Speech Acts for Dialogue Agents*

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

1 Introduction

A *dialogue agent* is one that can interact and communicate with other agents, in a coherent manner, not just with one-shot messages, but with a sequence of related messages all on the same topic or in service of an overall goal. Following the basic insights of *speech act theory*, these communications are seen not just as transmitting information but as actions which change the state of the world. Most of these changes will be to the mental states of the agents involved in the conversation, as well as the state or context of the dialogue. As such, speech act theory allows an agent theorist or designer to place agent communication within the same general framework as agent action. In general, though, communicative action requires a more expressive logic of action than is required for something like the single-agent blocks world, familiar in classical AI planning. For one thing, there are multiple agents, and there is also a possibility of simultaneous and fallible action.

In studying speech acts, the focus is on pragmatics rather than semantics — that is, how language is used by agents, not what the messages themselves mean in terms of truth-conditions in a model (see [66] for a good introduction to issues in natural language pragmatics). As with other aspects of pragmatics such as implicature and presupposition, an important concern is what can be inferred (perhaps only provisionally or with a certain likelihood) as a result of the performance of a speech act. While much of speech act work has been analyzing interaction in natural language, speech acts are also a convenient level of analysis for artificial communication languages. While the rules for interpreting whether a particular act has been performed will be different for an artificial language (presumably simpler, with less concern about vague and

*Appears in *Foundations of Rational Agency*, M. Wooldride and A. Rao, Editors, pp 169-201. Kluwer, 1999.

ambiguous utterances), the range of actions that agents will wish to perform will still be roughly the same (although perhaps not as many fine gradations will be necessary for simpler agents).

In this chapter, we consider the role of speech acts in allowing an agent to participate in a dialogue with another agent. The bulk of the paper will be a review of some of the most important work on speech acts and its implication for dialogue agents. In Section 2, we review the foundational work on speech act theory, stemming from work in the philosophy of language. In Section 3, we consider work in Artificial Intelligence aimed at providing computational accounts of speech acts, primarily accounts that could be used by agents to engage in communication — both performing and interpreting such acts. In most of that work, the emphasis is on the performance or recognition of single utterances. We therefore turn, in Section 4, to work on expanding speech act theory to also include other aspects of dialogue. In Section 5, we consider some work in which the agent communication language itself is formulated in terms of speech act concepts. Such languages make identification of speech acts somewhat easier, but still contain many of the same issues as languages in which speech acts have a less transparent relation to surface form. In Section 6, we reconsider some of the issues confronting speech act theory, in light of the presented work. Finally, in Section 7, we begin to sketch a particular theory of speech acts (that presented in [95], pointing out some of the extra elements required beyond a BDI-theory of agency.

2 Foundational Philosophical Speech Act Work

Work on speech acts begins with philosophers of language interested in issues in natural language pragmatics. In particular, difficulties with the primary view of the meanings of sentences as merely truth-functions. Pragmatics (as opposed to semantics) concentrates on *utterances* (the performance) rather than sentences (the objects produced), and also is generally concerned with other aspects of the context of production and interpretation.

2.1 Austin

Austin observed that utterances are not just descriptions of states of affairs, but are used to *do* things [16]. Under felicitous circumstances, utterances can change the mental and interactional state of the participants. Speaking is *acting*, and in speaking the speakers are performing *speech acts*. There are multiple types of action performed in speaking, and Austin distinguished several. *Locutionary acts* are the act of saying something, including a *phonetic act* – producing certain noises, a *phatic act* – producing words belonging to a vocabulary in a constructions conforming to a grammar, and a *rhetic act* – using the product of the phatic act with a particular sense and reference.

Illocutionary acts are those acts performed *in* saying something, for example, asking or answering a question, giving some information, etc. Most of the subsequent work on speech acts has been on illocutionary acts. Illocutionary acts are taken to be composed of an *illocutionary force*, which specifies the type of action (e.g., requesting, suggesting, warning, apologizing, informing), and a *propositional content* which specifies the details of the action (e.g., what it is that the hearer is being requested to do). Illocutionary acts are not always directly deducible from the locutionary acts which generate them. *Indirect speech acts* are those in which the act performed is other than what would be expected from a compositional account of the content. Austin gives a bridge example in which an utterance of "I bid three clubs" is used to inform a partner that the speaker has no diamonds. As this example illustrates, utterances can be used to achieve multiple functions. In this case the direct (conventional) *bid* act is performed, as well as the indirect *inform*.

Perlocutionary acts are those which are performed *by* saying something – actions which achieve effects which are special to the particular situation of utterance rather than the conventional nature of the communication. Examples include persuasion, surprise, and deterrence. Although a speaker might intend these, it generally takes more than just the utterance to achieve them.

Austin spent much of his effort on *performatives*, in which the utterance itself creates a result. *Explicit performatives* are those utterances in which the result is described in the utterance. In English, the word "hereby" often accompanies performative utterances. Not all speech act verbs can be used performatively, however, e.g., we don't have the conventional form in English, "I insult you." As others (e.g., [32]) have noticed, some such declarations would be self-defeating, such as "lie". Austin considers the relation of performative utterances and other pragmatic phenomena such as *implicature* and *presupposition*, in which other facts can be derived, but not through the monotonic entailments of ordinary logic.

Austin classifies the illocutionary acts into several categories, based primarily on illocutionary force: *verdictives, exercitives, commissives, expositives,* and *behavitives*. *Verdictives*, which are basically giving a (perhaps preliminary or tentative) verdict about something. *Exercitives* exercise a right, such as appointing, or nominating. *Commissives* commit the speaker to do something. *Behavatives* all have something to do with attitudes and social behavior, such as apologizing and congratulating. *Expositives* have to do with how an utterance fits into a conversation or argument, such as clarifying or postulating.

2.2 Searle

Searle extends and refines Austin's work on illocutionary acts. He observes that Austin's decomposition of speech acts into *illocutionary force* and *propositional content* shows up in the different kinds of negation that can be performed in a sentence [87]. For example, the sentence "I promise to come" has two negations:

a *propositional negation*, "I promise not to come", in which the illocutionary act is the same (promise), but in which the content is negated, and "I do not promise to come", in which the propositional content is the same, but the illocutionary act is no longer a promise, but a *refusal* to make a promise. Searle further decomposes the *propositional content* of an act as a combination of *predicating* and *reference* acts. Neither of these stand alone but are performed only in concert with illocutionary acts.

Probably the most important contribution was an attempt to provide necessary and sufficient conditions for the performance of illocutionary acts. He presented these as *constitutive rules* (like the rules which define the games of football or chess) of various sorts. *Normal input-output conditions* concern the conditions of intelligible speaking and understanding, including knowing the conventions of languages, paying attention, etc. *Propositional content conditions* describe restrictions on the content, e.g., for a promise the content must be a future action. *Preparatory conditions* involve the constraints on the world that make the speech act useful. *Sincerity conditions* involve alignment of the speaker's actual attitudes (belief, desire, etc.) with the attitudes expressed by the act. *Essential conditions* involve the speaker's intentions in performing the act – what she was *trying* to do. Searle also adds a condition based on Grice's notion of non-natural meaning [45], that the effect of the act is in part produced by the hearer's recognition that the utterance is intended to produce this effect by means of the recognition of the intention.

Searle also improves Austin's classification of types of illocutionary acts mentioned above [88]. Austin's classification is fairly haphazard, based more on similarity of illocutionary verbs than the acts themselves. Searle points out that there is no necessary correspondence between illocutionary acts, and the illocutionary verbs that a particular language chooses to describe these acts. Searle proposes an alternate taxonomy based on the purposes of the acts. To do this, Searle presents an analysis of several dimensions along which speech acts can vary, as outlined below:

1. **The point (purpose) of the act.** These correspond to the *essential conditions* described above, such as attempting to get the hearer to do something (request), represent the world (statement), or commit (promise). The point is the main, although not only component of illocutionary force. The force also includes other notions such as strength of the point, and the other dimensions.

2. **Direction of fit between words and the world.** Statements try to fit word to the way the world is, whereas commitments and requests try to fit the (future state of the) world to the way the words are. Searle uses an example (credited to [14]) of a list of groceries, and differences in how it might be constructed and used. If it is a shopping list given by a spouse, it is used by the shopper to fit the world (the groceries purchased) to the

words (the list). If, on the other hand, it is made by an observer of the shopper, indicating what was actually bought, it would be a fit of word to the world. Although identical lists might be produced, the difference is critical for repairing mistakes. In the case of word to world fit, if the shopper really bought pork chops rather than bacon, the observer can simply cross off "bacon" and write "pork chops" on his list. The shopper, however, must go back to the store and exchange his purchases, to repair an error.

3. **Expressed psychological state** Different acts will express different mental attitudes on the part of the speaker. An assertion expresses a belief, while a request expresses a desire for the addressee to act. This dimension relates most closely to Searle's *sincerity condition* – a speaker might not actually have these attitudes, but is nonetheless expressing them in performing the speech acts.

4. **Strength of point** This is related to the above – attitudes can generally be held with varying degrees of strength. Thus one can have a tentative or firm belief, and could suggest, request, or command.

5. **Status of the participants** The social relationship of the speakers may also play a role in the strength of the force. Thus a general can order a private, but a private may only suggest or request. These differences and the next also relate to Searle's *preparatory conditions*.

6. **Interest of the participants** This would distinguish boasts from laments, or promises from threats. While Searle separates this from condition (3), we can see this as relation to (expressed or assumed) desires of the speaker and addressee.

7. **Relations to the rest of the discourse** This is how discourse relations such as *answer* or *deduce* can be distinguished from other statements.

8. **Propositional content** Aspects of the propositional content that play a role in illocutionary force, e.g., that a promise or request can only involve future action, while a statement could be about the past. These correspond to Searle's *propositional content conditions*, though they are also closely related to dimension (2).

9. **Speech Acts vs. general acts** Some things must be performed as speech acts, whereas others might be performed by speech acts or other kinds of action. Examples include "classify", "estimate" and "diagnose".

