Grounding and the expression of belief

Benoit Gaudou and **Andreas Herzig** and **Dominique Longin**
Université Paul Sabatier – IRIT/LILaC, Toulouse, FRANCE
{gaudou,herzig,longin@irit.fr}

Abstract

In this paper we investigate the logic of speech acts and groundedness. A piece of information is grounded for a group of agents if it is publicly expressed and established by all agents of the group. Our concept of groundedness is founded on the expression of the sincerity condition of speech act theory.

We formalize groundedness within an extended BDI (Belief, Desire, Intention) logic where belief is viewed as a kind of group belief. We show that our logic permits to reconcile the mentalist approaches on the one hand, and the structural and social approaches on the other, which are the two rival research programs in the formalization of agent interaction. Although groundedness is thus linked to the standard mental attitude of belief, it is immune to the critiques that have been put forward against the mentalist approaches, viz. that they require too strong hypotheses about the agents' mental states such as sincerity and cooperation: just as the structural approaches, groundedness only bears on the public aspect of communication.

In our extended BDI logic we study communication between heterogeneous agents. We characterize inform and request speech acts in terms of preconditions and effects. We demonstrate the power of our solution by means of two examples. First, we revisit the well-known FIPA Contract Net Protocol. As a second example, we show how Walton & Krabbe's commitments can be redefined in term of groundedness.

Keywords: modal logic, grounding, dialogue, speech acts, commitment, BDI logic.

Introduction

Two main approaches have been followed to formalize and produce dialogues. The mentalist approach (often based on a Belief-Desire-Intention (BDI) logic) (for example (Cohen & Levesque 1990b; Rao & Georgeff 1991; 1992; Sadek 1992)) considers that a dialogue is function of the agents' mental states. It has great predictive power but uses very strong hypotheses on the agents' internal architectures and on their mental states (like sincerity, cooperation,...). It is often criticized (cf. e.g. (Singh 2000; Fornara & Colombetti 2002)) that these hypotheses do not apply to open systems with heterogenous agents.

To get round this problem, the conventional approaches take into account only what is public in the dialogue (for example (Conte & Castelfranchi 1995; Walton & Krabbe 1995; Singh 2000; Fornara & Colombetti 2002; Verdicchio & Colombetti 2003)) and describe dialogue through the notion of commitment. Although many commitment-based frameworks have been defined, the logic of commitment has not been entirely clarified yet. Moreover a public 'layer' in terms of commitments does not always allow to avoid reference to some private mental states, in particular when we want to formalize the deliberative capacities of agents. And the link between these two layers is not characterized in conventional approaches.

We propose to bridge the gap between these two approaches by extending a BDI-like logical framework with an operator formalizing what is public in the dialogue. This corresponds to the concept of groundedness.

We view grounded information as information that is *publicly expressed and accepted as being true by all the agents participating in a conversation*. A piece of information might be grounded even when some agents privately disagree, as long as they do not publicly manifest their disagreement.

Our notion stems from speech act theory, where Searle's *expression of an Intentional state* (Searle 1983) concerns a psychological state related to the state of the world. Even if an utterance was unsincere an Intentional state has been *expressed*, and that state corresponds to a particular belief of the speaker.

Vanderveken (Vanderveken 1990; 1991) has captured the subtle difference between *expressing* an Intentional state and *really being in* such a state by distinguishing *success conditions* from *non-defective performance conditions*, thus refining Searle's felicity conditions (Searle 1969; 1979; Searle & Vanderveken 1985). According to Vanderveken, when we assert p we *express* that we believe p (success condition), while the speaker's belief that p is a condition of non-defective performance.

The notion of groundedness is also behind Moore's paradox, according to which one cannot successfully assert "p is true and I do not believe p". The paradox follows from the fact that: on the one hand, the assertion entails expression of the sincerity condition about p (the speaker believes p); on the other hand, the assertion expresses the speaker believes he ignores that p. If we accept introspection then this expresses that the speaker does not believe p, and the assertion is contradictory (if we accept that beliefs are consistent).

Although groundedness is related to mental states because it corresponds to the expression of Intentional states, groundedness in a group is not an Intentional state: it is neither a belief, nor a goal, nor an intention.

Groundedness is an objective notion: it refers to what can be observed, and only to that. It is different from other objective notions such as that of social commitment of (Singh 1998; 2000; Fornara & Colombetti 2002; Verdicchio & Colombetti 2003). To see this consider the speech act where agent i asks agent j if j can pass the salt to him. Thereafter it is established (if we assume that the speech act is well and completely understood) that i wants to know whether j is able to pass him the salt (literal meaning), or that i wants j to pass him the salt (indirect meaning). In a commitment-based approach this typically leads to a conditional commitment (or precommitment) of j to pass the salt, which becomes an unconditional commitment upon a positive reaction.

In our approach we do not try to determine whether j *must* do such or such action or not: we just establish the facts, without any hypothesis on the agents' beliefs, goals, intentions, . . . or commitments.

In a previous paper (Gaudou, Herzig, & Longin 2006), we presented a modal logic of belief and choice augmented by the modal operator G to express the notion of grounding. But this operator was a bit too restricted: GA expresses that A is publicly grounded, where "publicly" means for all agents. Thus in a given group of agents, we cannot distinguish a private dialogue between two agents from a public debate. In the former a piece of information could be grounded between only two agents and stay secret for the other agents of the group.

We will formalize the extended notion of grounding and introduce it into a logic of belief, intention and action (Section *Logical Framework*), where belief is viewed as a kind of group belief. Afterwards we will show some applications of our new notion (Section *Applications*). We believe that such a notion is interesting because it fits the public character of speech act performance (Traum 1994) (Subsection *ACLs and grounding*). We can apply it to formalize dialogue games *à la* Walton & Krabbe (Subsection *Social commitment and dialogue games*) and dialogue protocols (Subsection *FIPA Contract Net Protocol*). As far as we are aware the logical investigation of such a notion has neither been undertaken in the social approaches nor in the conventional approaches. A very close notion has been proposed very recently in (Nickles, Fischer, & Weiss 2006) where the idea is to formalize the notion of manifested opinion in the sense of ostensible belief and ostensible intentions. We show in the *Discussion* Section that our logical framework captures these notions.

