

We would also like to convince our readers that much can be achieved with rather simple technical tools. We propose to show that certain pragmatic aspects of the common ground–specifically, the shared information about the DISCOURSE SITUATION (Barwise & Perry 1983), i.e., which speech acts have been performed in a conversation and their hierarchical organization–can be captured using technical tools that are essentially those of Discourse Representation Theory (DRT) (Kamp & Reyle 1993; Muskens 1994), whose 'dynamic' properties are also preserved. The theory of the common ground we propose gives us the tools to specify both semantic and (some) pragmatic aspects of the

[1] See, e.g., (Kamp & Reyle 1993) for the details of a semantic treatment of anaphoric reference, and the papers in (Cohen, Morgan, & Pollack 1990) for theories about the effect of speech acts on the mental state of agents.

common ground, and can therefore be used both to capture the meaning of utterances like *okay* and to formulate theories of interpretation. We have shown elsewhere that this minimal account can be extended to capture additional pragmatic aspects of the common ground (Poesio & Traum 1997).

Our theory of context can be sketchily characterized as follows. Our starting point is Stalnaker's observation ((Stalnaker 1979), p. 323):

> The fact that a speaker is speaking, saying the words he is saying in the way he is saying them, is a fact that is usually accessible to everyone present. Such observed facts can be expected to change the presumed common background knowledge of the speaker and his audience in the same way that any obviously observable change in the physical surroundings of the conversation will change the presumed common knowledge.

In other words, that the common ground includes information about which speech acts (both locutionary and illocutionary) took place: we will call this the discourse situation. This is, of course, the assumption behind much formal work on context in artificial intelligence. The question is, what is the connection between this work and work on context in linguistics? Our first hypothesis is that the discourse situation can be characterized by a DRS, recording (shared information about the) speech acts that took place. Our second hypothesis is that as speech acts are events, this information can be captured using the tools for representing events introduced in (Kamp & Reyle 1993). The discourse situation resulting from the dialog in (1) can thus be characterized as in (2).

(1)
A: There is an engine at Avon.
B: *It* is hooked to a boxcar.

(2)

$$\boxed{\begin{array}{l} ce1\ ce2\ s\ s'\ s'' \\ \hline ce1: \ \mathbf{asrt}(A,B, \boxed{\begin{array}{l} x\ w\ e \\ \hline \mathbf{engine}(x) \\ \mathbf{Avon}(w) \\ e: \boxed{\mathbf{at}(x,w)} \end{array}}\ (s)\ (s')) \\ ce2: \ \mathbf{asrt}(B,A, \boxed{\begin{array}{l} y\ u\ e' \\ \hline \mathbf{boxcar}(y) \\ e': \boxed{\mathbf{hook}(y,u)} \\ u\ \mathbf{is}\ x \end{array}}\ (s')\ (s'')) \end{array}}$$

We assume the theory of speech acts in (Traum & Hinkelman 1992), described briefly in the next section, according to which utterances generate different kinds of speech acts, some concerned with the task proper, other concerned with synchronization, grounding, and so forth; (2) shows only so-called CORE speech acts, which correspond approximately to the traditional illocutionary speech acts of speech act theory (Austin 1962): **inform**, **suggest**, etc. The DRS in (2) represents a discourse situation in which two conversational events occurred, *ce1* and *ce2*. The two conditions assert that *ce1* is an event of A asserting to B that the described situation of *ce1*, s', extends a previous situation s, and includes an engine located at Avon.

This example also illustrates our third hypothesis: that the 'dynamics' of discourse entities (i.e., the fact that discourse entities are made available for anaphoric reference in subsequent utterances) is achieved because the DESCRIBED SITUATION of a speech act–which, in the case of *ce2* in our example, is s''–extends the described situation (s') of a previous speech act (*ce1*). We assume with most literature on dynamic logics that the 'propositional content' of an utterance is a relation between two state-like object; our own specific assumptions are (i) that these two 'state-like' objects are situations, which we take to be ⟨world,assignment⟩ pairs, and (ii) that DRSs denote such relations. This gives us the desired dynamics of discourse referents: In (2), for example, the described situation s'' of the conversational event *ce2* coming after *ce1* is an extension of the described situation s' of *ce1*, hence it contains all of the discourse referents of s'.

