# A Timed Epistemic Logic for Formalizing Cooperation among Groups of Agents

Stefania Costantini[1], Andrea Formisano[2,**] and Valentina Pitoni[1]

[1]DISIM, Università di L'Aquila, Italy
[2]DMIF, Università di Udine, Italy

**Abstract**
In the multi-agent setting, it is relevant to model group dynamics of agents, and logic has proved to be an excellent tool. We have proposed in previous work an epistemic logic that allows one to formalize the new beliefs formed or removed by a group of agents, where several groups can co-exist and where an agent can pass from one group to another. A novelty introduced in this paper is that an agent can be lent by a (willing) group to another one in case of need. Another distinguished feature we introduce in this paper is time and temporal instants/intervals to express the time periods in which agents' beliefs hold.

**Keywords**
Multi-Agent Systems, Modal Logic, Epistemic Logic

## 1. Introduction

In the research we are reporting in this paper, we are interested in the fruitful application of Multi-Agent Systems (MAS) not only in purely technological applications but also in human/social sciences. There are applications where problem-solving in MAS can profit from the ability to represent group dynamics, understand the behavior of other agents, and perform reasoning by impersonating another agent. To this aim an agent needs to be able to represent aspects of *"Theory of Mind"* [1]. This can be described as the set of social-cognitive skills involving the ability to attribute mental states (such as desires, beliefs, knowledge) to oneself and to other agents, and, moreover, the capability to reason about such mental states. Consequently, agents with these abilities can make predictions and formulate interpretations of other agents' behaviors. However, in real-world settings, agents are situated in time and their beliefs, desires, and goals are timed, in the sense that they hold or make sense in a certain time interval. This is an important feature, usually missing from other approaches in that it is not easy to model and axiomatize, which we introduce in this paper.

We build on our previous work, where we introduced a logical framework based on an epistemic logic called *L-DINF*, which allows the representation of knowledge and beliefs of

*WOA 2023: 24th Workshop From Objects to Agents, November 6–8, Rome, Italy*
** Corresponding author.
The authors contributed equally.
stefania.costantini@univaq.it (S. Costantini); andrea.formisano@uniud.it (A. Formisano); valentina.pitoni@univaq.it (V. Pitoni)
0000-0002-5686-6124 (S. Costantini); 0000-0002-6755-9314 (A. Formisano); 0000-0002-4245-4073 (V. Pitoni)

agents, in order to reason about the mental actions of agents, taking into account the costs of actions and budgets that are available to afford such costs. The logical framework *L-DINF* allows the modeling of group dynamics of cooperative agents: one can model agents that can form groups and agents of the same group can support each other in performing collective actions. In addition, in *L-DINF* agents can take into consideration the preferences that each agent may have for what concerns performing each action [2, 3], preferences that the group will take into account when devising and executing plans.

The logical framework also encompasses the possibility for agents to have *roles* within their group of agents. Roles determine which actions each agent is enabled by its group to perform. The action is considered executable if at least one agent of the group can perform the action, with the approval of the group and on behalf of the group. An agent can join or leave a group whenever it wants, and, clearly, the role of an agent may change as it joins another group.

The adequacy of the framework in modeling aspects of the Theory of Mind is discussed in detail in [4], to which we also refer the interested reader for an extensive discussion on related works (which we cannot include here due to space limitations)

Initial attempts to enrich *L-DINF* by introducing a notion of time have been explored in [5, 6]. In this paper, we go further in this direction and propose an extension of *L-DINF* in which the specification of each belief of an agent includes the interval of time in which the belief holds. The *temporalization* of *L-DINF* is obtained by defining a “time module” suitable to add time in an easy way into logic representations of agents. Thus, our proposal finds potential applicability beyond our specific approach. In essence, this module is a particular kind of function (cf., function $T$ in Sect. 2.1), that assigns a “timing" to atoms and formulas.

We also extend this timed framework so that any agent of a group can be authorized by its group (“lent”, in a sense) to perform an action “as if it were a member of another group”. This possibility of “lending an agent” to another group is particularly relevant to model situations where in a group there is no agent that can perform a specific action (for various reasons, e.g., there is not enough budget to pay the cost of the action or the group members are not enabled to perform it, etc.). In such cases, the group can be helped by an agent of another (willing) group that performs the action. For the resulting extension of *L-DINF* we provide fully-defined semantics, an axiomatization that turns out to be strongly complete.

Notice that the contents of this paper do not correspond to a small addition to our existing work. Rather, we are proposing an extensive reworking and improvement, where the new logic features relevant new capabilities, never seen in related work.

The paper is organized as follows. In Sect. 2 we introduce syntax and semantics of *L-DINF*. In Sect. 3 we present the axiomatization of our logic. A completeness result for *L-DINF* exploiting the notion of canonical model, defined in Sect. 4, can be found in Appendix A. In Sect. 5 we discuss a significant example of the use of our logic, emphasizing the usefulness of the newly introduced features. Finally, in Sect. 6 we conclude.

## 2. Logical Framework

Agents (or, groups of agents) can perform two types of actions: *mental actions* and *physical actions*. The latter kind has to be intended as the usual actions that an agent performs to interact with the environment in which it operates. So-called mental actions are the elementary steps of inference that an agent takes to develop its reasoning activity. The effect of mental actions is to modify the beliefs of the agent. Moreover, it is always a sequence of mental inferences that triggers the actual execution of a physical action. That is to say, the agent's internal reasoning makes it aware that an action can be performed.

Let us start by introducing the syntax of the underlying logic *L-DINF*, which consists of a static component and a dynamic component. The former, *L-INF*, is a logic of explicit beliefs and background knowledge. The dynamic part extends the static one with operators capturing the consequences of agents' mental actions on their explicit beliefs.

We are interested in modeling the knowledge and beliefs of agents that may hold in specific time instants/intervals. Basic pieces of knowledge/belief will be represented by atomic propositions. To model a notion of time we adopt a specific convention concerning the arguments of such atomic propositions. More specifically, let be given a collection of predicate symbols $\Pi = \{p, q, h, \ldots\}$, we consider the atomic propositions of the form $p(t', t'')$, where $p \in \Pi$ and $t', t'' \in \mathbb{N}$, such that $t' \leq t''$, are two *time instants*. (By some abuse of notation, we also consider the special cases in which $t'' = \infty$.) Then, let $Atm = \{p(t', t'') \mid p \in \Pi \text{ and } t', t'' \in \mathbb{N}\}$ denote the set of all atomic propositions. Here, an atomic proposition of the form $p(t_1, t_2)$ stands for *"p is true from the time instant* $t_1$ *to* $t_2$*"* with $t_1 \leq t_2$ which is the *Temporal Representation* of the external world. Atomic propositions of the form $p(t, t)$, stand for *"p is true in the time instant t"*.

