Cooperation, dialogue and ethics

JENS ALLWOOD

Department of Linguistics, Göteborg University, Box 200, SE 405 30, Göteborg, Sweden.

DAVID TRAUM†

UMIACS, University of Maryland, College Park, MD 20742, USA.

KRISTIINA JOKINEN

Centre for Evolutionary Language Engineering, S.AI.L Trust VZW, Sint Krispijnstraat 7, Ieper 8900, Belgium.

This paper describes some of the basic cooperative mechanisms of dialogue. Ideal cooperation is seen as consisting of four features (cognitive consideration, joint purpose, ethical consideration and trust), which can also to some extent be seen as requirements building on each other. Weaker concepts such as "coordination" and "collaboration" have only some of these features or have them to lesser degrees. We point out the central role of ethics and trust in cooperation, and contrast the result with popular AI accounts of collaboration. Dialogue is also seen as associated with social activities, in which certain obligations and rights are connected with particular roles. Dialogue is seen to progress through the written, vocal or gestural contributions made by participants. Each of the contributions has associated with it both expressive and evocative functions, as well as specific obligations for participants. These functions are dependent on the surface form of a contribution, the activity and the local context, for their interpretation. We illustrate the perspective by analysing dialogue extracts from three different activity types (a travel dialogue, a quarrel and a dialogue with a computer system). Finally, we consider what kind of information is shared in dialogue, and the ways in which dialogue participants manifest this sharing to each other through linguistic and other communicative behaviour. The paper concludes with a comparison to other accounts of dialogue and prospects for integration of these ideas within dialogue systems.

© 2000 Academic Press

1. Introduction

It is uncontroversial to view dialogue as a joint activity requiring coordinated action from multiple participants. What is more challenging is to explicate the nature of the coordinated interaction in a productive manner, both for studying and understanding the nature of dialogue, and to build artificial agents that can engage in natural dialogue behaviour. In examining coordinated behaviour (for simplicity, between two actors, although larger coordinated groups are possible), one can draw two ends of a spectrum—one extreme has the two actors as completely subordinate to a joint controller,

†Now affiliated with the Institute for Creative Technologies, University of Southern California, 13274 Fiji Way, Marina del Rey, CA 90292, USA. *email: traum@ict.usc.edu*

e.g. someone coordinating left and right hands to lift a heavy object. On the other extreme, one can see the two actors as completely independent as to their motivations and actions, the coordination resulting merely from (inter)acting in the same environment, as *emergent behaviour* of robots, or perhaps animals that only look coordinated from the outside. Human dialogue does not fall on either of these extreme ends of the spectrum of coordinated behaviour, since it seems clear that people are independent entities, capable of independently motivated action, and it also seems that they (often) take each other's actions, motivations and other mental attitudes into consideration when acting, particularly for tasks such as dialogue.

Many authors (e.g. Allwood, 1976, 1997; Searle, 1990; Grosz & Sidner, 1990; Cohen & Levesque, 1991; Roschelle & Teasley, 1995) have claimed that this "taking into consideration" involves more than just reasoning about the internal state or likely future actions of the other agent (e.g. as in a chess game), namely actually building some joint socio-cognitive structure that plays a role in influencing the future behaviour of the agents. The specific proposals differ, however, as to what exactly is involved or how best to view these kinds of coordination. While different notions can each have their own role in explaining various aspects of coordination, a central question for the designers of dialogue systems is what kinds of structures are most useful in achieving coherent understanding and participation in dialogue. In this paper, we summarize and explore one such proposal which among other things involves, the notions of "ideal cooperation" and the theory of "activity-based communication analysis" described in a series of papers, see e.g. Allwood (1976, 1995*a*), and show how its components are essential ingredients in satisfactorily explaining both fully cooperative and competitive types of dialogue.

The paper is organized as follows: the first part of the paper introduces the concepts and the theoretical basis of the approach. We start by discussing the concept of *cooperation* (Section 2) and its relation to concepts like *coordination* and *collaboration* (Section 3). We continue by comparing the analysis put forward with other approaches to cooperation and collaboration (Section 4). We then turn to dialogue as a part of social activity (Section 5), relating this to the "Activity based Communication Analysis" approach (Section 6). We also discuss some cooperative features of dialogue and dialogue progression, in terms of the relationship between evocative and expressive functions of communicative contributions and the obligations they give rise to for sender and receiver (Section 7). In Sections 6 and 7 we also analyse two dialogues, one in an institutional cooperative setting and the other in a situation of conflict, to exemplify the co-operative nature of normal dialogue as part of a social activity, concerning features such as repetition, ellipsis and adjacency pairs. We summarize the main claims in Section 8, and continue with a discussion of the implications of our approach for computational agents. We analyse one of the TRIPS dialogues to demonstrate how the analysis can also be applied to a situation of artificial human–computer interaction (Section 9), and then give an overview of the application of the activity-based communication analysis in computer dialogue systems (Section 10). This is followed by a discussion of dialogue as involving the sharing of information and purposes, in comparison with the "common ground" approach (Section 11). Finally, we summarize some of the traits of the analysis of cooperation presented, and compare the present treatment to other approaches focusing on the importance of ethics for cooperation (Section 12). The paper concludes with future prospects and research.

Part I: cooperation

2. Cooperation—a matter of degree

The nature of cooperation has been debated at least since Peter Kropotkin's (1902) treatise on mutual aid. In the field of language and communication, an important contribution was made in Grice (1975), where an explication of communication as cooperative was made through the proposal of four maxims of rational communication which Grice, inspired by Kant, called the maxims of quality, relation, quantity and manner. Our point of departure is the definition of cooperation given in Allwood (1976). In this definition, cooperation is claimed to be a feature of the interaction between "motivated rational agents" engaged in a joint activity. The notion of "motivated rational agenthood" is characterized by the following seven principles.

1. Typical human beings are motivated rational agents.

Agenthood

2. The intentionally controllable behaviour of an agent is intentional and purposeful.
3. The actions of an agent are not performed against his own will.

Motivation

4. The actions of an agent are not performed against his own will.
5. Normal agents do not act so as to decrease their pleasure or increase their pain.

Rationality

6. The actions of a rational agents are selected so as to provide the most adequate and efficient way of achieving the purpose for which they are intended.
7. The actions of a rational agent are performed only if he thinks it is possible to achieve their intended purpose.

Cooperation is also claimed to be a matter of degree, definable in terms of four requirements that would be needed to achieve *ideal cooperation*. Thus, two or more parties interact cooperatively to the extent that they in their actions

(1) take each other into cognitive consideration,
(2) have a joint purpose,
(3) take each other into ethical consideration,
(4) trust each other to act in accordance with (1)–(3).

The analysis of cooperation as a matter of degree is further reinforced by recognizing that over and above variation in the extent to which the above four requirements are met, there are several possible ways of specifying each of the components of ideal cooperation, "cognitive consideration", "joint purpose", "ethical consideration" and "trust". We will now briefly describe each of these components in turn.

2.1. COGNITIVE CONSIDERATION

Taking X into cognitive consideration means roughly "thinking about X". There are a number of ways in which things (and especially agents) can be thought about. Prediction of and reasoning about the actions and reactions of another agent are important parts of cognitive consideration, but it also includes intuitions and other less

rational ways of thinking. Cognitive consideration is necessary, for example, in order to make sure that the information reaches the other person and is shared.

Cognitive consideration can conceptually be further refined in a number of ways. For example, the level of awareness and intentionality at which the reactions and processes which are involved occur can be specified. Cognitive consideration can also be specified as to whether these processes involve the considering agent's perception, emotion, cognitive attitudes or factual reasoning. Further, the object that is cognitively considered can be specified as to whether it is inanimate, animate or capable of agency. When using cognitive consideration in order to analyse what is involved in cooperation, the object will normally either be an entity capable of agency (like another person) or some property, behaviour or object directly connected with such an entity. Thus, we normally do not cooperate with stones, assuming that they are not capable of agency, but with persons.

2.2. JOINT PURPOSE

If we subject the notion of "joint purpose" to further analysis we can relate it to the following dimensions.

1. Degree of mutual contribution to shared purpose: Has the shared purpose been contributed by one, several or all involved parties?
2. Degree of mutual awareness of shared purpose: Have all concerned parties perceived and understood that other parties have the same purpose, i.e. one might well have the same purpose as another person without being aware of this fact?
3. Degree of agreement made about purpose: Have the parties entered into an agreement concerning working toward the purpose?
4. Degree of dependence between purposes: Does the achievement of one party's purpose depend on another party also achieving their purpose?
5. Degree of antagonism involved in the purposes: To what degree do the parties have a purpose to be antagonistic toward each other?

Variation in what we can mean by a joint purpose can occur along all five dimensions. We will consider the cases one by one. First, if A gives B an order to help him and B decides to comply, the two parties share a unilaterally contributed purpose and could be said to engage in one-way cooperation. In cases of this type, it is not uncommon to say that B is being cooperative, i.e. complying with the purpose contributed by A. This use of "cooperative" might even be extended to inanimate objects so that we might say, for example, that a car is being uncooperative, if it will not start.

Secondly, variation might also occur according to how aware we are of sharing a purpose with other agents. Having the same purpose as another person does not mean that we are automatically aware of this fact. However, without awareness of the purposes of another, it will not be possible to be as complete and accurate in the cognitive and ethical consideration of others. Likewise, without confidence that the other is aware of your own purpose it is more difficult (and risky) to trust others to do a good job of cognitively and ethically considering you. Moreover, if the purposes are assumed to be the same, cognitive and ethical consideration of the other party will partly coincide with what the agent has to consider on his or her own behalf.

Thirdly, there is a distinction to be made between being mutually aware of having the same purpose and having agreed to have the same purpose. You and I might both be aware that we both want to terminate a discussion without having agreed to this.

Fourthly, we can distinguish cases where we can both achieve the same purpose independently of the other person (e.g. you and I can independently leave a meeting) from cases where my achieving my purpose is dependent on you achieving the same purpose (e.g. two-way communication is only possible if both parties share this purpose).

Fifthly, purposes can be more or less antagonistic. For example, if A and B share the abstract purpose of wanting to hurt the other person ("I want to hurt you"), even if the abstract formulation is shared, the deictically anchored versions of this purpose will be antagonistic and different.

It is a matter of study to determine exactly what degree of each of these dimensions is required for there to be a sufficient "joint purpose" for ideal cooperation. At a minimum, both parties should be aware of and should have adopted the purpose. The purpose may, however, be contributed by only one of the parties. Further, there may be no mutual awareness of sharing the purpose, and there may be no explicit mutual agreement about the purpose. However, antagonistic or independent purposes are not sufficient since they do not constitute a joint purpose.

2.3. ETHICAL CONSIDERATION

Ethical consideration implies that the parties act according to the following maxims (cf. Allwood, 1976, 1995a). The formulation in brackets gives a slightly stronger version of each ethical maxim.

(1) They try not to force each other (make it possible for the other party to act freely).
(2) They try not to prevent each other from pursuing their own motives (help the other person to pursue his/her motives). Since the urge to escape pain and to seek pleasure is perhaps the strongest of all human motives, this means that they should try not to hurt each other (make it possible for the other party to seek pleasure).
(3) They try not to prevent each other from exercising rationality successfully (make it possible for the other party to exercise rationality successfully). Since correct information, at least in the long run, is a precondition of successful rational action they should not lie but give each other adequate and correct information.

Ethical consideration involves having ethical goals, making ethical decisions and acting ethically. It can also be seen as connected with general *obligations* to act in this manner. Even if the desire to behave in this manner changes, the obligations may persist, sometimes bringing with them social sanctions for failure to comply, cf. Allwood (1994).

Let us now examine the properties of ethical consideration in some more detail. In order to make it possible for others to be agents, we should refrain from imposing on them too much, instead leaving them the freedom to act according to their own will and intention. This is one of the sources of politeness. It is usually more polite to make an indirect request such as *could you pass me the salt* than a direct request like *pass me the salt* since the former in its formulation, even if not actual usage, gives the interlocutor somewhat greater degrees of freedom than the latter.

Making it possible for others to be agents is also what allows us to claim that "brainwashing" and many kinds of propaganda are unethical. They are unethical since they remove the recipient's possibilities to exercise his/her own critical judgement and in this way reduce the recipient's possibilities of being an agent.

Secondly, ethical consideration implies that we should make it possible for others to pursue their own motives. A very fundamental type of motive is related to pain and pleasure. Another is related to power, cf. Allwood (1980). People generally want to escape pain and seek pleasure. Thus, in our communicating, we should not unnecessarily hurt people but, if possible, rather give them joy. Here we have another source of politeness strategies. It is usually more polite to compliment people than to insult them. We also see one of the reasons why so much of communication has a consensus orientation. We do not wish to hurt others, especially if they have more power than we do. The implementation of ethics is, in this case and in others, strongly aided by the reliance on the golden rule in its positive and negative formulations, "do unto others what you would have them do unto you", and "do not do unto others what you would not have them do unto you". In other words, if you hurt others they might hurt you later on, whereas if you treat them nicely, they might treat you nicely.

Thirdly, ethical consideration in the analysis proposed here, involves making it possible for others to be rational, i.e. they should be able to act adequately and competently. In order to successfully act adequately and competently, we must have correct information, otherwise we cannot judge if the appropriate preconditions for a certain course of action are present. This directly implies that we should not lie or mislead. If we do this, the other person's possibilities of obtaining the desired outcomes through the exercise of his/her rationality are radically diminished. This does not mean that rational action has to be based on correct information. Rational action can be based on both correct and incorrect information. It only means that the likelihood of successfully achieving one's goal, in the long run, is greater if action is based on correct rather than incorrect information.

2.4. TRUST

If the parties trust each other, this means that they believe that the other communicators are cognitively and ethically considering them as well in trying to achieve common understanding or other joint purposes. Just like the concepts of "cognitive consideration" and "joint purpose" discussed earlier, the concept of "trust" can be specified in several ways. Two of these are the scope of trust and the motives for trust.

Concerning the scope of trust we might distinguish partial trust from more holistic trust. A can trust B as a carpenter without trusting him more holistically as a human being. It is this latter sense we have in mind in the definition of "ideal cooperation", even though the partial sense will also be involved in many cases.