10. **Extra-linguistic institutions** This includes features of the social and institutional context, going beyond language. Thus, only individuals with a certain position within an institution can marry people or bid three notrump. Searle says that (perhaps contra-Austin) not all acts need such

an institutional position, claiming that one need only obey the rules of language to make a statement or promise. However, if one considers more general agents, one could say that even for these actions, an institutional position is required: though merely as an agent within society. It is difficult to know what exactly Searle means by "the rules of language", but clearly syntactic and semantic well-formedness is not sufficient to perform a promise – there must be some notion of the language producer having undertaken a commitment.

11. **performativeness** Whether there is an illocutionary verb that has a performative use. "boast" and "threaten", for example, do not.

12. **style of performance** This distinguishes announcing from confiding.

Searle uses these features, mainly (1), (2), and (3) to produce a more consistent speech act taxonomy. Searle's categories include: *representatives*, which commit the speaker to the truth of an expressed proposition, *directives*, which involve getting the hearer to do something, *commissives*, which involve committing the speaker to some course of action, *expressives*, which convey a psychological state of the speaker, and *declarations*, which bring about the correspondence of the world to the word.

Searle's speech act theory was later formalized within a possible worlds semantics, and unified with a Tarskian tradition, along with Vanderveken [90, 103, 104].

3 AI Models of Speech Acts

A difficulty with the early philosophical work on speech acts was that there did not yet exist very formal accounts of action and mental states that could be used to design more precise definitions for speech acts. Moreover, such definitions can be rather difficult to test in the abstract, without observing what is really at stake in actual communication episodes. With the rapid progress in the 1970s on reasoning about actions and plans, it became possible to give a more operational account of speech acts, in terms of AI planning operators.

Bruce was the first to try to account for Speech Act theory in terms of AI work on actions and plans [20]. He defined natural language generation as *social action*, where a *social action* is one which is defined in terms of beliefs, wants, and intentions. He also presented *Social Action Paradigms* which showed how speech acts could be combined to form larger discourse goals. He showed how acts such as **Inform** or **Request** could be used in achieving intentions to change states of belief.

3.1 Allen, Cohen, and Perrault

Probably the most influential AI work on speech acts to date has been the work at University of Toronto by Allen, Cohen, and Perrault. Cohen and Perrault [38] tackled the problem of language production, defining speech acts as plan operators which affect the beliefs of the speaker and hearer.

In setting up the program they set forth some important goals for work on speech acts. They write that any account of speech acts should answer the following questions:

- Under what circumstances can an observer believe that a speaker has sincerely and successfully performed a particular speech act in producing an utterance for a hearer?
- What changes does the successful performance of a speech act make to the speaker's model of the hearer, and to the hearer's model of the speaker?
- How is the meaning (sense/reference) of an utterance x related to the acts that can be performed in uttering x?

They also suggest that a theory of speech acts based on plans should specify the following:

- A planning system: a formal language for describing states of the world, a language for describing operators, a set of plan construction inferences, a specification of legal plan structures. Semantics for the formal languages should also be given.
- Definitions of speech acts as operators in the planning system. What are their effects? When are they applicable? How can they be realized in words?

Cohen and Perrault's models of mental states consist of two types of structures: *beliefs* and *wants*. *Beliefs* are modal operators which take two arguments: an agent who is the believer, and a proposition which is believed. They also follow Hintikka [57], augmenting the belief structure to include quantified propositions. Thus an agent can believe that something has a value without knowing what that value is, or an agent can believe another agent knows whether a proposition is true, without the first agent knowing if it's true or not. *Wants* are different modal operators which can nest with beliefs. *Wants* model the goals of agents.

Perrault and Cohen then proceeded to make a first stab at satisfying these issues. The planning system they use is a modified version of STRIPS [43]. They maintain STRIPS's method of dealing with the frame problem: assuming that nothing can change the world except the explicit changes mentioned by the effects of an operator. They describe two different types of preconditions,

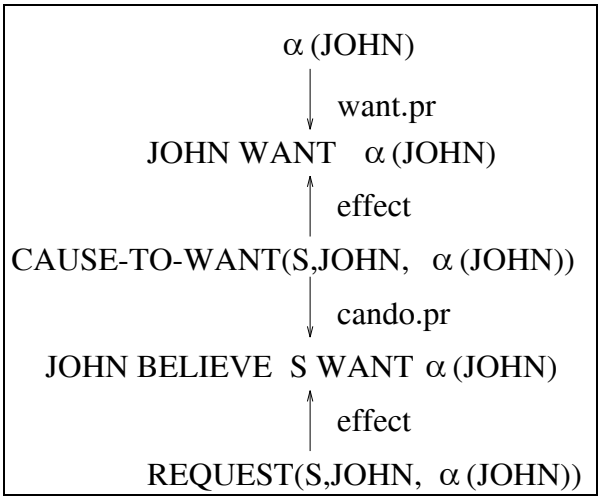

Figure 1: Cohen & Perrault (79) Plan for Request

both of which must hold for the action to succeed. *cando* preconditions indicate propositions which must be true for the operator to be applicable. *Want* preconditions are meant to cover sincerity conditions. In order to successfully perform an action, the agent (speaker) must want to do that action. They model the speech acts REQUEST and INFORM, within their planning system, including various forms of questions as requests for information (or to inform), and consider the relations of requests to informs of desires.

Figure 1 shows a slightly simplified example of the kinds of plan structures that an agent might build in deciding to issue a request, given the desire for some action to be performed. Here, one agent, S, wants anther agent, John, to perform some action α. S realizes that if John has a desire to perform α, then John may perform it. Now, S must find and perform some action that will have as an effect that John wants to perform α. Here, an assumption of cooperativity is required, such that if John realizes that S wants him to do α, John himself will want to do α, as well. And it is assumed that the direct effect of the request is a belief that S wants John to perform α. The upshot is that S can perform this request and then rely on John's rational processes of perception, reasoning, and deliberation to intend and perform α.

This model has a wide range of applicability in cooperative situations, and has also been used to explain why an answer follows a question (i.e., if α is the performance of some speech act).

Allen and Perrault [8] use essentially the same formalism as Cohen and Perrault, but for a slightly different purpose. They investigate the role of plan in-

REQUEST(speaker,hearer,act)
body: MB(hearer, speaker,
speaker WANT hearer DO act)
effect: hearer WANT hearer DO act

SURFACE-REQUEST(speaker,hearer,act)
effect: MB(hearer, speaker,
speaker WANT hearer DO act)

Figure 2: Allen (83) Speech Act Operators

ference and recognition in a cooperative setting. They show how the techniques of recognizing another agent's plans is important in interpreting utterances. For one thing, it can allow one to recognize an indirect speech act in a coherent and relevant manner. For example, Uttering (1) as a request for the information about when the train leaves.

(1) Do you know when the Windsor train leaves?

To do this, they bring in a level of *surface speech acts*, roughly comparable to Austin's locutionary act, which corresponds to the surface form, whereas the illocutionary act corresponds to the speaker's actual intention. Illocutionary acts are realized by surface speech acts, with an additional requirement that the hearer must recognize the speaker's intent. The planning system is again, basically a STRIPS system. There are preconditions and effects, and a body, which is a specification of the operator at a more detailed level, for instance decomposing an illocutionary act into a surface speech act.

The same logic was also used to provide an account of cooperative responses, where a hearer can deduce the speaker's plan and provide help even when it was not requested. For instance, providing the next step in a deduced plan, or detecting obstacles in the speaker's plan that may be unknown to the speaker, and providing help.

This work was later expanded and improved in [3], where the speech act operators in Figure 2 were proposed.

While the Allen and Perrault model nicely handled the role of plan and intention recognition in recognizing indirect speech acts, it was not particularly helpful in relating the role of various aspects of the surface form. In fact, some linguists have proposed that indirect speech acts like (1) are really conventional. Hinkelman elegantly combined both linguistic and plan-based information [56, 55] to yield the best hypothesis for a current utterance. This was done by using linguistic cues to produce partial speech act templates, while plan inference was used to rule out certain possibilities and choose among those that were remaining.

3.2 Non-monotonic Theories of Speech Acts

Perrault [77] takes as a starting point the problem that the utterance itself is insufficient to determine the effects of a speech act. All effects of utterance actions are based in part on the prior context, including the mental states of the agents, as well as what was actually uttered. However, formalizing the precise conditions which must hold is a tricky endeavor, because of the many possible contingencies. Thus an axiom stating the effects of an utterance in declarative mood must take account of the possibilities of lies, failed lies, and irony as well as standard information-giving acts. Perrault's approach is to state the effects in terms of Default Logic [80], so that the simple, most common effects can be derived directly, unless there is some defeater. He has a simple axiomatization of belief, intention and action, along with some normal default rules, including a *Belief Transfer* rule which says that if one agent believes that another agent believes something the first agent will come to believe it too, and a *Declarative* rule, which states that if an agent said a declarative utterance, then it believes the propositional content of that utterance. This simple schema allows Perrault to derive expected consequences for the performance of a declarative utterance in different contexts.

Although the formalization is simple and elegant, it still contains a number of serious difficulties. Foremost is the lack of a serious treatment of belief revision. Although intuitively, speech acts are used to *change* beliefs, Perrault's framework can only handle the case of new beliefs being added. As well as not allowing the kind of discourses in which one agent would try to change the beliefs of another, the logic also has the strange property that one agent can convince itself of anything it has no prior beliefs about merely by making an utterance to that effect in the presence of another agent! The logic also does not lend itself to a computational implementation, since one would need a complete, inductive proof scheme to make all of the necessary deductions.

Appelt and Konolige [15] reformulate Perrault's theory in terms of Hierarchic Autoepistemic Logic [59]. This reformulation has the advantages of implementability and the ability to order the defaults to overcome the problems that Perrault had with normal default logic, but it also loses the simplicity of Perrault's framework. It is hard to see whether Appelt and Konolige are trying to describe something from the point of view of an ideal observer or from a participant in the conversation. In formulating their theory, they also resort to some unintuitive devices such as the beliefs of an utterance.

