Commitment-based and dialogue-game based protocols —new trends in agent communication languages

N. Maudet
IRIT, Université Paul Sabatier,
Toulouse, France

B. Chaib-draa
Dépt. Informatique, Pav. Pouliot,
Université Laval, Ste-Foy, PQ, Canada

July 4, 2002

Abstract

This survey introduces existing approaches to agent communications languages (ACLs) and particularly, conversation policies (CPs) which can be viewed as general constraints on the sequence of semantically coherent messages leading to a goal. Then limitations of these CPs are discussed in detail, particularly limitations on flexibility and specification. Finally, ACLs are viewed from the dialectic point of view, and some approaches are introduced in this context: some focusing on commitment-based protocols and others on dialogue-game based protocols.

1 Introduction

In recent years the interest in multi-agent systems (MAS) has grown tremendously, and today multi-agent technology is being used in a large range of important industrial application areas (Moulin and Chaib-draa, 1996). These applications range from information management through industrial process control to electronic commerce. The principle is that agents can function more effectively in groups that are characterized by cooperation and division of labor. Agent programs are designed to autonomously collaborate with each other in order to satisfy both their internal goals and the shared external demands generated by virtue of their participation in agent societies. All these applications have one thing in common. Agents must be able to communicate to each other to decide what action to take and how this action can be coordinated with others' actions. Stemming from the need of information and knowledge sharing among agents in distributed computing environments (beyound mere remote procedure

calls), an agent communication language (ACL) is an important part of a suitable framework that allows heterogeneous agents to interact, to communicate with meaningful statements that convey information about their environment or knowledge.

This survey introduces existing approaches to agent communications languages (ACLs) and particularly, conversation policies (CPs) which can be viewed as general constraints on the sequence of semantically coherent messages leading to a goal. Then limitations of these CPs are discussed in detail, particularly limitations on flexibility and specification. Finally, the paper argues for an approach to ACLs based on dialectic and to see a protocol as a "web" of commitments or as a dialogue-game.

The rest of this paper is structured as follows: Section 2 is devoted to an introduction to agent communication languages, with a particular focus on KQML and FIPA-ACL. The demand for conversation policies is presented and discussed in section 3. Definitions, objectives and limitations for these conversations are clearly explained in section 3 and more specifically in sub-section 3.2. This is essential to understand the motivations of the section 4 which aims at furnishing the basic material of formal dialectic. In this section precisely, we concentrate on notions of social commitments and dialectical systems. The two following sections overview the different approaches inspired by these notions, that we conceptually separate in two: commitment-based protocols (section 5) and dialogue-game based protocols (section 6). The final discussion evaluates the current advances w.r.t. the limitations of the traditional conversation policies, and raises some open questions.

2 Agent communication languages

In open multi-agent systems, the possible heterogeneity of the agents requires the use of a standardised ACL. That is why an important part of the multi-agent community currently works at defining such a standard.

2.1 From communicating agents...

The two main agent communication languages (ACLs) —both from theoretical and practical use— are KQML (Finin et al., 1997) and FIPA-ACL (FIPA, 2001). Both borrow from traditional speech act theory (Austin, 1962; Searle, 1969): messages of the agents are considered as actions with consequences on the environment.

KQML was the first to be developed, in the context of DARPA research. A KQML message is conceptually divided into three levels (Labrou et al., 1999): (1) the **communication** level which specifies the sender and receiver agents; (2) the **message** level which mainly specifies the type of message (or performative): affirmation, question,.., but also the knowledge representation language[1] and the

[1] Several candidates for such logical languages exist, for instance PROLOG, or KIF, *Knowledge Interchange Format*, which is proposed as a standard.

used ontology and finally; (3) the **content** level, which specifies the message content. The following example illustrates a KQML message where $Agent_1$ "informs" $Agent_2$ that the price of Zidane (well known soccer player) is 542, using the Prolog language, and a soccer-specific ontology (which specifies, for instance, that the price unit is in millions of euros...)

```
(tell
     :sender    Agent1
     :receiver  Agent2
     :language  Prolog
     :ontology  Soccer
     :content   price (zidane 542))
```

As we see, this defines the syntax of the ACL, but nothing is said about the meaning (or the semantics) of KQML messages. In fact, KQML was first proposed without any precise semantics, and this gave rise to several critics, see *e.g.* (Cohen and Levesque, 1995).

More recently another effort to come to a standard ACL has started through the Foundation for Intelligent Physical Agents (FIPA) initiative. This foundation is a nonprofit association whose objective consists in promoting the success of emerging agent-based technology. It operates through the open international collaboration of companies and universities active members in the field. FIPA assigns tasks (ontologies, semantics, architectures, gateways, compliance) to technical committees, each of which has primary responsibility for producing, maintaining, and updating the specifications applicable to its tasks. FIPA has recently proposed a new ACL, called FIPA-ACL, having a precise semantics based on a formal language called SL. This formal language is a quantified, multimodal logic with modal operators for beliefs (B), desires (D), uncertain beliefs (U) and persistent goals (PG). In this case, the semantics in each "communicative act" (CA) is specified as a set of SL formulae that describe the act's feasibility conditions and its rational effects. The feasibility conditions describe the necessary conditions for the sender of the CA, whereas the rational effects describe the effects that an agent can expect to occur as result of the CA. Figure (1) illustrates for instance the semantics proposed by FIPA-ACL for "inform" (FIPA, 2001), which states that the informing agent believes the proposition p, believes that the partner does not already know whether p (he believes that the feasibility conditions are fulfilled), and has the intention that its partner eventually believes p (he has the intention to fullfill the rational effects).

$\langle i, inform(j,p)\rangle$	p propositional formula.
Feasibility conditions	$B_i p \wedge \neg B_i(Bif_j p \vee Uif_j p)$
Rational effects	$B_j p$

Figure 1: FIPA semantics for the inform act.

Recently, in the same vein, a semantics for KQML messages has been proposed in terms of pre, post, and completion conditions (Labrou and Finin, 1998). As this semantics, as well as the FIPA-ACL one, refer to the mental states of the agents, they are commonly called **mentalistic** (Singh, 1998).

2.2 ...to conversing agents

In current ACLs however, we cannot pretend that agents take part in a dialogue, simply by exchanging messages: not all conversations are coherent. How can agents build together such a complex object? In other words, what ability should we offer to software agents to allow them to actually dialogue?

The strictly mentalistic approach considers conversational structures as simple epiphenomena. The idea is merely that, given an ACL with a defined semantics, the structures of conversation will emerge from the succession of communicative acts, and especially from the consequences of these states on the participants' mental states. Take for instance the well-known question/answer adjacency pair structure: when $Agent_1$ asks $Agent_2$ a question about p, $Agent_2$ will recognize from $Agent_1$ the intention to know whether p is true. Given this, and assuming some cooperative behavior, $Agent_2$ will form the intention to know p himself, and in turn to inform his partner about this point. Thus, the occurrence of this adjacency pair seems to be correctly explained. If this approach is really attractive because of the great flexibility induced by the independence of a particular structural model, it soon appears to be problematic in the context of communicating software agents, both for practical and theoretical reasons.

>From a practical point of view, the semantics of communicative acts is *so* rich that it is far too complex to determine the possible answers by just inferring others' mental states, simply because there are too many semantically coherent dialogue continuations. It appears therefore that the necessity to restrict the space of possible answers to a given act in a context of communication is generally a hard problem. This can lead to a great deal, and therefore to an extensive deductive machinery, from the agents that must interpret context, relevance, etc. in order to choose the "right" continuation.

>From a theoretical point a view, mentalistic approaches assume two main controversial hypotheses:

1. Agents' mental states are "verifiable".
2. Agents are sincere.

Both hypotheses are problematic, especially in an open environment. For the first point, we cannot be sure what is the mental state of any other agent and how to accede it (since agents are generally heterogenous (Singh, 1998) in open environments). Whether an agent indeed conforms to such semantics (believes or not what he says, for instance) is not verifiable —in the sense that it cannot be determined by an independant observer (Wooldridge, 2000).

For the second point, it is obvious that the hypothesis of sincerity (agents believe what they say, intend what they promise, etc.) cannot hold in several dialogue contexts which are not fully cooperative —as for instance negotiation or persuasion.

Until recently ACL research issues have primarily related to the generation and interpretation of individual ACL messages. Nowadays researchers on ACLs try to address the gap between these individual messages and the extended message sequences, or *conversations*, that arise between agents. As part of its program code, every agent must implement tractable decision procedures that allow that agent to select and produce ACL messages that are appropriate to its intentions. This is not purely a problem of matching ACL semantics to agent intention: except in the most limited of agent systems, these decision procedures must also take into consideration the context of prior ACL messages and other agent events. Paradoxically, taking this context into account can actually *simplify* the computational complexity of ACL message selection for an agent. By engaging in preplanned or stereotypical conversations, much of the search space of possible agent responses can be eliminated, while still being consistent with the ACL semantics. The specification of these conversations is accomplished via **conversation policies** (CPs).

