

# A Modal Semantics for an Argumentation-Based Pragmatics for Agent Communication

Jamal Bentahar<sup>1</sup>, Bernard Moulin<sup>1</sup>, John-Jules Ch. Meyer<sup>2</sup>,  
and Brahim Chaib-draa<sup>1</sup>

<sup>1</sup> Laval University, Department of Computer Science and Software Engineering,  
Ste Foy, QC, G1K 7P4, Canada

jamal.bentahar.1@ulaval.ca  
{bernard.moulin, brahim.chaib-draa}@ift.ulaval.ca

<sup>2</sup> University Utrecht, Department of Computer Science,  
PO Box 80.089 3508TB, Utrecht, The Netherlands  
jj@cs.uu.nl

**Abstract.** In this paper we present a modal semantics for our approach based on social commitments and arguments for conversational agents. Our formal framework based on this approach uses three basic elements: social commitments, actions that agents apply to these social commitments and arguments that agents use to support their actions. This framework, called Commitment and Argument Network (CAN), formalizes the agents' interactions as a network in which agents manipulate commitments and arguments. More precisely, we propose a logical model (called  $DCTL^*_{CAN}$ ) based on  $CTL^*$  and on dynamic logic for this framework. The advantage of this logical model is to bring together social commitments, actions, argumentation relations, and the relations existing between these three elements within the same framework. Our semantics makes it possible to represent the dynamics of agent communication. It also allows us to establish the important link between social commitments as a deontic concept and arguments. The final objective of this paper is to propose a unified framework for pragmatics and semantics of agent communication by defining logic-based protocols.

## 1 Introduction

In the domain of agent communication, semantics is one of the most important aspects, particularly in the current context of open and interoperable multi-agent systems (MAS) [7]. Although a certain number of significant research works were done in this field [13, 22, 24, 25], the definition of a clear and global semantics is an objective yet to be reached. Agent communication pragmatics is another important aspect to be addressed. While semantics is interested in the meaning of communication acts, pragmatics deals with the way of using these acts. Pragmatics is related to the dynamics of agent interactions and to the way of relating the isolated acts to build conversations. Pragmatics was also addressed by several researchers [9, 18, 20, 21]. However, only few attempts have been made to address these two facets of agent communication in the same framework.

The objective of this paper is to propose a general framework to capture pragmatic and semantic issues of an approach based on social commitments and arguments for agent communication. Indeed, this work is a continuation of our previous research in which we addressed in detail the pragmatic aspects [2, 3]. Thus, the paper highlights the semantic issues of our approach and the link with pragmatic ones. The semantics that we define here deals with all the aspects used in our approach.

In addition to proposing a unified framework for pragmatic and semantic issues, this work presents two results: 1) it semantically establishes the link between social commitments and arguments; 2) it uses both a temporal logic (CTL\* with some additions) and a dynamic logic to define a complete and unambiguous semantics.

Paper overview. In Section 2 we address the pragmatic aspects by introducing the main ideas of our approach. In Section 3 we present the syntax and the semantics of the main elements of our logical model. Other details will be described in an extended version of the paper. In Section 4 we define protocols using our logical model. In Sections 5 and 6 we compare our approach to related work and conclude the paper.

## 2 Social Commitments and Argument-Based Approach

### 2.1 Social Commitments

A social commitment SC is a commitment made by an agent (called the *debtor*), that some fact is true [5]. This commitment is directed to a set of agents (called *creditors*). The social commitment content is characterized by time  $t_\varphi$ , which is generally different from the utterance time denoted  $t_u$  and from the time associated with the social commitment denoted  $t_{sc}$ .  $t_\varphi$  is the time described by the utterance, and thus by the content  $\varphi$ . Time  $t_{sc}$  refers to the time during which the social commitment holds. When it is an interval, this time is denoted  $[t_{sc}^{\text{inf}}, t_{sc}^{\text{sup}}]$ . If the social commitment is satisfied or violated we have  $t_{sc}=[t_u, t_\varphi]$ . However, if the social commitment is withdrawn, we have:  $t_{sc}=[t_u, t_w]$ , with  $t_w$  the withdrawal time. Time  $t_{sc}$  indicates the time during which the social commitment holds, i.e. the time during which the social commitment is *active* (we will return to this notion later). Time  $t_\varphi$  indicates the moment at which the social commitment must be satisfied.

In order to model the dynamics of conversations, we interpret a speech act SA as an action performed on a social commitment or on the content of a social commitment. A SA is an abstract act that an agent, the speaker, performs when producing an utterance  $U$  and addressing it to another agent, the addressee. The actions that an agent can perform on a social commitment are:  $\text{Act} \in \{\text{Create, Withdraw, Reactivate, Violate, Satisfy}\}$ . The actions performed on the content of a social commitment are  $\text{Act-content} \in \{\text{Accept-content, Refuse-content, Challenge-content, Change-content}\}$ . Thus, a SA leads either to an action on a social commitment when the speaker is the debtor, or to an action on a social commitment content when the speaker is the debtor or the creditor. Formally, in our framework a SA can be defined in BNF form as follows:

**Definition 1.**  $SA(i_k, Ag_1, Ag_2, t_u, U) =_{def}$   
 $Act(Ag_1, t_u, SC(id_n, Ag_1, Ag_2, t_{sc}, \varphi, t_\varphi))$   
 $\mid Act-content(Ag_k, t_u, SC(id_n, Ag_i, Ag_j, t_{sc}, \varphi, t_\varphi))$

where  $i, j \in \{1, 2\}$  and  $(k = i \text{ or } k = j)$ ,  $=_{def}$  means “is interpreted by definition as”,  $i_k$  is the identifier of the SA. The definiendum  $SA(i_k, Ag_1, Ag_2, t_u, U)$  is defined by the definiens  $Act(Ag_1, t_u, SC(id_n, Ag_1, Ag_2, t_{sc}, \varphi, t_\varphi))$  as an action performed by the debtor  $Ag_1$  on its social commitment. The definiendum is defined by the definiens  $Act-content(Ag_k, t_u, SC(id_n, Ag_i, Ag_j, t_{sc}, \varphi, t_\varphi))$  as an action performed by an agent  $Ag_k$  (the debtor or the creditor) on the social commitment content.

## 2.2 Taxonomy

In this section, we explain the various types of social commitments we use in the logical model:

**A. Absolute Commitments (ABCs):** They are social commitments whose fulfillment does not depend on any particular condition. Two types can be distinguished:

**A1. Propositional Commitments (PCs):** They are related to the state of the world and expressed by assertives.

**A2. Action Commitments (ACs):** They are always directed towards the future and are related to actions that the debtor is committed to carrying out. This type of social commitments is typically conveyed by promises.

**B. Conditional Commitments (CCs):** In several cases, agents need to make social commitments not in absolute terms but under given conditions. CCs allow us to express that if a condition  $\beta$  is true, then the creditor will be committed towards the debtor to making  $\gamma$  or that  $\gamma$  is true.

**C. Commitment Attempts (CTs):** The social commitments described so far directly concern the debtor who commits either that a certain fact is true or that a certain action will be carried out. These social commitments do not allow us to explain the fact that an agent asks another one to be committed to carrying out an action. To solve this problem, we propose the concept of CT. We consider a CT as a request made by a debtor to push a creditor to be committed.

We notice that there is no explicit relation between PCs and ACs. When the current state of the world does not satisfy a PC we speak about a violation of this social commitment. There is no rule indicating that the agent develops an AC to make the content of a PC true when this PC becomes violated.

## 2.3 Argumentation and Social Commitments

In the domain of agent communication, several argumentation-based approaches have been put forward, for example [1, 15]. An argumentation system essentially includes a logical language  $L$ , a definition of the argument concept, a definition of the attack relation between arguments and finally a definition of acceptability [1]. In our model, we adopt the following definition from [10]. Here  $\Gamma$  indicates a possibly inconsistent

knowledge base with no deductive closure.  $\vdash$  Stands for classical inference and  $\equiv$  for logical equivalence.

