



HAL
open science

Logical formalization of social commitments: Application to agent communication languages (long version of AAMAS 2009)

Benoit Gaudou, Andreas Herzig, Dominique Longin

► To cite this version:

Benoit Gaudou, Andreas Herzig, Dominique Longin. Logical formalization of social commitments: Application to agent communication languages (long version of AAMAS 2009). [Research Report] IRIT/RR-2009-14-FR, IRIT : Institut de Recherche Informatique de Toulouse. 2009. hal-03474455

HAL Id: hal-03474455

<https://hal.science/hal-03474455>

Submitted on 10 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Logical formalization of social commitments: Application to agent communication languages

Benoit Gaudou¹ — Andreas Herzig² — Dominique Longin²

¹ UMI 209 UMMISCO, Institut de Recherche pour le développement (IRD)
Bondy, F-93143, France

et

MSI, Institut de la Francophonie pour l'Informatique (IFI), Ha Noi, Viet Nam
benoit.gaudou@alumni.enseiht.fr

² IRIT, LILAC staff
{herzig,longin}@irit.fr

May 26, 2009

Abstract

Nowadays most of the Agent Communication Languages are no longer defined in terms of the agents' mental attitudes, but in terms of social commitments, which avoids strong hypotheses on the agents' internal states and thus allows verification of the semantics. While social approaches have attracted a lot of attention, two drawbacks remain. First, the notion of commitment does not have a clear and unambiguous characterization. Second, commitments are completely unrelated to the agents' reasoning. The aim of this paper is to propose a solution for these two problems based on a combination of a BDI (Belief, Desire, Intention) logic with a logic of what is publicly grounded between agents.

Keywords

Agent communication languages, modal logic, grounding, commitments, BDI logic, speech act theory

Contents

Abstract & keywords	1
1 Introduction	5
2 What are commitments?	5
2.1 Commitment to do	6
2.1.1 Definition	6
2.1.2 Satisfaction of social commitments	7
2.1.3 Commitment life-cycle	7
2.2 Propositional commitment	8
2.2.1 Characterization	8
2.2.2 Condition of satisfaction	8
2.2.3 Life-cycle	9
2.3 Links between the two notions	9
3 The grounding logic	10
4 Formalization of commitments in the grounding logic	11
4.1 Propositional commitments	11
4.1.1 <i>Pending</i> commitment	12
4.1.2 <i>Canceled</i> commitment	12
4.1.3 <i>Violated</i> commitment	13
4.1.4 <i>Fulfilled</i> commitment	13
4.2 Commitments to do	14
4.2.1 <i>Pending</i> commitment	14
4.2.2 <i>Unset</i> commitment	15
4.2.3 <i>Canceled</i> commitment	16
4.2.4 <i>Fulfilled</i> commitment	16
4.2.5 <i>Violated</i> commitment	17
5 Example	17
6 Conclusion	18
References	19

1 Introduction

Agent Communication Languages (ACLs) have been the object of many investigations since two decades, particularly in two directions: the mentalist approach and the social approach. The former one is based on Bratman’s philosophical studies on human reasoning [1], characterized with mental states like belief and intention. His work has inspired the development of BDI (Belief, Desire, Intention) logics [19, 3, 20], afterwards used for the characterization of speech acts semantics. To resolve drawbacks of these mentalist approaches (*e.g.* the need of strong hypotheses on agents, like sincerity or cooperation and the non-verifiability), Singh [23] argues that ACLs semantics should be expressed in term of social and objective notions. This introduces the latter approach in which semantics is defined in term of social commitments [10, 8, 18].

[5] observes that, albeit recent ACLs are mostly based upon social commitments, only few has been done for the formal characterization of this concept. The aim of this paper is thus to give a characterization of social commitments in terms of social and individual attitudes and apply it to give a new semantics to commitment-based ACLs (for example Colombetti *et al.*’s one [10]). We argue that our reductionist approach of social commitments has many benefits. It expresses social commitments in terms of well-grounded mental attitudes, which brings a strong explicatory power. Moreover it provides a clear and unambiguous definition of social commitments but also conditions of its dynamic. Finally it allows to capture and manage in a unified account the social aspect of commitments and the private mental attitudes of agents that have incurred these commitments.

To meet this goal, we use the logic of grounding introduced in [11]. It is a BDI-like logic extended with an operator capturing what has been publicly established in a group during a conversation. It has been used to give a new semantics to FIPA ACL [7] in terms of what has been publicly expressed instead of in terms of private mental attitudes. Moreover it is able to capture Walton and Krabbe’s notion of propositional social commitment [26] in persuasive dialogue games. Beyond the formalization of commitment, we will be able to link commitment to public mental attitudes and thus bridge the gap between both ACL approaches.

We first study the notion of commitment (section 2) and present the grounding logic (section 3). Then we propose a formalization of social commitment and of its life-cycle (section 4). Finally we apply our work to a short example (section 5) before concluding.

2 What are commitments?

Commitment has various senses in the AI literature. First of all, it is important to make a distinction between internal commitment and social commitment. Internal commitment “refers to a relation between an agent and an action” [2]. In particular, commitment is used in this sense in the famous “Intention is choice with commitment” [3]¹. It captures the persistence of an agent’s choice in performing an action. Note that the agent can be either an individual agent or a group of agents (handled as a whole). This entails a subdistinction between personal and collective commitment. The latter one should not be confused with

¹But [3] does not only consider intention to do but also intention that a proposition is true, and thus commitment as the relation between an agent and a proposition.

social commitment which refers to a particular relation between agents. We note that it can be incurred by individual but also by collective agents. In this paper we focus only on the formalization of individual social commitments. Personal commitment is taken into account *via* the notion of intention defined below.

In Castelfranchi's view [2], social commitments are a relation between two agents only about an action: they can be named *commitments in action* or *commitments to do*. Following and extending Hamblin's work [15], [26] have described another kind of social commitment: propositional social commitments. Both notions are described in the sequel in a unified account.