Logical Framework

In this section, we present a light version of the logic of belief, choice and action we developed in (Herzig & Longin 2004) which builds on the works of Cohen & Levesque (Cohen & Levesque 1990a) and Sadek (Sadek 1992), and augments it by a modal operator expressing groundedness in a group. We show that groundedness for the single-agent group $\{i\}$ corresponds to belief of i. Thus a particular individual belief operator is superfluous. We neither develop here temporal aspects nor dynamics between action and mental attitudes.

Semantics

Let $AGT = \{i, j, \ldots\}$ be a finite set of agents. A *group of agents* (or a *group* for short) is a nonempty subset of AGT. We use $I, J, K, \ldots$ to denote groups. When $I' \subseteq I$ we say that I' is a subgroup of I. Let $ATM = \{p, q, \ldots\}$ be the set of atomic formulas. Complex formulas are denoted by $A, B, \ldots$.

A model includes a set of possible worlds W and a mapping $V : W \longrightarrow (ATM \longrightarrow \{0,1\})$ associating a valuation V_w to every $w \in W$. Models moreover contain accessibility relations that will be detailed in the sequel.

Grounding. $G_I A$ reads "it is publicly grounded for group I that A is true" (or for short: "A is grounded for I"). When I is a singleton, $G_{\{i\}} A$ reads "A is grounded for (agent) i" and $G_{\{i\}}$ is identified with the standard belief operator Bel_i à la Hintikka. We write $G_i A$ for $G_{\{i\}} A$.

To each world w and each non-empty $I \subseteq AGT$, we associate the set of possible worlds $\mathcal{G}_I(w)$ that are consistent with all propositions grounded in world w for the group I. $\mathcal{G}_I(w)$ contains those worlds where all grounded propositions hold.

The truth condition for G_I stipulates that A is grounded in w, noted $w \Vdash G_I A$, iff A holds in every world that is consistent with the set of grounded propositions:

$$w \Vdash G_I A \text{ iff } w' \Vdash A \text{ for every } w' \in \mathcal{G}_I(w).$$

Every mapping $\mathcal{G}_I$ can be viewed as an accessibility relation, and we assume that:

❶ $\mathcal{G}_I$ is serial.

Thus, groundedness is rational: if a proposition holds in every world that is consistent with the set of grounded propositions, then at least one such a world exists.

Furthermore we postulate the following constraints on accessibility relations, for groups I and I' such that $I' \subseteq I$:

❷ if $u\mathcal{G}_{I'}v$ and $v\mathcal{G}_I w$ then $u\mathcal{G}_I w$;

❸ if $u\mathcal{G}_{I'}v$ and $u\mathcal{G}_I w$ then $v\mathcal{G}_I w$;

❹ if $u\mathcal{G}_I v$ and $v\mathcal{G}_{I'} w_1$ then there is w_2 such that $u\mathcal{G}_I w_2$ and

- $V(w_1) = V(w_2)$,
- $\mathcal{G}_K(w_1) = \mathcal{G}_K(w_2)$ for all K such that $K \cap I = \emptyset$,
- $\mathcal{C}_k(w_1) = \mathcal{C}_k(w_2)$ for all k such that $k \notin I$, where $\mathcal{C}$ is the accessibility relation for choice to be defined below;

❺ $\mathcal{G}_I \subseteq \bigcup_{i \in I} \mathcal{G}_I \circ \mathcal{G}_i$.

Constraint ❷ stipulates that agents of a subset I' of the set I are aware of what is grounded in the group I: whenever w is a world for which it is grounded for I' that all I-grounded propositions hold in w, then all I-grounded propositions indeed hold in w. This is a kind of *attention* property: each subgroup taking part in a conversation is aware of what is grounded in the group.

Similarly ❸ expresses that subgroups are aware of what is ungrounded in the group, too.

❷ and ❸ together make that if $u\mathcal{G}_{I'}v$ then $\mathcal{G}_I(u) = \mathcal{G}_I(v)$, *i.e.* if $u\mathcal{G}_{I'}v$ then what is grounded for I at u is the same as what is grounded for I at v. From ❷ and ❸ it also follows that $\mathcal{G}_I$ is transitive and euclidian.

❹ says that if an information "about something outside group I" (see the definition in the following subsection) is grounded for I then it is grounded for I this information is grounded for every subgroup of I.

❺ says that if it is grounded for a set I that a proposition is established for every agent then it is grounded for I, too.

Choice. Among all the worlds in $\mathcal{G}_i(w)$ that are possible for agent i, there are some that i prefers. Semantically, these worlds are identified by yet another mapping $\mathcal{C} : AGT \longrightarrow (W \longrightarrow 2^W)$ associating an accessibility relation $\mathcal{C}_i$ to each $i \in AGT$. $\mathcal{C}_i(w)$ denote the set of worlds the agent i prefers.

Ch_iA reads "agent i chooses that A". Choice can be viewed as a preference operator and we sometimes also say that "i *prefers* that A". Note that we only consider individual choices, group choices being beyond the scope of the present article.

The truth condition for Ch_i stipulates that $w \Vdash Ch_iA$ if A holds in all chosen worlds:

$$w \Vdash Ch_iA \text{ iff } w' \Vdash A \text{ for every } w' \in \mathcal{C}_i(w).$$

We assume that:

❻ $\mathcal{C}_i$ is serial, transitive, and euclidian.

(See (Herzig & Longin 2004) for more details about the logic of choice, and the definition of intention from choice.)

Choice and grounding. As said above, an agent only chooses worlds he considers possible:

❼ $\mathcal{C}_i(w) \subseteq \mathcal{G}_i(w)$.

