Unifying the intentional and institutional semantics of speech acts

Carole Adam[1], Andreas Herzig[2], Dominique Longin[2], and Vincent Louis

[1] RMIT University, Melbourne, VIC, Australia
carole.adam.rmit@gmail.com
[2] Université de Toulouse, CNRS, IRIT (UMR 5505), France
firstname.surname@irit.fr

Abstract. Research about the semantics of agent communication languages traditionally sees the opposition between the mentalist and social approaches. In this paper we adopt a mixed approach since we propose a logical framework allowing us to express both the intentional and institutional dimensions of a communicative action. We use this framework to give a semantics for some speech acts representing each of Searle's categories except expressives. This semantics relaxes the criticized constraints imposed in FIPA-ACL and also extends this standard with new speech acts and new institutional features to characterise them. It has been implemented in an extension of the Semantic Add-on for the JADE agent development platform, and used in an industrial application in the context of automated B2B exchanges.

1 Introduction

Designing efficient Agent Communication Languages (ACL) is an essential issue in Multi-Agent Systems in order to standardise exchanges between the agents. Research about the semantics of ACL sees the subscribers of the social approach [17, 8, 27] criticize mentalist approaches [24, 15] for only grounding on the agents' private mental attitudes. But one can similarly reproach to social approaches to provide a semantics only based on the agents' public commitments, independently of their mental attitudes. Now these "social attitudes" are mainly descriptive, while mental attitudes allow one to predict the agents' behaviour. Moreover mental attitudes allow agents to reason about social notions. It is thus essential to consider both mental and social attitudes. Some researchers thus propose a mixed approach based both on public and private aspects [18]. But they do not formalise institutional speech acts like declarations. Now such speech acts are essential in new application fields involving communication about norms, roles or powers of agents, for instance in electronic commerce or automated business to business exchanges.

In this paper we thus want to propose an alternative to the well-known standard FIPA-ACL [16] through the following changes: relaxed feasibility preconditions to allow a more flexible utilisation of the speech acts in various contexts;

new institutional speech acts like declarations and promises; an institutional interpretation of speech acts coupled with their classical intentional interpretation. Therefore we adapt an existing logical framework for the formalisation of institutional notions like roles, powers and norms [10]. We then formalise in this logical framework the institutional interpretation of some specific communicative actions, each one representing one of Searle's categories of speech acts (except expressive ones) [25]. Our notion of institution is very large: it is a set of rules and facts adopted by a group of agents, like the rules of a game, or the laws of a country; it covers formal, legal institutions as well as informal ones (*e.g.* social rules in a group...).

The paper is structured as follow. Section 2 discusses some other social semantics of speech acts. Section 3 briefly describes the syntax, semantics and axiomatics of our logical framework. The core of the paper (Section 4) is dedicated to the unified semantics of speech acts. We are then able to compare our semantics of ACL with some related ones in more details (Section 5). Finally we conclude about the future prospects opened by this work (Section 6).

2 State of the art

The mentalist approach consists in grounding the semantics of speech acts on the agents' internal mental attitudes. These are represented by belief, desire and intention modalities provided by BDI logics, that are classically used to formalise the reasoning of autonomous agents [23, 28]. This resulted in the design of several standards of agent communication languages like KQML [15] or FIPA [16], this latter one grounding on Sadek's rational interaction theory [24].

These approaches were criticised a lot for being only based on private concepts (mental attitudes) instead of public verifiable notions (like commitments). Therefore some work exist aiming at enriching BDI logics with deontic operators like obligation [13, 3] or with institutional operators like *count as* or institutional power [21], in order to formalise the institutional interpretation of speech acts exchanged by the agents. In previous work we used such an extended BDI framework to express the semantics of speech acts with institutional effects [11] but we were limited to declarative speech acts, and the intentional and institutional dimensions were quite blended.

Various other work aim at providing an institutional semantics for speech acts. For example Dignum and Weigand [14] propose a logical framework combining illocutionary and deontic logic to study and model the norms resulting from communication between agents; however, they only consider directive speech acts. Boella *et al.* [1] propose a role-based semantics allowing them to combine social commitments and mental attitudes to express the semantics of speech acts in the context of persuasion dialogues. Actually they rewrite the FIPA feasibility precondition and rational effects of speech acts but replace the private mental attitudes involved by public mental attitudes attributed to the agents' roles instead of the individual agents. This solves the flaw of mentalist approaches, criticised for grounding on unverifiable mental attitudes, but finally

there is no distinct institutional interpretation of speech acts, that could differ from one institution to another. In the following subsections we give some details about two approaches: Fornara and Colombetti's approach based on the notion of commitments, and Lorini *et al.* 's approach based on the notion of acceptance.

2.1 Fornara and Colombetti: semantics in terms of social commitments

As opposed to the mentalist approach, the social one [26, 27, 8] assumes that private mental attitudes are not verifiable and thus grounds on the concept of public (thus verifiable) commitments [7] to express the semantics of speech acts. All the commitments taken by the agents are stored for possible future reference. The semantics of speech acts is expressed only in terms of such commitments.

For example Fornara and Colombetti [17] ground on Castelfranchi's notion of commitment [7] to define a library of communicative acts. From the classification of speech acts into four categories (assertives, directives, commissives and declaratives) inspired from Searle's work [25], they redefine for each category the semantics of its speech acts in terms of social commitments. Thanks to this library, they provide a communication tool based on social commitments, alternative to the FIPA-ACL standard. This tool allows rational agents to reason about the underlying rules of communication and to respect them in order for the system to behave well.

