A Common Ontology of Agent Communication Languages:
Modeling Mental Attitudes and Social Commitments using Roles

Guido Boella [a,*], Rossana Damiano [a], Joris Hulstijn [b] and Leendert van der Torre [c]

[a] *Università di Torino, Italy - email: {guido,rossana}@di.unito.it*
[b] *Vrije Universiteit, Amsterdam, The Netherlands - email: jhulstijn@feweb.vu.nl*
[c] *University of Luxembourg, Luxembourg - email: leendert@vandertorre.com*

Abstract. There are two main traditions in defining a semantics for agent communication languages, based either on *mental attitudes* or on *social commitments*. These traditions share speech acts as operators with preconditions and effects, and agents playing roles like speaker and hearer, but otherwise they rely on distinct ontologies. They refer not only to either belief and intention or various notions of social commitment, but also to distinct speech acts and distinct kinds of dialogue. In this paper, we propose a common ontology for both approaches based on *public mental attitudes* attributed to *role instances*. Public mental attitudes avoid the unverifiability problem of private mental states, while reusing the logics and implementations developed for FIPA compliant approaches. Moreover, a common ontology of communication primitives allows for the construction of agents which do not need separate reasoning modules to participate in dialogues with both mental attitudes and social commitments compliant agents. Moreover, a common ontology of communication primitives allows for the construction of agents participating in and combining the full range of dialogues covered by the individual approaches without having to redefine the existing protocols to cope with new dialog types. We illustrate how to extend the ontology to a semantics for agent communication and how to define mappings from existing semantics to the new one.

Keywords: Agent Communication Languages, dialogue, speech acts, communication primitives, roles

1. Introduction

Agents are components in modern software that can be added to a dynamic and complex environment. Agents can sense the state of the environment and alter it by acting. Agents are autonomous, in the sense that they direct there own behavior, according to a set of goals or tasks. Agents may be designed to achieve goals and tasks of different stakeholders. Moreover, the agents in an environment can be heterogeneous in the sense that they are built in different ways. But although agents may have different goals and may be built in different ways, they can interact with other agents in the surrounding environment. The interaction mechanism is based on an *agent communication language*. Because an agent communication language should make heterogeneous agents able to communicate, there is a need for standardization of both the agent communication languages and the languages for expressing the content of the messages that are communicated.

In this paper, the *ontology* of agent communication languages is studied for standardization. Agent communication languages share the idea that agents communicate by performing speech acts (Austin, 1962; Searle, 1969), and that these actions can be modeled as a kind of planning operators with preconditions and effects (Cohen and Perrault, 1979; Allen and Perrault, 1980). Longer stretches of dialogue can be explained by plans, that structure attempts of participants to reach their joint goals, e.g. (Lochbaum, 1998). However, many variations of agent communication languages have been designed on these common speech act foundations, thus making the standardization effort difficult.

---

*Corresponding author: Guido Boella, Università di Torino, Italy - email: guido@di.unito.it

The partial ontology visualized in Figure 1 highlights the distinctions in terms of communication primitives between two main traditions in agent communication. To the left of the agent are his *mental attitudes* like beliefs and intentions, which are included in preconditions and effects of speech acts used in information seeking dialogues based on speech acts such as *inform* and *request*. This side has been popularized by the standardization efforts of the Foundation for Intelligent Physical Agents (FIPA, 2002a). To the right of the agent, we find the social semantics tradition, developed as a reaction to FIPA (Singh, 2000). The agent is related to his social commitments both as creditor and as debtor, which are adopted in the semantics of speech acts used in negotiation or persuasion. Though both kinds of dialogue rely on the terminology of commitment, requesting an action in negotiation means something else than defending a proposition in persuasion after a challenge. In Section 5 we therefore distinguish action commitment to give a semantics of speech acts like *request* or *propose* in negotiation (Singh, 2000; Fornara and Colombetti, 2004), from propositional commitment used for speech acts like *assert* or *challenge* in persuasion (Hamblin, 1970; Walton and Krabbe, 1995).
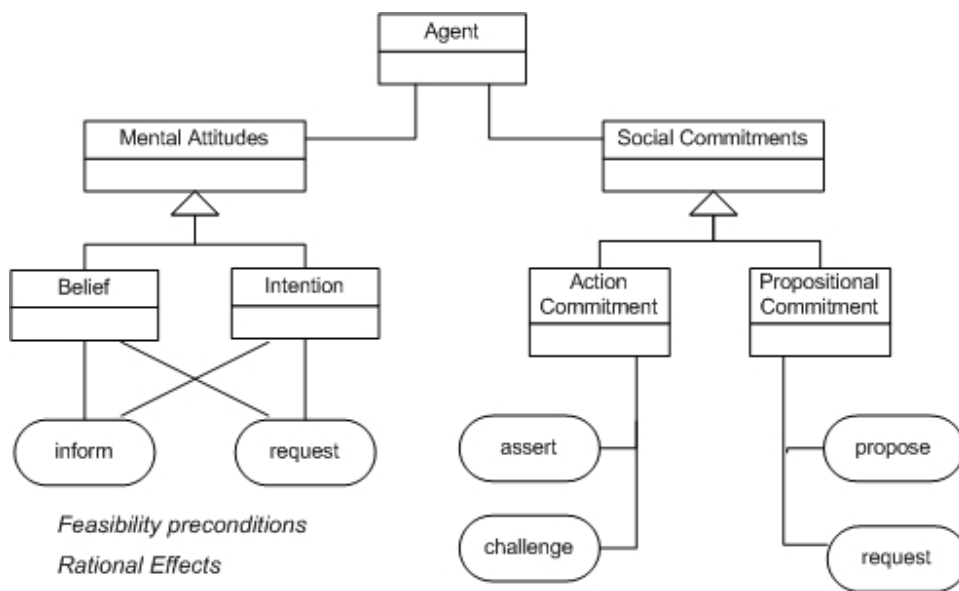

Fig. 1. Partial ontology of *Mental Attitudes* (FIPA) and *Social Commitments* based approaches

The comparison given in Table 1, further discussed in Section 7 on related research, illustrates that both traditions are distinct in most aspects. Not only the speech acts and dialogue types of the mental attitudes and social commitment based approaches are distinct, but they are based also on distinct kinds of semantics, agent attitudes, and underlying assumptions. Consequently, building a bridge between them to create a common ontology is a challenging problem. Moreover, it illustrates that for combined dialogues this problem has to be addressed. For example, consider a negotiation or persuasion dialogue that is intertwined with information seeking dialogues, like requesting the details of the object under negotiation. The separation of the two traditions makes it difficult to interpret the interaction among negotiation and information seeking in such a mixed dialogue.

The approaches that bridge the two traditions are based on a reinterpretation of the beliefs and intentions in Figure 1. The idea that meaning cannot be private, but is inter-subjective or based on a common ground, has been accepted for a long time in the philosophy of language. This has been discussed, for example, by Lewis (1969) or Stalnaker (2002) from the analytic tradition, who stress the importance of the common ground. The crucial point is that a private semantics does not make it possible for a language user to be objectively wrong about an interpretation from a third person point of view. In contrast to FIPA, beliefs and intentions are not interpreted as private mental attitudes, but as a kind of public mental attitudes, for example as grounded beliefs (Gaudou et al., 2006a) or as ostensible beliefs, or public opinions (Nickles et al., 2006).

	Mental Attitudes (FIPA)	Social Commitments
Attitudes	beliefs, intentions	propositional commitments, action commitments
Speech acts	inform, request	propose, assert, challenge
Semantics	model theoretic semantics	operational semantics
Dialogue types	information seeking	negotiation, persuasion
Attitude	cooperative	competitive
Assumptions	sincerity, cooperativity	fear of social disapproval or sanction

Table 1

Differences between approaches based on *Mental Attitudes* (FIPA) and *Social Commitments*

In this paper we build an ontology for public mental attitudes in agent communication centered on the concept of *role*. Communication in multi-agent systems is often associated with the roles agents play in the social structure of the systems (Zambonelli et al., 2003; Ferber et al., 2003), but existing approaches to the semantics of agent communication languages do not take roles into account, and thus far roles have not been associated with public mental attitudes in agent communication languages. Roughly, the basis of the idea to build the ontology on roles is that we associate beliefs and intentions with a role instance played by an agent, and that these role-based mental attitudes have a public character. The various alternatives for public mental attitudes can be compared only when it is shown how existing approaches can be mapped onto them. In this paper we show how this can be done for our own role-based semantics. In particular, we answer the following four questions:

1. What is a role-based ontology for agent communication and how can it be extended to a semantic language?
2. How can the agent communication language of FIPA be mapped onto the role-based language?
3. How can languages based on action commitments be mapped onto the role-based language?
4. How can languages based on propositional commitments be mapped onto the role-based language?

In this paper, a mapping from one communication language to another, means that for all expressions in the first language, the semantic representation for the first language can be replaced by a semantic representation in the representation format that is normally used for the second language. Ideally, this mapping should respect the intended meaning of the speech acts, but in some cases, we will see that information may get lost in the process. Therefore we do not aim for a complete mapping, in the sense that the original semantics is reflected completely in our role-based semantics, but we aim to maintain the spirit of the other language. The success criteria are that all dialogues in the original agent communication languages can be given a corresponding meaning in our role-based language.

We explain the main challenges in these four questions below. The ontology of agent communication answering the first question is a list of concept definitions and relations among them to describe the semantics of agent communication. As an illustrative example, in Figure 2 we visualize the speech act *inform*.

This figure visualizes two agents playing respectively roles of type r_1 and r_2 in a dialogue game, and should be read as follows. Since agents can play the same role several times, either simultaneously or sequentially, a role instance is characterized by an agent x, y, a role type r_1, r_2 and a session s of a dialogue game type. The lefthand side visualizes agent x (bottom) with role instance xr_1s (top), and the righthand side visualizes agent y (bottom) with role instance yr_2s (top). Thus, compared to most existing models, we distinguish the private state from the public face, for each agent playing a role. We associate beliefs (B) and intentions (I) with the role instances too; these are the public mental attitudes. The beliefs and intentions of the role instance are constructed by the introduction or removal of beliefs and intentions by the speech acts, according to the rules describing the dialogue game type, called constitutive rules. Since the performance of a speech acts is public for those participating in a dialogue, also the changes are public.

Agent x in role r_1 informs agent y in role r_2 that φ, and therefore φ becomes part of his public beliefs, as specified by the feasibility precondition of *inform* in FIPA, under the assumption that this belief persists. Moreover, the reason that he informs the other agent is that he wants him to believe φ, and therefore
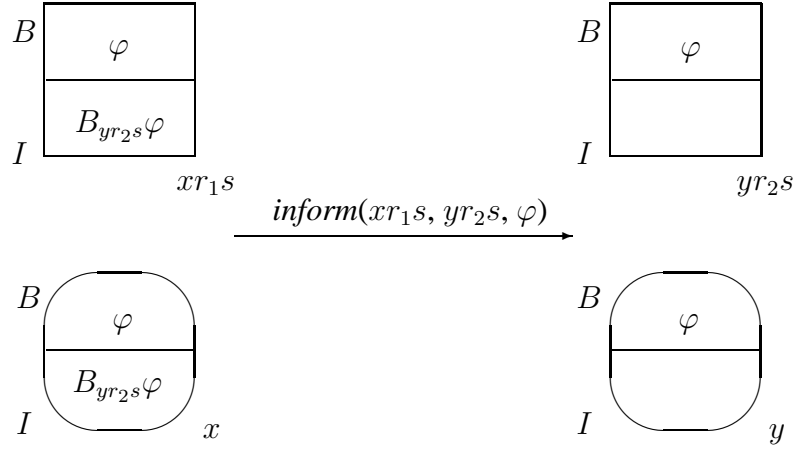
Fig. 2. Private and public mental attitudes.

$B_{yr_2s}\varphi$ is part of public intentions of agent x in role r_1, according to the constitutive rules of the dialogue game type. If agent y (publicly) believes that agent x is reliable, then his public beliefs will contain φ too, as visualized by φ in the top right square of the Figure 2.

The public beliefs and intentions of a dialogue participant in a particular role may differ in interesting ways from his private beliefs and intentions, but if the agents are sincere, then they will privately believe φ too, as visualized by φ in both the left and bottom right squares of Figure 2. The figure is explained in more detail in Section 3.1. The ontology has to clearly define the notion of a role, given the many interpretations of roles in computer science and in other sciences, and it has to distinguish role from role instance, and from the agent itself. The challenge in defining roles is to explain how mental attitudes can be attributed to role instances, and how these mental attitudes can be made public. As part of this problem, it has to be clarified which mental attributes should be assigned to the role instance and which to the agent. Finally, the set of concepts has to be extended to a semantic language, with all the constructs and variables needed to describe a semantics of speech acts.

A challenge in defining the mapping from FIPA to our role-based semantic language to answer the second question is to deal with the notion of precondition. For example, if an agent x informs another agent that φ, for example with $inform(xr_1s, yr_2s, \varphi)$ in Figure 2, then in the original FIPA semantics, agent x has a so called feasibility precondition, that he should already believe the proposition φ. If not, x would be insincere. With public mental attitudes, however, it is *not* the case that an agent must already publicly believe φ in his role r_1. To see why, assume that this precondition would hold. When two agents meet there are no public attitudes besides the common knowledge associated with the roles of the agents being played. So in that case, an agent would not be able to make an *inform*. Moreover, if the information were believed by the roles and thus were already public, then there would be no need for an *inform* in the first place. Clearly, the notion of precondition has to be reinterpreted in the context of public, role-based semantics.

Another challenge of this mapping is that in FIPA notions like reliability, cooperativity and sincerity are built into the conceptual system. We have to relax these properties if we want to integrate the social commitment approach, and deal for example with non-cooperative types of dialogue like argumentation or negotiation. Therefore, we have to make properties like reliability, cooperativity and sincerity explicit in the formal language, to distinguish between different contexts of use, and we have to explain how these properties can be associated with the role instances. The following formula explains the use of a reliability concept. Detailed explanation of this rule is given in Section 3.1. Since all modal operators are indexed by the same session s, we leave the session out of the rules.

$$B_{yr_2}(I_{xr_1}B_{yr_2}\varphi \wedge reliable(xr_1, \varphi)) \rightarrow B_{yr_2}\varphi \tag{RL16}$$

An initial problem of the mapping from the social commitment approach to answer the last questions is that there is not a single social commitment approach. There are approaches based on action commitment, and approaches based on propositional commitment. Moreover, the social commitment approaches focusing on action commitments (Singh, 2000; Fornara and Colombetti, 2004), are usually given an operational semantics. An additional challenge in how to use the proposed bridges between the two traditions to compare them is that we need either to translate one approach into the other, for all possible pairs of alternatives, or to have a common framework of reference. Moreover, it is necessary to be able to compare the individual speech acts which are not always defined for precisely the same purpose, to understand which are more general or specific in one approach or the other, and to make the hidden assumptions of the models explicit. Finally, we need to define not only commitments but also their evolution and the rules describing their evolution.

For extending our ontology into a semantic language we take a logical language that is directly derived from the FIPA semantic language SL (FIPA, 2002a), which we extend with roles. We call this language Role-SL. We extend it also with two rules to express that mental attitudes are public. We isolate a minimal notion of role, needed for agent communication. This minimal notion of role possesses what we call the characteristic property of roles: the fact that we can attribute mental attitudes to it. We take mental attitudes of a role as our ontological primitive, and derive role from it. The ontological question is then no longer "what is a role?", but rather "what are role-based mental attitudes, and how do they evolve?" To answer this question we consider why multiple sets of mental attitudes are needed, for each role the agent plays, instead of a single set for the agent itself. We define role-based mental attitudes, and explain why they are public. We discuss how it builds upon the notion of role we developed in our earlier work on normative multiagent systems (Boella and van der Torre, 2004). For the mapping from FIPA to the role-based ontology we define translation schemes for FIPA speech acts to Role-SL. For mapping of social commitment based approaches to a role-based ontology we take a similar approach, translating the states of the operational semantics, to sets of public mental attitudes in Role-SL.

The paper is structured as follows. In Section 2 we introduce our role-based ontology for agent communication, and we extend it to a semantic language. In Section 3 and 4 we map respectively the FIPA agent communication language and a social semantics (Fornara and Colombetti, 2004) into role-based semantics. In Section 5 we deal in particular with the difference between action commitments and propositional commitments, and we show how we can account for persuasion dialogues in our approach. In Section 6 we translate an instance of the propose interaction protocol from FIPA and from social semantics into our role model, to compare the two approaches. The paper ends with a discussion of related research and conclusions.

2. Role-based ontology for agent communication

Whereas usually in agent communication only a semantic language is given to describe the meaning of speech acts, we distinguish between an ontology and such a language. In Section 2.1 we define the ontology we use in this paper, that is, the set of concepts and the relations among them. In Section 2.2 the set of concepts is related to the semantic language SL used by FIPA, extended with the distinction between private mental attitudes of agents and public mental attitudes associated with role instances, and extended with rules reflecting that mental attitudes associated with roles are public.

2.1. The ontology

An agent can be part of multiple communication sessions at the same time. For example, consider an agent who informs another agent that his web services can be used by all other agents, informs a third agent that his services can only be used by agents with the appropriate certificates, requests from a fourth agent a document, and informs a fifth agent that he does not have an intention to obtain this document. Intuitively, we can make another 'copy' of the agent for each session, and we may refer to each 'copy' as

a role playing agent or a role enacting agent (Dastani et al., 2003). In practice an agent may be required to maintain a minimal consistency among these 'copies', but in this paper we do not consider this consistency problem, because it depends on the application domain of the language.

For each role playing agent, we distinguish between the private beliefs and intentions of the agent, and the public beliefs and intentions attributed to the role instance. For example, when a role playing agent believes that the secret key of the safe used when playing a role in the organization enrolling him is "123456", then this belief is not known to everybody and therefore is not attributed to the role instance, but is part of private mental attitudes of the agent.

The public beliefs and intentions attributed to the role instance are known to everybody overhearing the conversation and aware of the rules of the dialogue game. They represent what the agent is publicly held responsible for. If an agent does not adhere to his role by for example refusing what he has said before, then he can be sanctioned or blamed. Consequently, once an agent starts lying, he has to continue the dialogue consistently with these lies, independently of his real beliefs. We consider beliefs and intentions only, because knowledge may be considered as a strong kind of belief, and goals have a very minor role in the agent communication languages we consider in this paper such as the FIPA language.

Besides the uniqueness of the private mental attitudes and the variance of the public mental attitudes according to the roles an agent is playing, there are various other good reasons why one might want to distinguish the mental attitudes of a role instance from the mental attitudes of the agents themselves:

- If we wish to model that an agent is bluffing (telling something he is not sure of) or lying (telling something he believes is false), then we have to account for divergences between the private mental attitudes and the public ones;
- In a negotiation situation the agents only have the intention of reaching an agreement in common, but they do not have to communicate truthful information or their real preferences. For example, in a negotiation a bid does not need to correspond to the actual reservation price of the agent.
- In persuasion, if the aim is to win a dispute, an agent can adopt a point of view "for the sake of the argument", in order to show its inconsistency or counterintuitive consequences, without actually endorsing the view himself (Caminada (2004)).
- It is necessary to keep apart the motivations for playing a role from the rules of the game which affect the state of the roles. For example, an agent may be sincere, in the sense that he really acts as expected from his role, for pure cooperativity or for the fear of a sanction.