Customarily, we may also admit predicate symbols of higher arity and hence atomic propositions involving more than two arguments. In such case, we assume that the first two arguments are those that identify a time interval (for example, the atomic proposition $turn\_on(1, 4, pc)$ may be used to state that *"the agent knows that the* $pc$ *is turned on from time 1 to time 4"*).

Let the set $Atm_A$ denote the collection of the physical actions that an agent can perform, including "active sensing" actions (such as, for instance, "let's check if the oven is on", etc.).

In what follows, we will use the symbols $I$, $J$, to denote MTL "time-intervals" [7]. Each time interval is either a closed finite interval $[t, l]$, for any pair of expressions/values $t, l$ such that $0 \leq t \leq l$, or an infinite interval $[t, \infty)$ (for some expression/value $t$).

Let $Agt = \{1, 2, \ldots, n\}$ be the set of $n$ agents (identified for simplicity by integer numbers). Let $Grp$ be the set of groups of agents, i.e., $\forall G \in Grp, G \subseteq Agt$ where $G$ is not empty. The language $\mathcal{L}_{\textit{L-DINF}}$ of *L-DINF* is defined as follows (where $p(t_1, t_2)$ ranges over $Atm$, $d \in \mathbb{N}$, $i, j \in Agt$, and $H, G \in Grp$):

$$
\begin{array}{lll}
\varphi, \psi & ::= & p(t_1, t_2) \mid \neg\varphi \mid \varphi \wedge \psi \mid \mathbf{B}_{i,j}\varphi \mid \mathbf{K}_i\varphi \mid \Box_I\varphi \mid \\
& & do_i(\phi_A, I) \mid can\_do_i(\phi_A, I) \mid intend_i(\phi_A, I) \mid exec_i(\alpha) \mid \\
& & do_G(\phi_A, I) \mid can\_do_G(\phi_A, I) \mid intend_G(\phi_A, I) \mid exec_G(\alpha) \mid \\
& & [G{:}\alpha]\,\varphi \mid pref\_do_i(\phi_A, d, I) \mid maxpref\_do_G(i, \phi_A, I) \mid lend_G(i, H, \phi_A, I) \\
\alpha & ::= & +_j\varphi \mid \vdash_j(\varphi,\psi) \mid \cap_j(\varphi,\psi) \mid \downarrow_j(\varphi, \psi) \mid \dashv_j(\varphi, \psi)
\end{array}
$$

Other Boolean operators are defined from $\neg$ and $\wedge$ as usual. The language of *mental actions* of type $\alpha$ is denoted by $\mathcal{L}_{\mathsf{ACT}}$. The static part *L-INF* includes only those formulas not having sub-formulas of the form $[G : \alpha]\,\varphi$.

Before defining the formal semantics of *L-DINF*, let us informally describe the intended meaning of the basic formulas. We are interested in modeling the reasoning of agents that form coalitions/groups and act cooperatively. Agents may leave their group to join or form another group. Hence, we consider the set $Agt$ of all agents as partitioned in groups: each agent $i \in Agt$ always belongs to a unique group $G \in Grp$. For simplicity, we assume that all agents initially belong to the same group. Any agent $i$, at any time $t$, can however perform a (physical) action $join_A(i, j, t)$, for $j \in Agt$, to change its group and join $j$'s group. The case of $i{=}j$ denotes the action that allows agent $i$ to leave its current group and form a new singleton group $\{i\}$.

Formulas of the form $\mathbf{K}_i\varphi$ express background knowledge of agent $i$: $\mathbf{K}_i$ is nothing but the S5 well-known modal operator. Background knowledge is considered to be a stable, irrevocable, reliable, and non-contradictory *knowledge base* which includes facts known by the agent from the beginning and it is assumed to satisfy *omniscience* principles. Instead, the agent's working memory stores its belief set. Each agent has its own belief set and also owns a version of the belief sets of all the other agents. We impose that all agents in the same group $G$ share all their beliefs (also concerning the version of the belief sets of agents outside $G$) and that whenever an agent $j$ joins a group $G$, it implicitly replaces all its beliefs with those of the group.[1] Beliefs concerning what other agents believe are expressed through the modality $\mathbf{B}_{i,j}$. The formula $\mathbf{B}_{i,j}\varphi$ expresses the fact that the formula $\varphi$ is in the version of the belief set of agent $j$ owned by agent $i$. (The case $i = j$ corresponds to the current working memory of $i$.) As we will see, $i$ can update its belief set by performing mental actions.

The formula $intend_i(\phi_A, I)$ indicates the intention of agent $i$ to perform the physical action $\phi_A$ in the interval $I$, in the sense of the BDI agent model [8]. These formulas can be part of the agent's knowledge base from the beginning or can be derived later. In this paper we do not cope with the formalization of BDI, for which the reader may refer, e.g., to [9]. Hence, we will deal with intentions rather informally, also assuming that $intend_G(\phi_A, I)$ holds whenever all agents of group $G$ intend to perform $\phi_A$ at a point in the time interval $I$.

The formula $do_G(\phi_A, I)$ (and similarly for $do_i(\phi_A, I)$) indicates the *actual execution* of action $\phi_A$ by $G$, i.e., by any agent $i \in G$, at a time point in $I$. Note that, we do not provide an axiomatization for $do_G$. In fact, we assume that in any concrete implementation of the logical framework, $do_G$ is realized by means of a *semantic attachment* [10], that is, a procedure which connects an agent with its external environment in a way that is unknown at the logical level. The axiomatization only concerns the relationship between doing and being enabled to do.

The formula $can\_do_i(\phi_A, I)$ is an enabling condition for $i$ to perform the physical action $\phi_A$ in the interval $I$, while $pref\_do_i(\phi_A, d, I)$ indicates the level $d$ of preference/willingness of agent $i$ to perform $\phi_A$ in $I$. $maxpref\_do_G(i, \phi_A, I)$ indicates that $i$ exhibits the *maximum level* of preference on performing $\phi_A$ within all group members (in the time interval $I$). The formula $lend_G(i, H, \phi_A, I)$, where $H$ and $G$ are two disjoint groups and $i \in H$, states that $H$ can lend $i$ to $G$, agent $i$ can perform $\phi_A$ and it is entitled and authorized by $H$ to perform $\phi_A$. Lending agents relates to formulas of the form $can\_do_G(\phi_A, I)$, that mean that either an agent in $G$ can perform $\phi_A$ or $G$ can borrow an agent $j$ from another group and $j$ performs $\phi_A$.