However these authors are limited to the institutional dimension of speech acts and neglect their relations with the agents' mental attitudes. Yet agents must be able to reason autonomously about the institution before making their decision to perform a given speech act. Moreover no specific institution is explicit in their commitments, making it impossible to have different commitments in different institutions; therefore it is also impossible for speech acts to have different effects depending on the institution within which they are interpreted. For example the action of nodding one's head is interpreted in the context of French gestural language as meaning "yes", while in the context of Bulgarian gestural language it is interpreted as meaning "no". In this example the considered institutions are the respective sets of communicative rules in these two cultures.

2.2 Lorini *et al.* : semantics in terms of group acceptance

Lorini *et al.* [22] define a new semantics for speech acts using Gaudou *et al.* 's Acceptance Logic [19]. $\mathscr{AL}$ is a modal logic extended with the notion of acceptance, representing what a group of agents willingly accept to consider as true (even if some (or all) members of the group believe the opposite) in a given institutional context (and that they can refuse in another context). Acceptances in an institutional context influence the agents' behaviour and utterances in this context. They are represented with the operator $[C : x]\varphi$ reading "agents in group C accept that φ while functioning as members of this group in the institutional context x".

Institutional notions are not primitive but defined from this notion of acceptance. Thereby a fact is an *institutional fact* (that it, a fact that is only valid in an institutional context, but not objectively valid) if and only if, for every non-empty set of agents, the agents in this set accept this fact as true while functioning as group members in this institutional context. In the context of ACLs, this may be a particular rule of the specific *Ordinary Communication* institution that these authors consider.

The authors then consider the speech act Promise in the institutional context of Ordinary Communication (OC). According to them, if i informs j that he is going to perform action α for him, and j intends i to perform this action for him, this counts as a promise at the next instant. The consequence of this promise is that i is obliged to perform action α for j. Moreover the acceptance by these two agents i and j while functioning as a group in institution OC that i has promised to perform action α for j and that j intends him to do so implies a social commitment of i towards j to perform α for him. This framework is interesting but Lorini *et al.* have only formalised the promise yet. Moreover they do not seem to make a clear distinction between the intentional and institutional preconditions to perform a speech act.

3 Our logical framework

We adapt here an existing logical framework for norms, institutional powers and roles defined in [10]. It is a multi-modal logic with modal operators of belief, intention, obligation, institutional facts and consequences, and action.

3.1 Syntax

Let $AGT = \{i, j, ...\}$ be a finite set of agents. Let $ACT = \{\alpha, \beta...\}$ be the set of actions. We suppose that some actions in ACT are of the form $i{:}\alpha$, where i is the author of action α (the agent who performs it). Let $ATM = \{p, q, ...\}$ be the set of atomic formulas. Let $INST = \{s, t, ...\}$ be the set of institutions. Complex formulas are denoted by φ, ψ... The language of our logic is defined by the following BNF grammar:

$$\varphi ::= p|\neg\varphi|\varphi \vee \varphi|B_i\varphi|Ch_i\varphi|I_i\varphi|D_s\varphi|\varphi \Rightarrow_s \varphi|O\varphi|before_\alpha\varphi|after_\alpha\varphi$$

where p ranges over ATM, α over ACT, i over AGT, and s over $INST$. The classical boolean connectives $\wedge$, $\rightarrow$, $\leftrightarrow$, $\top$ (tautology) and $\bot$ (contradiction) are defined from $\vee$ and $\neg$ in the usual manner. The operators $done_\alpha\varphi$, $happens_\alpha\varphi$, $P\varphi$, $F\varphi$ and $power(i, s, \varphi, \alpha, \psi)$ will be defined as abbreviations.

3.2 Semantics and axiomatics

We only give here the informal meanings of our operators. It is sufficient to know that they have a Kripke semantics in terms of possible worlds. We also give some

useful axioms. This framework is adapted from Demolombe and Louis' logic of norms, roles and institutional powers [10]. But please notice that actually, the details of the semantics of operators is not important, and any other institutional logic would work.

Belief, intention and action $B_i p$ means that agent i believes that p. $Ch_i p$ means that agent i prefers p to be true. These two normal operators have a standard KD45 axiomatics. $I_i p$ means that agent i intends that p. I_i are operators defined in a KD normal modal logic. Their axiomatics is that defined for FIPA by Sadek [24]. In particular intention is linked with belief by the following mix axioms:

- introspection: $I_i p \leftrightarrow B_i I_i p$
- automatic dropping of achieved intentions: $I_i p \rightarrow \neg B_i p$

$before_\alpha$ and $after_\alpha$ are normal modal operators defined in standard tense logic in linear time version [6]. $done_\alpha \varphi = \neg before_\alpha \neg \varphi$ means that action α has just been performed, and φ was true before. $happens_\alpha \varphi = \neg after_\alpha \neg \varphi$ means that action α is about to be performed and φ will be true just after.

Institutional modalities Finally this framework also provides some specific operators to formalise institutional concepts. These operators have a parameter s specifying the institution within which they are valid. Here we consider an institution as a set of institutional facts and rules that a group of agents (the "members" of this institution) adopt. This is a general view that can account for various institutional contexts, be they formal institutions or informal ones: the law of a country, a contract between two parties in a business relationship, a social structure, the rules of a game...