Since the mental attitudes of role instances are the foundation of our ontology, they have to be precisely defined. In particular, we have to clearly distinguish the mental attitudes attributed to a role instance from mental attitudes attributed to the agent playing the role. The mental attitudes attributed to a role instance consist of two parts. First, they contain those public beliefs and intentions following from uttering speech acts. Second, the mental attitudes of a role consist of commonly held beliefs about the attitudes of roles. For example, a buyer in negotiation is expected to prefer lower prices. Moreover, agents in an organization may be expected to be sincere, cooperative and reliable when communicating with other agents in the same organization. No other mental attitudes are attributed to role instances.

Moreover, we distinguish two kinds of mental attitudes attributed to role instances as the result of speech acts. For example, consider an agent informing another one that "it is raining". We can derive that the agent publicly believes that it is raining, and that he has the public intention that the other agent believes it is raining, because these mental attitudes are presupposed when making such an action. More precisely, the agent's belief that it is raining is presupposed, and the belief that the other agent will believe it is the desired effect.

Definition 1 *Mental attitudes of role instances (R) consist of beliefs (B) and intentions (I), and are precisely:*

1. *Consequences of speech acts on the mental attitudes of agents overhearing the uttering of the speech act, representing what the agent is publicly held responsible for, distinguishing between:*

(a) *Presuppositions of the speech act; these are public mental attitudes from which we can infer that they held already before the speech act was uttered;*

(b) *Other public mental attitudes following from the speech acts; from these we cannot infer that they held already before the speech act was uttered.*

2. *Common knowledge associated with the roles being played by the agents.*

We build our ontology on the concept of "mental attitudes of a role instance" in Definition 1, rather than on the concept of role instance or role. However, the ontology has to clarify also what we mean by role instance, given the many interpretations of roles in computer science and in other sciences. This is a non-trivial problem, since there are numerous incompatible notions of role available in the literature (Baldoni et al., 2007; Loebe, 2007; Masolo et al., 2005; Herrmann, 2007). In this paper we are interested in role instances mostly, where a role is a generic class and a role instance is associated with an agent playing the role in a dialogue session.

Definition 2 *The characteristic property of a role instance is that it is channel end of communication, and contains mental attitudes representing the state of the session of interaction. It may or may not refer to other properties of roles such as expertise needed to play the role, powers associated with the role, expectations associated with the role, and so on.*

The starting point of the common ontology for agent communication consists of the ontology of mental attitudes approach to agent communication: agents, beliefs, intentions, utterances, speech act types, also called performatives, and dialogue types. Then we make two decisions. The first decision is to include sessions rather than leaving them out of the ontology. We could leave them out of the ontology, since the rules of the game always refer to a single session only, and we do not consider cross-session rules. Though leaving the sessions out of the ontology would simplify our ontology a little, it obscures the fact that agents can play multiple sessions of the same dialogue type at the same time. In particular, we like to compare the entities to which we attribute the private and public mental attitudes. Since, at some moment in time, a private mental attitude is attributed to an agent, and a public mental attitude is attributed to an agent playing a role in a session, we include the sessions in the ontology too.

The second decision concerns the representation of distinct abstractions, in particular role types and role instances, dialogue types and dialogue session instances, and speech act types and utterances. As usual there are various ways to represent them in a conceptual model. We choose to represent them all as first class citizens in our model to clarify the precise relations among them, and again to express that mental attitudes are attributed to agents and role instances. Therefore, we extend the basic ontology with dialogue sessions and roles to represent role instances with their public mental attitudes. We add several relations between these concepts. A role instance is played by an agent in a session, a dialogue session has a dialogue type, and agents and role instances are described by their beliefs and intentions. Moreover, utterances are made by role instances, directed towards role instances, and have a speech act type.

Moreover, we introduce rules defining the effects of speech acts on the mental attitudes of role instances. We call them the constitutive rules or norms of the dialogue game played by the agents. The constitutive rules describe a speech act in terms of the beliefs and intentions of agents, in the context of a dialogue type. Therefore, a speech act may mean something else in sessions of another dialogue type. Finally, in the ontology we do not consider the content of speech acts.

We have chosen to exclude various other concepts and relations from our ontology. For example, the belief of an agent that it is raining, and the belief of another agent that the first agent has this belief, are related. These complex relations among beliefs and intentions are *not* represented in the ontology. To detail such relations, a modal language is used in the more detailed semantic languages describing the mental attitudes. Moreover, also the content of the beliefs and intentions, and the constitutive norms are described in a semantic language referring to beliefs and intentions.

Definition 3 *A role-based ontology for agent communication is a tuple*

$$\langle A, RT, S, R, PL, DT, SDT, RDT, B, I, AD, SA, U, USA, URS, URA, CR, CRD, CRM \rangle$$

where:

- *A is a set of agents, like x and y.*
- *RT is a set of role types, like r, r_1, r_2, . . .*
- *S is a set of sessions, like s, s_1, s_2, . . .*
- *R is a set of role instances, like i or j.*
- *PL : $A \times RN \times S \mapsto R$ is a role playing function, such that for each role instance i there is precisely one agent x, role type r and session s with $i = PL(x, r, s)$. We also say that $agent(i) = x$, $roletype(i) = r$, and $session(i) = s$.*
- *DT is a set of dialogue types, such as information seeking, negotiation and persuasion.*
- *SDT : $S \mapsto DT$ if a function giving the dialogue type of a dialogue session.*
- *RDT $\subseteq RT \times DT$ if a relation describing the role types that can participate in a dialogue session of a dialogue type.*
- *B and I are sets of beliefs and intentions respectively, either of individual agents or of role instances. We write $M = B \cup I$ for the set of all mental attitudes.*
- *AD : $A \cup R \mapsto B \cup I$ is a function associating mental attitudes with agents and role instances.*
- *SA is a set of speech act types: $inform$, $request$, $propose$, etc.*
- *U is a set of utterances or messages, i.e., instances of uttering a speech act.*
- *USA : $U \mapsto SA$ is a function associating an utterance with its speech act type.*
- *URS : $U \mapsto R$ is a function associating an utterance with the role instance of the speaker.*
- *URA : $U \mapsto 2^R$ is a function associating an utterance with the (set of) role instances of the addressees.*
- *CR is a set of constitutive rules of dialogue games.*
- *CRD : $CR \mapsto SA \times DT$ is a function that associates with each constitutive rule a speech act and dialogue type.*
- *CRM : $CR \mapsto 2^M$ is a function detailing which mental attitudes are affected by a constitutive rule.*

The ontology is visualized in Figure 3. For readability we have also added the full names of the concepts, and ways to read some of the relations. It is instructive to compare our ontology in Figure 3 with the one in Figure 1. Besides the extension with roles and sessions, we have now only beliefs and intentions, and no longer commitments. Moreover, the class of speech acts in Figure 3 and the individual speech acts in Figure 1 represents that we have a single uniform way to provide a semantics for them, and therefore a uniform way to formalize dialogue sessions of varying dialogue types.

Finally, we explain in which sense the mental attitudes of role instances of the kind defined in Definition 3 are public. This is based on our notion of public communication, defined as follows. Communication is public for every agent who overhears the speech acts, and has the same prior knowledge of roles and dialogue types. Consequently, for public communication, the mental attitudes of role are public. The public beliefs and intentions attributed to the role instances are known to everybody overhearing the conversation according to the constitutive rules or norms. Constitutive rules explain how roles can be autonomous, in the literal sense of the term: auto-nomos, the behavior of roles is described by their own norms. The constitutive rules exist only since roles as descriptions are collectively accepted by the agents. Moreover, since the constitutive rules which change the beliefs and intentions publicly attributed to roles are accepted by all agents, the beliefs and intentions resulting from the application of a constitutive rule after a speech act are public too. This issue of the construction of social reality is discussed in (Boella and van der Torre, 2006a).
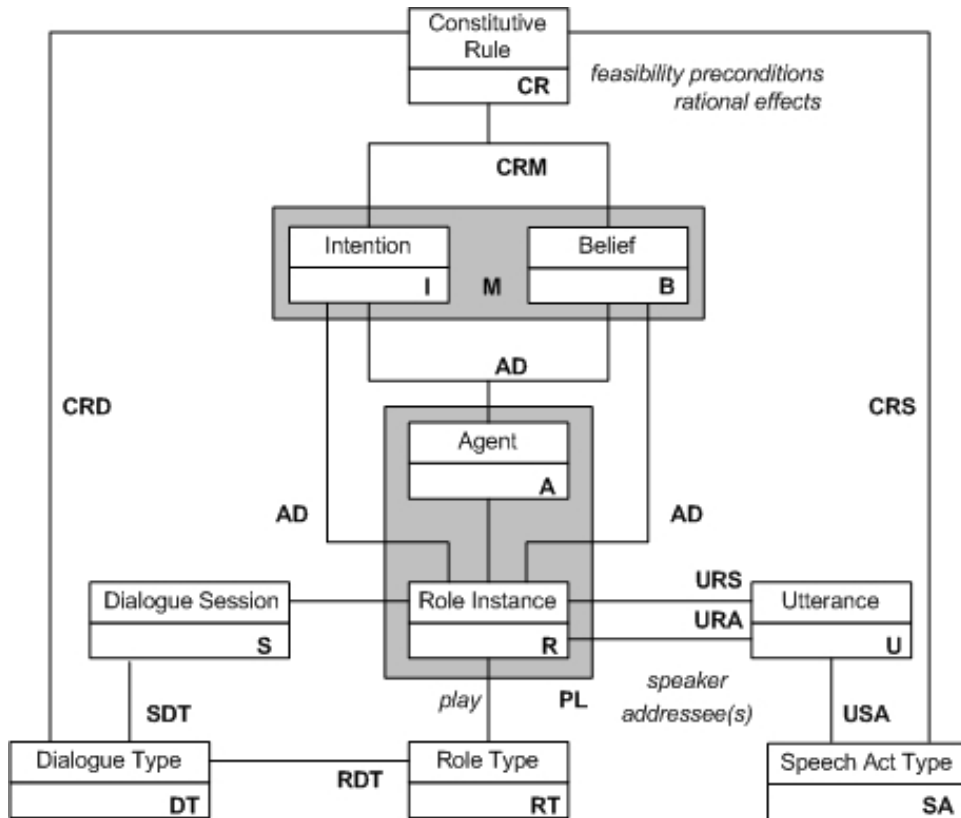
Fig. 3. The role-based ontology of communication.

Definition 4 *Communication is public in a group $S \subseteq A$, if and only if:*

1. *every agent in S hears the speech acts that are uttered, and has the same knowledge of roles and dialogue types.*
2. *every agent in S knows which role is played by the utterers of the speech acts.*
3. *every agent in S knows which beliefs are typically attributed to roles types.*

The constitutive rules may describe also the rules of the protocol the role is defined for, or the roles agents play in the social structure of the systems. See Section 7.2 for a discussion.

2.2. The semantic language Role-SL

We now sketch how a BDI logic can be designed based on our ontology, which can be used as a semantic language. In principle, we can use any existing modal and temporal predicate logic, but we have to introduce the following two important extensions.

1. We have to introduce role instances, and therefore roles and sessions. More precisely, we have to replace the agent identifiers in the BDI logic by a superclass consisting of agents and role instances. Another way to look at this, is that the set of agents in the BDI logic is partitioned into agents and role instances.
2. We have to make mental attitudes of role instances public. This is done by adding rules that state that if a role instance has a belief or intention, then the other role instances will believe this.

The distinction between agents and role instances raises the question how the agent and its role instances are related. Since this issue is relevant for our mapping of FIPA speech acts to our Role-SL, where we have to formalize FIPA property of sincerity. We discuss this issue in Section 3.1.

There are various BDI logics which can be used, or we can design a new language from scratch. Since our aim in this paper is to show the generality of our ontology, we show how to extend FIPA's specification language (FIPA, 2002a) and its predecessor (Sadek, 1991; Sadek et al., 1997; Sadek, 1999). However, this language has some limitations, which have been discussed in the literature (Pitt and Mamdani, 1999; Wooldridge, 2000). The three main problems are that the logic is based on a rather complex relation between desires, goals and intentions, that there is no explicit representation of a conditional or rule (simulated by material implication according to (FIPA, 2002a)), and that there is no explicit representation of time. In particular, the absence of explicit time has the consequence that there is no formal way to reason about the persistence of mental attitudes over time, for example which preconditions persist after the speech act has been uttered. Though solutions to some of these problems have been proposed (Herzig and Longin, 2000), there is no commonly accepted standard formalism, and we therefore continue to use FIPA's SL. In this paper we are not concerned with these limitations, our aim is to show only how to adapt this language for our ontology.

2.2.1. Basis of the Semantic Language Formalism

In FIPA-SL, propositions about mental attitudes and actions are formalized in a first order modal logic with identity. We follow the exposition in Annex A of (FIPA, 2002a) and refer to (Sadek, 1991) for details of this logic. The components of the formalism used in our Role-SL are as follows, where we extend the set of agents with role instances. Since we use the variables i and j to refer to role instances, we introduce new variables e and f to refer to the superclass of agents and role instances.

- $p, q, \ldots$ are taken to be closed formulas denoting propositions,
- φ and ψ are formula schemas, which stand for any closed proposition,
- $e, f \in A \cup R$ are schematic variables denoting agents and role instances, and,
- $\models \varphi$ means that φ is valid.

The mental model of an agent or role instance in SL or Role-SL is based on the representation of three primitive attitudes: belief, uncertainty and choice (or, to some extent, goal). They are respectively formalized by the modal operators B, U, and D. FIPA-SL uses the letter C for the latter, but to avoid confusion with commitments later in this paper, we use the letter D. Formulas using these operators can be read as follows:

- $B_e p$: e (implicitly) believes (that) p,
- $U_e p$: e is uncertain about p but thinks that p is more likely than $\neg p$, and,
- $D_e p$: e desires that p currently holds.

The logical model for the belief operator B in both SL and Role-SL is a KD45 Kripke structure (see, for example, (Halpern and Moses, 1985)) with the fixed domain principle (see, for example, (Garson, 1984)).

To enable reasoning about action, the universe of discourse of FIPA-SL involves, in addition to individual objects and agents, sequences of events. To talk about complex plans, events (or actions) can be combined to form *action expressions*, noted as a:

- $a_1; a_2$ is a sequence in which a_2 follows a_1
- $a_1 \mid a_2$ is a nondeterministic choice, in which either a_1 happens or a_2, but not both.

FIPA-SL introduces the operators $Feasible$, $Done$ and $Agent$ to enable reasoning about actions:

- $Feasible(a, p)$ means that a can take place and if it does p will be true just after that,
- $Done(a, p)$ means that a has just taken place and p was true just before that,
- $Agent(e, a)$ means that e denotes the only subject (agent or role) that ever performs (in the past, present or future) the actions which appear in action expression a,
- $Single(a)$ means that a denotes an action that is not a sequence. Any individual action is $Single$. The composite act $a; b$ is not $Single$. The composite act $a \mid b$ is $Single$ iff both a and b are Single.

From belief, choice and events, the concept of persistent goal is defined. An agent or role instance e has p as a persistent goal, if e has p as a goal and is self-committed toward this goal until e comes to believe that

the goal is achieved or to believe that it is unachievable. Intention is defined as a persistent goal imposing the agent to act. Formulas as PG_ep and I_ep are intended to mean that "e has p as a persistent goal" and "e has the intention to bring about p", respectively. The definition of I entails that intention generates a planning process. See Sadek (1992); Cohen and Levesque (1990) for the details of a formal definition of intention.

The FIPA documentation notes that there is no restriction on the possibility of embedding mental attitude or action operators. For example, formula $U_eB_fI_fDone(a, B_ep)$ informally means that agent or role instance e believes that, probably, agent or role instance f thinks that f has the intention that action a be done before which e has to believe p.

FIPA-SL also uses various abbreviations, including the following ones:

1. $Feasible(a) = Feasible(a, True)$
2. $Done(a) = Done(a, True)$
3. $Done(j, a) = Done(a) \wedge Agent(j, a)$.
4. $Possible(\varphi) = (\exists a)Feasible(a, \varphi)$
5. $Bif_e\varphi = B_e\varphi \vee B_e\neg\varphi$

The mapping of our ontology on Role-SL is as follows:

– The set of agents A and role instances R are mapped to indexes of modal operators, and can be used in the description of actions. The sets of roles types RT and sessions S can be used in the description of role instances using the function PL. We write xrs for $PL(x, r, s)$.
– The set of dialogue types DT and the function SDT giving the dialogue type of a dialogue session are represented by a special proposition $dialogue(s, d)$, and similarly for the other relations. The relation between roles and dialogue types can be used as a constraint: if an agent plays a role in a session, then the role must be allowed given the dialogue type of the session. Similarly other kinds of constraints can be expressed in Role-SL.
– The beliefs B and intentions I as well as the function $AD : A \cup R \mapsto B \cup I$ associating mental attitudes with agents and role instances are represented by modal formulas.
– Speech act utterances U are included in the set of actions.
– The constitutive rules CR are represented by formulas of Role-SL.
– The function CRD associating a speech act type and dialogue type with each constitutive rule, is represented by prefixing all rules with the relevant speech act type and dialogue type.
– The function $CRM : CR \mapsto 2^M$ detailing which mental attitudes are affected by a constitutive rule is made explicit by the description of the mental attitude in the constitutive rule.

Moreover, the semantic model of FIPA-SL leads to the properties of FIPA-SL in Table 2. In Role-SL, we have to decide whether we accept the properties for the role instances only, or also for the agents. Property 0, 1 and 2 are related to the logic of action and we therefore accept it both for role instances as for the agents, e or f. Property 3, 4 and 5 are related to the pre and postconditions of speech acts, and we therefore accept them only for role instances, i and j. The following three properties are of particular interest for the role model.

– Each role has correct knowledge about its own mental states, in particular, its beliefs about its intentions are correct. These rules correspond to FIPA's (2002a) schema $\varphi \leftrightarrow B_e\varphi$, (called Property 0 in Table 2, since it has no explicit name in FIPA) where φ is starting by an operator formalizing a mental attitude of agent or role instance e.
– Property 1 and 5 hold only under some conditions, and we therefore add a condition Φ to the properties of FIPA-SL. For example, Property 5 holds only for feasibility preconditions that persist. We do not discuss these conditions in this paper, see the FIPA documentation for details. Properties 3, 4 and 5 are discussed in Section 3, when we present the speech acts together with their feasibility preconditions FP and rational effects RE.