In the formula $\Box_I\phi$ the MTL interval "always" operator is applied to a formula, meaning that $\phi$ is always true during the interval $I$. (For simplicity, we will sometimes write $\Box_{[0,\infty)}$ as $\Box$.)

[1] An alternative possibility would be to admit that the agent $j$ joining $G$ contributes to the shared contents of the working memory of agents in $G \cup \{j\}$, by adding (part of) its belief set, possibly activating some form of belief fusion/revision. For the sake of simplicity, in this paper, we will not deal with this possibility.

The formulas $exec_i(\alpha)$ and $exec_G(\alpha)$ express executability of mental actions either by agent $i$, or by a group $G$ (which means that $\alpha$ is executable for any member of $G$).

The formula $[G{:}\alpha]\,\varphi$, where $\alpha$ is a mental action, states that "$\varphi$ holds after action $\alpha$ has been performed by an agent $i \in G$ and all agents in $G$ have common knowledge about this fact". Namely, all agents in $G$ have full visibility of this action and it is as if they had performed the action themselves. We distinguish five types of mental actions that allow us to capture the dynamic properties of explicit beliefs. These actions, performed by an agent $i$, form explicit beliefs via inference with respect to the version of $j$'s belief set owned by $i$ and update such representation of $j$'s working memory. (The case in which $i = j$ corresponds to reasoning steps performed by agent $i$ on the basis of its own beliefs.)

**$+_j(\varphi)$:** this mental action represents the action of learning a perceived belief. The formula $\varphi$ is assumed to be an atomic proposition $p(t_1, t_2)$. This mental operation is used to form a new belief from the perception $\varphi$. A perception may become a belief whenever an agent becomes "aware" of the perception and takes it into explicit consideration. Hence, after this action has been executed, agent $j$ believes $\varphi$ (as far as $i$ knows);

**$\downarrow_j(\varphi, \psi)$:** this action infers $\psi$ from $\varphi$, where $\psi$ is an atom, say $p(t_1, t_2)$, if $\varphi$ is believed by $j$ (as far as $i$ knows and in some suitable time interval including $[t_1, t_2]$) and, according to agent $i$ background knowledge, $\psi$ is a logical consequence of $\varphi$ and $j$ starts believing that $p(t_1, t_2)$ is true;

**$\cap_j(\varphi,\psi)$:** this action closes under conjunction the beliefs $\varphi$ and $\psi$ of $i$'s version of $j$'s belief set;

**$\dashv_j(\varphi, \psi)$:** this action is a simple form of belief revision, i.e., assuming $\varphi$ and $\psi$ are $p(t_1, t_2)$ and $q(t_3, t_4)$, respectively, $j$ believing $p(t_1, t_2)$ and according to $i$'s background knowledge that $p(t_1, t_2)$ implies $\neg q(t_3, t_4)$ , removes the timed belief $q(t_3, t_4)$ if the intervals match. Note that, should $q$ be believed in a wider interval $I$ such that $[t_1, t_2] \subseteq I$, the belief $q(.,.)$ is removed concerning intervals $[t_1, t_2]$ and $[t_3, t_4]$, but it is left for the remaining sub-intervals, hence, it is "restructured";

**$\vdash_j(\varphi,\psi)$:** this action infers $\psi$ from $\varphi$ if, where $\psi$ is an atom, say $p(t_1, t_2)$, as far as $i$ knows about $j$'s beliefs, $\varphi$ is believed by $j$ and according to $j$'s working memory $\psi$ is a logical consequence of $\varphi$ and $j$ starts believing that $p(t_1, t_2)$ is true.

## 2.1. Semantics of *L-INF*

Definition 1, to be seen, introduces the notion of *L-INF model*, which is then used to define the semantics of *L-INF*. The notion of *L-INF model* involves a "*time function*" $T$ that associates a time interval $I$ with each formula. Such time function is defined as follows:

- $T(p(t_1, t_2)) = [t_1, t_2]$ and $T(\neg\varphi) = T(\varphi)$;
- $T(\varphi \wedge \psi) = J$ where $J$ is the unique smallest interval including $T(\varphi)$ and $T(\psi)$;
- $T(\mathbf{B}_{i,j}\varphi) = T(\varphi)$ and $T(\mathbf{K}_i\varphi) = T(\varphi)$;
- $T(\Box_I\varphi) = I$;
- for formulas of the form $T([G{:}\alpha]\varphi)$ we consider different cases depending on $\alpha$:

1. $T([G:+_j(\varphi)]\,\varphi) = T(\varphi)$;
2. $T([G:\downarrow_j(\varphi,\psi)]\,\psi) = T([G:\vdash_j(\varphi,\psi)]\,\psi) = T(\psi)$;
3. $T([G:\cap_j(\varphi,\psi)]\,(\varphi\wedge\psi)) = T(\varphi\wedge\psi)$;
4. $T([G:\dashv_j(\varphi,\psi)]\,\psi) = J$, where $J$ is the smallest interval including the "restructured" set of time instants in which $\psi$ is believed after the belief update action;

- $T(do_i(\phi_A, I)) = I$ and $T(do_G(\phi_A, I)) = I$;
- $T(can\_do_i(\phi_A, I)) = I$ and $T(can\_do_G(\phi_A, I)) = I$;
- $T(intend_i(\phi_A, I)) = I$ and $T(intend_G(\phi_A, I)) = I$;
- $T(pref\_do_i(\phi_A, d, I)) = I$ and $T(maxpref\_do_G(i, \phi_A, I)) = I$;
- $T(lend_G(i, H, \phi_A, I)) = I$;
- $T(exec_i(\alpha)) = T(exec_{\{i\}}(\alpha))$ while for $T(exec_G(\alpha))$, depending on $\alpha$, we have:
  1. $T(exec_G(+_j(\varphi))) = T([G:+_j(\varphi)]\,\varphi)$;
  2. $T(exec_G(\downarrow_j(\varphi,\psi))) = T([G:\downarrow_j(\varphi,\psi)]\psi)$;
  3. $T(exec_G(\cap_j(\varphi,\psi))) = T([G:\cap_j(\varphi,\psi)](\varphi\wedge\psi))$;
  4. $T(exec_G(\dashv_j(\varphi,\psi))) = T([G:\dashv_j(\varphi,\psi)]\psi)$;
  5. $T(exec_G(\vdash_j(\varphi,\psi))) = T([G:\vdash_j(\varphi,\psi))]\psi)$.