An **institutional fact** is a fact that is recognised to be valid in the context of a given institution, but that can make no sense in itself; *i.e.* it is not a physically observable fact (what Searle calls a "brute fact") but something written in the registry of this institution. For example the fact that two people are married, or that one is authorised to drive a truck, is only valid *w.r.t.* the law of a country; all deontic facts should also be encapsulated in an institutional fact to make the institution in which they hold explicit. We represent these institutional facts with the operator $D_s \varphi$ meaning that for institution s, it is officially established that φ holds. In particular if φ is an agent's mental attitude, then $D_s \varphi$ can be understood as this agent's commitment (either a propositional commitment if φ is a belief, or a commitment in action if φ is an intention). For instance, $D_{FrenchLaw}\, votingAgeis18$ means that following the French law, voting age is reached at 18 years; $D_{EU}\, euroOfficialMoney$ means that in the European Union, the official money is Euro.

Institutional facts can be deduced from other facts thanks to the rules of the institution. For example the presentation of an invoice by a provider to his client *counts as* an obligation for the client to pay it. The existence of the invoice is

physically observable, while the obligation is only valid in an institutional context. We represent these **normative consequences** with the primitive operator $p \Rightarrow_s q$, meaning that according to the norms holding in institution s, p entails q. This operator is known in the literature as *count as*, and has been first formalised by Sergot and Jones [21]. The following mix axioms explicit the link between institutional facts and normative consequences:

$$(\varphi \Rightarrow_s \psi) \rightarrow D_s(\varphi \rightarrow \psi) \quad \text{(SD)}$$
$$(\varphi \Rightarrow_s \psi) \rightarrow (\varphi \rightarrow D_s\varphi) \quad \text{(SC)}$$

From these axioms and the properties of D_s (see [10, p.8] for details) we can deduce:

$$(\varphi \Rightarrow_s \psi) \rightarrow (\varphi \rightarrow D_s\psi) \quad \text{(SP)}$$

A particular case of normative consequence concerns the consequences of the performance of an official procedure. Actually some agents can have the power when performing a given procedure under some conditions to create new institutional facts. We represent these **institutional powers** as an abbreviation $power(i, s, cond, \alpha, \varphi) = ((done_{i:\alpha}\top \wedge cond) \Rightarrow_s \varphi)$. Intuitively this means that i has the power in institution s, by performing action α and if condition *cond* holds, to see to it that φ becomes officially true in institution s. For example a mayor has the power in the law of the French Republic, by performing a declaration, and on condition that the two people agree, to marry them. Obviously these powers result from the agent's role in the institution, but this is not the focus of this paper so we will not remind how roles are formalised in the original framework (the interested reader can refer to [10] for details on this point).

Deontic modalities We have a modality for impersonal *obligation to be*: $O\varphi$ reads "it is obligatory that φ", and its axiomatic is that of the Standard Deontic Logic [20], that is KD. *Obligations to do* can be expressed as obligations to be in a state where the obliged action has been performed. Obligations are impersonal since no agent is explicitly responsible for their fulfilment, but such an agent can implicitly appear in their content. For instance $O done_{i:\alpha}\top$ means that it is obligatory (for no one in particular) to be in a state where i has just performed action α; this can be understood as "i has the obligation to perform action α".

Permissions and interdiction are defined from obligations in a standard way: $P\varphi = \neg O\neg p$ means that it is permitted that φ, and $F\varphi = O\neg\varphi$ means that it is forbidden that φ.

Please notice that no institution is explicit as a parameter of this obligation modality. But such obligations will be encapsulated in institutional facts to express the institution in which they are valid. For example $D_s O\varphi$ means that "in institution s, it is obligatory that φ".

4 Semantics of speech acts

4.1 Preliminary remarks

Intentional and institutional dimensions The FIPA-ACL standard [16] defines features allowing one to give an **intentional dimension** to the observation and interpretation of a communicative action: the feasibility precondition (the appropriate mental attitudes to perform the speech act) and the rational effect (this is a formula φ representing the content of the speaker i's intention that he intends the receiver j to know; so the performance of the speech act allows any observer k to deduce this corresponding intentional effect: $B_k I_i B_j I_i \varphi$). Please notice that, following [24], the performance of the speech act does **not** automatically allow one to deduce its rational effect, but only its intentional effect, meaning that any agent k believes that the speaker i intends the hearer j to recognize its (i's) intention to achieve the rational effect φ. However, nothing ensures that i indeed achieves φ, his speech act may fail, for example the hearer may not obey an order, or may not believe an assertion. Thus the rational effect can only be deduced under some constraining hypotheses such as the sincerity and competence hypotheses used in FIPA.

In a similar way, we want to provide here the **institutional dimension** of the observation and interpretation of a communicative action relative to one or several institutions. This institutional interpretation is composed of the following features:

- a *permission condition* that is necessary and sufficient for the speaker to be allowed to perform this speech act;
- a *power condition* that also needs to be true for the speech act to have an institutional effect;
- an *explicit institutional effect* that is obtained when the speech act is performed while permission and power conditions were true.