These ideas have been implemented by modifying the logic TY_2^4, a version of Situation Theory proposed in (Muskens 1989); in short, we added to TY_2^4 some of the ideas from (Muskens 1994), the resulting logic still has the crucial properties of Muskens' system in (Muskens 1994), the Merging Lemma and the Unselective Binding Lemma. By adopting a compositional formulation of DRT such as that of Muskens, we will also be able to specify the update potentials of different kinds of utterances with respect to our 'mixed' notion of the common ground, including those utterances that do not contribute to the semantic content of larger utterances, such as *okay*.

Conversation Acts and Discourse Structure

Many theories of discourse structure have been proposed in the literature. We will adopt a speech act-based account; a theory of this kind gives us the tools to account not only for the organization of dialogs according to the domain task just discussed, but also for other kinds of structure observable in spoken dialogs, such as the structure of turn-taking and the structure of grounding. In addition, we will also see that information about speech acts is a crucial ingredient of accounts of interpretation that take into account the fragmentary nature of spoken input.[2]

[2] The DIT model being developed by Bunt (1995) is also based on speech acts. This dependence on speech acts is the main difference between our model and the STDRT model of context developed by Asher, Lascarides, Oberlander and others (see, e.g., (Lascarides & Asher 1991;

Most classic theories of speech acts concentrate on the actions performed by the conversational participants as a way of 'getting the job done'—e.g., instructions to the other conversant, requests for information necessary to accomplish the task, etc. But these actions are only a part of what happens in conversations; the conversants spend a lot of their time making sure they do not talk over each other and ensuring that 'informational' coordination is achieved. Recent theories of speech acts (e.g., (Novick 1988; Kowtko, Isard, & Doherty 1992; Traum 1994; Bunt 1995)) are built on the assumption that a good theory of the actions involved in these aspects of a conversation is as important to a dialogue system as a good theory of task-oriented acts.

We adopt here the multi-level CONVERSATION ACTS theory, presented in (Traum & Hinkelman 1992). This theory maintains the classical illocutionary acts of speech act theory (e.g., **inform**, **request**), now called CORE SPEECH ACTS. These actions are, however, reinterpreted as multi-agent collaborative achievements, taking on their full effect only after they have been *grounded*, i.e., acknowledged. Rather than being actions performed by a speaker to a hearer, the core speech acts are joint actions; the initial speaker and the hearer (called hereafter *initiator* and *responder*, respectively) each contribute actions of a more basic type, the result being the common ground assumed to be the effects of core speech acts.

In addition, Conversation Acts (CA) theory also assumes that three other kinds of speech acts are performed in conversations: acts for TURN-TAKING, GROUNDING, and more complex acts called ARGUMENTATION ACTS that involve more than one core speech act—for example, to perform an elaboration. The four kinds of acts of CA theory are displayed in Table 1. The acts from top to bottom are typically realized by larger and larger chunks of conversation: from turn-taking acts usually realized sub-lexically, to grounding acts which are realized within a single utterance unit (UU),[3] to core speech acts which are only completed at the level of a completed discourse unit (DU) to argumentation acts which can span whole conversations. The table also shows some representative acts for each class.

Rather than the Acts of the

Below, we present some key effects of the performance of core speech acts, using, however the Forward-looking acts of the provisional DRI dialogue act scheme (Discourse Resource Initiative 1997; Allen & Core Draft 1997), rather than those used in the TRAINS system (Allen *et al.* 1995), as was used in (Traum & Hinkelman 1992).

Mental States

Facts about the mental states of agents play an important role in speech act recognition, reference resolution, and dialog management. Information about the (mutually known) intentions, beliefs, perceptual input, and obligations (Traum & Allen 1994) of the conversants is also part of the discourse situation. This information can be represented in the language introduced in the previous section by conditions of the form s : K, asserting that STATE s of type K is part of the discourse situation, where a state is a particular type of situation with different properties from events (see, e.g., (Kamp & Reyle 1993) for a characterization of states and events). Continuing the example of (2), the fact that A intends boxcar y to be at Bath in situation s' which extends s (and therefore 'inherits' all the individuals that occur in s) can be thought of as a state. The occurrence of this state in a discourse situation can be represented by including in the Root DRS a condition that expresses the presence in the common ground of an intention *i1* of A, as follows:

(3)

$\ldots s'\, i1\ tj$
$\ldots$
$i1$: **intend-that**(A, s' : $\boxed{\textbf{at}(y, Bath)}$)