**Definition 2.** An argument is a pair  $(H, h)$  where  $h$  is a formula of  $L$  and  $H$  a sub-set of  $\Gamma$  such that : i)  $H$  is consistent, ii)  $H \vdash h$  and iii)  $H$  is minimal, so that no subset of  $H$  satisfying both i and ii exists.  $H$  is called the support of the argument and  $h$  its conclusion.

The link between social commitments and arguments enables us to capture both the public and reasoning aspects of agent communication. This link is explained as follows. Before committing to some fact  $h$  being true (i.e. before creating a social commitment whose content is  $h$ ), the speaker agent must use its argumentation system to build an argument  $(H, h)$ . On the other side, the addressee agent must use its own argumentation system to select the answer it will give. For example, an agent  $Ag_1$  accepts the social commitment content  $h$  proposed by another agent  $Ag_2$  if it is able to build an argument which supports this content from its knowledge base. If  $Ag_1$  has an argument neither for  $h$ , nor for  $\neg h$ , then it must ask for an explanation.

The argumentation relations that we use in our model are thought of as actions applied to social commitment contents. The set of these relations is:  $\{Justify, Defend, Attack, Contradict\}$ .

We used this approach in [2] to propose a formal framework called Commitment and Argument Network (CAN). The idea is to reflect the dynamics of agent communication by a network in which agents manipulate social commitments and arguments. In the following section we propose a formal semantics of this formalism in the form of a logical model.

### 3 The Logical Model

#### 3.1 Syntax

In this section we specify the syntax of the main elements we use in our framework. The details of the other elements are described in an extended version of the paper. Our formal language  $\mathcal{L}$  is based on an extended version of CTL\* [11] and on dynamic logic [14]. We use a branching time for the future and we suppose that the past is linear. We also suppose that time is discrete. Let  $\Phi p$  be the set of atomic propositions and  $\Phi a$  the set of action symbols. The set of the agents is denoted  $A$  and the set of time units is denoted  $TU$ . The various types of social commitments, the actions of the agents on social commitments and on their contents and the argumentation relations are introduced as modal operators. We denote  $\mathcal{L}_{sc}$  a sub-language of  $\mathcal{L}$  for social commitments. To simplify the notation, a social commitment, independently of its type, is denoted:  $SC(Id_0, Ag_1, Ag_2, \varphi)$ .  $Id_0 \in N$  is the social commitment identifier,  $Ag_1$  and  $Ag_2$  are two agents and  $\varphi$  the social commitment content. The language  $\mathcal{L}$  can be defined by the following syntactic rules.

##### 3.1.1 Propositional Elements

**R1.**  $\forall \phi \in \Phi p, \phi \in \mathcal{L}$ : Atomic formula

**R2.**  $p, q \in \mathcal{L} \Rightarrow p \wedge q \in \mathcal{L}$ : Conjunction

**R3.**  $p \in \mathcal{L} \Rightarrow \neg p \in \mathcal{L}$ : *Negation*

**R4.**  $p, q \in \mathcal{L} \Rightarrow p \therefore q \in \mathcal{L}$ : *Argumentation*

This means that  $p$  is an argument for  $q$ . We can read this formula:  $p$ , so  $q$ . At this level, our definition of the argument does not take into account the defeasible aspect. This aspect will be introduced into our model by the argumentation relations (Section 3.1.6).

**R5.**  $p \in \mathcal{L} \Rightarrow ?p \in \mathcal{L}$ : *Is  $p$  true?*

**R6.**  $p \in \mathcal{L} \Rightarrow \text{Ap} \in \mathcal{L}$ : *Universal path-quantifier*

**R7.**  $p \in \mathcal{L} \Rightarrow \text{Ep} \in \mathcal{L}$ : *Existential path-quantifier*

**R8.**  $p, q \in \mathcal{L} \Rightarrow p \text{ U }^+ q \in \mathcal{L}$ : **Until (in the future)**

Informally,  $p \text{ U }^+ q$  ( $p$  until  $q$ ) means that on a given path from the given moment, there is some future moment in which  $q$  will eventually hold and  $p$  holds at all moments until that future moment.

**R9.**  $p \in \mathcal{L} \Rightarrow \text{X}^+p \in \mathcal{L}$ : **Next moment (in the future)**

$\text{X}^+p$  holds at the current moment, if  $p$  holds at the next moment.

**R10.**  $p, q \in \mathcal{L} \Rightarrow p \text{ U }^- q \in \mathcal{L}$ : **Since (in the past)**

The intuitive interpretation of  $p \text{ U }^- q$  ( $p$  since  $q$ ) is that on a given path from the given moment, there is some past moment in which  $q$  eventually held and  $p$  holds at all moments since that past moment.

**R11.**  $p \in \mathcal{L} \Rightarrow \text{X}^-p \in \mathcal{L}$ : **Previous moment (in the past)**

$\text{X}^-p$  holds at the current moment, if  $p$  held at the previous moment.

### 3.1.2 Actions

**R12.**  $p \in \mathcal{L}/\mathcal{L}_{sc}, \alpha \in \Phi_a \Rightarrow \text{Perform}(\alpha)p \in \mathcal{L}$ : **Action performance (about propositions)**

$\text{Perform}(\alpha)p$  is an operator from dynamic logic. It indicates that the achievement of action  $\alpha$  makes the proposition  $p$  true.

**R13.**  $\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi) \in \mathcal{L}_{sc}, \alpha \in \Phi_a \Rightarrow$

$\text{Perform}(\alpha)\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi) \in \mathcal{L}$ : **Action performance (about social commitments)**

This indicates that the achievement of action  $\alpha$  makes the social commitment  $\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$  true in our model.

### 3.1.3 Social Commitments

**R14.**  $p \in \mathcal{L}/\mathcal{L}_{sc} \wedge \text{Id}_0 \in \mathbb{N} \wedge \{\text{Ag}_1, \text{Ag}_2\} \subseteq A \Rightarrow$

$\text{PC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, p) \in \mathcal{L}_{sc}$ : **Propositional commitment**

**R15.**  $\alpha \in \Phi_a \wedge p \in \mathcal{L}/\mathcal{L}_{sc} \wedge \text{Id}_0 \in \mathbb{N} \wedge \{\text{Ag}_1, \text{Ag}_2\} \subseteq A \Rightarrow$

$\text{AC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \alpha)p \in \mathcal{L}_{sc}$ : **Action commitment**

**R16.**  $\beta \in \mathcal{L}/\mathcal{L}_{sc} \wedge \gamma \in \mathcal{L}/\mathcal{L}_{sc} \cup \Phi_a \wedge Id_0 \in N \wedge \{Ag_1, Ag_2\} \subseteq A \Rightarrow CC(Id_0, Ag_1, Ag_2, \beta \Rightarrow \gamma) \in \mathcal{L}_{sc}$ : *Conditional commitment*

In order to formally introduce the notion of CT we need some definitions from first order logic.

**Definitions.**

*TerC*: a set of constant terms. A constant term can be a number, a name, etc.

*Var*: a set of variables.

*Val*:  $Var \mapsto TerC$ : a valuation function associating a variable to a constant term.

Let  $\mathcal{E}_{val}$  be a substitution that makes it possible to substitute each free variable  $x$  that appears in a formula  $\varphi$  by a constant term, i.e. by  $Val(x)$ . We denote a formula  $\varphi$  in which appears a sequence of free variables  $X$  by  $?X\varphi$ . The expression  $?X\varphi \bullet \mathcal{E}_{val}$  indicates the formula  $\varphi$  in which each variable  $x$  of the sequence of free variables  $X$  is substituted by a corresponding value (i.e. by  $Val(x)$ ). Thus, we can define the syntax of a CT as follows:

**R17.**  $?X\varphi \in \mathcal{L}/\mathcal{L}_{sc} \wedge Id_0 \in N \wedge \{Ag_1, Ag_2\} \subseteq A \Rightarrow$

$CT(Id_0, Ag_1, Ag_2, ?X\varphi) \in \mathcal{L}_{sc}$ : *Commitment attempt*

**3.1.4 Actions Applied to Commitments**

**R18.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow$

$Create(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *Creation of a social commitment*

**R19.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow$

$Withdraw(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *Withdrawal of a social commitment*

**R20.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow$

$Satisfy(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *Satisfaction of a social commitment*

**R21.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow$

$Violate(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *Violation of a social commitment*

**R22.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow Active(SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *An active social commitment*

**3.1.5 Actions Applied to Commitment Contents**

**R23.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow$

$Accept-content(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *Acceptation of a social commitment content*

**R24.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow$

$Challenge-content(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *Challenge of a social commitment content*

**3.1.6 Argumentation Relations**

**R25.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \wedge \varphi' \in \mathcal{L}/\mathcal{L}_{sc} \Rightarrow$

$Justify-content(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi') \in \mathcal{L}_{sc}$ : *Justification*

**R26.**  $SC(Id_0, Ag_1, Ag_2, \varphi) \in \mathcal{L}_{sc} \Rightarrow$

$Contradict-content(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi)) \in \mathcal{L}_{sc}$ : *Contradiction*. This relation means that an agent contradicts the content of its social commitment.