We will thus be able to combine the intentional and institutional dimensions of communicative actions (formalised as speech acts [25]), both essential to fully characterise their interpretation. Lorini *et al.* have also investigated such a unified approach but they have only formalised the interpretation of a promise in the context of ordinary communication; we aim at being much more generic. In particular we formalise one speech act from each of Searle's categories of illocutionary forces, except the expressive one.

Actually we have relaxed some of the (widely criticised) strong constraints imposed by FIPA-ACL semantics on the appropriate context of performance of speech acts. Instead of imposing these conditions as strong constraints, we have moved them into the permission preconditions of the speech act. The agents are thus physically able to disobey these constraints, but it is forbidden by the interaction norms, and they may incur sanctions for such violations. For example, relaxing the sincerity hypothesis physically allows the agents to lie, but this will be interpreted by other agents as a violation of communicative norms.

Notations In the sequel we use the following abbreviations:

- FP = feasibility preconditions
- RE = rational effect
- PermC = (institutional) permission condition
- PowC = power condition
- EE = institutional explicit effect

Speech acts are actions of the form $Force(sp, ad, inst, content)$ where $sp \in AGT$ is the speaker, $ad \in AGT$ is the addressee, $inst \in INST$ is the institutional context, *content* is the propositional content and can be any formula of our language, and $Force \in \{inform, promise, command, declare\}$ is the illocutionary force.

Action laws We now explain how the intentional and institutional dimensions of actions interact by providing the action laws governing the performance of speech acts.

We notice that FP is a factual executability precondition, while $PermC$ is an ideal one. But even ideal worlds are submitted to physical world laws, *i.e.* $PermC$ is not sufficient for the action to be executable, FP also has to be true. For example a mayor has the permission to marry people by making a declaration, but the declaration must be executable; thus if he is voiceless one day, he will be unable to marry anyone.

We thus have the following executability laws. The factual executability law (FEL_α) means that an action happens iff its feasibility precondition is true and the agent chooses to perform it. The ideal executability law (IEL_α) means that ideally, an action should happen only if it is permitted.

$$happens_\alpha \top \leftrightarrow (FP(\alpha) \wedge Ch_i happens_{i:\alpha} \top) \qquad (FEL_\alpha)$$

$$O(happens_\alpha \top \rightarrow PermC(\alpha)) \qquad (IEL_\alpha)$$

We also have the following effect laws. The rational effect law (REL_α) means that if the power precondition of an action is false, then only its rational effect can be deduced after its performance. The power effect law (PEL_α) means that if the power condition of an action is true, then both its rational and institutional effects can be deduced after its performance.

$$\neg PowC(\alpha) \rightarrow after_\alpha RE(\alpha) \qquad (REL_\alpha)$$

$$PowC(\alpha) \rightarrow after_\alpha(RE(\alpha) \wedge EE(\alpha)) \qquad (PEL_\alpha)$$

From these laws we can deduce the following theorems clarifying the factual executability and effects of α depending on the different combinations of its feasibility and power preconditions. If $FP(\alpha)$ is false then α is not executable.

$$\neg FP(\alpha) \rightarrow after_\alpha \bot$$

If i chooses to perform α when $FP(\alpha)$ is true but $PowC(\alpha)$ is false, then α is about to happen after which its rational effect will be true.

$$(Ch_i happens_{i:\alpha}\top \wedge FP(\alpha) \wedge \neg PowC(\alpha)) \rightarrow (happens_\alpha \top \wedge after_\alpha RE(\alpha))$$

Finally if i chooses to perform α when both $FP(\alpha)$ and $PowC(\alpha)$ are true, then α is about to happen after which both its rational and institutional effects will be true.

$$(Ch_i happens_{i:\alpha}\top \wedge FP(\alpha) \wedge PowC(\alpha)) \rightarrow (happens_\alpha \top \wedge after_\alpha (RE(\alpha) \wedge EE(\alpha)))$$

Please notice that $PermC(\alpha)$ does not appear in these last two theorems, since it does not influence the feasibility of α. Indeed an agent can choose to perform a forbidden action. If we had specified an explicit sanction $S(\alpha)$ for forbidden performance in the institutional interpretation of α, then we could write the following theorem:

$$Ch_i happens_{i:\alpha}\top \wedge FP(\alpha) \wedge \neg PermC(\alpha) \rightarrow (happens_\alpha \top \wedge after_\alpha S(\alpha))$$

However we did not specified such a sanction because it depends on many contextual parameters.

4.2 Assertives: inform

The assertive speech act *Inform* commits the speaker to the truth of a proposition. The notation $inform(i, j, s, \varphi)$ reads "agent i informs j in institution s that φ is true".

Intentional interpretation As we said before we have relaxed FIPA constraints on the executability of speech acts. We thus impose no feasibility precondition here.

$$FP(inform(i, j, s, \varphi)) = \top$$

The rational effect (the content of the speaker i's intention that he intends the receiver j to know) is that j believes the promised proposition φ to be true:

$$RE(inform(i, j, s, \varphi)) = B_j \varphi$$

Institutional interpretation The permission precondition to inform j that φ in institution s includes the constraints removed from the factual feasibility preconditions: the speaker should not believe that the hearer already knows if φ, and he should not be already committed on $\neg\varphi$ in the same institution.

$$PermC(inform(i, j, s, \varphi)) = \neg D_s B_i Bif_j \varphi \wedge \neg D_s B_i \neg\varphi$$