McRoy and Hirst use a uniform *abductive* approach for the production and recognition of speech acts, given aspects of the context as well as an (assumed) identification of surface form [75]. This approach lets them reason about mistakes in the interpretation process in a manner that is also amenable for reasoning about cancellable presuppositions and implicatures.

3.3 Dynamic Logic Approaches

One of the problems with work based on STRIPS is that it does not provide a fine enough characterization of time and change. Some researchers have thus moved to other logics of action which provide more powerful representational abilities, as well as a clearer semantics. The most explicit of these so far are based on variants of dynamic logic [53].

3.3.1 Cohen and Levesque

Cohen and Levesque have been attempting to solve a number of problems relating to formal characterizations of speech acts, through the use of a detailed logic of action and mental attitudes. [33] lays out the framework of the basic theory of rational action. It is based on a dynamic modal logic with a possible worlds semantics. They give axiomatizations for modal operators of beliefs and goals, and then derive intentions as persistent goals, those to which an agent is committed to either bring about or realize are unachievable.

Cohen and Levesque [34] use this logic to show how the effects of illocutionary acts can be derived from general principles of rational cooperative interaction. They claim, contrary to [90], that communicative acts are not primitive. They define what it means for an agent to be sincere and helpful, and give characterizations of imperatives and requests. They claim that recognizing the illocutionary force of an utterance is not necessary, that all that is important is that the hearer do what the speaker want, not that he recognize which act the speaker performed as a part of this process. They thus appear to be claiming that illocutionary acts should be seen as descriptive models of action, not as resources for agents. They conclude with a description of how Searle and Vanderveken's conditions on acts can be derived from their rational agent logic.

[32] extends the framework to handle Performatives. They define all illocutionary acts as attempts. Performatives like "I request you to ..." are acts which have a request component and an assertion component, and the assertion component is made true merely by the *attempt*, not the success of the action. Thus *request* is a performative verb, while *frighten* is not (because it requires a successful attempt and the success is beyond the control of the speaker), and *lie* is paradoxical when used performatively, because the explicit mention defeats the aim.

3.3.2 Sadek

Sadek uses a similar logic of rational action [82] as the basis for semantics of speech acts [81, 83]. The basic logic differs from that of Cohen and Levesque in that possible worlds are seen as time points, not whole sequences of time points. Also, Sadek aims not only at a logic that can act as a meta-theory of the rational action of agents, but one that could be used by agents themselves in reasoning to accomplish their desires and intentions. The logic also

$\langle i, REQUEST(j, a) \rangle$

FP: FP(a)[i\j] $\wedge$ K(i,Agent(j,a)) $\wedge\neg$K(i,PG(j,Done(a)))
PE: Done(a)

Figure 3: Sadek (91) Speech Act Operator

allows planning acts to occur as a result of simple reasoning about the agent's own mental state, without any extra-logical procedures. This rational agent architecture and application to communicative acts has been implemented as the basis for the dialogue management components of the AGS spoken dialogue system [84, 19].

Sadek's communicative actions are formulated in terms of their feasibility preconditions, and intended perlocutionary effects. Each of these can lead to further planning (i.e., reasoning). The intended perlocutionary effects of an act are not (always) the actual effects of the performed action, but more the goal for which the action is planned. Some of these goals will also require other actions before being satisfied. In some ways, thus, these acts are *dialogue acts* rather than just simple speech acts. An example of Sadek's formalization of agent i requesting agent j to perform action a is shown in Figure 3.[1]

3.3.3 Discussion

While the approaches of Cohen and Levesque, and Sadek have the great benefit of clarity and semantic precision, they still contain some features which make them difficult to use as a general basis for modeling the use of speech acts. For one thing, the model of belief and other attitudes is based on a model logic that incorporates omniscience of logical consequence. While this might be acceptable in a metalogic, it is clearly not sufficient for an agent of limited means. Bretier and Sadek [19] get around this by using an incomplete forward chaining reasoner which is able to handle the important inferences relating to dialogue act planning. While this is adequate for the simple kinds of interaction used in the AGS system – querying a knowledge base about potential information servers, it remains to be seen if this method is powerful enough for more general interactions, such as more general collaborative problem solving.

Another difficulty is that the notion of time and sequence of events is not rich enough to model dialogue situations. The model of action inherited from

[1] FP(a)[i\j] means that all the feasibility conditions on the requested action a are included, where i and j are switched and only i's mental state is kept. PG means persistent goal. See [81] for precise details.

dynamic logic allows only sequences of events. It also relies on predicates like **done(a)**, which means that action **a** was the last thing that happened. These together preclude adding temporally overlapping actions, since there would no longer be a clean action to serve as a transition from one world (or state) to the next.

A final difficulty is that the logic, being based on a notion of belief, goals, choice, and derived intention, is not adequate to model the whole range of speech acts. One thing that is missing is a notion of obligation or commitment to another, that is the main ingredient of commissives, for instance. Cohen and Levesque's *commitment* operator is more like a goal or intention than social commitment.

4 Extending Speech Acts to Dialogue

The most interesting use of speech acts and pragmatics generally, is not the analysis of particular utterances, but the way that utterances are combined in a coherent interaction to further some other purposes. Here language is often intertwined with other kinds of action. While engaging in dialogue was always the aim of most of the speech act work, much of the early emphasis, particularly in the computational work was on the interpretation of isolated sentences (in context). Here we describe some work that extends the analysis towards dealing with dialogue phenomena.

4.1 Dialogue Function as Action

Litman and Allen extend Allen and Perrault's work to include dialogues rather than just single utterances, and to have a hierarchy of plans rather than just a single plan [68, 69]. They describe two different types of plans: domain plans and discourse plans. Domain plans are those used to perform a cooperative task, while discourse plans, such as *clarification* and *correction*, are task-independent plans which are concerned with using the discourse to further the goals of plans higher up in the intentional structure. They also use a notion of *meta-plan* to describe plans (including discourse plans) which have other plans as parameters. Using these notions, Litman and Allen are able to account for a larger range of utterances than previous plan-based approaches, including sub-dialogues to clarify or correct deficiencies in a plan under discussion.

Carberry and Lambert [25, 64, 63], make a further distinction, adding *problem solving* plans to the discourse and domain plans of Litman and Allen. Problem solving plans are ways of building up actual domain plans, and are usually achieved by way of discourse plans.

Cohen and Levesque extend the work described in Section 3.3.1, above, into a theory of *joint intention* and multi-agent action. [35] presents an analysis of why confirmations appear in task-oriented dialogue. Using their theory of joint

intentions developed in [65], they state that the participants in one of these task-oriented dialogues have a joint intention that the task be completed. It is part of the definition of joint intention that if one party believes the object of intention to be already achieved or to be unachievable, that party must strive to make the belief mutual. It is this goal of mutual belief which drives the agent to communicate a confirmation. Although this is perhaps the first attempt in the computational literature which is explicitly concerned with a plan-based account of the generation of confirmations, it is noticeably lacking in several respects. It has no mention of how the intention to make something mutually believed turns into an intention to perform a confirmation. There is also some distance still from the logic to actual utterances. It is also not explained just what would count as a confirmation, and how one might recognize one.

4.2 Multiple Levels of Interactions

Several researchers have attempted to model the different kinds of phenomena in dialogue with different strata of action. Realizing that there is a range of action phenomena in dialogue, including but not limited to the sentence-level actions that have been concentrated on in speech act work, they extend the model to include multiple strata, even within the discourse act set. We consider some of those here.

One of the influential early classifications comes from the analysis of classroom dialogues [93], later modified in [39] and [94]. This was a system of 4 *ranks*[2], from smallest to largest: act, move, exchange, transaction. Moves are speech acts used toward particular purposes in dialogue. Exchanges are a set of acts, including some by each speaker, such as a question, answer, and feedback unit. Transactions are whole subdialogues including multiple exchanges aimed toward achieving a particular task.

Several researchers have followed up on this line of work, calling the exchange structure instead a *dialogue game*, e.g., [71, 60, 26, 1]. This has a better parallelism with the idea of moves, and generally the distinction between moves and acts is dropped. Agents generally plan at the level of games, and then execute moves to further those games along the path they desire.

Novick, realizing that the traditional speech acts are insufficient for modeling dialogue control, introduces several levels of what he calls *meta-locutionary acts*, to contrast with Austin's locutionary, illocutionary, and perlocutionary acts. The meta-locutionary acts include levels for turn-taking, repair of mutual models, reference/information, and attention. Agents used plans at each of these levels to collaborate on a task.

Traum and Hinkelman introduced a scheme which incorporates some features of these other approaches. Rather than being based on ranks, this included four

[2]according to the terminology of [51], *ranks* are hierarchical, such that an item in one rank can be built up of constituents at a lower level, like phrase and sentence in linguistics. *Levels*, on the other hand, cover different types of phenomena.

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: Traum & Hinkelman (92) Conversation Act Types

levels of acts, including turn-taking acts to coordinate who is speaking, *grounding acts* to coordinate the flow of mutual understanding, the traditional *core speech acts*, and *argumentation acts*, to handle higher-level coherence. These levels are summarized in Table 1. One difference from the other dialogue act taxonomies is that here, following [30], even the traditional speech acts are seen as multi-agent actions, since they require participation by both agents (by performance of grounding acts), before they have their full effects, such as a mutual belief or an established obligation.

Bunt [23, 24] describes several kinds of *dialogue acts*. Dialogue acts are seen as functions which update the dialogue context. Context is also divided into several types, including *linguistic* – the surrounding utterances in the dialogue, *semantic* – the underlying task and domain, *physical* – the location of the participants and the interaction in space and time, *social* – the type of situation and roles the participants play with respect to each other, including rights and obligations, and *cognitive* – the mental states of the participants. Each of these types has both *global* and *local* aspects – the former are roughly constant throughout the dialogue, while the latter can change, through the performance of dialogue acts, among other things. Bunt then divides dialogue acts into *task-oriented acts* and *dialogue control acts*. All acts affect the linguistic and cognitive context, but the former also changes the semantic context, while dialogue control acts affect the social or physical context. Bunt only considers those task-oriented acts involved in information-seeking dialogues, namely information seeking, and information providing acts [22].