Hence what is grounded for an agent must be chosen by him, and choice is a mental attitude that is logically weaker than groundedness.

We moreover require that worlds chosen by i are also chosen from i's 'grounded worlds', and *vice versa*.

❽ if $w\mathcal{G}_iw'$ then $\mathcal{C}_i(w) = \mathcal{C}_i(w')$.

This constraint means that agent i is aware of his choices.

Action. Let $ACT = \{\alpha, \beta \ldots\}$ be the set of actions. Sometimes we write $(i{:}\alpha)$ to denote that i is the author of (*i.e. performs*) the action α.

The model contains a mapping $R : ACT \longrightarrow (W \longrightarrow 2^W)$ associating an accessibility relation R_α to every $\alpha \in ACT$. $R_\alpha(w)$ is the set of worlds accessible from w through the execution of α. Just as Cohen and Levesque we suppose here that there is at most one possible execution of α. Hence R_α can also be viewed as a partial function on W.

The formula $After_\alpha A$ reads: "A holds after every execution of α". As there is at most one possible execution of α, the dual operator $Happens_\alpha A \stackrel{def}{=} \neg After_\alpha \neg A$ reads: "α is happening and A is true just afterwards". Hence $After_\alpha \bot$ expresses that α does not happen, and $Happens_\alpha \top$ that α happens. We often write $Happens(\alpha)$ for $Happens_\alpha \top$. The truth condition is:

$$w \Vdash After_\alpha A \text{ iff } w' \Vdash A \text{ for every } w' \in R_\alpha(w)$$

The formula $Before_\alpha A$ reads: "A holds before every execution of α". The dual $Done_\alpha A \stackrel{def}{=} \neg Before_\alpha \neg A$ expresses that the action α has been performed before which A held. Hence $Done_\alpha \top$ reads: "α has just happened".

The accessibility relation for $Before_\alpha$ is the converse of the above relation R_α. The truth condition is:

$$w \Vdash Before_\alpha A \text{ iff } w' \Vdash A \text{ for every } w' \in R_\alpha^{-1}(w).$$

As said above, we do not detail here the relationship between action and mental attitudes and refer the reader to (Herzig & Longin 2004).

Action and grounding. We consider in this paper that actions are public for attending agents, in the sense that are completely and soundly perceived by them.

For example, when agent i performs an assertive speech act only towards agent j then j will perceive the assertion. If no other agent perceives this action then the attentive group is limited to $\{i, j\}$, and the action is public for exactly this group.

Let α be an action performed by a agent i in front of attentive group K (of which i is a member). The property of public actions (for group K) corresponds to the constraints:

❾ $R_\alpha^{-1}(w) = \emptyset$ if and only if $(\mathcal{G}_K \circ R_\alpha^{-1})(w) = \emptyset$

Axiomatics

Grounding. The logic of the grounding operator is a normal modal logic of type KD:

$$G_IA \to \neg G_I \neg A \qquad (\mathrm{D}_{G_I})$$

(D_{G_I}) expresses that grounded information in a group are consistent: it cannot be the case that both A and $\neg A$ are simultaneously grounded.

In accordance with the preceding semantic conditions the following logical axioms respectively correspond to the constraints ❷ and ❸. Thus, for each $I' \subseteq I$:

$$G_IA \to G_{I'}G_IA \qquad (\mathrm{SR}_+)$$

$$\neg G_IA \to G_{I'}\neg G_IA \qquad (\mathrm{SR}_-)$$

The axioms of strong rationality (SR_+) and (SR_-) express that if a proposition A is grounded (resp. ungrounded) for group I then it is grounded for each subgroup that A is grounded (resp. ungrounded) for I. This is due to the public character of the grounding operator.

The next axiom must be restricted to particular formulas, viz. objective formulas for a group, that we define as follows.

Definition. The set of formulas that are objective for a group I is defined inductively to be the smallest set such that:

- every atomic formula p is objective for I;
- $G_K A$ is objective for I if $K \cap I = \emptyset$, for every formula A;
- $Ch_j A$ is objective for I if $j \notin I$, for every formula A;
- if A and A' are objective for I then $\neg A$, $A \wedge A'$ are objective for I.

With respect to the semantic constraint ❹, our third axiom of weak rationality stipulates that if I' is a subgroup of I and A is objective for I then:

$$G_I A \rightarrow G_I G_{I'} A \tag{WR}$$

(WR) expresses that if A is objective for group I and grounded for I then it is necessarily grounded for I that for each subgroup I' the formula is grounded.

Note that this does not imply that for every subgroup A is actually grounded, *i.e.* (WR) does not entail $G_I A \rightarrow G_{I'} A$. In particular, the fact that A is grounded for group I does not imply that the members of I believe that A. Note that (WR) concerns only formulas A that are objective for I. Indeed, if we applied (WR) to some mental states of an agent of the group, we would restrict the agents' autonomy.

For example, when an agent i asserts to another agent j that A in presence of group I, he publicly expresses that he believes A (Searle 1969; Vanderveken 1990) and thus he socially commits himself on the fact that he believes A, as we will develop in Subsection *ACLs and grounding*. Thus his belief that A is immediately and without discussion grounded for the group.
Now if agent i asserts that $G_j A$ in presence of group I, then the formula $G_I G_i G_j A$ holds afterwards, and if (WR) applied unrestrictedly then j could not express later that he ignores whether A, or believes $\neg A$. If he made this last speech act, the formulas $G_I G_j \neg A$ and, thanks to (WR), $G_I G_i G_j \neg A$ would hold, which is inconsistent with the above formula $G_I G_i G_j A$ (Gaudou, Herzig, & Longin 2006).

This restriction highlights that a formula A can be grounded in two different manners: either A is objective for group I and it must be discussed by all the agents of I, or A is not and it is grounded directly by being expressed.