**R27.**  $SC(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi) \in \mathcal{L}_{sc} \wedge \varphi' \in \mathcal{L}/\mathcal{L}_{sc} \Rightarrow$

Attack-content( $\text{Ag}_2, SC(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi), \varphi'$ )  $\in \mathcal{L}_{sc}$ : *Attack of a social commitment content*

**R28.**  $SC(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi) \in \mathcal{L}_{sc} \wedge \varphi' \in \mathcal{L}/\mathcal{L}_{sc} \Rightarrow$

Defend-content( $\text{Ag}_2, SC(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi), \varphi'$ )  $\in \mathcal{L}_{sc}$ : *Defense of a social commitment content against an attacker*

**Abbreviations:** We use in our model the following abbreviations:

**A1.**  $p \vee q$  (disjunction) is the abbreviation of  $\neg(\neg p \wedge \neg q)$

**A2.**  $p \Rightarrow q$  (implication) is the abbreviation of  $\neg p \vee q$

**A3.**  $F^+p$  (sometimes in the future) is the abbreviation of  $\text{true } U^+ p$

**A4.**  $G^+p$  (globally in the future) is the abbreviation of  $\neg F^+ \neg p$

**A5.**  $F^-p$  (sometimes in the past) is the abbreviation of  $\text{true } U^- p$

**A6.**  $G^-p$  (globally in the past) is the abbreviation of  $\neg F^- \neg p$

## 3.2 Semantics

In this section, we define the formal model in which we evaluate the well-formed formulas of our framework. Thereafter, we give the semantics of the different elements that we specified syntactically in the previous section.

### 3.2.1 The Formal Model

Let  $S$  be a set of states. A path  $Pa$  is an infinite sequence of states  $\langle s_0, s_1, \dots \rangle$  where  $T(s_0) < T(s_1) < \dots$ . The function  $T$  gives us for each state  $s_i$  the corresponding moment  $t$  (this function will be specified later). Generally, for all  $i$  and  $j$  of  $N$ , if  $i < j$  and  $s_i$  and  $s_j$  belong to the same path  $Pa$ , then  $T(s_i) < T(s_j)$ . We denote the set of all paths by  $\sigma$ . The set of all paths starting from the state  $s_i$  are denoted:  $\sigma^i$ . In our vision of branching future, we can have several states at the same moment. Only along a given path (for example the real path) there is one and only one state at one moment. Indeed, in our framework,  $s_i$  does not indicate (necessarily) the state at moment  $i$ . Therefore, it is necessary to specify the state  $s$  and the moment  $t$  i.e. a couple  $(s, t) \in S \times TU$ . According to this formalization, we can use the notation:  $M, s_i, T(s_i) \vdash \psi$  to indicate that  $\psi$  is satisfied in the Kripke model  $M$  at the state  $s_i$  at the moment  $T(s_i)$ . To simplify this notation, we will use in the rest of the paper the following notation:  $M, s_i \vdash \psi$ . In this notation:  $M, s_i \vdash \psi$  there is a "hidden" time.

Following this simplification we can write:

$M, s_i, T(s_i) \vdash \psi$  iff  $M, s_i \vdash \psi$ .

The formal model for  $\mathcal{L}$  is defined as follows:

$M(S, A, Np, Np?, Fap, Rpc, Rac, T)$  where

$S$  : a nonempty set of states.

$A$  : a nonempty set of agents.

$Np : S \mapsto 2^{\mathcal{A}ap}$  : function relating each state  $s \in S$  to the set of the atomic propositions that are true in this state.

$Np? : S \mapsto 2^{\Phi_{ap}}$  : function relating each state  $s \in S$  to the set of the atomic propositions that are neither true nor false in this state (i.e. we do not know if they are true or false).

$Fap : S \times \Phi_a \mapsto 2^S$  : function that gives us the state transitions caused by the achievement of an action.

$Rpc : A \times A \times S \mapsto 2^S$  : function producing the accessibility modal relations for PCs.

$Rac : A \times A \times S \mapsto 2^S$  : function producing the accessibility modal relations for AC.

$T : S \mapsto TU$  : function associating to any state  $s_i$  the corresponding time.

The functions  $Rpc$  and  $Rac$  give us the states that correspond to the time  $t_\varphi$ , i.e. the states in which the social commitment created by an agent  $Ag_1$  towards another agent  $Ag_2$  must be satisfied. These functions allow us to define a deadline for determining whether a violation or a satisfaction occurs. They give us all the states corresponding to the time  $t_\varphi$  on all paths starting from the state at moment  $t_u$ . The fact that these two functions give us a set of states means that the social commitment must be satisfied whatever the future. Since there is only one real path, the social commitment is satisfied or is violated only in one state of the set given by  $Rpc$  and  $Rac$ . Indeed, the outputs of the functions  $Rpc$  and  $Rac$  are known only after the creation of the social commitment. Thus, this depends on the state in which the social commitment is created. For example, if we have:  $s_j \in Rpc(Ag_1, Ag_2, s_i)$ , then this means that at moment  $T(s_i)$  agent  $Ag_1$  is committed towards agent  $Ag_2$  to satisfy a certain social commitment at moment  $T(s_j)$ . We can see that  $Rpc$  depends on the current moment  $T(s_i)$ .

The algebraic properties of these two relations are as follows:

1-  $Rac$  is not reflexive, i.e.:

$$\forall Ag_1, Ag_2 \in A, \forall s_i \in S, s_i \notin Rac(Ag_1, Ag_2, s_i)$$

The reason is that this accessibility relation defines a deadline and that action commitments are always directed towards the future. For the same reason, we have:  $Rpc^f$  is not reflexive and  $Rpc^p$  is reflexive, where:  $Rpc^f$  is the restriction of  $Rpc$  to the propositional commitments directed to the future, and  $Rpc^p$  is the restriction of  $Rpc$  to the propositional commitments directed towards the past and the present.

2-  $Rpc$  and  $Rac$  are serial, i.e.:

$$\forall Ag_1, Ag_2 \in A, \forall s_i \in S : \exists s_j \in R(Ag_1, Ag_2, s_i)$$

where  $R = Rpc$  or  $R = Rac$

This property fits with the notion of infinite path in CTL\*.

3-  $Rpc$  and  $Rac$  are transitive, i.e.:

$$\forall Ag_1, Ag_2 \in A, \forall s_i, s_j, s_k \in S : s_j \in R(Ag_1, Ag_2, s_i) \wedge s_k \in Rac(Ag_1, Ag_2, s_j) \\ \Rightarrow s_k \in R(Ag_1, Ag_2, s_i)$$

where  $R = Rpc$  or  $R = Rac$ .



Consequently, social commitments in our model are *S4* modal logic operators. The interpretation of this property is as follows: if an agent commits that a proposition is true, or so that an action will be performed, this implies that the agent commits so that it commits that the proposition is true or so that the action will be performed.

4- Because *Rpc* and *Rac* allow us to define a deadline, these relations are *not symmetric*, i.e.:  $\forall Ag_1, Ag_2 \in A, \exists s_i, s_j \in S : s_j \in R(Ag_1, Ag_2, s_i) \wedge s_i \notin R(Ag_1, Ag_2, s_j)$  where  $R = R_{pc}$  or  $R = R_{ac}$ .