Property 0	$((B_e I_e p) \leftrightarrow I_e p) \wedge ((B_e \neg I_e p) \leftrightarrow \neg I_e p)$	(RL1)
	$((B_e B_e p) \leftrightarrow B_e p) \wedge ((B_e \neg B_e p) \leftrightarrow \neg B_e p)$	(RL2)
Property 1	$\Phi \wedge I_e p \rightarrow I_e Done(act_1 \mid \ldots \mid act_n)$	(RL3)
Property 2	$I_e Done(act) \rightarrow B_e Feasible(act) \vee I_e B_e Feasible(act))$	(RL4)
Property 3	$I_i Done(act) \rightarrow I_i RE(act)$	(RL5)
Property 4	$B_i((Done(act) \wedge Agent(j, act)) \rightarrow I_j RE(act))$	(RL6)
Property 5	$B_i(\Phi \wedge Done(act) \rightarrow FP(act))$	(RL7)

Table 2

The properties of FIPA-SL

2.2.2. Extensions for public mental attitudes

In our model, we have to add axioms which are necessary due to the separation of role instances from agents. Adding axioms to a logical system raises the question whether the system is still consistent, and even when the system remains consistent, whether the new axioms do not interfere with the existing properties. This is the main logical issue in this paper; the other issues are of a more conceptual nature. We do not give a formal proof of the consistency of our system, because the full logical details of FIPA-SL and therefore also Role-SL are beyond the scope of this paper, but we rely on the following intuitive argument. The axioms which make mental attitudes public are the most challenging ones, since they imply properties not found in FIPA: if a role instance believes some mental attitude of a role instance, then also the other role instances engaged in the same session will believe it. First note that SL and therefore Role-SL only models what holds after a speech act; there is no reference to the past (or, in general, reference to time). Then, if we consider what can be derived from speech acts only, these properties can be interpreted in FIPA too. *It means that an agent has communicated all its beliefs, and that agents believe of other agents only what has been communicated to them.* Consequently, since our model can be interpreted as a special case in which agents immediately tell all they know to the other agents, and FIPA is generally applicable, so also to this special case, our axioms do not lead to inconsistency. The remaining axioms are harmless, since they only derive beliefs for particular kinds of formulas.

The first two properties concern the public character of mental attitudes when they are attributed to roles. Since the constitutive rules CR have as effect public mental attitudes, all inferences on the basis of speech acts are public. This motivates rules (RL8) and (RL9), expressing that each role instance has complete knowledge about the other role instances' beliefs and intentions. For example, from left to right, if an agent y in role r_2 in session s has the belief that p, expressed by $B_{yr_2s}p$, then this implies that any other agent x playing any role r_1 in the same session s believes this, $B_{xr_1s}B_{yr_2s}p$. Since properties always refer to the same session, we leave out the session to increase readability, and we write $B_{yr_2}p$ implies $B_{xr_1}B_{yr_2}p$, or that $B_i \neg p$ implies $B_j B_i \neg p$, where $i = xr_1$ and $j = yr_2$. The reading from right to left is analogous. If role instance yr_2 believes something about xr_1's beliefs, then role instance xr_1 actually believes that. Note that if $xr_1 = yr_2$, the rule coincides with Property 0 of FIPA-SL in Table 2 reflecting positive and negative introspection:

$$(B_{yr_2}p \leftrightarrow B_{xr_1}B_{yr_2}p) \wedge (\neg B_{yr_2}p \leftrightarrow B_{xr_1}\neg B_{yr_2}p) \tag{RL8}$$

$$(I_{yr_2}p \leftrightarrow B_{xr_1}I_{yr_2}p) \wedge (\neg I_{yr_2}p \leftrightarrow B_{xr_1}\neg I_{yr_2}p) \tag{RL9}$$

Another aspect of the fact that communication is public, is that the agents are aware of the communication. So if a belief is associated with a role instance, then it is believed by all agents participating in the same dialogue session. Under the same conditions as in the previous two rules, if an agent y in role r_2 in session s has the public belief that p, expressed by $B_{yr_2s}p$, then this implies that any other agent x playing any role r_1 in the same session s believes this privately, $B_x B_{yr_2s}p$.

$$(B_{yr_2}p \wedge play(x, r_1) \leftrightarrow B_x B_{yr_2}p) \wedge (\neg B_{yr_2}p \wedge play(x, r_1) \leftrightarrow B_x \neg B_{yr_2}p) \tag{RL10}$$

$$(I_{yr_2}p \wedge play(x, r_1) \leftrightarrow B_x I_{yr_2}p) \wedge (\neg I_{yr_2}p \wedge play(x, r_1) \leftrightarrow B_x \neg I_{yr_2}p) \tag{RL11}$$

In FIPA the architecture used for the exchange of messages is not part of the logical formalism, but it is described in a separate document, FIPA (2002b). A speech act is put inside a message which is exchanged between agents on an agent platform. We need the following rule to link the logical model with such a message architecture and information about the roles of agents. The action $send()$ corresponds to (part of) the FIPA message structure specification, with parameters speech act type ($sa \in SA$), sender ($x \in A$), receiver ($y \in A$), content (p), and conversation-id ($s \in S$). Information about the roles that agents are playing in the current session s is stored in relation PL.

$$Done(send(sa, x, y, p, s, ...)) \land PL(x, r_1, s) \land PL(y, r_2, s) \rightarrow Done(sa(xr_1s, yr_2s, p)) \quad \text{(RL12)}$$

3. From FIPA to roles

For the role-based semantics for FIPA communicative acts we have to define the constitutive rules CR of our dialogue game. Using Role-SL, the mapping from FIPA to roles may seem straightforward, since we have defined Role-SL as an extension of FIPA-SL. The constitutive rules are therefore simply the instantiations of property (RL6) and (RL7) of FIPA. However, we have to define the sincerity and cooperativity assumption of FIPA, and we have to show that the resulting formalization using public mental attitudes instead of private mental attitudes has the expected behavior in the FIPA protocols.

In Section 3.1 we formalize the sincerity, cooperativity and reliability assumption in Role-SL, in Section 3.2 we compare the formalization the FIPA speech acts in Role-SL with their formalization in FIPA-SL, and finally in Section 3.3 we illustrate the formalization by an example.

3.1. The formalization of sincere cooperative agents, and other agent types

In this section we formalize sincerity, cooperativity and reliability in Role-SL. In the following Section 3.2, we consider in more detail how these kinds of properties are modeled in FIPA-SL. Roughly, sincerity is explicit in FIPA-SL, but cooperativity and reliability are represented more implicitly.

FIPA assumes *sincerity*, in the sense that an agent can make an *inform* speech act only when it believes the content is true, and it requests or proposes something only when it has the intention that it be achieved. This roughly corresponds to the version of sincerity given by Sadek, inspirator of the FIPA approach.

> "*Sincerity*: An agent i cannot have the intention that an agent j comes to believe a proposition p without believing p herself or without having the intention to come to believe it. (Sadek, 1999, p. 182)"

In our role-based model, we model a weaker kind of sincerity. In our model, sincerity means that the public beliefs or intentions of the role instances are attributed also to the agents themselves. This is a common assumption in everyday life (Goffman, 1959). In our visualization of role-based communication in Figure 2, this kind of sincerity represents an implication from the top level to the bottom level. We model it in three steps. First, we have preconditions and effects, attached to speech acts. These generate public beliefs about role instances, similar to the original sincerity properties for agents in FIPA. Second, we add an explicit property of sincerity, modeled by rule (RL13) and (RL14). Third, we assume that whenever an agent is a FIPA agent, then it must be sincere. Using these steps, an agent can be sincere in one role, or in one session, and insincere in another. As before, we do not explicitly represent the session in the properties, so it is implicitly assumed that all references to role instances refer to the same session.

$$B_{xr}\varphi \land sincere(x, r, \varphi) \rightarrow B_x\varphi \quad \text{(RL13)}$$

$$I_{xr}\varphi \land sincere(x, r, \varphi) \rightarrow I_x\varphi \quad \text{(RL14)}$$

However, this does not imply that our model of Role-SL, together with the sincerity assumption, coincides with the FIPA model, since we do not have an implication the other way around. In other words, like FIPA we model the implication of public mental attitudes towards the private ones, but in the constitutive rules of the game we refer only to public mental attitudes. We do *not* assume that private mental attitudes

are made public, and therefore we do not assume that the private mental attitudes can be reused to check the preconditions of speech acts. In other words, the preconditions of speech acts are exclusively used to infer the public mental attitudes attributed to role instances and not to plan speech acts. In FIPA, preconditions are used to plan a speech act. Insincere speech acts are not possible since the speaker must believe the preconditions before executing a speech act. In our approach, instead, the private mental state of the speaker is kept separate from the public mental attitudes, unless sincerity is explicitly assumed. In this way it is possible to perform a speech act where neither preconditions are believed nor effects are intended by the speaker.

Cooperativity is implicit in FIPA. Although there is no formal property in FIPA-SL or in the speech acts discussed in the next section, if agents are not cooperative, there is no communication. It is assumed that agents recognize the intentions of other agents, and try to help them in reaching their intentions. Historically, the implicit assumption of cooperativity makes sense, because part of the FIPA framework was developed originally for designing cooperative natural language dialogue systems (Sadek, 1999), on the basis of the plan-based theory of speech acts, and a theory of rational agents (Cohen and Perrault, 1979; Allen and Perrault, 1980; Lochbaum, 1998). Such approaches propose general principles for plan- or intention recognition: based on the actions of the speaker, the hearer can try to recognize the underlying intention or plan, and subsequently adjust her (cooperative) response. The most obvious cooperative principle is intention adoption:

> "The *minimal principle of cooperation* (or the *intention adoption principle*) states that agents must not only react when they are addressed but, more than that, they must adopt the interlocutor's intention whenever they recognize it, and if there have no objection to adopt it." (Sadek, 1999, p. 182)

However, when agent communication languages are used in applications on the internet, or for non-cooperative interaction types like negotiation and debate, cooperativity can no longer be assumed as a general principle. Only in specific circumstances it makes sense to assume cooperativity, for example among agents that are known or trusted.

In our role-based model, we can make implicit cooperativity assumptions explicit. In particular, a successful *request* may add an intention to the intentions of the hearer, but only when we have a rule such as (RL15), that specifies that the hearer is cooperative. Cooperativity is a role-role relationship, which depends on the social context in which an interaction takes place.

$$B_{yr_2}(I_{xr_1}Done(act) \land cooperative(yr_2, xr_1, Done(act))) \rightarrow I_{yr_2}Done(act) \qquad \text{(RL15)}$$

Besides sincerity and cooperativity, the FIPA documentation also mentions *reliability*, though it neither assumes nor formalizes reliability:

> "Whether or not the receiver does, indeed, adopt belief in the proposition will be a function of the receiver's trust in the sincerity and reliability of the sender." (FIPA, 2002a)

We explain the role of reliability in FIPA, using our model of role-based communication in Figure 2. The constitutive rules of the communication game operate on the mental states of the role instances. Because of the feasibility precondition, an *inform* creates a proposition φ in the beliefs of the speaker xr_1, and a intention that yr_2 will come to believe φ. Only if the hearer yr_2 believes that the speaker is reliable, rule (RL16) will transfer the belief to the hearer's beliefs:

$$B_{yr_2}(I_{xr_1}B_{yr_2}\varphi \land reliable(xr_1, \varphi)) \rightarrow B_{yr_2}\varphi \qquad \text{(RL16)}$$

In particular, we assume that all roles are assumed to be reliable about their own mental states; however, this assumption does not extend to the private beliefs of agents, but remains confined to the beliefs of the roles they play. In our framework, when a speaker says something about its own beliefs or intentions, the belief that it believes what has said will be immediately added to the role of the hearer as a belief of the speaker role, without implying that the agent playing the hearer role believes it as well.

For more about issues of trust and reliability, we refer to Demolombe (2001); Liau (2003); Dastani et al. (2004).

$\langle x, inform(y, \varphi)\rangle$	FP:$B_x\varphi \wedge \neg B_x(Bif_{\,y}\varphi \vee Uif_{\,y}\varphi)$	
	RE:$B_y\varphi$	(1)
$\langle x, request(y, act)\rangle$	FP: $FP(act)[x \setminus y] \wedge B_x Agent(y, act) \wedge \neg B_x I_y Done(act)$	
	RE: $Done(act)$	(2)
$\langle x, agree(y, \langle x, act\rangle, \varphi)\rangle \equiv$	FP: $B_x\alpha \wedge \neg B_x(Bif_{\,y}\alpha \vee Uif_{\,y}\alpha)$	
$\langle x, inform(y, I_x Done(\langle x, act\rangle, \varphi))\rangle$	RE: $B_y\alpha$,	
	$\alpha = I_x Done(\langle x, act\rangle, \varphi)$	(3)
$\langle x, refuse(y, \langle x, act\rangle, \varphi)\rangle \equiv$	FP: $B_x\neg feasible(\langle x, act\rangle) \wedge B_x(B_j feasible(\langle x, act\rangle)\vee$	
$\langle x, disconfirm(y, feasible(\langle x, act\rangle))\rangle;$	$\quad U_x feasible(\langle x, act\rangle) \wedge B_y\alpha \wedge \neg B_x(Bif_{\,y}\alpha \vee Uif_{\,y}\alpha)$	
$\quad \langle x, inform(y, \varphi \wedge \neg Done(\langle x, act\rangle) \wedge$	RE: $B_y(\neg feasible(\langle x, act\rangle) \wedge B_y\alpha)$	
$\quad \neg I_x Done(\langle x, act\rangle)))\rangle$	$\alpha = \varphi \wedge \neg Done(\langle x, act\rangle) \wedge \neg I_x Done(\langle x, act\rangle)$	(4)
$\langle x, propose(y, \langle x, act\rangle, \varphi)\rangle \equiv$	FP: $B_x\alpha \wedge \neg B_x(Bif_{\,y}\alpha \vee Uif_{\,y}\alpha)$	
$\langle x, inform(y, I_y Done(\langle x, act\rangle, \varphi)$	RE: $B_y\alpha$	
$\quad \rightarrow I_x Done(\langle x, act\rangle, \varphi))\rangle$	$\alpha = I_y Done(\langle x, act\rangle, \varphi) \rightarrow I_x Done(\langle x, act\rangle, \varphi)$	(5)
$\langle x, accept\ proposal(\langle y, act\rangle, \varphi)\rangle \equiv$	FP: $B_x\alpha \wedge \neg B_x(Bif_{\,y}\alpha \vee Uif_{\,y}\alpha)$	
$\langle x, inform(y, I_x Done(\langle y, act\rangle, \varphi))\rangle$	RE: $B_y\alpha$	
	$\alpha = I_x Done(\langle y, act\rangle, \varphi)$	(6)
$\langle x, reject\ proposal(y, \langle y, act\rangle, \varphi, \psi)\rangle \equiv$	FP: $B_x\alpha \wedge \neg B_x(Bif_{\,y}\alpha \vee Uif_{\,y}\alpha)$	
$\langle x, inform(y, \neg I_x Done(\langle y, act\rangle, \varphi) \wedge \psi)\rangle$	RE: $B_y\alpha$	
	$\alpha = \neg I_x Done(\langle y, act\rangle, \varphi) \wedge \psi$	(7)

Table 3

The definition of speech acts in FIPA-SL

3.2. *The translation from FIPA to roles*

FIPA-SL contains twenty-two speech acts. Most of them are defined in terms of *inform* and *request*. The purpose of an *inform* act for the speaker is to get the hearer to believe some information. The purpose of a *request* is to get the hearer to do some act. This requires that the speaker does not already believe that the hearer has such an intention.

In this paper we consider seven speech acts. The FIPA definitions of speech acts in terms of feasibility preconditions (FP) and rational effects (RE) are summarized in Table 3. The ones we have selected are the most basic and often used ones, and we choose these to illustrate both the translation and two protocols, the request interaction protocol (FIPA, 2002c), and the propose interaction protocol. The request protocol uses *request*, *agree* and *refuse*, and the propose protocol uses *propose*, *accept proposal* and *reject proposal*. The difference between a *request* and a *propose* is that a *request* concerns an action of the hearer, whereas a *propose* concerns an action of the speaker.

Note that while in Table 3 the definitions are given on agents x and y, in Table 4 they are given on role instances i and j.

The definitions we use in Role-SL are given in Table 4, which are based on the following simplifications:

1. We have left out preconditions concerning uncertain beliefs, as represented by the modal operator Uif in FIPA-SL, since they are not relevant for the general principle, and in most uses of FIPA-SL they are removed. For example, the feasibility precondition of *inform* in the FIPA documentation is $B_i\varphi \wedge \neg B_i(Bif_{\,j}\varphi \vee Uif_{\,j}\varphi)$, but we left out the $Uif\varphi$ (and we wrote $Bif_{\,i}\varphi$ in full). However, if needed for some application, the extension with uncertain beliefs is straightforward.
2. In *request*, we drop the felicity precondition from FIPA that i thinks it feasible for j to perform act.
3. We use $Done(j, act)$ rather than $Done(act)$ as RE, which means that $Agent(j, act)$ is a RE rather than a FP.
4. Note that unlike the speech act definitions based on *inform*, *request* does not have a sincerity precondition: $B_i I_i Done(j, act)$. This does not matter, because agents are considered to be reliable about

$\langle i, inform(j, \varphi)\rangle$	FP:$B_i\varphi \wedge \neg B_i(B_j\varphi \vee B_j\neg\varphi)$ RE:$B_j\varphi$	(8)
$\langle i, request(j, act)\rangle$	FP: $\neg B_i I_j Done(j, act)$ RE: $Done(j, act)$	(9)
$\langle i, agree(j, act)\rangle \equiv$ $\langle i, inform(j, I_i Done(i, act))\rangle$	FP: $B_i I_i Done(i, act) \wedge$ $\quad \neg B_i(B_j I_i Done(i, act) \vee B_j\neg I_i Done(i, act))$ RE: $B_j I_i Done(i, act)$	(10)
$\langle i, refuse(j, act)\rangle \equiv$ $\langle i, inform(j, \neg I_i Done(i, act))\rangle$	FP: $B_i\neg I_i Done(i, act) \wedge$ $\quad \neg B_i(B_j\neg I_i Done(i, act) \vee B_j I_i Done(i, act))$ RE: $B_j\neg I_i Done(i, act)$	(11)
$\langle i, propose(j, act)\rangle \equiv$ $\langle i, inform(j, I_j Done(i, act) \rightarrow I_i Done(i, act))\rangle$	FP: $B_i(I_j Done(i, act) \rightarrow I_i Done(i, act)) \wedge$ $\quad \neg B_i(B_j(I_j Done(i, act) \rightarrow I_i Done(i, act)) \vee$ $\qquad B_j\neg(I_j Done(i, act) \rightarrow I_i Done(i, act)))$ RE: $B_j(I_j Done(i, act) \rightarrow I_i Done(i, act))$	(12)
$\langle i, accept\ proposal(j, act)\rangle \equiv$ $\langle i, inform(j, I_i Done(j, act))\rangle$	FP: $B_i I_i Done(j, act) \wedge$ $\quad \neg B_i(B_j I_i Done(j, act) \vee B_j\neg I_i Done(j, act))$ RE: $B_j I_i Done(j, act)$	(13)
$\langle i, reject\ proposal(j, Done(j, act))\rangle \equiv$ $\langle i, inform(j, \neg I_i Done(j, act))\rangle$	FP: $B_i\neg I_i Done(j, act) \wedge$ $\quad \neg B_i(B_j\neg I_i Done(j, act) \vee B_j I_i Done(j, act))$ RE: $B_j\neg I_i Done(j, act)$	(14)

Table 4

The definition of speech acts in Role-SL

their own intentions (see RL16), so we can derive this condition whenever it is needed from the rational effect.