The concept of trust is also dependent on what motives lie behind the expected behaviour of the trustee. Is he/she motivated by fear of sanctions, by self-interest or by a wish to behave cooperatively for ethical reasons? In the case of ideal cooperation it is this latter motive which is the focus, even though in cases of less ideal cooperation the other motives might well be involved.

Thus, it might be very reasonable to trust (in the partial sense) even an enemy or neutral party to act in a certain way, to the degree that it is in that party's own interest (assuming the party is rational). However, such trust would most rationally dissolve in the case of a change in that party's interests. One might also trust in some agent behaving in a socially acceptable or ethically considerate way, even against that agent's own desire, given a fear of sanctions or a sense of moral responsibility. Trust in the more holistic ethically ideal sense, would be involved if we trust another person to further our goals out of a sense that the person is truly cooperative, even in the absence of self-interest or fear of sanctions.

2.5. RELATIONSHIPS BETWEEN THE FOUR COMPONENTS OF IDEAL COOPERATION

The four requirements of ideal cooperation are related to each other in the following way. Cognitive consideration is basic and requires no other requirement to be met. Similarly, the parties can have the same purpose without being aware of this, i.e. without cognitively considering this aspect of the other person's attitudes and behaviour with each other. Ethical consideration, however, requires cognitive consideration but adds attitudes and behaviour not given by cognitive consideration alone. Similarly, trust requires a certain amount of cognitive consideration but adds expectations not necessarily given by cognitive consideration alone (or for that matter ethical consideration and a joint purpose).

We may now also consider the question of whether the conditions we have associated with trust could be achieved (e.g. by cognitive consideration alone) without adding an explicit requirement of trust on ideal cooperation. To discuss this let us again consider the relationship between trust and the first three levels of cooperation (cognitive consideration, joint purpose and ethical consideration). "Trust" requires not only that the parties themselves meet conditions (1)–(3) but also that they expect and rely on other parties to do the same. This requirement, for example, excludes a situation where A and B by cognitively considering each other arrive at the conclusion that the other party is not cognitively considering them. This situation would, however, be allowed without the requirement of "trust". Similarly, it excludes a situation in which A and B are both working for the same purpose while believing that the other party is not, or a situation where both parties are ethically considering each other while believing that the other party is not doing so. Thus, if these possibilities are to be excluded as incompatible with ideal cooperation and/or ideal communication, a requirement of trust is needed.

On the other hand, it is also true that A through cognitive consideration *could* arrive at the conclusion that B considers him/her cognitively or ethically or shares the same purpose. But this conclusion is not necessary. It only follows given a number of extra assumptions about the nature of B's actions (e.g. that he is ethical or has the same purpose as A). If A is able to make these assumptions, he/she probably also will trust B and meet condition (4) of ideal cooperation.

In the next section, we will examine some of the degrees of cooperation, in relation to the terms "coordination" and "collaboration".

3. Coordination, collaboration and cooperation

We would now like to use what has been said to briefly discuss the relation between the terms *cooperation*, *collaboration* and *coordination*. First, it should be said that the use of the three terms overlaps in ordinary language, especially between the two terms *cooperation* and *collaboration*. This means that any attempt to distinguish them must be partly stipulative.

It seems clear that *coordination* involves the least requirements on mutual relations between the interacting parties. On the scale proposed above, the lowest degree of cooperation (1) taking each other into cognitive consideration, could therefore be called *coordination*. *Cooperation* in the sense of *coordination* is even presupposed for most types of conflict since the parties need to cognitively consider each other to hurt each other. Turning to *collaboration*, this term can also be said to pose fewer requirements on the mutual relations between the interacting parties than *cooperation*. People can arguably be "collaborators" without trusting each other, and, without being ethically committed to each other. This can be seen in the use of the term *collaborator* to denote people who are working with an oppressive regime e.g. *Nazi collaborators*, or the use of the term by a boss to designate his employees or business contacts. We therefore suggest that *collaboration* could be used for interaction which combines criteria (1) and (2), i.e. cognitive consideration and having a joint purpose or mutual goal. We can see this also in some of the proposed explicit definitions of collaboration in the AI literature, for example, Grosz and Sidner (1990) and Cohen and Levesque (1991). Some, e.g. Clark and Wilkes-Gibbs (1986) and Grosz and Sidner (1990), however, use the term *collaboration* rather than *cooperation* to describe the driving force behind dialogue. While some notion of collaboration or *teamwork* (Cohen & Levesque, 1991) could be sufficient to account for much of the interaction in task-oriented dialogues focused on a joint purpose, a stronger notion, including also ethical commitment and trust, is needed to account for the more general level of coordination and coherence in dialogue, which to some extent can be present even in situations of conflict.

The term *cooperation* more naturally than *collaboration* seems to encompass states where the parties are ethically committed to each other, trust each other and have a more or less egalitarian (in terms of social power) relationship to each other. If this suggestion can be accepted, the term *cooperation* would function similarly to adjectives like *long*, *old* and *big* which cover both a dimension as a whole and a specific part of that dimension. You can ask of a baby how old he/she is and get the answer—3 days old. But you can also say of a person who has passed middle age that he/she is old. In the same vein, *cooperation* would cover *coordination* and *collaboration* but also be the preferred term for types of interaction involving more demanding aspects of the scale, such as ethical commitments, trust and egalitarian social relationships.

4. Other approaches to cooperation and collaboration

At this point it is worth comparing the account of ideal cooperation presented here to other accounts of coordinated activity in the AI and dialogue literatures. There we find that "cognitive consideration" has received by far the most attention, with several well-developed formalisms for the actions and reasoning of agents about their own actions and the actions of others. In particular, plan recognition techniques have been

used to predict agent's goals and plans from observed actions and likely future actions from goals and plans. In fact, Allen and Perrault (1980), showed how this could be used in combination with a variant of ethical maxim two, above, to understand indirect speech acts, and provide cooperative rather than literal replies. The belief ascription literature, (e.g. Ballim and Wilks, 1991) has also focused on how to derive new beliefs from the beliefs of others.

More recently, there have been several attempts within this literature to explicate a notion like joint purpose. Some, e.g. Searle (1990), have seen such a notion as primitive, while others have tried to formulate it in terms of other primitives, with reference also to cognitive consideration. The concept of joint purpose, together with cognitive consideration, has been used to develop a notion of collaboration. We will briefly consider two very influential and well-developed accounts here, joint intentions, as developed by Cohen and Levesque (1990a, b, 1991), and SharedPlans, as developed by Grosz and Sidner (1990), and later developed by Grosz and Kraus (1993, 1996, 1998). Numerous other variants have also been introduced.

4.1. COHEN AND LEVESQUE'S TEAMWORK

Cohen and Levesque start from a dynamic logic model theory of individual agent's actions, beliefs, goals and commitments and intentions (formalized as persistent goals of various sorts), as developed in Cohen and Levesque (1990a, b). The definitions of an individual persistent goal and an individual intention are generalized to those of a group. The basic notion of a joint purpose is a joint persistent goal. A group of agents is defined as having a joint persistent goal to achieve p (relative to an escape clause q, specifying conditions in which the team will drop the goal), if and only if the following conditions hold.

1. They mutually believe that p is currently false (since it is an achievement goal).
2. They mutually know that they all want p to eventually be true.
3. It is true and mutual knowledge that until they mutually believe that either p is true or can never be true or q is false, they will continue to mutually believe that they each have a weak achievement goal relative to q, with respect to the rest of the team.

An agent is defined as having a weak achievement goal if either of the following conditions holds.

(1) The agent has a normal goal.
(2) In the cases in which the agent believes p is true, will never be true or q is false, the agent has the goal that this status be mutually believed by the whole team.

The notion of teamwork is then formalized as a joint intention, which is defined as a joint persistent goal of having done the action, mutually believing throughout that the team is doing it. Collaboration is seen as teamwork: a joint action by a group of agents who share a joint intention to perform a (composite) action together. The team's commitments are related to the individual team members' commitments, but they are separate entities in the world. Thus, if one agent privately discovers that the (joint) goal is already achieved or impossible to achieve, the agent has a goal to make this fact mutually believed by communicating the changed status of the goal to the partner.

This notion of teamwork explicitly brings in cognitive consideration in a number of ways, especially in the form of mutual beliefs, which require beliefs about the beliefs of others. There are also some facets of ethical consideration, in terms of giving correct information (via the goals of achieving mutual belief). The other maxims of ethical behaviour are also partially present through the definitions of commitment, in which, following Bratman (1987), agents will not act in such a way as to make it impossible to achieve their intentions. So, insofar as the goals and freedom of action of an agent relate to performing jointly intended actions, other team members will allow the agent to proceed. However, there are two things missing with respect to actual ethical considera- tion. First, there is no sensitivity here to the actual goals of the agent, only to the assumed team goals. Secondly, this sensitivity is manifested only in the form of (persistent) goals, and not in any form of external commitment. Consider the case in which an agent insincerely pretends to joint a team. According to the definitions here, no such team (or joint) intention holds, since the requisite mutual beliefs and goals did not hold. There is no representation here of obligations that hold as a result of engaging in a contract, regardless of personal goals.

Likewise, trust is present here in only a limited fashion. Insofar as an agent can know that it is a member of a team, it can use cognitive consideration to realize the goals of the other agents, and can trust those agents to also have cognitive consideration and ethical consideration with regard to providing information (to achieve mutual belief). The problem, again, is knowing whether such a team is actually formed, and no guidelines are laid out for which an agent can know that this condition holds.

The teamwork model concentrates on the maintenance of a joint goal and does not deal with cooperative planning or the negotiation process through which a shared commitment is reached. The requirement for communication also seems to account for only limited kinds of dynamic cooperative communication: the expected effect of the individual's observation that the joint goal has already been or cannot be achieved. Furthermore, as Grosz and Kraus (1996) point out, Cohen and Levesque's agents cannot opt out from their commitment without telling the partner so. We believe that this requirement more directly follows from an ethical and cognitive consideration of the partners than from an agent's steadfast commitment to a joint intention. A commitment to act as part of a team, when no longer valid, does not by itself commit the agent to any particular action toward the partners.

4.2. GROSZ AND KRAUS' SHAREDPLAN

Like the Cohen and Levesque formalism, the SharedPlan formalism (Grosz & Sidner, 1990; Grosz & Kraus, 1996, 1999) is formulated in terms of mutual beliefs, as well as individual, beliefs and intentions. The SharedPlan formulation concentrates on the conditions under which one might rationally be said to have an achievable plan, or for a group of agents, an achievable shared plan. Strong rationality constraints are put on the agents in order for them to be said to have particular attitudes such as intentions or individual or shared plans. These constraints include things like knowing how to do an action or at least how to determine how to do it. The original SharedPlan framework said that a group of agents had to mutually believe that each action was executable, would contribute to the plan, and that the agent of the action intended the action (and

intended it to contribute to the plan), as well as that the agents actually had those intentions. Interestingly, one way of furthering the attitudes needed to establish the shared plan (Conversational default rule 1) included a condition that the agents were cooperative, without cooperativity being, itself defined.

Later work on SharedPlans (Grosz & Kraus, 1996, 1998) focused on elaborating the many ways that plans could be achieved and the attitudes needed for each of the types of decomposition. One interesting extension is the ability to contract out some actions to other agents or groups. While the SharedPlan formalism does go further in terms of specifying the behaviour of planning and collaborating agents, it suffers from much the same problems as the Cohen and Levesque formalism, with respect to the ideas of Ideal Cooperation. Obligations are not explicitly represented in the framework, which does specify certain conditions under which agents will adopt intentions and act in an ethical manner. However, this is all subordinated to types of cognitive consideration, i.e. whether the behaviour will lead to furtherance of the agent's actual goals. There is no way to reason within the formalism between what one desires and what one is obliged to do. Likewise, even if a certain amount of trust is available, coming from the mutual beliefs that agents have certain attitudes and act rationally, there really is no way to model trusting an agent to behave ethically (nothing prevents the agent from dropping the commitment to engage in the shared plan), or more general trust. However, there is nothing, in principle, which prevents adding these other aspects of ideal cooperation to the SharedPlan formalism.

Like the account of "Ideal Cooperation" presented in Section 2, the AI accounts considered in this section were, at least in part, formulated to shed light not just on multi-agent interaction in general, but also on the specific issue of agents communicating via dialogue. It is in such realms that the components of ideal cooperation are truly needed to explain observable behaviour. In the next part, we will therefore show how ideal cooperation is an important component (along with social activities, and particular circumstances) in accounting for dialogue behaviour.

Part II: dialogue and social activity

One of the main claims of the current perspective is that dialogue, while an instance of cooperation, is not usually a completely independent and distinct activity but is rather a result of cooperative behaviour within some activity to which it has an instrumental relationship. The dialogue aspect of an activity can be more or less self-sufficient, for example, casual conversation vs. giving instructions for a haircut. Dialogue behaviour is not seen as solely the result of individual action, but a composite of multiple factors including: individual motivations, general cooperativity, as outlined in the previous part of the paper and activity specific features, including rights and obligations pertaining to roles. In this part of the paper, we present an outline of this framework which has been given the name "Activity-based Communication Analysis."

5. Cooperation and communication

Normal communication is cooperative since it requires at least the following two of the conditions to be met.

(1) the parties in their actions take each other into cognitive consideration in order to achieve

(2) the joint purpose of understanding.

Taking the other person into cognitive consideration is necessary in order to make sure that the information reaches him/her and is shared. Communication can involve "one-way sharing", as in radio broadcasts. It can also be two-way, where both parties cognitively consider each other. However, two-way consideration alone would not be sufficient if both parties made each other understand totally different things. This would be what might be called "double one-way communication". The goals of both parties have to be related to each other. For this, the parties must take each other's purpose into account and manifest this in their responses. Two-way communication is thus cooperative in the sense of involving two-way cognitive consideration and the goal of shared understanding of a minimum of at least two related contributions. Two-way communication occurs even if the second responding contribution is negative since a negative reply also shows cognitive consideration of the other part and a sharing of the goal of joint understanding and, through the very fact that a response was made, provides a starting point for ethical consideration and trust.