On the other hand, Bunt considers a wide range of classes of dialogue control acts. Following [13], there are several types of feedback acts. Bunt considers *perception* and *interpretation* acts, which correspond to the grounding acts of [101], but also evaluation acts, including acceptance and agreement, and *dispatch* acts, such as the fulfillment of a request, which appear at the argumenta-

Communicative Act type	**Expressive function**	**Evocative function**
Statement	belief	belief judgment
Question	desire for information	the desired information
Request	desire for X	X

Table 2: Allwood (95) Functions of Some Communicative Acts

tion level. Bunt notes, (see also [29]) that giving positive feedback at one level implies problems at a higher level, while giving negative feedback at one level implies success at lower levels. Bunt also has classes of *social obligations management* acts, which place or relieve pressure on one party to do something, and a range of *interaction management acts*, including acts for turn-management, own communication management, time management, contact management, and discourse structuring.

While all of these functions certainly occur in dialogue, it is sometimes difficult to see what the principles are for clustering them. In fact, Bunt's own classification of some of these types changes, e.g., from [23] to [24]. For instance, Bunt's contact management function could easily be seen as the lowest level of feedback, as in [13, 28], while *own communication management* (the production of speech repairs) is not necessarily closely related to interaction, while feedback clearly *is*. Likewise, a variety of acts introduce and dispose of obligations [99, 10].

Allwood has also written extensively about *Communicative Acts*, generalizing speech acts to cover other, non-verbal actions that work in a similar manner [9, 11]. Allwood rejects many of the distinctions in Austin's speech act theory, preferring, like [34], to reason directly from utterance form and context to function and effect.

Allwood distinguishes two kinds of effects of speech acts: *expressive*, which serve to indicate things about the speaker or previous state of the world, and *evocative*, which tend to (or are aimed at) bring about changes in the hearer or the world. Most utterances will have both of these kinds of effects. Table 2 shows some of the acts with the main expressive and evocative functions.

In addition, each utterance will be associated with a number of obligations, both for the sender and the receiver [10]. The sender related obligations are generally like the sincerity conditions of Searle's speech act theory [87]. For example, if an agent sends a request for information, that agent is obliged to actually wonder what the answer is, need the information, and believe that the receiver can provide it. The receiver related obligations are similar to those described in [95] and in Section 7.2.2, below – mainly to evaluate the utterance and report on the evaluation. Allwood considers several types of feedback, based on four basic hierarchical communicative functions [13]: *contact*: whether the

interlocutor is willing and able to continue the interaction, *perception*: whether the interlocutor is willing and able to perceive the message, *understanding*: whether the interlocutor is willing and able to understand the message, and *attitudinal reaction* - covering responses to the message such as acceptance or rejection.

The communicative function is not restricted to the content (termed main message), but also covers other aspects such as speech management (or own communication management) [12], and interactive communication management.

5 Speech Act Based Agent Communication Languages

Speech acts, as the abstract level of how utterances change the interactional state, are thus a good level of description for describing any agent communication, not just communication in natural language. As such, some researchers have also decided that a language based on speech acts would, by itself, be a good agent communication language. Thus the sometimes difficult process of speech act interpretation is bypassed, as the agents directly identify the speech act they are performing as a part of the message itself. Some researchers studying pragmatic aspects of natural language processing (e.g., [26, 76, 106]) have used this method as a way of testing their systems without having to do natural language interpretation. Several schemes based on speech acts have also been proposed, however, as agent communication languages in their own right, e.g. [92].

5.1 KQML

A recent proposal for a universal agent communication language is the *Knowledge Query and Manipulation Language* (KQML) [50, 62]. This was devised over several years by a committee of researchers from different groups interested in having heterogeneous agents communicate in a flexible manner. Some of the requirements for an agent communication language and how KQML attempts to meet these needs are described in [72]. The underlying agency model is one of knowledge-based agents that manipulate knowledge bases, but can send queries and answers to other agents, or broker services to other providers. The main idea of this language is based on Austin's *performatives*. Each message sent or received is a performative, with an identifier stating which kind of action it is, as well as other parameters identifying aspects of the content, such as what language and ontology is being used. Agents can then pass messages around without actually having to understand any details of the content language. KQML is also purposely underspecified in terms of the range of performatives allowed: in order to be KQML compliant, an agent must only be able to send and receive messages of the required syntax, and if it does use any of the

reserved performatives, they have to have the same meaning. Designers are free to add extra performatives or not implement some of the reserved set. Some of the performatives are very familiar, such as *ask-if* and *tell*, whereas others are concerned more with the way messages are sent (all at once or one by one in a stream), or the way other agents can act as intermediaries for messages.

While KQML has been used in a number of groups as an agent communication language [44], there are still some problems with the specification. First it is lacking a precise semantics, which may cause divergences in the way performatives are used by different agents/groups. A first step towards this is provided by [61], but this only covers a few performatives. Cohen and Levesque [37] point out some further problems, in particular, vagueness of some performatives, mis-identified performatives (*achieve and broker* instead of *request*), and missing some performatives, including especially commissives. While KQML is flexible enough so that agents could avoid using the mis-identified performatives while adding the missing ones, it is disappointing that some of the basic building blocks of communication are left out.

5.2 Speech Acts as an Interface Language for Computer-mediated Collaboration

Some researchers have also advocated a limited interface based on speech acts even for human-human communication. There are generally two reasons for this: first, the computer mediator is doing more than just passing messages; it is also trying to do some analysis or updating of state, and thus the designers want to restrict the flexibility of communication to what the interface can handle (e.g., [17]). Another reason is to actually force explicitness, under the theory that this will make cooperation more productive [108].

5.3 Discussion

A number of issues arise in the attempt to bypass language interpretation and determine the speech act type directly from surface features of the message, by having agents explicitly signal the speech act type. As Sadock notes [85], the enterprise of speech act theory can be described as explicating the connections in (2).

(2) FORM $\overleftrightarrow{\text{grammar}}$ CONTENT $\overleftrightarrow{\text{pragmatics}}$ EFFECT

Thus, researchers are trying to trivialize the left relation by having agents communicate with forms precisely correlated with contents. While such limited languages can certainly reduce the ambiguity and vagueness of interpretation, it is arguable whether they can cut out such things as indirect speech acts and perlocutionary effects. Thus, even though a certain form is given a precise semantics in terms of it's direct effect, given that agents know something about

how other agents operate, they could use particular forms for ulterior purposes. Other agents could even recognize such behavior, and novel conventions of interpretation can spring up (e.g., according to the theories of [67]). Such is almost undoubtedly the case in the computer-mediated circumstances, where humans have a limited channel in which to express all of their communicative and social desires. The only thing that might stop it in the case of artificial agents is the limitations of the agents themselves.

6 Speech Act Theory Reconsidered

Given the above discussion, there are a number of issues that arise in evaluating how best to view language as action, and whether speech act theory is a useful way to view the communication of agents. The most basic question is whether speech acts should be afforded some sort of ontological status as an intermediary between performances (e.g., language utterances) and effects (e.g., changes in mental state). In spite of their initial attractiveness, some have questioned the utility of this intermediary, because of a number of factors, especially the fact that there is (still) no small commonly agreed upon set of acts that can circumscribe the range of communicative interaction. Another difficult issue is the *multifunctionality* of utterances. As described above, a single utterance can introduce a number of different kinds of change. If one were to introduce a different act for each bundle of changes, there would be a large number of acts indeed! One can get around this by adopting a multi-stratal approach of hierarchical and/or orthogonal acts, such as those described above, but still, even within a particular stratum, a single utterance may perform multiple functions. E.g., someone might be simultaneously informing about the current state while suggesting a future course of action.

Another issue is whether speech acts are to be viewed objectively – as things that actually happened, according to some specified occurrence constraints, or as subjective entities, corresponding to a particular agent's views of what has been performed. Related to this is the question of what conditions beyond just the intention to perform an act are required to have the act actually be achieved. In other words, what distinguishes an attempt from successful performance of the act.

While such a level as speech acts might not be strictly necessary, they are still useful for agents in planning communication. It may be easier for agents to plan and recognize concrete bundles of function than trying to move directly from utterance features to context change. This is true even if agents do not even have the same precise ontology of action, or if it is not always possible to discriminate which action has been performed.

7 Speech Acts in a Multi-Agent Action Theory

One of the main stumbling blocks for adequately formalizing speech acts is that it requires a very complex expressive logic of action and interaction. Speech acts always have at least two participating agents, and thus require a multi-agent logic. Moreover, in the most general case, action can proceed synchronously with other action (e.g., speech by the other agent). In addition, providing conditions and effects on performance of actions requires sophisticated reasoning about the mental states of agents. We sketch aspects of such a theory, using the conversation acts of [101, 95].

7.1 Reasoning about Multi-agent Action

A speech act theory which can account for conversations must include at least the following extensions to classical planning (e.g. STRIPS):

- temporal reasoning, including reasoning about overlapping and simultaneous actions
- uncertainty: attempted actions may fail to achieve their desired results, unexpected results may follow
- multiple agents, each with individual mental states
- cooperation among agents
- real-time resource-bounded reasoning
- integration of planning and acting

There is a large amount of research dedicated to addressing these problems, much more than can be summarized here. [97] explores some of the complexities involved in reasoning and acting in a richer environment. The annual European workshops on Modeling Autonomous Agents in a Multi-Agent World (MAAMAW) (reprinted in e.g., [40, 41, 107]) contain a variety of approaches to these problems.

7.2 Social Attitudes

One of the key features of *speech acts* as opposed to physical actions is that their main effects are on the mental and interactional states of agents, rather than on the state of some external domain. The attitudes of belief, desire, and intention are familiar to agency theories, (see for example chapters [79]s in this volume) but for a comprehensive theory of speech acts we must also go beyond these and consider *social attitudes*, which require component parts on the behalf of multiple agents. A big issue for these attitudes is whether they are basic or can be composed from more primitive individual attitudes such as belief and

intention. While the holding of individual attitudes can certainly be deduced from holding of social attitudes, we will treat Mutual Belief and Obligations as basic, while proposing a compositional definition of executing a joint plan.