5- *Rpc* and *Rac* are not *euclidean*, i.e.:

$$\forall Ag_1, Ag_2 \in A, \exists s_i, s_j, s_k \in S : s_j \in R(Ag_1, Ag_2, s_i) \wedge s_k \in R(Ag_1, Ag_2, s_i) \\ \wedge s_k \notin R(Ag_1, Ag_2, s_j) \wedge s_j \notin R(Ag_1, Ag_2, s_k)$$

where  $R = R_{pc}$  or  $R = R_{ac}$ .

Therefore, the *negative introspection* schema *S5* is not verifiable in our model.

We notice here that we do not impose a model to be *asymmetric*, but we only emphasize the fact that *Rpc* and *Rac* are *not symmetric*. For this reason we use the existential quantifier in 4.

As in CTL\*, we have in our model path formulas and state formulas. We propose to evaluate the static formulas (the different types of social commitments) as state formulas. These formulas can also be interpreted on paths in which case one considers satisfaction in the first state of a path. On the other hand, we propose to evaluate dynamic formulas (the actions on social commitments) on paths. These path formulas can become state formulas if they are true on all the paths starting from a given state.  $M, s_i \vdash \psi$  indicates that the formula  $\psi$  is evaluated in the state  $s_i$  of the model  $M$ .  $M, Pa, s_i \vdash \psi$  indicates that the formula  $\psi$  is evaluated on the path  $Pa$  starting from the state  $s_i$  of the model  $M$ . We can now define the semantics of the elements of  $\mathcal{L}$ .

### 3.2.2 Propositional Elements

**S1.**  $M, s_i \vdash \psi$  iff  $\psi \in \text{Np}(s_i)$  with  $\psi \in \Phi_p$

**S2.**  $M, s_i \vdash p \wedge q$  iff  $M, s_i \vdash p$  &  $M, s_i \vdash q$

**S3.**  $M, s_i \vdash \neg p$  iff  $\neg(M, s_i \vdash p)$

**S4.**  $M, s_i \vdash p \therefore q$  iff  $M, s_i \vdash p$  &  $(\forall j : M, s_j \vdash p \Rightarrow M, s_j \vdash q)$

In *S4* we add the first clause  $(M, s_i \vdash p)$  to capture the following aspect: when an agent presents an argument  $p$  for  $q$  (i.e.  $p \therefore q$ ) for this agent  $p$  is true and if  $p$  is true then  $q$  is true. Indeed,  $p$  so  $q$  is stronger than just stating that both  $p$  and  $q$  are true. The implication is much stronger since it holds in all the states of the model  $M$ . The idea is to express that  $p$  is the support of the conclusion  $q$ .

**S5.**  $M, s_i \vdash ?p$  iff  $p \in \text{Np}?(s_i)$ .

**S6.**  $M, s_i \vdash Ap$  iff  $(\forall Pa : Pa \in \sigma^{s_i} \Rightarrow M, Pa, s_i \vdash p)$

**S7.**  $M, s_i \vdash Ep$  iff  $(\exists Pa \in \sigma^{s_i} \& M, Pa, s_i \vdash p)$

**S8.**  $M, Pa, s_i \vdash p$  iff  $M, s_i \vdash p$ : *Propositional path formulas*

**S9.**  $M, Pa, s_i \vdash p \wedge q$  iff  $M, Pa, s_i \vdash p$  &  $M, Pa, s_i \vdash q$

- S10.**  $M, Pa, s_i \vdash \neg p$  iff  $\neg(M, Pa, s_i \vdash p)$   
**S11.**  $M, Pa, s_i \vdash p \cup^+ q$  iff  $(\exists j : i \leq j \ \& \ M, Pa, s_j \vdash q$   
 $\& (\forall k : i \leq k < j \Rightarrow M, Pa, s_k \vdash p))$   
**S12.**  $M, Pa, s_i \vdash X^+ p$  iff  $M, Pa, s_{i+1} \vdash p$   
**S13.**  $M, Pa, s_i \vdash p \cup^- q$  iff  $(\exists j : j \leq i \ \& \ M, Pa, s_j \vdash q$   
 $\& (\forall k : j < k \leq i \Rightarrow M, Pa, s_k \vdash p))$   
**S14.**  $M, Pa, s_i \vdash X^- p$  iff  $M, Pa, s_{i-1} \vdash p$ )

### 3.2.3 Actions

- S15.**  $M, Pa, s_i \vdash \text{Perform}(\alpha)p$  iff  $\forall s_j : s_j \in \text{Fap}(s_i, \alpha) \wedge s_j \subset Pa \Rightarrow M, Pa, s_j \vdash p$ .  
 where  $s_j \subset Pa$  indicates that  $Pa, s_j$  is a *suffix* of  $Pa, s_i$ .  
**S16.**  $M, s_i \vdash \text{Perform}(\alpha)p$  iff  $\forall Pa : Pa \in \sigma^{\text{si}} \Rightarrow M, Pa, s_i \vdash \text{Perform}(\alpha)p$ .

#### *Action performance (related to social commitments)*

- S17.**  $M, Pa, s_i \vdash \text{Perform}(\alpha)\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$  iff  
 $\forall s_j : s_j \in \text{Fap}(s_i, \alpha) \wedge s_j \subset Pa \Rightarrow M, Pa, s_j \vdash \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$ .

This formula indicates that the achievement of action  $\alpha$  makes the social commitment true in all the accessible states from the state  $s_i$ . As for S15, the accessible states are defined by the function *Fap*. The evaluation of this operator in a state is given by the following formula:

- S18.**  $M, s_i \vdash \text{Perform}(\alpha)\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$  iff  
 $\forall Pa : Pa \in \sigma^{\text{si}} \Rightarrow M, Pa, s_i \vdash \text{Perform}(\alpha)\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$ .

### 3.2.4 Social Commitments

#### *Social commitment as a path formula*

- S19.**  $M, Pa, s_i \vdash \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$  iff  $M, s_i \vdash \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$   
**S20.**  $M, s_i \vdash \text{PC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, p)$  iff  $(\forall s_j : s_j \in \text{Rpc}(\text{Ag}_1, \text{Ag}_2, s_i) \Rightarrow M, s_j \vdash p)$   
**S21.**  $M, s_i \vdash \text{AC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \alpha)p$  iff  
 $\forall s_j : s_j \in \text{Rac}(\text{Ag}_1, \text{Ag}_2, s_i) \Rightarrow M, s_j \vdash \text{Perform}(\alpha)p$ .

The formula S21 indicates that agent  $Ag_1$  is committed towards agent  $Ag_2$  to do  $\alpha$  and that in all accessible states  $s_j$  performing  $\alpha$  makes  $p$  true. According to formulas S20 and S21, the semantics we give to the social commitments requires their fulfillment. Thus, if it is created, a social commitment must be held. However, it is always possible to violate or withdraw such a social commitment. For this reason, these two operations (violation and withdrawal) are explicitly included in our framework. Thus, it is possible to have wrong social commitments in the model. The reason is that *Rpc* and *Rac* give us the states that correspond to the states in which the social commitment *must* be satisfied. These states are not conceived as merely "possible", but as states when the content of a social commitment *must* be true.

We notice that although *Rpc* and *Rac* are dynamic functions, we do not need to change the Kripke model  $M$  to capture this dynamics. This way of modeling is different from that used for example in KARO framework [19]. In our model which fits in naturally with CTL\* the whole dynamics is represented in one unique model.

**S22.**  $M, s_i \vdash \text{CC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \beta \Rightarrow \gamma)$   
iff  $(M, s_i \vdash \text{EF}^+\beta \Rightarrow M, s_i \vdash \text{ABC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \gamma))$

This formula indicates that agent  $\text{Ag}_i$  commits to perform  $\gamma$  (or that  $\gamma$  is true) only if the condition  $\beta$  is true (or is satisfied).

