

Time and Defeasibility in FIPA ACL Semantics

Guido Boella *University of Turin, Italy*

Guido Governatori *NICTA, Australia*

Joris Hulstijn *Free University Amsterdam, The Netherlands*

Régis Riveret *Independent researcher*

Antonino Rotolo *University of Bologna, Italy*

Leendert van der Torre *University of Luxembourg*

Abstract

Inferences about speech acts are often conditional, non-monotonic, and involve the issue of time. Most agent communication languages, however, ignore these issues, due to the difficulty to combine them in a single formalism. This paper addresses such issues in defeasible logic, and shows how to express a semantics for ACLs in order to make non-monotonic inferences on the basis of speech acts.

Keywords: Non-monotonic reasoning, Agent communication languages, FIPA

1. Introduction

FIPA Semantic Language or FIPA SL is a popular language introduced by the Foundation for Intelligent Physical Agents (FIPA) to describe the meaning of speech acts in terms of pre- and postconditions on mental attitudes [1]. Alternative traditions define the pre- and postconditions on social commitments, or unify both approaches using the same FIPA SL to describe the dynamics of so-called public mental attitudes associated with the communication roles [2]. However, the drawback of FIPA SL is that it only partially specifies the meaning of the speech acts. For example, it does not specify how mental attitudes persist in time, and the informal specification refers to agent properties like trustworthiness, but these properties are not represented in the formal language FIPA SL. Therefore other semantic languages have been proposed to describe the meaning of speech acts, such as epistemic dynamic logic [3].

The research question of this paper is to fully describe the pre- and postconditions of speech acts in a formal language, such that if an agent observes a sequence of speech acts, this agent can use the formal language and calculate the resulting state. Thus the formal language should be computationally efficient and unambiguous. For example, if the trustworthiness of agents is relevant for the resulting state, then the trustworthiness of the agents should be incorporated in the language. Moreover, the language should be expressive enough to reason about the most common patterns, in particular it should represent and reason about temporal phenomena, and it should be able to deal with common sense or non-monotonic reasoning.

We propose a defeasible logic (DL) extended with time and modal operators as representation language and reasoning mechanism to formalise speech acts. The main advantages of DL are that

- DL is based on a logic programming-like language and it is a simple, efficient but flexible formalism capable of dealing with many different intuitions of non-monotonic reasoning [4]; hence, it allows us to capture the non-monotonic behaviour of speech acts;
- DL has a linear complexity [5] and has also efficient implementations [6, 7, 8]; this computational result can be extended to DL with time and modalities, since it is possible to compute the extension of a temporal defeasible theory in time linear to the size of the theory;
- Significant extensions of DL have incorporated various modal operators in order to model cognitive agents [9, 10];
- DL has been extended to capture the temporal aspects, such as persistence, of several specific phenomena, like legal positions [11], legal modifications [12, 13, 14], and deadlines [15]; this framework for temporal persistence can be adjusted to model the persistent and transient nature of speech acts and their effects.

To some extent some of these features (i.e., non-monotonicity and temporal persistence) are also enjoyed by Event Calculus [16]. However, DL has some advantages over Event Calculus. It is not possible for Event Calculus to compute all conclusions of a theory in time linear to the size of the theory. Furthermore, efficient implementations exist for defeasible logic [6, 8], also when modalities [7, 8] and time [17, 18] are added.

We consider the following success guidelines for our formal language and the representation of speech acts:

Calculate the effect of speech acts such that agents can determine the state in which they are, given the history of a sequence of speech acts;

Define agent types and dialogue types as respectively a set of formulas and a set of rules of the language;

What if analysis of speech acts for the planning of agents of new speech acts (dialogue generation);

Modularity of the formalization of the consequences of speech acts, distinguish public from private mental attitudes, define modules for agent types and dialogue types.

FIPA Communicative Act Library Specification [1] defines not only the speech act itself, it additionally models the content, e.g.: a *request* speech act has an action in the content field. In *request-when* the content field holds another speech act. The expressivity of such content fields easily reaches first-order logic (plus modal operators). This is a problem if an agent wants to do reasoning about such a message, but this topic is outside the scope of the paper.

Moreover, in this paper we are not going to develop from scratch a new logic, but we use and adjust an existing one (DL with time). Moreover, we are not going to introduce a new approach to agent communication, but we use an existing one (role based communication unifying mental attitudes and social commitments approaches). Our contribution in this paper is in the use of the language to represent and reason about agent communication. We do not consider full planning of speech acts, which is left for further research.

The paper is organised as follows. In Section 2, we introduce some motivations and examples from different applicative scenarios which are not easy to deal with in the current approaches. In

Section 3, a variant of Temporal Defeasible Logic (TDL) is proposed to model FIPA speech acts. Section 4 shows how the TDL formalism can express the semantics for ACLs. Related work is discussed in Section 5. Finally, we conclude and present some directions for future work.

2. Non-monotonicity, persistence and rules in dialogues

In this section we introduce some motivations and examples from different applicative scenarios which are not easy to deal with in the current approaches.

It is widely accepted that the FIPA semantics for Agent Communication Languages (ACL) has some flaws [19]. First, the semantics is expressed in terms of private mental attitudes of agents. Therefore, it cannot be verified given common assumptions about agent systems. Second, the sincerity condition assumed in FIPA may be acceptable in cooperative circumstances, but is clearly wrong for persuasion and negotiation dialogues. Two solutions for these flaws have been discussed:

- An ACL semantics can be given in terms of social attitudes, like commitments [20, 21]. Social attitudes are maintained in public, as a kind of common data structure.
- An ACL semantics can be given in terms of public mental attitudes instead of private ones. This preserves the characteristics of the BDI-approach that underlies the FIPA semantics, without incurring the verifiability problems. Thus, [22, 23], for instance, use the notion of common ground, [24] refers to ostensible beliefs and goals and [25, 2] introduce public roles.

These solutions fail to consider other problems, which emerge when new application domains are considered: in particular, *ACL semantics*

- do not specify how to deal with *non-monotonicity* in communicative actions;
- mostly ignore the problem of *time*, i.e., the possibility of timing the different aspects of communication, and the *persistence* of preconditions and effects.

Non-monotonicity.

1. Consider an interactive application to persuade buyers, arguing with a user:

Seller: Why don't you buy a Mercedes A class? It is safe.

Buyer: I heard that it may roll over.

Seller: Yes, maybe it is not safe in that sense, but I know that you are not a driver who takes risks.

In this example, the seller has to retract his assertion on the safety of the product and to make a weaker one. Retraction makes inference about communicative actions non-monotonic. When retraction is possible, it is no longer true that every larger set of speech acts supports the same conclusions as its subsets. In this sense, FIPA-ACL semantics is monotonic, while social commitment approaches rarely deal with the propositional content of the speech acts.

2. Consider a debate with a politician, staged on Second Life. The application will record the assertions of the politician and will check automatically its coherence, to coerce him not to contradict himself. Among other things, the politician says that he is against alcohol abuse. The day after, you read on the news forum, that he was caught while driving drunk. You have to revise your conclusion that he really believed what he said online, but his public claims cannot be revised. Similar issues arise when deciding to trust a vendor in e-commerce; initially you may trust a vendor at face value, but after some unpleasant experiences, you will have to revise your beliefs.

FIPA's main problem is the unverifiability of private mental states on which the semantics is based. To solve this, approaches like [22, 23, 24, 26, 25, 2] distinguish public and private mental attitudes, but they fail to consider the defeasible connection between them. Hence, the concept of non-monotonicity plays a role here, too.

Monotonic inferences about effects and preconditions of communicative acts are only possible regarding the public beliefs or goals of the participants (if at all), while inferences about the private beliefs of the participants can be made only by default, under the assumption of cooperativity, trust or sincerity. It is not possible to model a dialogue on the basis of private beliefs and goals only: public mental states are needed since in many situations dialogues go on correctly, even if the participants are not cooperative, sincere or do not trust each other.

3. ACL semantics like FIPA, do not in general allow explicitly to make inferences about the success of a rational effect, even though informally the FIPA specification (p.10) says: "Whether or not the receiver does, indeed, adopt belief in the proposition will be a function of the receiver's trust in the sincerity and reliability of the sender." In many circumstances the speaker can in fact attribute to the hearer the proposition he has asserted, though in a defeasible way. For example, in collaborative situations when he knows that the hearer considers him reliable, or in persuasion dialogues when the hearer does not challenge claim. However, when a claim is in fact challenged, or when the circumstances turn out to be hostile instead of cooperative, such initial inferences of 'success', have to be retracted.

Time and persistence.

4. In an e-dispute, a legal case is debated. Suppose Ian sells at time 2 a faulty children's chair to John, causing John's child to get hurt. It is legally presumed that when a vendor sells a product, he acts in good faith, i.e., as if he were informing the client that the product is without faults. A witness, Kay, reveals at time 3 to John that the vendor had independently informed her at time 1 that the chair was faulty. Assume that the sincerity precondition of an 'inform' act persists into the future, unless there is evidence to the contrary. Given this assumption, we have at times 2 and 3 two conflicting vendor's beliefs about the quality of the chair and two conflicting client's beliefs about the same.