3 Conversation policies

Conversation policies (CPs) are *"general constraints on the sequences of semantically coherent messages leading to a goal"* (Greaves et al., 2000). Coherence of the dialogue is thus ensured by these constraints. This greatly facilitates the task of computing the possible answers to a given message. But the use of CPs is also theoretically interesting: the linkage between the ACL's semantics theory and its account of conversation is one of the overriding current questions in the ACL field. On the one hand, as we explained above, it seems obvious that large-scale properties of agent conversations, such as overall information flow and the establishment of commitments, are a consequence of the individual meanings of the messages that make up the conversation. In this view, the ACL semantics is primary, and every conversational property logically derives from the composition of some collection of semantic properties of the individual messages and their sequence. On the other hand, there is a significant thread of research that takes conversational sequences themselves to be semantically primitive, and the precise meaning of the individual messages is nuanced by their role in the overall conversation. Basically, the idea is to consider the meaning of the communicative act at an objective level, as given by the possible answers to a given message. In this view, the conversational semantics are primary, and because of the dependence of ACL semantics on context, the same message might have slightly different meanings when used in the context of different agent conversations. In the context of this view precisely, an ACL is viewed as a sort of conversation between software agents and not as a set of speech acts. Its semantics is the semantics of a conversation and it cannot be reduced to

the conjunction or composition of semantics of its speech acts (Vongkasem and Chaib-draa, 2000).

3.1 What are conversation policies?

The definition of (Greaves et al., 2000) voluntarily abstracts from any precise computational model. Usually, we model CPs as finite state machines (FSMs). It is also common to name these models *protocols*, even though the term is often used in a generic way, as a synonym of CPs —a usage that we follow in this paper. States of the automaton maps the possible state of the conversation after a given message by the participants. Carefully designed and highly complex conversation policies have been proposed using these techniques in the literature, and implemented in real applications, see for instance COOL (Barbuceanu and Fox, 1995). Winograd and Flores's example (see Fig. 2) "request for action" illustrates this (Winograd and Flores, 1986).

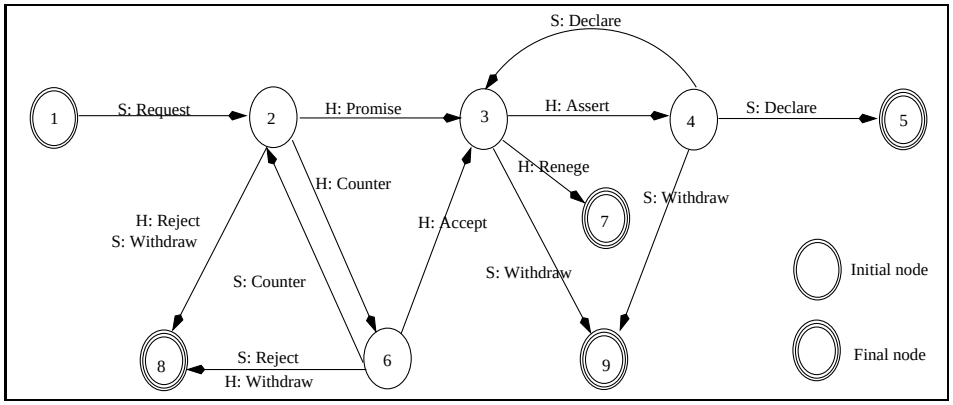

Figure 2: The request for action protocol.

Let us briefly describe the dialogue behavior expected with such a protocol. Conversation begins in the initial state (1), by the request from speaker S. In state (2), the dialogue can successfully be followed by the promise from H to realize the requested action, or come into a "negotiation cycle" with a counter-proposal from H, or fail with a reject leading to state (8). At state (3), the addressee will signal that the task has been achieved, or eventually decide to renege (leading to the final state (7)) and S will in turn positively (state(5)) or negatively evaluate this (state(3)). Note that the protocol says nothing about the content of the communicative acts.

Notice that other formalisms can be used to model conversation policies: notably Petri nets, possibly colored (Ferber, 1995; Cost et al., 2000; Lin et al., 2000), particularly well-suited to parallelized conversations (with more than two participants in conversation); and Dooley graphs (Parunak, 1996), which may offer a compact and precise representation of the conversation.

3.2 Limitations of conversation policies

Let us reconsider as an illustration the previous request for action protocol proposed by Winograd and Flores. Remarks about this protocol are manyfold. To start with, as explained above, the coherence of the conversation is ensured by the constraints imposed by the protocol on all participants —messages not expected in the protocol will simply not be considered. Also, when considered carefully, the protocol seems to be composed of different "phases" (or small protocols), not identified as such: firstly, the agents will try to negotiate a task for H to do. Next S and H will discuss the correct achievement of this task. These phases or small protocols are not specific to the particular case of the request for action. In addition, we have no information on how agents have agreed to use such a protocol.

In light of the above considerations and also of the critics of (Greaves et al., 2000; Vongkasem and Chaib-draa, 2000; Singh, 1998; Guerin and Pitt, 2001), we can identify two main issues that the forthcoming generations of CPs must address: flexibility and specification.

Flexibility The aim of conversation policies is basically to constrain the conversational behavior of the participants, but there is a delicate equilibrium to be found between this normative aspect and the flexibility expected in most multi-agent communications. Different points may participate in reaching this objective:

1. Adopt a formalism which allows more flexibility than FSMs, being more dependant on the state of the conversation than on the previous messages (Singh, 2000).
2. Consider unexpected (or exceptional) messages within the CPs (Singh, 1998). The objective is here to give appropriate follow-ups to such messages.
3. Prefer various small CPs, ideally those that we can compose, than a single large one (Greaves et al., 2000; Singh, 2000). The fact that we find recurrent similar structures of dialogue in various CPs is clearly a strong argument for this point. This possibilities of composition should be clearly defined.
4. Study how agents come to an agreement on the CP currently used — this point is crucial if the precedent objective is achieved (Vongkasem and Chaib-draa, 2000).

Specification The specification of the CPs is the second important challenge that we identify here. Indeed, specifications of the CPs often require ad-hoc formalism and are only semi-formally stated. This does not allow really taking all the profit from the specification, or formally verifing some expected properties of the model In general, the objectives are the following:

1. Specify CPs at a high level of abstraction. Such specifications should be relatively independent of the specificities of the agents involved in communication, particularly the private mental states of these agents (Flores and Kremer, 2002; Greaves et al., 2000; Pitt and Mamdani, 2000; Singh, 1998; Singh, 2000).
2. Adopt a declarative approach in order to explicitly declare the rules composing the CP (Singh, 2000). Clarity and expressibility follow from this point.
3. Provide formal properties of the CPs proposed. In particular, the problem of the termination of the CP is highly relevant.
4. Try to optimise the CPs. A precise specification might indeed identifying shortcuts that might take the dialogue participants in the CP without modifying the meaning of the interaction (Singh, 2000).

Finally, the use of conversation policies to guide agent communicative behavior engenders a host of practical questions. How should conversation policies be implemented in agent systems? Should CPs be downloaded from a common library, prebuilt into the agent's program, or derived from conversational axioms at runtime? How can conversation policies be negotiated, and unfamiliar policies learned? And finally, how can conversation policies be integrated with the other policies, plans, and rules which define an agent's behavior?

There is an alternative to the mental agency and to conversation policies: the social agency. This promising approach considers communicative acts as part of ongoing social interaction. In this case, even if we can't determine whether agents have a specific mental states, we are sure that communicating agents follow some social laws that sustain conversations. In this specific context of social agency, agent designers have usually assumed that the networks of obligation and power relationships that characterize human social behavior are not relevant to multi- agent systems. In practice, however, idiosyncratic social conventions have ended up being embedded in agent architectures and interaction protocols, with the result that different agent systems exhibit significant incompatibilities in this area. More research is needed into characterizing these fundamental communicative concepts in a multi-agent systems context. This includes concepts such as "commitment," "obligation," "convention," "power" (in the sense of hierarchical relations), and so forth (Kone et al., 2000; Moulin, 1983). Once these concepts are clarified, it then becomes possible to build a unified ACL semantics and pragmatics that takes account of these concepts.

Recently, researchers have begun to address the issues raised by conversations in general and by conversation policies in the context of ACL. Most of them have been inspired by the works of formal dialecticians (Hamblin, 1970; Walton and Krabbe, 1995). This research really now forms a field of dialectical models of interaction, as suggested by some authors (Gordon, 1996; Reed, 1998). Before turning to these models, we briefly present in the next section the basics of formal dialectic.

4 Basics of formal dialectic

Dialectic is the field of research concerned with the study of the interactive process of argumentation, *i.e.* the study of the dialectical contexts within which arguments are put forward (Hamblin, 1970). In particular, dialectic studies the different kinds of fallacies involved in argumentation —and formal dialectic has introduced formal notions and tools to deal with it. Among these various notions introduced or discussed by formal dialecticians, two have proved to be of first importance for those concerned with communication in multi-agent systems: **social commitments** and **dialectical systems**.

4.1 Social commitments

First of all, social commitment should not be confused with psychological commitment, which commonly captures some persistence of intentions —notably in the theory of rational interaction described in (Bratman, 1987; Cohen and Levesque, 1990). Crucially, social commitments are commitments that *binds* a speaker to a community. This motivates of course a distinction between the creditor and the debtor of the commitment. Notice that this notion is also clearly different from that of belief or intention and, more generally, social commitments are distincts of the *private* states of the agents.

Amongst social commitments, a classical distinction is also established between propositional or action commitments, see *e.g.* (Jennings, 1993). Thus, in the following examples (where A stands for the speaker),

(1) A: Ottawa is the capital of Canada.

(2) A: I will visit Ottawa next summer.

we would say that (1) is a propositional commitment which bids the agent A towards the audience (let us say B), while (2) is an action commitment which bids A towards B. Before concluding that A believes what he says, we can express that A is committed to the audience, and there are consequences related to such commitment. In particular, the audience will certainly penalize A if he makes an ulterior statement contradictory to p (meaning *Capital(Ottawa,Canada)*). In this case for example, the audience might consider A as a non-credible agent.