In order to define the semantics of CTs, we define the binary relation  $\vDash^{\text{EVal}}$  between a pair  $(M, s_i)$  and a formula  $\text{CT}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, ?X\varphi)$  as follows:

**S23.**  $M, s_i \vDash^{\text{EVal}} \text{CT}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, ?X\varphi)$  iff  
 $(M, s_i \vdash \text{EX}^+\text{F}^+\text{ABC}(\text{Id}_0, \text{Ag}_2, \text{Ag}_1, ?X\varphi \bullet \Xi_{\text{Val}}))$   
 $\vee (\exists \beta \in \mathcal{L}/\mathcal{L}\text{sc} : M, s_i \vdash \text{EX}^+\text{F}^+\text{CC}(\text{Id}_0, \text{Ag}_2, \text{Ag}_1, \beta \Rightarrow ?X\varphi \bullet \Xi_{\text{Val}}))$

This formula indicates that a CT whose content is  $?X\varphi$  is satisfied in the model  $M$  according to a substitution  $\Xi_{\text{Val}}$  iff the creditor (i.e.  $\text{Ag}_2$ ) will commit that a content  $?X\varphi \bullet \Xi_{\text{Val}}$  is true. In other words, the CT is satisfied iff the interlocutor will commit that the substitution  $\Xi_{\text{Val}}$  for the sequence  $X$  of free variables appearing in the formulae  $\varphi$  is true. The social commitment of the interlocutor can be absolute (ABC) or conditional (CC). We suppose here that agents are "dialogically" co-operative in so far as an agent accepts to offer a substitution  $\Xi_{\text{Val}}$  for the sequence  $X$ .

### 3.2.5 Actions Applied to Commitments

**S24.**  $M, \text{Pa}, s_i \vdash \text{Create}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi))$  iff  
 $\exists \alpha \in \Phi_a \ \& \ M, \text{Pa}, s_i \vdash \text{Perform}(\alpha)\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi) \wedge G^- \neg \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$

This formula indicates that the creation of a social commitment is satisfied in the model  $M$  along a path  $\text{Pa}$  iff there is an action  $\alpha$  whose performance makes true the social commitment (i.e. the social commitment holds after the performance of the action  $\alpha$ ) and if in the past (before the creation moment of the social commitment), the social commitment was never satisfied in this model. This formula highlights the fact that the creation of a social commitment is an action in itself. Indeed, the action  $\alpha$  corresponds to the agent's utterance which creates the social commitment.

**S25.**  $M, \text{Pa}, s_i \vdash \text{Withdraw}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi))$  iff  
 $\exists \alpha \in \Phi_a, M, \text{Pa}, s_i \vdash X^-\text{F}^-\text{Create}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi))$   
 $\wedge \text{Perform}(\alpha) \neg \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)$ .

This formula indicates that an agent withdraws its social commitment for  $\varphi$  iff: (1) The agent has already created this social commitment. (2) The agent performs an action  $\alpha$  so that this social commitment does not hold at the current moment.

The semantics of the satisfaction operation depends on the type of the social commitment. In this paper we give only the semantics of the satisfaction of a PC as follows:

**S26.**  $M, \text{Pa}, s_i \vdash \text{Satisfy}(\text{Ag}_1, \text{PC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, p))$  iff  
 $\exists j : j \leq i \ \& \ M, \text{Pa}, s_j \vdash \text{Create}(\text{Ag}_1, \text{PC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, p))$   
 $\wedge M, \text{Pa}, s_i \vdash p \wedge s_i \in \text{Rpc}(\text{Ag}_1, \text{Ag}_2, s_j)$ .

A PC is satisfied iff it was already created and the propositional content is true in the moment that corresponds to the moment where the social commitment must be satisfied. This moment is denoted by  $s_i$  that defines the deadline. For example, if an agent commits at 14PM that it will rain at 16PM, we say that the social commitment is satisfied if it really rains at 16PM, if not, the social commitment is violated.

We can think of satisfaction and violation as two dual relations. Hence, we can express the relation between satisfaction and violation for any social commitment type. For example, for a PC this relation is specified by the formula:

$$\begin{aligned} \mathbf{S27.} \quad & M, Pa, s_i \vdash \text{Violate}(\text{Ag}_1, \text{PC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \text{ iff} \\ & \exists j : j \leq i \ \& \ M, Pa, s_j \vdash \text{Create}(\text{Ag}_1, \text{PC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \wedge s_i \in \text{Rpc}(\text{Ag}_1, \text{Ag}_2, s_j) \\ & \wedge M, Pa, s_i \vdash \neg \text{Satisfy}(\text{Ag}_1, \text{PC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)). \end{aligned}$$

This formula expresses the following property: If an agent violates its social commitment in the state  $s_i$  (which represents the deadline) along the path  $Pa$ , then this agent does not satisfy this social commitment in this state along this path and vice versa.

After introducing the different actions that the debtor can apply to its social commitments, we can define the semantics of an active social commitment as follows:

$$\begin{aligned} \mathbf{S28.} \quad & M, Pa, s_i \vdash \text{Active}(\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \text{ iff} \\ & M, Pa, s_i \vdash ((\neg \text{Violate}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \\ & \wedge \neg \text{Satisfy}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \wedge \neg \text{Withdraw}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi))) \\ & \cup \text{Create}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi))) \end{aligned}$$

This property indicates that a social commitment is active iff: (1) This social commitment was already created. (2) Until the current moment, the social commitment was neither violated, withdrawn nor satisfied. Therefore, once the social commitment is satisfied, violated or withdrawn, it becomes inactive.

### 3.2.6 Actions Applied to Commitment Contents

$$\begin{aligned} \mathbf{S29.} \quad & M, Pa, s_i \vdash \text{Accept-content}(\text{Ag}_2, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \text{ iff} \\ & M, Pa, s_i \vdash \text{Active}(\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \wedge \text{Create}(\text{Ag}_2, \text{SC}(\text{Id}_1, \text{Ag}_2, \text{Ag}_1, \varphi)) \end{aligned}$$

This formula indicates that the acceptance of the social commitment content  $\varphi$  by agent  $Ag_2$  is satisfied in the model  $M$  along a path  $Pa$  iff: (1) The social commitment is active on this path because we cannot act on a social commitment content if the social commitment is not active. (2) Agent  $Ag_2$  creates a social commitment whose content is  $\varphi$ . Therefore,  $Ag_2$  becomes committed towards the content  $\varphi$ .

$$\begin{aligned} \mathbf{S30.} \quad & M, Pa, s_i \vdash \text{Challenge-content}(\text{Ag}_2, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \text{ iff} \\ & \exists \alpha \in \Phi a, \exists \varphi' \in \mathcal{L}/\mathcal{L}sc \ \& \ M, Pa, s_i \vdash \text{Perform}(\alpha)\text{PC}(\text{Id}_1, \text{Ag}_2, \text{Ag}_1, ?\varphi) \\ & \wedge \text{Active}(\text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi)) \wedge \text{EX}^+\text{F}^+\text{Justify-content}(\text{Ag}_1, \text{SC}(\text{Id}_0, \text{Ag}_1, \text{Ag}_2, \varphi), \varphi') \end{aligned}$$

This formula indicates that the challenge of the social commitment content  $\varphi$  by an agent  $Ag_2$  is satisfied in the model  $M$  along a path  $Pa$  iff: (1) Agent  $Ag_2$  commits that  $?\varphi$ . Indeed,  $\text{PC}(\text{Id}_1, \text{Ag}_2, \text{Ag}_1, ?\varphi)$  states that “ $Ag_2$  does not know  $\varphi$  but it would like to know it”. (2) The challenged commitment is active on this path. (3) Agent  $Ag_1$

justifies in the future its social commitment for  $\varphi$ . Indeed, when we challenge a statement, we expect an answer from the speaker. Thus, in our semantics the fact that there is a possibility of having an answer is included in the meaning of the challenge. The operator E in  $(EX^+F^+ \text{Justify-content}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi'))$  allows us to capture the concept of possibility i.e. that there is a path along which  $Ag_1$  will justify its social commitment. This formula highlights the fact that the challenge of a social commitment content is an action in itself. As for the creation operation, the action  $\alpha$  corresponds to the production of the utterance that challenges the social commitment content.