As already mentioned, cognitive consideration will usually be involved even in conflict. This is required in order to effectively hurt the other party. Without shared understanding and a relevant response, we would, however, have one-way conflictual communication. In order to have conflictual two-way communication, as shown for example, in the analysis of a quarrel (in Sections 6 and 7, below), joint understanding and a relevant response must also be involved. In this type of conflict, the purpose seems to be that both parties should understand that (and how) the other party has been hurt. However, even conflictual dialogue can be partly cooperative in the sense of involving ethical consideration and trust in some respects or for parts of the dialogue.

6. Social activity and activity-based communication analysis

Most communication takes place as a means for pursuing some joint activity like a commercial transaction, a negotiation or a lecture. This activity involves a joint purpose of its own for which communication serves as a means, i.e. joint understanding must be sought in the service of the purpose of the activity. This means that communication is not normally seen as a separate social activity but as a part of a more encompassing activity.

Most social activities also involve some degree of cooperation in the sense explicated above, i.e. the participants in the activity have a joint purpose, they cognitively and ethically consider each other and often trust each other as well. The participants mostly participate by occupying the more or less stereotypical roles which are involved in the activity. To use the examples given above, they may act as buyers or sellers, negotiators, lecturers or audience. Each role carries with it certain requirements on competence and certain rights and obligations with regard to communication and other kinds of action. The requirements are usually conventionalized but are still functional ways of making sure that a given activity can be pursued cooperatively. The obligations going with the specific roles of the activity will, in general, be subject both to constraints given by the

general ethical obligations, discussed above, as well as to constraints generated by the intersection of the ethical constraints with basic types of communicative actions to be discussed below, in Section 7. However, the structure and function of communication in a given social activity is not shaped merely by the constraints given by the activity and the environment. Besides being dependent on the activity, structure and function are dependent on what the individuals who happen to occupy the roles of the activity actually do, which in its turn, to a large extent, will be dependent on the physical, biological, psychological as well as social and cultural backgrounds of these individuals. Thus, actual role behaviour will be a result of the combined influence of such factors as interest, personality and socio-cultural identity (see also Allwood, 1984, 1995a, c). The analysis of dialogue and communication as parts of social activities follows the parameters for "activity-based communication analysis" described in e.g. Allwood (1980, 1984, 1995a) and Ahlsén (1995). According to this analysis, an activity can thus be classified as to the following four parameters: the *purpose* for engaging in the activity, the *roles* involved, the *artefacts* which are prominent in the activity, and the *physical and social environment* in which this activity takes place. We will exemplify these parameters in the dialogue examples discussed in Section 6.1.

The analysis proposed thus claims that much communication takes place under multiple requirements of cooperation. One type of requirement is the general ethical requirement imposed by being a cooperative sender or receiver. Another type of requirement is imposed by occupying a particular role in an activity, the obligations of which shape the ensuing dialogical cooperation between the parties to a considerable extent. A third type of requirement is imposed by the nature of the communicative acts which are used in the interaction. In order to discuss this in somewhat more detail, we must first briefly examine some of the units, functions and mechanisms of dialogue. The framework presented basically follows Allwood (1978, 1995a). For a critical discussion and motivation of the framework, see Allwood (1976, 1977a).

6.1. EXAMPLES

In order to make the concepts which have just been introduced more concrete and show how they can be applied to more or less cooperative settings, we will now illustrate the approach outlined above by analysing parts of two recorded and transcribed dialogues which occupy different positions along the continuum of cooperation. The first one is the beginning of a dialogue from a travel agency. The second one is part of a quarrel between two sisters. We start by giving an analysis of the nature of the two social activities. First, the travel agency dialogue and then the quarrel will be analysed. In Section 7, we will analyse the same dialogues in terms of expressive and evocative functions and speaker obligations. The dialogues are both taken from the Göteborg Spoken Language Corpus where about 25 different types of social activity have been recorded and transcribed (see http://www.ling.gu.se/SLSA/SLcorpus.html).

In the transcriptions, numbered square brackets (e.g. [1]1) indicate overlap. Slashes /, //, /// indicate pauses of different lengths. Angular brackets ⟨ ⟩ enclose comments about a part of the utterance indicated by the brackets. Angular brackets are also used to indicate unconventional sounds which are part of an utterance (as exemplified in the dialogue fragment in Section 6.1.2). Spellings of words follow the MSO standard.

6.1.1 The travel agency dialogue—flight to Paris

Swedish		English	
C1:	hup	C1:	hup
A1:	[1a] 1	A1:	[1 yes]1
C2:	[1örm]1 //flyg ti Paris	C2:	[1uhm]1 // flights to Paris
A2:	mm/ska [2 du ha] 2 en retur biljett	A2:	mm/do [2 you want] 2 a return ticket
C3:	[2 ö:]2	C3:	[2eh]2
C4:	va sa du	C4:	what did you say
A3:	ska du ha en tur å retur	A3:	Do you want a return
C5:	ja/ö	C5:	yes/eh
A4:	vilken månad ska du åka	A4:	which month are you going

1. *Purpose.* The purpose of the activity is information about travelling (e.g. for a given destination, information concerning means of transportation, times, prices and lodging) and to enable the ordering and selling of trips.

2. *Roles.* The activity typically involves two roles, travel agent and customer. There can be several customers and occasionally several travel agents. Each role can be characterized in terms of obligations, rights, and competence requirements. The travel agent's obligations which are based on meeting certain competence requirements required for the job include giving information about types of transportation to different destinations, prices, times and hotels. It also includes making reservations for journeys, rental cars and hotels as well as selling tickets. A travel agency dialogue is asymmetrical since one party provides services to another. This means that the obligations of the agent are the rights of the customer who him/herself has few obligations over and above making use of the agent's services and observing general obligations connected with ethics and politeness.

3. *Artefacts.* Travel agency dialogues are usually supported by telephones and computers. Both of these tools enable and restrict various features of the dialogue. We can, for example, predict that there will be utterances reporting results from using the computer or dealing with periods of waiting for such results.

4. *Physical and social environment.* Other factors that might influence the dialogue are conditions of lighting, sound level and seating arrangements.

On the basis of the analysis of purpose and roles of an activity, it is usually possible to make fairly good predictions about what type of communicative acts will occur in the activity and about what roles they are associated with. Thus, in a travel dialogue it is reasonable to expect the following acts to occur. For the customer—requests for information about destinations, means of transportation, prices, times, etc. For the travel agent—answers giving the required information, and perhaps suggestions of options. Both parties will engage in greetings and farewells and both parties will make requests for specifications and clarifications and reply to such requests. Both parties will also engage in various types of communication management [see Allwood, Nivre & Ahlsén

(1990) Allwood (1995*a*, *b*) as well as below, in Section 7] concerning feedback, turn taking or reports.

6.1.2. A quarrel between two sisters

Swedish		English	
D1:	men herregud < clicking sound > /	D1:	but good Lord ⟨clicking sound⟩
S1:	kan du låta bli min freestyle eller	S1:	could you leave my walkman alone
D2:	nä	D2:	no
S2:	⟨a men släng inte ner den nu⟩ ⟨yelling⟩	S2:	⟨yes but don't throw it down now⟩ ⟨yelling⟩
D3:	⟨ingen fara⟩/den e ändå så gammal ⟨very slowly⟩	D3:	⟨no danger⟩/ it is so old anyway ⟨very slowly⟩
S3:	vadå gammal två dar	S3:	old? two days
D4:	aa//	D4:	yeah//
S4:	⟨sings⟩	S4:	⟨sings⟩
D5:	sluta du e ÄCKLI	D5:	stop you are DISGUSTING

We can analyse the quarrel between the two sisters in a manner similar to the way in which we analyse the travel bureau dialogue.

1. *Purpose.* There is no conventional shared purpose over and above informal conversation. Two sisters are having an informal conversation and are irritating each other in various ways. The first transcribed disagreement occurs when one of the sisters (D) picks up a Walkman belonging to the other one (S). This turns out not to be acceptable to (S) and a quarrel starts.

2. *Roles.* The roles are assumed to be tailored to the purpose, i.e. informal conversation. So the roles of the two participants are those of participants in informal conversation. Underlying these roles, however, are the roles of being the elder and younger sister. The effects of these roles are combined in the conversation, so that we get the roles of two sisters in informal conversation. Since this is not a conventionalized activity with a clear purpose, nothing substantial can really be said about the rights and obligations corresponding to the two roles. This means that general ethical obligations will have a stronger influence. The influence is, however, mitigated by the fact that sisters have a strong mutual bond which, at least, in Swedish culture makes it possible for them to neglect certain ethical and politeness considerations and thus have access to a larger interactive communicative repertoire than would be the case for people who are less well acquainted.

3. *Artefacts.* Artefacts do not have a standardized conventional role in the activity. They do, however, play a more accidental role as, for example, being objects of attention, see below the role of the Walkman as an object of the quarrel.

4. *Physical and social environment.* The physical and social environment is the home of the sisters. This probably means that there is a fairly relaxed atmosphere.

6.2. DISCUSSION

It follows form what has been said that far fewer predictions can be made about the type of communicative acts that will occur in the quarrel than in the case of the travel

bureau. In the quarrel, we are not dealing with a well-defined and conventionalized activity. Perhaps all that can be said is that the range of possible communicative acts will be much greater than in a functionally well-defined activity and that these communicative acts, will probably be informally delivered, often be elliptical and contain a rich amount of presuppositions and implicatures.

A comparison of the two types of dialogue, thus, shows that in the case of the travel agency, functional and conventional requirements will restrict the kind of communication to a greater extent than in the case of the quarrel between the two sisters. We might say that conventional activity requirements play a much greater role in the travel bureau dialogue than in the case of the quarrel, where there is no conventional activity that is being pursued. In the quarrel, instead, general human requirements concerning ethics and requirements generated by the speech acts will tend to predominate, mitigated by the enablements and requirements given by the role of being adolescent sisters.

For computer dialogue systems, this implies that activities with clear functional and conventional requirements will be easier to design systems for, since many of the requirements can be "hardwired" into the system. However, such an approach loses generality, and cannot easily cope with a change of activity or even with trying to do unexpected things within a particular activity. A more promising approach might be to maintain the requirements as special cases of a more general conversational competence, so that a system could more flexibly shift activities or reason about the requirements, when necessary.

7. Some features of dialogue and dialogue progression

7.1. CONTRIBUTIONS, EXPRESSIVE AND EVOCATIVE FUNCTIONS

Following Grice (1975), Allwood *et al.* (1990) and Allwood (1995*a*), the basic units of dialogue are gestural or vocal *contributions*† from the participants. The term *contribution* is used instead of *utterance* in order to cover also gestural and written input to communication. Verbal contributions can consist of single morphemes or be several sentences long. The term *turn* is used to refer to the right to contribute, rather than to the contribution produced during that turn. One may make a contribution without having a turn and one may have the turn without using it for an active contribution, as demonstrated in the example below, in which B's first contribution involves giving positive feedback without having the turn (square brackets indicate overlap) and his second contribution involves being silent and doing nothing while having the turn.

A: look ice cream [would] you like an ice cream
B1:   [yeah]
B2: (silence and no action)

---

†The term *contribution* has been used in various ways. Clark and Schaefer (1989), use the term in a more restricted sense to refer to what they call "grounded" contributions. They use the term *presentations* for single-agent contributions that may or may not have been "grounded".

In accordance with Allwood (1976, 1978, 1995*a*), each contribution is viewed as having both an *expressive* and an *evocative* function. The *expressive* function lets the sender express beliefs and other cognitive attitudes and emotions. What is "expressed" is made up of a combination of reactions to the preceding contribution(s) and novel initiatives. The *evocative* function is the reaction the sender intends to call forth in the hearer. Thus, the evocative function of a statement normally is to evoke a belief in the hearer, the evocative function of a question is to evoke an answer, and the evocative function of a request is to evoke a desired action. For a discussion of the relations between these functions and Bühler's (1934) symptom, symbol and signal functions as well as the Austin (1962) locutionary, illocutionary and perlocutionary functions see Allwood (1976, 1977*a*, 1978). The notion of evocative function is also similar to the notion of "intended perlocutionary function" of Sadek (1991).

Each contribution to a dialogue is associated with the following default evocative functions, cf. Allwood (1987, 1995*a*). A contribution is intended to make the receiver.

(1) Continue (C).
(2) Perceive (P).
(3) Understand (U).
(4) React in accordance with main evocative function (R).

The receiver now has to evaluate whether he or she can/wants to continue, perceive, understand and go along with the evocative intention of the preceding utterance. The result of the evaluation will be an important part of the expressive function of the response to this utterance and can be given in explicit or implicit form (see below). Using these concepts, we now turn to an analysis of the cooperative use of the expressive and evocative aspects of contributions. We can provide a more detailed analysis of the cooperative goal of communication into four subgoals, related to the four evocative/expressive functions, one of which is the joint understanding we have already discussed.

(1) Continued interaction until both parties agree to halt.
(2) Joint perception and awareness.
(3) Joint understanding.
(4) Cooperative achievement of evocative intentions.

7.2. OBLIGATIONS

If the four subgoals mentioned above are to be cooperatively pursued, whether it be in the service of some activity or not, they impose certain obligations on both sender and receiver. With regard to both expressive and evocative functions, the sender should take the receiver's perceptual, cognitive and behavioural ability into cognitive and ethical consideration and should not mislead, hurt or unnecessarily restrict the freedom of the receiver. The receiver should reciprocate with an evaluation of whether he/she can hear, understand and carry out the sender's evocative intentions and signal this to the interlocutor. Without reasons to the contrary, the sender and receiver should also trust the other to behave in this manner.

The sender's and receiver's obligations can be summarized as follows (see also Allwood, 1994).

Sender

1. *Sincerity*. The sender should, unless she/he indicates otherwise, have the attitude normally associated with a particular type of communicative act, e.g. statement—belief, request—desire (cf. Allwood, 1976, 1995*a*).

2. *Motivation*. Normally, communicative action, like other actions should be motivated.

3. *Consideration.* If communicative action is to be cooperative and ethical it must take the other person into cognitive and ethical consideration.

Receiver

1. *Evaluation.* The receiver should evaluate the preceding utterance with regard to whether he/she can continue the interaction, perceive and understand and accept its main evocative intention.

2. *Report.* After having evaluated, the receiver should report the result verbally or non-verbally.