7.2.1 Mutual Belief and Grounding

Mutual belief is one commonly assumed social attitude. Most of the theories of speech acts as plans described above have as some of the main effects of speech acts the addition of some new *mutual beliefs*. Mutual beliefs are also taken to be some of the prerequisites for felicitous utterance of speech acts. But just what are mutual beliefs? This section reviews some of the proposals for how to represent the properties of mutual beliefs in terms of simpler beliefs, and how one could acquire new mutual beliefs. As with *belief* and *knowledge*, there have been a variety of names for a cluster of related concepts, including "mutual knowledge", "common knowledge", "mutual belief", and "shared information". In the discussion below, we generally use the term used by the author under discussion, but treat these terms as synonymously meaning *mutual belief*, in which what is mutually believed by a group of agents is not necessarily actually true.

Formulations of Mutual Belief

While people agree for the most part about the intuitions underlying the phenomenon of mutual belief, there have been a variety of different ways proposed of modeling it. Barwise [18] compares model theories for three different formulations.

Schiffer uses what Barwise calls "the *iterate* approach" [18, p. 202]. He defines mutual knowledge between two agents A and S of a proposition p, $K^*_{SA}p$ as [86, p. 30]: $K_Sp \wedge K_Ap \wedge K_SK_Ap \wedge K_AK_Sp \wedge K_SK_AK_Sp \wedge K_AK_SK_Ap \wedge \cdots$ It is thus an infinite conjunction of nested beliefs. This approach has since been adopted by many others, including Allen [3] and Perrault, who provides an elegant default logic theory of how to obtain each of these beliefs given prior knowledge and a conversational setting [77].

Barwise credits Harman with the *fixed-point approach*. Harman formulates mutual knowledge as "knowledge of a self-referential fact: A group of people have mutual knowledge of p if each knows p *and we know this*, where *this* refers to the whole fact known" [54, p. 422]. As Barwise points out, the fixed-point approach is strictly stronger than the iterate approach, because it includes as well the information that the common knowledge is itself common knowledge. It also replaces an infinite conjunction with a self-referential one.

The final approach discussed by Barwise is the *shared-situation approach*. He credits it to Lewis. Lewis formulates rules for common knowledge as follows [67, p. 56]:

Let us say that it is *common knowledge* in a population **P** that **X** if and only if some state of affairs **A** holds such that:

1. Everyone in **P** has reason to believe that **A** holds.
2. **A** indicates to everyone in **P** that everyone in **P** has reason to believe that **A** holds.
3. **A** indicates to everyone in **P** that **X**.

This schema is also used by Clark and Marshall, and is apparently the one which Barwise himself endorses.

[31] uses a *belief spaces* approach to model belief. Each space contains a set of propositions believed by an agent. Nested belief is represented by nested spaces. There is a space for the system's beliefs (SB) which can contain a space for the system's beliefs about the user's beliefs (SBUB) which in turn can contain a space for the system's beliefs about the user's beliefs about the system's beliefs (SBUBSB). If Cohen were to adopt the iterated approach directly, it would require an infinity of belief spaces. Instead, he takes the space one deeper than the deepest which contains any non-mutual beliefs, and points it to its parent space, thus creating a loop, where each even nesting is the same as every other even nesting. Now each of the nested beliefs in the iterated approach can be generated or seen to be present in his belief spaces, by iterating through the loop. This approach shares some features with the fixed-point approach (the self-referentiality) and it allows quick determination of whether mutual belief exists (by searching for a loop), unlike the iterated approach, but it is in fact not as strong as the fixed-point approach because the higher-order implications of the fixed-point approach, such as mutual belief about the mutual belief, cannot be represented.

A slight modification is to add a separate kind of space, a *mutual belief* space to represent mutual beliefs. This is the approach taken by Bruce and Newman [21]. The Rhetorical knowledge representation system [7] also uses a mutual belief space, but disallows nested beliefs within a mutual belief space, giving essentially the power of Cohen's system. This also seems to be the approach used by Maida [70].

How can Mutual Belief be Achieved?

While it is still somewhat controversial how best to formally model mutual belief, an interesting question is how such mutual belief gets established among interacting agents. While it has long been known that is impossible to guarantee the establishment of mutual knowledge in an environment where message transmission could fail [52], most formulations of speech acts have gone to the other extreme, and assumed mutual belief after the performance of any utterance within a shared situation (e.g., as the effect of a single agent speech act). However, examining actual conversation reveals that there is a process of feedback that accompanies initial utterances, and, in task-oriented spoken language,

it is generally only after some sort of acknowledgment that an assumption of mutual belief is made. Furthermore, lack of understanding can be signaled with some sort of repair or request for repair. In cases in which the original speaker does not receive any feedback, one can observe requests for acknowledgment or repetitions and refashionings of the original contribution in an attempt to elicit some kind of feedback.

While the assumption of mutual belief resulting directly from a single utterance can be seen simply as an idealization (like modeling belief with a modal logic such as S5, with the resulting properties of logical omniscience), it is one with unfortunate consequences if used as the basis for models of speech acts and inter-agent communication. The consequences are twofold: first, a reasoner will make incorrect predictions about the mental state at particular times, and more importantly, the agent will be unable to recognize the relevance of or necessity for performing the feedback acts which actually establish the mutual understanding.

Clark and Shaefer refer to the process of reaching mutual belief (or common ground) as *grounding* [30]. They present a descriptive model, in terms of presentation and acceptance phases that allow them to track the augmentation of common ground as the conversation proceeds. Their model is not well-suited for an on-line agent involved in dialogue, however, since it requires examination of subsequent spans of text in order to determine the boundaries of these phases.

In [98, 95], we developed a computational account of the grounding process. This account was based on speech act theory, using actions to introduce, acknowledge, and repair material. Traditional speech acts, such as *inform*, and *request* are now seen as multi-agent actions, which require participation by both parties to have their full effects (such as the mutual belief that one speaker wanted the other speaker to believe something).We introduced a level of dialogue structure called *discourse units* (DUs), at which these core speech acts are completed. These DUs are built up by single-utterance *grounding acts*. Recognizing the fact that multiple types of action occur in conversation, we extended speech act theory to the multi-level *conversation act* theory, described in [101]. As well as the grounding and core speech acts, there are also levels to model turn-taking behavior and higher order coherence of dialogue. A finite automaton was used to track the state of a DU, given a sequence of grounding acts in conversation. This model could also be used to predict possible subsequent acts as well as determine which act(s) must be performed in order to have a grounded DU (which would thus realize the effects of the constituent core speech acts).

7.2.2 Obligations

We claim that *Obligations* are necessary for modeling many social situations including natural language conversation. For example, they are necessary for capturing the effects of some speech acts, such as requests [99]. Obligations represent what an agent *should* do, according to some set of norms; its formal

source of obligation	**obliged action**
S_1 Accept or Promise A	S_1 achieve A
S_1 Request A	S_2 address Request: accept **or** reject A
S_1 YNQ whether P	S_2 Answer-if P
S_1 WHQ P(x)	S_2 Inform-ref x
utterance not understood or incorrect	repair utterance
S_1 Initiate DU	S_2 acknowledge DU
Request Repair of P	Repair P
Request Acknowledgment of P	acknowledge P

Table 3: Traum & Allen (94) Sample Obligation Rules

aspects are examined using Deontic Logic (e.g., [105, 73]). Generally, obligation is defined in terms of a modal operator often called *permissible*. An action is *obligatory* if it is not permissible not to do it. An action is *forbidden* if it is not permissible.

Just because an action is obligatory with respect to a set of rules R does not mean that the agent will perform the action. So we do not adopt the model suggested by [91] in which agents' behavior cannot violate the defined social laws. If an obligation is not satisfied, then this means that one of the rules must have been broken. We assume that agents generally plan their actions to violate as few rules as possible, and so obligated actions will usually occur. But when they directly conflict with the agent's personal goals, the agent may choose to violate them (and perhaps suffer the consequences of not meeting its obligations). Obligations are quite different from and cannot be reduced to intentions and goals. In particular, an agent may be obliged to do an action that is contrary to its goals (for example, consider a child who has to apologize for hitting her younger brother). [42] use obligations in a similar way, noting also that authority (such as a pre-existing hierarchical relationship) can be important in the ability to force obligations on others.

In [99] we argued that obligations play an important role in accounting for many of the interactions in dialog. For example, Table 3 shows the obligations resulting from the performance of speech acts. Obligations do not replace the plan-based model of speech acts (e.g., [38, 8]) but augment it. The resulting model more readily accounts for discourse behavior in adversarial situations and other situations where it is implausible that the agents adopt each others' goals. The obligations encode learned social norms, and guide each agent's behavior without the need for intention recognition or the use of shared plans at the discourse level. While such complex intention recognition may be required in some interactions, it is not needed to handle the typical interactions of everyday discourse.

The deliberation process in a social situation must take obligations into account, in addition to goals and intentions. In forming new intentions, sometimes an agent will choose to pursue its obligations rather than its goals. It is important for the agent to reason about both of these notions, so that this choice can be made explicit (either in the agent design or by the agent itself). There is also an illuminating relationship between this deliberation process and the notion of initiative in dialogue. Following the initiative of the other can be seen as an *obligation-driven* process, while leading the conversation will be *goal-driven*.

There is another related concept that also occurs as the result of many speech acts: a commitment to certain facts being true [10]. These relate to the sincerity conditions postulated by Searle. Thus when an agent makes a claim, the agent is committed to the truth of that claim. Unlike the obligations considered above, there is no direct connection between this and future actions. The agent is obliged to keep these commitments to others accurate, but, if the agent comes to believe these are not accurate, there is no special preference for making them true – an equally valid approach is to just inform the other agent that the reported facts no longer hold, assuming the question is still of interest. This is to be contrasted with the kind of commitment to (or *intending that*) future states that is related to intention, and requires some action when the agent believes the state will not hold.