ACL approaches like FIPA generally do not model the persistence of preconditions into the future, let alone into the past.

5. Consider an e-commerce application, in which an agent makes an offer or a proposal. How long does the offer persist? Is the offer still valid after a refusal? Commitment based approaches like [27] deal with this by means of a finite state machine describing the state of a commitment. Typically, a proposal will persist, until it is either accepted or rejected. FIPA, instead, ignores whether the effect of a speech act will persist or not.

6. Consider instead an application to make a user cooperate. What happens to a request? Does it persist like a proposal? Probably it does not: if a request is not ‘taken on’, it just passes away. For the same reason time is made explicit in the Contract Net Protocol. There is a deadline on responses to the *call for proposals* (CFP) speech act which requests a proposal from bidders. So apparently, the CFP only ‘persists’ until the deadline.
7. Finally, in a newsfeed service we would like to have some quality of service warranties about the timing of the information. Suppose I want to be informed about all stock exchange data of Air France. I expect an inform about any changes in the value of Air France, as soon as possible after the actual change occurred. Else, we get a kind of ‘temporal insincerity’ of the newsfeed agent.

3. Temporal Defeasible Logic

This paper provides a computational-oriented approach to describe the behaviour of FIPA speech acts. We are interested in modelling their non-monotonic character, their temporal aspects and the persistence of their preconditions and effects. In this section we propose a variant of Temporal Defeasible Logic (TDL), which extends and modifies the logic of [11] in order to model FIPA speech acts.

3.1. Language

Standard DL is based on a formal language which is usually propositional and consists of a set of literals, namely a set of propositional statements and their negations. In other words, once defined a set $\text{Prop} = \{p, q, \dots\}$ of propositional atoms, a literal l is a member p of Prop or its negation $\neg p$. By convention, given a literal l , the *complement* $\sim l$ is $\neg p$ if $l = p$, and p if $l = \neg p$.

To accommodate the cognitive and temporal aspects of the speech acts we need to develop a suitable variant of DL (Temporal Defeasible Logic, TDL) which enriches the basic language of standard DL with the following elements:

- a discrete totally ordered set of instants of time $\mathcal{T} = \{t_0, t_1, t_2, \dots\}$;
- a set of agents $\text{Ag} = \{a, b, \dots\}$;
- a set of agent roles $\text{Role} = \{r_1, r_2, \dots\}$ (given $x \in \text{Ag}$ we use $r(x) \in \text{Role}$ to denote the role played by agent x);
- the families of modal operators $\text{Bel} = \{B_x\}_{x \in \text{Ag} \cup \text{Role}}$ (for belief) and $\text{Goal} = \{G_x\}_{x \in \text{Ag} \cup \text{Role}}$ (for goal);
- a set of speech act types $ST = \{\text{inform}, \text{promise}, \dots\}$.

Based on the above elements a *modal literal* is an expression $X^t l$ (or its negation, i.e., $\neg X^t l$) where X is a modal operator, $t \in \mathcal{T}$ and l is either a literal or a modal literal. For example, the expression $B_a^t G_{r(b)}^t p$ means that agent a believes at t that the goal of the role played by agent b at time t' is (was, will be) p .

A *speech act* is an expression $st_{i,j}(s,t)$ or the negation of it (i.e., $\neg st_{i,j}(s,t)$), where $i, j \in \text{Ag} \cup \text{Role}$, s is either a literal, a speech act, or a modal literal, $st \in ST$, and $t \in \mathcal{T}$. For example, the expression $\text{inform}_{r(a),r(b)}(p,t)$ means that agent a , in her role, informs at time t agent b , in

his role, that p is the case. (Note that we assume that the iteration of modal operators and speech acts is finite.)

On account of the above extension, the notion of *complement* must be extended as well. The complement of a literal l will denote the set of literals that are in conflict with l . If l is any type of literal, $\mathcal{C}(l)$ denotes the complement of l , i.e., the set of literals that cannot hold when l does. Thus $\mathcal{C}(l)$ is defined as follows:

- if $l = p \in \text{Prop}$, then $\mathcal{C}(l) = \{\neg p\}$;
- if $l = \neg p \in \text{Prop}$, then $\mathcal{C}(l) = \{p\}$;
- if $l = X^t m$, such that $X \in ST \cup \text{Bel} \cup \text{Goal}$ and m is either a literal or a modal literal, then $\mathcal{C}(l) = \{X^t n \mid n \in \mathcal{C}(m)\} \cup \{\neg X^t m\}$;
- if $l = \neg X^t m$, such that $X \in ST \cup \text{Bel} \cup \text{Goal}$ and m is either a literal or a modal literal, then $\mathcal{C}(l) = \{X^t n \mid n \in \mathcal{C}(m)\} \cup \{X^t m\}$.

Hereafter, we will use $X^t p$ to denote any modal literal holding at time t or any speech act $st_{i,j}(p,t)$. A *rule* is an expression $lbl : A \hookrightarrow_x m$, where lbl is a unique label of the rule, A is a (possibly empty) set of speech acts and modal literals, $\hookrightarrow \in \{\rightarrow, \Rightarrow, \rightsquigarrow\}$, m is either a speech act or a modal literal, and x is either π or τ signaling whether we have a *persistent* or *transient* rule. *Strict rules*, marked by the arrow \rightarrow , support indisputable conclusions whenever their antecedents, too, are indisputable. *Defeasible rules*, marked by \Rightarrow , can be defeated by contrary evidence. *Defeaters*, marked by \rightsquigarrow , cannot lead to any conclusion but are used to defeat some defeasible rules by producing evidence to the contrary. A *persistent* rule is a rule whose conclusion holds at all instants of time after the conclusion has been derived, unless interrupting events occur; *transient* rules establish the conclusion only for a specific instant of time. Thus, $ex_1 : B_a^5 p \Rightarrow_\pi G_{r(a)}^6 q$ means that if agent a at time 5 believes p , then, defeasibly, the role played by agent a has the goal q at time 6 and the goal continues to hold after 6 until some event overrides the goal of q . If we change π into τ the resulting rule $ex_2 : B_a^5 p \Rightarrow_\tau G_{r(a)}^6 q$ means that $r(a)$ has the goal q at time 6, but we do not know whether the goal will persist after 6. Note that we assume that defeaters are only transient: if a persistent defeasible conclusion, such as $G_a^5 p$ is blocked at time 6 by a transient defeater $A \rightsquigarrow_\tau \neg G_a^6 p$, such a conclusion no longer holds after 6 unless another applicable rule reinstates it (for the details, see below, Section 3.2).

We will use some abbreviations. Given a rule r and a set R of rules, $A(r)$ denotes the antecedent of r while $C(r)$ denotes its consequent; R^π denotes the set of persistent rules in R , and $R[\psi]$ the set of rules with consequent ψ . R_s , R_{sd} and R_{df} are respectively the set of strict rules, the set of strict and defeasible rules, and the set of defeaters in R .

3.2. Proof Theory

There are in TDL three kinds of features: facts, rules, and a superiority relation among rules. Facts are indisputable statements, represented by modal literals and speech acts. The superiority relation (\succ) provides information about the relative strength of rules, i.e., about which rules can overrule which other rules. A knowledge base that consists of these items is called a TDL theory.

Definition 1. A TDL theory is a structure (F, R, \succ) , where

- F is a finite set of facts,

- R is a finite set of rules, and
- \succ is an acyclic binary relation over R .

TDL is based on a constructive inference mechanism based on tagged conclusions. Proof tags indicate the strength and the type of conclusions. The strength depends on whether conclusions are indisputable (the tag is Δ), namely, obtained by using facts and strict rules, or they are defeasible (the tag is ∂). The type depends on whether conclusions are obtained by applying a persistent or a transient rule: hence, conclusions are also tagged with π (persistent) or τ (transient).

Provability is defined below and is based on the concept of a derivation (or proof) in a TDL theory D .

Definition 2. Given a TDL theory D , a proof P from D is a finite sequence of tagged modal literals or speech acts such that:

- (1) Each tag is one of the following: $+\Delta^\pi$, $-\Delta^\pi$, $+\partial^\pi$, $-\partial^\pi$, $+\Delta^\tau$, $-\Delta^\tau$, $+\partial^\tau$, $-\partial^\tau$;
- (2) The proof conditions Definite Provability (Positive), Definite Provability (Negative), Defeasible Provability (Positive), and Defeasible Provability (Negative) given below are satisfied by the sequence P .

Given a proof P we use $P(n)$ to denote the n -th element of the sequence, and $P[1..n]$ denotes the first n elements of P .