3. *Action.* In some activities and roles, a positive evaluation of the ability to carry out the main evocative intention also obligates the listener to carry out the action associated with this intention.

Since perception and understanding mostly function as a means for the sharing of the expressive and evocative functions of each contribution, a cooperative response usually consists in one of the following responses, used separately or in combination.

(1) Overtly signalling the result of the listener"s evaluation through the use of an explicit positive or negative feedback expression, such as a head nod, a head shake or a verbal expression like *m*, *what*, *yes*, *no* or *OK*, after a statement or request.
(2) Direct verbal action, as when a question is answered.
(3) Direct non-verbal action, as when a window is closed after a request to do so.
(4) Implicitly accepting an evocative intention by contributing a response that implies acceptance, as when you accept a stated belief by exploring one of its consequences.

Since the main thrust of a dialogue revolves around evocative intentions which are aimed at achieving more than mere perception and understanding, a cooperative response that signals only perception and understanding usually occurs only in the following circumstances: (1) when a message can be perceived and understood but no commitment is made to its evocative function or (2) a message cannot be perceived or understood. In the first case, often low key feedback expressions like *m* or *well* are used and in the second, we find instead negative feedback expressions such as *pardon* or *what*. These issues are explored further in Allwood, Nivre and Ahlsén (1992).

7.3. EXAMPLES—EXPRESSIVE AND EVOCATIVE FUNCTIONS AND OBLIGATIONS

We now turn again to the travel dialogue and the quarrel to illustrate what expressive and evocative functions and obligations might be involved in dialogues of these types.

Every utterance, unless otherwise coded, either implicitly or explicitly expresses CPU (contact, perception and understanding). CPU are only coded when they are part of the main evocative or expressive function of an utterance. When they are not, another expressed attitude such as acceptance or belief will imply CPU which therefore will not be indicated. Similarly, "acceptance" of information will only be indicated if it is part of what is mainly expressed. If a question is followed by an answer, the answer to the question will be taken to imply acceptance of the task of answering. If a request is followed by the required action, the action will be taken to imply acceptance of carrying out the task and if a statement is followed by a comment which presupposes what is stated to be true, the comment will be taken to imply acceptance of the information expressed by the statement. In all these cases, acceptance will not be coded. A comment is also needed about statements. Statements can be implied or explicit. Answers to questions often contain implicit statements. A *yes* or a *no* answer to a yes/no question, for example, implies an affirmed or negated statement of what is queried in the question. If a statement is implicit, we will code its related expressive functional commitment as an expression of and a commitment to the (propositional) information in the statement. If it is explicit, we will code the statement as an expression of and a commitment to a belief containing that propositional information. In terms of commitments the two will be equivalent, but the former code has the advantage that an informational object can be shared between questions, answers and requests. We ask for and request information rather than beliefs, even though what a conversational interaction will provide are beliefs containing such information.

The next step is to link the expressive and evocative functions with utterance and dialogue act-related obligations, which can now be added as modifications of the role-related obligations we have discussed above. In coding obligations we will, for the speaker, normally indicate commitment to whatever attitude and motive has been expressed by the speaker. Unless it is relevant, we will normally not indicate that the utterance also should be based on cognitive and ethical consideration of the listener. For the listener, the fundamental obligations are never more than evaluation and response (report) but if circumstances are such that a positive evaluation takes place and the role relation is such that the listener, ceteris paribus, is obliged to act in conformity with the speakers main evocative intentions, we will also, in brackets, indicate this action as part of the listener's obligation. In such cases, we will leave out the "respond" obligation which, however, in case the evaluation is negative, will become the main obligation.

In the tables below, "/" means simultaneous functions, ";" means functions occurring sequentially, + means linked obligations (often by virtue of a means—ends relationship). Variables such as X, Y are used as shorthand for the actual information, and utterances are referred to by speaker and number.

7.3.1. The travel agency dialogue—flight to Paris

Travel dialogue—expressive and evocative functions and obligations

Contribution	Expressive and Evocative Function	Obligations Introduced	
C1: hup	expr: presence/desire for contact evoc: CP/start interaction	speaker: commitment to interest in contact listener: evaluate + respond	
A1: [1a]1	expr: CPU Acc (C1) evoc: state request	speaker: commitment to contact listener: evaluate + respond	
C2: [1öm]1// flyg ti Paris:	expr: hesitation; desire for info [X] evoc: give info [X]	speaker: commitment to interest in info (X) listener: evaluate + (give info (X))	
A2: mm/ ska [2 du ha]2 en returbiljett	expr: accept evoc[C2]; desire for info [Y] evoc: give info [Y]	speaker: commitment to need for info (Y) listener: evaluate + (give info (Y))	
C3: [2 ö]2	expr: C hesitation evoc: C	obligations irrelevant	
C4: va sa du	expr: not PU[A2]/ desire for information [Z] evoc: give info [Z]	speaker: commitment to need for info (Z) listener: evaluate + (give info (Z))	
A3: ska du ha en tur & retur	expr: info [Z] evoc: accept info[Z]/give info [Y]	speaker: commitment to info Z + need for info (Y) listener: evaluate + (give info (Y))	
C5: ja/ö	expr: info [Y]; hesitation evoc: continue existing purposes	speaker: commitment to info (Y) listener: evaluate + (continue give info (X))	
A4: vilken månad ska du åka	expr: desire for info [W] evoc: give info [W]	speaker: commitment to need for info (W) listener: evaluate + (give info (W))	

We can see how expectations related to evocative functions and to obligations connected with role and dialogue act influence the interpretation of the utterances and the progression of the dialogue. The first utterance C1 *hup* is not a conventional word of Swedish but a sound which, for example, could be used by a solitary speaker as an expression of surprise or fear. In this context, however, given the purpose of the activity and the roles of the interacting parties, it functions as a summons for contact and perception (attention) and a way of initiating the interaction. In utterance A1, it has been assumed that CPU (contact, perception, understanding) as well as acceptance is part of

what is being mainly expressed which is why CPU and acceptance have been coded while they have been left out in most other utterances. In C4 *va sa du* (what did you say) has been coded as expressing non-perception/understanding of utterance A2. Turning to "acceptance", we can see that it is left out except in A1 and A2. In C2, thus, acceptance of the task of making a request is implied by the fact that C2 can be construed as such a request, and in C5, acceptance of A3 as being a clarification of A2 is implied by the fact that C5 answers the yes/no question contained in A3. Utterances A3 and C5 are both implicit statements expressing beliefs. In the case of A3, it is the belief that A3 is a clarification of A2 and in C5 it is C's belief that he/she wants a return ticket. However, in accordance with what was said above, we code the commitments going with A2 and C5 as commitments to the information. Since the *yes*-answer in C5 is also an implied positive statement that "C wants a return ticket", it carries the default evocative function (cf. Section 7.1) that A should share this belief (not coded). In A4, A does not object but continues his task which then implies that he, in fact, accepts this belief, i.e. that C wants a return ticket.

In utterance C2, the NP *flyg ti Paris* (flight to Pairs), because it is uttered by the customer at the beginning of the activity (after A1 which is a sort of permission to initiate relevant interaction), can function as a request for information giving rise to an obligation for the agent to furnish that information. The reason this is an obligation rather than just a hoped for action from A is that a positive evaluation on A's part can be expected since A, by his role is obligated to provide relevant services. Also, since C has entered the role of customer he/she is, in turn, obliged to provide sufficient information for A to do his/her job. Similarly, the requests for specification in (A2, A3 and A4) and clarification in (C4) give rise to obligations to furnish information which are relevant in the activity and motivated by the roles of the two interlocutors.

7.3.2. The quarrel between two sisters
For comparison, we will now analyse the quarrel in a similar way by first giving an analysis of expressive and evocative functions and then turning to obligations.

In this dialogue, CPU is less taken for granted than in the travel bureau dialogue. In utterances D1, D2, S3 and D4, PU or CPU have been included as main evocative functions since getting the other sister to listen and understand seems to be a main evocative intention which can be less taken for granted in quarrel then in a travel agency dialogue. Another difference is that the roles of the two sisters are such that there is no expectation that positive evaluation carries with it an obligation to act. Thus, the only obligation D has after utterances S1 and S2 is to evaluate whether she is willing and able to cease the action S requests not to be done. A third difference is that since utterances D3 and D5 contain explicit statements, we have used the predicate "belief" to code the expressive function and the commitments generated by this. In S3, which is an implicit statement, we have, like in the travel dialogue, used the predicate "information".

Quarrel—expressive and evocative functions and obligations

	Contribution	Expr. and evocative function		Obligations introduced	
D1:	men herregud ⟨clicking sound⟩/	expr:	irritation	speaker:	commitment to being upset for some reason
		evoc:	PU/irritation	listener:	evaluate + respond
S1:	kan du låta bli min freestyle eller	expr:	desire for cessation of action/irritation	speaker:	commitment to expressed desire
		evoc:	cessation of action	listener:	evaluate + respond
D2:	nä	expr:	refusal	speaker:	commitment to refusal
		evoc:	PU/irritation	listener:	evaluate + respond
S2:	⟨a men släng inte ner den nu⟩⟨yelling⟩	expr:	desire for cessation of action/irritation	speaker:	commitment to expr. desire
		evoc:	cessation of action	listener:	evaluate + respond
D3:	⟨ingen fara⟩/den e ändå så gammal ⟨very slowly⟩	expr:	belief	speaker:	commitment to belief
		evoc:	irritation	listener:	evaluate + respond
S3:	vadå gammal två dar	expr:	protest/info	speaker:	commitment to protest + info
		evoc:	CPU	listener:	evaluate + respond
D4:	aa//	expr:	CPU	speaker:	commitment to PU
		evoc:	CPU	listener:	evaluate
S4:	⟨sings⟩	expr:	disdain	no relevant obligation	
		evoc:	irritation		
D5:	sluta du e ÄCKLI	expr:	desire for cessation of action + belief + irritation	speaker:	commitment to desire and belief
		evoc:	cessation of action	listener:	evaluate + respond

7.4. DISCUSSION

It is fairly clear that the quarrel is different in nature from the travel agency dialogue. The conventional expectations associated with the role of teenage sister and the roles of customer and agent are of a different kind. In the travel agency dialogue, the roles allow fairly good predictions about what communicative acts it is reasonable to expect and about what the obligations of the two parties are, but this is much more uncertain in the quarrel. The interaction between the sisters is in a sense free of clear role obligations. Instead, there is probably a kind of fundamental trust between the two sisters which allows for a breach of some ethical and politeness considerations of obligations as well as for a neglect of obligations generated locally by the communicative acts used by the other party. Thus, D does not placate S by assenting to requests or by trying to lessen the irritation S expresses. Rather, she seems to want to tease S, in order to make her more

irritated. When S starts to sing, by ignoring D she irritates D. D then answers by insulting S. Ethical considerations involving trying not to hurt the other party are thus diminished and some of the features of what in "Conversation Analysis" is called preference organization do not seem to be present.

If we consider to what extent the two dialogues exhibit features of cooperation, we see that in the travel agency dialogue, the two parties clearly take each other into cognitive consideration. They also cooperate in trying to achieve the common purpose of giving and receiving information about travelling.

In addition, they seem to show each other some ethical consideration. The agent, for example, tells the customer to hold on when he is using his computer to find relevant information. Probably, this is also connected with some mutual trust between the parties. Each expects the other party to treat him/her in a way that is correct given her/his roles as customer and agent. This leads to a kind of harmony between communication based on role obligations and communication based on obligations generated by the communicative acts that are used. The travel dialogue exhibits what we might call professional cooperation or cooperation strongly influenced by roles in a conventionalized social activity.

Turning to the quarrel, cooperation is both less obvious and of a different kind, if it exists at all. The two sisters cooperate at least to the extent that they take each other into cognitive consideration. This is shown by the coherence of their interaction. They might also be said to cooperate in the sense that they share the purpose of achieving some kind of mutual understanding. This is shown by the fact that they do seem to interpret each other's utterances in a reasonable way and respond to them coherently. More controversially, one might also claim that they, after a while, come to share the purpose of irritating each other, which, as we can see, has consequences for how they respond to each other's utterances. Whether or not quarrelling or mutual irritation can be accepted as a joint purpose depends on whether the resulting interaction merely is the outcome of two individual purposes (where one person wishes to irritate the other person) or whether it has features indicating a joint purpose such as, for example, mutually licensing neglect of various obligations and commitments. The question of whether the two sisters show each other trust and ethical consideration beyond coherent responses, is an even more complicated issue. They are irritating each other and thus being unethical. However, the pain seems to be kept within certain limits. Therefore, it can perhaps be claimed that even though their interaction is not ethically ideal, there is a sense of trust between the two which means that there will be limits to how much the other party can be hurt—a kind of mutual bond of tolerance up to a point. This kind of fundamental trust might be what often differentiates a quarrel between people who are bonded by, for example, siblinghood, marriage or friendship from a quarrel between strangers or enemies.

In sum, we may therefore conclude that both interactions exhibit cooperation, albeit of different kinds and magnitude. What has been presented so far is an account of some of the main cooperative dimensions of dialogue. The goal of a dialogue is to allow the participants to share awareness and understanding while at the same time attempting to influence each other. In doing so, the dialogue participants often express (and thereby often clarify) their attitudes and emotions. These goals are realized through communicative actions which are guided by cognitive consideration and often also by different types of ethical consideration and trust. The goals are often further reinforced by being linked

to the functional role requirements of a particular activity. The dialogue successively progresses as the senders provide expressive and evocative information, which the receiver(s) either explicitly confirm by the use of the feedback system of a language (Allwood, 1988; Allwood *et al.*, 1992) or implicitly confirm (as being jointly perceived, understood or accepted) by contributing new information building on the previous contribution.

These cooperative mechanisms of dialogue now allow us to explain why there should be such phenomena as "adjacency pairs" (Schegloff & Sacks, 1973), "exchange structures" (Sinclair & Coulthard, 1975), "dialogue grammars" (Moeschler, 1989) or "dialogue games" (Kowtko & Isard, 1991). According to Schegloff and Sacks, adjacency pairs occur as a kind of conventional pairing of one speech act with another and it belongs to linguistic competence (backed up by "preference organization") to know how to respond to a given type of speech act. The problem with this approach is that it does not explain what happens when people respond coherently in unexpected ways. Responses such as *shut up* or *why do you say that*, etc., are always possible, the question is why they do not occur, more frequently than they do.