Following (Walton and Krabbe, 1995), we do not regard this distinction as relevant here: by simply uttering (1), A is committed in a way that constrains its subsequent actions. Precisely, depending upon context, A may then become committed to a number of things, for example, holding that p, defending that p (if challenged), not denying that p, giving evidence that p, arguing that p, proving or establishing that p and so on. Ultimately, A becomes committed to some (sets of) partial strategies. But the reason that we speak of *propositional* commitment here is that all A's commitments (as defined by the strategies he is committed to) center on the proposition p. We shall say in this case that A *is committed to* p, meaning that A is committed to a set of **partial strategies** centering on p (Walton and Krabbe, 1995).

4.2 Dialectical systems

We can trace the notion of dialogue game to the philosophical tradition of (Aristotle, 1928) —a tradition pursued by several philosophers in the Middle Ages, and, more recently, by the work of Paul Lorenzen and his colleagues (Lorenzen and Lorenz, 1978) or the Pragma-Dialectic school of (van Eemeren and Grootendorst, 1992). In formal dialectic, as explained before, the objective is to study fallacious reasoning or argumentation. The work of (Hamblin, 1970) has proposed a popular notion of **dialectical systems** as normative models of persuasive dialogue which mainly consists of (i) a set a moves —*e.g.* challenge, assertion, question; (ii) one commitment store (CS) for each conversant; (iii) a set of dialogue rules regulating the moves; (iv) a set of commitment rules defining the effect of the moves on the commitment stores. The dialogue correctly proceeds when the participants in conversation conform to the dialogue rules, and eventually ends when the proponent has withdrawn his thesis, or (symmetrically) that the opposant has conceded the proponent's claim. These termination conditions might be different for other dialogue types, as proposed by (Walton and Krabbe, 1995) and discussed further below. The commitment stores of the participants thus reflect the state of the dialogue (choose your own favorite term: dialogue/conversational/... score/record/context/profile/...).

Table 1 reflects how the commitments stores (CS) evolve for each participant (CS_P for P and CS_C for C) during the piece of dialogue (3), in a dialogue game style proposed by (MacKenzie, 1979). $\neg u$ stands for "the Kyoto Protocol is useless", m stands for "majority is required", and $\neg a$ for "most countries do not agree to ratify it". Note in particular, that challenged propositions are stored in this model, in order to avoid some circular dialogues considered as fallacies. This prevents for instance player P from answering C_4 by asserting that the protocol is useless, a fact still challenged at this point of the discussion.

(3) P_1: The Kyoto protocol is useless.
C_2: Why?
P_3: Because it needs to be widely approved to be useful and most countries are reluctant to ratify it.
C_4: Why are these countries reluctant to ratify it?

Turn	**Player**	**Move**	CS_P	CS_C
1	P	assert($\neg u$)	$\neg u$	$\neg u$
2	C	challenge($\neg u$)	$\neg u$	$?\neg u$
3	P	assert($m \wedge \neg a$)	$\neg u, m \wedge \neg a$	$?\neg u, m \wedge \neg a$
			$m \wedge \neg a \rightarrow \neg u$	$m \wedge \neg a \rightarrow \neg u$
4	C	challenge(m)	$\neg u, m \wedge \neg a$	$?\neg u, ?m, \neg a$
			$m \wedge \neg a \rightarrow \neg u$	$m \wedge \neg a \rightarrow \neg u$

Table 1: How commitment stores evolve during dialogue (3).

At each turn of the dialogue, the set of permitted moves is given by the dialogue rules. These dialogue rules depend on the previous moves of the conversation, or on the state of the conversation as detailed in the participants' CS.

4.2.1 Dialogue types

The hypothesis clearly stated in (Walton and Krabbe, 1995) is that dialectical systems can model various types of dialogue. The authors identify six types of dialogue (not exhaustively). This classification is based upon three factors: the information available to the participants, the goal of the dialogue itself, and the individual goals of the participants. Table 2 summarizes these types, and shows the goals (of the dialogue and the participants).

Type of dialogue	**Goal of the dialogue**	**Initial situation**
Persuasion	resolution of conflict	conflicting point of view
Negotiation	making a deal	conflict of interest
Deliberation	reach a decision	need for action
Information-seeking	spreading knowledge	personal ignorance
Inquiry	growth of knowledge	general ignorance
Eristic	accommodation in relationship	antagonism

Table 2: The six primary types of dialogue.

These six types may be refined in subtypes, simply by specifying more elaborate conditions on the dialogues (*e.g.* the type of conflict, the degree of rigidity of the rules). Thus for example, a dispute is a subtype of persuasion, where each participant tries to defend its own point of view. Very importantly, Walton and Krabbe argue that *"types of dialogue coincide with particular dialectical systems or dialogue games"* (Walton and Krabbe, 1995, p.67). In all their types of dialogue, both parties *ideally* have a recorded log of their individual commitments to any given point to which the dialogue has progressed: the commitment store (CS) as previously described.

4.2.2 Dialectical shifts

However, dialogues are usually not of a single type from beginning to end For instance, it is common to start an inquiry dialogue, to realize during the dialogue that there is a controversial issue at stake, to enter into a dispute sub-dialogue, and to eventually resume the inquiry dialogue when the issue has been resolved. This means that most dialogue types are complex (*i.e.* composed of different types). The notion of dialectical shift has been introduced for that aspect (Walton and Krabbe, 1995). A dialectical shift is a change in the context of dialogue during a conversation from one type of dialogue to another. The shift might be constructive and agreed to by all parties, and it is a licit shift. In this case, the second dialogue may even be functionally related to the goal of the original dialogue: it is an "embedded dialogue". For instance, considering the made-up

example (4), we could see it as a negotiation sub-dialogue embedded in a request for action dialogue.

(4)

S:	Will you ratify Kyoto?	begin Request for Action
H:	No, but we'll reduce our gas.	begin Negotiation
S:	Can you do 40% ?	
H:	No, 30%.	
S:	OK.	end Negotiation
H:	We will let you know when the objective is reached.	
S:	Thank you	end Request for Action

In other occasions, the shift might be concealed or inappropriate because it is an illicit type of shift. Illicit shifts in our context of dialogue based on argumentations are frequently associated with fallacies (a fallacy can be viewed as an argument which *appears to be correct but is not*). In some dialogues, we might have a cascading effect. In this case, there is a series of shifts from one type of dialogue to another, and then to yet another, and so forth. Thus, one can begin by a negotiation, and then shift to a critical discussion and then shift to a negotiation that might shift to a quarrel, etc.

4.3 Discussion

Formal dialectic has developed several carefully designed systems, especially for persuasion and critical discussion. Now, the wide use of the term dialogue games calls for some clarification: the objective of these models clearly varies in the different disciplines they are used In the philosophy of argumentation, they provide proof theories for formal logics. In computational linguistics, they are used to explain human dialogue and to generate artificial dialogues which humans can understand (Levin and Moore, 1980; Moore, 1993; Hulstijn, 2000b): games are thus based on some observations of the conventions that operate in actual discussions —they are first descriptive before being generative.

More significant for our purpose, within the field of multi-agent systems the models developed in formal dialectic are now used to define agent-to-agent conversation policies —and there is no reason that these games describe some natural conversations. The natural link expected with some underlying argument-based models of (non-monotonic) reasoning is also an attractive feature of this approach: agents are supposed to have some minimum argumentative capability (*e.g.* giving some reasons supporting a belief or an intention) (Parsons et al., 1998; Prakken, 2001; Prakken, 2000; Brewka, 2001). This leads to richer discussions than just exchanges of unsupported locutions, specially in the cases of persuasion or negotiation (Parsons and Jennings, 1996). For instance, it is possible for an agent to explain why he cannot accept a request of the other agent. Interestingly, current research in formal dialectic is in turn influenced by research in MAS. This interplay has already proven fruitful, and feeds a large amount of current literature in these fields (Reed et al., 2002).

More precisely, we argue that the influence of formal dialectic has in fact produced two sorts of approaches: **commitment-based protocols** and **dialogue-game based protocols**. Commitment-based protocols (Sect. 5) aims at defining semantics of the communicative acts in terms of public notions, *e.g.* social commitments. Dialogue-game protocols (Sect. 6), in addition, consider that protocols are captured within appropriate structures (games), and that these structures can be composed in different ways to form the global structure of dialogue.

5 Commitment-based protocols

The notion of social commitment has long been recognized as an important topic in the field of multiagent systems. Social commitments are in fact key notions to allow coordination and communication in multi-agents systems (Jennings, 1993). Nevertheless, the first attempts to use these notions to ground the communicative theories are more recent —and essentially motivated by the requirement of verifiability. This section surveys some representative approaches in these directions to date, discussing in particular how commitments and protocols interplay.

5.1 Singh

Singh was certainly the first to clearly emphasize the need to define the semantics of ACLs in terms of "social notions" (Singh, 1998). He proposed himself a **social semantics** for agent communication language in (Singh, 2000) based on Habermas' view. He precisely proposed to define three levels of semantics for each act, corresponding to three validity claims made with each communication acts: the objective claim (that the communication is true), the subjective claim (that the communication is sincere), the practical claim (that the communication is justified). For instance, by informing j that p, i gets committed towards j that p holds (objective claim), that he believes that p (subjective claim), and towards the whole group that he has reasons to believe p (practical claim). By doing so, Singh admits the mentalistic approach (at the subjective level), but embedded within a social attitude (the practical claim that leads to a commitment made by the speaker towards the audience). It is one thing to say, following the purely subjective school, that agent is sincere when he asserts the fact p. It is another thing to say that the agent is socially committed to be sincere about the fact p towards the audience, as proposed by Singh.