7.2.3 Coordinated Activity, Shared Plans and Joint Intention

Another common social attitude is that of *Joint intention* [35] or *shared plan* [47]. These concepts are used to model the propensity of a collaborative team to act. The intuition here is that it is more than just a collection of individual intentions and beliefs that is responsible for the coordinated teamwork activity. Although most agree that some sort of collective intentional attitude is useful in formalizing an account of cooperative behavior, it is still fairly controversial what the properties of such an attitude should be and how this collective intention is related to the individual intentions of the participants. Important questions include: how do shared intentions guide individual action, and how can individual beliefs and intentions come together to form shared intentions? This section reviews some proposals for such attitudes.

Lewis [67] defined a *Convention* as a situation in which there is some regularity R in behavior in a population, and everyone conforms to R, everyone expects everyone else to conform to R, and everyone prefers to conform to R, given that everyone else will. A typical example is which side of the road to drive on. In England and Japan it is the left side, in America and Europe, the right. It doesn't really matter to the drivers which side to drive on, as long as everyone agrees. Coordinated activity is thus seen as individual intention in a state of mutual knowledge about norms. Knowledge of conventions serve to make it in the mutual self-interest of each of the members of the population to follow along.

Grosz and Sidner [47] take basically the same viewpoint. They extend Pollack's definition of a Simple Plan [78], to that of a *SharedPlan*, which is formalized as a set of mutual beliefs about the executability of actions and the intentions of particular agents to perform parts of that action. They also present some conversational default rules based on cooperativeness to use communication to add to the shared beliefs. The shared plan formalism is further developed in [48, 49].

Searle [89] starts with the intuition that collective intention is not just a summation of individual intentions. He wants to distinguish between just following a convention and actual cooperative activity. He postulates that *we-intentions* are a primitive form of intentionality, not reducible to individual intentions. There is still a problem of how *we-intentions* can produce the individual intentions necessary for an individual to act.

Cohen and Levesque present their own theory, not in terms of individual intentions (which also aren't primitive in their theory) but in terms of mutual belief and weak mutual goals [65, 36]. Their formulation says that the individuals each have the goals to perform the action until they believe that either it has been accomplished or becomes impossible. Also in the event of it becoming completed or impossible, the agents must strive to make this belief mutual. This framework is also used to explain certain types of communicative behavior such as confirmations as described above.

In [95], we developed a similar notion, that of agents executing a multi-agent plan. This is an extension of the notion of executing a plan described in [100]. A group of agents $\{A_i\}$ is executing a multi-agent plan **MP** iff:

1. Each A_i is executing a single-agent plan $\mathbf{MP}_i$, which has as its actions A_i's actions from **MP**, and as its constraints the constraints of **MP**, as well as the occurrence of all actions by other agents (thus A_i will be committed to the occurrence of the actions of others).

2. Each A_i is obliged to the other agents to perform her own actions as part of the multi-agent plan.

This formulation has several differences from the other works mentioned. First, no mutual belief is stipulated as a necessary component of the multi-agent plan execution. While mutual beliefs may sometimes be important for collaboration, and particularly for decisions about adopting plans and repairing plan executions, they are not strictly necessary for this kind of teamwork. In this framework, it is the personal commitment to the occurrence of the actions of others, and the obligations *to* those others (as well as the personal intention to perform the action) that forms the glue binding the collaborating team together. While any agent may break the team at any time by dropping these commitments and intentions, the obligations will remain until the agent is released, and it is this which motivates such actions as letting another know that an action has been performed or is deemed impossible. Even with notions

of mutual belief of intentions (as in the SharedPlan formalism of [47, 48]), or commitment to inform an agent if an action is performed or impossible (as in the joint intentions of [35]), it is hard to see, in practice, what keeps an agent adhering to these commitments when its personal goals diverge.

7.3 Defining Speech Acts

Given the rich logic of mental state and action partially developed above, how can we use this to give precise definitions of speech acts? There are actually several important components for speech acts. One is the essential characteristics of what *is* and *isn't* a particular act. Another important criteria is how one can recognize whether such an act has been performed. Since speech acts have as part of their defining characteristics (both enabling conditions and effects) aspects of the mental states of communicating agents, a particular observer (even a participant) may not have access to all the necessary information, and so may not be able to tell (with certainty) whether a particular act has occurred, or whether a particular bit of dialogue is the performance of a particular kind of act. A third component is the use of speech acts in planning – how can agents plan to use speech acts to accomplish their (non-linguistic) goals, and how are these plans related to actual behavior and the realization of such acts? We consider each of these in turn for some of the speech act types of [101], shown in Table 1.

The turn-taking acts can be easily described using an interval temporal logic, such as that of [2, 4, 5]. Of interest are periods of each speaker's speech and silence, and the turn, itself. Events such as starting, continuing and stopping speaking, as well as other features such as prosody and expectations will be interpreted as turn-taking acts and serve to delimit individual turns. Even these simple acts would be difficult with a simple dynamic logic, since overlapping intervals of the two agents speaking must be considered.

Since, according to the theory of [101], all core speech acts must be grounded to have their full effect, it will be instructive to first look at the grounding acts. We use the formulation of executing a multi-agent plan described in Section 7.2.3 to formally model the grounding process described earlier in that section. Agents involved in a task oriented dialogue are assumed to be executing a specialization of the abstract plan recipe shown in Figure 4. This abstract recipe will be called CR (for Communication Recipe). Our claim is that successful execution of (a specialization of) this recipe will result in the (mutually assumed) mutual belief between the two agents that INITIATOR(CR) has communicated CONTENT(CR) to RESPONDER(CR). Agents engaged in conversation can be modeled as executing multi-agent plans that are specializations of this recipe. We will call any plan which has as its recipe a specialization of CR a *conversation plan*.

The acts of presenting and acknowledging the content are broken into some indeterminate number of conceptual sub-acts, about at the granularity of the

Actions(CR) = $\{\text{present}_i \mid \text{present}_i = \text{Present}(\text{Agent1}_i, \text{Content}_i, \text{Recipient1}_i, \text{t1}_i)\}$
$\cup\{\text{ack}_i \mid \text{ack}_i = \text{Ack}(\text{Agent2}_i, \text{Content}_i, \text{Recipient2}_i, \text{t2}_i)\}$

Constraints(CR) =	{
Temporal constraints	$\forall i(\text{Before}(\text{t1}_i, \text{t2}_i))$,
Agent constraints	$\forall i(\text{Agent1}_i = \text{Recipient2}_i = \text{INITIATOR(CR)},$
	$\text{Agent2}_i = \text{Recipient1}_i = \text{RESPONDER(CR)})$,
Object Constraints	$\text{CONTENT(CR)} = \cup_i \text{Content}_i\}$

Figure 4: Traum (94) Plan Recipe for Communication (Recipe CR)

propositions in [58]. The only constraints on performance of these is that for a particular piece of the content, the presentation must come before the acknowledgment, and both must occur to achieve mutual belief. Constraints on exactly what types of executions can present and acknowledge the above contents will be determined by conventions of the particular language used and the communicative contexts. For any given execution, some parts of the content will be expressed explicitly as part of the compositional conventional meaning of the utterance, and others will be presented implicitly by conventions of situated meaning and Gricean implicatures. In [95], the grounding acts from [98] were given formal specifications as executions related to conversation plans. As well as the components for executing a multi-agent plan, additional definitions are given in a situation logic for actions of starting, continuing, completing, and repairing the execution of a plan.

The content of the grounding actions described above will include the speech act type (illocutionary force) of the actions as well as the contents of core speech acts. Thus, the performance of each core speech act could also be decomposed as the performance of a conversation plan, with appropriate constraints on the contents. All core speech acts thus include the effects in (3). In addition, some of the other effects are shown in Table 4. Here, *Committed* is meant to be a social commitment to a state of affairs, rather than the individual commitment of [33].

(3) MB(A,B,occurs(act)) $\wedge$ MB(A,B,holds(effect(act))).

The actual effects of interest will generally come from reasoning about the reasoning and deliberation of other agents, such as B coming to believe ϕ, or even B coming to believe that A believes B, given that A commits herself to its truth.

The definitions of the argumentation acts will be sequences of core speech acts, with constraints on the timing and content. E.g., that the answer actually provides information asked for by the question.

act	effects
Promise(A, B, Act)	Obliged(A, B, Act)
Accept(A, B, Act)	Committed(A, B, Cando(A, Act))
Request(A, B, Act)	Obliged(B, A, Do(A, Accept(Act) $\vee$ Reject(Act)))
	Committed(A, B, Believe(A, Cando(B, Act)))
	Committed(A, B, Desire(A, Do(B, Act)))
YNQ(A, B, ϕ)	Committed(A, B, $\neg$Knowif(A, ϕ))
	Committed(A, B, Desire(A, Knowif(A, ϕ)))
Assert(A, B, ϕ)	Committed(A, B, ϕ)

Table 4: Sample Speech Act Effects

7.4 Planning Speech Acts

Definitions of speech acts only capture part of the reasoning necessary for planning speech acts and utterances. No matter how exhaustive the list of effects, agents may still care not about the direct effects, but the extended perlocutionary effects springing from the performance of the act in context. Likewise, agents may choose to perform acts which have only a certain probability of producing the effect desired. Acts can also be planned at different levels – one may plan a whole argumentation act (or dialogue game), or merely a single speech act.

Because of the on-line, interactive nature of dialogue, it is generally not very useful to plan very far in advance. Often, an agent can not predict the future actions of others with very much accuracy. Thus it usually only pays to plan only a few utterances in advance, while reacting to the actions of others in a manner so as to move towards achievement of the agent's goals. Some kinds of dialogues, such as arguments or negotiations may require more involved planning, while often casual conversation can be completely unplanned except for the very next utterance.