Some properties of mental states follow from the fact that states are just one kind of situation; such properties include, for example, 'downward persistence' properties, i.e., the fact that if agent A is in a state MS, and if MS spans the temporal interval I, then A is in that state at all intervals I′ such that I′ is contained in I. These basic properties should be complemented by axiomatizations of the relevant states; we will not do so here. The reader can assume her/his own favorite formalization of mental attitudes, of which there are many around (e.g., (Cohen & Levesque 1990) or (Konolige & Pollack 1993)). Below, we present some of the aspects of mental and social state which seem necessary for capturing the forward-looking DRI acts. We would like to emphasize that the facts represented as conditions in the root DRS correspond to *mutually known* facts in other theories; e.g., a condition of the form **believe**(A,φ) in the root DRS corresponds to a fact of the form **bmb**(A,B,**believe**(A,φ)) in Cohen and Levesque's (1990) formalism.[4]

We will assume that all properties of the discourse situation can be expressed as properties of some state or event included in the discourse situation. Thus the

Asher 1993)). One reason for the difference is that Asher *et al.* are concerned with texts rather than conversations.

[3]Utterance Units roughly correspond to intonation phrases, although long pauses are also taken as unit boundaries. See (Traum & Heeman 1997) for an empirical investigation of the appropriate utterance unit boundaries for grounding.

[4]Assuming here that we are representing the discourse situation from A's point of view. If we we modeling from B's point of view, this would be equivalent to **bmb**(B,A,**believe**(A,φ)).

Discourse Level	Act Type	Sample Acts
Sub UU	Turn-taking	**take-turn**, **keep-turn**, **release-turn**, **assign-turn**
UU	Grounding	**initiate**, **continue**, **ack**, **repair**, **ReqRepair**, **ReqAck**, **cancel**
DU	Core Speech Acts	**inform**, **ynq**, **check**, **eval** **suggest**, **request**, **accept**, **reject**
Multiple DUs	Argumentation	**elaborate**, **summarize**, **clarify** **q&a**, **convince**, **find-plan**

Table 1: Conversation Act Types

basic format for representing the conditions of mental states is shown in (4).

(4) s : $\mathbf{MS}(A,X)$

In the next sections we discuss how we propose to capture facts about the evolving structure of discourse. In addition, facts about the interactional state of the dialogue, such as which conversant has the turn or initiative at a given time, can be represented in a similar manner.[5]

A first cut at some important aspects of mental state types for the endeavor of understanding speech acts is described below:

Bel(A,$K(s)$) Agent A Believes the DRS K is true of situation S.

Int(A,α) A intends to perform an action of type α

Option(A,α) α is an action that A is aware that she can perform. Of the range of these actions, some are chosen as intentions. Options are also used in this fashion in the IRMA architecture (Bratman, Israel, & Pollack 1988; Pollack 1992).

In addition, there are some important *social* attitudes, which relate an agent not only to a course of events or action, but also to a social group. In addition to mutual belief, these include:

SCCOE(A,B,$K(s)$) this stands for Socially Committed to a Course of Events. It is the public counterpart to individual belief. It means that A is committed to B to K being the case (whether or not she actually privately believes it).

Obliged(A,B,α) A is obligated to B to perform an act of type α(whether or not she actually intends to).

Forward-looking Act Effects

The main forward-looking acts from the DRI dialogue act coding scheme are shown in (5). The tabbing represents sub-classes of the main act type. Sub-acts will maintain all of the properties of the parent act, while also adding additional information about the act. The DRI scheme is meant to be a very general one that can serve as a point of comparison for individual coding efforts. Each of these acts can be further specified, when required. For instance, inform could be defined as an assertion in which the addressee does not know the asserted condition (and thus receives the information from the speaker). Similarly, one could break info-request into specific types of question.

[5] For a different view of how information about belief and intentions could be incorporated in DRT, see (Kamp 1990).

(5) Statement
- Assert
- Reassert
- Other

Future-Addressee-Action
- Open-option
- Directive
    - Info-request
    - Action-Directive

Future-Speaker-Action
- Offer
- Commit

In order to model these acts within the current logic, we say that a (conversational) event is characterized by an event type corresponding to the appropriate act. The fact that such an event occurred also has implications about other aspects of the mental and social state. For instance, these acts can bring about new obligations, SCCOE's, and options. To describe these effects, and definitional conditions, we introduce some further bits of notation. The most general case of the discourse state being updated by the performance of a speech act is shown in (6). Here, DS and DS$'$are the old and updated discourse situation, while ce is the conversational event itself. K, the content of the speech act may also contain described situations.