The meaning of the proof tags is as follows:

- $+\Delta^\pi X^{t_p} p$ (resp. $+\Delta^\tau X^{t_p} p$): we have a definite derivation (i.e., a proof using only strict rules and facts) of Xp holding from time t_p onwards (resp. of Xp holding at t_p);
- $-\Delta^\pi X^{t_p} p$ (resp. $-\Delta^\tau X^{t_p} p$): we can show that it is not possible to have a definite derivation of Xp holding from time t_p onwards (resp. of Xp holding at t_p);
- $+\partial^\pi X^{t_p} p$ (resp. $+\partial^\tau X^{t_p} p$): we have a defeasible derivation of Xp holding from time t_p onwards (resp. of Xp holding at t_p);
- $-\partial^\pi X^{t_p} p$ (resp. $-\partial^\tau X^{t_p} p$): we can show that it is not possible to have a defeasible derivation of Xp holding from time t_p onwards (resp. of Xp holding at t_p).

Let us now provide the proof conditions corresponding to the above proof tags.

Definite Provability (Positive).

If $P(n+1) = +\Delta^x X^{t_p} p$, then

- 1) $X^{t_p} p \in F$ if $x = \tau$; or
- 2) $\exists r \in R_s^x[X^{t'_p} p]$ such that $\forall Y^{t_a} a \in A(r) : +\Delta^y Y^{t_a} a \in P[1..n]$.

where:

- (a) $y \in \{\pi, \tau\}$;
- (b) if $x = \pi$, then $t'_p \leq t_p$;

(c) if $x = \tau$, then $t'_p = t_p$.

If the conclusion is transient (if $x = \tau$), the above conditions are basically the standard ones for definite proofs in DL, which are just monotonic derivations using forward chaining. If the conclusion is persistent ($x = \pi$), Xp can be obtained at t_p or, by persistence, at any time t'_p before t_p . Finally, notice that facts lead to strict conclusions, but are taken not to be persistent.

Example 1. Consider this theory:

$$(F = \{\text{inform}_{r(a),r(b)}(p,t_p)\}, R = \{\text{inform}_{r(a),r(b)}(p,t_p) \rightarrow \pi B_{r(a)}^{t_p} p\}, \succ = \emptyset)$$

We can derive $+\Delta^\pi B_{r(a)}^{t_p} p$, and $+\Delta^\pi B_{r(a)}^{t_p} p$ for $t > t_p$.

Defeasible Provability (Positive).

If $P(n+1) = +\partial^x X^{t_p} p$, then

- 1) $+\Delta^x X^{t_p} p \in P[1..n]$ or
- 2) $\forall l \in \mathcal{C}(X^{t_p} p)$, $-\Delta^x l \in P[1..n]$, and
 - 2.1) $\exists r \in R_{sd}^x[X^{t_p} p]$ such that
 - $\forall Y^{t_a} a \in A(r) : +\partial^y Y^{t_a} a \in P[1..n]$, and
 - 2.2) $\forall s \in R^y[l]$, where $l \in \mathcal{C}(X^{t_p} p)$, either
 - 2.2.1) $\exists Y^{t_b} b \in A(s)$, $-\partial^y Y^{t_b} b \in P[1..n]$ or
 - 2.2.2) $\exists w \in R^y[X^{t_p} p]$ such that
 - $\forall Y^{t_c} c \in A(w) : +\partial^y Y^{t_c} c \in P[1..n]$ and $w \succ s$, and
 - 2.3) $\forall s \in R^y[\neg X^{t_{\sim p}} p] \cup R^y[X^{t_{\sim p}} n]$, where $n \in \mathcal{C}(p)$, either
 - 2.3.1) $\exists Y^{t_b} b \in A(s)$, $-\partial^y Y^{t_b} b \in P[1..n]$ or
 - 2.3.2) $\exists w \in R^y[X^{t_{\sim p}} p]$ such that
 - $\forall Y^{t_c} c \in A(w) : +\partial^y Y^{t_c} c \in P[1..n]$ and $w \succ s$.

where

- (i) $y \in \{\pi, \tau\}$;
- (ii) if $x = \pi$, then $t'_p \leq t_{\sim p} \leq t_p$;
- (iii) if $x = \tau$, then $t'_p = t_{\sim p} = t_p$.

Defeasible derivations run in three phases. In the first phase we put forward a supported reason (rule) for the conclusion we want to prove. Then in the second phase we consider all possible (actual and potential) reasons against the desired conclusion. Finally in the last phase, we have to rebut all the counterarguments. This can be done in two ways: we can show that some of the premises of a counterargument do not obtain, or we can show that the argument is weaker than an argument in favour of the conclusion. If $x = \tau$, the above conditions are essentially those for defeasible derivations in DL. If $x = \pi$, a proof for $X^{t_p} p$ can be obtained by using a persistent rule which leads to Xp holding at t_p or at any time t'_p before t_p . In addition, for every instant of time between the t'_p and t_p , p should not be terminated. This requires that all possible attacks were not triggered (clauses 2.2.1 and 2.3.1) or are weaker than some reasons in favour of the persistence of Xp (clauses 2.2.2 and 2.3.2). Notice that clause (2.2) takes into account all possible attacks at the present time, while clause (2.3) considers attacks in the past. The idea of

(2.3) is that the conclusion was proved as persistent in the past and that all the attacks from then to now have been defeated at the time of the attack.

The inference conditions for negative proof tags ($-\Delta$ and $-\partial$), here below, are derived from the inference conditions for the corresponding positive proof tags by applying the Principle of Strong Negation introduced by [28]. The strong negation of a formula is closely related to the function that simplifies a formula by moving all negations to an innermost position in the resulting formula and replaces the positive tags with the respective negative tags and viceversa.

Definite Provability (Negative).

If $P(n+1) = -\Delta^x X^{t_p} p$, then

- 1) $X^{t_p} p \notin F$ if $x = \tau$; and
- 2) $\forall r \in R_s^x[X^{t_p} p] : \exists Y^{t_a} a \in A(r) : -\Delta^y Y^{t_a} a \in P[1..n]$.

where:

- (a) $y \in \{\pi, \tau\}$;
- (b) if $x = \pi$, then $t'_p \leq t_p$;
- (c) if $x = \tau$, then $t'_p = t_p$.

Defeasible Provability (Negative).

If $P(n+1) = -\partial^x X^{t_p} p$, then

- 1) $-\Delta^x X^{t_p} p \in P[1..n]$ and either
- 2) $\exists l \in \mathcal{C}(X^{t_p} p), +\Delta^x l \in P[1..n]$ or
 - 2.1) $\forall r \in R_{sd}^x[X^{t_p} p] :$
 $\exists Y^{t_a} a \in A(r) : -\partial^y Y^{t_a} a \in P[1..n]$, or
 - 2.2) $\exists s \in R^y[l]$, where $l \in \mathcal{C}(X^{t_p} p)$, such that
 - 2.2.1) $\forall Y^{t_b} b \in A(s), +\partial^y Y^{t_b} b \in P[1..n]$ and
 - 2.2.2) $\forall w \in R^y[X^{t_p} p]$
 $\exists Y^{t_c} c \in A(w) : -\partial^y Y^{t_c} c \in P[1..n]$ or $w \neq s$, or
 - 2.3) $\exists s \in R^y[\neg X^{t_{\sim p}} p] \cup R^y[X^{t_{\sim p}} n]$, where $n \in \mathcal{C}(p)$, such that
 - 2.3.1) $\forall Y^{t_b} b \in A(s), +\partial^y Y^{t_b} b \in P[1..n]$ and
 - 2.3.2) $\forall w \in R^y[X^{t_{\sim p}} p] :$
 $\exists Y^{t_c} c \in A(w) : -\partial^y Y^{t_c} c \in P[1..n]$ or $w \neq s$.

where

- (i) $y \in \{\pi, \tau\}$;
- (ii) if $x = \pi$, then $t'_p \leq t_{\sim p} \leq t_p$;
- (iii) if $x = \tau$, then $t'_p = t_{\sim p} = t_p$.

As we have already said the conditions for the negative proof tags are derived from the conditions for the corresponding positive proof tag. In general the inference conditions for a negative proof tag explore all the possibilities to derive a literal (with a given proof strength) before stating that the literal is not provable (with the same proof strength). Thus conclusions with these tags are the outcome of a constructive proof that the corresponding positive conclusion cannot be obtained.