More technically, new modal operators are introduced, and the semantics of these commitments themselves is expressed using a branching-time logic based on the Computation Tree Logic (CTL) —quantifications over the different paths are needed to account for notions such as "p will eventually be true in some/all/the real path". Obviously, social semantics for the communicative acts does not replace the protocol as argued by Singh. Indeed, although the commitment-based semantics can tell us the result of composing some commu-

nications, it is the protocols that tell us what composition is more appropriate. It is therefore important to combine a commitment-based semantics with protocols for more flexible ACLs.

5.2 Amgoud *et al.*

Directly inspired by a dialectical system proposed by MacKenzie (MacKenzie, 1979), and following (Maudet and Evrard, 1998), Amgoud *et al.* (Amgoud et al., 2000a; Amgoud et al., 2000b) have first defined *dialogue rules* and *update rules* for the different types of locutions supported by their dialogue model. Dialogue rules clearly defines the protocol (sequences of moves allowed or not), while updates rules captures the effect of the communicative acts on the state of the dialogue (*i.e.* the commitment stores of the conversants). Indeed, the authors assume that each player has a knowledge base, accessible to both players, containing the commitments made in the dialogue[2]. These commitment stores are denoted CS_A and CS_B respectively. On the one hand, it is worth noting that the dialogue rules do not refer in any case to these commitment stores (this differs from (MacKenzie, 1979)): the flexibility of the dialogue is not enhanced. On the other hand, the model extends the set of communicative acts detailed in (MacKenzie, 1979) in order to capture the different types of dialogue introduced in (Walton and Krabbe, 1995).

The authors have detailed in (Amgoud et al., 2000a; Amgoud et al., 2000b) the cases of persuasion and of negotiation. The example of figure 3 shows how they instantiate a communicative act of the negotiation protocol: the "acceptance of a request". The update rule simply captures that the proposed commitment will be included in the acceptee commitment store in case of acceptance. This does not traduce however a general policy on commitment store: different

$\langle accept(p) \rangle$	p is a propositional formula.
dialogue	the other player can play any move, except a refusal.
update	$CS_A^i = CS_A^{i-1} \cup \{p\}$ and $CS_B^i = CS_B^{i-1}$

Figure 3: Acceptance of a request act.

policies are employed in the model. As the example above shows, some moves require acceptance of the partner before being included in the commitment store. The authors have however followed MacKenzie's *de facto* commitment idea in the case of assertions (meaning that you are committed on the content of the speaker's assertion, unless you explicitly retract it).

[2]The authors also define a subjective semantics for the locutions, based on a powerful reasoning capability offered by an argumentation system. Both players are indeed equipped with an argumentation system (AS), a sort of triple $\langle \mathcal{A}(\Delta), Undercut, Pref \rangle$ such that $\mathcal{A}(\Delta)$ is a set of the arguments built from Δ, $Undercut$ is a binary relation representing defeat relationship between arguments and, $Pref$ is a preordering on $\mathcal{A}(\Delta) \times \mathcal{A}(\Delta)$. In the context of the conversation, the set Δ is the union of the private database of the agent and of the commitment stores.

5.3 Colombetti

Colombetti has recently proposed ALBATROSS (*Agent Language Based on the Treatment of Social Semantics*) (Colombetti, 2000b; Colombetti, 2000a) as an approach in the same vein as Singh. He has in particular, introduced the notion of **pre-commitment** as a kind of conditional commitment (*e.g.* a request pre-commits the agent to which it is addressed, meaning that this agent will be committed in case of acceptance). Like Singh, his semantics of social commitment is also expressed in term of a temporal logic based on CTL*. In this logic, the author introduced a new type of expression ($\alpha, \beta, \ldots$), called *action formulas*, which describe action types. If i is an agent name and α is an action formula, an expression of the form $D_i\alpha$ is a state formula meaning that agent i (i.e., the actor) has just completed (in the sense of "done") the execution of a token α. Then, the author introduced the expressions $mc(i, j, \lambda)$ and $mp(i, j, \lambda)$ as action formulas meaning "make commitment" and "make pre-commitment" respectively. In this case, the author obtains the following axioms:

1. $D_i mc(i, j, \lambda) \rightarrow C_{ij}\lambda$
2. $D_i mp(j, i, \lambda) \rightarrow P_{ji}\lambda$

The first axiom means that the acceptance transforms a pre-commitment into commitment (C_{ij} is the modal operator for "commitment"). The second axioms means that a pre-commitment (as a commitment) should persist.

Now the question is how to define the semantics of messages in terms of Colombetti's logic and in particular in terms of *mc* and *mp*. In fact a speech acts as "assert" is expressed here as:

```
(assert
     :sender    i
     :receiver  j
     :content   Lambda)
```

sending such a message implies : $D_i mc(i, j, \lambda)$. In this way, the actions allowed like *mc* and *mp* can be performed indirectly, through messages that realize a predefined repertoire of speech act types. The other way proposed by Colombetti consists of giving agents a more direct access to "commitments". The idea is to allow an agent to make a commitment, a precommitment, etc. by a *declaration*. Thus, an agent may commit to the truth of statement by sending a message of the form:

```
(declare
     :sender    i
     :receiver  j
     :content   (make-commitment
                 :debtor       i
                 :creditor     j
                 :condition    Lambda))
```

Contrary to Singh's work, this approach does not necessarily need explicit additional protocols, since the protocol itself may be expressed of (possibly pre-commitments). This is the choice made in some cases, as the following rule shows:

$$askIf(x, y, \phi) = request(x, y, Done(y, assert(x, y, \phi)|assert(x, y, \neg\phi)/now+k))$$

This is elegant and leads to more flexibility, since agents are supposed to answer as expected by the commitments, but are not *obliged* to. However Colombetti, when he claims that *"the semantics of the message should be independent of the nature of conversation"*, seems reluctant to defend this as a general policy.

5.4 Flores and Kremer

Flores and Kremer (Flores and Kremer, 2002) have also recently proposed an ACL approach which aims at defining conversation protocols and speech acts semantics in a unifying model based on the notion of social commitment. Social commitments to actions, are here engagements to a course of action taken by an agent relative to another agent on whose behalf actions are done. Contrary to the approaches described so far, it is emphasized that these commitments are also the binding element of the conversations themselves: the notion of **conversational commitment** is precisely introduced to distinguish those commitments that entail the utterance of a dialogue act.

During conversation, each agent maintains a private database in which these shared commitments are simply added or deleted—similar to the commitment stores introduced in (Hamblin, 1970). Defining how these operations on social commitments are negotiated through the dialogue is devoted to the so-called **protocol for proposals**. Using dialogue acts such as *proposal, accept, reject* or *counter* (considered as an alternative proposal), agents jointly[3] decide what operations on commitments to apply on their databases. For instance, an agent may propose to an other agent to add a social commitment to its own database. When accepted, the corresponding operation is applied to the set of commitments of the proposed agent. Thus, *proposal* or *accept* are just generic types of moves, which can be split in different utterances, depending on the nature of the debtor, of the creditor, and of the operation on the commitment —see Table 3 for the utterances corresponding to the *propose* illocutionary point.

According to their approach and as it is shown in Table 3, *request, offer, release and discharge* are all based on the illocutionary point *propose*. Thus, a *request* is a proposal to adopt a social commitment for action in which the speaker is the creditor and the addressee is the debtor. Similarly, an *offer* is a proposal to adopt a commitment in which the speaker is the debtor and the addressee is the creditor, a *release* is a proposal to discharge a commitment in which the speaker is the creditor and the addressee the debtor, and finally, a

[3] Of note the fact that any operation is submitted to the acceptance of the partner. This defines very strong joint commitments, a point that may be discussed for some specific utterances.

Utterance	Illocutinary Point	Operation	Creditor	Debtor
Request	Propose	Add	Speaker	Addressee
Offer	Propose	Add	Addressee	Speaker
Release	Propose	Del	Speaker	Addressee
Discharge	Propose	Del	Addressee	Speaker

Table 3: Utterances with a "propose" illocutionary point.

discharge is a proposal to discharge a commitment in which the speaker is the debtor and the addressee the creditor.

Finally, the authors formally define a small set of fundamental policies to support the negotiation of shared commitments to action, thus allowing agents to advance the state of their joint activities. The model is exemplified through the famous *Contract Net Protocol*, which requires six different CPs. In some aspects, this approach is close to the dialogue-game approach. However, even though the use of different policies is clearly stated, these policies are not considered as structures, and no particular kind of composition is studied, except the sequential composition.

5.5 Guerin and Pitt

Also influenced by the work of (Hamblin, 1970), Guerin and Pitt share the view of an ACL based on the notion of social commitments (Guerin and Pitt, 2001). In some aspects, this work goes further that the approaches detailed so far.

Firstly, they use an extended version of the commitment stores which contains the (public) expressed mental attitudes (see also Singh) of the participants, conversational options and obligations (to perform subsequent dialogue acts) and the history of messages uttered. Also, the authors are very keen to precise that the commitment stores are part of a larger **social state** which may also contain persistant social relations (*e.g.* long term commitments).

More importantly, Guerin and Pitt adress the practical engineering problem of how to link the declarative statements representing ACL messages and the mechanism underlying the commitment stores —most of the approaches described so far leaved this issue unresolved. Basically, ACL messages are considered as declarative statements that are given a procedural interpretation by means of a **denotational semantics**. More precisely, an ACL specification has three parts (i) the converse function —gives conversational options and obligations w.r.t. the current state of the commitment store, (ii) the protocol semantics —gives the meaning of the dialogue act w.r.t. the current protocol, and (iii) the dialogue act semantics —gives the core meaning of the act The denotational semantics compiles such an ACL specification into a function from social state to social state. We do not introduce formal details here: what is important is that each agent equipped as described can convert any ACL message into an update function which can be applied to (a copy of) its social state.