(6) $$\boxed{\begin{array}{l} ce \\ \hline ce:\ \mathbf{SA}(A,B,K) \end{array}}(\mathrm{DS})(\mathrm{DS}')$$

In addition, some other event types are needed to give a characterization of the speech act effects. These described informally below:

e : **Try**(A,α) means that e is characterized by A trying to perform an act of type α (present-directed intention).

e : **Do**(A,α) means that e is characterized by A performing an act of type α (regardless of whether he actually tried to or not).

e : **Achieve**(A,K) means that e is characterized by the act of A bringing about the satisfaction of DRS K.

e : **Address**(A,e') means that e is characterized by A considering and responding to e$'$.

e : **Accept**(A,e') (a backward-looking speech act) means that e is characterized by A accepting act e$'$.

The formulations of acts and their main effects or defining conditions are given in Table 2. The enclosing DRS used to represent the discourse situation update, as in (6) is assumed, and the markers DS, DS$'$ used as needed. A speaker is committed to the veracity of her statement. If the statement is used to try to achieve the belief of the addressee (regardless of it's success, or the prior belief of the addressee), then it is an assertion. If the speaker was already previously so committed, then it is a reassert. An Other statement is a statement that is not an assert or reassert. A future-addressee-action act constrains the discourse situation to contain an option for the addressee. Open-option does only this. More precisely, it does not count as an attempt to get the addressee to actually do the mentioned act, merely allows it as a possibility for consideration. A directive, on the other hand, *does* count as such an attempt. Directives also impose an obligation to address the directive itself (though not necessarily to perform the requested action) (Traum & Allen 1994). There are also two kinds of directives, depending on what kind of action is directed. If it is a statement, then the action is an info-request. Otherwise it is an action-directive. A future-Speaker-action act mentions an option of the speaker. A Commit act means that the speaker has an obligation to perform the action. An offer is a conditional commitment: if the addressee accepts, then the speaker is committed.

It is also useful to consider default inferences which can be drawn about the mental states of agents, given the performance of acts or existence of other states. For instance, generally, and for the case of an honest agent, one can draw a correspondence between SCCOE and actual belief, as in (7).

(7) $\forall$,b,K,s,i
$$\boxed{s:\ \mathbf{SCCOE}(a,b,K)}(\mathrm{i}) \Rightarrow \exists\ \mathrm{s}'\ \boxed{s':\ \mathbf{Bel}(a,K)}(\mathrm{i})$$

Similar rules can be written to aid in dialogue management. For instance, if an agent is committed to K, and yet believes that K is not true, a sincere agent can adopt the intention (and form a plan) to make a statement that K is not the case.

Conversational Threads, Argumentation Acts, and Discourse Scripts

It is a basic fact about the way humans interpret events that they tend to be grouped into larger 'stories' or, as we will call them here, THREADS (Nakhimovsky 1988; Webber 1988; Kameyama, Passonneau, & Poesio 1993). A thread is itself an event, that decomposes hierarchically into its constituent events (Kautz 1987). The hierarchical organization of speech acts into larger units or discourse segments (associated with more general discourse purposes) is just an instance of this more general phenomenon of events being grouped into threads, and the relations between DSPs assumed by Grosz and Sidner (Grosz & Sidner 1986) are those generally assumed to hold between actions (e.g., in Kautz's theory). We will use the term CONVERSATIONAL THREADS for threads of conversational events when we want to distinguish between this 'technical' notion of discourse segment from the intuitive notion.[6]

Our theory of event structure is fairly standard. We assume that events can be decomposed into smaller events; the relation between events and the threads of which they are a part of corresponds to the domination relation in theories such as Kautz's. We also assume that each event in a thread has an immediately preceding and immediately following event. Finally, we assume that the perspective from which we view a thread changes over time, i.e., we assume that each thread has a 'current-event' ('now point') at any time t.