The view described above, rather than merely invoking a conventional mechanism, instead suggests that relevant pairings of utterances occur because speakers are cooperative, i.e. to some extent consider each other's contributions both cognitively and ethically, share purposes and trust each other.

Thus, in evaluating another person's contribution it would not be cooperative to just ignore it or to reject it out of hand without reason. Instead, we usually try to at least perceive, understand and continue and, if we have no reasons against, comply with the main evocative intention, When such compliance occurs a successful "adjacency pair" is produced. However, what has occurred is not merely an instance of a conventional mechanism but rather a voluntary ethically motivated action.

The regular and expectable features of dialogue should be seen as an outcome of cooperation in which expressive and evocative features of contributions, on the basis of obligations, are evaluated and responded to by new contributions with new expressive and evocative features. In this process, a large part of the bond and coherence between utterances is provided by meeting the obligations given by general ethics, activity roles and particular communicative acts. Since it has further been claimed that cooperation is a matter of degree, which is based on the willingness and ability of the participants, regular dialogue features can, at any moment, be modified, changed or interrupted. The fact that this does not happen more often than it does is a sign of the strength of the role that cooperation plays in human social life.

8. Summary of the claims made in ACA

Activity-based Communication Analysis relates, characterizes and /or defines several different concepts related to cooperation and communication. Among these are "cooperation", "communication" and "social activity". Below, some of the main points of the theory related to these concepts are summarized.

(1) Cooperation is a matter of degree.
(2) Ideal cooperation involves at least four requirements: cognitive consideration, joint purpose, ethical consideration and trust.
(3) Ideal cooperation as well as ideally cooperative communication require these four criteria to be met.
(4) Actual cooperation and communication which is based on the actual backgrounds and interests of particular interacting parties, however, might not meet all four requirements.
(5) Whereas the lowest degree of cooperation itself only requires cognitive consideration, communication always requires (even in conflict) a little more, i.e. cognitive consideration and the joint purpose of mutual comprehension.
(6) Cooperation is primarily connected with joint social activity through the criteria of joint purpose and cognitive consideration.
(7) Actual communication in different social activities is oriented toward ideal cooperation through different roles, whereby also ethical consideration and trust will be activated in a way which is adapted to the role in question.
(8) Actual communication is the outcome of many influencing constraints and enablements. Some of these depend on individual backgrounds and interests whereas others are of a more social (often normative) nature.
(9) Some constraints and enablements take the form of social obligations influencing communication. These are at least of three kinds.

  (1) General obligations (based on the ethics of motivated rational agency).
  (2) Role-dependent obligations.
  (3) Obligations induced by particular communicative acts.

(10) In actual communication all three types of social obligations interact to influence actual communicative behaviour.

9. Implications for computational agents

The theories of cooperation, social activity and dialogue presented above primarily concern human cooperation and communication. An important question is to what degree this kind of theory also applies to computational agents. Computers can easily fit into the above analysis in the form of *artefacts* (e.g. in the travel agency activity). While this may be an important aspect of some activities and individual dialogues (e.g. about a computer), this is not really the sense we are concerned with. Rather, the concern is whether computer systems can be *agents*, acting as dialogue participants, with individual attitudes, cooperative behaviour, rights and responsibilities and filling roles in an activity.

On the one hand, computer systems (at least those currently existing) are certainly *not* agents in the same way that people are. They do not have primary goals of seeking pleasure and avoiding pain, the same sorts of physical embodiment, the same kinds of competence concerning language usage and commonsense reasoning, or the same embedding within the social structure. They are also, for the most part, designed as (complex) tools for particular purposes. For all of these reasons, in many ways they play more the role of an artefact than the role of a participant.

On the other hand, there are a number of reasons for viewing some computer systems as agents, with many of the same attributes (at a suitably abstract level) as humans. For one thing, agent-systems are currently a very popular computational metaphor (Jennings, 1999). Designers of such systems certainly view them as agents, as do some users. Such systems can have some of the basic aspects of agency, including autonomy, goals, and independent action, especially as viewed from the outside. Likewise, many speak of extended human-computer or machine-machine interaction as "dialogue". One might debate to what extent it is appropriate to apply the same concepts (e.g. purpose, cognition, ethics, trust and obligations) to computer systems as to human beings. While we will not take any deep theoretical stand on this question here, we will merely point out that there are several reasons to use the same kinds of analysis for dialogue with computational rather than biological agents. First, since many of these computer agents are meant to interact with humans, it would be useful for them to be able to model the way humans engage in the process of dialogue. Second, as McCarthy (1990) points out, humans often use mentalistic metaphors in talking and thinking about the operations of complex machines. Computer systems can often fill some of the same roles as humans in traditional human activities, e.g. as a bank teller or an information provider. Finally, this is how many AI researchers, themselves, conceive of the systems they design. Thus, mentalistic terminology such as beliefs, goals and intentions are part of the formal language of computer dialogue systems, regardless of the actual nature of the relation of these terms to the human attitudes or natural language words.

Thus, we feel it is important to avoid both extremes in the debate and implications over the nature of computational dialogue participants. We should not ignore the agentive properties of such systems and deny the similarity of human–computer to human–human dialogue (especially when using human means of communication, such as natural language utterances), and so have to build a completely new theory of artificial dialogue, from the ground up. Aside from the wasted effort, such an enterprise has dubious utility, since computer systems may be as different from each other as they are from human beings. On the other hand, we should not blindly assume that (current generation) computer systems are no more different from humans than one human from another, and apply all aspects of a theory of cooperation and dialogue for humans. For example, even though the same banking function can be performed by interacting with a human bank teller or an ATM, the activities are not at all the same. In the case of interacting with an information providing system, the differences are slightly more subtle, but present nonetheless. Instead, we feel the right approach is to use the same abstract theory, applying it specifically to the differences, thus yielding different predictions and analysis of specific cases.

In the rest of this section, we will briefly discuss some of the most important differences between computer dialogue systems and humans with respect to the issues discussed in the paper. These points will be illustrated with an analysis of a small portion of a dialogue between a person and a computer system, parallel to the analysis of the human-human dialogues in the previous sections.

9.1. INDIVIDUAL AGENCY

In a broad sense a computer agent may be described as having some of the main individual features of agency, including purposes, beliefs, intentions, rationality,

motivated action and ability to consider the other party in various ways. Some significant differences include: lack of an over-riding pleasure–pain principle motivating action, lack of sophistication of interlocking motivations, and lack of bodily limitations (attention, hunger, tiredness, etc.), but the presence of other physical limitations such as electricity as well as very narrow and limited understanding and competence. Going through the seven principles of motivated rational agenthood from Allwood (1976), presented above in Section 2, we can briefly explore some of the most relevant differences. The first, general principle ("Typical human beings are motivated rational agents") clearly does not apply, since computer systems are not humans. It is also probably unwise to assume a companion principle that all computer systems (even those termed as "agents") are motivated rational agents, in the sense characterized by the more specific principles, since many systems are clearly not.

Concerning the second principle ("The intentionally controllable behavior of an agent is intentional and purposeful"), a question comes up as to how much of a dialogue system's behaviour is controllable. One might claim that none of it is, since a system is programmed. Leaving that issue aside (since it might also cause one to bring up the debate about free will among humans), we could stipulate that intentionally controllable behaviour for systems is a situation in which the system can consider and choose from a set of alternative behaviors. According to this definition, only a subset of the dialogue behaviour of a system would be considered intentionally controllable. For example, a dialogue system might be able to choose whether to respond to a question with "no" or "please repeat", but might not be able to formulate different kinds of negative or qualified answers (or choice of words or politeness) due to limitations in its representational or language generation capacity. On the other hand, a system might be able to consciously deliberate and choose whether to perform some behaviour which in humans would be completely reflexive (e.g. some kinds of low-level feedback behaviour). Thus, while the principle applies equally to humans and machine agents, the implications are very different as to how to ascribe particular behaviour. The third principle ("The actions of an agent are not performed against his own will") also has more limited application to computer systems, since they are generally not "in danger by not performing an action". There is still the question, considered above, of to what degree the systems have a choice in their behaviour.

The fourth principle ("The actions of a normal agent are motivated") can be applied to computer agents since they can also be motivated in the sense of choosing means that will tend to achieve their ends. For many systems, it is still a question of whether this choice is done by the system itself, or by the designer of the system, in choosing a particular programming. As stated above, the fifth principle ("Normal agents do not act so as to decrease their pleasure or increase their pain") does not really apply, since few, if any, systems operate on the basis of pleasure/pain.

The rationality principles six and seven ("The actions of a rational agent are selected so as to provide the most adequate and efficient way of achieving the purpose for which they are intended" and "The actions of a rational agent are performed only if he thinks it is possible to achieve their intended purpose") seem as directly applicable to computer agents as to people: the actions are selected to be the most adequate and efficient for achieving the intended purpose, and the actions are performed only if the agent thinks it is possible to achieve the purpose. There are still issues about how well actual dialogue systems meet these rationality goals, due to different kinds of skills at using language.

9.2. COOPERATION

Given that computer systems can be (but are not necessarily) agents (though different in kind and degree from human agents), the next question is to what extent they can be cooperators. They certainly can be said to be cooperative, in the weak, instrumental sense of facilitating one's goals (as pointed out in Section 2.2). However, we are more interested in the extent to which they can be cooperative partners by participating in the aspects of ideal cooperation, presented in Section 2. Weak degrees of cognitive consideration are not a problem for computer systems, given any internal representation of a partner. Of course, not all systems will have or use such representations. Traum and Allen (1991) present various levels of interacting agents in terms of what kind of cognitive consideration was needed, concerning the agency of another agent. User models are also popular ways of designing systems with a capacity for cognitive consideration. This does not mean, of course, that a given dialogue system will actually engage in (much) cognitive consideration of its interlocutor. Many systems are designed to react to specific linguistic inputs with fixed responses, without any consideration to the user.

Computer systems can also have joint purposes. Often dialogue systems are designed specifically to achieve a particular purpose. Knowing this purpose, a user can easily achieve a fairly high degree of joint purpose, when this is shared. Some systems are also set up to help a user achieve some purpose in a particular domain. Given potential limitations on cognitive consideration, a system might not actually contribute to mutual awareness of this shared purpose.

Computer systems may also have some aspects of ethical consideration, as represented by the ethical maxims in Section 2.3. Concerning maxim (1), it is certainly possible for a system which has the capability to weigh alternative actions and cognitively consider another agent to adhere to maxim (1) and not force the other agent. Many dialogue systems achieve this, to some degree, by giving initiative to the user. Some systems are quite frustrating to use/interact with, precisely because of ignoring this option, i.e. only allowing the user limited choices, which do not include the desired one (even when such an option is logically and technologically possible). Likewise, maxims (2) and (3) seem fine as normative principles for computer systems to adhere to, especially in relation to human agents. A more difficult question is to what degree a human should behave according to these maxims concerning computational agents. Certainly, in a user-system relationship, there may be no over-riding need for the user to allow the system to function as an agent, by adherence to the principles. In this case, the system functions more as an instrument, regardless of the dialogue-like interface. Adherence to such principles might follow directly from self-interest (e.g. giving the system adequate information, so it can correctly calculate the answer to the user's question). While this may be appropriate behaviour toward a system (though not toward a person), we would not say that such behaviour is very cooperative. Another instance where one might want to be more cooperative even towards a system is when it functions as an "agent" (here in the sense of a factotum, rather than of an actor, which has been the general use of "agent" in this paper). In such a case, it may violate ethical obligations to the person or group for which the system is employed to violate the ethical maxims toward that system.

Computer systems also exhibit some degree of "trust", although it can be difficult to determine exactly how to characterize this or to what degree trust must be involved. At

a minimum, computers and humans must trust each other to use language appropriately to express meanings and to act according to principles of rational motivated agency (although this trust might be implicit rather than a result of conscious deliberation). More problematic is the degree to trust the other to engage in cognitive and ethical consideration and to have a joint purpose.

9.3. SOCIAL ACTIVITIES AND DIALOGUE FEATURES

The main components of activity analysis are equally applicable to both human and computer participants. In both cases, we expect the same aspects of purpose, roles, artefacts and physical and social environment to be important in classifying the activity. However, as stated above, there will be differences in activities depending on human or computer participants. Some of these go beyond the fact that a computer system plays one of the roles, to differences in the activities themselves. *Purposes* in many cases will be more or less the same for human-computer activities as for activities involving two humans. As in human–human dialogues, purposes may be related not just to the ends of an activity but also to the process of engaging in the activity itself (to reach those ends). Thus (part of) the purpose of a user may be to use the system, rather than just to achieve the aims of the interaction. *Artefacts* may be the same in human–human and human–computer activities, a constraint being that the computer interaction media must be present, while many human activities are face to face. Similar things may be said about the physical and social environment, with, usually, a significant difference in the social environment, both in terms of the computer being more of a servant or tool, and not being part of the human social structure. The biggest differences will be in the roles played by computer and human participants. As regards these, there will generally be similar task-related competence requirements, however, there may be lesser or different requirements on the computer concerning general abilities and language competence. Likewise, for computer participants, there are generally fewer rights and more obligations to follow the purposes of the user than would be the case for a participant in a human–human activity.

   The main broad dialogue features will be the same for human and computer dialogue participants. That is, for both, there will be both expressive and evocative functions, default functions of CPUR, (R = acceptance) and obligations on speaker and hearer for sincerity, motivation and consideration, and evaluation, report and action, respectively. However, there will be differences in the particular functions and obligations of utterances, springing mainly from the different competences and social relations of computer participants.

9.4. EXAMPLE: HUMAN–COMPUTER DIALOGUE

In order to illustrate some of the differences between human–human and human-computer dialogue, we will now present an analysis of a human–computer dialogue from the TRIPS system (Ferguson & Allen, 1988) in a manner similar to the presentations of the human–human dialogues in Sections 6 and 7.