2.1 Commitment to do

2.1.1 Definition

Social commitment to do links at least two agents: the agent who is committed (the *debtor* [9]) and the agent to whom the debtor is committed (the *creditor*). A third part can be involved in a commitment: the *witness*. Castelfranchi [2, p. 3] proposes following definition for the social commitment of the debtor i to the creditor j w.r.t. the action α :

“ i and j mutually know that i intends to do α and this is j 's goal, and that as for α j has specific rights on i (j is entitled by i to α).”

In opposition to Singh's [22] or Colombetti's [4] works about commitments in ACLs, we consider with Castelfranchi that social commitment is not primitive but can be defined in terms of agents' mental attitudes with additional deontic concepts. In the sequel, we discuss this definition.

Firstly, this definition imposes the mutual knowledge on i 's intention which logically implies that i 's intention actually holds. But as Castelfranchi says himself, the actual intention of i to perform the action α is neither necessary nor sufficient for his social commitment to do α : the entailment link between social commitment and individual intention is represented by the hypothesis that the agent is honest. We can thus weaken the definition by substituting the mutual knowledge with a notion capturing only the public feature of i 's intention: if i has a social commitment toward j to do α then it must be publicly grounded for both agents that i has the intention to do α . This notion of public ground should have as property that the public ground of an attitude does not implies that this attitude holds actually and that, in the case where agents are honest, it should entail mutual knowledge.

Secondly, the action α to which i is committed should be a goal of the agent j . We argue that this hypothesis is also too strong in general cases. Compare this to the speech act theory ([25, p. 182–183]) and in particular consider commissive acts. On the one hand, a promise produces a commitment of the speaker toward the hearer and requires what is promised to be good for the hearer, but on the other hand a threat has the same social result (creation of a commitment) but requires in contrary the object of the threat not to be good for the hearer. Thus, we need to distinguish between what we call desirable social commitment (when the fact that α be performed is a goal of the creditor), and undesirable social commitment (when the fact that α be performed is not a goal of the creditor). And the social commitment *à la* Castelfranchi corresponds to a desirable social commitment.

We do not need to define social commitment in such a restrictive way. Hence we do not specify in our definition of social commitment if the creditor j has or has not the goal that α be performed.

Finally even though we admit the importance of the deontic aspect in the characterization of social commitments, in this first attempt of formalization oriented toward ACLs application, we omit the deontic part of the commitment. In particular we will not reason about the various rights and obligations linking both commitment proponents.

To conclude, we consider that social commitment to do can be characterized by a public ground of both agents on the intention of one agent to perform an action α . Castelfranchi argues that commitments are created by communicative acts. This suggests to consider speech act theory and formalize not only commitment but also speech acts creating and managing them.

2.1.2 Satisfaction of social commitments

Once a commitment has been incurred, it is important to characterize cases where it is fulfilled. Castelfranchi [2] distinguishes two kinds of satisfaction: subjective satisfaction (when j believes that the action has been performed) and objective satisfaction (when the action has actually been performed). Objective satisfaction is clearly necessary. If j wrongly believes that α has been done, i stays with risks of blame hanging over him (by ongoing rights of j). But it is not sufficient. j needs to be aware of the objective satisfaction, or more precisely, if i can inform j of the commitment satisfaction, he should do it (we consider that the creditor has the right to be informed).

This condition of satisfaction has been quite simplified in ACLs: only the objective condition remains to avoid subjectivism [24]. But this condition can be verified only by an omniscient agent. We consider that this is problematic in multi-agent systems where agents have incomplete knowledge and interaction with other agents which should not depend on the validation of a central agent.

Thus in the sequel we will impose a public condition in order that every creditor is aware when a commitment is satisfied. Note that in the case where every action is public (that is soundly and completely perceived), both conditions are equivalent.

Colombetti *et al.* among others refine Castelfranchi's account by adding additional states in the commitment life-cycle.

2.1.3 Commitment life-cycle

ACLs semantics resting on commitments are indeed based on their life-cycle: each speech act is characterized by the commitment it produces, or more precisely by the changes in commitment states it induces.

After its creation by a particular communicative act the social commitment is in the state *pending* [10] (also named *created* in [8, 18]). This state corresponds to the generic notion of social commitment described above. Afterwards it becomes either *fulfilled* [10, 8] (*discharged* [18]) when satisfaction condition holds or *violated* [10, 8] if the action has not been performed. It can also be *canceled* by i or *released*² by j [8, 18]. Colombetti

²In the sequel, we stay close to Colombetti *et al.*'s account and do not take into account this state.

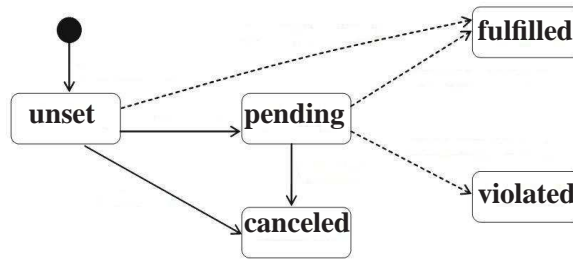


Figure 1: Action commitments life-cycle [10]

et al. introduce an intermediate *unset* state corresponding to a commitment request of *i* waiting for being accepted by *j*.

Figure 1 summaries Colombetti *et al.*'s account [10]. We can note that there is no arrow from initial state to pending state. Indeed solid arrows correspond to low-level actions *create commitment*, which creates a *unset* commitment and *set commitment*, which alters the state of a commitment. A *pending* commitment is thus the result of the creation of a new commitment and the change of its state to pending.

All these considerations were about commitment to do. We now adapt this account to take into account the second kind of social commitments that are propositional commitments.

2.2 Propositional commitment

2.2.1 Characterization

Following Walton & Krabbe [26], propositional commitments occur in a dialogue and correspond to what has been asserted by a party and then must be defended, for example by arguments, when it is attacked. At any stage of the dialogue, they correspond to a kind of ground on which will be settled the next steps. They serve as glue to maintain dialogue coherence. We base our analysis of propositional commitments on the one presented in [12] for persuasion dialogues.

Colombetti *et al.* argue that while propositional commitments and commitments to do are different in many aspects, his account can manage both kinds of commitments. We show in the sequel which modifications need to be added.