In Conversation Act theory, certain kinds of threads are singled out. We assume that rhetorical 'relations' such as **elaboration** or **explanation** (Mann & Thompson 1987) are in fact a particular form of conversation act involving multiple core speech acts, called ARGUMENTATION ACTS. These acts implicitly involve domination, satisfaction-precedence, and other relations between the component events, depending on the type of rhetorical relation.[7] For example, the in-

[6]The reason for this term is that 'discourse segments'—introduced to account for reference facts—are only one type of conversational threads. Another are the *Discourse Units* used by the participants to build common ground. It is still an open question as to the precise relationship between these two types of threads, which represent the structure of topic management, and the structure of sharing of information, respectively.

[7]A similar position is taken in recent work on rhetorical structure in the generation field (Moore & Paris 1993).

Name	act condition	defining effect
Statement	$e:\ \mathbf{stmt}(A,B,K(s)(s'))$	$DS':\ \mathbf{SCCOE}(A,B,K(s))$
Assert	$e:\ \mathbf{asrt}(A,B,K(s)(s'))$	$e:\ \mathbf{Try}(A,\mathbf{Achieve}(A,DS':\ \mathbf{Bel}(B,K(s))))$
Reassert	$e:\ \mathbf{rsrt}(A,B,K(s)(s'))$	$DS:\ \mathbf{SCCOE}(A,B,K(s))$
Future-Addr-Act	$e:\ \mathbf{FutOA}(A,B,e')$	$DS':\ \mathbf{Option}(B,e')$
Open-option	$e:\ \mathbf{OpOp}(A,B,e')$	$\neg(e:\ \mathbf{Try}(A,\mathbf{Achieve}(A,\exists\ s,K\ (K(DS')(s) \wedge s:\ \mathbf{Do}(B,e')))))$
Directive	$e:\ \mathbf{Dir}(A,B,e')$	$e:\ \mathbf{Try}(A,\mathbf{Achieve}(A,\exists\ s,K\ (K(DS')(s) \wedge s:\ \mathbf{Do}(B,e'))))$
		$DS':\ \mathbf{Obliged}(B,A,\mathbf{address}(B,e))$
Info-Request	$e:\ \mathbf{Q}(A,B,e')$	$e':\ \mathbf{stmt}(B,A,K'(s')(s''))$
Future-Speaker-Act	$e:\ \mathbf{FutSA}(A,B,e')$	$DS':\ \mathbf{Option}(A,e')$
Commit	$e:\ \mathbf{ComSA}(A,B,e')$	$DS':\ \mathbf{Obliged}(A,B,e')$
Offer	$e:\ \mathbf{Offer}(A,B,e')$	$DS':\ \mathbf{Do}(B,\mathbf{Accept}(B,e)) \rightarrow \mathbf{Obliged}(A,B,e')$

Table 2: Act Definitions

terpretation of *A: There is an engine at Avon. B: It is attached to a boxcar* in which the second utterance constitutes an elaboration of the first would result in the following discourse situation, in which the speech act *ce2* elaborates the speech act *ce1*:

$$(8)\quad \boxed{\begin{array}{l} ce1\ ce2\ aa1\ s\ s'\ s'' \\ \hline ce1:\ \mathbf{asrt}(A,B, \boxed{\begin{array}{l} x\ w\ e \\ \hline \mathbf{engine}(x) \\ \mathbf{Avon}(w) \\ e:\ \boxed{\mathbf{at}(x,w)} \end{array}}\ (s)\ (s')) \\ ce2:\ \mathbf{asrt}(B,A, \boxed{\begin{array}{l} y\ u\ e' \\ \hline \mathbf{boxcar}(y) \\ u\ \mathbf{is}\ x \\ e':\ \boxed{\mathbf{hook}(y,u)} \end{array}}\ (s')\ (s'')) \\ aa1:\ \mathbf{elaborate}(B,A,ce1,ce2) \end{array}}$$

More generally, we assume that people know a lot about the structure of certain kinds of threads, and use this information to predict what's going to happen next, as well as to 'fill in' holes in the description. The idea that this information about SCRIPTS—threads whose structure and roles are known in advance, such as the sequence of events that takes place in a restaurant—plays a crucial role in natural language processing was explored in well-known work by Schank and associates (Schank & Abelson 1977). Similarly, people know a lot about the organization of events in conversations, both at a broad level (e.g., what generally happens in a conversation) and at a more local level (e.g., what to expect after a question); the so-called micro- and macro-structure of conversations has been studied in the field of CONVERSATION ANALYSIS (Schegloff & Sacks 1973; Sacks, Schegloff, & Jefferson 1974). The connection between work on scripts and work in conversation analysis was explored in AI work on 'dialogue games' (Power 1979; Kowtko, Isard, & Doherty 1992; Airenti, Bara, & Colombetti 1993) and 'discourse scripts' (Poesio 1991; Turner 1989).