9.4.1. *Activity analysis*

The activity in which the TRIPS dialogue is a part is actually a rather complex one. At the most direct level, it involves transportation planning. However, it is not a natural activity, like the previous two dialogues, but a "laboratory" or "classroom" task, in which the participants are "pretending" to do the activity for the purposes of the experiment, rather than actually engaging in it because of a desire or need for transportation plans. Moreover, the setting is a simulated area, rather than the real world, so some aspects of the task may bear only an abstract relationship to the actual task being simulated. On top of this, part of the task is to show off the abilities of the computer system to participate in this task. Thus, above and beyond the differences and limitations of computer systems described in the previous section, we should expect some further differences from a naturalistic activity of the same sort.

1. *Purpose.* As the activity itself is complex, so is the purpose a combination of the purposes of these sub-tasks. The purpose of the object-level task is to solve a transportation planning problem—in this case to evacuate people from an Island (called Pacifica) by transporting them all to the city Delta where presumably they will be flown to safety. One level up, the purpose is to simulate being engaged in such an activity, and finally, there is a purpose of observing and evaluating the abilities of the system to engage in this kind of joint activity.

2. *Roles.* The activity has two roles: a system that provides information about and executes requests concerning logistics and a user who is using the system to find information and accomplish the given goal. The roles are also complicated because of the complex cluster of sub-activities. At the direct level, the system has the competence requirements of knowing about the domain, and general conversational competence (limited by the abilities of the system). The "obligations" of the system are to provide relevant information and execute given directives. The user's competence requirements include knowing what the task is, and engaging in dialogue to be able to solve the task. These requirements are, of course, within the "game" of solving the task, so in a sense there are obligations on both to "act naturally" rather than with more real motivations and actions. Since the dialogue is also meant to test and showcase the abilities of the computer, the user has obligations to treat the system as a dialogue partner, rather than a tool, regardless of how effective that might be toward achieving the task.

Since the system is a computer, it is, however, questionable if it has any rights or obligations except in a metaphorical sense. Turning to the user, it could be said that rights and obligations require a relation to another agent who also has rights and obligations. If the system has no rights and obligations, this requirement is not met and the user has no rights and obligations either (at least not in relation to the computer). However, behind the system there are agents who have created and provided the system and in relation to these agents, the user has rights and obligations. The user role could perhaps then be characterized as follows: with regard to competence the user should have a sufficient degree of "computer literacy" to use the system. As for rights and obligations, the user has the right to expect the system to be courteous and be competent as far as planning "logistics" goes and he/she has the obligation to act/communicate in a way which is relevant to planning transport logistics.

3. *Artefacts.* The computer's hardware is itself an artefact and thus the computer acquires the double status of both means and ends for the dialogue. It is fairly obvious that this influences the dialogue in powerful ways. On a more superficial level, factors such as microphone, type of monitor and seating arrangements will have an influence. In addition, there are maps that the computer can display, and signals of its degree of understanding of the speech, as a sort of feedback communication, if the user attends to these displays.

4. *The physical and social environment.* The environment is a user in front of a terminal communicating with the system. In the actual dialogue instance below the setting for this is a university research environment.

9.4.2. *Expressive and evocative functions and obligations*

We now proceed to an analysis of the functions of the utterances in the TRIPS dialogue presented below. Utterance U.1.1 is similar to utterance C1 in the travel dialogue, but uses a greeting rather than an unconventional sound. It is uttered by a user who knows this is a way to start interaction. The code "social recognition" is taken to be what would in non-computer circumstances be expressed by a greeting. It is hypothesized that this function is carried over to the computer setting. However, it is doubtful whether the user thinks he/she can really evoke social recognition from the computer. Just in case, however, the system's reply S.2.1 *hi* could be said to express social recognition. Since both system and user "know" that the way to initiate task interaction is for the user to make a request, S.2.1 can have this evocative function. Likewise, if the user can have obligations, the role of a user now creates the obligation to initiate a relevant interaction. In U.3.1 a new feature is explicit coding of the presumably responsible user's cognitive consideration of the system's competence or ability to carry out the desired action. As in the travel dialogue but distinct from the quarrel, the roles of system and user are asymmetrical and carry mutual obligations for actions which should be carried out if evaluation is positive. This can be seen, for example, in the way the expressive function of S.2.1 meets the requirements of the evocative function of U.1.1.

For U.5.1, we provide two alternatives since the system wrongly analyses the user's utterance *where are the people* as *clear the people*. We first give the system's interpretation and then the intended interpretation. In S.6.1. we have let the regret normally expressed by such a phrase be carried over to the system, even though it could be objected that the system has no such feeling, and this is merely a "canned" bit of text to express negative feedback of understanding in a naturalistic way. Utterance U.7.1 is similar to utterance A3 in the travel dialogue, but whereas A3 gives information explicitly asked for in C4, U.7.1 provides a clarification implicitly evoked by S.6.1 (*I'm afraid I didn't understand you*). Utterance S.4.2 is interesting since it introduces a multimodal communicative act —a display of a map, which the system has committed itself to in S.4.1 and the user requested in U.3.1.

Contribution	Expressive and Evocative Function		Obligations Introduced	
U.1.1 HELLO	expr:	social recognition, desire to interact	speaker:	commitment to socially recognizing listener
	evoc:	social recognition	listener:	evaluate + respond
S.2.1 Hi	expr:	social recognition willingness to interact	speaker:	commitment to willingness to interact
	evoc:	state request	listener:	evaluate + continue relevantly
U.3.1 SHOW ME A MAP OF PACIFICA	expr:	desire for action	speaker:	commitment to expressed desire and consideration of systems ability to perform action
	evoc:	perform action	listener:	evaluate + perform action
S.4.1 OK	expr:	acceptance to perform action	speaker:	commitment to perform action
	evoc:	CPU	listener:	evaluate + C
S.4.2 DISPLAYS MAP	expr:	info	speaker:	commitment to info
	evoc:	CPU	listener:	evaluate + respond
U.5.1 CLEAR THE PEOPLE	*ALT*1 expr	desire for action	speaker:	commitment to desire + consider system competent
	evoc:	perform action	listener:	evaluate + perform action
	ALT 2 expr:	desire for information	speaker:	commitment to desire + consider system competent
	evoc:	give information	listener:	evaluate + give information
S.6.1 I'm afraid I didn't understand you	expr:	belief/regret not U	speaker:	commitment to expressed attitudes + consider speaker capable of clarification
	evoc:	clarification	listener:	evaluate + give clarification
U. 7.1 WHERE ARE THE PEOPLE	expr:	clarification + desire for information	speaker:	commitment to desire consider system competent
	evoc:	give information	listener:	evaluate + give information

9.5. HUMAN-HUMAN AND HUMAN-COMPUTER DIALOGUES

As our discussion above has illustrated, the framework of activity-based communication analysis is applicable to both human–human and human–computer dialogues. In human–computer dialogues, an analysis of the type presented can be used to facilitate, for example, interpretation, response planning and generation. A problem in extending the framework to human–computer analysis is that certain concepts like "rights" and "obligations" have to be taken in a transferred, metaphorical sense when applied to a system. A similar problem appears when attributing attitudes such as regret or social recognition to a computer. The problem reappears in applying the parameters of cooperation. Can a computer system really take a user into ethical consideration or trust him/her? Certainly, it will have somewhat different attitudes from the familiar human ones. This conceptual difficulty here might be what lies behind the fact that most earlier work concerning cooperation with computers has been based only on cognitive consideration and shared purpose. However, as we have tried to argue, we believe this to be insufficient and see no reason why a computer system could not engage in some degree of ethical consideration and trust. In principle, this is no more or less strange than equipping a computer with beliefs, intentions and desires.

10. Ideal cooperation and activity analysis in the design of computer dialogue systems

The first attempt to apply the theory of ideal cooperation and communicative activity analysis to a computational framework was done within the project Pragmatics-based Language Understanding System (PLUS). PLUS was a large European project which aimed at studying the use of context and pragmatics in human–computer dialogues and building a robust and cooperative interface to an electronic Yellow Pages (Black *et al.*, 1991; Cunningham, Black, Underwood & Jokinen, 1993). Robustness was understood as the system's ability to behave adequately in a wide range of situations (to cope with syntactic deficiency and semantic vagueness, as well as be flexible enough to allow a real dialogue with the user), and to be achieved by studying natural language as dependent on activity with the essential characteristics of conveying meaning that is both appropriate and relevant in context.

The PLUS discourse model employed two complementary pragmatic theories, ACA (Allwood, 1976, 1978, 1992; Allwood & Haglund, 1991), and dynamic interpretation theory (Bunt, 1990), the latter concentrating on the dynamic aspects of dialogue interpretation and regarding communicative acts as functions from one dialogue state to another. The dialogues were understood as linguistically expressed complexes of communicative actions, described in terms of their effects on the context. The system"s reasoning used three sets of rules: pragmatic rules which encoded the system's knowledge about cooperative, rational behaviour and world model rules which dealt with the system's knowledge about the actions and objects that populate its world, and the reasoning rules to update and relate the world model entities.

PLUS had ambitious aims for the use of pragmatic and contextual reasoning in the building of a robust dialogue system. While the project contributed to better understanding of dialogue systems that would exhibit cooperative and flexible behaviour, the ideas

of pragmatics-based language understanding were not fully realized in the form of "PLUS technology", and more work is needed to spell out the power and advantage of such an approach.

In the framework set up by PLUS, Jokinen (1994, 1996*a–c*) developed a computational model which explicitly takes into account the ideas of activity-based communication analysis. This *Constructive Dialogue Model* (CDM) includes both aspects of traditional AI accounts of rational agency, as well as the aspects of ideal cooperation, formulated above. CDM models the basic attitudes of beliefs and intentions, and then updates the context of a dialogue with the expressive and evocative itentions. It emphasizes the agents' knowledge of the domain and of appropriate dialogue strategies, rather than their predefinded dialogue behaviour. Navigation through domain-knowledge is controlled by topic shifting rules, while the speaker's intentions guide expectations of the next utterance. Consequently, the "constructiveness" refers to the construction of mutual context via a repeated cycle of interpretation, evaluation and reaction to the partner's contributions.

Most of the focus in the model is on utterance/response planning, which starts from the intentions of the system, and applies filtering rules which consider cooperativity and sender and receiver obligations. The rich context model and constraint-based processing model allows integration of content and form planning in natural language generation. However, the interaction between different obligations and their formalization as general kinds of constraints was left implicit, and a preliminary attempt to formalize these notions is presented in Jokinen (1997). The implemented system also only considered factual information-seeking activities and roles of the user and the system therein.

11. What is shared in dialogue?

If we put together the features of cooperation with the basic communicative requirements of contact, perception, understanding, attitudinal and behavioural reactions we get an analysis of what is shared in dialogue. We will see that "common ground" in the sense of a set of shared beliefs, e.g. Clark and Wilkes-Gibbs (1986) will turn out to be a special case of what can be shared in dialogue. Instead, we get a more complex picture with a build up of sharing and mutuality on many levels. Basically, however, there are two phenomena: (1) build up of what we might call shared inclinations, and (2) use of linguistic and other communicative behaviour which manifests what is being shared by establishing links from one utterance to another.

In what follows, we will attempt to connect these two phenomena by first discussing what is to be shared and then discussing how this is cooperatively achieved by different types of linguistic and communicative behaviour. Although we will be treating the various kinds of shared information one by one, it is important to note that in normal dialogue they are present simultaneously, mutually supporting each other.

11.1. CONTACT AND SHARED WILLINGNESS TO CONTINUE

A very basic requirement on communication is that the parties are in contact and are willing to continue the contact. If A attempts to communicate with B, he/she can, by virtue of cognitive and ethical consideration as well as trust, expect B to respond, at least

by indicating that no contact is possible. In fact, any response from B is enough to manifest contact. To build up willingness (motivation) to continue, a continuation must not be painful but rather pleasant. This can be achieved by friendliness, removing tension (smiles, laughter), not imposing, giving support and politeness. Unless other bonds exist between the parties, manifest links of these types are usually necessary means to maintain contact. This can be seen in the travel dialogue where the first (unconventional) utterance *hup* is responded to by *a* (yes) from the travel agent showing his willingness to interact.

11.2. SHARED AWARENESS AND UNDERSTANDING

Another basic requirement, which perhaps is the main purpose of communication and dialogue, is shared awareness and understanding. In order to evaluate how we should react to another person's evocative intentions, we must understand what they are. I can understand your claims without necessarily taking a stand on whether I believe them or not. In fact, we mostly continue talking about topics without sharing all beliefs about them. Attempting to reach agreement is one of the things that drives dialogue forward. So even though we are not always building up shared beliefs by agreeing on them, we are building shared awareness and shared activation of information.

There are many ways in which participants in dialogue can manifest shared awareness and shared activation of information to each other. Let us look at some of them. One of the simplest and clearest ways of manifesting shared awareness is to share expressions with another speaker, i.e. to repeat fully or partially the words used by another speaker. This serves as an indicator to both parties that the information activated by the words is jointly activated. Because of its convenience, repetition is a standard device in most language communities to show awareness of a common referent or topic and it is used even if the purpose is to contradict the claim of another speaker as in the following example:

A: I like snow
B: snow is too cold for me

Similarly, repetition is a standard device for giving positive feedback in many languages (cf. Allwood *et al.*, 1992).

A similar function to that of repetition can be achieved by reformulation, perhaps most strikingly by pronominal reformulations which require knowledge of either previously activated information or (in deictic usage) of information directly available in the context. Shared awareness of jointly activated information is also a precondition for the use of the definite article which requires previous shared acquaintance with whatever is being referred to (Strawson, 1950, 1964; Allwood, 1977*b*; Clark & Marshall, 1981).

11.3. SHARED PURPOSES

Shared purposes are a third basic requirement of dialogue. We have already mentioned such shared purposes as "willingness to continue contact" and "mutual awareness and understanding". Another kind of purpose comes into play when we consider the purpose and obligations connected with various types of communicative acts. For example, if

A wants to make a statement and B shares information, B can help A to finish his statement in the following fashion:

A: tomorrow I will go
B: to London

A and B here coproduce the statement, *tomorrow I will go to London*, thus exhibiting a joint purpose in coproducing a statement. Normally, as in this case, coproduction of a communicative act will presuppose not only a shared purpose but also shared activated background information.