And finally, corresponding to the semantic constraint ❺, we have the last axiom of common grounding:

$$(\bigwedge_{i \in I} G_I G_i A) \rightarrow G_I A \tag{CG}$$

It expresses that if a proposition is established for every agent in AGT, then it is grounded. Together, (WR) and (CG) stipulate that for formulas A that are objective for I we have $(\bigwedge_{i \in I} G_I G_i A) \leftrightarrow G_I A$.

From axioms (SR$_+$) and (SR$_-$), we can prove that we have the modal axioms (4) and (5) for G_I operators as theorems of our logic:

$$G_I A \rightarrow G_I G_I A \tag{4_{G_I}}$$

$$\neg G_I A \rightarrow G_I \neg G_I A \tag{5_{G_I}}$$

Thus operator G_I is in a normal modal logic of type KD45. Hence for individual groundedness we obtain the standard logic of belief KD45.

We can moreover show that if $I' \subseteq I$ then:

$$G_I A \leftrightarrow G_{I'} G_I A \tag{1}$$

$$\neg G_I A \leftrightarrow G_{I'} \neg G_I A \tag{2}$$

These theorems express that subgroups of a group are aware of what is grounded in the group. The formula $(\bigwedge_{I' \subseteq I} G_I G_{I'} A) \rightarrow G_I A$ is provable from our axiom (CG).

Moreover we can prove that:

$$G_I A \leftrightarrow G_I G_{I'} G_I A \tag{3}$$

$$\neg G_I A \leftrightarrow G_I G_{I'} \neg G_I A \tag{4}$$

These theorems say that if A is (not) grounded for a group, it is grounded for this group that it is grounded for every subgroup of this group that A is (not) grounded for the group.

Even if I' is a subgroup of I we do not necessarily have $G_I A \rightarrow G_{I'} A$. Such a principle would be too strong because it would restrict the autonomy of subgroups I' of I: a proposition can be grounded for a group I while there is a dissident subgroup I' of I, *i.e.* a group where the contrary is grounded: $G_I A \wedge \neg G_{I'} A$ is consistent in our logic even if $I \cap I' \neq \emptyset$.

Choice and intention. With respect to the semantic constraint ❻, we have the axioms (D$_{Ch_i}$), (4$_{Ch_i}$) and (5$_{Ch_i}$):

$$Ch_i A \rightarrow \neg Ch_i \neg A \tag{D_{Ch_i}}$$

$$Ch_i A \rightarrow Ch_i Ch_i A \tag{4_{Ch_i}}$$

$$\neg Ch_i A \rightarrow Ch_i \neg Ch_i A \tag{5_{Ch_i}}$$

We define intention in a way similar to Cohen and Levesque as:

$$Int_i A \stackrel{def}{=} Ch_i \Diamond G_i A \wedge \neg G_i A \wedge \neg G_i \Diamond G_i A \tag{Def_{Int_i}}$$

where $\Diamond$ is an operator of linear temporal logic LTL. Hence i intends that A if and only if in i's preferred worlds i will believe A at some world in the future, i does not believe A holds now (i.e., A is an achievement goal), and it is not the case that i believes he will come to believe A anyway (A is not self-realizing). For more details on this intention operator, see (Herzig & Longin 2004). We often write $Int_i \alpha$ for $Int_i Done_\alpha \top$.

Choice and Grounding. Due to the semantic constraint ❼ we have the following axiom:

$$G_i A \rightarrow Ch_i A \tag{5}$$

which means that every formula grounded for agent i must necessarily be chosen by this agent.

Our semantics also validates the principles:

$$Ch_i A \leftrightarrow G_i Ch_i A \tag{6}$$

$$\neg Ch_i A \leftrightarrow G_i \neg Ch_i A \tag{7}$$

that correspond with constraint ❽. This expresses that agents are aware of their choices.

Action. As the relation $R_\alpha^{-1}(w)$ is the converse of R_α, for the completeness of the logic we also have the conversion axioms:

$$A \to After_\alpha Done_\alpha A \qquad (\mathrm{I}_{After_\alpha,\, Done_\alpha})$$
$$A \to Before_\alpha Happens_\alpha A \qquad (\mathrm{I}_{Before_\alpha,\, Happens_\alpha})$$

Action and grounding. As we have said above we only consider public actions and α be an action performed by a agent i in front of attentive group K (of which i is member). Thus we have following axioms of public actions corresponding to the semantic constraint ❾, for each group K observing an action α:

$$G_K Done_\alpha \top \leftrightarrow Done_\alpha \top \qquad (\mathrm{PA}_{K,\alpha})$$

Example

To highlight our proposal for the semantics of grounding consider the following example where there are three agents $AGT = \{0, 1, 2\}$:

1. Let agent 0 (privately) believe that 2 is smart, formally written $G_0\ smart_2$.
2. Now suppose that in private conversation agent 0 tells 1 that 2 is not smart. The illocutionary effect is $G_{\{0,1\}} G_0 \neg\ smart_2$.
3. If 1 publicly adopts $\neg\ smart_2$ (e.g. by confirming publicly that $\neg smart_2$) we moreover obtain $G_{\{0,1\}} \neg smart_2$.
4. Then agent 2 joins in the conversation, and later on 0 informs 1 and 2 that 2 is smart: the illocutionary effect is $G_{\{0,1,2\}} G_0\ smart_2$.
5. Then if both 1 and 2 publicly adopt $smart_2$ we moreover obtain $G_{\{0,1,2\}} smart_2$.

This illustrates that even for nested groups $J_0 = \{0\} \subset J_1 = \{0, 1\} \subset J_2 = \{0, 1, 2\}$ we might have states of public groundedness for the different groups which are about propositions that are mutually inconsistent, viz. here:

$$G_{J_0}\ smart_2$$
$$G_{J_1} \neg smart_2$$
$$G_{J_2}\ smart_2$$

Applications

A lot of Agent Communication Languages (ACLs for short) are based on mental states: speech acts are described with beliefs, goals, and intentions. The most popular ACLs are FIPA-ACL (FIPA 2002a) and KQML (Finin, Labrou, & Mayfield 1997).