### 3.2.7 Argumentation Relations

**S31.**  $M, Pa, s_i \vdash \text{Justify-content}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi')$  iff

$M, Pa, s_i \vdash \text{Active}(SC(Id_0, Ag_1, Ag_2, \varphi)) \wedge \text{Create}(Ag_1, SC(Id_1, Ag_1, Ag_2, \varphi' \therefore \varphi))$ .

This formula indicates that the justification of the social commitment content  $\varphi$  by an agent  $Ag_1$  is satisfied in the model  $M$  on a path  $Pa$  iff: (1) This social commitment is active on this path. (2) This agent creates on this path a social commitment whose content is  $\varphi'$  that supports the conclusion  $\varphi$ . In other words, an agent's social commitment towards another agent to make a content  $\varphi$  true is justified (by means of  $\varphi'$ ) iff the social commitment exists (has been created) and moreover a social commitment is created to establish an argument ( $\varphi', \varphi$ ), where  $\varphi'$  is committed to be true because according to the definition of the connector ( $\therefore$ ),  $\varphi'$  is true for  $Ag_1$ . The fact that this operator is included in the social commitment indicates that the agent is committed that  $\varphi'$  is true and then  $\varphi$  is true, i.e.  $\varphi$  is true because  $\varphi'$  is true. Indeed, agents have knowledge bases and the propositions that are not challenged can be used for justification (i.e. as supports of arguments). Hence, to end the chain of argumentation, agents use PCs that are not challenged any further. The justification operation is the basis of other argumentation operations. As shown by the following properties (S33 and S34), this is due to the fact that all the other operations are defined using this operation.

**S32.**  $M, Pa, s_i \vdash \text{Contradict-content}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi))$  iff

$(\exists \varphi' \in \mathcal{L}/\mathcal{L}sc: (M, Pa, s_i \vdash \text{Active}(SC(Id_0, Ag_2, Ag_1, \varphi))$

$\wedge \text{Create}(Ag_1, SC(Id_1, Ag_1, Ag_2, \varphi')) \wedge (\varphi' \therefore \neg\varphi))$

This formula indicates that an agent contradicts its previous social commitment whose content is  $\varphi$  if it creates another social commitment whose content is a logical conclusion of  $\neg\varphi$ , whereas its social commitment for  $\varphi$  is still active.

#### Properties:

**S33.**  $M, Pa, s_i \vdash \text{Attack-content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi')$  iff

$M, Pa, s_i \vdash \text{Active}(SC(Id_0, Ag_1, Ag_2, \varphi))$

$\wedge \text{Justify-content}(Ag_2, SC(Id_1, Ag_2, Ag_1, \neg\varphi), \varphi')$

This formula indicates that the attack of the social commitment content  $\varphi$  by an agent  $Ag_2$  is satisfied in the model  $M$  along a path  $Pa$  iff: (1) This social commitment

is active on this path. (2) This agent justifies along this path its social commitment whose content is  $\neg\varphi$ .

**S34.**  $M, Pa, s_i \vdash \text{Defend-content}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi')$  iff  
 $\exists\varphi'' \in \mathcal{L}/\mathcal{L}sc \ \& \ M, Pa, s_i \vdash \text{Active}(SC(Id_0, Ag_1, Ag_2, \varphi))$   
 $\wedge X^F \text{Attack-content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi'')$   
 $\wedge \text{Attack-content}(Ag_1, SC(Id_1, Ag_2, Ag_1, \varphi''), \varphi')$

This formula indicates that the defense of the social commitment content  $\varphi$  by an agent  $Ag_1$  is satisfied in the model  $M$  along a path  $Pa$  iff: (1) This social commitment is active on this path. (2) This agent attacks the attacker of the content of its social commitment.

### 3.2.8 Link Between Commitments and Arguments

Until now we gave the semantics of the main elements of our formalism. We can now formally establish the link between social commitments and arguments. This link is shown by the two following formulas:

**S35.**  $A(\text{Create}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi))) \Rightarrow$   
 $((\neg(F^+ \text{Contradict-content}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi))))$   
 $\wedge (F^+(\text{Challenge-content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi))) \Rightarrow \exists\varphi':$   
 $AX^+ F^+ \text{Justify-content}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi'))$   
 $\wedge (F^+ \text{Attack-content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi') \Rightarrow \exists\varphi'':$   
 $AX^+ F^+ \text{Defend-content}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi), \varphi''))))$

This formula provides the conditions generated by the creation of a social commitment on all paths. The agent must be in a position to check these conditions before creating a social commitment. Indeed, if an agent creates a social commitment, then it should not contradict itself during the conversation. It must also be able to justify its social commitment if it is challenged and to defend it if it is attacked. By establishing the link between social commitments and arguments, this formula reflects the deontic aspect of social commitments. These conditions are also valid for withdrawal, acceptance and refusal because their semantics is expressed in terms of the creation operation. On the other hand, an agent challenges a social commitment content if it has no argument for or against this content. Therefore, An agent challenges a social commitment content if it cannot accept or refuse it. Formally:

**S36.**  $A((\text{Active}(SC(Id_0, Ag_1, Ag_2, \varphi)) \wedge \neg \text{Accept-content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi))$   
 $\wedge \neg \text{Refuse-content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi)))$   
 $\Rightarrow \text{Challenge-content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi)))$

## 4 Logic-Based Protocols

Until now we defined a modal semantics for our approach in order to give a meaning to the different communicating actions. The purpose behind the definition of this semantics using temporal and dynamic logic is to be able to verify the correctness of the agent communication protocols. A protocol is correct iff it satisfies given

properties specified using our logic. Thus, the correctness problem is a model-checking one. In this section, we show how we can define these protocols on the basis of our approach. This definition enables us to establish the link between the semantics and the pragmatics.

---


$$\begin{array}{l}
R1 \wedge: \frac{\psi_1 \wedge \psi_2}{\psi_1 \psi_2} \quad R2 \vee: \frac{\psi_1 \vee \psi_2}{\psi_1 \psi_2} \quad R3 \vee: \frac{E(\psi)}{\psi} \quad R4 \neg: \frac{\neg\psi}{\psi} \quad R5 ? : \frac{?\psi}{\psi} \\
R6 \neg: \frac{A(\Phi)}{E(\neg\Phi)} \quad R7 < \alpha_p >: \frac{E(\Phi, \text{Perform}(\alpha)\varphi)}{E(\Phi, \varphi)} \\
R8 < C >: \frac{E(\Phi, \text{Create}(Ag_1, SC(Id_0, Ag_1, Ag_2, \varphi)))}{E(\Phi, SC(Id_0, Ag_1, Ag_2, \varphi))} \\
R9 < S_{PC} >: \frac{E(\Phi, \text{Satisfy}(Ag_1, PC(Id_0, Ag_1, Ag_2, \varphi)))}{E(\Phi, \varphi)} \\
R10 < V_{PC} >: \frac{E(\Phi, \text{Violate}(Ag_1, PC(Id_0, Ag_1, Ag_2, \varphi)))}{E(\Phi, \neg\varphi)} \\
R11 < Ch >: \frac{E(\Phi, \text{Challenge} - \text{content}(Ag_2, PC(Id_0, Ag_1, Ag_2, \varphi)))}{E(\Phi, PC(Id_1, Ag_2, Ag_1, ?\varphi))} \\
R12 < Acc >: \frac{E(\Phi, \text{Accept} - \text{content}(Ag_2, SC(Id_0, Ag_1, Ag_2, \varphi)))}{E(\Phi, SC(Id_1, Ag_2, Ag_1, \varphi))} \\
R13 < Att >: \frac{E(\Phi, \text{Attack} - \text{content}(Ag_2, PC(Id_0, Ag_1, Ag_2, \varphi), \varphi'))}{E(\Phi, PC(Id_1, Ag_2, Ag_1, \varphi' : \neg\varphi))} \\
R14 < Def >: \frac{E(\Phi, \text{Defend} - \text{content}(Ag_1, PC(Id_0, Ag_1, Ag_2, \varphi), \varphi'))}{E(\Phi, PC(Id_1, Ag_1, Ag_2, \varphi' : \varphi))} \\
R15 [PC_{Ag_1}]: \frac{E(\Phi, PC(Id_0, Ag_1, Ag_2, \varphi))}{E(\Phi, \varphi)} \quad R16 < \diamond >: \frac{E(\Phi, l)}{l, E(\Phi)} \\
R17 \wedge: \frac{E(\Phi, \varphi_1 \wedge \varphi_2)}{E(\Phi, \varphi_1, \varphi_2)} \quad R18 \vee: \frac{E(\Phi, \varphi_1 \vee \varphi_2)}{E(\Phi, \varphi_1) \ E(\Phi, \varphi_2)} \quad R19 ? : \frac{E(\Phi, ?\varphi)}{E(\Phi, \varphi)} \\
R20 X^+ : \frac{E(\Phi, X^+ \varphi_1, \dots, X^+ \varphi_n)}{E(\Phi, \varphi_1, \dots, \varphi_n)} \quad R21 \wedge : \frac{E(\Phi, \varphi_1 : \varphi_2)}{E(\Phi, \varphi_1, X^+(\neg\varphi_1 \vee \varphi_2))} \\
R22 \vee: \frac{E(\Phi, \varphi_1 U^+ \varphi_2)}{E(\Phi, \varphi_2) \ E(\Phi, \varphi_1, X^+(\varphi_1 U^+ \varphi_2))}
\end{array}$$