5. We also drop the conditions under which requests or proposals are agreed to be carried out. These are better dealt with in Section 4 below, where we explain conditional proposals, which are easier to understand starting from the Social Commitments approach.

6. In the definition of *refuse* we also leave out some additional idiosyncrasies of FIPA, like the fact that a refusal is expressed as a *disconfirm*, which is similar to *inform*.[1] In FIPA, the reason for encoding refusal as an *inform* lies in the fact that FIPA assumes that a refusal should be motivated by the infeasibility of the requested act, and that the agent who cannot perform the requested act should inform the requester about the reasons for the infeasibility. In our perspective, this move is in line with the fact that we dropped this feasibility requirement from *requests*. In fact we have reinterpreted *refuse*, as the opposite of an *agree*.

Despite the compact representation in Table 3 and 4, often it is easier to consider the instantiations of FIPA properties as they appear in this Table 3. We therefore list the instantiations of the speech acts in Table 5.

In the instantiation it must be noted that:

– The consequent of the instantiated rules starts always with a public mental attitude of a role: this means that every role believes the effect of the rule. For this reason we leave out the belief modality in the speech acts in instances of Property (RL6) and (RL7).

– Φ in precondition: these properties hold only if they persist in time, as in FIPA.

– We assume reliability about the mental states of roles. For example, for our interpretation of the RE of *inform* above in (CR19), we translated the RE to an intention of the speaker. In that case, we would end up with $B_j Done(agree(i, j, Done(i, act))) \rightarrow I_i B_j I_i(Done(i, act))$. So it is the communicative intention of i, that j will believe that i wants to do the action. Following Sadek et al. (1997)), we can simplify this kind of $I_i B_j I_i(...)$ -formula into $I_i...$, if we assume that the hearer takes

---

[1] The difference depends on the precondition of the hearer being uncertain about the proposition or its negation.

inform	FP	$Done(inform(i,j,\varphi)) \rightarrow B_i\varphi$	(CR17)
	FP	$Done(inform(i,j,\varphi)) \rightarrow \neg B_i(B_j\varphi \vee B_i\neg\varphi)$	(CR18)
	RE	$Done(inform(i,j,\varphi)) \rightarrow I_i B_j\varphi$	(CR19)
request	FP	$Done(request(i,j,Done(j,act))) \rightarrow \neg B_i I_j Done(j,act)$	(CR20)
	RE	$Done(request(i,j,Done(j,act))) \rightarrow I_i Done(j,act)$	(CR21)
agree	FP	$Done(agree(i,j,Done(i,act))) \rightarrow B_i I_i Done(i,act)$	(CR22)
	FP	$Done(agree(i,j,Done(i,act))) \rightarrow \neg B_i\,(B_j I_i Done(i,act) \vee$ $B_j\neg I_i Done(i,act))$	(CR23)
	RE	$Done(agree(i,j,Done(i,act))) \rightarrow I_i Done(i,act)$	(CR24)
refuse	FP	$Done(refuse(i,j,Done(i,act))) \rightarrow B_i\neg I_i Done(i,act)$	(CR25)
	FP	$Done(refuse(i,j,Done(i,act))) \rightarrow \neg B_i\,(B_j\neg I_i Done(i,act) \vee$ $B_j I_i Done(i,act))$	(CR26)
	RE	$Done(refuse(i,j,Done(i,act))) \rightarrow \neg I_i Done(i,act)$	(CR27)
propose	FP	$Done(propose(i,j,Done(i,act))) \rightarrow B_i(I_j Done(i,act) \rightarrow I_i Done(i,act))$	(CR28)
	FP	$Done(propose(i,j,Done(i,act))) \rightarrow$ $\neg B_i(\,(B_j I_j Done(i,act) \rightarrow I_i Done(i,act) \vee$ $B_j\neg(I_j Done(i,act) \rightarrow I_i Done(i,act))))$	(CR29)
	RE	$Done(propose(i,j,Done(i,act))) \rightarrow B_j(I_j Done(i,act) \rightarrow I_i Done(i,act))$	(CR30)
accept *proposal*	FP	$Done(accept\ proposal(i,j,Done(j,act))) \rightarrow B_i I_i Done(j,act)$	(CR31)
	FP	$Done(accept\ proposal(i,j,Done(j,act))) \rightarrow \neg B_i(\,B_j I_i Done(j,act) \vee$ $B_j\neg I_i Done(j,act))$	(CR32)
	RE	$Done(accept\ proposal(i,j,Done(j,act))) \rightarrow I_i Done(j,act)$	(CR33)
reject *proposal*	FP	$Done(reject\ proposal(i,j,Done(j,act))) \rightarrow B_i\neg I_i Done(j,act)$	(CR34)
	FP	$Done(reject\ proposal(i,j,Done(j,act))) \rightarrow \neg B_i(\,B_j\neg I_i Done(j,act) \vee$ $B_j I_i Done(j,act))$	(CR35)
	RE	$Done(reject\ proposal(i,j,Done(j,act))) \rightarrow \neg I_i Done(j,act)$	(CR36)

Table 5

Rules expressing preconditions and effects of speech acts in the FIPA translation

the speaker to be *reliable* concerning reports about its own mental states. By rule (RL16) we derive (CR24) above.

Analogously in rule (CR30), the effect of a *propose* is directly believed by the addressee j, since it concerns a mental attitude of role i, who is thus considered reliable about it. Note the contrast with for example the effect of *agree* in (CR24). This effect is directly expressed by a public intention of the speaker i, which is, thus also believed by the hearer j. In contrast, the effect of a *propose* is expressed as a material implication between two intentions, since it represent the way FIPA-SL allows to model conditional intentions. For this reason, the implication must be nested inside a belief operator.

– Moreover, the second precondition is usually canceled by the intended effect if it is successful.

3.3. Example

We have made substantial changes to the FIPA semantics, and we therefore have to illustrate that our mapping works. We illustrate it using an example.

To illustrate the rules, we give a example of a short negotiation about household tasks, that we will use to illustrates the distinction between the FIPA approach and the social commitment approach.

    x-1: Could you do the dishes?
    y-1: Only if you put the garbage out.
    x-2: OK.

The dialogue can be interpreted as follows. Initially, agent x makes a *request* in role $i = xr_1s$. Then y makes a counter proposal in role $j = yr_2s$, but because there are no speech acts to represent counter

proposals, we model this as an implicit refusal, followed by a conditional proposal from y. Note that the original FIPA definition of *propose* does not allow conditional proposals, however, as shown in the next section, this is the interpretation of proposals given by social commitments approach. So in order to make the two languages comparable we drop the condition that the agent of the proposed action is the speaker and we allow to propose joint plans which are sequences of actions by different agents. We discuss this issue in more depth in Section 6.2.

In order to deal with composite actions, and actions executed by various agents, we have to make explicit some rationality assumptions which are assumed in the FIPA model (p. 38). Typically, such assumptions depend on the agent planning architecture:

- A rationality constraint concerning the intention that other agents perform an action: if i wants that action act is done by j, then i has the intention that j has the intention to do act:

$$I_i Done(j, act)) \rightarrow I_i I_j Done(j, act)) \tag{RL37}$$

- Some FIPA axioms concern the execution of complex actions. For example Property 1 in Table 2. In a similar vein, we add two rules concerning the distribution of tasks. Taking inspiration from (Cohen and Levesque, 1990), we add a rule to express that if an agent intends a joint action, then it intends that each part is done at the right moment. If $act = act_1; ...; act_n$, where ";" is the sequence operator, act is a joint action:

$$I_i Done(act) \rightarrow (Done(act_1(i_1); \ldots; act_k(i_k)) \rightarrow I_i Done(act_{k+1}(i))) \tag{RL38}$$

If $act_{k+1}(j)$ and $i \neq j$, then by axiom RL37:

$$I_i Done(act) \rightarrow (Done(act_1(i_1); \ldots; act_k(i_k)) \rightarrow I_i I_j Done(act_{k+1}(j))) \tag{RL39}$$

Each agent in the group wants that the others do their part.

In our example, we introduce a composite plan which is the concatenation of j's action of doing the dishes and the action by i of moving the garbage out: $kitchen(i, j) \equiv dishes(j); garbage(i)$.

Example 1

x-1: Could you do the dishes?	$request(i, j, Done(j, dishes))$
y-1: Only if you put the garbage out?	$refuse(j, i, Done(j, dishes))$
	$propose(j, i, Done(kitchen(i, j)))$
x-2: OK.	$accept\ proposal(i, j, Done(kitchen(i, j)))$

We start with an empty set of facts about the mental attitudes of x and y. In Table 6 we depict the evolution of the public mental states of agents x and y in their roles i and j.

After turn x-1 we apply rules (CR20) and (CR21). So that it adds to the public mental state of i the precondition of a *request* and its rational effect. Since the mental states are public, from $I_i Done(j, dishes)$ (by RL37), we could also write $I_i I_j Done(j, dishes)))$ we can derive that also j is aware of this intention $B_j I_i Done(j, dishes)$ (RL9) and from Proposition 0 (RL1) that also i is aware of his own intentions $B_i I_i Done(j, dishes)$.

If we assume persistence for the sincerity precondition, and for the mental attitudes which are not immediately contradicted by later updates, rules (CR25), (CR26) and (CR27) produce the next state of affairs after the implicit refusal of y-1. The second precondition from (CR26) is canceled by the success of the effect. A proper representation of this kind of reasoning requires an explicit notion of time and persistence, and rules about changes as a result of actions and events. See Boella et al. (2007) for an initial attempt based on defeasible logic with temporal operators.

The implicit proposal of y-1 is interpreted using rules (CR28) – (CR30). It adds a public belief of j about a relation between the intentions of i and j of doing the kitchen work together.

$$B_j(I_i Done(kitchen(i, j)) \rightarrow I_j Done(kitchen(i, j)))$$

	Role instance i	Role instance j
x-1: request	BELIEFS $\neg B_i I_j Done(j, dishes)$	BELIEFS
	INTENTIONS $Done(j, dishes)$	INTENTIONS
y-1: refuse	BELIEFS	BELIEFS
	INTENTIONS $Done(j, dishes)$	INTENTIONS $\neg I_j Done(j, dishes)$
y-1: propose	BELIEFS $I_i Done(kitchen(i, j)) \rightarrow I_j Done(kitchen(i,j))$	BELIEFS $I_i Done(kitchen(i, j)) \rightarrow I_j Done(kitchen(i,j))$
	INTENTIONS $Done(j, dishes)$	INTENTIONS
x-2: accept proposal	BELIEFS $I_i Done(kitchen(i, j)) \rightarrow I_j Done(kitchen(i, j))$ $I_i Done(kitchen(i, j))$ $I_j Done(kitchen(i, j))$ $I_j Done(j, dishes))$ $Done(j, dishes) \rightarrow I_i Done(i, garbage)$	BELIEFS $I_i Done(kitchen(i, j)) \rightarrow I_j Done(kitchen(i,j))$ $I_i Done(kitchen(i,j))$ $I_j Done(kitchen(i,j))$ $I_j Done(j, dishes)$ $Done(j, dishes) \rightarrow I_i Done(i, garbage)$
	INTENTIONS $Done(kitchen(i, j))$	INTENTIONS $Done(kitchen(i,j))$ $Done(j, dishes)$

Table 6

An example of dialogue interpretation in Role-SL. Note that $kitchen(i, j) \equiv dishes(j); garbage(i)$
.

Since this belief concerns the mental states of j, and every role is considered reliable about his own mental states (CR30), then also i publicly adopts this belief:

$$B_i(I_i Done(kitchen(i, j)) \rightarrow I_j Done(kitchen(i, j)))$$

Note the implication here. The implication derives from the fact that the proposal still needs to be accepted by i, before j will really adopt the conditional intention.

Finally, after acceptance x-2 by rules (CR31) – (CR33) the state of the public mental attitudes is updated. Note that rule (CR33) implicitly implements rule (RL16) so that the intention of conditionally moving the garbage is immediately public:

$$I_i Done(kitchen(i, j))$$

At this point more information can be inferred than the one provided by those rules.

First of all, by *modus ponens*, given the intention of i, it is possible to derive the intention of j to do the kitchen work:

$$I_j Done(kitchen(i, j))$$

Moreover, it is possible to infer from the intention of j of doing the kitchen work that j has the intention to do his part of the plan, since it is the right moment (from rule RL38):

$$I_j Done(j, dishes(j))$$

Moreover, from rule (RL37), we can infer that after j does the dishes ($Done(j, dishes)$), j will adopt the intention of moving the garbage out:

$$Done(j, dishes) \rightarrow I_i Done(i, garbage)$$

Note that $\neg I_j Done(j, dishes)$, which persists from the refusal, is now overridden by the new inference $I_j Done(j, dishes)$.
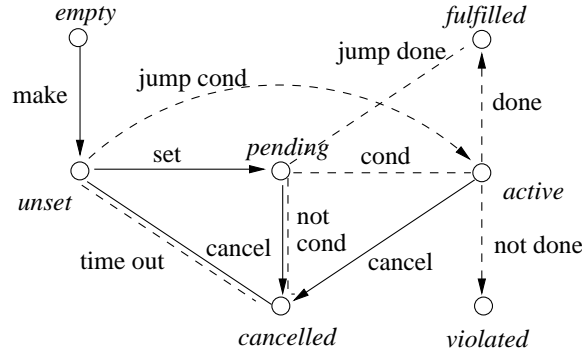
Fig. 4. Commitment State Automaton (Fornara and Colombetti, 2004).

Action	Change		
make	$empty$	$\rightarrow$	$C(unset, i, j, \varphi \mid \psi)$
set	$C(unset, i, j, \varphi \mid \psi)$	$\rightarrow$	$C(pending, i, j, \varphi \mid \psi)$
cancel	$C(-, i, j, \varphi \mid \psi)$	$\rightarrow$	$C(canceled, i, j, \varphi \mid \psi)$

Table 7
Commitment Change Rules, solid lines in Figure 4 (Fornara and Colombetti, 2004).

The example shows the interplay between mental attitudes and how the public character of roles' mental attitudes is used to make inferences in the interpretation process. Note that from the example nothing can be inferred about the private mental states of agents x and y. To make these inferences it is necessary to assume explicitly the sincerity of both agents.

4. From commitments to roles

In this section we show how to define a social semantics in the role-based model. We selected the version of Fornara and Colombetti (2004), because it is worked out in detail.

4.1. Social Commitment Semantics

The social commitment approach uses an operational semantics, meaning that the semantics of the speech acts of an agent communication language are expressed as changes in the state a commitment can be in. A social commitment is a relationship that gets established between two agents: the *debtor*, i.e., the agent who makes the commitment, and a *creditor*, i.e., the agent to which the commitment is made. The commitment is to achieve or maintain some state of affairs, expressed by a proposition. In many applications, commitments are made under certain conditions. Therefore, the conditions under which commitments are made are also expressed. Representation $C(state, i, j, \varphi \mid \psi)$ expresses a social commitment of status $state$, between debtor i and creditor j, to accomplish or maintain φ, under condition that ψ.

A commitment can have different states: *unset* (i.e., to be confirmed), *pending* (i.e., confirmed, but its condition is not true), *active* (i.e., confirmed and its condition is true), *fulfilled* (i.e., its content is true), *violated* (i.e., the content is false even if the commitment was active), or *canceled* (e.g., the debtor does not want to be committed to the action). This state can be modified by speech acts of the participants or by external events, like the execution of an action.

The state changes can be depicted in the form of a finite state automaton (Figure 4). The effects of external events and of speech acts on the status of a commitment are represented in Table 7 and Table 8 respectively. Using these semantics, Fornara and Colombetti define speech acts as given in Table 9. They

Rule	Label	Event	Status Change		
1.	done	$v(\varphi) := 1$	$C(active, i, j, \varphi \mid \top)$	$\rightarrow$	$C(fulfilled, i, j, \varphi \mid \top)$
2.	not done	$v(\varphi) := 0$	$C(active, i, j, \varphi \mid \top)$	$\rightarrow$	$C(violated, i, j, \varphi \mid \top)$
3.	condition	$v(\psi) := 1$	$C(pending, i, j, \varphi \mid \psi)$	$\rightarrow$	$C(active, i, j, \varphi \mid \top)$
4.	not condition	$v(\psi) := 0$	$C(pending, i, j, \varphi \mid \psi)$	$\rightarrow$	$C(canceled, i, j, \varphi \mid \psi)$
5.	jump condition	$v(\psi) := 1$	$C(unset, i, j, \varphi \mid \top)$	$\rightarrow$	$C(active, i, j, \varphi \mid \top)$
6.	jump done	$v(\varphi) := 1$	$C(pending, i, j, \varphi \mid \psi)$	$\rightarrow$	$C(fulfilled, i, j, \varphi \mid \psi)$
7.	time out			$C(unset, i, j, \varphi \mid \psi)$	$\rightarrow$	$C(canceled, i, j, \varphi \mid \psi)$

Table 8

Update Rules: dashed lines in Figure 4 (Fornara and Colombetti, 2004).

Precondition	Definition
	$inform(i, j, P) := \{C_i().make(i, j, P, \top).set(pending)\}$
	$request(i, j, P, Q) := \{C_i().make(j, i, P, Q)\}$
	$promise(i, j, P, Q) := \{C_i().make(i, j, P, Q).set(pending)\}$
$\exists\, C_i(unset, i, j, P \mid Q)$	$accept(i, j, C_i(unset, i, j, P \mid Q)) := \{C_i(unset, i, j, P \mid Q).set(pending)\}$
$\exists\, C_i(unset, i, j, P \mid Q)$	$reject(i, j, C_i(unset, i, j, P \mid Q)) := \{C_i(unset, i, j, P \mid Q).set(cancelled)\}$

Table 9

Speech acts of Fornara and Colombetti

use *unset* for the state resulting from a *request* and *pending* for conditional commitments, whose condition is not yet true.

Our mapping is only partial. For example, we do not have timeouts. There is no translation of the update rules, because we model them by means of the logic, e.g., using *modus ponens* in case of conditionals.

4.2. Translating Commitment States

In order to perform the translation, we adopt the following methodology: first, we map each commitment state to certain beliefs and intentions of the roles. Second, according to the constitutive rules (CR) of the role model, a speech act directly changes those beliefs and intentions in such a way to reflect the commitment introduction or change of state.

In a social commitments semantics in general, the difference between propositional and action commitment lies only in the type of their content: an action or a proposition. Fornara and Colombetti (2004) do not have a separate type of action in their language. They can only distinguish the kind of proposition: an informative proposition , or a proposition that represents a desired state of affairs. Therefore there is no substantial difference between an *inform* and a *promise* in Table 9. By contrast, according to Walton and Krabbe (1995), a propositional commitment can be seen as a kind of action commitment to defend it. In the mapping between social commitments and the role model a new distinction emerges: rather than having commitment stores, we model propositional commitments as beliefs of the role and action commitments as intentions. How to capture the idea of a commitment to defend one's position is discussed in Section 5. In this section we focus on action commitment only.