In Subsection *ACLs and grounding*, we characterize in our framework the main speech acts of FIPA: inform and request acts. In the two subsections *Social commitment and dialogue games* and *FIPA Contract Net Protocol* we study briefly two distinct ways to produce dialogues: Walton & Krabbe dialogue games (Walton & Krabbe 1995) and the FIPA Contract Net Protocol (CNP). We formalize them with the notion of grounding.

ACLs and grounding

In our framework, speech acts are just particular actions: they are 5-tuples of the form $\langle i, J, K, FORCE, A\rangle$ where $i \in AGT$ is the *author* of the speech act (*i.e.* the speaker), $K \subseteq AGT$ the *group* of agents attentive to the conversation, $J \subseteq K \setminus \{i\}$ the set of its *addressees*, $FORCE$ its illocutionary force, and A a formula denoting its propositional content. We must have $i \in K$ and $J \neq \emptyset$.

The distinction between the addressees J of a speech act and the group of agents K taking part in the conversation improves the usual FIPA-like characterization of speech acts: from the speech act theory standpoint, when a speaker talks to a subgroup J of K then the *success conditions* (Searle 1969; Vanderveken 1990) apply only to J (but are evaluated from the point of view of the entire group). Nevertheless, effects also obtain for the entire group K. This motivates that the addressees and the group must be distinguished, and must both be a parameter of the speech act.

The inform act. One of the simplest speech act is $\langle i, J, K, \mathsf{Inform}, A\rangle$ which means: "agent i informs group J among the attentive group K that A is true". In FIPA-ACL, agent i can perform such an act (restricted to only one addressee j) only if i believes A is true and if i does not believe j has an opinion about A. This is expressed in FIPA-ACL by $Bel_i A \wedge \neg Bel_i(BelIf_j A \vee Uif_j A)$, where $Uif_j A$ reads "either A or $\neg A$ is probable for j".

As the FIPA preconditions are private mental attitudes, we do not keep them here. The preconditions of our actions are of two types: public relevance and public rationality. The *relevance precondition* of $\langle i, J, K, \mathsf{Inform}, A\rangle$ is that i has not already expressed he believes A, and the same for J (that is: $\neg G_K G_i A \wedge \neg G_K G_J A$), and that J has not expressed that he does not believe A (formally: $\neg G_K \neg G_J A$ — otherwise the speech act would not be an inform act but a convince act). The *rationality precondition* corresponds to the fact that an agent must be publicly consistent, and means the agent has not expressed he does not believe A (formally: $\neg G_K \neg G_i A$). Hence we define:

$$\mathbf{Prec}(\langle i, J, K, \mathsf{Inform}, A\rangle) \stackrel{def}{=}$$
$$\neg G_K G_i A \wedge \neg G_K G_J A \wedge$$
$$\neg G_K \neg G_J A \wedge \neg G_K \neg G_i A$$

In FIPA-ACL, the *rational effect* (RE) roughly corresponds to the expected perlocutionary effect of the act. The RE is not directly added to the mental state of the addressee, but if this effect can be derived from the mental state (after the act performance) then the author of the act has achieved his aim. In fact, sincerity and credulity hypotheses always entail the rational effect. Thus the FIPA rational effect of $\langle i, j, \mathsf{inform}, A\rangle$ is that the addressee believes what is asserted, i.e. $Bel_j A$.

But in fact, we can never guarantee such perlocutionary effects because we cannot control other agents' mental states.[1]

[1] From this point of view, Searle (Searle 1969) shows that what we could name "perlocutionary act" cannot be a speech act (in the

However speech act theory says we cannot perform an action without **necessarily expressing** sincerity and preparatory conditions (Searle 1983). The preparatory condition roughly corresponds to the relevance precondition of FIPA-ACL. (Note that we adopt here a public point of view and do not impose the agent is sincere and has checked the preparatory conditions. Usual BDI logics cannot capture this aspect of communication.)

Thus, **expression** of such conditions is an **effect** of the act. When i informs J that A, he expresses that he believes A and that he intends J believes A (formally: $G_K G_i A \wedge G_K Int_i G_J A$) which is the expression of the sincerity condition.

One might think that it is too strong that only by performing a speech act an agent can ground a formula for the whole group. Moreover $G_K G_i A$ can hold while neither agent i believes A nor at least one agent of the group K believes that i believes A. This situation could appear to be hypocrite.

But in fact by asserting that A, agent i commits himself in front of the whole group to his belief that A. Thus formula $G_K G_i A$ characterizes the acceptance by the group of the commitment. While members of the group can think privately that i has lied, they cannot deny that he has incurred a commitment. The acceptance is consequently implicit and immediate and does not require any discussion. An agent incurs a social commitment by performing an informing speech act in front of an attentive group of agents.

The speaker also expresses the preparatory condition: he believes A is not grounded for J yet (formally: $G_K G_i \neg G_J A$).

Putting all these effects together we get:

$$\textbf{Effect}(\langle i, J, K, \mathsf{Inform}, A\rangle) \stackrel{def}{=} G_K G_i A \wedge G_K Int_i G_J A \wedge G_K G_i \neg G_J A$$

What about inform-actions whose propositional content commits the hearers to some belief? We have the following theorem:

Theorem. The action $\langle i, J, K, \mathsf{Inform}, G_{K'} A\rangle$ is inexecutable, for each K' such that $K \subseteq K' \subseteq AGT$.