7.5 Recognizing Speech Acts

An agent will combine features of an input utterance with aspects of the current context (including the mental states of agents, as well as the visual field and the set of previous utterances) to decide what acts have been performed.

One big question for recognizing speech acts is whether one should be trying just to recognize the acts themselves or the intentions (e.g., plans) that motivated the other agent to produce the act. Often, as in the case of indirect speech acts, one cannot easily separate these questions – one may come to decide that an agent has performed a request by deciding that the agent wanted one to do something. Still, intentions and performed actions can sometimes diverge,

e.g., an agent can mistakenly or unintentionally perform an action, which may carry with it all of the responsibilities and other effects as if it were intended. Likewise, intentions which were not (fully) realizes will not carry the burden of imposing obligations and social commitments.

Sometimes, especially when agents are cooperative, deducing the intention is actually more important/useful than what was actually done. Moreover, deducing *why* an agent performed the act can lead to helpful behavior (or avoidance of traps in a competitive dialogue).

The question arises, however, as to how much of the plan it is necessary to infer. McRoy, for instance, challenges the utility of the deep intention recognition in [8], claiming instead that considering the possible actions and their immediate effects (including expectations set up for future utterances) is sufficient, especially when combined with a facility to repair erroneous conclusions [74, 75].

While recognizing the intentions and plans of a speaker can play a useful role in act interpretation, and may in fact be crucial to an agent's actual success in some domains, one must also remember the on-line nature of dialogue in this case as well. More prominence should be placed on interpreting the act itself, and performing this interpretation in real time. Since the speaker is easily accessible by the interpreting agent, if there are difficulties understanding either the act itself, or the motivations, the agent can query to find out what was meant or why it was said, through the grounding process.

8 Conclusions

In summary, we believe that speech acts are a good link between the mental states of agents and purposeful communication. While a comprehensive theory of speech acts will strain most contemporary theories of rational agency, requiring a very expressive theory of action and mental state, it also provides a good testbed for a theory of agency in a multi-agent world. While many of the specific concerns of natural language speech act work, such as the interpretation of ambiguous utterances, may be alleviated by using more restricted agent communication languages, many of the same issues will arise for cooperative systems that must infer intentions of others.

Acknowledgments

The author's ideas on speech acts reported above have been greatly aided by conversations with many others, especially: James Allen, Jens Allwood, Harry Bunt, Peter Heeman, Elizabeth Hinkelman, Massimo Poesio, Teresa Sikorski, and Richmond Thomason. The author was partially supported while writing this paper by the U.S. Army Research Office under contract/grant number

DAAH 04 95 10628 and the U.S. National Science Foundation under grant IRI-9311988. Some of the work described above was developed in collaboration with James Allen and supported by ONR/DARPA under grant number N00014-92-J-1512, by ONR under research grant number N00014-90-J-1811, and by NSF under grant number IRI-9003841.

References

[1] Gabriella Airenti, Bruno G. Bara, and Marco Colombetti. Conversation and behavior games in the pragmatics of dialogue. *Cognitive Science*, 17:197–256, 1993.

[2] James F. Allen. Maintaining knowledge about temporal intervals. *Communications of the ACM*, 26(11):832–843, November 1983.

[3] James [F.] Allen. Recognizing intentions from natural language utterances. In Michael Brady and Robert C. Berwick, editors, *Computational Models of Discourse*. MIT Press, 1983.

[4] James F. Allen. Time and planning. In R. Pelavin J. Allen, H. Kautz and J. Tenenberg, editors, *Reasoning About Plans*. Morgan Kaufmann, 1991.

[5] James F. Allen and George Ferguson. Actions and events in interval temporal logic. *Journal of Logic and Computation*, 4(5), 1994.

[6] James [F.] Allen, James Hendler, and Austin Tate, editors. *Readings In Planning*. Morgan Kaufmann, 1990.

[7] James F. Allen and B. W. Miller. The RHET system: A sequence of self-guided tutorials. Technical Report 325, Department of Computer Science, University of Rochester, July 1991.

[8] James F. Allen and C. Raymond Perrault. Analyzing intention in utterances. *Artificial Intelligence*, 15(3):143–178, 1980.

[9] Jens Allwood. *Linguistic Communication as Action and Cooperation*. PhD thesis, Göteborg University, Department of Linguistics, 1976.

[10] Jens Allwood. Obligations and options in dialogue. *Think Quarterly*, 3:9–18, 1994.

[11] Jens Allwood. An activity based approach to pragmatics. Technical Report (GPTL) 75, Gothenburg Papers in Theoretical Linguistics, University of Göteborg, 1995.

[12] Jens Allwood, Joakim Nivre, and Elisabeth Ahlsen. Speech management: On the non-written life of speech. Technical Report (GPTL) 58, Gothenburg Papers in Theoretical Linguistics, University of Göteborg, 1989.

[13] Jens Allwood, Joakim Nivre, and Elisabeth Ahlsen. On the semantics and pragmatics of linguistic feedback. *Journal of Semantics*, 9, 1992.

[14] G. E. M. Anscombe. *Intention*. Basil Blackwell, 1957.

[15] Douglas Appelt and Kurt Konolige. A nonmonotonic logic for reasoning about speech acts and belief revision. In *Proceedings of Second International Workshop on Non-Monotonic Reasoning*, pages 164–175, 1988.

[16] J. A. Austin. *How to Do Things with Words*. Harvard University Press, 1962.

[17] M. Baker and C. Lund. Flexibly structuring interaction in a CSCL environment. In *Proceedings of the European Conference on AI in Education.*, pages 401–407, 1996.

[18] Jon Barwise. *The Situation in Logic*, chapter 9: On the Model Theory of Common Knowledge. CSLI Lecture Notes: Number 17. Center for The Study of Language and Information, 1989.

[19] P. Bretier and M. D. Sadek. A rational agent as the kernel of a cooperative spoken dialogue system: Implementing a logical theory of interaction. In J. P. Müller, M. J. Wooldridge, and N. R. Jennings, editors, *Intelligent Agents III — Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages (ATAL-96)*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Heidelberg, 1996.

[20] Bertram C. Bruce. Generation as a social action. In *Theoretical Issues in Natural Language Processing-1*, pages 64–67, 1975. Also appears in [46], pp. 419-422.

[21] Bertram C. Bruce and Denis Newman. Interacting plans. *Cognitive Science*, 2:195–233, 1978.

[22] H. C. Bunt. Information dialogues as communicative action in relation to partner modelling and information processing. In M.M Taylor, F. Neel, and D. G. Bouwhuis, editors, *The Structure of Multimodal Dialogue*. Elsevier Science Publishers B.V., 1989.

[23] Harry Bunt. Context and dialogue control. *Think Quarterly*, 3:19–31, 1994.

[24] Harry Bunt. Interaction management functions and context representation requirements. In *Proceedings of the Twente Workshop on Langauge Technology: Dialogue Management in Natural Language Systems (TWLT 11)*, pages 187–198, 1994.

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

[26] Jean Carletta. *Risk-taking and Recovery in Task-Oriented Dialogue*. PhD thesis, University of Edinburgh, 1992.

[27] Herbert H. Clark. *Arenas of Language Use*. University of Chicago Press, 1992.

[28] Herbert H. Clark. Managing problems in speaking. *Speech Communication*, 15:243 – 250, 1994.

[29] Herbert H. Clark and Edward F. Schaefer. Collaborating on contributions to conversation. *Language and Cognitive Processes*, 2:1–23, 1987.

[30] Herbert H. Clark and Edward F. Schaefer. Contributing to discourse. *Cognitive Science*, 13:259–294, 1989. Also appears as Chapter 5 in [27].

[31] Phillip R. Cohen. *On Knowing What to Say: Planning Speech Acts*. PhD thesis, University of Toronto, 1978. Reproduced as TR 118 Department of Computer Science, University of Toronto.

[32] Phillip R. Cohen and Hector J. Levesque. Performatives in a rationally based speech act theory. In *Proceedings ACL-90*, pages 79–88, 1990.

[33] Phillip R. Cohen and Hector J. Levesque. Persistence, intention, and commitment. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*. MIT Press, 1990.

[34] Phillip R. Cohen and Hector J. Levesque. Rational interaction as the basis for communication. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*. MIT Press, 1990.

[35] Phillip R. Cohen and Hector J. Levesque. Confirmations and joint action. In *Proceedings IJCAI-91*, pages 951–957, 1991.

[36] Phillip R. Cohen and Hector J. Levesque. Teamwork. *Nous*, 35, 1991.

[37] Phillip R. Cohen and Hector J. Levesque. Communicative actions for artificial agents. In *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, pages 65–72, June 1995.

[38] Phillip R. Cohen and C. R. Perrault. Elements of a plan-based theory of speech acts. *Cognitive Science*, 3(3):177–212, 1979.

[39] M. Coulthard, M. Montgomery, and D. Brazil. Developing a description of spoken discourse. In M. Coulthard and M. Montgomery, editors, *Studies in Discourse Analysis*, pages 1–50. Routledge & Kegan Paul, 1981.

[40] Y. Demazeau and J. P. Muller, editors. *Decentralized A.I.* Elsevier Science Publishers B. V., 1990.

[41] Y. Demazeau and J. P. Muller, editors. *Decentralized A.I. 2.* Elsevier Science Publishers B. V., 1991.

[42] F. Dignum and B. van Linder. Modeling social agents: Communication as action. In J. P. Müller, M. J. Wooldridge, and N. R. Jennings, editors, *Intelligent Agents III — Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages (ATAL-96)*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Heidelberg, 1996.

[43] Richard E. Fikes and Nils J. Nilsson. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208, 1971. Also appears in [6].

[44] Tim Finin, Richard Fritzson, Don McKay, and Robin McEntire. Kqml as an agent communication language. In *Proceedings of the Third International Conference on Information and Knowledge Management (CIKM'94)*, 1994.

[45] H. P. Grice. Meaning. *Philosophical Review*, 66:377–88, 1957.

[46] Barbara J. Grosz, Karen Sparck Jones, and Bonnie Lynn Webber, editors. *Readings In Natural Language Processing.* Morgan Kaufmann, 1986.