---

**Fig. 1.** Some tableau rules for DCTL\*<sub>CAN</sub> logic

Agent communication protocols are specified as a set of rules describing the entry condition, the dynamics and the exit condition of these protocols[4]. Using our logic,

these rules can be specified as action formulas: actions on social commitments, actions on social commitment contents and argumentation relations. These protocols can be specified using transition systems. The purpose of these transition systems is to describe not only the sequence of the allowed actions (like classical transition systems), but also the semantics of these actions. The semantics we use here is a tableau semantics [6] that we can consider as a simplification of the semantics defined in Section 3.2. This semantics is specified in terms of the decomposition of action formulas to sub-formulas using a set of inference or *proof rules* called *tableau rules*. The tableau rules are designed so that the formula is true if all the sub-formulas are true. The tableau semantics enables us to define *top-down proof systems*. The idea is: given a formula, we apply a proof rule and determine the sub-formulas to be proven. Fig. 1 shows some examples of tableau rules of our  $DCTL^*_{CAN}$  logic.  $\Phi$  is a set of path formulas  $\varphi_i$  and  $\psi$  is a state formula. The definition of the tableau rules is based on the semantics defined in Section 3.2.

The states of the transition systems are sub-transition systems that we call *semantic transition systems*. These automata describe the semantics of the actions labeling the entry transitions. Defining protocols using transition systems in such a way allows us to verify:

- 1- The correctness of the protocol (if the model of the protocol satisfies the properties that the protocol should specify).
- 2- The compliance to the semantics (if the specification of the protocol respects the semantics).

The definition of the transition systems of agent communication protocols is given by the following definitions:

**Definition 3.** A semantic transition system  $T'$  describing the semantics of an action formula is a 6-uple  $\langle S', F, L', R, \rightarrow^R, s'_0 \rangle$  where:  $S'$  is a set of states,  $F$  is a sub-set of the set of formulas from  $DCTL^*_{CAN}$  ( $F$  does not include the action formulas),  $L' : S' \mapsto F$  is the labeling state function,  $R \in \{\wedge, \vee, \neg, ?, \langle, X^+, X^-, PC_{Ag}, AC_{Ag}\}$  is the set of rule labels,  $\rightarrow^R \subseteq S' \times R \times S'$  is the transition relation,  $s'_0$  is the start state.

Intuitively, states  $s'$  contain the sub-formulas of the action formulas, and the transitions are labeled with operators associated with the formula of the source state. Semantic transition systems enable us to describe the semantics of formulas using sub-formulas connected by logical operators. Thus, there is a transition between states  $s'_i$  and  $s'_j$  iff  $L'(s'_j)$  is a sub-formula or a semantically equivalent formula of  $L'(s'_i)$ .

**Definition 4.** A transition system  $T$  for an agent communication protocol is a 6-uple  $\langle S, \wp, L, Act, \rightarrow^{Act}, s_0 \rangle$  where:  $S$  is a set of states,  $\wp$  is a set of semantic transition systems,  $L : S \rightarrow T'$  is the function associating a state  $s \in S$  to a semantic transition system  $T' \in \wp$  describing the semantics of the action labeling the entry transition,  $Act \in \{Create, Withdraw, Satisfy, Accept-content, Refuse-content, Challenge-content, Justify-content, Defend-content, Attack-content\}$  is the set of actions,  $\rightarrow^{Act} \subseteq S \times Act \times S$  is the transition relation,  $s_0$  is the start state.



The transitions are labeled with the actions applied to social commitments and to social commitment contents and the argumentation actions. We write  $s \rightarrow^* s'$  in lieu of  $\langle s, *, s' \rangle \in \rightarrow$  where  $* \in Act$ .

#### 4.1 Logical Properties to be Verified

The properties to be verified in the protocols specified by  $DCTL^*_{CAN}$  are action and temporal properties. For example we can verify if a model of an agent communication protocol satisfies the following property:

$$AG^+(Challenge-content(Ag_2, PC(Id_0, Ag_1, Ag_2, \varphi)) \Rightarrow \\ \exists \varphi': F^+Justify-content(Ag_1, PC(Id_0, Ag_1, Ag_2, \varphi), \varphi'))$$

This property indicates that if an agent  $Ag_2$  challenges the content of an  $Ag_1$ 's propositional commitment (PC), then  $Ag_1$  will justify this content. Another property capturing the deontic notion of social commitments is given by the following formula:

$$AG^+(Attack-content(Ag_2, PC(Id_0, Ag_1, Ag_2, \varphi), \varphi') \Rightarrow \exists \varphi'': \\ (F^+Defend-content(Ag_1, PC(Id_0, Ag_1, Ag_2, \varphi), \varphi'') \\ \vee F^+Attack-content(Ag_1, PC(Id_1, Ag_2, Ag_1, \varphi'), \varphi'')) \\ \vee F^+Accept-content(Ag_1, PC(Id_1, Ag_2, Ag_1, \varphi')))$$

Thus we can verify if a protocol satisfies the fact that if an agent  $Ag_2$  attacks the content of an  $Ag_1$ 's propositional commitment, then  $Ag_1$  will defend its commitment content, attack the  $Ag_2$ 's argument or accept it.

We are currently developing a model checking technique to verify these properties and the underlying semantics using a combination of tableau-based and automata-based model checking technique. This technique enables us to verify temporal and action properties by exploring the product graph of a labeled tableau automata representing the logical property and the transition system describing the protocol. The advantage of this technique is that the state space is explored in a need-driven fashion. The algorithm searches only the part of the state space that needs to be explored to prove or disprove a certain formula.

## 5 Related Work

Semantical considerations for agent communication have recently begun to find a significant audience in the MAS community. We can distinguish three kinds of semantics:

1- *Mentalistic semantics*: This subjective semantics is based on so-called agent's mental states. The best-known formalisms describing it are [8, 16, 19, 22]. KQML [12] and FIPA-ACL use this semantics to define a pre/post conditions semantic of communication acts. The advantage of this semantics is its compatibility with the formalisms used for reasoning about rational agents. However, the verification of such a semantics is not possible if we cannot have access to the agents' programs. In addition, this pre/post condition semantics offers no dynamic or operational description of agent communication. Because our approach is based on public and

argumentative concepts, the compliance verification can be made without having access to the agents' programs. The satisfaction and the violation of agents' social

commitments make it possible to determine if the agent respects our semantics. In addition, the agents' ability to justify their social commitments facilitates this verification. In addition, our semantics treats more explicitly the dynamic aspect of agent communication using the agents' actions on social commitments and on their contents.