Proof. We will prove that the preconditions of this act are inconsistent. Let K' be an supergroup of K, *i.e.* $K \subseteq K' \subseteq AGT$. In particular we have:

$$\models \textbf{Prec}(\langle i, J, K, \mathsf{Inform}, G_{K'} A\rangle) \rightarrow \neg G_K \neg G_i G_{K'} A \wedge \neg G_K G_i G_{K'} A$$

From theorems (1) and (2), we can prove the equivalences, for $i \in K$ and $K \subseteq K'$:

$$\models \neg G_K \neg G_i G_{K'} A \leftrightarrow \neg G_K \neg G_{K'} A$$
$$\models \neg G_K \neg G_{K'} A \leftrightarrow G_{K'} A$$

Similarly from theorem (1), we can show that:

$$\models \neg G_K G_i G_{K'} A \leftrightarrow \neg G_K G_{K'} A$$
$$\models \neg G_K G_{K'} A \leftrightarrow \neg G_{K'} A$$

The precondition of the action $\langle i, J, K, \mathsf{Inform}, G_{K'} A\rangle$ is inconsistent and thus this kind of acts is inexecutable.

□

If an agent could perform the act $\langle i, J, K, \mathsf{Inform}, G_{K'} A\rangle$, one of its effects would be $G_K G_i G_{K'} A$, which is equivalent to $G_{K'} A$. This theorem highlights an important property of our logic: if an agent could perform such a speech act he could ground a formula A for the whole group without possible discussion.

Moreover this theorem sheds a new light on the seemingly too powerful theorem (1) ($G_I A \leftrightarrow G_{I'} G_I A$) and its counterpart (2). In particular, the implication $G_{I'} G_I A \rightarrow G_I A$ says that when it is grounded for a subgroup I' of I that $G_I A$ then *de facto* it is grounded for I that A, and seems to give too much power to a subgroup. But the above theorem expresses that no agent of I' can express a formula in the scope of operator G_I, *i.e.* he cannot establish by discussion and consensus formulas such as $G_{I'} G_I A$. Thus such a formula can only hold if $G_I A$ holds, which is a quite intuitive property.

The request act. Another simple FIPA speech act is the request act. Let us assume i is the author of a request. $\langle i, J, K, \mathsf{Request}, A\rangle$ means "agent i requests a subset J of group of agents K to perform some action having A as effect, K attending". The *relevance precondition* is: it is not grounded for K that:

1. i intends that A,
2. J intends that A, and
3. J does not intend A (otherwise the act would be close to a persuasion speech act).

The *rationality precondition* is that it is not grounded for K that i does not intend A.

The *effects* are:

1. i intends that A (expression of the sincerity condition), and
2. i expresses he believes that J does not intend that A be true (expression of the preparatory condition).

We did not define what group intention is. Here, we only consider individual actions, whose authors are individual agents which do not need other agents (versus group actions, group intention, teamwork…as e.g. studied in (Cohen & Levesque 1994)). Thus, to say that "group J intends A" means "there is at least one agent which intends A to be true" (that is: $\bigvee_{j \in J} Int_j A$). Futhermore, due to our definition of intention, here the negation of a choice is more appropriate than a negation of an intention. (See (Herzig & Longin 2004).) Thus:

$$\textbf{Prec}(\langle i, J, K, \mathsf{Request}, A\rangle) \stackrel{def}{=} \neg G_K Int_i A \wedge (\neg G_K \bigvee_{j \in J} Int_j A) \wedge (\neg G_K \bigwedge_{j \in J} \neg Ch_j A) \wedge \neg G_K \neg Ch_i A$$

$$\textbf{Effect}(\langle i, J, K, \mathsf{Request}, A\rangle) \stackrel{def}{=} G_K Int_i A \wedge (G_K G_i \bigwedge_{j \in J} \neg Ch_j A)$$

speech act theory sense). It was just a mistake of Austin (Austin 1962).

In this way, the semantics of every FIPA-ACL speech act can be redefined. This is subject of ongoing work.

Social commitment and dialogue games

Our notion of grounding is close to the notion of commitment in dialogue developed by Walton and Krabbe (Walton & Krabbe 1995). Like them we contract commitments for example by performing speech acts (like asserting or conceding). In previous work we suggested a formalization of the persuasion dialogue type PPD_0.

Our formalism allows to describe the two kinds of commitments used in PPD_0 dialogues: *assertion* (or strong commitment) incurred after speech acts like assert, argue... and *concession* (or weak concession) incurred after speech acts like concede. . . For a group of agents I:

$$SC_i^I A \stackrel{def}{=} G_I G_i A \qquad (\mathrm{Def}_{SC_i^I})$$

$$WC_i^I A \stackrel{def}{=} G_I \neg G_i \neg A \qquad (\mathrm{Def}_{WC_i^I})$$

Second, we characterized speech acts in terms of preconditions and effects, and have shown how this constrains the agents' options for the choice of actions (as well as their order), and thus drives them to follow the dialogue game.

For example, we have shown that after an assertion, under certain conditions the hearer will perform either a challenge act or a concede act (there are only two agents s et h):

Theorem (Gaudou, Herzig, & Longin 2006).

$$\begin{aligned} &LAWS \models After_{\langle s,\{h\},\{s,h\},\mathsf{Assert},p\rangle}(\\ &\quad (\neg WC_h^{\{s,h\}} p \wedge \neg Done_{\langle h,\{s\},\{s,h\},\mathsf{Challenge},p\rangle}\top) \rightarrow \\ &\quad G_{\{s,h\}}(Happens(\langle h,\{s\},\{s,h\},\mathsf{Challenge},p\rangle) \\ &\quad \vee Happens(\langle h,\{s\},\{s,h\},\mathsf{Concede},p\rangle))) \end{aligned}$$

FIPA Contract Net Protocol

Similarly to Walton & Krabbe dialogue games we can formalize the well-known FIPA Contract Net Protocol (CNP) (FIPA 2002b). The CNP uses acts defined in the FIPA-ACL library: this library gives for some speech acts a semantics *i.e.* expresses their Feasibility Preconditions and Rational Effects in a BDI-logic.