Here, we represent conditional commitments $C(state, i, j, p \mid q)$ as a conditional intention p, for example, of role i in case q is true: $B_i(q \rightarrow I_i(p))$. Note that since the condition q is outside the intention operator, the whole implication is nested in a (public) belief of the role, as we did with the translation of FIPA. We believe that conditional attitudes are better accounted for in a conditional logic, like Input/Output logic (Makinson and van der Torre (2000)). As in Section 3, we stick to FIPA's solution for the sake of clarity, while being aware of its limitations. Moreover, like in Section 3, we use a simplified notation $Done(i, act)$ to indicate that agent i is supposed to perform the action act. The $Done()$ construct turns an action into a proposition, expressing a state of affairs.

Given these preliminaries, the state translations are as follows:

- An *unset* commitment corresponds to a conditional intention believed by the creditor.

$$C(unset, i, j, Done(i, act) \mid q) \equiv B_j(q \rightarrow I_j I_i Done(i, act)) \tag{CR40}$$

In the antecedent of this rule, the commitment condition q becomes a condition on the intention assumed by the creditor. At this stage, no mental attitude is attributed to the debtor: it has not publicly assumed any intention, but has only been publicly requested to.

- A commitment is *pending* when it is a conditional intention of the creditor and the debtor of the commitment conditionally wants to perform the action provided the associated condition q is true, and the creditor has this as a belief:

$$C(pending, i, j, Done(i, act) \mid q) \equiv B_j(q \rightarrow I_j I_i Done(i, act)) \wedge \\ B_i(q \rightarrow I_i Done(i, act)) \wedge B_j(q \rightarrow I_i Done(i, act)) \tag{CR41}$$

In the automaton of Figure 4, the *unset* and *pending* states are disjoint. In the above formulation however, *pending* implies *unset*. So if it is necessary to mirror the exact definition, the *unset* state must be augmented, e.g.:

$$C(unset, i, j, Done(i, act) \mid q) \equiv B_j(q \rightarrow I_j I_i Done(i, act)) \wedge \\ (\neg B_i(q \rightarrow I_i Done(i, act)) \vee \neg B_j(q \rightarrow I_i Done(i, act))) \tag{CR40'}$$

- A commitment is *active* when it is both an intention of the debtor and of the creditor, and the pending condition is true:

$$C(active, i, j, Done(j, act) \mid \top) \equiv I_i Done(i, act) \wedge I_j Done(i, act) \tag{CR42}$$

Note that to make a pending commitment active, it is sufficient that the condition q is believed to be true, since from

$$B_i(q \wedge (q \rightarrow I_i Done(i, act)))$$

we can derive $I_i(Done(i, act))$ with rule (RL1).

- Commitments are *violated* or *fulfilled* when they are intentions of the creditor and the content of the commitment is respectively true or false according to the beliefs of the creditor (abstracting here from temporal issues):

$$C(fulfilled, i, j, Done(i, act) \mid \top) \equiv B_j Done(i, act) \wedge I_j Done(i, act) \tag{CR43}$$

$$C(violated, i, j, Done(i, act) \mid \top) \equiv B_j \neg Done(i, act) \wedge I_j Done(i, act) \tag{CR44}$$

Since roles are public, fulfilment and violation are not dependent on what the agents subjectively believe about the truth value of the content of the commitment, but on roles' public beliefs.

- Finally, a commitment is *canceled* if the creditor does not want the intention to be achieved anymore, regardless if the debtor still wants it:

$$C(canceled, i, j, Done(i, act) \mid q) \equiv \neg I_j Done(i, act) \tag{CR45}$$

4.3. Translating Speech Acts

Given the definition of the commitment state in terms of the mental states of the roles, we can provide the following translation of the speech acts semantics defined by Fornara and Colombetti (2004). In this section, we deal only with speech acts introducing action commitments, like *request*, *propose*, *promise*, *accept* and *reject*, while speech acts introducing propositional commitments, such as assertions, are discussed in Section 5.

The translation of the speech act can be directly done using the translation of commitments above and applying it to the definition of speech acts in Fornara and Colombetti (2004). However, as in the case of

promise	$Done(promise(i,j,Done(i,act)),q)) \rightarrow (q \rightarrow I_j I_i Done(i,act)) \wedge$ $B_i(q \rightarrow I_i Done(i,act)) \wedge$ $B_j(q \rightarrow I_i Done(i,act))$	(CR46)
request	$Done(request(i,j,Done(j,act)),q)) \rightarrow (q \rightarrow I_i I_j Done(j,act))$	(CR47)
accept	$Done(accept(i,j,Done(i,act)),q)) \rightarrow (q \rightarrow I_j I_i Done(i,act)) \rightarrow$ $(B_i(q \rightarrow I_i Done(i,act)) \wedge$ $B_j(q \rightarrow I_i Done(i,act)))$	(CR48)
reject	$Done(reject(i,j,Done(j,act)),q)) \rightarrow (q \rightarrow I_j I_i Done(i,act)) \rightarrow$ $(B_i \neg I_i Done(i,act) \wedge B_j \neg I_i Done(i,act))$	(CR49)
propose	$Done(propose(i,j,Done(j,p)),q)) \rightarrow (Done(i,q) \rightarrow I_i I_j Done(j,p)) \wedge$ $(B_i(s \rightarrow I_i Done(i,q)) \wedge$ $B_j(s \rightarrow I_i Done(i,q))$ where $s \equiv B_i(Done(i,q) \rightarrow I_j(Done(j,p))) \wedge B_j(Done(i,q) \rightarrow I_j(Done(j,p)))$	(CR50)

Table 10

Speech acts and the translation of commitments

FIPA, we also provide the rules for translating speech acts into mental attitudes; they are summarized in Table 10.

Even if the translation is straightforward, some consideration are needed:

– First of all, in contrast with the FIPA translation, speech acts affect both the beliefs and intentions of speaker and hearer and not only of the speaker. This represents the fact that a commitment is seen as a relationship, which is publicly established.
– Second, the interpretation of *promise* and *accept* in Fornara and Colombetti (2004) needs some discussion. According to Guerini and Castelfranchi (2007) the definition above is too weak: a *promise* needs to be explicitly accepted, and then kept. It is not enough that j has indicated to prefer i to do the action, and that i and j believe that i has an intention to do the action; what is missing is j's belief, as a result of the acceptance, that i will actually do it. Nevertheless, we think that the conditional nature of a promise, i.e., that it requires explicit acceptance, is well covered by the combination of rules (CR46) and (CR48). Rule (CR46) covers only an initiative to make a promise; the promise is only complete when accepted, and fulfilled when successful.
– A *request* introduces an unset commitment with the receiver as debtor, i.e., to perform the requested action (Rule CR40). It does not create any intention for the addressee until he has accepted it.
– *Accept* and *reject* change the state of an existing unset commitment to *pending* and *canceled* respectively. In order to account for this fact, we insert in the antecedent of the rules for *accept* and *reject* the reference to the configuration of beliefs and intentions that represent an existing commitment.
– A *propose* is a complex speech act composed by a *request* and a conditional *promise*; it introduces an unset commitment with the receiver as debtor and a pending commitment with the speaker as debtor. Since a *propose* is used in a negotiation, q and p refer respectively to an action of the speaker and of the receiver.

$$propose(i,j,Done(j,p),Done(i,q)) \equiv$$
$$request(i,j,Done(j,p),Done(i,q)); promise(i,j,Done(i,q),s) \tag{CR51}$$

where $s \equiv B_i(Done(i,q) \rightarrow I_j Done(j,p)) \wedge B_j(Done(i,q) \rightarrow I_j Done(j,p))$, i.e., p is a pending commitment of the receiver.

4.4. Example

We now return to the running example, to illustrate some of the details of our interpretation of the social commitment approach. We now use notation $\varphi \mid \psi$ for conditional commitments, as in the original. The result is shown in Table 11.

x-1:	$request(i, j, Done(j, dishes) \mid \top)$	$C(unset, j, i, Done(j, dishes) \mid \top)$
y-1:	$reject(j, i, Done(j, dishes) \mid \top)$	$C(canceled, j, i, Done(j, dishes) \mid \top)$
	$propose(j, i, Done(i, garbage) \mid Done(j, dishes))$	$C(unset, i, j, Done(i, garbage) \mid Done(j, dishes))$
		$C(pending, j, i, Done(j, dishes) \mid \sigma)$
		$\sigma \equiv C(pending, i, j, Done(i, garbage) \mid Done(j, dishes))$
x-2:	$accept(i, j, Done(i, garbage) \mid Done(j, dishes))$	$C(pending, i, j, Done(i, garbage) \mid Done(j, dishes))$
		$C(pending, j, i, Done(j, dishes) \mid \top)$ (resolving σ)

Table 11

The running example, in the social commitments approach

> x-1: Could you do the dishes?
> y-1: Only if you put the garbage out
> x-2: OK.

We start from state *empty*. According to Fornara and Colombetti (2004), the *request* from x in role r_1, $(xr_1 = i)$, creates an unset commitment waiting for acceptance. Now y in role $r_2, (yr_2 = j)$ responds with a conditional acceptance, which could be seen as a kind of counter proposal. However, there are two issues for discussion.

First, a *request* can only be accepted or rejected, but not be adjusted by a counter proposal. So just like in Section 3 we must interpret j's response here as rejection, followed by a conditional proposal.

Second, in Fornara and Colombetti's account, rendered in rule (CR50), conditional proposals for $\varphi \mid \psi$ are seen as a combination of a *request* for φ with a conditional *promise* to make sure that ψ.[2] The *promise* is again conditional on a statement σ which implies the conditional commitment $\varphi \mid \psi$ under discussion. A *promise* is in turn interpreted as combination of a *make* and a *set* action, resulting in the *pending* state. So under this interpretation a *promise* does not need acceptance[3].

Our interpretation of the dialogue above, would be as in Table 11. Remember that $i = xr_1, j = yr_2$ for some roles r_1, r_2. So we end up with the following set of commitments.

$$\{ C(pending, i, j, Done(i, garbage) \mid Done(j, dishes)), \ C(pending, j, i, Done(j, dishes) \mid \top) \}$$

This is a conditional commitment of i to put the garbage out, provided j does the dishes, and a commitment of j to indeed do to the dishes, given the context of this other commitment. Note that once we have resolved σ the conditional nature of the commitment of j to do the dishes, is lost.

Using the translation rules, in particular (CR41), we get the situation described in Table 12. Note that due to the public nature of commitments, as defined in terms of intentions and beliefs, it is not necessary to resort to the reliability rule like in Section 3.3. Consider for example the case of the *reject*, which is defined in terms of beliefs of both debtor and creditor.

---

[2]Thanks to the formalization of conditional proposal, it is possible to write $propose(j, i, Done(i, garbage) \mid Done(j, dishes))$; by contrast, in FIPA, we had to resort to the notion of joint plan to approximate the same notion, establishing an equivalence between $kitchen(i, j)$ and $dishes(j); garbage(i)$.

[3]We believe this interpretation of proposals is conceptually mistaken. First, proposals are concerned with actions to be done by the speaker; requests deal with actions to be done by the hearer. Second, implicit requests are better made explicit in the agent communication protocol. Third, promises do need acceptance. Nevertheless we continue with Fornara and Colombetti's interpretation here for the sake of the argument

	Role i	Role j
x-1: *request*	BELIEFS $\top \to I_i(I_j(Done(j, dishes)))$	BELIEFS
	INTENTIONS $I_j(Done(j, dishes))$	INTENTIONS
y-1: *reject*	BELIEFS $\top \to I_i(I_j(Done(j, dishes)))$ $\neg I_j(Done(j, dishes))$	BELIEFS $\neg I_j(Done(j, dishes))$
	INTENTIONS	INTENTIONS $\neg I_j(Done(j, dishes))$
y-1: *propose*	BELIEFS $\top \to I_i(I_j(Done(j, dishes)))$ $\neg I_j(Done(j, dishes))$ $s \to I_j(Done(j, dishes))$ where $s \equiv B_j(Done(j, dishes) \to I_i(Done(i, garbage))) \wedge$ $\qquad B_i(Done(j, dishes) \to I_i(Done(i, garbage)))$	BELIEFS $\neg I_j(Done(j, dishes))$ $Done(j, dishes) \to$ $\quad I_j(I_i(Done(i, garbage))))$ $s \to I_j(Done(j, dishes))$
	INTENTIONS	INTENTIONS $\neg I_j(Done(j, dishes))$
x-2: *accept*	BELIEFS $\top \to I_i(I_j(Done(j, dishes)))$ $s \to I_j(Done(j, dishes))$ $Done(j, dishes) \to I_i(Done(i, garbage))$ s $I_j(Done(j, dishes))$	BELIEFS $s \to I_j(Done(j, dishes))$ $Done(j, dishes) \to I_i(Done(i, garbage))$ s $I_j(Done(j, dishes))$
	INTENTIONS	INTENTIONS $Done(j, dishes)$

Table 12

The running example in the translated social commitments approach

5. Persuasion dialogues in the social commitments approach

The purpose of this section is to extend social commitment approaches to persuasion dialogues to show that our role-based semantics is general enough. A similar extension could be made for FIPA, but this extension is more interesting, due to the ambiguity of the term commitment.

We distinguish action commitments from propositional commitments, because an action commitment is fulfilled when the hearer believes that the action is performed, whereas a propositional commitment is fulfilled only when the hearer concedes to the proposition. Using the role-based semantics, we will show that propositional commitments are related to public beliefs, and action commitments to public intentions.

Consider for example the following two sentences:

(1) A promises B to deliver the goods before Friday.
(2) A informs B that Al Gore would have been a better president than Bush.

The first sentence commits agent A to the delivery of goods before Friday and the second sentence commits agent A to defend the argument that Al Gore would have been a better president than Bush. We say that the first sentence leads to an *action commitment* and the second sentence leads to a *propositional commitment*. Researchers in the social commitment approach to agent communication (Castelfranchi, 1995; Singh, 2000; Fornara and Colombetti, 2004) focus on the former, because they are interested in task oriented dialogues and negotiation. Researchers in the argumentation tradition on the other hand (Hamblin, 1970; Walton and Krabbe, 1995), focus on the latter:

"to assert a proposition may amount to becoming committed to subsequently defending the proposition, if one is challenged to do so by another speaker" (Walton and Krabbe, 1995).

Despite these differences, someone could argue that a promise and an assertion have the same effect: they create a commitment of the speaker, respectively an action commitment or a propositional commitment. This is coherent with the idea that the meaning of a speech act is defined in terms of the attitudes of the speaker, without any effects on the hearer. But according to Kibble (2005), the meaning of a speech act must also be defined in terms of the effects on the hearer. In particular, an assertion that goes unchallenged may count as a concession for the hearer. This corresponds to the 'silence means consent' principle, already studied by Mackenzie (1979). Walton and Krabbe argue that in case of a concession, the hearer becomes weakly committed to the proposition: the hearer can no longer make the speaker defend the proposition by challenging him, albeit the hearer does not have to defend the proposition himself if challenged. After an assertion the hearer can make a concession explicitly, or implicitly by not challenging the assertion. So we reinforce Kibble's claim that speech acts also have an effect on the attitudes of the hearer. To show the feasibility of the semantics, we model persuasion dialogues loosely inspired by the PPD_0 protocol of Walton and Krabbe (1995). We illustrate the approach by a dialogue that involves a mixture of propositional and action commitments. Despite the differences, action commitment and propositional commitment have more than accidental similarities too. This needs some argument, because

"... the word commitment and the corresponding verb commit are used in many different ways, and there is no reason to suspect that there is a common core meaning in all of them." (Walton and Krabbe, 1995, p 13).

Our general intuition is that through making a commitment, the number of future options becomes restricted. For action commitment, this can be explained by referring to the well-known slogan that "intention is choice with commitment" (Cohen and Levesque, 1990). Having the stability provided by commitments makes sense when re-planning is costly, and when certain resources must to be reserved in advance, e.g. time slots in an agenda (Bratman, 1987). The same holds for commitments made to other agents. For example, by agreeing to meet on Friday at noon, future options to do other things on Friday become restricted. Analogously, conceding to an action of another agent means agreeing not to prevent him from doing the action by making other conflicting commitments. For example, a concession to do an action which requires a car, implies that the conceder will not use the car for other purposes. The relation between commitment and concession is similar to the one between obligation and permission.

If we view argumentation or persuasion as a kind of game in which players make moves, a propositional commitment can also be said to limit the future possible moves a player can make. In particular, whenever an agent, the *proponent*, makes an assertion, he is committed to uphold that proposition. This means that all moves which would enable the other player, the *opponent*, to force the proponent to retract the proposition, must be avoided. Here too, commitments restrict the set of future options.

Social commitments are contrasted with the notion of commitment of Cohen and Levesque (1990) used in the definition of intention, due to the interactional character of social commitments. Thus, at first sight, our argument above seems to confuse the two levels. However, when we equate social commitment with either intention (for action commitment) or belief (for propositional commitment) it must be noted that we refer to public intentions and public beliefs, i.e., to the kind of mental attitudes attributed to roles. In this way, we capture both the traditional view of commitment as limiting options and the interactional nature of social commitments.

Just like one needs a specific logic of practical reasoning to explain action commitment, we need a specific persuasion protocol to explain propositional commitment. Here we will take a simple persuasion protocol. Although this protocol is simplified, we believe it is sufficient to illustrate the notion of propositional commitment. The protocol is depicted in Figure 5. The idea behind the protocol corresponds to the following quote, which is often cited to relate propositional commitment to a kind of action commitment:

"Suppose X asserts that P. Depending on context, X may then become committed to a number of things, for example, holding that P, defending that P (if challenged), not denying that P, giving
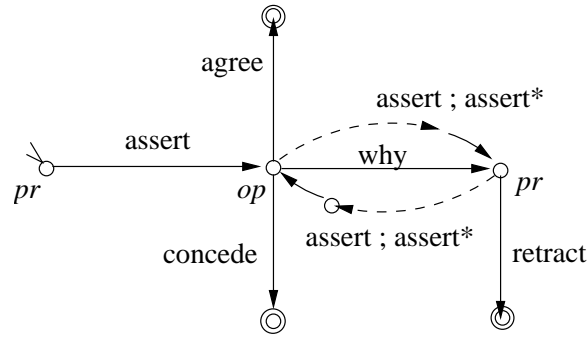
Fig. 5. A protocol for persuasion dialogues.

evidence that P, arguing that P, proving or establishing that P and so on" (Walton and Krabbe, 1995, p23).

The protocol is defined as follows. For each instantiation of the protocol, there are participants in two roles: the *proponent* (pr), and the *opponent* (op). By definition, the proponent is the agent who makes the initial assertion. Proponent and opponent have different burden of proof. Note that the agent who is the opponent of some proposition p, may very well become the proponent of another proposition q later in the dialogue. For this reason, it is crucial that we have an explicit representation of roles.

Following an assertion by the proponent, the opponent can respond by a challenge like a why-question, which essentially requests the proponent to come up with some arguments to support the assertion, or by a rebutting argument. The opponent may also agree with the asserted proposition, or concede the asserted proposition. Agreeing means not only relieving the proponent from the burden of proof, but also adapting the burden of proof towards third parties. Conceding means only that the opponent has given up the right to challenge the proposition, and has relieved the proponent of the burden of proof. Thus agreement implies concession. Alternatively, we can say that not challenging means a concession. In this way we capture Kibble's (2005) idea that entitlement to a commitment comes by default after an assertion.

In response to a challenge, the proponent must give some argument in support of the proposition, which itself consists of one or more assertions, or retract the original assertion. Following the assertion in response to a challenge, the opponent may either concede the proposition, or challenge or concede any of the assertions made by the proponent during the argument in support.