Definition 1, below, also depends on a given *set of worlds* $W$ and on a *valuation function* $V : W \longrightarrow 2^{Atm}$. For each world $w \in W$, let $t_1$ the minimum time instant of $T(p(t_1, t))$ where $p(t_1, t) \in V(w)$ and let $t_2$ be the supremum time instant (we can have $t_2 = \infty$) w.r.t. the atoms $p(t, t_2)$ in $V(w)$. When useful, we will make explicit these two time instants $t_1$ and $t_2$ by denoting the world $w$ as $w_I$ with $I = [t_1, t_2]$.

**Definition 1.** *A L-INF model $M$ is composed of set of worlds (or situations) $W$, a collection $\mathcal{R}=\{R_i\}_{i\in Agt}$ of equivalence relations on $W$ (i.e., $R_i \subseteq W\times W$ for each $i \in Agt$), and a collection of semantic functions as follows:*

- *A* neighborhood *function* $N : Agt \times Agt \times W \longrightarrow 2^{2^W}$ *such that these conditions hold:*
  **(C1)** $\forall i, j \in Agt,\ \forall w_I \in W,\ \forall X \subseteq W\ \big(X \in N(i, j, w_I) \to X \subseteq \{v_I \in W \mid w_I R_i v_I\}\big)$,
  **(C2)** $\forall i, j \in Agt,\ \forall w_I, v_I \in W\ \big(w_I R_i v_I \to N(i, j, w_I) = N(i, j, v_I)\big)$;
- *An* executability *function for mental actions* $E : Agt \times W \longrightarrow 2^{\mathcal{L}_{\mathsf{ACT}}}$ *such that*
  **(C3)** $\forall i \in Agt,\ \forall w_I, v_I \in W\ \big(w_I R_i v_I \to E(i, w_I) = E(i, v_I)\big)$;
- *A* budget *function* $B : Agt \times W \longrightarrow \mathbb{N}$ *such that*
  **(C4)** $\forall i \in Agt,\ \forall w_I, v_I \in W\ \big(w_I R_i v_I \to B(i, w_I) = B(i, v_I)\big)$;
- *A* cost *function* $C : Agt \times \mathcal{L}_{\mathsf{ACT}} \times W \longrightarrow \mathbb{N}$ *such that*
  **(C5)** $\forall i \in Agt,\ \forall w_I, v_I \in W,\ \forall \alpha \in \mathcal{L}_{\mathsf{ACT}}\ \big(w_I R_i v_I \to C(i, \alpha, w_I) = C(i, \alpha, v_I)\big)$;
- *An* executability *function for physical actions* $A : Agt \times W \longrightarrow 2^{Atm_A}$ *such that*
  **(C6)** $\forall i \in Agt,\ \forall w_I, v_I \in W\ \big(w_I R_i v_I \to A(i, w_I) = A(i, v_I)\big)$;
- *A* preference *function for physical actions* $P : Agt \times W \times Atm_A \longrightarrow \mathbb{N}$ *such that*
  **(C7)** $\forall i \in Agt,\ \forall w_I, v_I \in W,\ \forall \phi_A \in Atm_A\ \big(w_I R_i v_I \to P(i, w_I, \phi_A) = P(i, v_I, \phi_A)\big)$;
- *An* enabling *function for physical actions* $U : Agt \times W \longrightarrow 2^{Atm_A}$ *such that*
  **(C8)** $\forall i \in Agt,\ \forall w_I, v_I \in W\ \big(w_I R_i v_I \to U(i, w_I) = U(i, v_I)\big)$;
- *A* lending *function* $L : Agt \times Grp \times Grp \times Atm_A \times W \longrightarrow \{\mathsf{true}, \mathsf{false}\}$ *such that*
  **(C9)** $\forall G, H \in Grp,\ \forall i \in H,\ \forall w_I, v_I \in W,\ \forall \phi_A \in Atm_A$
  $\big(w_I R_i v_I \to L(i, H, G, \phi_A, w_I) = L(i, H, G, \phi_A, v_I)\big)$;

- *The* time function $T$ *defined before;*
- *A* valuation *function for atomic propositions in* $Atm \quad V : W \longrightarrow 2^{Atm}$*;*
- *A* valuation *function* $S : W \longrightarrow 2^{\{do_G(\phi_A, I) | \phi_A \in Atm_A, G \in Grp, I \subseteq \mathbb{N}\}}$ *for formulas of the form* $do_G(\phi_A, I)$.

The set $R_i(w_I) = \{v_I \in W \mid w_I R_i v_I\}$ is the *epistemic state* of agent $i$ at $w_I$. It identifies the situations that $i$ considers possible at world $w_I$. In cognitive terms, $R_i(w_I)$ can be conceived as the set of all situations that $i$ *can* retrieve from its long-term memory and reason about.

While $R_i(w_I)$ concerns background knowledge, $N(i, i, w_I)$ is the set of all facts that agent $i$ explicitly believes at world $w_I$, a fact being identified with a set of worlds. Hence, if $X \in N(i, i, w_I)$ then, the agent $i$ has the fact $X$ under the focus of its attention and believes it. While $N(i, i, w_I)$ is the explicit *belief set* of agent $i$ at world $w_I$, the set $N(i, j, w_I)$, for $i \neq j$, represents the version known by $i$ of the belief set of agent $j$.

$E(i, w_I)$ is the set of mental actions that $i$ can execute at $w_I$ and $B(i, w_I)$ is the budget that $i$ has available to perform mental actions. Similarly, $C(i, \alpha, w_I)$ is the cost to be paid to execute $\alpha$.

As concerns physical actions, $A(i, w_I)$ is the set of actions that $i$ can execute at $w_I$, while its preference on executability of physical actions is determined by the function $P$. The integer value $d = P(i, w_I, \phi_A)$ is the degree of willingness of $i$ to execute $\phi_A$ in world $w_I$. $U(i, w_I)$ instead is the set of physical actions that $i$ is enabled by its group to perform in $I$. Hence, function $U$ defines the *role* of an agent in its group, via the actions that it is allowed to execute.