The so-called adjacency pairs are a second case of shared purpose. The listener accepts the obligation to evaluate the speaker's main evocative intentions and then evaluates whether he/she is able to comply with this intention, e.g. answer a question, agree to a statement or carry out a request, and then responds accordingly. Shared purposes are also involved in so-called elliptical utterances. Consider the following example:

A: what is the weather like
B: cold

Both parties are aware of the evocative intention of A's question. This enables B to accept the obligation of providing an answer by making use of the jointly activated information in A's question and just supplying a relevant weather predicate. The referent of the predicate and the functional role of the predicate as an answer is all shared information given by the previous utterance. "Ellipsis", thus, works by making efficient use of shared cooperative assumptions about coherence and relevance of an "elliptical" utterance to already jointly activated information.

In the same way, feedback utterances like *yes*, *no* and *m*, get most of their actual meaning from the content and purpose that have been activated by the previous utterance. A *yes* can be used merely to signal that the purpose of shared perception and understanding has been achieved. It can, however, in other contexts acquire further functions. After an invitation, it becomes a signal of acceptance. After a statement it can become a signal of agreement, etc. Again, shared awareness of the content and evocative intention of the previous utterance are the determining factors. See also the use of feedback words in the two analysed dialogue excerpts in Sections 6 and 7.

It is important to note that although the upholding of shared purposes has relied on manifestation of communicatively relevant behaviour in a local and adjacent context, in the examples we have discussed so far, i.e. coproduction, adjacency pairs, ellipsis and feedback, this need not always be the case. Purposes can also be global, cf. Allwood (1984). This could be the case, for example, if participants in dialogue were making a joint plan. The joint purpose of plan making could then be used to over-ride the purpose of an adjacent utterance (like gossip or jokes), to bring the dialogue back to its main purpose, as predicted also by models of the intentional structure of task-oriented dialogue, such as Grosz and Sidner (1986).

As we can see, sharing of purposes is a striking feature of dialogue cooperation. Since it is very natural to connect this feature to the notion of dialogue relevance, i.e. an utterance becomes relevant by being related to some local or global purpose or prerequisite of a dialogue (cf. Allwood, 1984, 1995a). Because of its ubiquitous usefulness in dialogue, we can perhaps now understand why the notion of relevance has become so popular.

Probably, the intuitive appeal of the notion of "dialogue relevance" is a reason for the popularity of non-dialogue oriented analyses of relevance, e.g. Sperber and Wilson (1986).

11.4. SHARED ATTITUDES

Besides shared activation of information and shared purposes, dialogue often also involves a build up of shared attitudes. Perhaps the most important attitudes are beliefs. Beliefs are, for example, what are shared when agreement is reached in a dialogue. But beliefs are not the only attitudes that can be shared. Other cognitive and emotional attitudes such as surprise, hopes and joy can also be shared. When a belief is shared, this can be signalled, for example, by using feedback words like *yes* or by a combination of a feedback word and a pronominal reformulation of whatever is being agreed with.

As has already been discussed, the shared awareness of the evocative intention of a preceding statement will mean that the feedback word and/or pronominal reformulation are sufficient to serve as manifestations of a shared belief. Shared beliefs are also manifested through coherence and non-contradiction between the contributions of participants in a dialogue. If what others are saying is in agreement with what I have said and what I believe, I think we share beliefs.

As a consequence, shared beliefs are often implicit. They are probably more often activated by implications of various kinds, e.g. implicatures or presuppositions than they are by explicit statements. Since awareness of implied beliefs is dependent on a shared background, this means that a shared background which can be jointly activated is an essential precondition for the build up of a shared space of information in dialogue. Another precondition is, of course, the willingness to do so which is deeply connected with the presence of ethical consideration and trust.

12. Overview and comparison to other approaches

Let us now summarize some of the traits of the analysis of cooperation presented so far. Cooperation in dialogue and communication is seen as a function of many interacting factors, the most important of which will be mentioned below. The factors are causally influential but can also be seen as requirements of a specific type which, when related to social activities, can become obligations to behave in a specific way.

(1) Individual motives of several kinds.
(2) Functional requirements on coordinated action, to a large extent this is contained in what is called cognitive consideration in the definition of cooperation (given in Section 2).
(3) Activity and role requirements, as illustrated in the travel agency and quarrel examples in Section 7. They provide much of the empirical content of the "joint purposes" in the definition of cooperation. They also influence the above-mentioned functional requirements.
(4) General ethical requirements and trust, see the discussion in Section 2.
(5) Requirements posed by the interaction of ethical requirements both with the communicative acts actually produced and the actual intentions and interpretations of these, as described and illustrated in Section 7.

None of these factors alone are sufficient to determine or predict the nature of cooperation. They all influence each other to yield the type of cooperation that results. Thus, the kind of cooperation that follows a request depends on the activity in which it is made, the role of the requester (e.g. officer or private), the general ethical requirements of the situation, as well as individual motives and functional requirements. In some situations, where some of these aspects can be held constant, it might be possible to simplify the analysis somewhat so that, for example, only the last three types of requirement (i.e. requirements based on ethics, activity and communicative acts) or even the last two types of requirements (i.e. activity and communicative acts) are taken into account. This is often done, for example, in constructing a dialogue system, where the system will only work on a pre-determined problem, using pre-determined means. Even in such circumstances, the simplified models may prove brittle when trying to adapt to new domains, or in cases where different communication situations are actually present than were designed for.

12.1. COMPARISON TO OTHER APPROACHES

If we compare the present treatments with treatments of cooperative interaction found, for example, in the accounts presented by Grice, Clark or more implicitly in conversation analysis (CA), some differences can be noted. The account given here is very different in focusing on the importance of ethics for cooperation. It is also significantly different in stressing the importance of social activity and the roles in social activity for cooperation. Here, however, the increasing amount of analysis of institutional talk within CA will probably lead to a similar view.

This analysis also provides a contrasting view of the type of "grounding behaviour" described in Clark and Wilks-Gibbs (1986) and Clark and Schaefer (1989). According to Clark's formulation, "grounding" is viewed as subordinate to the joint collaborative purpose, and subject to an economy principle of "least collaborative effort". In the analysis given here, this might coincide with the central role given to the purpose of the activity for deciding what is an acceptable level of perception and understanding. This level will vary in virtue of different activity types and of how central the information is to the activity. The analysis, however, also has an additional mechanism which underlies "grounding behaviour", namely, the obligations to evaluate, mentioned above. This general obligation helps explain the presence of "grounding" even in cases of casual conversation or conflict, where no clear purpose above socializing or pursuing conflict can be related to the information expressed.

When it comes to the contribution to cooperation given by communicative acts, the present approach is similar to Grice in being open to the importance of actual intention and actual interpretation. However, it is different from a Gricean view by giving the linguistic form and the context a more important role (at least compared to Grice, 1975). It differs from the Clarkian view in recognizing communicative acts as having occurred even if the intended listener reaction has not taken place. A communicative act does not have to be completely successful in order to occur. The approach also differs from CA in allowing reference to intentional features like expressive and evocative function or even the notion of obligation.

Most of the AI dialogue work has focused on cognitive consideration, to the exclusion of the other aspects of cooperation presented here. This kind of work, which views dialogue participation as considered, rational behaviour, in support of goals and with reference to beliefs has many advantages in flexibility and general applicability over more hardwired, purely "reactive" approaches to dialogue. However, it is very difficult (or perhaps impossible) to provide a sufficient explanatory and productive account of dialogue behaviour, purely in terms of individual motivations. For example, in terms of dialogue, the main way of viewing communication has been as the performance of speech acts [based on the ideas of Austin (1962) and Searle (1969)]. Using only the notions of cognitive consideration, theorists were forced to describe the enabling conditions and effects for these actions purely in terms of beliefs, goals and plans. While even in the early work (e.g. Allen & Perrault, 1980; Cohen & Perrault, 1979), cooperation was seen as an important ingredient in the explanatory process of dialogue coherence, it was usually left unspecified as to precisely what it entailed or how to judge whether cooperativity was present.

Later work, including that described in Section 4, has focused on also explicating the nature of joint purpose, and an account of what the cognitive implications are for engaging in a collaboration on a joint task. Such accounts have also been the basis for more promising accounts of speech act meaning and dialogue participation (e.g. Cohen et al., 1990; Sadek, 1991, 1994; Lochbaum, 1994, 1998; Chu-Carroll & Carberry 1998, 1999). Cooperative processes, such as negotiation have sometimes been based on iterative plan reasoning and performance of meta-acts (e.g. propose–evaluate–modify cycle proposed by Chu-Carroll (1996), Chu-Carroll and Carberry (1998) to provide cooperative responses). Bringing in not only notions such as Mutual Belief, but also the notions of cooperation/collaboration themselves and how they relate to the attitudes of the partners involved, gives a more coherent account of the dialogue participation process. However, excluding explicit notions of ethics and trust makes some of these theories more complex and indirect than they need to be. Some of these theories, e.g. the formulations of joint intentions and SharedPlans described above, try to enforce ethical and trusting behaviour by reconceptualizing these as individual goals or mutual beliefs about intentions. It is questionable whether these formulations can actually achieve appropriate behaviour in all cases (e.g. when agents change their goals or intentions from what is normatively still required), but, in any case, it seems that a more direct representation of the social aspects of dialogue can yield a more straight-forward account.

There have been several attempts to augment the "Belief, Desire, Intentions" AI models of dialogue participation with other, more social aspects of dialogue, which yield models closer to the framework of Ideal Cooperation presented in this paper. These accounts take inspiration from theorists such as Sacks and Clark, in addition to the ideas of Searle and Grice used by the more traditional AI accounts. McRoy, for instance, designed a system to perform speech act identification using abductive reasoning about the motives of another speaker, in combination with the expectations of dialogue coming from Conversational Analysis accounts of adjacency pairs (McRoy, 1995; McRoy & Hirst, 1995). This account provided a natural way to detect and repair misunderstandings, with less detailed analysis of overall motives than was required by theories based solely on plan recognition (e.g. Allen & Perrault, 1980). Traum (1994), Traum and Allen

(1992) presented a computational account of Clark and Shaeffer's grounding process as the performance of dialogue acts related to the understanding level discussed here, which allowed the maintenance of illocutionary act effects such as mutual belief, while still providing an account of the function of feedback utterances.

These accounts of dialogue also considered some of the social pressure effects of utterances in dialogue (cf. Bunt, 1990). For example, Traum and Allen (1994) used obligations as some of the direct effects of utterances, in a manner similar to that presented in Section 7. This use of obligation as a basic motivation for action (along with personal goals) provided both a simpler account and more explanatory power for behaviour in a wider range of circumstances than those relying on cognitive consideration alone (cf. Traum, 1996).

The formulation of Traum (1994, 1997) treats obligations as primitives within the ontology, and takes as a definition of collaboration the notion of agents executing a multi-agent plan. In addition to the conditions of agents intending to do sub-actions, and being committed to the performance of the actions of others, the agents were also obliged to the others to perform their parts. Thus, even if the intentions and personal commitments were dropped, the obligations would remain for the agent to deal with. While this formulation did more directly represent the responsibility of an agent to do its part, it did not include other broader obligations to engage in ethical consideration or obligations deriving from roles within an activity (as described in Section 6), that would be part of ideal cooperation. Likewise, trust was not explicitly modelled, but left to the reasoning about others. While one could trust on the basis of obligations as well as personal goals and intentions, there was no mechanism for including higher levels of trust.

13. Conclusions

In this paper, we have suggested that cooperation should be analysed as a matter of degree involving at least four features which can be seen as requirements building on each other. We have also suggested that even though the terms *cooperation*, *collaboration* and *coordination* overlap in everyday usage, there is some support for relating the three terms using the four features so that coordination is related to the primary requirement of cognitive consideration, collaboration to also having a joint purpose and ideal cooperation to ethical consideration and trust. Since ideal cooperation also presupposes the two former, it can, like the adjective "old" which also covers "young", be used as a cover term for all four requirements.

We exemplified the cooperation features by analysing two naturally occurring human–human dialogues. We also studied the basic assumptions of cooperation behind human–computer interaction, and showed how the activity-based communication analysis can be applied to these situations too by analysing a task-oriented dialogue between a human and a computer system. Furthermore, we compared the approach to other current frameworks of cooperation and collaboration.

Our main point is that more of the notions discussed in this paper (e.g. trust and obligation) should be brought into the design and reasoned behaviour of these systems. Directly reasoning about ethical concerns, obligations and trust in the manner suggested here may lead to both more comprehensive and direct abilities to engage in a range of

dialogue behaviour than trying to mimic this behaviour with more complex application of a smaller set of cognitive primitives. More work of the kind performed in e.g. Jokinen (1994), and Traum (1994), and in the PLUS project is clearly needed, to formalize these other ideas to a degree sufficient for computational systems to reason about. Such formalizations will also help to clarify some notions left implicit in previous formulations of ideal cooperation, e.g. Allwood (1976, 1995a), and thus lead to a more refined analysis of human dialogue, as well.

References

AHLSEN, E. (1995). *Pragmatics and aphasia.* Gothenburg Papers in Theoretical Linguistics, Vol. 77. Department of Linguistics, Göteborg University.

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

ALLWOOD, J. (1976). *Linguistic communication as action and cooperation.* Gothenburg Monographs in Linguistics, Vol. 2. Department of Linguistics, Göteborg University.

ALLWOOD, J. (1977a). A critical look at speech act theory. In Ö. DAHL, Ed. *Logic, Pragmatics and Grammar.* pp. 53–69. Lund: Studentlitteratur.

ALLWOOD, J. (1977b). Negation and the strength of presuppositions. In Ö. DAHL, Ed. *Logic, Pragmatics and Grammar.* pp. 11–52. Lund: Studentlitteratur.

ALLWOOD, J. (1978). On the analysis of communicative action. In M. BRENNER, Ed. *The Structure of Action*, pp. 168–191. Oxford: Basil Blackwell.

ALLWOOD, J. (1980). On power in communication. In J. ALLWOOD & LJUNG, Eds. *ALVAR—a festschrift to Alvar Ellegård*. SPELL I, Department of English, University of Stockholm.

ALLWOOD, J. (1984). On relevance in spoken interaction. In G. KJELLMER & S. BÄCKMAN, Ed. *Papers on Language and Literature*, pp. 18–35. Acta Universitatis Gothoburgensis.