A why-challenge does not add any new material. But an opponent can also challenge propositions by so called rebutting arguments, that provide an independent argument for the opposite assertion. Rebutting arguments can be handled as assertions by the opponent, which trigger another instantiation of the same protocol, with a role reversal. So for rebutting arguments, the burden of proof lies with the rebutter.

The dialogue ends either when the proponent has run out of arguments to support his assertion; in this case he is forced to retract the proposition, or when the opponent has run out of challenges; in this case he is forced to concede. An agreement of the opponent, is essentially an assertion by the opponent. This end condition corresponds to Walton and Krabbe's win and loss rules of PPD_0 persuasion dialogues (Walton and Krabbe, 1995, p.152).

5.1. Propositional commitment

In our formalization of the persuasion protocol, propositional commitments of the proponent are modeled as public beliefs. Open challenges are modeled as public negated belief of the opponent. In the definition of the update rules in Table 13, the Before and After fields indicate respectively the condition necessary for the utterance to have an effect on the public beliefs, and the effect that results. They should not be confused with the FP and RE conditions of FIPA.

In the translation, we use the principle that a concession PP is represented by the fact that the debtor i does not believe that $\neg p$. This makes it impossible for the debtor to raise further challenges.

Before	speech act	After
$--$	$assert(pr, op, p)$	$B_{pr}p$
$B_{pr}p$	$why(op, pr, p)$	$\neg B_{op}p$
$B_{pr}p$	$concede(op, pr, p)$	$\neg B_{op}\neg p$
$B_{pr}p$	$agree(op, pr, p)$	$B_{op}p$
$B_{pr}p, \neg B_{op}p$	$retract(pr, op, p)$	$\neg B_{pr}p$

Table 13

Updates rules in Persuasion protocol; *pr* is the proponent and *op* the opponent, for $x, y \in A$ and $s \in S$.

$$PP(i, j, p) \equiv \neg B_i \neg p$$

As usual, i and j are the role instances associated with the two interacting agents.

A propositional commitment PC is active when the debtor i believes the proposition, while nothing is required of the creditor j:

$$PC(active, i, j, p) \equiv B_i p \tag{CR52}$$

A propositional commitment PC is fulfilled when the creditor concedes the proposition, and therefore cannot challenge it anymore:

$$PC(fulfilled, i, j, p) \equiv \neg B_j \neg p \tag{CR53}$$

A propositional commitment PC is violated when the debtor's beliefs are in contradiction, due to the failure of defending some previous commitment.

$$PC(violated, i, j, p) \equiv \neg B_i p \wedge B_i p \tag{CR54}$$

Note that a proper treatment of this issue requires a detailed mechanism for dealing with temporal issues which is missing in FIPA-SL. It is not clear whether a conditional propositional commitment is different from a propositional commitment about a conditional, and it is also unclear what it means for a propositional commitment to be *unset* or *pending*. Thus we do not define such states here, nor do we define cancelation, which in addition requires to introduce time.

5.2. Speech acts

Given the definition of the commitment states in terms of the mental states of the roles, we can provide the following translation of the speech acts semantics. Speech acts affect both the beliefs and intentions of speaker and hearer and not only of the speaker. This represents the fact that in social commitments agents are publicly committed to the mental attitudes attributed by the constitutive rules to the roles they play. No cooperativity or sincerity assumptions are necessary, by contrast to FIPA.

In Table 14 we report the constitutive rules of the speech acts for persuasion:

– An assertion introduces an active propositional commitment of the speaker and, if it is not challenged, it also introduces a concession of the hearer: CR55, CR56.
– Both implicit and explicit concessions are modeled as the absence of the belief to the contrary. This is similar to weak commitment (Gaudou et al., 2006a, eq 17 p.128): CR56, CR59.
– Agreement simply means that the hearer becomes committed too. So agreement implies concession: CR60.
– Asserting an argument against p counts as a rebut-challenge. We simplify here for space reasons the notion of argument: CR57
– A why-challenge asks arguments to support the assertion. It indicates that the opponent is not yet convinced: CR61.
– Not replying to a why challenge with a supporting argument, counts as a retraction: CR62.
– Putting forward an argument in support of the original assertion is a way to reply to a why challenge: CR58.

assert	$Done(assert(i, j, p)) \rightarrow B_i p$	(CR55)
	$Done(assert(i, j, p)) \wedge (\neg Done(why(j, i, p)) \vee \neg Done(rebut(j, i, p))) \rightarrow \neg B_j(\neg p)$	(CR56)
	$B_j p \wedge Done(assert(i, j, (q \rightarrow \neg p) \wedge q)) \rightarrow Done(rebut(i, j, p))$	(CR57)
	$\neg B_j p \wedge Done(assert(i, j, (q \rightarrow \neg p)) \wedge q) \rightarrow Done(support(i, j, p))$	(CR58)
concede	$B_j p \wedge Done(concede(i, j, p)) \rightarrow \neg B_i(\neg p)$	(CR59)
agree	$B_j p \wedge Done(agree(i, j, p)) \rightarrow B_i(p)$	(CR60)
why	$B_j p \wedge Done(why(i, j, p)) \rightarrow \neg B_i p$	(CR61)
	$Done(why(i, j, p)) \wedge \neg Done(support(j, i, p)) \rightarrow \neg B_j(p)$	(CR62)

Table 14

Speech acts for persuasion

Replying to a rebut challenge with a counter argument is also compulsory, but because the rebut challenge is performed by means of a set of assertions, this is already accounted for by rule CR58.

Once a concession has been introduced it prevents the agent from committing itself to the opposite proposition, since this would lead to a contradiction: $\neg B_i(\neg p) \wedge B_i p$. So now i cannot assert $\neg p$ nor challenge p, since a challenge is performed by informing about an argument for $\neg p$. Avoiding a contradiction explains also why an agent would challenge an assertion, if he previously committed himself to the contrary.

5.3. Example

In Table 15 we show the interpretation of the following dialogue. For each turn, we report the beliefs and intentions which are created and those which persist from the previous turn.

Example 2

x-1:	Tomorrow the University is open.	$Done(assert(i, j, open))$
x-2:	Can you give the exams for me?	$Done(request(i, j, Done(j, give\text{-}exam)))$
y-1:	Isn't it closed for the Olympic games?	$Done(assert(j, i, (games \rightarrow \neg open) \wedge games))$
		$Done(rebut(j, i, open))$
x-3:	Not for exams.	$Done(assert(i, j, (exam \wedge games \rightarrow open) \wedge exam))$
		$Done(rebut(i, j, (games \rightarrow \neg open) \wedge games))$
y-2:	I see.	$Done(agree(j, i, open))$
y-3:	OK	$Done(accept(j, i, Done(j, give\text{-}exam)))$
y-4:	[j gives the exam]	$Done(j, give\text{-}exam)$

After i's assertion, if j would not reply, j would be publicly considered to concede that the university is open, by rule CR56:

$$\neg(Done(why(j, i, open)) \vee Done(rebut(j, i, open))) \rightarrow \neg B_j(\neg open)$$

So j must reply by doing a *rebut* challenge, after which i will get stuck in the contradiction, if he does not defend $B_i(open \wedge \neg open)$, due to rule CR56:

$$\neg(Done(why(i, j, (games \rightarrow \neg open) \wedge games)) \vee$$
$$Done(rebut(i, j, (games \rightarrow \neg open) \wedge games))) \rightarrow \neg B_i(open)$$

Since i replies with a rebutting argument, j will implicitly concedes if he does not reply (CR56):

$$\neg(Done(why(j, i, (exam \wedge games \rightarrow open) \wedge exam)) \vee$$
$$Done(rebut(j, i, (exam \wedge games \rightarrow open) \wedge exam))) \rightarrow \neg B_j(\neg open)$$

However, in y-2, he agrees with i and adopts his belief.

	Proponent i	Opponent j
x-1: *assert*	BELIEFS	BELIEFS
x-2: *request*	*open*	
	INTENTIONS	INTENTIONS
	$I_j Done(j, \textit{give-exam})$	
y-1: *assert*	BELIEFS	BELIEFS
rebut		$(games \rightarrow \neg open) \wedge games)$
		$\neg open$
	open	
	INTENTIONS	INTENTIONS
	$Done(j, \textit{give-exam})$	
x-3: *assert*	BELIEFS	BELIEFS
rebut	$(exam \wedge games \rightarrow open) \wedge exam)$	
	open	$\neg open$
	INTENTIONS	INTENTIONS
	$Done(j, \textit{give-exam})$	
y-2: *agree*	BELIEFS	BELIEFS
y-3: *accept*	*open*	*open*
	INTENTIONS	INTENTIONS
	$Done(j, \textit{give-exam})$	$Done(j, \textit{give-exam})$
y-4: *give-exam*	BELIEFS	BELIEFS
	$Done(j, \textit{give-exam})$	$Done(j, \textit{give-exam})$

Table 15

The interpretation of the dialogue in Example 2.

6. Comparisons

In this section we give two examples of a comparison between FIPA and social commitments using our role semantics, to assess the feasibility of the role semantics as an intermediate language. We want to illustrate that no information gets lost in the translation from the social commitment approach to the beliefs and intentions of roles.

6.1. Comparing Inform

We first consider the case of *inform*. We show that by first translating a state described in terms of commitments into public beliefs and intentions, and then applying a speech act as defined in Role SL, we get the same results as when we first apply the social commitments version of a speech act, and then translate the resulting state in public beliefs and intentions. The general idea is shown in Figure 6.

Suppose we have a state s_1 with no commitments: state *empty*. According to Fornara and Colombetti (2004), the meaning of an *inform* is the combination of a *make* and a *set* of an unconditional commitment. The result is:

$$C(pending, i, j, \varphi \mid \top)$$

Because it concerns an unconditional commitment, there is no difference between active and pending commitments. So we use 'jump condition' rule number 5 (See Table 4 in Section 4). We also leave the condition '$\mid \top$' out of the notation. That produces:

$$C(active, i, j, \varphi)$$

Now we translate according to rule (CR42), which produces

$$B_i \varphi$$

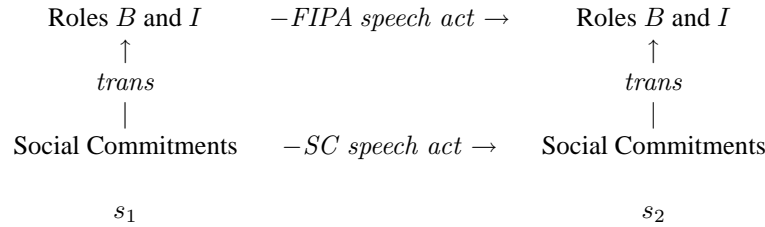So the translation of s_2 becomes $\{B_i \varphi\}$. That completes the first half of the diagram.

$$\begin{array}{ccc}
\text{Roles } B \text{ and } I & -FIPA\ speech\ act \rightarrow & \text{Roles } B \text{ and } I \\
\uparrow & & \uparrow \\
trans & & trans \\
| & & | \\
\text{Social Commitments} & -SC\ speech\ act \rightarrow & \text{Social Commitments} \\
& & \\
s_1 & & s_2
\end{array}$$

Fig. 6. Comparison diagram.

For the role translation, the empty state corresponds to a set of no beliefs or intentions: {}, Now, the meaning of the *inform* speech act, is captured by rules (CR17) – (CR19). So, in the role version of FIPA, s_2 becomes

$$\{B_i\varphi, \neg B_i(B_j\varphi \vee B_j\neg\varphi), I_i B_j\varphi\}$$

Clearly, this is more detailed than the translation of the social commitments approach. The FIPA version contains presupposed information about what the speaker does not know beforehand, and about hist intentions. However, we *can* say that going from social commitments to roles, no information is lost.

The difference in the amount of detail between the approaches, becomes even more clear when we compare the outcomes of the garbage example, discussed at the end of Section 3 and Section 4. Here too the social commitments approach is less elaborate.

But in addition to the different level of detail (feasibility preconditions vs operational semantics), there are also important conceptual differences in the way proposals are modeled.

6.2. Propose Interaction Protocol

As a second example, we have chosen the Propose interaction protocol of FIPA (2002a). This protocol consists of a *propose* followed by an acceptance or a refusal.

For a specific instantiation of the protocol, we compare the set of beliefs and intentions produced by the translation of FIPA and social commitments into the role-based semantics to assess whether the intentions concerning executable actions are the same in the two approaches – the agents would act in a similar fashion – and whether it is possible to find in FIPA the same commitments as in the social commitments approach.

The main difficulty in mapping FIPA onto social commitments concerns the communicative act *propose*. In social commitments approaches, it is viewed as a way to negotiate a joint plan: "If I do q, then you do p". This can model auctions for example (Fornara and Colombetti, 2004). Instead, the FIPA definition of *propose* refers to one action only.

Example 3

x-1:	*propose*	This bike costs only $50!
y-1:	*accept*	All right.
x-2:	*give*	[gives bike]
y-2:	*pay*	[pays $50]

| y-1′: | *reject* | No, that is too much. |

We are inspired by the example in FIPA (2002a), which reports the action of selling an item for a given amount of money. Note that the action of selling is conceived of as a joint action composed of giving the item and paying the money, as in Section 3.3: $sell(i, j) \equiv give(i); pay(j)$.

We could say that an accepted *propose* act forms a plan to be performed by both agents. Once the intentions of both agents to perform the plan have been formed, the plan is distributed between the agents

	Sender i	Receiver j
x-1: *propose*	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$
	INTENTIONS $B_j(I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j)))$	INTENTIONS
y-1: *accept*	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ $I_j Done(sell(i,j))$ $I_i Done(sell(i,j))$ $I_i Done(give(i))$ $Done(give(i)) \rightarrow I_i Done(pay(j))$ $Done(give(i)) \rightarrow I_j Done(pay(j))$	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ $I_j Done(sell(i,j))$ $I_j Done(give(i))$ $Done(give(i)) \rightarrow I_j Done(pay(j))$
	INTENTIONS $Done(give(i))$	INTENTIONS $B_i I_j Done(sell(i,j))$
x-2: *give*	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ $Done(give(i)) \rightarrow I_j Done(pay(j))$ $Done(give(i)) \rightarrow I_i Done(pay(j))$ $I_i Done(give(i))$ $Done(give(i))$ $I_j Done(pay(j))$ $I_i Done(pay(j))$	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ $Done(give(i)) \rightarrow I_j Done(pay(j))$ $I_j Done(give(i))$ $Done(give(i))$ $I_j Done(pay(j))$
	INTENTIONS	INTENTIONS $Done(pay(j))$
y-2: *pay*	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ $Done(give(i)) \rightarrow I_j Done(pay(j))$ $Done(give(i)) \rightarrow I_i Done(I_j Done(pay(j)))$ $I_j Done(pay(j))$ $Done(pay(j))$	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j)))$ $Done(give(i)) \rightarrow I_j Done(pay(j))$ $I_j Done(pay(j))$ $Done(pay(j))$
	INTENTIONS	INTENTIONS
y-1′: *reject*	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ $\neg I_j Done(sell(i,j))$	BELIEFS $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ $\neg I_j Done(sell(i,j))$
	INTENTIONS	INTENTIONS $B_i \neg I_j Done(sell(i,j))$

Table 16

Example 3 with FIPA.

according to rules RL38 and RL39 which we introduced in Section 3.3, and the intentions concerning the steps of the plan are formed. By contrast, if the proposal is not accepted, as in y-1′, no intentions are formed according to the plan. The remainder of the translation is straightforward: at each turn, the rules for translating the semantics of FIPA and social commitments into the role-based semantics are applied (see Sections 3 and 4). Then, *modus ponens* and the rules are applied. We assume here that beliefs and intentions which are not affected by subsequent speech acts persist.

6.3. Example in FIPA

In FIPA (see Table 16), the proposal to sell *propose*$(i, j, Done(sell(i, j))$ is an *inform* that introduces in the role i the belief that the precondition $I_j Done(sell(i,j)) \rightarrow I_i Done(sell(i,j))$ is true. We skip for space reasons the other feasibility precondition, but the reader can easily check that it is true and consistent

with the state of the dialogue. The rational effect is an intention of the speaker, but since the speaker is reliable (it has correct beliefs about its own mental states), after the proposal, the receiver believes $I_j Done(sell(i, j)) \rightarrow I_i Done(sell(i, j))$ too.

The acceptance of the proposal by j in FIPA is an *inform* that j has the intention $I_j Done(sell(i, j))$. Again, the receiver believes the content of the *accept proposal* speech act because an agent in a role is reliable about its own mental states. Since the speaker believes $I_j Done(sell(i, j)) \rightarrow I_i Done(sell(i, j))$ and $I_j Done(sell(i, j))$, it believes also to have $Done(sell(i, j))$ as a intention (by *modus ponens*) and, by rule (RL1), it actually has the intention to sell. Most importantly, if an agent in a role has the intention to make a joint plan, by the rule RL38, then it has the intention to do its part at the right moment (the others know this) and the intention that the others do their part. The result of the distribution is: $B_i(Done(give(i)) \rightarrow I_i Done(pay(j))) \wedge B_j(Done(give(i)) \rightarrow I_j Done(pay(j)))$. Thus, when $Done(give(i))$ is true, we derive that both i and j intend j to do its part: $I_i Done(pay(j))$. Subsequently, this intention is carried out.

However, if the proposal is rejected, as in y-1′, this means that j has no intention for the joint plan $sell(i, j)$: $\neg I_j Done(sell(i, j))$. Consequently, no distribution of intentions is possible, and no individual intention to $give$ is formed by i.

6.4. Example in Social Commitments

The translation from social commitments ACL semantics to the role-based semantics is accomplished by applying the rules defined in the previous section (see Figure 17). Given the FIPA *propose* speech act, the corresponding speech act in social commitments ACL is $propose(i, j, Done(give(i)), Done(pay(j)))$. By applying the rule that translates this speech act in the role-based ACL semantics, we get to the state in which both i and j have the belief that $Done(give(i)) \rightarrow I_i I_j Done(pay(j))$ representing an unset commitment of j. Moreover, a pending commitment by i is represented by $s \rightarrow I_i Done(give(i))$, where $s \equiv B_i(Done(give(i)) \rightarrow I_j Done(pay(j))) \wedge B_j(Done(give(i)) \rightarrow I_j Done(pay(j)))$.

The *accept proposal* is modeled by $accept(j, i, C_i(unset, j, i, Done(pay(j)) \mid Done(give(i))))$ in social commitments semantics. This speech act, whose precondition is true, results in j's act of creating the belief of i that j believes $Done(give(i)) \rightarrow I_j Done(pay(j))$, thus making the condition s true. The application of *modus ponens* to the belief $s \rightarrow I_i Done(give(i)))$ and this new belief results in the introduction of an active commitment whose debtor is role i: $I_i Done(give(i)) \wedge I_j Done(give(i))$ When $give(i)$ is executed, the commitment of i to do $give(i)$ is fulfilled and the commitment of j to do $pay(j)$ is active: its condition $Done(give(i))$ is satisfied.