The expression $L(i, H, G, \phi_A, w_I)$ states that agent $i$ can be lent by its group $H$ to another group $G$ to perform $\phi_A$ (e.g., when no agents in $G$ can do it, cf. the semantics defined below).

Constraint (**C1**) imposes that agent $i$ can have explicit in its mind only facts that are compatible with its current epistemic state. According to (**C2**), if a world $v_I$ is compatible with the epistemic state of agent $i$ at $w_I$, then $i$ should have the same explicit beliefs at $w_I$ and $v_I$. In other words, if two situations are equivalent as concerns background knowledge, then they cannot be distinguished through the explicit belief set. This aspect of the semantics can be extended in future work to allow agents to make plausible assumptions. Properties analogous to (**C2**) are imposed by constraints (**C3**)–(**C9**). For instance, (**C3**) imposes that agent $i$ always knows which mental actions it can perform and those it cannot and that executability of mental actions cannot distinguish between $R_i$-equivalent situations.

**Remark 1.** *Notice that our way of defining a model is highly modular, aimed to facilitate implementation and to allow a designer to "tune" the practical behavior of a logical theory. In fact, the executability of mental and physical actions, costs, budget, preferences, etc., are all specified by means of specific special functions. Changing the definition of such functions will change the behavior of the agent that the underlying theory aims to define. Moreover, in implementation terms, a designer has total freedom about how these functions are realized.*

**Definition 2.** *Given a model* $M$ *as in Definition 1, the truth values for formulas are inductively defined as follows, where* $i, j \in Agt$, $G \in Grp$, $w_I \in W$, *and where we put* $||\varphi||^M_{i,w_I} = \{v_I \in W : w_I R_i v_I$ *and* $M, v_I \models \varphi\}$*:*

*(i)* $M, w_I \models p(t_1, t_2)$ *iff* $\forall t'_1, t'_2$ *such that* $t_1 \leq t'_1 \leq t'_2 \leq t_2$ *it holds that* $p(t'_1, t'_2) \in V(w_I)$ *and* $T(p(t_1, t_2)) \subseteq I$

(ii) $M, w_I \models \neg\varphi$ *iff* $M, w_I \not\models \varphi$ *and* $T(\neg\varphi) \subseteq I$
(iii) $M, w_I \models \varphi \wedge \psi$ *iff* $M, w_I \models \varphi \wedge M, w_I \models \psi$ *and* $T(\varphi), T(\psi) \subseteq I$
(iv) $M, w_I \models \mathbf{K}_i\, \varphi$ *iff* $T(\varphi) \subseteq I$ *and* $\forall\, v_I \in R_i(w_I)$ *it holds that* $M, v_I \models \varphi$
(v) $M, w_I \models \mathbf{B}_{i,j}\, \varphi$ *iff* $||\varphi||^M_{i,w} \in N(i, j, w_I)$ *and* $T(\varphi) \subseteq I$
(vi) $M, w_I \models \Box_J\varphi$ *iff* $T(\varphi) \subseteq J \subseteq I$ *and for all* $v_I \in R_i(w_I)$ *it holds that* $M, v_I \models \varphi$
(vii) $M, w_I \models exec_i(\alpha)$ *iff* $\alpha \in E(i, w_I)$ *and* $T(exec_i(\alpha)) \subseteq I$
(viii) $M, w_I \models exec_G(\alpha)$ *iff* $\exists i \in G$ *such that* $\alpha \in E(i, w_I)$ *and* $T(exec_G(\alpha)) \subseteq I$
(ix) $M, w_I \models lend_G(i, H, \phi_A, J)$ *iff* $G, H \in Grp$, $i \in H$, $L(i, H, G, \phi_A, w_I) = \mathsf{true}$, $\phi_A \in A(i, w_I) \cap U(i, w_I)$, $T(lend_G(i, H, \phi_A, J)) \subseteq I$, $G \cap H = \emptyset$ *and* $M, w_I \models maxpref\_do_H(i, \phi_A, J)$
(x) $M, w_I \models pref\_do_i(\phi_A, d, J)$ *iff* $P(i, w_I, \phi_A) = d$, $\phi_A \in A(i, w_I)$, $\phi_A \in U(i, w_I)$, *and* $T(pref\_do_i(\phi_A, d, J)) \subseteq I$
(xi) $M, w_I \models maxpref\_do_G(i, \phi_A, J)$ *iff* $T(maxpref\_do_G(i, \phi_A, J)) \subseteq I$, $i \in G$, *and* $M, w_I \models pref\_do_i(\phi_A, d, J)$ *for* $d = \max\{P(j, w_I, \phi_A) \mid j \in G \wedge \phi_A \in A(j, w_I) \cap U(j, w_I)\}$
(xii) $M, w_I \models do_G(\phi_A, J)$ *iff* $do_G(\phi_A, J) \in S(w_I)$ *and* $T(do_G(\phi_A), J) \subseteq I$
(xiii) $M, w_I \models do_i(\phi_A, J)$ *iff* $M, w_I \models do_{\{i\}}(\phi_A, J)$
(xiv) $M, w_I \models can\_do_i(\phi_A, J)$ *iff* $\phi_A \in A(i, w_I) \cap U(i, w_I) \wedge T(can\_do_i(\phi_A, J)) \subseteq I$
(xv) $M, w_I \models can\_do_G(\phi_A, J)$ *iff* $\big(\exists i \in G$ *such that* $\phi_A \in A(i, w_I) \cap U(i, w_I) \wedge T(can\_do_G(\phi_A, J)) \subseteq I\big) \vee \big(\exists\, H \in Grp,\ \exists j \in H,\ M, w_I \models lend_G(j, H, \phi_A, J)\big)$.

Point (i) establishes that if a predicate is true within an interval, it is also uniformly true in all sub-intervals. Note that a physical action $\phi_A$ can be performed by a group of agents if at least one agent of the group can do it. In this case, the agent which best prefers will execute the action. The last item of the above list, states that a group $G$ can perform in a certain situation (world $w_I$) and time interval $I$ an action $\phi_A$ in case an agent in the group has the right competencies (it is able, by the function $A$ and enabled by the function $U$), or if there exists another group $H$ including an agent $j$ with these competencies, and which is available to lend $j$ to $G$ (i.e., $L(j, H, G, \phi_A, w_I) = \mathsf{true}$).