ALLWOOD, J. (1987). A semantic analysis of understanding. In V. ROSEN, Ed. *Papers from the 10th Scandinvian Conference of Linguistics.* Department of Linguistics and Phonetics, University of Bergen, pp. 37–51.

ALLWOOD, J. (1988). Om det svenska systemet för språklig återkoppling. In P. LINELL, V. ADELSWÄRD, T. NILSSON & P. A. PETTERSSON, Eds. *Svenskans Beskrivning 16*, Vol. 1. (SIC 21a). University of Linköping, Tema Kommunikation, pp. 89–106.

ALLWOOD, J. (1992). *On dialogue cohesion.* Gothenburg Papers in Theoretical Linguistics, Vol. 65. Department of Linguistics, Göteborg University. *Proceedings of the 13th Scandinavian Conference of Linguistics*. Roskilde.

ALLWOOD, J. (1994). Obligations and options in dialogue. *Think Quartely*, **3,** 9–18.

ALLWOOD, J. (1995a). *An activity-based approach to pragmatics.* Gothenburg Papers in Theoretical Linguistics, Vol. 76. Department of Linguistics, Göteborg University.

ALLWOOD, J. (1995b). Reasons for management in dialog. In R. J. BEUN, M. BAKER & M. REINER, Eds. *Dialogue and Instruction*, pp. 241–250. Berlin: Springer-Verlag.

ALLWOOD, J. (1995c). Language communication and social activity. In K. JUNEFELT, Ed. *Special Session on Activity Theory. Proceedings of the XIVth Scandinavian Conference of Linguistics and the VIIIth Conference of Nordic and General Linguistics.* Gothenburg Papers in Theoretical Linguistics, Vol. 73, pp. 17–39. Department of Linguistics, Göteborg University.

ALLWOOD, J. (1997). Dialog as collective thinking. In P. PYLKKANEN, P. PYLKKO & A. HAUTAMÄKI, Eds. *Brain, Mind and Physics.* Amsterdam: IOS Press.

ALLWOOD, J. & HAGLUND, B. (1991). *Communicative activity analysis of a Wizard of Oz-experiment.* Technical Report PLUS Working Paper, Department of Linguistics, University of Gothenburg.

ALLWOOD, J., NIVRE, J. & AHLSÉN, E. (1990). *Speech management: on the non-written life of speech.* Gothenburg Papers in Theoretical Linguistics, Vol. 58. Department of Linguistics, University of Göteborg. Also in *Nordic Journal of Linguistics*, **13.**

ALLWOOD, J., NIVRE, J. A& AHLSEN, E. (1992). The semantics and pragmatics of linguistic feedback. *Journal of Semantics*, **9**, 1–26.

AXELROD, R. (1984). *Evolution of Cooperation*. New York: Basic Books.

AUSTIN, J. A. (1962). *How to Do Things with Words*. Harvard University Press.

BALKANSKI, C. T. (1990). *Modeling act-type relations in collaborative activity*. Technical Report TR-23-90. Center for Research in Computing Technology, Harvard University.

BALLIM, A. & WILKS, Y. (1991). *Artificial Believers*. Hillsdale, NJ: Lawrence Erlbaum Associates.

BLACK, W. J., BUNT, H. C., DOLS, F., DONZELLA, C., FERRARI, G., HAIDAN, R., IMLAH, W. G., JOKINEN, K., LAGER, T., LANCEL, J.-M., NIVRE, J., SABAH, G. & WACHTEL, T. (1991). A pragmatics-based language understanding system. Esprit Conference paper.

BRATMAN, M. (1987). *Intentions, Plans, and Practical Reason*. Harvard University Press.

BUNT, H. (1990). DIT—dynamic interpretation in text and dialogue. In L. KALMAN & L. POLOS, Eds. *Papers from the Second Symposium on Languages and Logic*. Budapest: Akademiai Kiado.

CHU-CARROLL, J. (1996). Response generation in collaborative dialogue interactions. In S. LUPER-FOY, A. NIJHOLT & G. VELDHUIJZEN VAN ZANTEN, Eds. *Dialogue Management in Natural Language Processing Systems. Proceedings of the* 11*th Twente Workshop on Language Technology*, pp. 177–186. Twente.

CHU-CARROLL, J. & CARBERRY, S. (1998). Collaborative response generation in planning dialogs. *Computational Linguistics*, **24**, 355–400.

CHU-CARROLL, J. & CARBERRY, S. (2000). Collaborative response generation in planning dialogs. *International Journal of Human-Computer Studies*, **53**, 969–1015.

CLARK, H. H. & MARSHALL C. R. (1981). Definite reference and mutual knowledge. In A. K. JOSHI, B. L. WEBBER & I. A. SAG, Eds. *Elements of Discourse Understanding*. Cambridge: Cambridge University Press.

CLARK, H. H. & SCHAEFER, E. F. (1989). Contributing to discourse. *Cognitive Science*, **13**, 259–294.

CLARK, H. & WILKES-GIBBS, D. (1986). Referring as a collaborative process. *Cognition*, **22**, 1–39.

COHEN, P. R. & LEVESQUE, H. J. (1990*a*). Persistence, intention, and commitment. In P. R. COHEN, J. MORGAN & M. E. POLLACK, Eds. *Intentions in Communication*, pp. 33–69. Cambridge, MA: MIT Press.

COHEN, P. R. & LEVESQUE, H. J. (1990b). Rational interaction as the basis for communication. In P. R. COHEN, J. MORGAN & M. E. POLLACK, Eds. *Intentions in Communication*, pp. 221–255. Cambridge, MA: MIT Press.

COHEN, P. R. & LEVESQUE, H. J. (1991). Teamwork. *Noûs*, **25**, 487–512.

COHEN, P. R. & PERRAULT, C. R. (1979). Elements of a plan-based theory of speech acts. *Cognitive Science*, **3**, 177–212.

CUNNINGHAM, H., BLACK, W. J., UNDERWOOD, N. & JOKINEN, K. (1993). Dialogue management in a pragmatics-based natural language information system. In T. MCENERY C. D. PAICE, Eds. *Proceedings of the British computer Society 14th Information Retrieval Colloquium*. London: Springer-Verlag.

FERGUSON, G. & ALLEN, J. F. (1998). TRIPS: An intelligent integrated problem-solving assistant. *Proceedings of the 15th National Conference on Artificial Intelligence* (*AAAI-98*), Madison, WI, 26–30 July, 1998, pp. 567–573.

GALLIERS, J. R. (1989). *A theoretical framework for computer models of cooperative dialogue, acknowledging multi-agent conflict*. Technical Report 172, Computer Laboratory, University of Cambridge.

GMYTRASIEWICZ, P. J. (1995). On rational reasoning about other agents. In M. FEHLING, Eds. *Rational Agency*: *Concepts*, *Theories*, *Models*, *and Applications. Proceedings of the AAAI Fall Symposium*, pp. 68–74. Boston: MIT.

GRICE, H. P. (1975). Logic and conversation. In P. COLE & J. L. MORGAN, Eds. *Syntax and Semantics*, *Vol. 3*: *Speech Acts*, pp. 43–58. New York: Academic Press.

GROSZ, B. J. & KRAUS, S. (1993). Collaborative plans for group activities. In *Proceedings IJCAI-93*, pp. 367–373.

GROSZ, B. J. & KRAUS, S. (1996). Collaborative plans for complex group action. *Artificial Intelligence*, **86**, 269–357. Also appears as Technical Report TR-20-95, Center for Research in Computing Technology, Harvard University.

GROSZ, B. J. & KRAUS, S. (1999). The evolution of Shared Plans. In A. RAO & M. WOOLDRIDGE, Eds. *Foundations and Theories of Rational Agencies*, pp. 227–262. Dordrecht: Kluwer Academic Publishers.

GROSZ, B. J. & SIDNER, C. L. (1986). Attention, intention, and the structure of discourse. *Computational Linguistics*, **12**, 175–204.

GROSZ, B. J. & SIDNER, C. L. (1990). Plans for discourse. In P. R. COHEN, J. MORGAN & M. E. POLLACK, Eds. *Intentions in Communication*, pp. 417–444. Cambridge, MA: MIT Press.

JENNINGS, N. (1999). Agent based computing: promise and perils. *Proceedings of IJCAI-99*, pp. 1429–1436. Stockholm.

JOKINEN, K. (1994). *Response planning in information-seeking dialogues*. Ph.D. Thesis, University of Manchester Institute of Science and Technology.

JOKINEN, K. (1995). Rational agency. In M. FEHLING, Ed. *Rational Agency: Concepts, Theories, Models, and Applications. Proceedings of the AAAI Fall Symposium*, pp. 89–93. Boston: MIT.

JOKINEN, K. (1996a). Reasoning about coherent and cooperative system responses. In G. ADORNI and M. ZOCK, Eds. *Trends in Natural Language Generation*. An Artificial Intelligence Perspective, pp. 168–187. Berlin: Springer-Verlag.

JOKINEN, K. (1996b). Cooperative response planning in CDM. In S. LUPERFOY, and A. NIJHOLT and G. VELDHUIJZEN VAN ZANTEN, Eds. *Dialogue Management in Natural Language Processing Systems. Proceedings of the 11th Twente Workshop on Language Technology*, pp. 159–168. Twente.

JOKINEN, K. (1996c). Goal formulation based on communicative principles. *Proceedings of the 16th COLING*, pp. 598-603. Copenhagen.

JOKINEN, K. (1997). Ideal cooperation in dialogues. *Proceedings of the IJCAI '97 Workshop on Cooperation, Collaboration and Conflict in Dialogue Systems*, pp. 53–57. Nagoya.

KOWTKO, J. & ISARD, S. (1991). Conversational games within dialog. *Proceedings of the ESPRIT Workshop on discourse Coherence*, University of Edinburgh, 4–6 April.

KROPOTKIN, P. (1902). *Mutual Aid. London*: Heinemann.

LEVESQUE, H. J., COHEN, P. R. & NUNES, J. H. T. (1990). On acting together. *Proceedings of AAAI-90*, pp. 94–99. Boston, MA.

LOCHBAUM, K. (1994). *Using collaborative plans to model the intentional structure of discourse*. Ph.D. Thesis, Harvard University. Available as Tech Report TR-25-94.

LOCHBAUM, K. (1998). A collaborative planning model of intentional structure. *Computational Linguistics*, **24**, 525–572.

MCCARTHY, J. (1990). Ascribing mental qualities to machines. In J. MCCARTHY, Ed. *Formalizing Common Sense*. Ablex.

MCROY, S. (1995). Misunderstanding and the negotiation of meaning using abduction. *Knowledge-Based Systems*, **8**, 2–3.

MCROY, S. W. & HIRST, G. (1995). The repair of speech act misunderstandings by abductive inference. *Computational Linguistics*, **21**, 5–478.

MOESCHLER, J. (1989). *Modélisation du dialog*. Paris: Hermès.

POLLACK, M. (1990). Plans as complex mental attitudes. In P. R. COHEN, J. MORGAN & M. E. POLLACK, Eds. *Intentions in Communication*, pp. 77–103. Cambridge, MA: MIT Press.

RAND, A. (1964). *On the Virtue of Selfishness, a New Concept of Egoism*. New American Library.

ROSCHELLE, J. & TEASLEY, S. (1995). The construction of shared knowledge in collaborative problem solving. In C. O'MALLEY, Ed. *Computer-Supported Collaborative Learning*. Berlin: Springer-Verlag.

SADEK, M. D. (1991). Dialogue acts are rational plans. *Proceedings of the ESCA/ETR Workshop on Multi-modal Dialogue*.

SADEK, M. D. (1994). Communication theory = rationality principles + communicative act models. *Proceedings of the AAAI Workshop on Planning for Interagent Communication*.

SEARLE, J. R. (1969). *Speech Acts*. Cambridge: Cambridge University Press.

SEARLE, J. R. (1990). Collective intentions and actions. In P. R. COHEN, J. MORGAN & M. E. POLLACK, Eds. *Intentions in Communication*, pp. 401–415. Cambridge, MA: MIT Press.

SINCLAIR, J. & COULTHARD, R. M. (1975). *Towards and Analysis of Discourse. The English Used by Teachers and Pupils*. London: Oxford University Press.

SCHEGLOFF, E. & SACKS, H. (1973). Opening up closings. *Semiotica*, **7,** 289–327.

SPERBER, D. & WILSON, D. (1986). *Relevance*. Oxford: Basil Blackwell.

STRAWSON, P. F. (1950). On referring. *Mind*, **59,** 320–344.

STRAWSON, P. F. (1964). Identifying reference and truth values. *Theoria*, **XXX,** 96–118.

TRAUM, D. R. (1994). *A computational theory of grounding in natural language conversation.* Ph.D. Thesis, Department of Computer Science, University of Rochester. Also available as TR 545, Department of Computer Science, University of Rochester.

TRAUM, D. R. (1996). Conversational agency: the TRAINS-96 Dialogue Manager. In S. LUPER-FOY, A. NIJHOLT & G. VELDHUIJZEN VAN ZANTEN, Eds. *Dialogue Management in Natural Language Processing Systems. Proceedings of the 11th Twente Workshop on Language Technology*, pp. 1–11. Twente.

TRAUM, D. R. (1997). A reactive-deliberative model of dialogue agency. In. P. MÜLLER, M. J. WOOLDRIDGE & N. R. JENNINGS, Eds. *Intelligent Agents III-Proceedings of the 3rd International Workshop on Agent Theories, Architectures, and Languages (ATAL-96)*, Lecture Notes in Artificial Intelligence. Heidelberg: Springer-Verlag.

TRAUM, D. R. & ALLEN, J. F. (1991). Causative forces in multi-agent planning. In. Y. DEMAZEAU & J. P. MULLER, Eds. *Decentralized A.I.2.*, pp. 89–105. Amsterdam: Elsevier Science Publishers B. V.

TRAUM, D. R. & ALLEN, J. F. (1992). A speech acts approach to grounding in conversation. *Proceedings* 2nd *International Conference on Spoken Language Processing* (ICSLP-92).

TRAUM, D. R. & ALLEN, J. F. (1994). Discourse obligations in dialogue processing. *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, pp. 1–8.