Example 2. Consider the following theory.

$$\begin{aligned}
(F &= \{\text{inform}_{r(a),r(b)}(p,t_1), G_a^{t_3}q, B_{r(a)}^{t_3}c, G_b^{t_4}B_a^{t_4}d\}, \\
R &= \{r_1 : \text{inform}_{r(a),r(b)}(p,t_1) \Rightarrow_{\pi} B_a^{t_1}p, \\
&\quad r_2 : G_a^{t_3}q \Rightarrow_{\pi} B_a^{t_3}\neg p, \\
&\quad r_3 : B_{r(a)}^{t_3}c \rightsquigarrow_{\tau} B_a^{t_3}p, \\
&\quad r_4 : G_b^{t_4}B_a^{t_4}d \Rightarrow_{\tau} \neg B_a^{t_4}p\}, \\
> &= \{r_3 > r_2, r_1 > r_4\}
\end{aligned}$$

At time t_1 , r_1 is the only applicable rule; accordingly we derive $+\partial^{\pi}B_a^{t_1}p$. At time t_2 no rule is applicable, and the only derivation permitted is the derivation of $+\partial^{\pi}B_a^{t_2}p$ by persistence. At time t_3 both r_2 and r_3 are applicable, but r_4 is not. If r_2 prevailed, then it would terminate B_ap . However, it is rebutted by r_3 , so we derive $+\partial^{\pi}B_a^{t_3}p$. Finally at time t_4 , rule r_4 is applicable, thus we derive $+\partial^{\tau}\neg B_a^{t_4}p$ and $-\partial^{\pi}B_a^{t_4}p$, which means that r_4 terminates B_ap . Notice that, even if r_4 is weaker than r_1 , the latter is not applicable at t_4 , thus it does not offer any support to maintain B_ap .

Proposition 1. Let D be a TDL theory. For every $\# \in \{\Delta, \partial\}$, $x, y \in \{\pi, \tau\}$:

- it is not the case that $D \vdash +\#^x X^l p$ and $D \vdash -\#^y X^l p$;
- if $D \vdash +\partial^x X^l p$ and $D \vdash +\partial^y l$, such that $l \in \mathcal{C}(X^l p)$, then $D \vdash +\Delta^x X^l p$ and $D \vdash +\Delta^y l$.

Proof sketch. The proof is a trivial variation of the ones for Theorems 1 and 2 in [10]: the properties of Proposition 1 are in fact a consequence of the use of the principle of strong negation for the definition of the proof conditions. \square

Proposition 1 shows the soundness of TDL: it is not possible to derive a tagged conclusion and its opposite, and that we cannot defeasibly prove both p and its complementary unless the definite (monotonic) part of the theory proves them; this means that inconsistency can be derived only if the theory we started with is inconsistent, and even in this case the logic does not collapse to the trivial extensions (i.e., everything is provable).

Definition 3. Let HB_D be the Herbrand Base for a TDL theory D . The extension of D is the 4-tuple $(\Delta^+, \Delta^-, \partial^+, \partial^-)$, where $\#^{\pm} = \{X^l p \mid Xp \in HB_D, D \vdash \pm\#^x X^l p, t \in \mathcal{T}\}$, $\# \in \{\Delta, \partial\}$, and $x \in \{\pi, \tau\}$.

Theorem 1. Given a TDL theory D , the extension of D can be computed in linear time, i.e., $O(|R| * |H_D| * |\mathcal{T}_D|)$, where \mathcal{T}_D is the set of distinct instants occurring in D .

Proof sketch. Theorem 1 follows from the result for the extension with modal operators of [9] (to handle the set of conflicting literals) and the result for the logic of [11] given in [29] to handle persistence of temporal literals. In particular, since the current logic does not devise any procedure for introducing modalities (they are derived insofar as they are already explicitly included in the head of rules) and temporal persistence is calculated focusing only on the temporal parameter labeling the modality in the outermost position, the computation of the theory extension falls within the case analyzed in [29] where the complexity is linear. \square

4. Reasoning about Speech Acts in FIPA Semantics

We now show how TDL formalism can express the semantics for ACLs. For ease of reference, let us focus on FIPA's *inform* and take [30] as a starting point. In FIPA, the meaning of communicative acts is defined in terms of rational effects (REs) and feasibility preconditions (FPs). The REs are the mental states the speaker wants to bring about in the hearer, and the FPs encode the appropriate conditions for issuing a communicative act. For instance, here is the FIPA definition of the *inform* communicative act:

$$\langle a, \text{inform}(b, p) \rangle \quad \begin{array}{l} \text{FP: } B(a, p) \wedge \neg B(a, B(b, p) \vee B(b, \neg p)) \\ \text{RE: } B(b, p) \end{array}$$

As a feasibility precondition speaker a must believe what he says and he must not believe that hearer b already has an opinion on the conveyed proposition. The rational effect is that hearer b comes to believe p .

Operators like this can be used to generate a dialogue directly, but they can also be used in the interpretation of the utterances of the interlocutor. In FIPA, this methodology relies on axioms ([30], Properties 4 and 5) according to which, when a communicative act is executed, its FPs are assumed to be true, and its RE is wanted by the speaker¹:

$$\begin{array}{l} B(a, \text{done}(\text{act})) \rightarrow FP(\text{act}) \\ B(a, \text{done}(\text{act}) \wedge \text{agent}(b, \text{act})) \rightarrow G(b, RE(\text{act})) \end{array}$$

FIPA and most other ACL semantics do not specify some important aspects of communicative acts, such as temporal persistence and non-monotonicity, also with reference to the role based semantics to deal with private vs public mental attitudes.

Persistence. In the case of *inform*, e.g., hearer b can infer not only that the precondition that a believes p is true at the moment of execution of the communicative act, b can also infer that this precondition held some time before the communicative act, and will hold afterwards. For instance, the content of an *inform* is supposed to persist, but a suggestion evaporates, when it is not taken on.

Non-monotonicity. An *inform* speech act has effects on the public beliefs of the speaker. However, the speaker can later retract its utterance, thus retracting its effects as well. This means that the utterance of an *inform* is done only in a defeasible way, so that it can be retracted later.

Roles. The rational effect of a speech act can be successful or not, which does not only depend on the speaker, but also on the hearer. ACL semantics like FIPA do not allow explicitly to make inferences about the success of a rational effect, since the mental attitudes that are modeled, are the private ones of the agents. To solve this, we adopt [26, 25, 2]'s *role-based* approach to agent communication, where mental attitudes are publicly attributed to dialogue participants, and can only change according to the rules of the dialogue. Public beliefs and goals represent the expected behavior of an agent, associated with a role in the dialogue. The public nature of the mental attitudes of roles solves the verifiability problem. Moreover, each participant can engage in different dialogues at the same time by playing different roles, and dialogues can obey different kinds of rules associated with each role. For example, in a persuasion dialogue, the proponent of a proposition can have a different 'burden of proof' than the opponent.

In the rules below, the beliefs and goals are therefore attributed to the roles (e.g., $B_{r(a)}$), and not only to the individual agents (e.g., B_a). Attitudes of individual agents can be unknown, or

¹In FIPA notation, act stands for any action, $\text{done}(act)$ is the proposition that expresses completion of act , and $\text{agent}(b, act)$ represents that b is the agent who executes action act .

can be different from the attitudes publicly attributed to their roles. Doing so, a lie is captured by deriving the conclusions $+ \partial B_{r(a)} \gamma$ (assuming that it is obtained as a FP of the inform act with respect to p) and $+ \partial B_a \delta$ where $\delta \in \mathcal{C}(\gamma)$.

We define different dialogue types and agent types [31], which can be used to model different types of dialogues. Concerning agent types, they regulate the way a hearer interprets the dialogue from the point of view of the interaction between public and private beliefs and goals of the speaker and the hearer. Speech acts have effects on the public mental attitudes of the speaker only. To infer that these effects (defeasibly) hold for the private mental attitudes of the speaker and/or the public or private beliefs and goals of the hearer further assumptions by the hearer are necessary:

- *Sincerity*: the speaker really believes or wants privately the beliefs or goals which can be publicly attributed to him on the basis of his utterances.
- *Reliability*: the public beliefs of the speaker become public beliefs of the hearer.
- *Trust*: the public beliefs of the hearer which have been adopted on the basis of the public beliefs of the speaker become the private beliefs of the hearer.
- *Cooperativity*: the public goals of the hearer which have been adopted on the basis of the public goals of the speaker become the private goals of the hearer.

Agent types are different combinations of these assumptions. Agent types are used in the interpretation of a speech act by a hearer, and in the planning phase of the speaker, who can foresee what will be the effects of his speech acts depending on the agent type the hearer attributes to him.

Different dialogue types are modelled as different sets of rules to interpret speech acts. As we show below, rules are used to derive what follows from a given speech act. However, not all the rules can be used at the same time, but only those which are appropriate to the situation. For example, we distinguish between information seeking dialogues, presupposing that the addressee of an inform does not have an opinion concerning what is said, from persuasive dialogues, where this assumption does not hold and the addressee is assumed to publicly believe an assertion if he does not challenge it.

These two mechanisms meet the second success guideline.

In the following we present the rules in a modular way, to satisfy the last success guideline of Section 1, distinguishing the following components:

1. Uttering and retracting,
2. Feasibility preconditions and rational effects,
3. Success conditions,
4. Bridge to private mental attitudes.