We have exhibited in (Gaudou, Herzig, & Longin 2005) pre- and postconditions of the CNP acts in terms of grounding. We have shown that the are sound and complete w.r.t. the CNP. For example, in the case of the *refuse* speech act $\langle i,\{j\},\{i,j\},\mathsf{refuse},(i{:}\alpha)\rangle$ (which means: "participant i refuses to the manager j to perform the action α")[2], the precondition is:

$$\begin{aligned} &\mathbf{Prec}(\langle i,\{j\},\{i,j\},\mathsf{refuse},(i{:}\alpha)\rangle) \stackrel{def}{=} \\ &\quad G_{\{i,j\}} Int_j(i{:}\alpha) \wedge \\ &\quad \neg G_{\{i,j\}} Int_i(i{:}\alpha) \wedge \neg G_{\{i,j\}} G_j Int_i(i{:}\alpha) \wedge \\ &\quad \neg G_{\{i,j\}} \neg G_j Int_i(i{:}\alpha) \wedge \neg G_{\{i,j\}} \neg Int_i(i{:}\alpha) \end{aligned}$$

[2] In the CNP, the refuse act is performed by one participant towards the manager, and the other participants are not aware of it. Thus, agent group K in this case is just the participant and the manager, and the addressee group J is just the manager.

and the effect is:

$$\begin{aligned} &\mathbf{Effect}(\langle i,\{j\},\{i,j\},\mathsf{refuse},(i{:}\alpha)\rangle) \stackrel{def}{=} \\ &\quad G_{\{i,j\}} Int_i(i{:}\alpha) \wedge G_{\{i,j\}} Int_i G_j Int_i(i{:}\alpha) \wedge \\ &\quad G_{\{i,j\}} G_i \neg G_j Int_i(i{:}\alpha) \end{aligned}$$

Discussion

Link with common belief.

The operator G_I expresses what may be called manifested, public common belief. Such common belief comes from public actions, whose correct perception by every member of the group is common knowledge. These hypotheses make that the properties of the modal operator G_I are much stronger than those of the standard notion of common belief.

In turn, the link between manifested common belief $G_J A$ and private individual belief $G_i A$ is weaker than in the case of standard common belief because the axiom $G_J A \rightarrow G_i A$ is not valid. The latter makes that the induction axiom $\bigwedge_{i \in J}(G_i A \wedge G_i(A \rightarrow G_J A)) \rightarrow G_J A$ is not valid. As we have argued in subsection *Axiomatics*, validity of such principles would violate the agents' autonomy. Nevertheless axioms $G_J A \rightarrow G_i G_J A$, for $i \in J$, and $(\bigwedge_{i \in J} G_i(A \wedge G_J A)) \rightarrow G_J A$ (which together with $G_J A \rightarrow G_i A$ make up the fixpoint axiom) are valid.

However, a strong link can be established if we substitute the notion of private individual belief $G_i A$ by that of manifested individual belief, as expressed by the formula $G_J G_i A$. Then both the fixpoint and the induction axiom can be proved in a rather straightforward way. The only restriction in this picture is that the first axiom $G_J A \rightarrow G_J G_i A$ applies to formulas that are *objective* for J only (see Definition of Subsection *Axiomatics*). This is due to our axiom (WR), and we have motivated the restriction in Subsection *Axiomatics*.

Beyond these principles, axioms (WR) and (CG) together give us $G_J A \leftrightarrow \bigwedge_{i \in J} G_J G_i A$ for objective A. This equivalence expresses that public common belief in a group is the same as the conjunction of manifested individual beliefs of the group members (i.e. the public version of the 'everybody believes' operator).

Link with Tuomela's group belief.

Tuomela has refined the notion of common belief and has investigated several forms of group belief (Tuomela 1992). He distinguishes *(proper) group beliefs* from *shared we-beliefs*. In the first case a group may typically believe a proposition while none of its members really believes it. In the second case, the group holds a belief which each individual agent really holds, too.

Our operator G_I is closer to Tuomela's *(proper) group beliefs* because the formula $G_I A \rightarrow G_i A$ is not valid. Thus, $G_I A$ means that a group I "(intentionally) jointly accept A as the view of I (...) and there is a mutual belief [about this]" (Tuomela 1992). We can consider that our operator G_I is a good approximation of the Group-Belief. Is is an approximation because we do not distinguish the agents contributing to the grounding of the group belief (the leaders) from those which passively accept it.)

Link with ostensible attitudes of Nickles et al.

Nickles et al. have proposed a logic of ostensible beliefs and intentions (Nickles, Fischer, & Weiss 2006; Nickles 2005). $Op(a_1, a_2, A)$ denotes "agent a_1 holds the ostensible belief A facing agent a_2". $OInt(a_1, a_2, A)$ denotes "agent a_1 facing agent a_2 exhibits the intention to make A true". They only give a basic semantics to their logic, on top of which some principles are stated axiomatically. For example, their axiom (2) is: $Op(a_1, a_2, A) \to \neg Op(a_1, a_2, \neg A)$.

Their notion of ostensible mental states is very close to our notion of grounding mental states, and their operators can be translated into our logic. $Op(a_1, a_2, A)$ corresponds to our $G_{\{a_1,a_2\}} G_{a_1} A$, and $OInt(a_1, a_2, A)$ to our $G_{\{a_1,a_2\}} Int_{a_1} A$.

For example, their axiom (2) becomes in our formalism $G_{\{a_1,a_2\}} G_{a_1} A \to \neg G_{\{a_1,a_2\}} G_{a_1} \neg A$. The latter is a theorem of our logic because the operator G_I satisfies the D-axiom for any group I.

Conclusion

The main contribution of this paper is the definition of a logic of grounding, extended to a group of agents. We have shown that this notion has its origins in speech act theory (Vanderveken 1990; 1991), philosophy of mental states (Searle 1983), and in philosophy of social action (Tuomela 1992), and is thus a philosophically well-founded notion. Its extension opens a large domain of exploration. We can formalize the performance of speech acts not only towards a single addressee but also towards a group. We have also added a distinction between the group of addressees and the group of attentive agents. This enables us to account for the interaction between two conversations, for example by reporting the sentences asserted in a previous conversation in a subsequent one.