Now the institutional effect of *Inform* is to retract possible opposite commitments contracted before and to assert a new commitment on φ. Indeed, even if

agent i was previously committed on $\neg\varphi$ (and therefore was not permitted to inform anyone that φ), he may violate that obligation. But these two commitments are inconsistent so the previous one must be retracted while asserting the new contradictory one. Though one can still detect that the opposite commitment was true when i performed the action and that he has thus violated the rules of the institution. Actually due to the seriality of D_s we have that $D_s B_i \varphi \rightarrow \neg D_s B_i \neg\varphi$. So the explicit institutional effect of inform is a new institutional fact that can be interpreted as i's commitment to the truth of φ:

$$EE(\mathit{inform}(i, j, s, \varphi)) = D_s B_i \varphi$$

This effect is always obtained and does not depend on particular powers of i, so the power condition is trivial.

$$PowC(\mathit{inform}(i, j, s, \varphi)) = \top$$

Example For example in the context of B2B exchanges, if a provider sends his catalogue to a client, this counts as information about the prices given in this catalogue. As an effect of this action, the provider is thus committed to these prices during the validity of his catalogue. In the specific institution s constituted by the contract between the provider and the client, we assume that we have a specific rule forbidding to contradict one's commitments, which takes the form $D_s O(D_s B_i p \rightarrow \mathit{after}_{\alpha} D_s B_i p)$, for every speech act α, where p is the proposition denoting that the price is 100. This means that according to the institutional contract s between i and j, it is obligatory that if an agent i is committed to believe that the price of an item is 100, then after any speech act he is still committed to this (in other words it is forbidden to retract this commitment by any speech act). From this we can deduce that the provider is obliged to respect the given prices, *i.e.* $D_s O(D_s B_i p \rightarrow \mathit{after}_{Inform(i,j,s,\neg p)} \bot)$ (it is obligatory that if i is committed to p, then the action of informing agent j that $\neg p$ is not feasible).

Please note that the provider i can set up different contracts with different clients, in particular with different prices. This is made possible by making the institution explicit in the semantics of speech acts, and thus allowing us to specify different semantics in different institutions.

4.3 Commissives: promise to

This commissive speech act commits the speaker on a course of action. The notation *promise-to*(i, j, s, α) reads "i promises to j in institution s to perform action α".

Intentional interpretation We begin with specifying the intentional dimension of this speech act, that is not given in FIPA-ACL. A promise-to is feasible

if the speaker believes that the hearer intends the concerned action to be performed[3]. For example it makes no sense that a child promises to his father to play (this is rather an assertive), while it makes sense to promise him to make his schoolwork. So:

$$FP(\textit{promise-to}(i,j,s,\alpha)) = B_i I_j \, done_\alpha \top$$

The rational effect pursued by the speaker is that the hearer be aware of his intention to perform the promised action:

$$RE(\textit{promise-to}(i,j,s,\alpha)) = B_j I_i \, done_{i:\alpha} \top$$

Institutional interpretation In an institutional context s, this promise to perform an action α is permitted on condition that the action $i{:}\alpha$ is not explicitly forbidden itself, and that the speaker is not committed to an opposite course of action. So the permission precondition is the following:

$$PermC(\textit{promise-to}(i,j,s,\alpha)) = \neg D_s O \neg happens_{i:\alpha} \top \wedge \neg D_s I_i \neg done_{i:\alpha} \top$$

The institutional effect consists in ratifying in institution s the speaker's intention to perform action α; so after $\textit{promise-to}(i,j,s,\alpha)$ the speaker has stored in the registry of s its intention to perform α, which is similar to him being committed in s to this course of action.

$$EE(\textit{promise-to}(i,j,s,\alpha)) = D_s I_i \, done_{i:\alpha} \top$$

This is thus similar to the $\textit{inform}(i,j,s,\varphi)$ speech act except that a promise stores a commitment in action while an inform stores a propositional commitment.

There is no power precondition, so the institutional effect of a (permitted) promise is always reached.

$$PowC(\textit{promise-to}(i,j,s,\alpha)) = \top$$

Example A client c promises to pay his provider p once the ordered goods have been delivered. The action to pay is denoted by α_{pay} This promise is valid in the context of a B2B exchange contract, that is a particular institution denoted $b2b$ here. So this promise is formalised as: $promise(c,p,b2b,\alpha_{pay})$. This promise is permitted since obviously the promised action to pay is not forbidden ($\neg D_{b2b} O \neg happens_{c:\alpha_{pay}} \top$) and the client is not committed not to pay ($\neg D_{b2b} I_c \neg done_{c:\alpha_{pay}} \top$). So when the client receives the delivery, his promise allows to deduce his commitment (or ratified intention) to pay: $D_{b2b} I_c \, done_{c:\alpha_{pay}} \top$, that is the institutional effect of this speech act.

[3] Please notice that threats such as "I promise that I will kill you" cannot be considered as promises in the sense of Searle.

4.4 Directives: command

This directive speech act is commonly used by the speaker to make the hearer perform some action. The notation $command(i, j, s, \alpha, cond)$ reads "i orders to j in institution s, in virtue of condition $cond$, to perform action α".