4.1. Inform

Rules $R_{inf} = \{i_1, i_2, i_3, \dots, i_{14}\}$ (see below) define the meaning of an *inform* communicative act, for a standard type of cooperative dialogue, like information exchange, where one agent is supposed to know more than the other. Adding also rule i_2 allows us to capture situations

where the knowledge is more symmetric. In what follows, a, b are agents, $r(a)$ and $r(b)$ the role-playing-agents in the dialogue, *inform* is a speech act type, s is either a literal or a modal literal, and $t < t'$ are time points in \mathcal{T} . Rules are prioritized as follows: $\succ_{inf} = \{i_2 \succ i_1, i_8 \succ i_7, i_{10} \succ i_9, i_{12} \succ i_{11}, i_{14} \succ i_{13}\}$.

Uttering and retracting an inform. Rule i_1 describes how an *inform* act is performed by an agent a through an utterance event. The rule is defeasible since the communicative act can be retracted later, as indicated in rule i_2 . There are different ways of handling retraction. The solution is to withdraw the original communicative act by means of a defeater. This means that all the consequences that can be inferred from the act, expressed in rules $i_3 - i_6$, are also withdrawn. If we would take the alternative solution of only retracting the content of the *inform*, we would need additional explicit rules to withdraw those consequences too.

$$\begin{aligned} i_1 \quad & utter_{a,b}(inform_{r(a),r(b)}(s,t),t) \Rightarrow_{\tau} inform_{r(a),r(b)}(s,t) \\ i_2 \quad & retract_{a,b}(inform_{r(a),r(b)}(s,t),t') \rightsquigarrow_{\tau} \neg inform_{r(a),r(b)}(s,t) \end{aligned}$$

The agent of actions *utter* and *retract* is the individual agent a and not his role $r(a)$. The rules i_1 and i_2 are used to connect individual agents to their roles. These rules are non-persistent, since the action of uttering only temporally coincides with execution of an *inform* communicative act. If the agent makes an *inform* at a given time, it is not possible to infer that it is making the *inform* again at the next time instant.

The defeasible character of i_1 is necessary in applications where an agent is trying to persuade the other one and the other one can reply (see item 1 in Section 2), or where two agents are trying to cooperatively find a solution and to advance arguments which are probed by the other one. In these types of dialogues the speaker can retract his assertions to avoid a contradiction. In information gathering scenarios like the ones considered by FIPA this is not necessary, since an agent is supposed to know more than the other.

Feasibility preconditions and rational effects. Rules $i_3 - i_5$ represent the FPs of *inform*. Following Properties 4 and 5 of [30], they are interpreted as strict rules. Only rule i_3 is persistent towards the future, since its effect is not affected by the consequent of other rules. Instead, since the *inform* possibly changes the beliefs of the hearer b , $i_4 - i_5$ are not persistent and they refer to the situation before the execution of the speech acts:

$$\begin{aligned} i_3 \quad & inform_{r(a),r(b)}(s,t) \rightarrow_{\pi} B_{r(a)}^t s \\ i_4 \quad & inform_{r(a),r(b)}(s,t) \rightarrow_{\tau} \neg B_{r(b)}^{t-1} s \\ i_5 \quad & inform_{r(a),r(b)}(s,t) \rightarrow_{\tau} \neg B_{r(b)}^{t-1} l \end{aligned}$$

where $l \in \mathcal{C}(s)$. Note that the beliefs in the consequent of the rules are not attributed to the agents as private beliefs, but to the roles they play: thus they have a public character².

Rule i_6 represents the RE of *inform*: its propositional content is embedded in a goal of the speaker that the hearer believes it:

$$i_6 \quad inform_{r(a),r(b)}(s,t) \rightarrow_{\pi} G_{r(a)}^t B_{r(b)}^t s$$

²The conclusions of FP rules could also persist from the past, since they are observations and are not caused by the speech acts. For example, we may infer that the speaker believed s in the past, and, unless other information are available, that this belief holds by persistence at t as well. This inference is non-monotonic. In addition, the past time from when this belief starts to hold cannot be determined by the time of *inform*, but comes from additional evidence, such as the time of another public belief of the speaker related to s . Facts like this can be easily embedded within our framework by adding further suitable rules expressing this evidence.

Success conditions. FIPA does not allow explicit inferences about the success of the RE, but in our model of cooperative information exchange R_{inf} , rule i_7 can represent that the hearer publicly adopts the information conveyed, if he does not believe that the speaker is unreliable:

$$\begin{aligned} i_7 & G_{r(a)}^t B_{r(b)}^t s \Rightarrow \pi B_{r(b)}^t s \\ i_8 & G_{r(a)}^t B_{r(b)}^t s, B_{r(b)}^t \neg \text{reliable}(a) \rightsquigarrow \tau \neg B_{r(b)}^t s \end{aligned}$$

Bridge to private mental attitudes. It is not necessary that the hearer privately believes what was said. Only if there is no evidence to the contrary, we assume that individual agents believe what their roles believe ($i_9 - i_{12}$). Rule i_9 assumes that the speaker individually believes what he says, unless he is believed to be insincere (i_{10}). Rule i_{11} assumes sincerity for goals in a similar way. Rule i_{13} assumes that a hearer believes what has been asserted, unless he is believed not to be a trusting character (i_{14}).

In all these cases, the cooperative behavior is the default, but it can be overruled by evidence to the contrary. Hence we have $i_{10} \succ i_9$, $i_{12} \succ i_{11}$ and $i_{14} \succ i_{13}$.

$$\begin{aligned} i_9 & \text{inform}_{r(a),r(b)}(s,t) \Rightarrow \pi B_a^t s \\ i_{10} & \text{inform}_{r(a),r(b)}(s,t), B_{r(b)}^t \neg \text{sincere}(a) \rightsquigarrow \tau \neg B_a^t s \\ i_{11} & \text{inform}_{r(a),r(b)}(s,t), G_{r(a)}^t B_{r(b)}^t s \Rightarrow \pi G_a^t B_b^t s \\ i_{12} & \text{inform}_{r(a),r(b)}(s,t), B_{r(b)}^t \neg \text{sincere}(a) \rightsquigarrow \tau \neg G_a^t B_b^t s \\ i_{13} & \text{inform}_{r(a),r(b)}(s,t), B_{r(b)}^t s \Rightarrow \pi B_b^t s \\ i_{14} & \text{inform}_{r(a),r(b)}(s,t), B_{r(b)}^t s, B_{r(a)}^t \neg \text{trusting}(b) \rightsquigarrow \tau \neg B_b^t s \end{aligned}$$

Strict inferences about REs and FPs of communicative acts are only possible regarding the public beliefs or goals of the participants, while inferences about the private mental states of the participants can be made only by default. Sincerity, trust and cooperativity are the assumptions to pass information from the public level to the private one.

In the following we present some examples of reasoning about time and persistence, to illustrate the mechanism before considering the rules of dialogue.

Example 3 (Time and persistence; see item 4 in Section 2). *Consider the agents i (Ian), j (John) and k (Kay): their the roles are $v(i)$ (vendor), $c(j)$ (client) and $w(k)$ (witness). Literal s means “the chair is without faults”. The theory contains R_{inf} , \succ_{inf} , plus the following facts and additional rules:*

$$\begin{aligned} \text{Facts: } & \text{utter}_{i,k}(\text{inform}_{v(i),w(k)}(\neg s, 1), 1), \text{sell}(i, j, 2), \\ & B_{c(j)}^3(\text{reliable}(k)), \text{utter}_{k,j}(\text{inform}_{w(k),c(j)}(\neg s, 3), 3) \end{aligned}$$

$$\begin{aligned} \text{Rules: } & l_1 : \text{sell}(v(i), c(j), 2) \Rightarrow \tau \text{inform}_{v(i),c(j)}(s, 2) \\ & l_2 : \text{sell}(v(i), c(j), 2) \Rightarrow \pi B_{c(j)}^2(\text{reliable}(i)) \\ & l_3 : \Rightarrow \pi B_j^2 \neg \text{reliable}(i) \\ & l_4 : B_j^2 \neg \text{reliable}(i) \Rightarrow \pi B_{c(j)}^2 \neg \text{reliable}(i) \end{aligned}$$

Priorities: $l_2 \succ l_4$.

l_1 states that when $v(j)$ sells the chair at 2, he acts as if he were informing $c(j)$ that the chair is without faults; l_3 assumes that j normally believes that i is unreliable and then, via rule l_4 , that j as a client believes so. However, j believes that i is reliable when he is a vendor selling items: this overrides the fact that in general j believes the opposite (rule l_2 is stronger than l_4).