The notion of grounding bridges the gap between mentalist and structural approaches. Just as the structural approaches to dialogue, it requires no hypotheses on the internal state of the agents, and formalizes for the observation of a dialogue by a third party. However, it also accounts for an objective viewpoint on dialogue because the logic also involves individual belief. And we have shown that we can formalize both dialogues where the speech acts semantics is defined with commitments (such as Walton & Krabbe's) as well as dialogues that are defined within a BDI-logic (FIPA-ACL).

Our characterization of speech acts is limited to the establishment of what must be true in order to avoid self-contradictions of the speaker. In further works we plan to refine this and define the FIPA-ACL library more precisely and from a public point of view.

We did not present a formal account of the dynamics. This requires the integration of a solution to the classical problems in reasoning about actions (frame problem, ramification problem, and belief revision). These technical aspects will be described in future work.

References

Austin, J. L. 1962. *How To Do Things With Words*. Oxford University Press.

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

Cohen, P. R., and Levesque, H. J. 1990b. Rational interaction as the basis for communication. In Cohen, P. R.; Morgan, J.; and Pollack, M. E., eds., *Intentions in Communication*. MIT Press.

Cohen, P. R., and Levesque, H. J. 1994. Preliminaries to a collaborative model of dialogue. *Speech Communication Journal'94, special issue on Spoken Dialogue* 15(3–4).

Conte, R., and Castelfranchi, C. 1995. *Cognitive and Social Action*. London: UCL Press.

Finin, T.; Labrou, Y.; and Mayfield, J. 1997. Kqml as an agent communication language. In Bradshaw, J., ed., *Software Agents*. AAAI/MIT Press.

FIPA. 2002a. FIPA Communicative Act Library Specification. http://www.fipa.org, Foundation for Intelligent Physical Agents.

FIPA. 2002b. FIPA Contract Net Interaction Protocol Specification. http://www.fipa.org, Foundation for Intelligent Physical Agents.

Fornara, N., and Colombetti, M. 2002. Operational Specification of a Commitment-Based Agent Communication Language. In Castelfranchi, C., and Johnson, L. W., eds., *Proc. First Int. Joint Conf. on Autonomous Agents and MultiAgent Systems (AAMAS-2002)*, volume 2 of *ACM Press*, 535–542.

Gaudou, B.; Herzig, A.; and Longin, D. 2005. A logic of grounding for interaction protocol analysis. Technical report, Institut de Recherche en Informatique de Toulouse, Toulouse, France. www.irit.fr/LILAC.

Gaudou, B.; Herzig, A.; and Longin, D. 2006. A logical framework for grounding-based dialogue analysis. In van der Hoek, W.; Lomuscio, A.; de Vink, E.; and Wooldridge, M., eds., *IJCAI International Workshop on Logic and Communication in Multiagent Systems (LCMAS 2005), Edinburgh, 1st of August, 2005*. To appear.

Herzig, A., and Longin, D. 2004. C&L intention revisited. In Dubois, D.; Welty, C.; and Williams, M.-A., eds., *Proc. 9th Int. Conf. on Principles on Principles of Knowledge Representation and Reasoning (KR2004) , Whistler, Canada*, 527–535. AAAI Press.

Nickles, M.; Fischer, F.; and Weiss, G. 2006. A Framework for the Representation of Opinions and Ostensible Intentions. In van der Hoek, W.; Lomuscio, A.; de Vink, E.; and Wooldridge, M., eds., *IJCAI International Workshop on Logic and Communication in Multiagent Systems (LCMAS 2005), Edinburgh, 1st of August, 2005*. To appear.

Nickles, M. 2005. Exposing the communication level of open systems: Expectations, ostensible attitudes and multi-source assertions. Technical report, University of Munich.

Rao, A. S., and Georgeff, M. P. 1991. Modeling rational agents within a BDI-architecture. In Allen, J. A.; Fikes, R.;

and Sandewall, E., eds., *Proc. Second Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'91)*, 473–484. Morgan Kaufmann Publishers.

Rao, A. S., and Georgeff, M. P. 1992. An abstract architecture for rational agents. In Nebel, B.; Rich, C.; and Swartout, W., eds., *Proc. Third Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'92)*, 439–449. Morgan Kaufmann Publishers.

Sadek, M. D. 1992. A study in the logic of intention. In Nebel, B.; Rich, C.; and Swartout, W., eds., *Proc. Third Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'92)*, 462–473. Morgan Kaufmann Publishers.

Searle, J. R., and Vanderveken, D. 1985. *Foundation of illocutionary logic*. Cambridge University Press.

Searle, J. R. 1969. *Speech acts: An essay in the philosophy of language*. New York: Cambridge University Press.

Searle, J. R. 1979. *Expression and Meaning. Studies on the Theory of Speech Acts*. Cambridge University Press.

Searle, J. R. 1983. *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.

Singh, M. P. 1998. Agent Communication Languages: Rethinking the Principles. *Computer* 31(12):40–47.

Singh, M. P. 2000. A Social Semantics for Agent Communication Languages. In Dignum, F., and Greaves, M., eds., *Issues in Agent Communication*, number 1916 in LNAI, 31–45. Springer-Verlag.

Traum, D. R. 1994. *Computational theory of grounding in natural language conversation*. Ph.D. Dissertation, Computer Science Departement, University of Rochester.

Tuomela, R. 1992. Group beliefs. *Synthese* 91:285–318.

Vanderveken, D. 1990. *Principles of language use*, volume 1 of *Meaning and Speech Acts*. Cambridge University Press.

Vanderveken, D. 1991. *Formal semantics of success and satisfaction*, volume 2 of *Meaning and Speech Acts*. Cambridge University Press.

Verdicchio, M., and Colombetti, M. 2003. A Logical Model of Social Commitment for Agent Communication. In *Proc. Second Int. Joint Conf. on Autonomous Agents and MultiAgent Systems (AAMAS-2003)*, 528–535. ACM Press.

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