Intentional interpretation According to the FIPA-ACL semantics, a request is feasible only if the speaker does not believe the hearer to already intend to perform the commanded action, and does believe that the part of the feasibility preconditions of the commanded action that concerns him (*i.e.* that are his mental attitudes) are valid. Here we consider that when α is an action of agent j then $FP(\alpha)$ is of the form $FP_i(\alpha) \wedge FP_{\neq i}(\alpha)$ where the former is "i's part of $FP(\alpha)$" (similar to FIPA-ACL notation $FP(\alpha)[i\backslash j]$, that is the part of $FP(\alpha)$ that are mental attitudes of agent i). But we do not impose this latter constraint on the feasibility of α as a feasibility precondition of the command. So:

$$FP(command(i, j, s, \alpha, cond)) = \neg B_i I_j \, done_{j:\alpha} \top$$

The rational effect of a command (*i.e.* the effect that i intends j to believe that i intends to achieve) is that j has performed the commanded action:

$$RE(command(i, j, s, \alpha, cond)) = done_{j:\alpha} \top$$

Institutional interpretation The permission precondition to command someone to perform an action is to be empowered to do so, *i.e.* to dispose of the institutional power to create the obligation to perform the commanded action by commanding it under some condition *cond* given as an explicit attribute of the command. [4]

An additional permission precondition is the constraint coming from FIPA feasibility precondition that we relaxed, that is that the part of the feasibility preconditions of α that depends on i hold (one should not command someone to perform an action whose preconditions are made false by his own mental attitudes). Finally, one is not permitted to command someone to perform a forbidden action.

$$\begin{aligned} PermC(command(i, j, s, \alpha, cond)) = \neg D_s \, O \neg happens_{j:\alpha} \top \wedge \\ power(i, s, cond, command(i, j, s, \alpha, cond), O \, done_{j:\alpha} \top) \wedge FP_i(\alpha) \end{aligned}$$

The explicit institutional effect of this power is to create two new institutional facts, corresponding to the obligation for j to perform α, and the recording of j's knowledge of his obligation. Actually this obligation could exist before, and in this case the command corresponds to a notification; but it can also be

[4] Institutional powers obviously depend on roles. This notion has been explored in previous work [10] but we will not enter in the details here since they are not in the scope of this paper.

created from scratch by the command (see the examples in the next paragraph for clarification).

$$EE(command(i, j, s, \alpha, cond)) = D_s O done_{j:\alpha} \top \wedge D_s B_j O done_{j:\alpha} \top$$

This explicit institutional effect is only deduced if the power applies in the current context, *i.e.* if its condition is true. So:

$$PowC(command(i, j, s, \alpha, cond)) = cond$$

Example For example a parent can command his children to clean his room. In this case, the action becomes obligatory through the command, because of the parent's authority over his son. In other words, his parent role gives him the institutional power to command his child to perform actions, under some conditions on the nature of the actions. Similarly a professor commanding his students to make some schoolwork creates the obligation for them to do so, on the strength of his role of professor. Indeed the role of professor gives an institutional power to command students to perform schoolwork, under the condition that it is related to the course.

But an order does not necessarily *create* an obligation, and may just put in focus an existing one. For example a bailiff can be sent to officially command an uncooperative client to pay an invoice. In this case the obligation already exists (and is attested by the invoice) so the bailiff only reminds the client of it[5]. He is permitted to perform such a command in virtue of his role of bailiff (which gives him the power to force clients to pay) and because he is sent by the provider (which constitutes the applicability condition of this power).

4.5 Declaratives: declare

This declarative speech act changes the institutional reality by creating a new institutional fact. The notation $declare(i, j, s, cond, \varphi)$ reads "i declares to j in institution s that given condition *cond*, the fact φ is now established". The condition usually bears upon the speaker's role that empowers him to perform such a declaration.

Intentional interpretation This intentional interpretation is partly inspired from the intentional interpretation of an $inform(i, j, s, D_s\varphi)$. The feasibility precondition of a declaration is that the speaker does not believe the declared fact to be already established (indeed a declaration must create a **new** institutional fact). The rational effect (*i.e.* the intended effect) has two parts: the first one is to make the declared institutional fact true; the second part is similar to the

[5] Actually this seems to be a notification rather than a command, but the aim is to make the client behave, while the aim of a notification is only to make the receiver officially aware of what is notified. In further work we expect to study into more details the links between declarations, commands and notifications.

rational effect of an inform about $D_s\varphi$, *i.e.* to make the hearer aware of this information. So:

$$FP(declare(i,j,s,cond,\varphi)) = \neg B_i D_s \varphi$$
$$RE(declare(i,j,s,cond,\varphi)) = B_j D_s \varphi \wedge D_s \varphi$$

Institutional interpretation The permission precondition to perform $declare(i,j,s,cond,\varphi)$ is that i really has the power to establish the declared fact φ by declaring it under the announced conditions *cond*. This power is locally granted by each specific institution to some agents depending on their role. For example the French republic grants the mayors the right to pronounce two people husband and wife, under the condition that they both consent to it. Thus an ordinary agent who is not mayor does not have this power, so that he is not allowed to pronounce marriages.

$$PermC(declare(i,j,s,cond,\varphi)) =$$
$$power(i,s,cond,declare(i,j,s,cond,\varphi),\varphi)$$

The explicit effect of a declaration is to store the declared fact in the institution, as well as the fact that the hearer is officially aware of this fact.

$$EE(declare(i,j,s,cond,\varphi)) = D_s\varphi \wedge D_s B_j D_s \varphi$$

This explicit effect is only obtained under the additional condition that *cond* is valid:

$$PowC(declare(i,j,s,cond,\varphi)) = cond$$

Example For example a country can declare war to another one, by the voice of its representative that is empowered to do so, and under some conditions like the agreement of some counsellors. A mayor is empowered by its country to pronounce weddings under some conditions that the people are of age and consenting.