6 Dialogue-game protocols

Basically, the dialogue-game approach considers CPs as abstract communication structures —structures that can be composed to reflect the whole dialogue. Convincing arguments for the use of such structures in the multi-agent context can also be found in (Reed, 1998). There are however, several differences between the current approaches based on dialogue-game based protocols and used in multiagent systems, when things are considered more carefully. In the following, we present these approaches while trying to answer the following questions:

- What kind of structure has the game? How are rules specified within the game?
- What kind of games' compositions are allowed? Particularly, how about embedding and sequencing as suggested by (Walton and Krabbe, 1995)?
- How are games grounded? In other words, how do conversants reach agreement on the current game, how are games opened or closed?

6.1 Reed

Reed has proposed the notion of **dialogue frames** (Reed, 1998) as abstract exchange structures directly based upon the work of (Walton and Krabbe, 1995) >From the dialogue types defined by these authors, he proposes the following set mapping the kind of dialogue type and the notions involved in the dialogue type.

$$\begin{aligned} D = \{ & < persuade, B >, \\ & < negotiate, C >, \\ & < inquire, B >, \\ & < infoseek, B >, \\ & < deliberate, P > \} \end{aligned}$$

Thus, persuasion, inquiry and info seek are concerned with beliefs (B), while negotiation is discussing contracts (C), and deliberation aims at determining a plan (P).

Structure: In Reed's approach, a dialogue frame is formally defined as a tuple with four elements

$$F = \langle \langle t, \Delta \rangle \in D, \tau \in \Delta, \{u^0_{x_0 \to y_0}, ... u^n_{x_n \to y_n}\} \rangle$$

where t is the type of this dialogue frame, τ is the *topic* of the dialogue frame, x_0, $y_0 \in A$ are the interlocutors and $u^j_{x_j \to y_j}$ refers to the jth utterance occurring in a dialogue between agents x_j and y_j, in which x_j is the originator of the utterance.

Dialogue frames are thus instances of a specific structure —of a particular type ($\langle t, \Delta \rangle \in D$) and about a specific topic ($\tau \in \Delta$). For instance, a negotiation

about the fact that the price of Zidane is 502, or a persuasion about the fact the this player is a good player. The frame specifies a protocol, simply considered as a sequence of utterances. Utterances made at the ith turn by speaker x towards y ($u^i_{x_i \to y_i}$) are supported statements (although the support may be empty, as we will see below). Turn-taking is just a matter of alternance. However, Reed does not inspect in detail the nature of the protocols actually used in the frames. Here a Reed's example of persuasion dialogue

$$u^0_{b\to a} : \langle\langle propose(persuade, has(c, information))\rangle, \{\emptyset\}\rangle$$
$$u^1_{a\to b} : \langle\langle accept(persuade, has(c, information))\rangle, \{\emptyset\}\rangle$$
$$u^2_{b\to a} : \langle\langle tell(has(c, information))\rangle, \{told\_by(has(c, information), d)\}\rangle$$
$$u^3_{a\to b} : \langle\langle tell(unreliable(d))\rangle, \{\emptyset\}\rangle$$
$$u^4_{a\to b} : \langle\langle tell(unreliable(d))\rangle, \{\emptyset\}\rangle$$
$$u^5_{b\to a} : \langle\langle concede(unknown(has(c, information)))\rangle, \{\emptyset\}\rangle$$

where b initiates a dialogue to *persuade* a that some third party, c, has information. b supports his claim by citing d as his source. a gives a counter-argument by pointing out the unreliability of d, and with no further supports available, b retracts his assertion with a *concede* which terminates the dialogue frame.

Grounding: As far as we know, Reed was the first to propose that these structures could be handled through meta-acts of communication (*propose*/*accept*). These moves, whose purpose is to open the frames, have an empty support. For instance, the move

$$u^0_{a\to b} : \langle\langle propose(negociate, \langle buy(b, Zidane), \{\langle price, 500\rangle\}\rangle), \{\emptyset\}\rangle$$

is well-formed, and captures a proposition from a to b to negotiate the price of Zidane, initially proposed at 500. Generally, the agent proposing the frame is the one beginning to play within the frame. Notice that no meta-move is required to quit the game. Instead, Reed assumes that moves of concession or acceptance by one of the agents automatically close the frame.

Composition: Reed considers two kinds of game composition, *sequencing* (canonical ordering) and *embedding*, and mainly concentrates on this latter case. Embedding is captured within the model without further complication of the structures. Indeed, since propositions to enter a frame are moves like any others (as we have seen above), they can be made within ongoing frames. When a new dialogue frame ϕ_1 is proposed at turn i by x, and accepted by y at $i+1$ while a frame ϕ_0 was open, he assumes that ϕ_0 is just suspended (ϕ_1 is then embedded in ϕ_0). When the frame terminates, ϕ_0 resumes where it was stopped. Generally, the speaker who concedes in the embedding frame is not the speaker who resumes in the embedding frame.

6.2 Dastani *et al.*

Dastani and his colleagues (Dastani et al., 2001; Dastani et al., 2000) have proposed a methodology for constructing flexible negotiation protocols based on dialogues games, following the work of (Hulstijn, 2000b). Although their negotiating agents may have competing interests, they share a common goal in order to coordinate their actions. To represent coordinated actions, the authors have used partial representations, *i.e.* recipes. A particular type of these recipes is formed by the so called dialogue games.

Even though this approach has focused on negotiation, the framework is supposed to be generic enough to accept different types of dialogue. Utterances are captured by dialogue acts, classically composed of a semantic content and of a communicative function. A conversational record or dialogue context ($CS_{a,b}$) keeps track of the statements and commitments made so far in the dialogue.

The communicative function of each dialogue act has a task-related and/or a interactive-related function, and this permits to define the key notion of **coherence**. An utterance or move in a negotiation dialogue is coherent with the dialogue context, if (i) it fits a plan that might achieve the apparent goals of the agent, and (ii) it fits the current interaction rules. Depending on the attitude of the agent and on who has the initiative, (i) or (ii) takes precedences. Hence, coherence is both a matter of task-level coherence and of interactive-level coherence. While task-level coherence may involve complex plan inference based on the datum available in the conversational record, interactive-level coherence is captured within pre-planned recipes for joint communicative actions (Hulstijn, 2000a), precisely the dialogue games introduced below.

Structure: At the interactive-level, a dialogue act is considered either initiative or reactive. The basic game structure, as shown in figure 4), simply captures that an initiative act may be followed by a positive, negative, or reformulating answer.

$exchange(a, b, \zeta) =$	$initiative(a, b, \eta);$ $[pos\_response(b, a, \zeta)$ $\mid neg\_response(b, a, \zeta)$ $\mid retry(a, b, \xi)],$ where $CS_{a,b} \models coherent(\eta, \zeta)$

Figure 4: The basic exchange structure.

Most exchanges have the form indicated in figure 4 where an exchange is allowed, given that the coherence constraint on the semantic contents of the initiative and response is met (noted $CS_{a,b} \models coherent(\eta, \zeta)$). In other words, the response must address the initiative. In case of a response, the content of the response, denoted by ζ above, is considered "grounded". For example,

the following example (taken from (Dastani et al., 2001)) shows an information seeking exchange.

$information\_seeking(a, b, \psi) =$	$question(a, b, ?\phi); (answer(b, a, \psi)$ $\mid clarification\_question(b, a, ?\chi))$, where $CS_{a,b} \not\models \neg\psi$ $(consistent)$ and $CS_{a,b} \not\models \psi$ $(informative)$ and $CS_{a,b} \models relevant(?\psi, \phi)$ and $CS_{a,b} \models licensed(?\psi, \phi)$.

This example shows that for an answer to a question, the authors require that it be consistent, informative, relevant to the question, and not over-informative with respect to the question.

Composition: Games can statistically be composed at the stage of definition by *sequencing* or *chaining*. Let us first consider the case of sequential combina-

$game(a, b, (\eta.\zeta)) =$	$exchange(a, b, \eta); game(a, b, \zeta);$ where $CS_{a,b} \models coherent(\eta, \zeta)$

Figure 5: Sequential combination of games.

tion (see Fig. 5). As expected, the result of the combined game depends upon the results of the composing games. The recursive nature of the definition indicates that it is possible to combine as many games as requested. Like in the basic exchange case, some coherence constraints are stated between the games' topic. For instance, to be combined, games have to share a common subject matter. Let us now turn to chaining. Constraints require the last dialogue act (reactive) of the first game being the first (initiative) of the second game. Canonical examples of such chaining structures are question/answer/evaluation or proposal/counter-proposal/... Games can also be dynamically composed during the dialogue. IR structures can be *embedded*, thus dialogue such as I(IR)R are well-formed. However, no constraints on such combinations are proposed.

Grounding: The authors mention that some negotiation phases may be useful to determine the current game, but no specific acts are proposed to account for this aspect.

6.3 McBurney and Parsons

Another approach explicitly using the game structure has recently been proposed by McBurney et Parsons (McBurney and Parsons, 2001a; McBurney and Parsons, 2002). Their ambition is to define a formalism for game representation allowing to represent the different types of dialogue of (Walton and Krabbe, 1995) and the meta-dialogues —*i.e.* dialogues about dialogue. To do this, they

propose a three layer model for agent dialogues: (i) the topic layer, (ii) the dialogue layer; and (iii) the control layer. The topic layer just defines the matters possibly discussed in the dialogue. Of main interest for us are the two remaining layers, which are concerned with games' structure, grounding and composition.