Time	Conclusion	Argument
1	1. $+\partial^\tau \text{inform}_{v(i),w(k)}(\neg s, 1)$	(a1): Facts, i_1
1	2. $+\partial^\pi B_{v(i)}^1 \neg s$	(a2): 1, i_3
2	3. $+\partial^\tau \text{inform}_{v(i),c(j)}(s, 2)$	(a3): Facts, l_1
2	4. $+\partial^\pi B_j^2 \neg \text{reliable}(i)$	(a4): l_3
2	5. $+\partial^\pi B_{c(j)}^2 (\text{reliable}(i))$	(a5): Facts, l_2
2	6. $-\partial^\pi B_{c(j)}^2 \neg \text{reliable}(i)$	(a6): conflict between l_2 and l_4
2	7. $+\partial^\pi G_{v(i)}^2 B_{c(j)}^2 s$	(a7): 3, i_6
2	8. $+\partial^\pi B_{c(j)}^2 s$	(a8): 7, i_7
2	9. $-\partial^\pi B_{v(i)}^2 \neg s$	(a9): conflict between (a2) and 3, i_3
2	10. $-\partial^\pi B_{v(i)}^2 s$	(a10): conflict between (a2) and 3, i_3
3	11. $+\partial^\tau \text{inform}_{w(k),c(j)}(\neg s, 3)$	(a11): Facts, i_1
3	12. $+\partial^\pi G_{w(k)}^3 B_{c(j)}^3 \neg s$	(a12): 11, i_6
3	13. $-\partial^\pi B_{c(j)}^3 \neg s$	(a13): conflict between (a6) and Facts, 12, i_7
3	14. $-\partial^\pi B_{c(j)}^3 s$	(a14): conflict between (a6) and Facts, 12, i_7

Table 1: The relevant derivations of Example 3

Consider the relevant derivations in Table 1. For conclusions 7-8 and 11-12, due to the persistence of $B_{v(i)}^1 \neg s$ and $B_{c(j)}^2 s$, these beliefs should also hold, respectively, at 2 and 3. But they are defeated by opposite arguments, which are in turn defeated by the former ones. (TDL is a sceptical non-monotonic formalism: with two conflicting defeasible conclusions DL refrains to take a decision.) 7-8 show that FPs of some inform acts are violated, 11-12 that some REs are not successful.

Example 4 (Multiple inform). Consider the situation where the agent a , in his role, informs at time 1 agent b in her role about the fact s . Later, at time 3, agent a , again in his role, informs another agent c in her role about the same fact s . Assume, however, that $r(c)$ believes that a is not sincere. An interesting aspect of this scenario can be modeled by the following theory, which contains two instances of the same rules i_9 and i_{10} ³:

$$\begin{aligned}
(F &= \{\text{inform}_{r(a),r(b)}(s, 1), \text{inform}_{r(a),r(c)}(s, 3), B_{r(c)}^3 \neg \text{sincere}(a)\}, \\
R &= \{i_9' : \text{inform}_{r(a),r(b)}(s, 1) \Rightarrow_\pi B_a^1 s, \\
&\quad i_{10}' : \text{inform}_{r(a),r(b)}(s, 1), B_{r(b)}^1 \neg \text{sincere}(a) \rightsquigarrow_\tau \neg B_a^1 s, \\
&\quad i_9'' : \text{inform}_{r(a),r(c)}(s, 3) \Rightarrow_\pi B_a^3 s, \\
&\quad i_{10}'' : \text{inform}_{r(a),r(c)}(s, 3), B_{r(b)}^3 \neg \text{sincere}(a) \rightsquigarrow_\tau \neg B_a^3 s\}, \\
> &= \{i_{10}' > i_9', i_{10}'' > i_9'', i_{10}' > i_9'', i_{10}'' > i_9'\}
\end{aligned}$$

³For brevity's sake, let us omit utterances involving rule i_1 and assume to have the inform speech acts as facts.

The first speech act triggers i'_9 , which is not blocked by i'_{10} since this last rule is not applicable. Hence, we get $+\partial^\pi B_a^1 s$ and, by persistence, also $+\partial^\pi B_a^2 s$. At time 3, rule i'_9 is triggered together with i'_{10} : since all instances of i_{10} are stronger than all instances of i_9 , we terminate the persistence of $B_a s$: hence, we have no longer sufficient reasons to say that a believes that s is the case, and so we get $-\partial^\pi B_a^3 s$. This should not look strange: c 's belief that a is not sincere is a publicly attributed belief (the belief regards $r(c)$) and, in presence of this subsequent public evidence, we have reasons to undermine the previous inference: before time 3 there were good reasons to think that a believed that s (there was no evidence to the contrary), but this does not hold from 3 onwards. Of course, different logical arrangements can be made, such as removing the general priority between i_9 and i_{10} (which affects all rule instances) and adjust things case by case and in each concrete context, in such a way that different instances of the pair i_9 - i_{10} may behave differently. For example, if we think that $r(c)$'s beliefs are not reliable, we could simply change \succ and rather have $i'_9 \succ i'_{10}$, with which we would still obtain $+\partial^\pi B_a^3 s$.

Example 5 (What if analysis and agent types). Assume agent i wants that j knows that s , not just as a public admission, but that j really believe s privately. To achieve this goal i plans an inform speech act, in the role s (speaker) addressed to receiver j . However, to understand whether he has possibilities to achieve it, i must consider also the agent type that j attributes to him. The agent type is $AT_{s(j),r(i)} = \{\text{reliable}(i), \text{sincere}(i), \text{trusting}(j)\}$. From this agent type, i can infer that j will accept privately that s .

Facts: $\text{utter}_{i,j}(\text{inform}_{s(i),r(j)}(s, 1), 1)$,
 $B_{r(j)}^1(\text{reliable}(i)), B_{r(j)}^1(\text{sincere}(i)), B_{s(i)}^1(\text{trusting}(j))$.

Suppose, however, that, similarly to what we had in Example 3, the following rules hold:

Rules: $x_1 : \Rightarrow_\pi B_j^1 \neg \text{reliable}(i)$
 $x_2 : B_j^1 \neg \text{reliable}(i) \Rightarrow_\pi B_{c(j)}^1 \neg \text{reliable}(i)$
 $x_3 : \Rightarrow_\pi B_j^1 \neg \text{sincere}(i)$
 $x_4 : B_j^1 \neg \text{sincere}(i) \Rightarrow_\pi B_{c(j)}^1 \neg \text{sincere}(i)$
 $x_5 : \Rightarrow_\pi B_i^1 \neg \text{trusting}(j)$
 $x_6 : B_i^1 \neg \text{trusting}(j) \Rightarrow_\pi B_{s(i)}^1 \neg \text{trusting}(j)$

Notice that the facts $B_{r(j)}^1(\text{reliable}(i)), B_{r(j)}^1(\text{sincere}(i)), B_{s(i)}^1(\text{trusting}(j))$ guarantee that i_{10} , i_{12} , and i_{14} are inapplicable, thus allowing us to use i_9 , i_{11} , and i_{13} and obtain i 's and j 's private beliefs that s is the case: consider some relevant derivations in Table 2 (let us omit there the failures in triggering i_{10} , i_{12} , and i_{14} and focus on the positive conclusions).

Example 6 (Example 5 reframed). We continue the previous example and reframe it a bit by adopting the point of view of the addressee j . Assume now that j privately and non-defeasibly believes that $\neg s$, thus rule i_{13} is defeated. At this point j decides to inform i about his beliefs. Since i trusts more j on s than his belief, he adopts s privately and retracts at time 3 his assertion.

Facts: $\text{utter}_{j,i}(\text{inform}_{r(j),s(i)}(\neg s, 2), 2)$
 $\text{retract}_{s(i),r(j)}(\text{inform}_{s(i),r(j)}(s, 1), 3)$,
 $B_{s(i)}^2(\text{reliable}(j)), B_{s(i)}^2(\text{sincere}(j)), B_{s(i)}^2(\text{trusting}(i))$

Time	Conclusion	Argument
1	1. $+\partial^\tau \text{inform}_{s(i),r(j)}(s, 1)$	(a1): Facts, i_1
1	2. $+\partial^\pi B_{s(i)}^1 s$	(a2): 1, i_3
1	3. $+\partial^\pi B_{r(j)}^1(\text{reliable}(i))$	(a3): Facts
1	4. $+\partial^\pi G_{s(i)}^1 B_{r(j)}^1 s$	(a4): 1, i_6
1	5. $+\partial^\pi B_{r(j)}^1 s$	(a5): 4, i_7
1	6. $+\partial^\pi B_i^1 s$	(a6): 1, i_9
1	7. $+\partial^\pi G_i^1 B_j^1 s$	(a7): 1, 4, i_{11}
1	8. $+\partial^\pi B_j^1 s$	(a8): 1, 5, i_{13}

Table 2: The relevant derivations of Example 5

Rules: $l_3 := \Rightarrow \pi B_i^1 s$
 $l_4 := \rightarrow \pi B_j^1 \neg s$

Consider the relevant derivations in Table 3.