Citizens have to declare their income to the public treasury in order to calculate the amount of tax that they will pay. This is a declaration since the effect is a new institutional fact officially establishing one's declared income as being believed by him. Any citizen is empowered to do so. Moreover the law imposes a constraint on the generated commitment, that is an obligation to believe this income to be true. Thereby if the declared income was false the citizen is liable for prosecution and sanctions.

4.6 Example of reasoning with our action laws

This example is situated in the context of a B2B exchange (in institution $b2b$) between a buyer b and a seller s. The seller intends potential clients to know the prices of his products, *e.g.* $I_s B_b p$. With our relaxed feasibility precondition, he can use an assertive speech act whatever the context. Though if the buyer has

already been informed of the prices before ($D_{b2b}B_sB_bp$), the seller is not permitted to inform him again. Thus if he informs him anyway, according to IEL_α he violates an obligation. This can be detected by other agents, and specific rules of the institution may specify sanctions to compensate this. Being aware of such pre-specified sanctions, an agent can deliberately choose to violate an obligation if the intended outcome (here, that clients be aware of the seller's offer) is more important than the incurred sanction. This shows the importance of having both intentional and institutional semantics of speech acts, to allow agents to reason about the relative importance of their goals and their obligations, in order to make an appropriate decision.

5 Detailed comparison with other work

In this section we compare our semantics of speech acts with those proposed by Fornara and Colombetti, and by Lorini *et al.* (that we have presented above).

5.1 Concept of commitment

We have shown before that what we mean by commitment in this work is a ratified mental attitude, *i.e.* a mental attitude (belief or intention) stored in the institution. This notion is similar to Fornara and Colombetti's commitment that is also a public concept, except that we have not made explicit its creditor. Actually the debtor is committed towards the whole institution, but an implicit creditor can sometimes be found in the content φ of the commitment. For example if agent i promises to j to pay him, he commits himself to a proposition involving agent j, expressing that j will be payed at some future instant. The creditor can sometimes be found in the sanction associated with the violation of the commitment, too; for example the obligation to pay damages to an agent.

Our notion of ratified mental attitude is also similar to Gaudou *et al.* 's notion of acceptance, because it must influence the agent's behaviour and utterances. Indeed, the agent's ratified mental attitudes are mental attitudes that he has expressed, that are stored in the institution, and to which he must conform while subsequently acting and speaking, even if they are not consistent with his real mental attitudes. For example an agent who promises that he has seen a given movie must then be able to talk about it in order to be consistent with his promise; if he is unable to narrate the end of the movie one can notice that he is contradicting his commitment.

5.2 Notion of institution

By institution we mean a set of rules and facts that are adopted by a group of agents (the members of the institution). This seems to be a more generic notion than Lorini *et al.* 's concept of informal institution, since it accounts for this particular kind of institutions but also for various other ones: laws of a country, rules of a game, contract between businesses, social norms of a culture... In

particular it allows to have institutional rules that are ignored by the members of the institution, what is the case for law for example, since one cannot be aware of the whole set of laws of his country, while he is one of its citizens. Fornara and Colombetti do not make explicit the institutional context in which their speech acts are interpreted, so we believe that they also consider a kind of "ordinary communication" institutional context.

In our view informal institutions are described by a specific set of facts and rules, determining their specific functioning. In particular the fact that all agents must accept a fact for it to become institutional is a particular institutional law. In other kinds of institutions, facts must be adopted by a majority of members (voting to create a law or to elect the president for example), or the opinion of one single member can suffice (the referee is always right). Thus we cannot adopt such an hypothesis in our account. Indeed on the contrary we consider the generic interpretation of speech acts in any institution s. More specific rules can be additionally specified in each particular interpretation, but the object of this paper is to identify for each category of speech acts the features that are common to their institutional interpretations whatever the institutional context.

6 Conclusion

In this work we have provided an expressive logical framework blending the agents' mental attitudes (beliefs, intentions) with their social attitudes (obligations, institutional facts and powers...). To illustrate its expressivity, please notice that our framework allows to represent some forms of contrary-to-duty obligations-to-do. Such obligations take the form:

$$O\mathit{after}_{\alpha}\bot \rightarrow \mathit{after}_{\alpha} O\mathit{done}_{\mathit{repair}_{\alpha}}\top$$

where repair_{α} is the contrary-to-duty obligation associated to the violation of the obligation to refrain from doing α. This means that if it is forbidden to perform α, then after α it is obligatory to perform a repairing action repair_{α}.

We have then used this framework to provide a semantics for an agent communication language based on FIPA-ACL but relaxing its widely criticised too constraining feasibility conditions, and adding permission preconditions. This way, agents can choose to perform forbidden speech acts but would then be liable to sanctions in the corresponding institution. Our ACL semantics also includes new speech acts (commissives and declaratives). It generalises existing approaches by unifying the intentional and institutional dimensions in one single framework, while strongly distinguishing them; moreover it allows to consider various kinds of institutional contexts; finally it provides action laws taking both dimensions into account.