We reformulate Grosz and Sidner's notion of discourse structure in terms of threads to establish an explicit connection between work on intention recognition using expectations and work based on the planning paradigm. A discourse script is simply a particular type of thread; by recognizing the thread of which a certain speech act is a part of, and the current position in that thread (as specified by the 'now' point of that thread), we can use expectations to recognize the type of speech act. What's more, our assumption that conversational events are organized into conversational threads is a more general assumption than Grosz and Sidner's idea that core speech acts are organized in discourse segments, since we allow for threads of turn-taking acts and grounding acts as well.

Current Work

The work described above is still in the early stages. We are continuing it along several fronts. First, we are examining some of the details of the semantic formulation, particularly constraints on the ways in which situations extend other situations. We are also starting to use this framework to code and analyze whole conversations in an attempt to isolate constraints on the way conversational threads are extended in practice. Finally, we intend to use this theory as the basis for conversational agents, in which the agent, using defeasible reasoning can track and participate in a conversation, building and using the representation as the conversation proceeds.

References

Airenti, G.; Bara, B. G.; and Colombetti, M. 1993. Conversation and behavior games in the pragmatics of dialogue. *Cognitive Science* 17:197–256.

Allen, J., and Core, M. Draft, 1997. Draft of damsl: Dialog act markup in several layers. available through the WWW at: http://www.cs.rochester.edu/research/trains/annotation.

Allen, J. F.; Schubert, L. K.; Ferguson, G.; Heeman, P.; Hwang, C. H.; Kato, T.; Light, M.; Martin, N.; Miller, B.; Poesio, M.; and Traum, D. R. 1995. The TRAINS project: a case study in building a conversational planning agent. *Journal of Experimental and Theoretical Artificial Intelligence* 7:7–48.

Allen, J. F. 1983. Recognizing intentions from natural language utterances. In Brady, M., and Berwick, R., eds., *Computational Models of Discourse*. Cambridge, MA: MIT Press. 107–166.

Asher, N. 1993. *Reference to Abstract Objects in English*. Dordrecht: D. Reidel.

Austin, J. L. 1962. *How to Do Things with Words*. Cambridge, MA: Harvard University Press.

Barwise, J., and Perry, J. 1983. *Situations and Attitudes*. The MIT Press.

Bratman, M. E.; Israel, D. J.; and Pollack, M. E. 1988. Plans and resource-bounded practical reasoning. *Computational Intelligence* 4:349–355.

Bunt, H. C. 1995. Dynamic interpretation and dialogue theory. In Taylor, M. M.; Néel, F.; and Bouwhuis, D. G., eds., *The Structure of Multimodal Dialogue 2*. Amsterdam: John Benjamins.

Carberry, S. 1990. *Plan Recognition in Natural Language Dialogue*. Cambridge, MA: The MIT Press.

Cohen, P. R., and Levesque, H. J. 1990. Persistence, intention and commitment. In Cohen, P.; Morgan, J.; and Pollack, M., eds., *Intentions in Communication*. Morgan Kaufmann. chapter 12.

Cohen, P. R.; Morgan, J.; and Pollack, M., eds. 1990. *Intentions in Communication*. Cambridge, MA: MIT Press.

Discourse Resource Initiative. 1997. Standards for dialogue coding in natural language processin. Report no. 167, Dagstuhl-Seminar.

Grosz, B. J., and Sidner, C. L. 1986. Attention, intention, and the structure of discourse. *Computational Linguistics* 12(3):175–204.

Kameyama, M.; Passonneau, R.; and Poesio, M. 1993. Temporal centering. In *Proc. of the Meeting of the Association for Computational Linguistics*, 70–77.

Kamp, H., and Reyle, U. 1993. *From Discourse to Logic*. Dordrecht: D. Reidel.