2- Social semantics: This objective semantics was proposed by Singh [23, 24] as an alternative to the mentalistic one. Singh used CTL to propose a formal language and a model in which the notion of social commitment is described. Verdicchio and Colombetti [25] proposed an interesting logical model of social commitments by extending CTL\*. This model is based on the fact that agent communication should be analyzed in terms of communicative acts. Mallya et al. [17] used the temporal commitment structure specified by [13] to define some constraints in order to capture some operations on social commitments. Our logical model uses some ideas of [25] and it belongs to this kind of semantics, but it differs from these propositions in the following respects: 1) In our approach the social commitment semantics is not defined as an abstract accessibility relation, but as an accessibility relation that takes into account the satisfaction of the social commitment. The semantics is defined in terms of the deadline at which the social commitment must be satisfied. This way is more intuitive than the semantics defined by Singh. 2) We differentiate social commitments as static structures evaluated in states from the operations applied to social commitments as dynamic structures evaluated on paths. This enables us to describe more naturally the evolution of the agent communication as a system of states / transitions which reflects the interaction dynamics. 3) In our model, the strength of social commitments as a basic principle of agent communication does not result only from the fact that they are observable, but also from the fact that they are supported by arguments. The social commitment notion we formalize is not only a public notion but also a deontic one. The deontic aspect is captured by the fact that social commitments are considered as obligations. The agent is obliged to satisfy its social commitments, to behave in accordance with these social commitments and to justify them. It is also obliged not to contradict its social commitment contents during the conversation. 4) We capture in our semantics not only PCs, but the various types of social commitments. This enables us to have a greater expressivity and to capture the different types of SAs.

3- Argumentation-based semantics: This semantics is defined in [1] to capture the meaning of certain communication acts. It is based upon an argumentation system and on the formal dialectics. This semantics has the advantages of being simple and of taking into account the argumentation aspect of agent communication. In addition to the fact that this semantics does not take into account temporal and dynamic aspects in its formalization, it is different from our approach on several points: 1) It is based on an informal logic. 2) It is described in terms of pre/post conditions and it does not offer the meaning of the different communication acts. 3) The commitment notion

used in this semantics captures only the propositions stated by the agents. 4) Contrary to our approach, the satisfaction, violation, cancellation, attack and defense notions do not appear.

## 6 Conclusion and Future Work

In this paper we developed a formal semantics for our approach based on social commitments and arguments to model agents' interactions. We proposed a logical model ( $DCTL^*_{CAN}$ ) based on a combination of  $CTL^*$  and dynamic logic. The model captures different social commitment types, different actions applied to these social commitments and various argumentation relations. We showed how we can define protocols using  $DCTL^*_{CAN}$  logic in order to be able to check the correctness and the semantic compliance using a model-checking algorithm. We used the tableau semantics as a simplified semantics for the compliance verification. The tableau rules based on this semantics are needed for the translation of the logical properties to be verified to a tableau automata. Thus, our CAN formalism includes both pragmatic and semantic issues of agent communication.

We plan as future work to develop efficient model-checking algorithm for logic-based protocols. The idea we are investigating is to use an automata theoretic model-checking based on the emptiness problem of graphs. We intend to implement this algorithm using the CWB-NC verification tool.

**Acknowledgments.** We would like to thank Josée Desharnais from Laval University for her valuable comments about the logical model. We would also like to thank the three anonymous referees of AAMAS'04 for their detailed and very interesting comments that allowed us to improve the quality of this paper.

## References

1. Amgoud, L., Maudet, N., and Parsons, N. An argumentation-based semantics for agent communication languages. 15th Euro. Conf. on AI (2002) 38-42.
2. Bentahar, J., Moulin, B., and Chaib-draa, B. Commitment and Argument Network: a formal framework for representing conversation dynamics Logic and Dialogue. Kluwer (2004) (to appear).
3. Bentahar, J., Moulin, B., and Chaib-draa, B. Commitment and argument network: a new formalism for agent communication. Dignum, F. (ed.). *Advances in Agent Communication. Lecture Notes in Artificial Intelligence*, Vol. 2922. Springer-Verlag (2003) 146-165.
4. Bentahar, J., Moulin, B., and Chaib-draa, B. Specifying and Implementing a Persuasion Dialogue Game using Commitments and Arguments. Rahwan, I., Moraitis, P., and Reed, C. (eds.). *Argumentation in Multi-Agent Systems*. Springer Verlag (2004), in this volume.
5. Castelfranchi, C. Commitments: from individual intentions to groups and organizations. *The Int. Conf. ICMAS*. (1995) 41-48.
6. Cleaveland, R. Tableau-based model checking in the propositional mu-calculus. In *Acta Informatica*, Vol. 27(8). (1990) 725-747.

7. Chaib-draa, B., and Dignum, F. Trends in agent communication language. *Computational Intelligence*, Vol. 18(2). Oxford (2002) 89-101.
8. Cohen, P.R., and Levesque, H.J. Persistence, intentions and commitment. Cohen, P.R., Morgan, J., and Pollack, M.E. (eds.). *Intentions in Communication*. Cambridge (1990) 33-70.
9. Dastani, M., Hulstijn, J., and der Torre, L.V. Negotiation protocols and dialogue games. *The Belgium/Dutch AI Conf. Kaatsheuvel* (2000) 13-20.
10. Elvang-Goransson, M., Fox, J., and Krause, P. Dialectic reasoning with inconsistent information. *The 9th Conf. on Uncertainty in AI* (1993) 114-121.
11. Emerson, E.A. Temporal and Modal logic. *Handbook of Theoretical Computer Science*. van Leeuwen, J. (eds). Elsevier Science Publishers (1990) 995-1072.
12. Finin, T., Labrou, Y., and Mayfield, J. KQML as an agent communication language. Bradshaw, J.M. (ed.). *Software Agent*. MIT Press (1997) 291-316.
13. Fornara, N., and Colombetti, M. Operational specification of a commitment-based agent communication language. *The First Int. Conf. on AAMAS* (2002) 536- 542.
14. Harel, D. Dynamic logic. *Handbook of Philosophical Logic*. Gabbay, D.M., and Guenther, F. (eds), Vol. 2 (1984) 497-604.
15. Rahwan, I., Ramchurn, S.D., Jennings, N.R., McBurney, P., Parsons, S., and Sonenberg, L. Argumentation-based negotiation. *Knowledge Engineering Review* (2004) (to appear).
16. Van der Hoek, W., and Wooldridge, M. Towards a logic of rational agency. *Logic Journal of the IGPL*, Vol. 11(2) (2003) 133-157.
17. Mallya, A.U., Yolum, P., and Singh, M. Resolving Commitments Among Autonomous Agents. Dignum, F. (ed.). *Advances in Agent Communication*. *Lecture Notes in Artificial Intelligence*, Vol. 2922. Springer Verlag (2003) 166-182.
18. McBurney, P., Parsons, S., and Wooldridge, M. Desiderata for agent argumentation protocols. *The First Int. Conf. on AAMAS* (2002) 402-409.
19. Meyer, J.-J. Ch., van der Hoek, W., and van Linder, B. A logical Approach to the dynamics of commitments. *AI Journal*, Vol. 113 (1-2) (1999) 1-40.
20. Parsons, S., Sierra, C., and Jennings, N. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, Vol. 8(3) (1998) 261-292.
21. Pitt, J., and Mamdani, A. Communication protocols in multi-agent systems: a development method and reference architecture. Dignum, F., and Greaves, M. (eds.). *Issues in Agent Communication*. *Lecture Notes in Artificial Intelligence*, Vol. 1916. Springer Verlag (2000) 160-177.
22. Rao, A.S., and Georgeff, M.P. BDI agents: from theory to practice. *The Int. Conf. ICMAS* (1995) 312-319.
23. Singh, M.P. Agent communication languages: rethinking the principles. *IEEE Computer*, Vol. 31(12) (1998) 40-47.
24. Singh, M.P., A social semantics for agent communication language. Dignum, F., and Greaves, M. (eds.). *Issues in Agent Communication*. *Lecture Notes in Artificial Intelligence*, Vol. 1916. Springer Verlag (2000) 31-45.
25. Verdicchio, M., and Colombetti, M. A logical model of social commitment for agent communication. *The Second Int. Conf. on AAMAS* (2003) 528-535.