The *reject proposal* communicative act in FIPA in social commitments semantics corresponds to the speech act $reject(j, i, C_i(unset, j, i, Done(pay(j)) \mid Done(give(i))))$. The reject speech act attributes to both i and j the belief that j does not have the intention $Done(pay(i))$, thus retaining i's pending commitment from becoming active and canceling the unset commitment from j.

6.5. Comparison

What are the main differences between these approaches? Table 16 and 17 show that, once translated into the role-based semantics, the resulting intentions of the two approaches coincide, with one significant exception.

The difference can be observed in the first row. After the *propose* act there is no equivalent in FIPA of the belief – publicly attributed to the proponent – that it has the intention that the addressee forms a conditional intention to pay the requested amount of money for the sold item, where the condition consists of the proponent giving the item. Such a belief can be found in the social commitments approach.

This difference is due to the definition of proposing in FIPA. FIPA does not express the advantage of the proposer in proposing the plan. For example, in the selling case, there is no clue of the reason why the joint plan is proposed, namely to receive the specified money. This is left implicit in the definition of the selling act. In social commitments approaches, reciprocity is expressed explicitly by the fact that a *propose*

	Sender i	Receiver j
x-1: *propose*	BELIEFS	BELIEFS
	$Done(give(i)) \rightarrow I_i I_j Done(pay(j))$	
	$s \rightarrow I_i Done(give(i))$	$s \rightarrow I_i Done(give(i))$
	where $s \equiv B_i(Done(give(i)) \rightarrow I_j Done(pay(j))) \wedge$	
	$\qquad B_j(Done(give(i)) \rightarrow I_j Done(pay(j)))$	
	INTENTIONS	INTENTIONS
y-1: *accept*	BELIEFS	BELIEFS
	$Done(give(i)) \rightarrow I_i I_j(pay(j))$	
	$s \rightarrow I_i Done(give(i))$	$s \rightarrow I_i Done(give(i))$
	$Done(give(i)) \rightarrow I_j Done(pay(j))$	$Done(give(i)) \rightarrow I_j Done(pay(j))$
	s	s
	$I_i Done(give(i))$	$I_i Done(give(i))$
	INTENTIONS	INTENTIONS
	$Done(give(i))$	
x-2: *give*	BELIEFS	BELIEFS
	$Done(give(i)) \rightarrow I_i I_j Done(pay(j))$	
	$s \rightarrow I_i Done(give(i))$	$s \rightarrow I_i Done(give(i))$
	$Done(give(i)) \rightarrow I_j Done(pay(j))$	$Done(give(i)) \rightarrow I_j Done(pay(j))$
	$I_i Done(give(i))$	$I_i Done(give(i))$
	$Done(give(i))$	$give(i)$
	$I_j Done(pay(j))$ [MP]	$I_j Done(pay(j))$
	INTENTIONS	INTENTIONS
		$Done(pay(j))$
y-2: *pay*	BELIEFS	BELIEFS
	$I_j Done(pay(j))$	$I_j Done(pay(j))$
	$Done(pay(j))$	$Done(pay(j))$
	INTENTIONS	INTENTIONS
y-1′: *reject*	BELIEFS	BELIEFS
	$Done(give(i)) \rightarrow I_i I_j Done(pay(j))$	
	$s \rightarrow I_i Done(give(i))$	$s \rightarrow I_i Done(give(i))$
	$\neg I_j Done(pay(j))$	$\neg I_j Done(pay(j))$
	INTENTIONS	INTENTIONS

Table 17

Example 3 with social commitments.

is composed of a conditional *promise* together with a *request* (see also the model in (Yolum and Singh, 2002)). This provides a way to express any kind of arrangement, even non conventional ones. In social commitments approaches, the subsequent acceptance presupposes the existence of an unset commitment having as debtor the participant to which the proposal was addressed. However, FIPA's *accept proposal* in the second turn fills the gap: when the addressee displays the intention to take part in the joint plan, the distribution of the tasks of giving and paying takes place, generating the appropriate intentions for the two roles.

7. Related work

Here we discuss the historical development of work on the semantics of agent communication languages, work on roles, as well as our own related work.

7.1. Agent communication languages

The linguistic and philosophical theory of speech acts has inspired a number of models of agent communication languages. Most models of communicative acts can be classified into two types, depending

on whether they rely on mental attitudes or social commitment, including both action commitment and propositional commitment.

7.1.1. Mental attitudes

The semantics of agent communication languages provided by the Foundation for Intelligent Physical Agents (FIPA, 2002a), is paradigmatic of the models based on mental attitudes. The FIPA standards and the work at France Telecom on which it is based (Sadek, 1992; Bretier and Sadek, 1997; Sadek et al., 1997), are based on earlier models for natural language interpretation (Cohen and Perrault, 1979; Allen and Perrault, 1980). FIPA has concentrated most on collaborative dialogue types, like information seeking. The sincerity condition assumed by FIPA makes sense for natural language dialogue, because such dialogue agents are designed to be cooperative. However, the generalization of the FIPA language to non-cooperative settings has proven to be problematic. The combination of speech act theory (Searle, 1969), with a general theory of planning and action, e.g. (Pollack, 1990), provides a general model of rational interaction. Essentially, the speech acts are treated as any other action, which can be combined into plans by means of their preconditions and intended effects. In FIPA, communicative acts are defined in terms of the mental states of the agent who issues them. Mental states are represented according to the well known Belief-Desire-Intention (BDI) framework (Bratman, 1987; Cohen and Levesque, 1990). The bridge between the communicative acts and the behavior of agents is provided by the notions of rational effect (RE) and feasibility preconditions (FP). The rational effect is the mental state that the speaker intends to bring about in the hearer by issuing a communicative act, and the feasibility preconditions encode the appropriate mental states for issuing a communicative act. To guarantee communication, the framework relies on intention recognition on the part of the hearer.

Despite the fact that FIPA standards are well known, they are not yet widely used in agent applications. One of the reasons might be that it is counterintuitive to generalize the FIPA approach to non-cooperative settings. By definition, the private mental states of agents cannot be inspected, and can therefore not be verified (Pitt and Mamdani, 1999; Wooldridge, 2000). Such an approach is not appropriate in situations in which agents may be insincere or non-cooperative, as in argumentation or negotiation settings. Moreover, a typical agent application involves stake holders with diverse or opposed interests. Typical agent applications will therefore often involve non-cooperative settings.

7.1.2. Social commitments: action commitment in negotiation

Agent communication languages based on social commitment constitute an attempt to overcome the mentalistic assumptions of FIPA by restricting the analysis to public communication (Castelfranchi, 1995; Singh, 2000; Colombetti et al., 2004; Fornara and Colombetti, 2004; Bentahar et al., 2004). Communicative acts are defined in terms of the social commitments they publicly determine for speaker and the hearer. According to Fornara and Colombetti (2004) commitment is "a social relationship between the speaker and the hearer". The notion of social commitment does not refer to intentionality; it is a primitive social notion, developed for modeling mutual commitments in contracts and electronic commerce, for example. A commitment in Fornara and Colombetti (2004) has a *debtor* and *creditor*, i.e., respectively, the agent who makes the commitment, and the agent to which the commitment is made. We could say that the creditor is interested in the content of the commitment. The social commitments approach has focussed on non-cooperative types of dialogue, like negotiation, but not on persuasion. Researchers working in the tradition of Winograd and Flores (1986), such as Singh (2000), are interested in cooperative dialogue and negotiation. Hence they tend to investigate the action commitments, that typically result from directives like *request* and *propose*.

The social commitment approach also has a number of drawbacks. A social commitment describes the status of an agreement about future action. Unlike intentions, as usually conceived, social commitments do not have a causal relation to action. So the use of social commitments to model the effects of communication relies on the notion of obligation to explain how commitments affect the behavior of the individual agents. As it appears, commitments require enforcement mechanisms of obligations, like sanctions (Pasquier et al., 2004). This makes sense for competitive environments, like argumentation dialogue or negotiation, but it does not make sense in cooperative environments, like information seeking or inquiry,

where a commitment can simply be interpreted as an expectation. Even though social commitments may be useful in the analysis and design of communication protocols for applications like electronic commerce, this does not mean that they provide an adequate semantics of the speech acts involved.[4] The social semantics is an operational semantics, which translates speech acts to transitions in a data-structure, the commitment automata.

7.1.3. Social commitments: propositional commitment in persuasion

The argumentation tradition originating with Hamblin (1970) and Walton and Krabbe (1995), focuses on propositional commitments: "to assert a proposition may amount to becoming committed to subsequently defending the proposition, if one is challenged to do so by another speaker" (Walton and Krabbe, 1995). However, this kind of propositional commitment is biased towards argumentation dialogue, with assertive speech acts like 'assert' or 'challenge', failing to be general enough for cooperative dialogues.

In many social semantics the distinction between the commitment to an action and the propositional commitments of assertives becomes blurred. This is a problem when dealing with argumentation and persuasion dialogues where the notion of propositional commitment is basic. According to Bentahar et al. (2004), the difference between an action commitment and a propositional commitment lies only in the type of content, and both kinds of commitment are fulfilled (or violated) if the content is true (false) in the world, albeit the debtor cannot do anything to make a propositional commitment true, whereas he can perform the action object of a commitment. At a closer analysis the definition of fulfillment is not correct for action commitment. According to these authors, a commitment is fulfilled if (at the deadline time) its content is true in the world. This objectivistic solution is too weak: fulfillment does not only depend on what is true in the world, but also on what is believed by the creditor. Thus the creditor can still claim to be entitled to the commitment, until he is convinced and the evidence is shared by both agents. Moreover, this view of fulfillment is not realistic for propositional commitment. For propositional commitments the problem is made worse by the fact that their content is not restricted to actions whose execution can be monitored in the world. Consider, e.g., a commitment towards the fact that Al Gore would have been a better president than Bush. There is no way to fulfill such a hypothetical commitment, unless one of the agents concedes on the basis of arguments. This is the general case in persuasion and argumentation, e.g., in a political debate or in a trial. An alternative solution is to define that a propositional commitment it fulfilled when the creditor also becomes committed to the proposition. But this is too strong, since not all assertions and informs aim to make the hearer believe them, and viceversa, not all informs and assertions aim to satisfy a goal of the creditor to know information. For example, in information seeking, the creditor wants to have reliable information but, in a dispute, he wants to win. In general, and in contrast with action commitment, there is not always a creditor who has the goal to have some information. In our approach, a propositional commitment is fulfilled when the hearer concedes, and, thus, cannot challenge the proposition anymore. Other possible goals, like that the hearer must come to believe the proposition, depend on particular types of dialogues and have their own fulfillment conditions. Our role-based semantics allows to associate different fulfillment conditions to the roles that belong to different dialogue games.

The social semantics approach – with the noticeable exception of Kibble (2005) – generally only makes the speaker committed to the propositional content of an assertive. It do not express the notion of communicative intention: an *inform* makes mutually believed the intention of the speaker that the addressee believes the proposition. Social semantics delegate to the protocol level the coherence of a dialogue, by dropping communicative intentions from the speech act level. Concerning concessions, inspired by Kibble (2005), we argue that an assertion creates by default a concession of the hearer, which can be in contradiction with his beliefs, unless he challenges the assertion. When the hearer challenges the assertion, the contradiction is passed to the speaker, and forces a retraction of the speaker, if not challenged in turn; and so on, until one of them does not have any arguments left to put forward, and concedes or retracts explic-

---

[4]We are aware of the difference between semantics in computer science, and in linguistics. In linguistics, a semantics of an agent communication language would rather be called a pragmatics, because it mainly concerns the correct use of these acts. See the insightful discussion in Coleman et al. (2006).

itly. Finally, considering the creation of a commitment as a result of an assertion, as social commitments semantics does, is different from the traditional interpretation, which sees propositional commitments as a kind of action commitment to defend the proposition. Instead, a commitment to defend, is typically created by the challenge.

7.1.4. Bridges

Recently, some other papers went in the same direction of redefining FIPA semantics, using public mental states e.g., (Nickles et al., 2006; Verdicchio and Colombetti, 2006; Gaudou et al., 2006a). However, these other solutions need to add to dialogue new theoretical concepts which are not always completely clear or diverge from existing work. In particular, Gaudou et al. (2006a) use an explicit grounding operator, which only partially overlaps with the tradition of grounding in theories of natural language dialogue. Opinions (Nickles et al., 2006) are introduced specifically for modeling dialogue, but with no relation with persuasion and argumentation. Finally, commitments in Verdicchio and Colombetti (2006) roughly overlap with obligations. Moreover, these other approaches relate to the well known FIPA semantics in different degrees: Gaudou et al. (2006a) and Nickles et al. (2006) try to stay close to the original semantics, as we do, while Verdicchio and Colombetti (2006) substitute it entirely with a new semantics, which, among other things, does not consider preconditions of actions.

7.1.5. Other work in agent communication

In this paper we do not consider the construction of communicative protocols out of individual speech acts. There exist first principle methodologies, in which the semantics of a protocol is defined completely in terms of the semantics of the individual speech acts. Alternatively, larger protocols may be construed from basic protocols consisting of speech acts that naturally belong together. For example, *propose*, *accept* and *reject* form such a basic interaction protocol, see Section 4.

7.2. Roles

The distinguishing feature of our approach is that the public mental attitudes attributed to agents during the dialogue are associated with roles. he importance of roles is recognized in multiagent systems and their function ranges from attributing responsibilities to assigning powers to agents in organizations. In this paper we exploit the notion of roles, as introduced in agent-oriented software engineering, e.g., GAIA (Wooldridge et al., 2000), TROPOS (Bresciani et al., 2004). We distinguish interactive roles, such as speaker, (over)hearer and addressee. Clearly, different constitutive rules apply to speaker and hearer. Further, we could add rules so that the effects of implicit acknowledgement differ between the addressee of a message, and a mere overhearer (Gaudou et al., 2006b). We also distinguish social roles, that belong to a particular type of dialogue. E.g., in argumentation, the 'burden of proof' is different for the proponent or the opponent of a claim. Or consider the fact that a *request* cannot be refused, if there is some authority relation between speaker and addressee. In fact, such a relationship turns a *request* into an order. Note that one agent can play different roles in different interactions with the same or different interlocutors. For example, the buyer of a good may become a seller in a second transaction. Because social roles are associated with dialogue types, each with a specific set of dialogue rules, roles allow us to reason about assumptions in different kinds of dialogues. Thus, sincerity could be assumed in cooperative dialogues, such as information exchange, but not in non-cooperative dialogues, such as persuasion or negotiation. Ostensible beliefs and the grounding operator can only distinguish interactive roles like speaker and addressee, and the effect on different groups of agents.

7.2.1. Roles as prescriptions

Roles in a social institution are traditionally used to determine the obligations, permissions and institutional powers of an agent. Roles prescribe which possible speech acts are allowed for an agent, and possibly which acts must be used to respond. This idea is most prominent in the ISLANDER system (Esteva et al., 2002), in which the roles of agents determine the interface with the environment. In our role model (Boella and van der Torre, 2004) roles are always associated to some kind of institution, see also e.g.

Kagal and Finin (2005), which is described by constitutive rules. In the case of dialogue, the institution is represented by the type of dialogue game played by the participants.

In the role-based semantics introduced in this paper, the moves available to the dialogue participants are based on the mental attributes of the roles only, not on the mental attitudes of the agent. Since an agent enters in a dialogue only in a certain role, the communicative actions at his disposal, depend on the role. Thus, agents participating in a dialogue in different roles can perform different kinds of communicative actions. This is due to the fact that the communicative action performed depends on the constitutive rules of the dialogue game: if an agent utters a sentence which is not recognized as a communicative action since the agent is not playing the right role, the communicative action is considered not to be performed.

A typical situation is represented by the Contract Net Protocol (Smith, 1980) where the initiator role and the participant role can perform different actions, e.g., call for proposal and proposal respectively. The initiator and participant role are already present in the Contract Net Protocol. The distinguishing property of our approach is that roles are not simply labels, but they are associated with instances representing the state of the participant in the interaction. The state is represented as a set of beliefs and intentions attributed to the role enacting agent. Such a state is modified by the communicative actions performed during the dialogue according to the constitutive rules of the dialogue game.

Each instance of a dialogue game is associated with instances of the roles played by the agents; so each agent is associated with a different state of the interaction in each dialogue he is participating in. For example, an agent who is playing the role of participant in a Contract Net negotiation, can at the same time participate as initiator in another negotiation to subcontract part of the task. In each of his roles the agent is associated with a set of beliefs and intentions representing the situation of the conversation thus far. Moreover, the price the agent as initiator can pay to a sub-contractor, depends on the price for which it undertook the task, as participant. So all the roles must be related to a common agent, who has to direct the different negotiations according to his private reservation price and to the outcome of the other interactions.

7.3. Roles as expectations

Agents can make predictions, and use these predictions to coordinate their behavior with an agent, due to the fact that the agent enacts a particular role. This predictive aspect is common in the social sciences, made famous by the restaurant script (Schank and Abelson, 1977) and emphasized in agent theory by Castelfranchi (1998). An example from human life, is the fact that the car of someone taking driving lessons, is clearly marked with a sign, like 'L' or 'E'. This sign does not change the prescriptive status – the traffic code applies just as much – but it signals to other drivers to be careful and more considerate.

In our role-based semantics, expectations are based both on the mental attitudes ascribed to the agent and to the role. To play a role, an agent is expected to act *as if* the beliefs and intentions of the role were his own, and to keep them coherent, as he does for his own mental attitudes. An agent should adopt the role's intentions and carry them out according to the beliefs of the role. This holds despite the fact that the model remains neutral with respect to the motivations that agents may have when playing a role; agents can adopt the mental attitudes attributed to their roles as a form of cooperation, or they can be publicly committed to their roles. The roles' attitudes represent what the agent is publicly held responsible for: if the agent does not adhere to his role, he can be sanctioned or blamed.

Expectations follow also from the objectives of the agents. For agent communication languages, the most convincing example has to do with bargaining. In protocols like the Contract Net, listed above, there are no constraints on the content of proposals. However, purely based on the apparent objectives of agents when entering a conversation in a particular role, we can infer a number of preferences:

– the initiator wants to have some task achieved, otherwise he would not send the call for proposals
– the initiator wants to give up as little as possible, in return for the accomplishment of the task
– a participant may either want to achieve the task, or not. In the first case, he will send some offer. In the last case he will send a reject, or fail to reply.

   – if a participant is interested, he will want as much as possible in return for doing the task

From such preferences, we can infer some coherence conditions on the content of proposals. For example, it will be very unlikely that a participant will first offer to do the task for 40, and later to do the task for 50, because the participant does not expect the initiator to accept the higher offer, after the lower offer was declined. More about coherence conditions can be found in Hulstijn et al. (2004).

7.4. Our own earlier work

This journal paper is a revised and extended version of various workshop and conference papers. Boella et al. (2005, 2006b) discuss a role-based semantics, which can deal with both mentalistic approaches and social commitments. Of this work, Boella et al. (2005) is based on our normative multi-agent systems framework (Boella and van der Torre, 2006a,b), which describes roles by means of the agent metaphor, and is formalized in Input/Output logic (Makinson and van der Torre, 2000). This work only addresses part of the problem; it focuses on persuasion dialogues. Boella et al. (2006a) provide a common role-based semantics for both approaches, using a representation format based on the FIPA semantic language. We model additional categories of speech acts, like commissives and directives, which can be applied in negotiation dialogues between cooperative agents.