2.2.2 Condition of satisfaction

A commitment to do is fulfilled (resp. violated) as soon as its content becomes true (resp. false) [10] and Colombetti *et al.* argue that the same mechanism can be applied to propositional commitment. But as a propositional commitment represents what has been uttered, it appears that the satisfaction of such a commitment cannot be related to the validity of its content. For example, a speaker can utter a false proposition but being convinced of its truth. We cannot say that this speaker has violated his commitment if in the sequel of the dialogue he stays coherent with it. Satisfaction and violation is rather determined by the capacity of the agent to maintain it, *i.e.* to stay coherent in his discourse by avoiding to express contradictions.

Moreover, satisfaction may be impossible to check: consider the following example where John tells to his wife Mary: “I find you are very beautiful tonight”. John has incurred a propositional commitment to the proposition: “John believes that Mary is beautiful tonight”. This commitment cannot be verified by any agent, because Mary cannot know John’s private mental attitudes. Defining an ACL based on commitment with such a flaw, would suffer from the same drawbacks as mentalist approaches.

2.2.3 Life-cycle

The life-cycle for commitments to do does not fully transfer to propositional commitments. In particular, the state *unset* is not reachable by propositional commitments because Colombetti’s only speech act involving creation of a propositional commitment is the Inform speech act creating a *pending* commitment (and no *unset* one). Moreover this state would result from any kind of directive speech act, which does not manage, by definition [25], propositions but only actions.

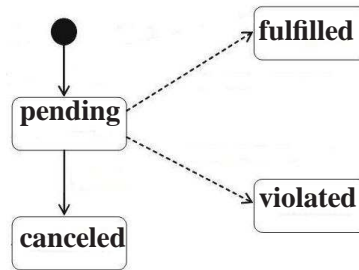


Figure 2: Propositional commitments life-cycle

The *unset* state is thus irrelevant for propositional commitment, and can be dropped from the life-cycle describing propositional commitment (Figure 2). A *pending* propositional commitment is created by an assertive speech act. It can be *canceled* also with a Cancel speech act.

As discussed above, the fulfilled and violated states are not as simple to describe as they appear. In particular the content truth value does not appear relevant to discriminate both above states. We rather consider Walton & Krabbe’s win/loss conditions [26]. In a PPD_0 dialogue, an agent wins the game if the opponent concedes his thesis and loses it if he retracts his own initial thesis. Otherwise, he has neither won nor lost the game.

2.3 Links between the two notions

As far as we are concerned, no formal link has been exhibited between propositional commitment and commitment to do. For example by uttering: “I promise to take out the garbage”, John has incurred a commitment to take out the garbage. Moreover, following Speech Act Theory [25], he has also expressed at least that he has the intention to perform the intention to take out the garbage. Thus he is also propositionally committed on his intention. He can be sanctioned if he does not perform this action, but also if he utters that he does not have this intention (in this case the sanction would be at the dialogue level: he

would appear to be incoherent and thus untrustworthy). As we will show in the formalization, by incurring a commitment to do α , an agent incurs *de facto* some propositional commitments. In this example, John is committed to his intention to perform the action.

We present in the next section the logical framework used to formalize social commitments.

3 The grounding logic

[11] introduces a logic capturing what has been publicly grounded in a group of agents by extending the logic of intention developed in [16] with an modal operator Grd . Let $AGT = \{i, j, \dots\}$ be a finite set of agents and ACT the set of actions. The expression $(i:\alpha)$ denotes that i is the author of action α .

For the sake of space, we will not present in detail the underlying usual possible worlds semantics and we present here only a selection of axioms ([11] presents the whole semantics and axiomatics). The logic is based on following five classes of primitive normal modal operators.

$Grd_I \varphi$ means that “ φ is publicly grounded for the group I ”. For a group I reduced to a singleton $\{i\}$, $Grd_{\{i\}}$ is identified with the usual belief operator *à la* Hintikka [17]: $Grd_{\{i\}} \varphi$ (also noted $Grd_i \varphi$ for convenience) means that i believes φ . The axiomatization is as follows ($I' \subseteq I \subseteq AGT$):

$$\begin{aligned}
 Grd_I \varphi &\rightarrow \neg Grd_I \neg \varphi && \text{(D}_{Grd_I}\text{)} \\
 Grd_I \varphi &\rightarrow Grd_{I'} Grd_I \varphi && \text{(4}_{Grd_I, Grd_{I'}}\text{)} \\
 \neg Grd_I \varphi &\rightarrow Grd_{I'} \neg Grd_I \varphi && \text{(5}_{Grd_I, Grd_{I'}}\text{)} \\
 Grd_I \varphi &\rightarrow Grd_I Grd_{I'} \varphi, \text{ with } \varphi \text{ objective for } i && \text{(WR)} \\
 (\bigwedge_{i \in I} Grd_I Grd_i \varphi) &\rightarrow Grd_I \varphi && \text{(CG)}
 \end{aligned}$$

Thus, Grd_I operators are defined as rational ($\text{(D}_{Grd_I}\text{)}$), public for every subgroup ($\text{(4}_{Grd_I, Grd_{I'}}\text{)}$) and ($\text{(5}_{Grd_I, Grd_{I'}}\text{)}$) and are formed by a joint agreement of each member ((WR) and (CG)):

In short, an objective formula for i is a formula for which the following equivalence does not hold: $\varphi \leftrightarrow Grd_i \varphi$. Thanks to Axioms ($\text{(D}_{Grd_I}\text{)}$) and ($\text{(4}_{Grd_I, Grd_{I'}}\text{)}$), it is public for a group I that φ is grounded in the sense that φ is grounded if and only if it is mutually believed [6] by the group I that it is grounded.