[47] Barbara J. Grosz and Candace L. Sidner. Plans for discourse. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* MIT Press, 1990.

[48] Barbara J. Grosz and Sarit Kraus. Collaborative plans for complex group action. *Artificial Intelligence*, 86(2):269–357, 1996.

[49] Barbara J. Grosz and Sarit Kraus. The Evolution of SharedPlans. In this volume.

[50] External Interfaces Working Group. Draft specification of the kqml agent-communication language. available through the WWW at: http://www.cs.umbc.edu/kqml/papers/, 1993.

[51] M. A. K. Halliday. Categories of the theory of grammar. *Word*, 17:241–92, 1961.

[52] J. Y. Halpern and Y. Moses. Knowledge and common knowledge in a distributed environment. *Journal of the ACM*, 37(3):549–587, 1990.

[53] D. Harel. *First Order Dynamic Logic.* Springer-Verlag, 1979.

[54] Gilbert Harman. Review of *linguistic behaviour* by jonathan bennett. *Language*, 53:417–424, 1977.

[55] Elizabeth A. Hinkelman. *Linguistic and Pragmatic Constraints on Utterance Interpretation*. PhD thesis, University of Rochester, 1990.

[56] Elizabeth A. Hinkelman and James F. Allen. Two constraints on speech act ambiguity. In *Proceedings ACL-89*, pages 212–219, 1989.

[57] Jaakko Hintikka. *Knowledge and belief; an introduction to the logic of the two notions*. Cornell University Press, 1962.

[58] Jerry Hobbs. Ontological promiscuity. In *Proceedings ACL-85*, pages 61–69, 1985.

[59] Kurt Konolige. Hierarchic autoepistemic theories for nonmonotonic reasoning: preliminary report. In *Proceedings of Second International Workshop on Non-monotonic Reasoning*, pages 42–59, 1988.

[60] Jacqueline C. Kowtko, S. Isard, and G. Doherty. Conversational games within dialogue. In *Proceedings of the ESPRIT Workshop on Discourse Coherence*, 1991.

[61] Yannis Labrou and Tim Finin. A semantics approach for kqml – a general purpose communication language for software agents. In *Proceedings of the Third International Conference on Information and Knowledge Management (CIKM'94)*, 1994.

[62] Yannis Labrou and Tim Finin. A proposal for a new kqml specification. Technical Report CS-97-03, Computer Science and Electrical Engineering Department, University of Maryland Baltimore County, 1997.

[63] Lynn Lambert. *Recognizing Complex Discourse Acts: A Tripartite Plan-Based Model of Dialogue*. PhD thesis, University of Delaware, 1993. Reproduced as TR 93-19 Department of Computer and Information Science, University of Delaware.

[64] Lynn Lambert and Sandra Carberry. A triparite plan-based model of discourse. In *Proceedings of the 29^{th} Annual Meeting of the Association for Computational Linguistics*, pages 47–544, 1991.

[65] Hector J. Levesque, Phillip R. Cohen, and Jose H. T. Nunes. On acting together. In *Proceedings AAAI-90*, pages 94–99, 1990.

[66] Stephen C. Levinson. *Pragmatics*. Cambridge University Press, 1983.

[67] David K. Lewis. *Convention: A Philosophical Study*. Harvard University Press, 1969.

[68] Diane J. Litman. *Plan Recognition and Discourse Analysis: An Integrated Approach for Understanding Dialogues.* PhD thesis, University of Rochester, 1985. Reproduced as TR 170 Department of Computer Science, University of Rochester.

[69] Diane J. Litman and James F. Allen. Discourse processing and common sense plans. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* MIT Press, 1990.

[70] Anthony S. Maida. Belief spaces: Foundations of a computational theory of belief. Technical Report Cs-84-22, Pennsylvania State University Department of Computer Science, December 1984.

[71] William C. Mann. Dialogue games: Conventions of human interaction. *Argumentation*, 2:511–532, 1988.

[72] J. Mayfield, Y. Labrou, and T. Fini. Desiderata for agent communication languages. In *Working Notes AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environment.*, March 1995.

[73] L. Thorne McCarty. Permissions and obligations: An informal introduction. Technical Report LRP-TR-19, Dept. of Computer Science, Rutgers University, 1986.

[74] Susan McRoy. *Abductive Interpretation and Reinterpretation of Natural Language Utterances.* PhD thesis, University of Toronto, 1993. Reproduced as TR CSRI-288 Department of Computer Science, University of Toronto.

[75] Susan W. McRoy and Graeme Hirst. The repair of speech act misunderstandings by abductive inference. *Computational Linguistics*, 21(4):5–478, 1995.

[76] David Novick. *Control of Mixed-Initiative Discourse Through Meta-Locutionary Acts: A Computational Model.* PhD thesis, University of Oregon, 1988. also available as U. Oregon Computer and Information Science Tech Report CIS-TR-88-18.

[77] C. Raymond Perrault. An application of default logic to speech act theory. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* MIT Press, 1990.

[78] Martha E. Pollack. Plans as complex mental attitudes. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* MIT Press, 1990.

[79] A. Rao and M. Wooldridge Foundations of Rational Agency. In this volume.

[80] R. Reiter. A logic for default reasoning. *Artifical Intelligence*, 13(1,2):81–132, April 1980.

[81] M. D. Sadek. Dialogue acts are rational plans. In *Proceedings of the ESCA/ETR workshop on multi-modal dialogue*, 1991.

[82] M. D. Sadek. A study in the logic of intention. In C. Rich, W. Swartout, and B. Nebel, editors, *Proceedings of Knowledge Representation and Reasoning (KR&R-92)*, pages 462–473, 1992.

[83] M. D. Sadek. Communication theory = rationality principles + communicative act models. In *Proceedings of the AAAI Workshop on Planning for Interagent communication*, 1994.

[84] M. D. Sadek, A. Ferrieux, and A. Cozannet. Towards an artificial agent as the kernel of a spoken dialogue system: A progress report. In *Proceedings of the AAAI Workshop on Integration of Natural Language and Speech Processing*, 1994.

[85] J. M. Sadock. Comments on Vanderveken and on Cohen and Levesque. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication*, pages 257–270. MIT Press, 1990.

[86] Stephen R. Schiffer. *Meaning.* Oxford University Press, 1972.

[87] John R. Searle. *Speech Acts.* Cambridge University Press, New York, 1969.

[88] John R. Searle. A classification of illocutionary acts. *Language in Society*, 5:1–23, 1976.

[89] John R. Searle. Collective intentions and actions. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* MIT Press, 1990.

[90] John R. Searle and Daniel Vanderveken. *Foundations of Illocutionary Logic.* Cambridge University Press, 1985.

[91] Yoav Shoham and Moshe Tennenholtz. On the synthesis of useful social laws for artificial agent societies. In *Proceedings AAAI-92*, pages 276–281, 1992.

[92] Candace L. Sidner. An artificial discourse language for collaborative negotiation. In *Proceedings of the forteenth National Conference of the American Association for Artificial Intelligence (AAAI-94)*, pages 814–819, 1994.

[93] J. M. Sinclair and R. M. Coulthard. *Towards an analysis of Discourse: The English used by teachers and pupils.* Oxford University Press, 1975.

[94] Anna-Brita Stenstrom. *Questions and Responses.* Lund Studies in English: Number 68. Lund : CWK Gleerup, 1984.

[95] David R. Traum. *A Computational Theory of Grounding in Natural Language Conversation.* PhD thesis, Department of Computer Science, University of Rochester, 1994. Also available as TR 545, Department of Computer Science, University of Rochester.

[96] David R. Traum. A reactive-deliberative model of dialogue agency. In J. P. Müller, M. J. Wooldridge, and N. R. Jennings, editors, *Intelligent Agents III — Proceedings of the Third International Workshop on Agent Theories, Architectures, and Languages (ATAL-96)*, Lecture Notes in Artificial Intelligence. Springer-Verlag, Heidelberg, 1996.

[97] David R. Traum and James F. Allen. Causative forces in multi-agent planning. In Y. Demazeau and J. P. Muller, editors, *Decentralized A.I. 2*, pages 89–105. Elsevier Science Publishers B. V., 1991.

[98] David R. Traum and James F. Allen. A speech acts approach to grounding in conversation. In *Proceedings 2nd International Conference on Spoken Language Processing (ICSLP-92)*, pages 137–40, October 1992.

[99] David R. Traum and James F. Allen. Discourse obligations in dialogue processing. In *Proceedings of the 32^{th} Annual Meeting of the Association for Computational Linguistics*, pages 1–8, 1994.

[100] David R. Traum and James F. Allen. Towards a formal theory of repair in plan execution and plan recognition. In *Proceedings of the 13th Workshop of the UK Planning and Scheduling Special Interest Group*, September 1994.

[101] David R. Traum and Elizabeth A. Hinkelman. Conversation acts in task-oriented spoken dialogue. *Computational Intelligence*, 8(3):575–599, 1992. Special Issue on Non-literal language.

[102] W. Van Der Hoek and B. Van Linder and J-J.CH. Meyer. An integrated Modal Appraoch to Rational Agents. In this volume.

[103] Daniel Vanderveken. On the unification of speech act theory and formal semantics. In P. R. Cohen, J. Morgan, and M. E. Pollack, editors, *Intentions in Communication.* MIT Press, 1990.

[104] Daniel Vanderveken. *Meaning and Speech Acts.* Cambridge University Press, 1990-1991.

[105] G. H. von Wright. Deontic logic. *Mind*, 60:1–15, 1951.

[106] Marilyn A. Walker. *Informational Redundancy and Resource Bounds in Dialogue*. PhD thesis, University of Pennsylvania, 1993.

[107] E. Werner and Y. Demazeau, editors. *Decentralized A.I. 3*. Elsevier Science Publishers B. V., 1992.

[108] Terry Winograd and Fernando Flores. *Understanding Computers and Cognition*. Addison-Wesley Publishing Company, Inc., 1986.