4.2. Request

Analogously to *inform*, we define the preconditions and effects of a *request* communicative act, used in deliberation dialogues, by rules $R_{req} = \{r_1, \dots, r_{13}\}$, with priority $r_2 \succ r_1$, $r_7 \succ r_6$, $r_9 \succ r_8$, $r_{11} \succ r_{10}$ and $r_{13} \succ r_{12}$. For space reasons, we do not explicitly model preconditions of actions, so compared to FIPA, we have to simplify the definitions. Again, the cooperative behavior is the default, which is overruled when the agent refuses the *request*, or when there is evidence that the agent is insincere, or non-cooperative.

Uttering and retracting request. As in the case of *inform*, a *request* after being uttered can be retracted by the speaker.

- r_1 $\text{utter}_{a,b}(\text{request}(s,t), t) \Rightarrow \tau \text{request}_{r(a),r(b)}(s,t)$
- r_2 $\text{retract}_{a,b}(\text{request}_{r(a),r(b)}(s,t), t') \rightsquigarrow \tau \neg \text{request}_{r(a),r(b)}(s,t)$

Feasibility preconditions and rational effect. The precondition of a *request* is that the speaker does not believe in advance that the hearer wants already to perform the requested action. The rational effect is that the speaker wants to achieve s and that the hearer achieves it.

- r_3 $\text{request}_{r(a),r(b)}(s,t) \rightarrow \pi G_{r(a)}^t s$
- r_4 $\text{request}_{r(a),r(b)}(s,t) \rightarrow \tau \neg B_{r(a)}^{t-1} G_{r(b)}^t s$
- r_5 $\text{request}_{r(a),r(b)}(s,t) \rightarrow \pi G_{r(a)}^t G_{r(b)}^t s$

Success conditions. If the hearer does not explicitly refuse the *request* he is assumed to publicly adopt the requested goal (but not necessarily privately). Again this is a defeasible assumption which can be retracted in front of an explicit refusal.

- r_6 $G_{r(a)}^t G_{r(b)}^t s \Rightarrow \pi G_{r(b)}^{t,\pi} s$
- r_7 $\text{request}_{r(a),r(b)}(s,t), \text{refuse}_{r(b),r(a)}(s,t') \rightsquigarrow \tau \neg G_{r(b)}^{t,\pi} s$

Time	Conclusion	Argument
2	1. $+\partial^\pi B_j^2 \neg s$	(a1): l_4 , persistence
2	2. $-\partial^\pi B_j^2 s$	(a2): conflict between l_4 and i_{13}
2	3. $+\partial^\tau \text{inform}_{r(j),s(i)}(\neg s, 2)$	(a3): Facts, i_1
2	4. $+\partial^\pi B_{r(j)}^2 \neg s$	(a4): 3, i_3
2	5. $+\partial^\pi B_{s(i)}^2(\text{reliable}(j))$	(a5): Facts
2	6. $+\partial^\pi G_{r(j)}^2 B_{s(i)}^2 s$	(a6): 3, i_6
2	7. $+\partial^\pi B_{r(j)}^2 s$	(a7): 6, i_7
2	8. $+\partial^\pi B_i^2 \neg s$	(a8): l_3 , persistence
3	9. $+\partial^\pi \text{retract}_{s(i),r(j)}(\text{inform}_{s(i),r(j)}(s, 1), 3)$	(a9): Facts
3	10. $-\partial^\tau \text{inform}_{s(i),r(j)}(s, 1)$	(a10): conflict between i_1 and i_2
3	11. $-\partial^\pi B_{s(i)}^1 s$	(a11): conflict between i_1 , i_3 and i_2
3	12. $-\partial^\pi G_{s(i)}^1 B_{r(j)}^2 s$	(a12): conflict between i_1 , i_6 and i_3
3	13. $-\partial^\pi B_{s(i)}^1 s$	(a13): conflict between i_1 , i_7 and i_2
3	14. $-\partial^\pi B_i^1 s$	(a14): conflict between i_1 , i_9 and i_2

Table 3: The relevant derivations of Example 6

Bridge to private mental attitudes. Depending on the agent type attributed by the hearer to the speaker different inferences can be made about public and private goals of both of them. If the speaker is believed to be sincere, then he is held to want also privately what he publicly requested. If the hearer is cooperative he will also adopt privately the goal of the requestee.

$$\begin{aligned}
r_8 \quad & \text{request}_{r(a),r(b)}(s,t) \Rightarrow_\tau G_a^t s \\
r_9 \quad & \text{request}_{r(a),r(b)}(s,t), B_{r(b)}^t \neg \text{sincere}(a) \rightsquigarrow_\tau \neg G_a^t s \\
r_{10} \quad & \text{request}_{r(a),r(b)}(s,t), G_{r(b)}^t s \Rightarrow_\pi G_b^t s \\
r_{11} \quad & \text{request}_{r(a),r(b)}(s,t), G_{r(b)}^t s, B_{r(a)}^t \neg \text{cooperative}(b) \rightsquigarrow_\tau \neg G_b^t s \\
r_{12} \quad & \text{request}_{r(a),r(b)}(s,t), G_{r(a)}^t G_{r(b)}^t s \Rightarrow_\pi G_a^t G_b^t s \\
r_{13} \quad & \text{request}_{r(a),r(b)}(s,t), B_{r(b)}^t \neg \text{sincere}(a) \rightsquigarrow_\tau \neg G_a^t G_b^t s
\end{aligned}$$

In this paper we do not discuss the limited persistence of effects of requests, as envisaged in scenario 6 which are not accepted by a hearer.

4.3. Abstract communicative acts: *inform-if*

Like in FIPA, we can also handle abstract communicative acts, like *inform-if* $_{a,b}(s,t)$ which is composed of the nondeterministic choice of $\text{inform}_{a,b}(s,t)$ and $\text{inform}_{a,b}(\neg s,t)$. Note that in FIPA *inform-if* is an abstract action which cannot directly be executed:

$$\begin{aligned}
ii_1 \quad & \text{inform}_{r(a),r(b)}(s,t) \rightarrow_\tau \text{inform-if}_{r(a),r(b)}(s,t) \\
ii_2 \quad & \text{inform}_{r(a),r(b)}(\neg s,t) \rightarrow_\tau \text{inform-if}_{r(a),r(b)}(s,t)
\end{aligned}$$

Thus, we can define *query-if* as a *request* to *inform-if*:

$$qi_1 \quad \text{queryif}_{r(a),r(b)}(s,t) \rightarrow_\tau \text{request}_{r(a),r(b)}(\text{inform-if}_{r(b),r(s)}(s,t'),t)$$

To satisfy the *request*, the receiver has to execute either an $\text{inform}_{r(b),r(s)}(s,t')$ or an $\text{inform}_{r(b),r(s)}(\neg s,t')$.

4.4. Persuasion dialogues

ACLs are usually studied in relation to specific *dialogue types*, such as cooperative information exchange, negotiation or persuasion (see, e.g., [32]). We can extend the previous rules with persuasion, by defining acts like *challenge* and *concede*, besides the *retract* discussed above. A single communicative act like *inform* may have different semantics, in different types of dialogue. This is due to different background assumptions, for example regarding sincerity, cooperativity, or trust. Thus, in non-cooperative dialogue types like persuasion or negotiation (see Example 1), it is possible to reverse the general principle of rules $i_6 - i_{14}$, that cooperative behavior is expected by default, but can be overruled by evidence to the contrary. Alternatively, we can follow the principle that “silence means consent”. In a persuasion dialogue, the hearer is assumed to believe what the speaker said (rule i'_7 below), unless he explicitly challenges the proposition (rule i''_7 below), thus defeating the conclusion that he believes the content of the *inform*.

$$\begin{aligned} i'_7 & \quad G_{r(a)}^t B_{r(b)}^t s \Rightarrow \pi B_{r(b)}^t s \\ i''_7 & \quad \text{inform}_{r(a),r(b)}(s,t), \text{challenge}_{r(b),r(a)}(s,t') \rightsquigarrow_{\tau} \neg B_{r(b)}^t s \end{aligned}$$

In addition to challenges, we can add explicit concessions [32]. If an agent concedes to s , it does not necessarily mean that he now believes s , but that he no longer believes the opposite. For example, the concession blocks the agent from performing an *inform* that $\neg s$ later in the interaction:

$$c_1 \quad \text{concede}_{r(a),r(b)}(s,t) \rightarrow \pi \neg B_{r(a)}^t \neg s$$

So for persuasion, the rules R_{inf} are altered as follows:

$$R_{persuasion} = (R_{inf} \setminus \{i_7, i_9, i_{10}, i_{11}, i_{12}, i_{13}, i_{14}\}) \cup \{i_2, i'_7, i''_7, c_1\}$$

where $i''_7 \succ i'_7$. Rules bridging public and private mental attitudes are removed, since from a debate it does not follow that the interactants believe what they were not able to defend with a counterargument.

5. Related work

Other papers went in the same direction of redefining FIPA semantics: e.g., [26, 25, 2, 22, 24, 20]. Like us, most of them distinguish between public and private mental attitudes. There are various differences.