$Choice_i \varphi$ means “ i chooses that φ ”. It is defined in a KD45 logic. We consider that agents are aware of their own mental attitudes. Moreover we admit strong rationality features for the choice. We have thus the following links between belief and choice:

$$\begin{aligned}
 Grd_i \varphi &\rightarrow Choice_i \varphi && \text{(SRat)} \\
 Choice_i \varphi &\leftrightarrow Grd_i Choice_i \varphi && \text{(PI}_{Choice_i}\text{)} \\
 \neg Choice_i \varphi &\leftrightarrow Grd_i \neg Choice_i \varphi && \text{(NI}_{Choice_i}\text{)}
 \end{aligned}$$

From choice, the intention can be defined: agent i has the intention that φ holds iff he chooses to believe φ in the future, he does not believe φ and he believes that he will not believe φ anyway (*i.e.* without acting himself):

$$Int_i \varphi \stackrel{\text{def}}{=} Choice_i F Grd_i \varphi \wedge \neg Grd_i \varphi \wedge \neg Grd_i F Grd_i \varphi \quad \text{(Def}_{Int_i}\text{)}$$

Due to the previous links between choice, belief and the definition of the intention, the following formulas are theorems of our logic:

$$\begin{aligned}
 \vdash Int_i \varphi \rightarrow \neg Grd_i \varphi & \quad (\text{Rat}_{Int_i}) \\
 \vdash Int_i \varphi \leftrightarrow Grd_i Int_i \varphi & \quad (\text{PI}_{Int_i}) \\
 \vdash \neg Int_i \varphi \leftrightarrow Grd_i \neg Int_i \varphi & \quad (\text{NI}_{Int_i})
 \end{aligned}$$

F (resp. P) are operators of linear temporal logic LTL. $F\varphi$ (resp. $P\varphi$) means that φ will be (resp. has been) true at some instant in the future (resp. in the past).

$After_\alpha$ and $Before_\alpha$ are dynamic operators: $After_\alpha$ (resp. $Before_\alpha$) reads: “ φ holds after (resp. held before) every execution of the action α ”. They are operators of tense logic \mathbf{K}_t . The dual $Done_\alpha \varphi \stackrel{\text{def}}{=} \neg Before_\alpha \neg \varphi$ expresses that the action α has been performed before which φ held. Note that $Done_\alpha \top$ means that α has been performed. Formula $Int_i Done_\alpha \top$ denotes i 's intention to do action α .

4 Formalization of commitments in the grounding logic

The aim of this section is to formalize both social commitments in their different states and actions allowing to reach these states. As mentioned above, we although make distinction between propositional commitments and commitment to do in the following formalization. Moreover links between those two kinds of commitment will be highlighted.

In the sequel, we stay close to Castelfranchi's view of links between commitments and actions: we base out our formalization mainly on speech act theory by characterizing commitments with help of speech acts inducing them. For example, assertive speech acts induce *pending* propositional commitments. We study the primitive one, that is Assert. Likewise directive ones induce *unset* action commitments and commissive speech acts *pending* action commitments. The primitives are Direct and Commit.

4.1 Propositional commitments

For description of propositional commitments, we reuse some definitions introduced in [12] for persuasion dialogues. Note that we are not limited to this particular kind of dialogue and thus all these definitions can be used in every type of dialogue.

We need to simplify Colombetti *et al.*'s general commitment life-cycle and adapt it to propositional commitments. In particular we drop the *unset* state that we do not consider relevant for them. We consider that transitions between states are not the result of low-level actions managing commitments, such as in [9], but rather of speech acts of which we will give the semantics of the primitive ones. Their semantics will be provided in terms of Feasibility Precondition (FP) and Intentional Effect (IE). Feasibility Precondition describes the condition under which an action can be performed; Intentional Effect describes the illocutionary intention that, when recognized by the addressee of the act, identifies the performed speech act. (This is the Gricean's view of speech act understanding.)

For convenience, we introduce the predicate $C(state, i, j, \varphi)$ to represent a propositional social commitment of i toward j in a state $state$ about a proposition φ . Moreover similar notation is used for commitment to do by replacing proposition φ by any action α .

4.1.1 Pending commitment

For [26], propositional commitments are produced by assertions. Following speech act theory [25], Assert is the primitive assertive speech act and thus the simplest one. It has “the preparatory condition that the speaker has reasons or evidence for the truth of the propositional content” and “the sincerity condition that the speaker believes the propositional content” [25, p. 125].

As illocutionary effect, we impose with Vanderveken [25] that the speaker expresses the preparatory and the sincerity condition. We simplify this account by combining both conditions³ ($Grd_{\{i,j\}} Grd_i \varphi$). Contrarily to Inform speech act defined in [13]⁴, Assert does neither suppose anything about the group of interlocutors (*i.e.* that it is not aware of φ), nor any speaker’s intention. Thus we only impose as precondition that the speaker stays consistent; this means that he has neither asserted nor conceded the contrary. In [26] and [12] terminology, to say that an agent concedes a proposition φ means that he makes public the fact that he envisages that φ holds ($Grd_{\{i,j\}} \neg Grd_i \neg \varphi$). The consistency condition can thus be written: $\neg Grd_{\{i,j\}} Grd_i \neg \varphi \wedge \neg Grd_{\{i,j\}} \neg Grd_i \varphi$. As the latter term implies the former, we define the Assert speech act by:

DEFINITION. $\langle i, j, \text{Assert}, \varphi \rangle$

FP: $\neg Grd_{\{i,j\}} \neg Grd_i \varphi$

IE: $Grd_{\{i,j\}} Grd_i \varphi$

Note that our account conforms to speech act theory in what concerns the links between Assert and Inform speech acts. Inform is an extension of Assert, in the sense that the precondition (resp. effect) of the former include at least all ones of the latter.

As mentioned above, a propositional *pending* commitment is the result of an assertion. We remark that we can thus identify *pending* propositional commitments with Strong Commitments defined in [12]. Thus we have the definition:

DEFINITION. $C(\text{pending}, i, j, \varphi) \stackrel{\text{def}}{=} Grd_{\{i,j\}} Grd_i \varphi$

Following theorem is a consequence of the fact that by performing a speech act an agent expresses the sincerity conditions of the act: when an agent is committed on φ , he is also committed to the fact that he believes it, and *vice versa*:

$$\vdash C(\text{pending}, i, j, \varphi) \leftrightarrow C(\text{pending}, i, j, Grd_i \varphi)$$

4.1.2 Canceled commitment

A *canceled* commitment is a *pending* commitment that has been retracted. It is thus the result of a Cancel speech act. (To cancel means to take back what one says.) This act drops the *pending* commitment ($Grd_{\{i,j\}} Grd_i \varphi$).