In future work we intend to improve the institutional and intentional semantics of speech acts by accounting for deadlines. Various researchers [4, 12, 9] have shown that an important feature of obligations to perform an action is the deadline before which this action must be performed, that is essential to be able to assess the violation or fulfillment of such obligations. Though for the sake of

simplicity we have omitted deadlines in this paper. An idea to manage them in future work could be to use existing formalisations of norms with deadlines, or to ground on linear temporal logic with until and since operators [5].

Finally we would like to mention that our framework for the institutional interpretation of speech acts has been successfully implemented into institutional agents that have been used in a prototype of industrial application: a multi-agent mediation platform for automated business to business exchanges [2].

References

1. G. Boella, R. Damiano, J. Hulstijn, and L. van der Torre. Role-based semantics for agent communication: embedding of the 'mental attitudes' and 'social commitments' semantics. In *AAMAS'06*, Hakodate, Hokkado, Japon, 2006.
2. F. Bourge, S. Picant, C. Adam, and V. Louis. A multi-agent mediation platform for automatic b2b exchanges. In *ESAW*, 2008. demonstration.
3. J. Broersen, M. Dastani, and L. van der Torre. Beliefs, obligations, intentions, and desires as components in an agent architecture. *International Journal of Intelligent Systems*, 20(9):893–919, 2005.
4. J. Broersen, F. Dignum, V. Dignum, and J.-J. C. Meyer. Designing a deontic logic of deadlines. In *DEON*, pages 43–56, 2004.
5. J. Burgess. Logic and time. *The Journal of Symbolic Logic*, 44(4):566–582, 1979.
6. J. Burgess. *Handbook of philosophical logic*, volume 7, chapter Basic tense logic, pages 1–42. Kluwer Academic Publishers, 2nd edition, 2002.
7. C. Castelfranchi. Commitments : From individual intentions to groups and organizations. In *ICMAS-95*, pages 41–48, San Francisco, 1995.
8. B. Chaib-draa, M.-A. Labrie, M. Bergeron, and P. Pasquier. Diagal : An agent communication language based on dialogues games and sustained by social commitments. *Journal of Autonomous Agent and Multi-Agent Systems*, 13:61–95, 2006.
9. R. Demolombe, P. Bretier, and V. Louis. Norms with deadlines in dynamic deontic logic. In *ECAI'06*, pages 751–752, 2006.
10. R. Demolombe and V. Louis. Norms, institutional power and roles: towards a logical framework. In *16th International Symposium on Methodologies for Intelligent Systems (ISMIS'06)*, volume LNAI 4203, pages 514–523. Springer, 2006.
11. R. Demolombe and V. Louis. Speech acts with institutional effects in agent societies. In *DEON'06*, 2006.
12. F. Dignum and R. Kuiper. Obligations and dense time for specifying deadlines. In *HICSS*, volume 5, pages 186–195. IEEE Computer Society, 1998.
13. F. Dignum, D. Morley, E. Sonenberg, and L. Cavedon. Towards socially sophisticated bdi agents. In *ICMAS'2000*, pages 111–118, 2000.
14. F. Dignum and H. Weigand. Communication and deontic logic. In R. Wieringa and R. Feenstra, editors, *Information Systems, correctness and reusability*, pages 242–260, Singapore, 1995. World Scientific.
15. T. Finin, R. Fritzson, D. McKay, and R. McEntire. KQML as an agent communication language. In *Int. conf. Information and knowledge management*, 1994.
16. FIPA. The foundation for intelligent physical agents. `http://www.fipa.org`.
17. N. Fornara and M. Colombetti. A commitment-based approach to agent communication. *Applied Artificial Intelligence*, 18(9-10):853–866, 2004.

18. B. Gaudou, A. Herzig, D. Longin, and M. Nickles. A new semantics for the fipa agent communication language based on social attitudes. In G. Brewka, S. Coradeschi, A. Perini, and P. Traverso, eds, *ECAI*, pp. 245–249. IOS Press, 2006.
19. B. Gaudou, D. Longin, E. Lorini, and L. Tummolini. Anchoring institutions in agents' attitudes: towards a logical framework for autonomous multi-agent systems. In *AAMAS'08*, 2008.
20. A. Jones and J. Carmo. *Handbook of philosophical logic*, volume 8, chapter Deontic Logic and Contrary-to-duties, pages 265–343. Kluwer Academic Publishers, 2002.
21. A. Jones and M. Sergot. A formal characterisation of institutionalised power. *Journal of the interest group in pure and applied logics*, 4(3), 1996.
22. E. Lorini, D. Longin, and B. Gaudou. The institutional dimension of speech acts: a logical approach based on the concept of acceptance. Research report, IRIT, 2008.
23. A. Rao and M. Georgeff. Modeling rational agents within a BDI-architecture. In *KR'91*, 1991.
24. D. Sadek. A study in the logic of intention. In *KR'92*, 1992.
25. J. R. Searle. Speech acts : An essay in the philosophy of language. In *Cambridge Univ. Press*, 1969.
26. M. Singh. An ontology for commitments in multiagent systems: Towards a unification of normative concepts. 1999.
27. M. P. Singh. A social semantics for agent communication languages. In F. Dignum and M. Greaves, eds., *Issues in Agent Communication*, pp. 31–45. Springer, 2000.
28. M. Wooldridge. *Reasoning about rational agents*. MIT Press, 2000.