A specific evaluation function $S$ deals with formulas of the form $do_G(\phi_A)$. These kinds of formulas are, nevertheless, left not axiomatized. This is because $do_G(\phi_A, I)$ refers to the practical execution of an action by some kind of actuator, where in a robotic application this action can have physical effects in time interval $I$. To find a way to account for such expression, we choose to resort to a concept that has been called by Weyhrauch in the seminal work [10] a *semantic attachment*, i.e., it is assumed that some device exists, which connects an agent with its external environment in a way that is unknown at the logical level. The aim of [10] was exactly to explain how formal systems could be used in AI by being "mechanized" in a practical way; thus, this work aimed to provide ideas about a principled though potentially running implementation of these systems. We assume that the function $S$ reflects at the semantic level the presence of such a semantic attachment mechanism. Hence, the semantics is concerned only with the relationship between doing and being enabled to do. A similar treatment is exploited for *join* actions. Performing $join_A(i, j, t)$ implies that agents $i, j$ are at time $t$ in the same group. We assume that the execution of $join_A(i, j, t)$ affects the contents of the working memories of the agents $i$ and $j$ (and, consequently, of the other members of their groups).

## 2.2. Semantics of mental actions

Given a model $M$ as in Definition 1 and a world $w_I$, we set:

*(xvi)* $M, w_I \models [G{:}\alpha]\varphi$ iff $M^{[G:\alpha]}, w_I \models \varphi$

where $M^{[G:\alpha]}$ is the model obtained from $M$ by replacing the two functions $N$ and $B$ with their *updated* versions $N^{[G:\alpha]}$ and $B^{[G:\alpha]}$, resp. (to be seen below). $M^{[G:\alpha]}$ represents the fact that the execution of $\alpha$ affects the sets of beliefs of agents in $G$ (i.e., the neighborhood function changes as mental actions are performed) and modifies the available budget.

We write $\models_{\textit{L-DINF}} \varphi$ to denote that $M, w_I \models \varphi$ holds for all worlds $w_I$ of every model $M$.

A key aspect in the definition of the logic is the following, which states under which conditions, and by which agents, an action may be performed:

$$\mathit{enabled}_{w_I}(i, G, \alpha) \;\leftrightarrow\; \big(i \in G \wedge \alpha \in E(i, w_I) \;\wedge\; \tfrac{C(i,\alpha,w_I)}{|G|} \leq \min_{h\in G} B(h, w_I)\big).$$

To handle the case of multiple enabled agents, we assume defined a predicate $\mathit{doer}_{w_I}(i, G, \alpha)$ to univocally select one among the enabled agents. Its definition might rely on any criteria, even involving background knowledge and belief sets. For simplicity, let us define such predicate as:

$$\mathit{doer}_{w_I}(i, G, \alpha) \leftrightarrow i = \min\{j | \mathit{enabled}_{w_I}(j, G, \alpha)\}.$$

This condition, as defined above, expresses the fact that a mental action is enabled when: at least an agent can perform it; and, the "payment" due by each agent, obtained by dividing the action's cost equally among all agents of the group, is within each agent's available budget (this choice is inherited from *L-DINF*). In case more than one agent in $G$ can execute an action, we implicitly assume the agent $j$ performing the action is the one corresponding to the lowest possible cost. Namely, $j$ is such that $C(j, \alpha, w_I) = \min_{h\in G} C(h, \alpha, w_I)$. Other choices might be viable, so variations of this logic can be easily defined by devising some other enabling condition and, possibly, introducing differences in neighborhood update.

Notice that the definition of the enabling function basically specifies the "role" that agents take while concurring with their own resources to actions' execution. Also, in the case of the specification of different resources, different corresponding enabling functions might be defined.

**Neighborhood updating** The updated neighborhood $N^{[G:\alpha]}$ resulting from execution of a mental action $\alpha$ is specified as follows (where $j$ is an agent not necessarily in $G$):

- If $\alpha$ is $+_j(\varphi)$, then, for each $i, h \in \mathit{Agt}$ and $w_I \in W$,

$$N^{[G:+_j(\varphi)]}(i, h, w_I) = N(i, h, w_I) \cup \{||\varphi||^M_{k,w_I}\}$$

  if $i{\in}G$ and $h{=}j$ and $\mathit{doer}_{w_I}(k, G, +_j(\varphi))$ holds and, moreover, $T([G{:}+_j(\varphi)]\varphi) \subseteq I$. Otherwise, the neighborhood function does not change (i.e., $N^{[G:+_j(\varphi)]}(i, h, w_I) = N(i, h, w_I)$).
- If $\alpha$ is $\downarrow_j(\psi, \chi)$, then, for each $i, h \in \mathit{Agt}$ and $w_I \in W$,

$$N^{[G:\downarrow_j(\psi,\chi)]}(i, h, w_I) = N(i, h, w_I) \cup \{||\chi||^M_{k,w_I}\}$$

  if $i{\in}G$ and $h{=}j$ and $\mathit{doer}_{w_I}(k, G, \downarrow_j(\psi, \chi))$ and $M, w_I \models \mathbf{B}_{k,j}\psi \wedge \mathbf{K}_k(\psi \rightarrow \chi)$ and $T([G{:}\downarrow_j(\psi, \chi)]\,\chi) \subseteq I$ hold. Otherwise, the neighborhood function does not change (i.e., $N^{[G:\downarrow_j(\psi,\chi)]}(i, h, w_I) = N(i, h, w_I)$).