DEFINITION. $\langle i, j, \text{Cancel}, \varphi \rangle$

³A reason to merge both conditions is that belief is shaped by evidence. (See [14] for instance.) As we do not need to exhibit these reasons, we can thus consider that they are embedded in the belief.

⁴In that work, for the Inform act: $\langle i, J, K, \text{Inform}, \varphi \rangle$

FP: $\neg Grd_K Grd_J \varphi \wedge \neg Grd_K Int_i Grd_J \varphi \wedge \neg Grd_K \neg Grd_i \varphi$

IE: $Grd_K Grd_i \varphi \wedge Grd_K Int_i Grd_J \varphi$

FP: $Grd_{\{i,j\}} Grd_i \varphi$

IE: $\neg Grd_{\{i,j\}} Grd_i \varphi$

Thus an agent has a commitment about φ in the *canceled* state when he has canceled (*i.e.* $\neg Grd_{\{i,j\}} Grd_i \varphi$ holds) a *pending* commitment incurred in the past ($P Grd_{\{i,j\}} Grd_i \varphi$), without $\neg\varphi$ having been conceded ($\neg Grd_{\{i,j\}} \neg Grd_i \varphi$, this condition distinguishing *canceled* and *violated* commitments as detailed in the following section):

DEFINITION. $C(\text{canceled}, i, j, \varphi) \stackrel{\text{def}}{=} P Grd_{\{i,j\}} Grd_i \varphi \wedge \neg Grd_{\{i,j\}} Grd_i \varphi \wedge \neg Grd_{\{i,j\}} \neg Grd_i \varphi$

Such a *canceled* commitment can be the result of the following action:

$$After_{\langle i,j, \text{Cancel}, \varphi \rangle} C(\text{canceled}, i, j, \varphi)$$

4.1.3 Violated commitment

Following discussion of Section 2.2.3, to characterize a *violated* commitment, we extend the loss **PPD**₀-rule: an agent loses the game when he retracts (*i.e.* cancels) his own thesis (say φ). As additional condition for a *violated* commitment, the agent should concede the converse ($Grd_{\{i,j\}} \neg Grd_i \varphi$). This represents the case of a public contradiction⁵.

DEFINITION. $C(\text{violated}, i, j, \varphi) \stackrel{\text{def}}{=} P Grd_{\{i,j\}} Grd_i \varphi \wedge \neg Grd_{\{i,j\}} Grd_i \varphi \wedge Grd_{\{i,j\}} \neg Grd_i \varphi$

Such a *violated* commitment can follow from the following sequence of actions:

$$After_{\langle i,j, \text{Cancel}, \varphi \rangle} After_{\langle i,j, \text{Assert}, \neg\varphi \rangle} C(\text{violated}, i, j, \varphi)$$

4.1.4 Fulfilled commitment

A proponent i wins a **PPD**₀ game when the opponent j concedes his thesis. Thus a propositional commitment becomes *fulfilled* when the target agent concedes its content ($Grd_{\{i,j\}} \neg Grd_j \neg\varphi$):

DEFINITION. $C(\text{fulfilled}, i, j, \varphi) \stackrel{\text{def}}{=} Grd_{\{i,j\}} Grd_i \varphi \wedge Grd_{\{i,j\}} \neg Grd_j \neg\varphi$

Such a *fulfilled* commitment can be the result of the following sequence of actions (among others):

$$After_{\langle i,j, \text{Assert}, \varphi \rangle} After_{\langle j,i, \text{Assert}, \varphi \rangle} C(\text{fulfilled}, i, j, \varphi)$$

Note that the formula $Grd_{\{i,j\}} Grd_i \varphi \wedge Grd_{\{i,j\}} \varphi$ is a particular case of fulfilled commitment (thanks to (**D** _{Grd_I}), (**WR**) and (**RN** _{G})).

⁵Our logical framework forbids contradictory commitments ($C(\text{pending}, i, j, \varphi)$ and $C(\text{pending}, i, j, \neg\varphi)$), thus to violate a commitment without becoming in contradiction, an agent has to drop his commitment before conceding the contrary.

4.2 Commitments to do

For commitments to do, we stay close to Clombetti's life-cycle as it appears to be well adapted for this kind of commitments, but we will not consider transitions via low-level actions, but via high-level speech acts.

4.2.1 Pending commitment

Following Colombetti *et al.*, there are two ways to incur a *pending* commitment to do: either by committing oneself spontaneously with *e.g.* a speech act of promise, or by accepting to commit on a directed action. Thus in both cases it results from a commissive speech act. (Following Vanderveken [25], the illocutionary force *accept* is a force of commissive type.) As above we consider the primitive commissive Commit speech act. A commissive speech act has “the condition that the propositional content represents a future course of action of the speaker”, “the preparatory condition that the speaker is capable of carrying out that action” and “the sincerity condition that he intends to carry it out” [25, p. 125–126]. Thus by performing a commissive speech act, an agent expresses at least that he has the intention to perform the action ($Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top$) and that he believes that he can carry out the action ($Grd_{\{i,j\}} \neg Grd_i \neg F Grd_i Done_{(i:\alpha)} \top$). Note that $Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top \rightarrow Grd_{\{i,j\}} \neg Grd_i \neg F Grd_i Done_{(i:\alpha)} \top$ is a theorem.⁶ As precondition, we only impose that the agent stays coherent (*i.e.* that the performance of this action does not involve inconsistency). Thus we have:

DEFINITION. $\langle i, j, \text{Commit}, Done_{(i:\alpha)} \top \rangle$

FP: $\neg Grd_{\{i,j\}} \neg Int_i Done_{(i:\alpha)} \top$

IE: $Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top$

We can thus characterize a *pending* commitment to do.

DEFINITION. $C(\text{pending}, i, j, \alpha) \stackrel{\text{def}}{=} Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top$

Due to the equivalence: $Int_i \varphi \leftrightarrow Grd_i Int_i \varphi$ (PI_{Int_i}), when an *unset* commitment is incurred, a propositional commitment appears, which allows us to represent the example of John and his promise to take out the garbage.

THEOREM.