Structure: The dialogue games are defined as classical dialectical systems (see Sect. 4) with the following components: (i) beginning rules, (ii) locution rules, (iii) dialogue rules, (iv) update rules, and (v) ending rules. These rules are defined at the dialogue layer.

Grounding: The grounding of games is captured at the control layer, where a specific **control-level dialogue** is defined —and it is assumed that conversants consent to play this meta-dialogue. Practically, agents engaged in a game at the dialogue level can jump to this layer by simply playing a

$$PROPOSE\_RETURN\_CONTROL$$

move. Within the control layer, the control-level dialogue handles how agents jointly select the games currently opened/closed. To account for this, meta-dialogue moves are proposed

$$(BEGIN(G(p)) \text{ or } END(G(p))$$

where G are games' types and p topic of the game. The control layer also defines the legal combination of games, that is, the composition.

Composition	Property
n-Iteration	Repetition of n dialogues G: each occurence is undertaken until closure and followed immediately by the next occurence.
Sequencing	Undertaking G until its closure and then immediately undertaking H
Embedding	Undertaking G until a legal sequence of moves Φ has been executed, then switching to dialogue H which is undertaken until its closure, whereupon G resumes from where it was interrupted and continues until closure.
Parallelization	Undertaking both G and H simultaneously, until each ar closed
Test	Dialogue to assess the truth-status of p: returns a truth-value for p

Table 4: Dialogues' compositions in McBurney and Parson's approach.

Composition: McBurney and Parsons sensibly increase the kind of game combinations traditionally considered in literature. Beside sequencing and embedding, they study iteration, parallelization and test. Table 4 summarizes the different combinations considered. Note that some of these combinations, *e.g.* parallelization, do not correspond to actual natural dialogue behavior (McBurney and Parsons, 2001b), but may prove useful in the context of communication between software agents. We assume in the table that uppercase G and H denote dialogues, and p a well-formed formula. We also denote as active(s) the dialogue(s) opened at the lower levels of embedding. Embedding does not entail suspension of the embedded dialogues: some rules of interaction between rules and commitments are proposed (any conflicting rules from higher dialogues are over-ridden by those of the lowest open dialogue). This differs from Reed's dialogue frames where the embedding game is supposed to be suspended during the embedded game.

6.4 Maudet and Chaib-draa

The currently in progress agent communication language DIAGAL (*DIAlogue-Game based Agent Language*) adapts the work of (Maudet, 2001; Maudet and Evrard, 1998) to communication between software agents. The model makes a distinction between dialogue games and communication games. While dialogue games are temporarily activated during the dialogue for a specific goal, communication games capture more general communicative requirements supposed to be valid for any type of dialogue. The model has recently been used to capture and enrich the classical request for action protocol as described in Sect. 3.1, see (Chaib-draa et al., 2002).

Structure: Games are bilateral structures defined by (i) entry conditions (which must be fulfilled at the beginning of the game), (ii) exit conditions (defining the goals of the participants when engaged in the game), and (iii) dialogue rules. Typically, those dialogue rules can specify what dialogue a participant is commited to play, in case of occurence of a given dialogue act by the partner. Notice that all these notions, even dialogue rules, are defined in terms of (possibly conditional) commitments: the model adopts a strict commitment-based approach within game structure, close to that of (Flores and Kremer, 2002). Technically, games are then conceived as structures capturing the different commitments created during the dialogue.

Grounding: Games can have different status during a dialogue: they can be opened, closed, or simply proposed. Among the communication games is the **contextualization game** which aims precisely at defining how games are handled and how their status evolve through the dialogue —the role devoted to the control-level dialogue in (McBurney and Parsons, 2002). This approach also assumes the use of meta-acts of dialoguue to handle game structure. A

Move	Effect
$prop.in(x, j)$	$create(x, C(y, acc.in(y, j) \mid ref.in(y, j)))$
$acc.in(x, j)$	$create(x, C(\{y, x\}, j))$ $create(x, C(y, cont(y, j)))$
$cont(x, j)$	$create(x, C(y, cont(y, j) \mid prop.out(x, j)))$
$prop.out(x, j)$	$create(x, acc.out(y, j) \mid ref.out(y, j))$
$acc.out(x, j)$	$cancel(x, C(\{x, y\}, j))$

Table 5: The contextualisation game.

first account of this contextualization game is reported in the table[4] 5. The proposal to enter in the game j ($prop.in$) makes the game *proposed* and creates the commitment for the partner to accept ($acc.in$) or refuse to enter ($ref.in$) in j. The notion of counter-proposal does not explicitly appear in the game, but may be captured as a refuse and an alternative proposal. The entry in the game *opens* the game and causes the creation of the collection of commitments of the game (noted $C(\{y, x\}, j)$). As long as the game is *open*, it is possible to continue to play within the current game —it is assumed that any expected (*i.e.* satisfying a commitment) move in the context of the current game plays at the same time a continuation move in the contextualization game. Finally, it is possible to propose to quit the game. Accepting to quit the current game *closes* this game and cancels the remaining commitments of the game. However, this simple contextualization game needs to be amended when combinations of games are considered.

Composition: Initially, two possibilities of games' composition are detailed in (Maudet, 2001): embedded and pre-sequenced games —others combinations are studied in (Chaib-draa et al., 2002). Intuitively, two games are pre-sequenced when the first game "prepares" the second game. A particular kind of pre-sequence is at stake when the prepared game is known by the two conversants (because it has been proposed). Such pre-sequencing games can be played to ensure that entry conditions of a forthcoming game are actually established — *e.g.* to make public a controversial issue before entering a persuasion game. Embedding is now familiar to us and implies that several games may be open at the same time. This suggests an appropriate structure on these active games — the simplest being a stack. Note again that in this framework the consequence of embedding is just that commitments of the embedded game have priority over commitments of the embedding game: in the absence of a conflict, those commitments contracted in embedding games can affect allowed or precluded moves in the current game.

[4] The notation is inspired from (Singh, 2000): $C(x, \alpha_1 | \alpha_2)$ stands for the commitment of x (towards y, ignored for lisibility) to do α_1 or α_2, and $C(x, \alpha_1 \Rightarrow \alpha_2)$ for the conditional commitment to do α_2 in case of the occurrence of the event α_1. Operations on the commitments used in this table are just creation and cancellation.

Pre-sequencing means that g_2 is opened while g_1 is proposed.
Conditions game g_1 is *proposed*.
Effects successful termination of game g_1 involves entering game g_2.
In case that the first game is not successful, the second game is simply ignored.

Embedding means that g_1 is now opened while g_2 was already opened.
Conditions game g_1 is *open*.
Effects (conversational) commitments of the embedded games are considered prioritary over those of the embedding game.

Much work needs to be done to incude these conditions and effects in an extended contextualization game.

7 Conclusion and perspectives

This survey has presented the current state of the art in the field of conversation policies, with a special focus on the growing field of dialectical models of dialogue: dialogue-game based and commitment-based approaches. These two approaches should not be viewed as two distinct alternatives since they share most of the underlying dialectical principles of interaction.

The commitment-based approach regards conversation as a social activity, and consequently aims at defining the semantics of ACLs in terms of some social notions. Technically, we have seen that this may be done alternatively through the addition of some modal operator, or through the introduction of commitment stores together with some (more or less explicit) procedural interpretation. The dialogue-game based approach can be considered as an extension of the commitment-based approach, which basically consists of considering the CPs as structures *per se*. This paper has illustrated and emphasized the differences between the particular models proposed in the literature, and most notably the expected interplay between the protocol and the commitments. More research is needed (both in theoretical aspects of commitments' mechanisms, and practical aspects of implementation) to unfold these issues, evaluate and compare more carefully these models.

Let us now return to the different objectives that we listed in section 3.2, and discuss how models based on dialectical notions can, or not, fullfill them. Recently, a number of desiderata for dialectical protocols have also been proposed (McBurney et al., 2002). Most of them meet the requirements that we list here —we refer to these desiderata in the discussion.

Flexibility :

1. *Adopt a formalism which allows more flexibility than FSMs.* Keeping track of the state of the dialogue, as proposed in the dialectical literature, is a key point to make the CPs more flexible. For instance, recording the commitments of the participants in a conversational

store makes possible ulterior references to these commitments instead of just on *the* previous move of the dialogue (*e.g.* backtracking replies). Notice that in the dialectic approach, the basic structure is just a set of commitments, but it may be useful to also store the questions currently challenged, or under discussion —see *e.g.* (Gordon, 1996). In these cases, it may be necessary to use some notion of relevance (Prakken, 2000; Dastani et al., 2001) to ensure the focus of the dialogue. Most of the current systems are also non-cumulatives, in the sense that they allow (under some conditions) the retraction of commitments (MacKenzie, 1979). This is important to enable the *self-transformation* (McBurney et al., 2002) of the participants through the dialogue: changing their point of view, preferences, and so forth.

2. *Permit unexpected messages within the CPs.* Protocols described as FSMs require agents to conform to the expected transitions. Commitments motivate agents to conform to some expected behavior (Excelente-Toledo et al., 2001), discourage disruption (McBurney et al., 2002) and thus facilitate the coordination between agents. However, agents remain autonomous and can act in a different manner if they have reason for doing so (Dignum, 1999). This is a strong argument supporting the definition of the conversational rules of the CPs in term of commitments. The CPs describe the expected conversational behaviors, but participants have the possibility of not following these rules (Hulstijn, p.c.). This raises the delicate question of the system policy on normative compliance and control. An elegant solution is to admit norm violation as long as nobody objects (Brewka, 2001) .