Like in [26, 25, 2], the distinguishing feature of our approach is that the public mental attitudes attributed to agents during the dialogue are associated with roles. However, we use roles to redefine the FIPA semantics in a non-monotonic framework based on TDL which allows us to extend FIPA to persuasion dialogue. We distinguish interactive roles, such as speaker, (over)hearer and addressee. Clearly, different constitutive rules apply to speaker and hearer. Further, we could add rules so that the effects of implicit acknowledgement differ between the addressee of a message, and a mere overhearer [23]. Because social roles are associated with dialogue types, with specific sets of dialogue rules, roles allow us to reason about assumptions in different kinds of dialogues. E.g., sincerity could be assumed in cooperative dialogues, such as information exchange, but not in non-cooperative dialogues, such as persuasion or negotiation. Ostensible beliefs and the grounding operator distinguish only interactive roles or different groups of agents.

The importance of roles is recognized in multiagent systems and their function ranges from attributing responsibilities to assigning powers to agents in organizations. Other solutions, instead, need to add to dialogue new theoretical concepts which are not always completely clear

or diverge from existing work. In particular, [22] use an explicit grounding operator, which only partially overlaps with the tradition of grounding in theories of natural language dialogue. Opinions [24] are introduced specifically for modelling dialogue, but with no relation with persuasion and argumentation. Finally, commitments in [20] overlap with obligations.

Moreover, the approaches relate to the well known FIPA semantics in different degrees: [23] and [24] try to stay close to the original semantics, as we do, while [20] substitute it entirely with a new semantics, which, among other things, does not consider preconditions of actions. [22] and [24] use modal logic, and [20] use CTL. They thus use more common frameworks, but they do not consider the computational properties of their proposals. Moreover, most of the other approaches do not consider the persistence of preconditions and effects, an essential point when dealing with actions and their effects. Time introduces most of the complexities in our formal system, but time is crucial for agent communication, because speech acts are uttered one after the other, but their effects on mental attitudes are persistent.

6. Summary

We used TDL to study non-monotonicity and time in role-based agent communication. Non-monotonicity occurs in reasoning about the persistence of FPs and REs of speech acts. Whereas FIPA makes strong assumptions about the private states, the alternative of using public mental attitudes does not make any assumptions about them; using non-monotonic reasoning we can make inferences about the private mental attitudes of the agents which hold only by default and can always be revised. Finally, non-monotonicity can be used for challenges, concessions, and retractions. E.g., an inform is accepted –its content becomes part of the public beliefs of the addressee– unless it is challenged.

Future work include passing from what-if analysis to planning, and increasing the expressiveness of the framework, for example to model nested rules to cope with speech acts like proposals which in FIPA is modelled as an *inform* with a conditional intention as content.

References

- [1] FIPA, FIPA ACL message structure specification, Tech. Rep. XC00061, Foundation for Intelligent Physical Agents (2001).
- [2] G. Boella, R. Damiano, J. Hulstijn, L. van der Torre, A common ontology of agent communication languages: Modelling mental attitudes and social commitments using roles, *Applied Ontology* 2 (3-4) (2007) 217–265.
- [3] H. van Ditmarsch, W. van der Hoek, B. Kooi, *Dynamic Epistemic Logic*, Springer, 2007.
- [4] G. Antoniou, D. Billington, G. Governatori, M. J. Maher, Representation results for defeasible logic, *ACM Transactions on Computational Logic* 2 (2) (2001) 255–287.
- [5] M. J. Maher, Propositional defeasible logic has linear complexity, *Theory and Practice of Logic Programming* 1 (6) (2001) 691–711.
- [6] N. Bassiliades, G. Antoniou, I. Vlahavas, A defeasible logic reasoner for the Semantic Web, *International Journal on Semantic Web and Information Systems* 2 (2006) 1–41.
- [7] E. Kontopoulos, N. Bassiliades, G. Governatori, G. Antoniou, A modal defeasible reasoner of deontic logic for the semantic web, *International Journal of Semantic Web and Information Systems* 7 (1) (2011) 18–43.
- [8] H.-P. Lam, G. Governatori, The making of SPINdle, in: Paschke et al. [33], pp. 315–322.
- [9] G. Governatori, A. Rotolo, Bio logical agents: Norms, beliefs, intentions in defeasible logic, *Journal of Autonomous Agents and Multi Agent Systems* 17 (2008) 36–69.
- [10] G. Governatori, A. Rotolo, A computational framework for institutional agency, *Artificial Intelligence and Law* 16 (1) (2008) 25–52.
- [11] G. Governatori, A. Rotolo, G. Sartor, Temporalised normative positions in defeasible logic, in: ICAIL’05, ACM Press, 2005, pp. 25–34.

- [12] G. Governatori, M. Palmirani, R. Riveret, A. Rotolo, G. Sartor, Norm modifications in defeasible logic, in: M.-F. Moens, P. Spyns (Eds.), JURIX 2005: The Eighteenth Annual Conference, IOS Press, 2005, pp. 13–22.
- [13] G. Governatori, A. Rotolo, R. Riveret, M. Palmirani, G. Sartor, Variations of temporal defeasible logic for modelling norm modifications, in: Proceedings of 11th International Conference on Artificial Intelligence and Law, ACM Press, New York, 2007, pp. 155–159.
- [14] G. Governatori, A. Rotolo, Changing legal systems: Legal abrogations and annulments in defeasible logic, *The Logic Journal of IGPL* 18 (1) (2010) 157–194.
- [15] G. Governatori, J. Hulstijn, R. Riveret, A. Rotolo, Characterising deadlines in temporal modal defeasible logic, in: M. A. Orgun, J. Thornton (Eds.), 20th Australian Joint Conference on Artificial Intelligence, AI 2007, Springer, Berlin, 2007, pp. 486–496.
- [16] R. Kowalski, M. Sergot, A logic-based calculus of events, *New Gen. Comput.* 4 (1986) 67–95. doi:10.1007/BF03037383.
- [17] R. Rubino, Una implementazione della logica defeasibile temporale per il ragionamento giuridico, Ph.D. thesis, CIRSFD, University of Bologna (2009).
- [18] R. Rubino, A. Rotolo, A Java implementation of temporal defeasible logic, in: Governatori et al. [33], pp. 297–304.
- [19] M. J. Wooldridge, Semantic issues in the verification of agent communication languages, *JAAMAS* 3 (1) (2000) 9–31.
- [20] M. Verdicchio, M. Colombetti, From message exchanges to communicative acts to commitments, in: Procs. LCMAS'05, 2006.
- [21] M. P. Singh, A social semantics for agent communication languages, in: Procs. IJCAI Workshop on ACLs, 2000.
- [22] B. Gaudou, A. Herzig, D. Longin, A logical framework for grounding-based dialogue analysis, in: Procs. of LCMAS'05, 2006.
- [23] B. Gaudou, A. Herzig, D. Longin, M. Nickles, A new semantics for the FIPA agent communication language based on social attitudes, in: Procs. ECAI'06, 2006.
- [24] M. Nickles, F. Fischer, G. Weiss, Communication attitudes: A formal approach to ostensible intentions, and individual and group opinions, in: Procs. LCMAS'05, 2006.
- [25] G. Boella, R. Damiano, J. Hulstijn, L. van der Torre, Role-based semantics for agent communication: Embedding of the mental attitudes and social commitments semantics, in: Procs. AAMAS'06, 2006, pp. 688–690.
- [26] G. Boella, R. Damiano, J. Hulstijn, L. van der Torre, Public mental attitudes for ACL semantics, in: Agent Communication II, no. 3859 in LNCS, Springer, 2006, pp. 30–44.
- [27] N. Fornara, M. Colombetti, A commitment-based approach to agent communication, *Applied Artificial Intelligence* 18 (9-10) (2004) 853–866.
- [28] G. Antoniou, D. Billington, G. Governatori, M. J. Maher, Embedding defeasible logic into logic programming, *Theory and Practice of Logic Programming* 6 (6) (2006) 703–735.
- [29] G. Governatori, A. Rotolo, On the complexity of temporal defeasible logic, in: Proceedings NMR 2010, 2010.
- [30] FIPA, FIPA ACL communicative act library specification, Tech. Rep. SC00037J, FIPA (2002).
- [31] J. Broersen, M. Dastani, J. Hulstijn, L. van der Torre, Goal generation in the BOLD architecture, *Cognitive Science Quarterly* 2(3-4) (2002) 428–447.
- [32] D. Walton, E. Krabbe, *Commitment in Dialogue: Basic Concepts of Interpersonal Reasoning*, 1995.
- [33] G. Governatori, J. Hall, A. Paschke (Eds.), Rule Interchange and Applications, International Symposium, RuleML 2009, Las Vegas, Nevada, USA, November 5-7, 2009. Proceedings, Vol. 5858 of Lecture Notes in Computer Science, Springer, 2009.