$\vdash C(\text{pending}, i, j, \alpha) \leftrightarrow C(\text{pending}, i, j, Int_i Done_{(i:\alpha)} \top)$

Remarks about the Promise It is interesting to consider the act Promise because it is often used to represent commissive speech acts. As particular commissive act, it inherits the preconditions and postconditions of Commit. Moreover it has its particular features (see [21] and [25, p. 182]):

⁶ $\vdash Int_i \varphi \rightarrow Choice_i F Grd_i \varphi$, by (Def_{Int_i})
 $\vdash Choice_i \varphi \rightarrow \neg Grd_i \neg \varphi$, by (SRat) and (D_{Choice_i})
 $\vdash Int_i Done_{(i:\alpha)} \top \rightarrow \neg Grd_i \neg F Grd_i Done_{(i:\alpha)} \top$
 $\vdash Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top \rightarrow$
 $Grd_{\{i,j\}} \neg Grd_i \neg F Grd_i Done_{(i:\alpha)} \top$

- the speaker is aware that the hearer prefers that action is performed ($Grd_{\{i,j\}} Grd_i Choice_j Done_{(i:\alpha)} \top$)⁷;
- it is not obvious to the speaker and the hearer that the speaker would perform the action without having promising to do it ($Grd_{\{i,j\}} Grd_i \neg Grd_j Done_{(i:\alpha)} \top$);
- “it involves a special kind of commitment, namely the explicit undertaking of an obligation that may remain tacit in other types of commitment”.

The last precondition can be formalized by the formula $Grd_{\{i,j\}} Obl_{i,\{j\}} F Done_{(i:\alpha)} \top$, with $Obl_{i,\{j\}}$ an operator of obligation directed from i toward j . Its integration is out of the scope of this paper.

To complete this characterization, we impose as preconditions that the agent stays consistent and thus that he has not expressed yet contrary intention and that he has no contrary obligation. Moreover we consider he is not committed yet to perform the action.

4.2.2 Unset commitment

An *unset* commitment of j toward i corresponds to the particular social relation that results from the performance by agent i of a request, an order or another directive speech act. To formalize such a commitment, we need to consider what is primitive in directive speech acts and thus to formalize Direct speech act. Directives are close to commissives: the distinction is only on the author of the action [25]. A directive speech act has “the condition that the propositional content represents a future course of action of the hearer”, “the preparatory condition that the hearer can carry out that action” and “the preparatory condition that the speaker desires or wants the hearer to carry it out” ([25, p. 126]). Thus by performing a directive speech act, an agent expresses at least that he has the intention that the hearer performs the action ($Grd_{\{i,j\}} Int_i Done_{(j:\alpha)} \top$) and that he believes that the hearer can carry out the action ($Grd_{\{i,j\}} \neg Grd_i \neg F Grd_i Done_{(j:\alpha)} \top$). As previously the first formula implies the second one. As precondition, we only impose that the agent stays coherent. Thus we have:

DEFINITION. $\langle i, j, \text{Direct}, Done_{(j:\alpha)} \top \rangle$

FP: $\neg Grd_{\{i,j\}} \neg Int_i Done_{(j:\alpha)} \top$

IE: $Grd_{\{i,j\}} Int_i Done_{(j:\alpha)} \top$

We can thus characterize an *unset* commitment. We note that contrarily to Colombetti *et al.*, it is directed from i toward j . We argue that after a request of i it is doubtful that agent j is the debtor of any commitment. We consider that in this case, i should be committed to something, *i.e.* his intention that j performs the action.

DEFINITION. $C(\text{unset}, i, j, \alpha) \stackrel{\text{def}}{=} Grd_{\{i,j\}} Int_j Done_{(i:\alpha)} \top$

As above, this commitment is equivalent to a propositional commitment:

THEOREM.

$\vdash C(\text{unset}, i, i, \alpha) \leftrightarrow C(\text{pending}, j, i, Int_j Done_{(i:\alpha)} \top)$

⁷This condition allows to distinguish Promise from Threat. The operator of choice represents what the agent prefers and by extension what is good for him.

For example, when Mary requests John to take out the garbage, she expresses that she wants John performs the action $takeOutGarbage$, i.e. $Grd_{\{m,j\}} Int_m Done_{(j:takeOutGarbage)} \top$. She is also committed about her expression of this intention: as described above, she cannot act as if she does not want John takes out the garbage. She has thus the propositional commitment: $C(pending, m, j, Int_m Done_{(j:takeOutGarbage)} \top)$

4.2.3 Canceled commitment

For Colombetti *et al.*, an *unset* commitment and a *pending* commitment can be canceled. Indeed if the cancelation of a *pending* commitment seems to be a genuine cancel action, the action inducing a *canceled* commitment from an *unset* one is rather a refusal. When Mary requests John to take out the garbage, he will refuse (and not cancel some commitment) to do this chore.

Refuse is also a commissive speech act. It is the negative counterpart of acceptance of a request. Thus a previous performance of a Request (i.e. $Grd_{\{i,j\}} Int_j Done_{(i:\alpha)} \top$) is a precondition for this speech act. By refusing to perform an action, an agent expresses that he does not want to perform the requested action (i.e. $Grd_{\{i,j\}} \neg Choice_i Done_{(i:\alpha)} \top$).

DEFINITION. $\langle i, j, Refuse, Done_{(i:\alpha)} \top \rangle$

FP: $Grd_{\{i,j\}} Int_j Done_{(i:\alpha)} \top \wedge \neg Grd_{\{i,j\}} \neg Choice_i Done_{(i:\alpha)} \top$

IE: $Grd_{\{i,j\}} \neg Choice_i Done_{(i:\alpha)} \top$

We can thus define the refused commitment to do:

DEFINITION. $C(refused, i, j, \alpha) \stackrel{def}{=} Grd_{\{i,j\}} \neg Choice_i Done_{(i:\alpha)} \top \wedge P Grd_{\{i,j\}} Int_j Done_{(i:\alpha)} \top$

But if John has accepted to take out the garbage, he has incurred a *pending* commitment ($Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top$). As he does not want anymore to perform this action, he must cancel his commitment. We argue that agents can use the same action to cancel both kinds of commitment. As preconditions of Cancel are verified ($Grd_{\{i,j\}} Grd_i Int_i Done_{(i:\alpha)} \top \leftrightarrow Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top$), he can perform this speech act to be disengaged (i.e. $\neg Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top$).