3. *Allow CP's composition.* Composition is not considered by the commitment-based approach, whereas this is clearly one of the main features of the dialogue-game approach. We have seen that different possibilities on games' composition have been carefully studied in literature. Among these structurations, "embedding" has received much attention. Following (Reed et al., 2002), we cite for instance negotiation within deliberation (the picture hanging case of (Parsons and Jennings, 1996)), info-seeking within deliberation, expert-consultation within persuasion, deliberation within critical discussion, negotiation within persuasion (the car sales case of (McBurney and Parsons, 2002), or the team formation case of (Dignum et al., 2000)), or inquiry within persuasion (the party case of (Maudet and Evrard, 1998)). Some important questions remain however: are all kind of structurations allowed with all types of dialogue? —for instance, some of the embeddings proposed above are controversial. What are the consequences of these structurations on the commitments and the games? —for instance, are the embedding games just suspended during execution of embedded games? Can we list all the

useful structurations? —see (McBurney and Parsons, 2002) for the most complete proposal in this respect to date.

4. *Study how agents come to an agreement on the current CP.* Within the dialogue game approach, this question is central (Maudet, 2002) Most of the interest of this approach lies in the expected possibility to shift from game to game during conversation. But few is said about how this is done in practice. We have seen that, following (Reed, 1998), most approaches now make use of some meta-acts of communication to handle game structure. Also the need to define a protocol for game negotiation has been identified —see the control layer of (McBurney and Parsons, 2002), or the contextualization game of (Maudet, 2001). However, the mechanisms underlying this game negotiation process is still ongoing research. A desiderata proposed in (McBurney et al., 2002) is that no agent should be precluded to participate in a dialogue game if he is qualified and willing to participate (*inclusiveness*). It may seem at first glance contradictory with the notion of game negotiation, but we should keep in mind that establishing the qualification of the agent may lead to a negotiation process.

Specification :

1. *Adopt public specifications.* Firstly, CPs are supposed to be defined as public conversational principles, independent of agents' internal specificities. Defining the CPs as a combination of some public protocol and conversational store is thus adequate. Also adequate is the "hardcore" commitment-based approach of (Flores and Kremer, 2002) or (Maudet, 2001), that defines protocol themselves in terms of public social commitments. This indicates that some approaches —*e.g.* (Amgoud et al., 2000a) or (Maudet and Evrard, 1998) are not sufficiently clear about this distinction between the internal and external aspects of their models, as also noted in (McBurney and Parsons, 2001b). Secondly, *all* the specifications must be public (*transparency*) —the *dialogue purpose*, *the dialogue rules*, the *underlying argumentation theory*— in order to ensure the *fairness* of the system (McBurney et al., 2002).

2. *Adopt a declarative approach.* The commitment-based approach is sometimes used in combination with some underlying protocol not declaratively defined, but only pictured as an automata. In a dialogue game, the (dialogue or commitment) rules are declared. This facilitates the task of designing games, and improves clarity and expressiveness. Eventually, this makes possible the definition of CPs or games in some standard explicit language, *e.g.* XML —see in this line the AgentFactory Toolkit of (Economou et al., 2001), or the Dialogue Game Simulator currently developped by Chaib-draa *et al.* (Bourget, 2001).

3. *Exhibit properties of the CPs.* Among the different properties relevant to be studied, the specific problem of termination has recently been addressed by some researchers. Wooldrige and Parsons have studied termination of some negotiation protocols (Wooldridge and Parsons, 2000), while (Parsons et al., 2002) have studied the termination and complexity of other types of dialogue as suggested by the Walton and Krabbe's typology (information-seeking, inquiry, and persuasion). The work of (Sadri et al., 2001) is also very relevant here. Termination, of course, involves the definition of rather rigid policies (*e.g.* by restricting the number of possible answers, or limiting the number of embedded challenges). Under this label, we could list the desideratas of *rule-consistency*, *system simplicity*, and *computational simplicity* of (McBurney et al., 2002).
4. *Optimise the CPs.* This issue hasn't yet received a comprehensive account. However, a recent paper (Economou et al., 2001) paves the way for such studies. This approach is not commitment-based, since commitments are not explicitly used in the definition of CPs —in so doing, they do not fully take advantage of flexibility as discussed above. The idea is rather to extract the (implicit) commitments contracted during the execution of CPs (described as FSMs). The authors propose an algorithm which isolates these "deontic states", and use this algorithm to verify the fitness of a CP w.r.t. to a given context. This kind of tool will certainly facilitate the optimisation of CPs, for instance by identifying useless paths that may be eliminated in a CP without consequences on the (commitment level) semantics.

Finally, we would like to raise the difficult question of the **agent architecture** which should fit these new languages of communication. In other words, how should we design agents to allow them to dialogue with the languages of communication described in this survey? Although not unique, the approach of normative/deliberative agents seems promising. In their deliberative process, such agents take into account their own intentions, belief and goals (following the classical theories of rational agency) but also the normative notions which motivate them to act as they are committed to (follow the rules of the dialogue, conform to the conversation policy). This supposes of course the use of a private mentalistic semantics in addition to the public notions described so far —see also the notion of layered semantics (Pitt and Mamdani, 2000). For instance, Amgoud *et al.* defines for each move some particular rationality rules, *i.e.* the conditions under which the moves may be played during the dialogue, wr.t. to the private argumentation system of each agent—see (Parsons et al., 2002) for a generalization of this work—.

But how should these different levels be combined? In fact, allowing the agents to reason about their social and private attitudes involves complex architectures. Recent works explore this aspect with attention: the BDOING (Belief, Desire, Obligations, Intentions, Norms and Goals) architecture (Castelfranchi

et al., 2000), the BOID (Belief, Obligations, Intentions, Desires) architecture (Broersen et al., 2001), the normative/deliberative model of agency (Boella and Lesmo, 2000). These numerous works make us confident as to the future developments of these models.

References

Amgoud, L., Maudet, N., and Parsons, S. (2000a). Modelling dialogues using argumentation. In *Proceedings of the 4th International Conference on Multi-Agent Systems (ICMAS00)*, pages 31–38, Berlin. IEEE Press.

Amgoud, L., Parsons, S., and Maudet, N. (2000b). Arguments, dialogue, and negotiation. In *Proceedings of the European Conference on Artificial Intelligence (ECAI)*, pages 338–342, Berlin. IOS Press.

Aristotle (1928). *Topics*. Clarendon Press, Oxford, UK. (W. D. Ross, Editor).

Austin, J. L. (1962). *How to Do Things With Words*. Oxford University Press: Oxford, England.

Barbuceanu, M. and Fox, M. (1995). Cool: A language for describing coordination in multiagent systems. In *Proceedings of the first International Conference on Multi-Agent Systems (ICMAS)*, pages 17–25.

Boella, G. and Lesmo, L. (2000). Deliberative normative agents. In *Proceedings of the Workshop on Norms and Institutions*, Barcelona, Spain.

Bourget, D. (2001). *Manual of the Dialogue Game Simulator*. Université Laval, Québec, Canada. Technical Report.

Bratman, M. (1987). *Intention, plans, and practical reason*. Harvard University Press, Cambridge, MA.

Brewka, G. (2001). Dynamic argument systems: a formal model of argumentation based on situation calculus. *Journal of logic and computation*, 11(2):257–282.

Broersen, J., Dastani, M., Hulstijn, J., Huang, Z., and van der Torre, L. (2001). The boid architecture: conflicts between beliefs, obligations, intentions and desires. In *Proceedings of Fifth International Conference on Autonomous Agents (Agents2001)*, pages 9–16, Montreal, Canada. ACM Press.

Castelfranchi, C., Dignum, F., Jonker, C., and Treur, J. (2000). Deliberative normative agents: principles and architecture. In *Intelligent Agents VI —LNAI 1757*, pages 364–378.

Chaib-draa, B., Maudet, N., and Labrie, M. A. (2002). Request for action reconsidered as a dialogue game based on commitments. In *Proceedings of the Workshop on Agent Communication Languages and Conversational Policies*, Bologna, Italy.

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

Cohen, P. R. and Levesque, H. J. (1995). Communicative actions for artificial agents. In *Proceedings of the International Conference on Multi-agent Systems (ICMAS95)*, pages 65–72.

Colombetti, M. (2000a). A commitment-based approach to agent speech acts and conversations. In *Proceedings of the Workshop on Agent Languages and Conversational Policies*, pages 21–29, Barcelona.

Colombetti, M. (2000b). Commitment-based semantics for agent communication languages. In *Proceedings of the First Workshop on the History and Philosophy of Logic, Mathematics and Computation (HPLMC00)*, Sansebastian.

Cost, R. S., Chen, Y., Finin, T., Labrou, Y., and Peng, Y. (2000). Issues in agent communication: an introduction. In *(Dignum and Greaves, 2000)*, pages 178–192.

Dastani, M., Hulstijn, J., and der Torre, L. V. (2000). Negotiation protocols and dialogue games. In *Proceedings of the Belgium/Dutch Artificial Intelligence Conference (BNAIC00)*, Kaatsheuvel.

Dastani, M., Hulstijn, J., and der Torre, L. V. (2001). Negotiation protocols and dialogue games. In *Proceedings of the 5th International Conference on Autonomous Agents (Agents2001)*, pages 180–181, Montreal, Canada. ACM Press.