8. Conclusions

Public mental attitudes have been proposed to bridge the gap between the two main traditions in defining a semantics for agent communication languages,: mental attitudes and social commitments. These traditions, even if they share the idea of speech acts as operators with preconditions and effects, and agents playing roles like speaker and hearer, rely on completely distinct ontologies. Not only does the mental attitudes approach refer to concepts like belief, desire, or intention and is the social commitment approach based on the notion of commitment, but the two approaches also refer to distinct speech acts, distinct types of dialogue games, and so on. Public mental attitudes avoid the problems of mentalistic semantics, such as the unverifiability of private mental states, and they allow the reuse of the logics and implementations developed for FIPA compliant approaches. The distinction between private and public mental attitudes allows us to model multiple communication sessions, as well as bluffing and lying agents in persuasion and negotiation.

We associate public beliefs and intentions, representing what the agent is publicly held responsible for, with a role instance. If the agent does not adhere to his role, then he can be sanctioned or blamed. The mental attitudes contain precisely the public beliefs and intentions which follow from speech acts being uttered – requirements as well as consequences – together with commonly held beliefs about the attitudes of roles.

The existence of such public mental attitudes is the only property of role instances we require, and this property therefore characterizes our notion of role in agent communication. This characteristic also forms the main distinction with most other role-based models, which typically allow private mental attitudes like secrets for role-enacting-agents or agents-in-a-role too.

The final constituent of our model is the set of constitutive rules which defines the effects of speech acts on the mental attitudes of role instances, for dialogue types like information seeking, negotiation and persuasion. Communication is public for every agent who overhears the speech acts, and has same prior knowledge of roles. Consequently, for public communication, the mental attitudes of a role are public. The challenge of agent communication languages whose semantics is based on public mental attitudes, such as our role-based semantics, is to define mappings from existing languages to the new one. In this paper we show how to map the two existing traditions to our role-based ontology of public mental attitudes and we show how it allows for a comparison.

The mapping from FIPA semantics to our role-based semantics interprets the feasibility precondition of FIPA as a kind of presupposition. For example, if an agent i informs another agent that φ with

$inform(i, j, \varphi)$, then in the original FIPA semantics, agent i should already believe the proposition φ. Otherwise, i would be insincere. With public mental attitudes, however, it is *not* the case that an agent must already publicly believe φ. Instead, it is presupposed that the agent believes φ, and therefore, the agent is publicly held responsible for the fact that φ holds. If φ does not hold, then the agent can be sanctioned or blamed. Moreover, FIPA notions like reliability and sincerity are no longer implicitly built into the formal system. By making such notions explicit for different contexts of use, we can integrate the social commitment approach, and deal both with cooperative and non-cooperative types of dialogue.

The first mapping from the social commitment semantics focuses on *action commitments* (Singh, 2000; Fornara and Colombetti, 2004), with an operational semantics. We introduce six update rules mirroring the commitment update rules of Fornara and Colombetti: an *unset* commitment corresponds to an intention of the creditor, a commitment is *pending* when it is an intention of the creditor and the debtor of the commitment conditionally wants to perform the action if the associated condition q is true, and the creditor has this as a belief, and so on. Then:

A *promise* introduces a pending commitment of the speaker;

A *request* introduces an unset commitment with the receiver as debtor, i.e., the agent of the requested action;

Accept and *reject* change the state of an existing unset commitment to *pending* and *canceled* respectively. In order to account for this fact, we insert in the antecedent of the rules for accept and reject the reference to the configuration of beliefs and intentions that represent an existing commitment.

A *propose* is a complex speech act composed by a *request* and a conditional *promise*; it introduces an unset commitment with the receiver as debtor and a pending commitment with the speaker as debtor. Since a *propose* is used in a negotiation, q and p refer respectively to an action of the speaker and of the receiver.

The second mapping from the social commitment semantics focuses on *propositional commitments*. We distinguish propositional and action commitment by means of the notion of concession, inspired by Walton and Krabbe (1995) and Kibble (2005). We map propositional commitments to the roles' beliefs, and action commitments to their intentions. The constitutive rules of a dialogue game represent the effects that speech acts have on the roles' mental attitudes, and, thus, indirectly on commitments. Concessions are introduced as the absence of a belief to the contrary, and prevent further challenges. This is analogous to action commitments, which prevent future actions that require the same resources. We illustrate propositional commitments by modeling a particular persuasion protocol, inspired by Walton and Krabbe. This protocol endorses the 'silence means consent' principle. Since under this principle concessions of the hearer can be made by default, i.e., by not challenging an assertion, we demonstrate that the semantics of speech acts should not only be expressed in terms of the effects on the attitudes of the speaker, but also on those of the hearer.

The research reported in this paper does have some limitations. One limitation is that we have chosen to stay relatively close to the original FIPA-SL language, so we inherit some of its well known problems. In Section 7 we have discussed various alternative representation languages for agent communication to overcome these problems. A limitation that our approach has in common with many other approaches, is the absence of explicit time variables in the representation language. This may become a problem, when we are modeling the feasibility preconditions of a speech act. For example, an *inform* only makes sense when the hearer does not already know the information. So, from the utterance of an *inform* act, we may infer that the speaker believes the hearer does not know the information. But this is immediately overruled by the rational effect of the *inform* act. Similar problems exist for the retraction of claims, as we discussed in Section 5. We believe that these problems may be overcome by the use of a defeasible logic with explicit time variables. An initial proposal in this direction is made by Boella et al. (2007).

The first issue for further research is the use of our role-based ontology for descriptions of expected behavior. To play a role, an agent is expected to act *as if* the beliefs and intentions of the role were his own, and to keep them coherent, as he does for his own mental attitudes, see also (Goffman, 1959). An agent should adopt to the intentions of its role, and carry them out according to his role's beliefs. The only

thing agents have to do when communicating is to respect the rules of the dialogue; otherwise they are out of the game. Agents thus have to respect the cognitive coherence of the roles they play or they get into a contradictory position. We may adopt Pasquier and Chaib-draa's (2003) view that dialogue arises from the need to maintain coherence of mental states: "two agents communicate if an incoherence forces them to do so. [...] Conversation might be seen [...] as a generic procedure for attempting to reduce incoherence". An agent engaged in a dialogue tries to avoid contradictions, not with his private mental states, but with the public image which his role constitutes. Here we do not consider the coherence of the private mental states of the agents, but the coherence of the mental attitudes of the roles. Referring only to the agents' beliefs would not be realistic, since there is no way to ensure that the addressee of an *inform* act, for example, accepts to believe a proposition. After all, agents are autonomous. This is one of the limitations of the traditional mentalistic approach. We overcome it using rules referring to public mental attitudes, but those of the roles, rather than the agents'. According to Goffman (1959), the pretense that roles and individual agents are identical, is very common in everyday practice. Although we suspect that others have beliefs and strategic goals different from the ones publicly expressed, we pretend not to see this. If we would, we would have to accuse the other openly of insincerity, often without proper evidence. The resulting conflict would be unwanted for all participants.

The second issue for further research is a formalization of the notion of responsibility, and the way in which the obligation to play a role consistently can be enforced. The introduction of obligations requires the reference to an explicit multi-agent normative system, as described in (Boella and van der Torre, 2004). Formalizing responsibility and enforcement starts with formalizing violations, which should be based on explicit evidence. Consider again the example of a liar, who once he starts lying, has to continue the dialogue consistently with what he said before, independently of his real beliefs. To formalize the consistent liar, we also have to define what it means that he no longer is able to lie, for example due to explicit evidence to the contrary.

References

Allen, J. F. and Perrault, C. R. (1980). Analyzing intention in utterances. *Artificial Intelligence*, 15(3):143–178.

Austin, J. (1962). *How to do things with words*. Harvard University Press.

Baldoni, M., Boella, G., and van der Torre, L. (2007). Interaction between objects in powerJava. *Journal of Object Technology*, 6(2):7–12.

Bentahar, J., Moulin, B., Meyer, J., and Chaib-draa, B. (2004). A modal semantics for an argumentation-based pragmatics for agent communication. In *Argumentation in Multi-Agent Systems (ArgMAS'04)*, LNCS 3366, pages 44–63, Berlin. Springer.

Boella, G., Damiano, R., Hulstijn, J., and van der Torre, L. (2006a). ACL semantics between social commitments and mental attitudes. In *Agent Communication II (AC'05 and AC'06)*, LNCS 3859, pages 30–44. Springer, Berlin.

Boella, G., Damiano, R., Hulstijn, J., and van der Torre, L. (2006b). Role-based semantics for agent communication: Embedding of the 'mental attitudes' and 'social commitments' semantics. In *Proceedings of the 5th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'06)*, pages 688–690. ACM Press.

Boella, G., Hulstijn, J., Governatori, G., Riveret, R., Rotolo, A., and van der Torre, L. (2007). FIPA communicative acts in defeasible logic. In *Proceedings of 7th IJCAI International Workshop on Nonmonotonic Reasoning, Action and Change (NRAC'07)*, pages 9–16. University of Technology.

Boella, G., Hulstijn, J., and van der Torre, L. (2005). A synthesis between mental attitudes and social commitments in agent communication languages. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'05)*, pages 358–364, Piscataway, NJ. IEEE Computer Society.

Boella, G. and van der Torre, L. (2004). Organizations as socially constructed agents in the agent oriented paradigm. In *Engineering Societies in the Agent World (ESAW'04)*, LNAI 3451, pages 1–13, Berlin. Springer.

Boella, G. and van der Torre, L. (2006a). A game theoretic approach to contracts in multiagent systems. *IEEE Transactions on Systems, Man and Cybernetics - Part C: Applications and Reviews*, 36(1):68–79.

Boella, G. and van der Torre, L. (2006b). Security policies for sharing knowledge in virtual communities. *IEEE Transactions on Systems, Man and Cybernetics - Part A: Systems and Humans*, 36(3):439–450.

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

Bresciani, P., Giorgini, P., Giunchiglia, F., Mylopoulos, J., and Perini, A. (2004). TROPOS: An agent-oriented software development methodology. *Autonomous Agents and Multi-Agent Systems*, 8(3):203–236.

Bretier, P. and Sadek, D. (1997). A rational agent as the kernel of a cooperative spoken dialogue system: implementing a logical theory of interaction. In Muller, J. P., Wooldridge, M. J., and Jennings, N. R., editors, *Intelligent Agents III*, LNAI 1193, pages 189–203. Springer, Berlin.

Caminada, M. W. A. (2004). *For the Sake of the Argument; explorations into argument-based reasoning*. Phd. thesis, Vrije Universiteit, Amsterdam.

Castelfranchi, C. (1995). Commitment: from intentions to groups and organizations. In *Proceedings of the First International Confernce on Multi-Agent System (ICMAS'95)*, pages 41–48. AAAI/MIT Press.

Castelfranchi, C. (1998). Modeling social action for AI agents. *Artificial Intelligence*, 103(1-2):157–182.

Cohen, P. R. and Levesque, H. J. (1990). Intention is choice with commitment. *Artificial Intelligence*, 42(2-3):213–261.

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

Coleman, J., O'Rourke, M., and Edwards, D. B. (2006). Pragmatics and agent communication languages. In *Proceedings of the ESSLLI Workshop on Formal Ontologies for Communicating Agents*. ESSLLI.

Colombetti, M., Fornara, N., and Verdicchio, M. (2004). A social approach to communication in multiagent systems. In *Declarative Agent Languages and Technologies (DALT'03)*, LNCS 2990, pages 191–220. Springer, Berlin.

Dastani, M., Dignum, V., and Dignum, F. (2003). Role-assignment in open agent societies. In *Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'03)*, pages 489–496. ACM Press.

Dastani, M., Herzig, A., Hulstijn, J., and Van der Torre, L. (2004). Inferring trust. In *Proceedings of 5th Workshop on Computational Logic in Multi-agent Systems (CLIMA V)*, LNAI 3487, pages 144– 160, Berlin. Springer.

Demolombe, R. (2001). To trust information sources: a proposal for a modal logical framework. In *Trust and Deception in Virtual Societies*, pages 111 – 124. Kluwer.

Esteva, M., de la Cruz, D., and Sierra, C. (2002). ISLANDER: an electronic institutions editor. In *Proceedings of the 1st International Joint Conference on Autonoumous Agents and Multiagent Systems (AAMAS'02)*, pages 1045 – 1052. ACM Press.

Ferber, J., Gutknecht, O., and Michel, F. (2003). From agents to organizations: an organizational view of multiagent systems. In *Agent-Oriented Software Engineering IV (AOSE'03)*, LNCS 2935, pages 214–230, Berlin. Springer.

FIPA (2002a). FIPA ACL communicative act library specification. Technical Report SC00037J, Foundation for Intelligent Physical Agents.

FIPA (2002b). FIPA ACL message structure specification. Technical Report SC00061G, Foundation for Intelligent Physical Agents.

FIPA (2002c). FIPA Propose interaction protocol specification. Technical Report SC00036H, Foundation for Intelligent Physical Agents.

Fornara, N. and Colombetti, M. (2004). A commitment-based approach to agent communication. *Applied Artificial Intelligence*, 18(9-10):853–866.

Garson, J. W. (1984). Quantification in modal logic. In *Handbook of Philosophical Logic*, volume II, pages 249–307. Reidel.

Gaudou, B., Herzig, A., and Longin, D. (2006a). A logical framework for grounding-based dialogue analysis. *Electronic Notes in Theoretical Computer Science*, 157(4):117–137.

Gaudou, B., Herzig, A., Longin, D., and Nickles, M. (2006b). A new semantics for the fipa agent communication language based on social attitudes. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI'06)*, pages 245–249. IOS Press.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday.

Guerini, M. and Castelfranchi, C. (2007). Promises and threats in persuasion. *Pragmatics and Cognition*, 15(2):277–311).

Halpern, J. Y. and Moses, Y. (1985). A guide to the modal logics of knowledge and belief: Preliminary draft. In *Proceedings of the 9th International Joint Conference on Artificial Intelligence (IJCAI'85)*, pages 480–490. Morgan Kaufmann.

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

Herrmann, S. (2007). A precise model for contextual roles: The programming language ObjectTeams/Java. *Applied Ontology*, 2(2):181–207.

Herzig, A. and Longin, D. (2000). Belief dynamics in cooperative dialogues. *Journal of Semantics*, 17(2):91–118.

Hulstijn, J., Dignum, F., and Dastani, M. (2004). Coherence in multi-agent interaction. In *Agent Communication (AC'2004)*, LNCS 3396, pages 134–152. Springer, Berlin.

Kagal, L. and Finin, T. (2005). Modeling conversation policies using permissions and obligations. In *Developments in Agent Communication (AC'04)*, LNCS 3396, pages 120–133, Berlin. Springer.

Kibble, R. (2005). Speech acts, commitment and multi-agent communication. *Computational & Mathematical Organization Theory*, 12(2-3):127–145.

Lewis, D. (1969). *Convention: A Philosophical Study.* Harvard University Press.

Liau, C. (2003). Belief, information acquisition, and trust in multi-agent systems – a modal formulation. *Artificial Intelligence*, 149:31–60.

Lochbaum, K. E. (1998). A collaborative planning model of intentional structure. *Computational Linguistics*, 24(4):525–572.

Loebe, F. (2007). Abstract vs. social roles - towards a general theoretical account of roles. *Applied Ontology*, 2(2):127–158.

Mackenzie, J. (1979). Question begging in non-cumulative systems. *Journal of Philosophical Logic*, 8:117–133.

Makinson, D. and van der Torre, L. (2000). Input-output logics. *Journal of Philosophical Logic*, 29(4):383–408.

Masolo, C., Guizzardi, G., Vieu, L., Bottazzi, E., and Ferrario, R. (2005). Relational roles and qua-individuals. In *Proceedings of the AAAI Fall Symposium on 'Roles, an Interdisciplinary Perspective'*, pages 103–112. AAAI Press, Menlo Park.

Nickles, M., Fischer, F., and Weiss, G. (2006). Communication attitudes: A formal approach to ostensible intentions, and individual and group opinions. *Electronic Notes in Theoretical Computer Science*, 157(4):95–115.

Pasquier, P. and Chaib-draa, B. (2003). The cognitive coherence approach for agent communication pragmatics. In *Proceedings of the 2nd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'03)*, pages 544–551, New York. ACM Press.

Pasquier, P., Flores, R., and Chaib-draa, B. (2004). Modelling flexible social commitments and their enforcement. In *Engineering Societies in the Agent World (ESAW'04)*, LNCS 3451, pages 139–151, Berlin. Springer.

Pitt, J. and Mamdani, A. (1999). Some remarks on the semantics of FIPA's agent communication language. *Autonomous Agents and Multi-Agent Systems*, 2(4):333–356.

Pollack, M. E. (1990). Plans as complex mental attitudes. In *Intentions in Communication*, pages 77–103. MIT Press.

Sadek, D. M. (1991). *Attitudes Mentales et Interaction Rationnelle: Vers une Théorie Formelle de la Communication*. PhD thesis, Université de Rennes I.

Sadek, D. M. (1992). A study in the logic of intention. In *Proceedings of the 3rd International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 462–473. Morgan Kaufmann.

Sadek, D. M. (1999). Design considerations on dialogue systems: From theory to technology - the case of ARTIMIS. In *Proceedings of the ESCA Tutorial and Research Workshop on Interactive Dialogue in Multi-Modal Systems, Kloster Irsee*, pages 173– 188. ISCA Archive.

Sadek, D. M., Bretier, P., and Panaget, E. (1997). ARTIMIS: Natural dialogue meets rational agency. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence (IJCAI'97)*, pages 1030–1035. Morgan Kaufmann.

Schank, R. C. and Abelson, R. P. (1977). *Scripts, Plans, Goals and Understanding: an Inquiry into Human Knowledge Structures*. Erlbaum.

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

Singh, M. P. (2000). A social semantics for agent communication languages. In *Issues in Agent Communication (AC'00)*, LNCS 1916, pages 31 – 45. Springer, Berlin.

Smith, R. G. (1980). The contract net protocol: High-level communication and control in a distributed problem solver. *IEEE Transactions on Computers*, 29(12):1104–1113.

Stalnaker, R. C. (2002). Common ground. *Linguistics and Philosophy*, 25:701–721.

Verdicchio, M. and Colombetti, M. (2006). From message exchanges to communicative acts to commitments. *Electronic Notes in Theoretical Computer Science*, 157(4):75–94.

Walton, D. N. and Krabbe, E. C. (1995). *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New York Press.

Winograd, T. and Flores, F. (1986). *Understanding Computers and Cognition*. Addison-Wesley.

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

Wooldridge, M., Jennings, N., and Kinny, D. (2000). The Gaia methodology for agent-oriented analysis and design. *Autonomous Agents and Multi-Agent Systems*, 3(3):285–312.

Yolum, P. and Singh, M. P. (2002). Flexible protocol specification and execution: Applying event calculus planning using commitments. In *Proceedings of the 1st International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS'02)*, pages 527–534. ACM Press.

Zambonelli, F., Jennings, N., and Wooldridge, M. (2003). Developing multiagent systems: The Gaia methodology. *IEEE Transactions of Software Engineering and Methodology*, 12(3):317–370.