DEFINITION. $C(canceled, i, j, \alpha) \stackrel{def}{=} \neg Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top \wedge P Grd_{\{i,j\}} Int_i Done_{(i:\alpha)} \top$

We can note that the distinction between a *canceled* propositional commitment and a *canceled* commitment to do comes from the difference between the satisfaction/violation conditions.

4.2.4 Fulfilled commitment

We simply consider that a *pending* commitment to do is fulfilled as soon as it is public that the debtor has performed the action. But we need to consider also an *unset* commitment immediately fulfilled by the performance of the requested action.

DEFINITION. $C(\text{fulfilled}, i, j, \alpha) \stackrel{\text{def}}{=} P(C(\text{pending}, i, j, \alpha) \vee C(\text{unset}, i, j, \alpha)) \wedge Grd_{\{i,j\}} Done_{(i:\alpha)} \top$

We stay close to Colombetti *et al.*'s fulfillment condition but we consider that each agent must be able to determine which is the current commitment state. We remark that for public actions (*i.e.* actions α such as $Grd_{\{i,j\}} Done_{(i:\alpha)} \top \leftrightarrow Done_{(i:\alpha)} \top$), a *pending* commitment is fulfilled as soon as the action has been performed.

Moreover as soon as it is public that the commitment has been fulfilled, the *unset* and *pending* commitments are dropped (if they exist). Indeed if the action α is public then every agent is aware of its performance ($Grd_i Done_{(i:\alpha)} \top \wedge Grd_j Done_{(i:\alpha)} \top$ holds) which implies that both intentions are dropped ($\neg Int_i Done_{(i:\alpha)} \top \wedge \neg Int_j Done_{(i:\alpha)} \top$). In the case where the action is not public, $Done_{(i:\alpha)} \top$ is an objective formula, and thus thanks to (WR) we have $Grd_{\{i,j\}} Done_{(i:\alpha)} \top \rightarrow Grd_{\{i,j\}} Grd_i Done_{(i:\alpha)} \top \wedge Grd_{\{i,j\}} Grd_j Done_{(i:\alpha)} \top$, which induces straight $Grd_{\{i,j\}} \neg Int_i Done_{(i:\alpha)} \top \wedge Grd_{\{i,j\}} \neg Int_j Done_{(i:\alpha)} \top$.

THEOREM.

$C(\text{fulfilled}, i, j, \alpha) \rightarrow \neg C(\text{unset}, i, j, \alpha) \wedge \neg C(\text{pending}, i, j, \alpha)$

4.2.5 Violated commitment

We consider that a commitment is violated by the debtor when it becomes grounded that he will never be able to perform it. It is the case when he admits publicly that he will never be able to perform it.

DEFINITION. $C(\text{violated}, i, j, \alpha) \stackrel{\text{def}}{=} PC(\text{pending}, i, j, \alpha) \wedge Grd_{\{i,j\}} Grd_i \neg F Grd_i Done_{(i:\alpha)} \top$

As previously, when a commitment is violated, the *pending* one is dropped .

THEOREM. $C(\text{violated}, i, j, \alpha) \rightarrow \neg C(\text{pending}, i, j, \alpha)$

In the following section, we illustrate our formalization on a dialogue example.

5 Example

We now illustrate the formalization of commitments in our logical framework with a case study. We give details of some key sentences of a simple car selling example [13]. We consider thus a dialogue between a seller s and a buyer b . We focus on commitments to do and in particular on both way (from an *unset* or a *pending* commitment) to have a fulfilled commitment.

$s \rightarrow b$: **Information about discount**

$\langle s, b, \text{Assert}, \neg \text{discounts} \rangle$

Effect:

$C(\text{pending}, s, b, \neg \text{discounts})$

The seller s has incurred a propositional commitment on the impossibility of discounts, while he may have an opposite private belief. We do not detail here the possible persuasion or negotiation dialogue in which s and b will enter if b does not believe that discounts are not possible and thus wants that s cancels his commitment.

$b \rightarrow s$: **Query if car type has high accident rate**

$\langle b, s, \text{Direct}, \langle s, b, \text{AssertIf}, \text{accHigh} \rangle \top \rangle$

Effect:

$C(\text{unset}, s, b, \langle s, b, \text{AssertIf}, \text{accHigh} \rangle \top)$

By asking whether the car type has a high accident rate, the buyer b has thus created a *unset* commitment for agent s to perform the *AssertIf* action.

$s \rightarrow b$: **Information about accident rate**

$\langle s, b, \text{Assert}, \neg \text{accHigh} \rangle$

Effect:

$C(\text{pending}, s, b, \neg \text{accHigh})$

Seller s has incurred a propositional commitment on $\neg \text{accHigh}$.

It is also interesting to remark that by having performed the requested action, s has fulfilled his *unset* commitment:

$C(\text{fulfilled}, s, b, \langle s, b, \text{AssertIf}, \text{accHigh} \rangle \top)$

$b \rightarrow s$: **Process of price bargaining**

$\langle b, s, \text{Direct}, \text{sell}(s, 10000\text{£}) \rangle$

Effect:

$C(\text{unset}, s, b, \text{sell}(b, 10000\text{£}))$

We consider that this act is the final step of the process of price bargaining: b requests s to lower the price of his car down to 10000 £. He has thus created an *unset* commitment. The next step of the dialogue is either s 's acceptance of the request or his refusal.

$s \rightarrow b$: **Acceptance of the price**

$\langle s, b, \text{Commit}, \text{sell}(s, 10000\text{£}) \rangle$

Effect:

$C(\text{pending}, s, b, \text{sell}(b, 10000\text{£}))$

s accepts to sell his car for the requested price and is thus committed to do this action. Indeed this commitment will be fulfilled once he has signed the selling contract with b . The contract makes the action public ($\text{Grd}_{\{s,b\}} \text{sell}(s, 10000\text{£})$) and thus fulfilled the *pending* commitment.