Dignum, F. (1999). Autonomous agents with norms. *AI and law*, 7(1):69–79.

Dignum, F., Dunin-Kęplicz, B., and Verbrugge, R. (2000). Agent theory for team formation by dialogue. In Castelfranchi, C. and Lespérance, Y., editors, *Pre-Proceedings of the Seventh International Workshop on Agent Theories, Architectures, and Languages (ATAL-2000)*, pages 141–156, Boston, USA.

Dignum, F. and Greaves, M., editors (2000). *Issues in agent communication*, volume 1916 of *Lecture Notes in Computer Science*. Springer-Verlag.

Economou, G., Tsvetovat, M., Sycara, K., and Paolucci, M. (2001). Implicit commitments through protocol-level semantics. In *Proceedings of the fifth International Conference on Autonomous Agents (Agents2001)*, pages 79–80, Montreal, Canada. ACM Press.

Excelente-Toledo, C., Bourne, R. A., and Jennings, N. R. (2001). Reasoning about commitments and penalties for coordination between autonomous agents. In *Proceedings of fifth International Conference on Autonomous Agents (Agents2001)*, pages 131–138, Montreal, Canada. ACM Press.

Ferber, J. (1995). *Les systèmes multi-agents: vers une intelligence collective.* InterEditions.

Finin, T., Labrou, Y., and Mayfield, J. (1997). KQML as an agent communication language. In Bradshaw, J., editor, *Software agents*, pages 291–316. AAAI/MIT Press.

Flores, R. F. and Kremer, R. C. (2002). A formal theory for agent conversations for actions. *Computational intelligence.* (submitted).

FIPA, Foundation for Intelligent Physical Agents (2001). Communicative act library specification (XC00037H). http://www.fipa.org/spec.

Gordon, T. F. (1996). Computational dialectics. In Hoschka, P., editor, *Computers as assistants - A new generation of support systems*, pages 186–203. Lawrence Erlbaum Associates.

Greaves, M., Holmback, H., and Bradshaw, J. (2000). What is a conversation policy? In *(Dignum and Greaves, 2000)*, pages 118–131.

Guerin, F. and Pitt, J. (2001). A denotational semantics for agent communication languages. In Müller, J. P., Andre, E., Sen, S., and Frasson, C., editors, *Proceedings of the Fifth International Conference on Autonomous Agents*, pages 497–504, New York, USA. ACM Press.

Hamblin, C. L. (1970). *Fallacies.* Methuen.

Hulstijn, J. (2000a). Dialogue games are recipe for joint action. In Cohen, R., editor, *Proceedings of the 4th Workshop on the semantics and pragmatics of dialogue (Gotalog00).*

Hulstijn, J. (2000b). *Dialogue models for inquiry and transaction.* PhD thesis, University of Twente, The Netherlands.

Jennings, N. R. (1993). Commitments and conventions: the foundation of coordination in multi-agent systems. *The knowledge engineering review*, 8(3):223–250.

Kone, M. T., Shimazu, A., and Nakajima, T. (2000). The state of the art in agent communication languages. *Knowledge and information systems*, 2:259–284.

Labrou, Y. and Finin, T. (1998). Semantics and conversations for an agent communication language. In Huhns, M. and Singh, M., editors, *Reading in agents*, pages 235–242. Morgan Kaufmann.

Labrou, Y., Finin, T., and Peng, Y. (1999). Agent communication languages: the current landscape. *IEEE Intelligent systems*, pages 45–52.

Levin, J. and Moore, J. (1980). Dialogue-games: meta-communication structure for natural language interaction. *Cognitive science*, 1(4):395–420.

Lin, F., Norrie, D. H., Shen, W., and Kremer, R. (2000). A schema-based approach to specifying conversation policies. In *(Dignum and Greaves, 2000)*, pages 193–204.

Lorenzen, P. and Lorenz, K. (1978). *Dialogische Logik*. Wissenschaftliche Buchgesellschaft, Darmstadt, Germany.

MacKenzie, J. (1979). Question-begging in non-cumulative systems. *Journal of philosophical logic*, 8:117–133.

Maudet, N. (2001). *Modéliser l'aspect conventionnel des interactions langagières: la contribution des jeux de dialogue*. PhD thesis, Université P. Sabatier, Toulouse.

Maudet, N. (2002). Negotiating games —a research note. *Journal of autonoumous agents and multi-agent systems*. (to appear).

Maudet, N. and Evrard, F. (1998). A generic framework for dialogue game implementation. In *Proceedings of the 2nd Workshop on Formal Semantics and Pragmatics of Dialogue*, pages 185–198, University of Twente, The Netherlands.

McBurney, P. and Parsons, S. (2001a). Agent ludens: games for agent dialogues. In Parsons, S. and Gmytrasiewicz, P., editors, *Procceedings of the Workshop on Game Theoretic and Decision Theoretic Agents (GTDT2001)*, AAAI Spring Symposium, pages 70–77, Menlo Park, CA. AAAI Press. Tech. Report SS-01-03.

McBurney, P. and Parsons, S. (2001b). Desiderata for dialogue-game protocols for agent interactions. In *Proceedings of the Workshop Adventures in Argumentation*, pages 44–47, Toulouse, France.

McBurney, P. and Parsons, S. (2002). Games that agents play: a formal framework for dialogue between autonomous agents. *Journal of Logic, Language, and Information —Special issue on logic and games*, 11(3). (to appear).

McBurney, P., Parsons, S., and Wooldridge, M. (2002). Desiderata for agent argumentation protocols. In Castelfranchi, C. and Johnson, W. L., editors, *Proceedings of the First International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2002)*, Bologna, Italy. AAMAS.

Moore, D. J. (1993). *Dialogue game theory for intelligent tutoring systems*. PhD thesis, Leeds Metroplitan University, England.

Moulin, B. (1983). The social dimension of interactions in multiagent systems. In Wobcke W., Zhang, C., editor, *Agent and multi-agent systems*, LNAI 1441. Springer Verlag.

Moulin, B. and Chaib-draa, B. (1996). An overview of distributed artificial intelligence. In O'Hare, G. and Jennings, N., editors, *Foundations of Distributed Artificial Intelligence*, pages 3–55. Wiley interScience.

Parsons, S. and Jennings, N. (1996). Negotiation through argumentation —a preliminary report. In *Proceedings of the 2nd International Conference on Multi Agent Systems (ICMAS96)*, pages 267–274. AAAI Press.

Parsons, S., Sierra, C., and Jennings, N. R. (1998). Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261–292.

Parsons, S., Wooldridge, M., and Amgoud, L. (2002). A study of formal inter-agent dialogues. In Castelfranchi, C. and Johnson, W. L., editors, *Proceedings of the first International Conference on Autonomous Agents and Multi-agent Systems (AAMAS2002)*, Bologna, Italy.

Parunak, H. V. (1996). Visualizing agent conversations: using enhanced dooley graphs for agent design and analysis. In *Proceedings of the Second International Conference on Multi-agent Systems (ICMAS96)*, pages 275–282. AAAI Press.

Pitt, J. and Mamdani, A. (2000). Communication protocols in multi-agent systems: a development method and reference architecture. In *(Dignum and Greaves, 2000)*, pages 160–177.

Prakken, H. (2000). On dialogue systems with speech acts, arguments, and counterarguments. In *Proceedings of the 7th European Workshop on Logic for Artificial Intelligence (JELIA)*, Lecture Notes in AI 1919, pages 239–253.

Prakken, H. (2001). Relating protocols for dynamic dispute with logics for defeasible argumentation. *Synthese*, 127:187–219.

Reed, C. (1998). Dialogue frames in agent communication. In *Proceedings of the Third International Conference on MultiAgent Systems (ICMAS98)*, pages 246–253, Paris, France. IEEE Press.

Reed, C., Norman, T., and Gabbay, D. (2002). *Symposium on Argumentation and Computation*. (to appear).

Sadri, F., Toni, F., and Torroni, P. (2001). Logic agents, dialogues and negotiation: an abductive approach. In Schroeder, M. and Stathis, K., editors, *Proceedings of the Symposium on Information Agents for E-Commerce, Artificial Intelligence and the Simulation of Behaviour Conference (AISB-2001)*, York, UK. AISB.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press: Cambridge, England.

Singh, M. P. (1998). Agent communication languages: rethinking the principles. *IEEE Computer*, pages 40–47.

Singh, M. P. (2000). A social semantics for agent communication language. In *(Dignum and Greaves, 2000)*, pages 31–45.

van Eemeren, F. H. and Grootendorst, R. (1992). *Argumentation, communication, and fallacies: a pragma-dialectical perspective.* Lawrence Erlbaum, London.

Vongkasem, L. and Chaib-draa, B. (2000). Acl as a joint project between participants. In *(Dignum and Greaves, 2000)*, pages 235–248.

Walton, D. and Krabbe, E. (1995). *Commitment in dialogue: basic concepts of interpersonal reasoning.* State University of New York Press, Albany, NY.

Winograd, T. and Flores, F. (1986). *Understanding Computers and Cognition: A New Foundation for Design.* Ablex Publishing Co., Norwood, New Jersey, 1986.

Wooldridge, M. (2000). Semantic issues in the verification of agent communication languages. *Journal of Autonomous Agents and Multi-Agent Systems*, 3(1):9–31.

Wooldridge, M. and Parsons, S. (2000). Languages for negotiation. In *Proceedings of the 14th European Conference on Artificial Intelligence (ECAI00)*, pages 393–397, Berlin, Germany.