Kamp, H. 1990. Prolegomena to a structural account of belief and other attitudes. In Anderson, C. A., and Owens, J., eds., *Propositional Attitudes—The Role of Content in Logic, Language, and Mind*. Stanford: University of Chicago Press and CSLI. chapter 2, 27–90.

Kautz, H. A. 1987. *A Formal Theory of Plan Recognition*. Ph.D. Dissertation, Department of Computer Science, University of Rochester, Rochester, NY. Also available as TR 215, Department of Computer Science, University of Rochester.

Konolige, K., and Pollack, M. E. 1993. A representationalist theory of intention. In *Proceedings IJCAI-93*.

Kowtko, J. C.; Isard, S. D.; and Doherty, G. M. 1992. Conversational games within dialogue. Research Paper HCRC/RP-31, Human Communication Research Centre.

Lascarides, A., and Asher, N. 1991. Discourse relations and defeasible knowledge. In *Proc. ACL-91*, 55–63.

Mann, W. C., and Thompson, S. A. 1987. Rhetorical structure theory: A theory of text organization. Technical Report ISI/RS-87-190, USC, Information Sciences Institute.

Moore, J. D., and Paris, C. L. 1993. Planning text for advisory dialogues: Capturing intentional and rhetorical information. *Computational Linguistics* 19(4):651–694.

Muskens, R. 1989. *Meaning and Partiality*. Ph.D. Dissertation, University of Amsterdam.

Muskens, R. 1994. A compositional discourse representation theory. In Dekker, P., and Stokhof, M., eds., *Proceedings of the 9th Amsterdam Colloquium*, 467–486.

Nakhimovsky, A. 1988. Aspect, aspectual class, and the temporal structure of narratives. *Computational Linguistics* 14(2):29–43.

Novick, D. 1988. *Control of Mixed-Initiative Discourse Through Meta-Locutionary Acts: A Computational Model*. Ph.D. Dissertation, University of Oregon. also available as U. Oregon Computer and Information Science Tech Report CIS-TR-88-18.

Perrault, C. R. 1990. An application of default logic to speech act theory. In Cohen, P. R.; Morgan, J.; and Pollack, M. E., eds., *Intentions in Communication*. Cambridge, MA: The MIT Press. chapter 9, 161–185.

Poesio, M., and Traum, D. R. 1997. Conversational actions and discourse situations. *Computational Intelligence* 13(3).

Poesio, M. 1991. Expectation-based recognition of discourse segmentation. In *Proc. AAAI Fall Symposium on Discourse Structure*.

Poesio, M. 1993. A situation-theoretic formalization of definite description interpretation in plan elaboration dialogues. In Aczel, P.; Israel, D.; Katagiri, Y.; and Peters, S., eds., *Situation Theory and its Applications, vol.3*. Stanford: CSLI. chapter 12, 339–374.

Pollack, M. E. 1992. The uses of plans. *Artifical Intelligence* 57(1):43–68.

Power, R. J. D. 1979. The organization of purposeful dialogue. *Linguistics* 17:107–152.

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

Schank, R. C., and Abelson, R. 1977. *Scripts, Plans, Goals and Understanding*. Lawrence Erlbaum.

Schegloff, E. A., and Sacks, H. 1973. Opening up closings. *Semiotica* 7:289–327.

Stalnaker, R. 1979. Assertion. In Cole, P., ed., *Syntax and Semantics*, volume 9. Academic Press. 315–332.

Traum, D. R., and Allen, J. F. 1994. Discourse obligations in dialogue processing. In *Proceedings of the 32th Annual Meeting of the Association for Computational Linguistics*, 1–8.

Traum, D. R., and Heeman, P. 1997. Utterance units in spoken dialogue. In Maier, E.; Mast, M.; and Luperfoy, S., eds., *Dialogue Processing in Spoken Language Systems — ECAI-96 Workshop*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Heidelberg. 125–140.

Traum, D. R., and Hinkelman, E. A. 1992. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence* 8(3). Special Issue on Non-literal Language.

Traum, D. R. 1994. *A Computational Theory of Grounding in natural language conversation.* Ph.D. Dissertation, University of Rochester, Department of Computer Science, Rochester, NY.

Turner, E. H. 1989. *Integrating Intention and Convention To Organize Problem Solving Dialogues.* Ph.D. Dissertation, Georgia Institute of Technology.

Webber, B. L. 1988. Tense as discourse anaphor. *Computational Linguistics* 14(2):61–73.