6 Conclusion

The aim of this paper was to use logical framework to describe social approaches of ACLs. In particular we formalized the notion of social commitment. We give description of both propositional commitments and commitments in action that are inspired by Walton & Krabbe's account of propositional commitments in persuasion dialogues and Colombetti *et al.*'s commitment life cycle. Contrarily to Colombetti *et al.*'s standpoint we argued that their commitment life-cycle should differ depending on the nature of the propositional content of the commitment. We have modified his life-cycle to take into account propositional commitments, which allowed to establish the link with the W&K account. We also discussed Colombetti *et al.*'s work on commitments in action, where we gave a formal characterization of each of its possible states. Moreover we linked formally both kinds of commitments.

Perspectives include among others the study of deontic aspects of commitments, to highlight links between commitments, obligations and rights, but also the institutional

dimension of these concepts.

References

- [1] M. E. BRATMAN. *Intention, Plans, and Practical Reason*. Harvard University Press, 1987.
- [2] C. CASTELFRANCHI. *Commitment: from individual intentions to groups and organizations*. In *Proc. of ICMAS*, ed. V. Lesser. MIT press, pp. 528–535, 1995.
- [3] P. R. COHEN, H. J. LEVESQUE. *Intention is Choice with Commitment*. *Artificial Intelligence Journal*, vol. 42(2–3), 1990.
- [4] M. COLOMBETTI. *A Commitment-Based Approach to Agent Speech Acts and Conversations*. In *Workshop on Agent Languages and Communication Policies*. Barcelona, pp. 21–29, 2000.
- [5] F. DIGNUM, R. M. VAN EIJK. *Agent communication and social concepts*. *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)*, vol. 14, pp. 119–120, 2007.
- [6] R. FAGIN, J. Y. HALPERN, Y. MOSES, M. Y. VARDI. *Reasoning about Knowledge*. MIT Press, 1995.
- [7] FIPA. *FIPA Communicative Act Library Specification*. <http://www.fipa.org/specs/fipa00037/>, Foundation for Intelligent Physical Agents, 2002.
- [8] R. A. FLORES, P. PASQUIER, B. CHAIB-DRAA. *Conversational Semantics with Social Commitments*. *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)*, vol. 14, pp. 165–186, 2007.
- [9] N. FORNARA, M. COLOMBETTI. *Operational specification of a commitment-based agent communication language*. In *Proc. AAMAS-2002*. ACM Press, Bologna, Italy, pp. 536–542, 2002.
- [10] N. FORNARA, F. VIGANÒ, M. COLOMBETTI. *Agent communication and artificial institutions*. *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)*, vol. 14, pp. 121–142, 2007.
- [11] B. GAUDOU, A. HERZIG, D. LONGIN. *Grounding and the expression of belief*. In *Proc. KR 2006*, eds. P. Doherty, J. Mylopoulos, C. A. Welty. AAAI Press, Windermere, UK, pp. 211–229, 2006.
- [12] B. GAUDOU, A. HERZIG, D. LONGIN. *A Logical Framework for Grounding-based Dialogue Analysis*. In *Proc. LCMAS 2005*, eds. W. van der Hoek, A. Lomuscio, E. de Vink, M. Wooldridge. Elsevier, UK, vol. 157 of *ENTCS*, pp. 117–137, 2006.
- [13] B. GAUDOU, A. HERZIG, D. LONGIN, M. NICKLES. *A New Semantics for the FIPA Agent Communication Language based on Social Attitudes*. In *Proc. ECAI 2006*. IOS Press, Trento, Italy, pp. 245–249, 2006.

- [14] R. HAKLI. *Group beliefs and the distinction between belief and acceptance*. *Cognitive Systems Research*, vol. 7, pp. 286–297, 2006.
- [15] C. L. HAMBLIN. *Fallacies*. Methuen London, 1970.
- [16] A. HERZIG, D. LONGIN. *C&L intention revisited*. In *Proc. KR2004*, eds. D. Dubois, C. Welty, M.-A. Williams. AAAI Press, Whistler, Canada, pp. 527–535, 2004.
- [17] J. HINTIKKA. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press, Ithaca, 1962.
- [18] A. U. MALLYA, M. P. SINGH. *An algebra for commitment protocols*. *Journal of Autonomous Agents and Multi-Agent Systems (JAAMAS)*, vol. 14, pp. 143–163, 2007.
- [19] A. S. RAO, M. P. GEORGEFF. *Modeling Rational Agents within a BDI-Architecture*. In *Proc. KR'91*, eds. J. A. Allen, R. Fikes, E. Sandewall. Morgan Kaufmann Publishers, pp. 473–484, 1991.
- [20] D. SADEK. *A Study in the Logic of Intention*. In *Proc. KR'92*, eds. B. Nebel, C. Rich, W. Swartout. Morgan Kaufmann Publishers, pp. 462–473, 1992.
- [21] J. R. SEARLE. *Speech acts: An essay in the philosophy of language*. Cambridge University Press, 1969.
- [22] M. P. SINGH. *Social and Psychological Commitments in Multiagent Systems*. Tech. Rep. TM-91-08, 1991.
- [23] M. P. SINGH. *Agent Communication Languages: Rethinking the Principles*. *Computer*, vol. 31(12), pp. 40–47, 1998.
- [24] M. P. SINGH. *A Social Semantics for Agent Communication Languages*. In *Issues in Agent Communication*, eds. F. Dignum, M. Greaves. Springer-Verlag, no. 1916 in LNAI, pp. 31–45, 2000.
- [25] D. VANDERVEKEN. *Principles of language use*, vol. 1 of *Meaning and Speech Acts*. Cambridge University Press, 1990.
- [26] D. N. WALTON, E. C. W. KRABBE. *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*. State University of New-York Press, NY, 1995.

ABSTRACT

Nowadays most of the Agent Communication Languages are no longer defined in terms of the agents' mental attitudes, but in terms of social commitments, which avoids strong hypotheses on the agents' internal states and thus allows verification of the semantics. While social approaches have attracted a lot of attention, two drawbacks remain. First, the notion of commitment does not have a clear and unambiguous characterization. Second, commitments are completely unrelated to the agents' reasoning. The aim of this paper is to propose a solution for these two problems based on a combination of a BDI (Belief, Desire, Intention) logic with a logic of what is publicly grounded between agents.

KEYWORDS

Agent communication languages, modal logic, grounding, commitments, BDI logic, speech act theory