- If $\alpha$ is $\cap_j(\psi,\chi)$, then, for each $i, h \in Agt$ and $w_I \in W$,
$$N^{[G:\cap_j(\psi,\chi)]}(i,h,w_I) = N(i,h,w) \cup \{||\psi \wedge \chi||^M_{k,w_I}\}$$
if $i \in G$ and $h = j$ and $doer_{w_I}(k, G, \cap_j(\psi,\chi))$ and $M, w_I \models \mathbf{B}_{k,j}\psi \wedge \mathbf{B}_{k,j}\chi$ and $T([G : \cap_j(\psi,\chi)]\,(\psi \wedge \chi)) \subseteq I$. Otherwise, the neighborhood function remains unchanged (i.e., $N^{[G:\cap_j(\psi,\chi)]}(i,h,w_I) = N(i,h,w_I)$).
- If $\alpha$ is $\vdash_j(\psi,\chi)$, then, for each $i, h \in Agt$ and $w_I \in W$,
$$N^{[G:\vdash_j(\psi,\chi)]}(i,h,w_I) = N(i,h,w_I) \cup \{||\chi||^M_{k,w_I}\}$$
if $i{\in}G$ and $h{=}j$ and $doer_{w_I}(k, G, \vdash_j(\psi,\chi))$ and $M, w_I \models \mathbf{B}_{k,j}\psi \wedge \mathbf{B}_{k,j}(\psi \rightarrow \chi)$ and $T([G{:}\vdash_j(\psi,\chi)]\,\chi) \subseteq I$ hold. Otherwise, the neighborhood function remains unchanged: $N^{[G:\vdash_j(\psi,\chi)]}(i,h,w_I) = N(i,h,w_I)$.
- If $\alpha$ is $\dashv_j(p(t_1,t_2), q(t_3,t_4))$, then, for each $i, h \in Agt$ and $w_I \in W$,
$$N^{[G:\dashv_j(p(t_1,t_2),q(t_3,t_4))]}(i,h,w_I) = \Big(\big(N(i,h,w_I) \setminus \{||q(x,y)||^M_{k,w}\}\big) \cup \{||q(x,t_3{-}1)||^M_{k,w_I}\} \cup \{||q(t_4{+}1,y)||^M_{k,w_I}\}\Big)$$
if $i{\in}G$, $h{=}j$, $doer_{w_I}(k, G, \dashv_j(p(t_1,t_2), q(t_3,t_4)))$, $M, w_I{\models}(\mathbf{B}_{k,j}(p(t_1,t_2)) \wedge \mathbf{K}_k(p(t_1,t_2){\rightarrow}\neg q(t_3,t_4)))$ and $T(q(x,y)){\subseteq}I$ hold, where $[x,y]$ is the smallest interval including $[t_3,t_4]$ and is such that $M, w_I{\models}\mathbf{B}_{k,j}(q(x,y))$ holds. If these conditions do not hold, the neighborhood function does not change (i.e., $N^{[G:\dashv_j(p(t_1,t_2),q(t_3,t_4))]}(i,h,w_I)$ $= N(i,h,w_I)$). Note that the possible update $||q(x,t_3{-}1)||^M_{k,w_I}$ (resp., $||q(t_4{+}1,y)||^M_{k,w_I}$) is added to the neighborhood only if the interval $[x,t_3{-}1]$ (resp., $[t_4{+}1,y]$) is not empty.

Notice that, after an agent $k \in G$ performed $\alpha$, all agents $i \in G$ see the same update in the neighborhood. Conversely, for any agent $i \notin G$ the neighborhood remains unchanged (i.e., $N^{[G:\alpha]}(i,h,w_I){=}N(i,h,w_I)$, for each $h$). However, even for agents in $G$ the neighborhood function remains unchanged if the required preconditions, on beliefs, knowledge, and budget (see the definitions of $doer_{w_I}$ and $enabled_{w_I}$), do not hold (and hence the action is not executed).

**Budget updating** Since each agent in $G$ has to contribute to cover the costs of execution by consuming part of its available budget, an update of the budget function is needed. Hence, for each $i \in Agt$ and each $w_I \in W$, whenever a mental action $\alpha$ is performed (recall that mental actions are performed by $k$ w.r.t. the $k$'s own version of the working memory of $j$), we set
$$B^{[G:\alpha]}(i,w_I) = B(i,w_I) - C(k,\alpha,w_I)/|G|,$$
if $i{\in}G$, $doer_{w_I}(k,G,\alpha)$ holds, and, depending on $\alpha$, the same conditions described before to enable neighborhood updates are satisfied. Otherwise, the budget is preserved, i.e., $B^{[G:\alpha]}(i,w_I){=}B(i,w_I)$. Clearly, the budget is preserved for those agents that are not in $G$.

**Joining a group** A comment is due concerning the action $join_A(i,j,t)$. As mentioned, we assume that whenever an agent $i \in G$ joins the group of another agent $j$ (by executing $join_A(i,j,t)$), the neighborhood function $N(i,h,w_I)$ becomes equal to $N(j,h,w_I)$, for each $h \in Agt$. In case $i \in G$ executes $join_A(i,i,t)$ (i.e., it leaves $G$ and forms a new singleton group) then it maintains its current neighborhood function, but without any binding with the belief set of the other agents in $G$.

**Remark 2.** *In the mental actions* $\vdash_i(\varphi,\psi)$ *and* $\downarrow_i(\varphi,\psi)$*, the new belief* $\psi$ *which is inferred can be of the forms* $can\_do_G(\phi_A, I)$*,* $can\_do_i(\phi_A, I)$*,* $do_G(\phi_A, I)$*, or* $do_i(\phi_A, I)$*, which denotes the actual (possibility of) execution of the physical action* $\phi_A$*, by* $i$ *or by* $G$*. The conclusion* $do_i(\phi_A, I)$ *(derived, for instance, from* $can\_do_i(\phi_A, I)$ *and possibly other conditions) implies that the physical action* $\phi_A$ *is performed by* $i$ *(and similarly for* $do_G(\phi_A, I)$*).*

*We assume that actions succeed by default. In a concrete system, in case of failure, a corresponding* failure event *would be perceived by the agent (again, we rely on* semantic attachment*).*

## 3. Axiomatization

In this section, we describe an axiomatic system for *L-DINF*. This yields a derivation notion $\vdash$ for *L-DINF*. The axiomatization is sound for the class of *L-INF* models and strongly complete (cf. Appendix A). The axiomatic system is composed of the following axioms and inference rules (where $G \in Gpr$ and $i, j, k \in Agt$), together with the usual axioms of propositional logic:

1. $(\mathbf{K}_i\varphi \wedge \mathbf{K}_i(\varphi \to \psi)) \to \mathbf{K}_i\psi$
2. $\mathbf{K}_i\varphi \to \varphi$
3. $\neg\mathbf{K}_i(\varphi \wedge \neg\varphi)$
4. $\mathbf{K}_i\varphi \to \mathbf{K}_i\mathbf{K}_i\varphi$
5. $\neg\mathbf{K}_i\varphi \to \mathbf{K}_i\neg\mathbf{K}_i\varphi$
6. $\mathbf{B}_{i,j}\varphi \wedge \mathbf{K}_i(\varphi \leftrightarrow \psi) \to \mathbf{B}_{i,j}\psi$
7. $\mathbf{B}_{i,j}\varphi \to \mathbf{K}_i\mathbf{B}_{i,j}\varphi$
8. $\dfrac{\varphi}{\mathbf{K}_i\varphi}$
9. $p(t_1, t_2) \to p(t'_1, t'_2)$ with $[t'_1, t'_2] \subseteq [t_1, t_2]$
10. $\Box_I\varphi \wedge \Box_I(\varphi \to \psi) \to \Box_I(\psi)$
11. $\Box_I\varphi \to \Box_J\varphi$ with $J \subseteq I$
12. $[G{:}\alpha]p(t_1, t_2) \leftrightarrow p(t_1, t_2)$
13. $[G{:}\alpha]\neg\varphi \leftrightarrow \neg[G{:}\alpha]\varphi$
14. $[G{:}\alpha](\varphi \wedge \psi) \leftrightarrow [G{:}\alpha]\varphi \wedge [G{:}\alpha]\psi$
15. $[G{:}\alpha]\mathbf{K}_i\varphi \leftrightarrow \mathbf{K}_i([G{:}\alpha]\varphi)$
16. $[G{:}+_j(\varphi)]\mathbf{B}_{i,j}\psi \leftrightarrow \big(\mathbf{B}_{i,j}([G{:}+_j(\varphi)]\psi) \vee \big(doer_{w_I}(k, G, +_j(\varphi)) \wedge \mathbf{K}_i([G{:}+_j(\varphi)]\psi \leftrightarrow \varphi)\big)\big)$
17. $[G{:}\downarrow_j(\varphi,\psi)]\mathbf{B}_{i,j}\chi \leftrightarrow \big(\mathbf{B}_{i,j}([G{:}\downarrow_j(\varphi,\psi)]\chi) \vee \big(doer_{w_I}(k, G, \downarrow_j(\varphi,\psi)) \wedge \mathbf{B}_{k,j}\varphi \wedge \mathbf{K}_k(\varphi \to \psi) \wedge \mathbf{K}_i([G{:}\downarrow_j(\varphi,\psi)]\chi \leftrightarrow \psi)\big)\big)$
18. $[G{:}\cap_j(\varphi,\psi)]\mathbf{B}_{i,j}\chi \leftrightarrow \big(\mathbf{B}_{i,j}([G{:}\cap(\varphi,\psi)]\chi) \vee \big(doer_{w_I}(k, G, \cap_j(\varphi,\psi)) \wedge \mathbf{B}_{k,j}\varphi \wedge \mathbf{B}_{k,j}\psi \wedge \mathbf{K}_i([G{:}\cap_j(\varphi,\psi)]\chi \leftrightarrow (\varphi \wedge \psi))\big)\big)$
19. $[G{:}\vdash_j(\varphi,\psi)]\mathbf{B}_{i,j}\chi \leftrightarrow \big(\mathbf{B}_{i,j}([G{:}\vdash(\varphi,\psi)]\chi) \vee \big(doer_{w_I}(k, G, \vdash_j(\varphi,\psi)) \wedge \mathbf{B}_{k,j}\varphi \wedge \mathbf{B}_{k,j}(\varphi \to \psi) \wedge \mathbf{K}_i([G{:}\vdash_j(\varphi,\psi)]\chi \leftrightarrow \psi)\big)\big)$
20. $[G{:}\dashv_j(\varphi,\psi)]\mathbf{B}_{i,j}\chi \leftrightarrow \big(\mathbf{B}_{i,j}([G{:}\dashv_j(\varphi,\psi)]\chi) \vee \big(doer_{w_I}(k, G, \dashv_j(\varphi,\psi)) \wedge \mathbf{B}_{k,j}\varphi \wedge \mathbf{K}_k(\varphi \to \neg\psi) \wedge \mathbf{K}_i\big([\dashv_j(\varphi,\psi)]\chi \leftrightarrow \neg\psi\big)\big)\big)$
21. $exec_G(\alpha) \leftrightarrow \bigvee_{i \in G} exec_i(\alpha)$
22. $can\_do_G(\phi_A, I) \leftrightarrow \bigvee_{i \in G} can\_do_i(\phi_A, I) \vee \bigvee_{H \in Grp, j \in H} lend_G(j, H, \phi_A, J)$
23. $pref\_do_i(\phi_A, d, I) \to can\_do_i(\phi_A, I)$
24. $maxpref\_do_{G \cup \{i\}}(i, \phi_A, I) \leftrightarrow pref\_do_i(\phi_A, d, I) \wedge \bigwedge_{j \in G}(pref\_do_j(\phi_A, d_j, I) \to d_j \leq d)$
25. $lend_G(i, H, \phi_A, I) \leftrightarrow (G \cap H = \emptyset) \wedge i \in H \wedge can\_do_i(\phi_A, I) \wedge maxpref\_do_H(i, \phi_A, I)$

26. $\dfrac{\psi \leftrightarrow \chi}{\varphi \leftrightarrow \varphi[\psi/\chi]}$

We write *L-DINF* $\vdash \varphi$ to denote that $\varphi$ is a theorem of *L-DINF*.

The above axiomatization is sound for the class of *L-INF* models. All axioms are valid and the inference rules (8) and (26) preserve validity. In particular, soundness of axioms (16)–(20) follows from the semantics of $[G : \alpha]\varphi$, as previously defined. As mentioned earlier in the paper, the axiomatization does not deal with formulas of the forms $do_G(\phi_A)$, as they are intended to be realized by a *semantic attachment*, that connects an agent with its external environment.

# 4. Canonical Model

In this section, we introduce the notion of canonical model of our logic (compare Definitions 3 and 1). We have developed a proof of strong completeness w.r.t. the proposed class of models, by means of a standard canonical-model argument, that can be found in Appendix A.

As before, let $Agt$ be a set of agents.

**Definition 3.** *The* canonical *L-INF model $M_c$ is specified by choosing the set $W_c$ of all maximal consistent subsets of $\mathcal{L}_{\textit{L-INF}}$ as a collection of situations and by defining the following components:*

- *For $w_I \in W_c$, $\varphi \in \mathcal{L}_{\textit{L-INF}}$ let $D_\varphi(i, w_I) = \{v \in R_{c,i}(w_I) | \varphi \in v\}$. Then, we put $N_c(i, j, w_I) = \{D_\varphi(i, w_I) | \mathbf{B}_{i,j}\, \varphi \in w_I\}$.*
- *$\mathcal{R}_c = \{R_{c,i}\}_{i \in Agt}$ is a collection of equivalence relations on $W_c$ such that, for every $i \in Agt$ and $w_I, v_I \in W_c$, $w_I R_{c,i} v_I$ if and only if (for all $\varphi$, $\mathbf{K}_i\, \varphi \in w_I$ implies $\varphi \in v_I$).*
- *$E_c : Agt \times W_c \longrightarrow 2^{\mathcal{L}_{\mathsf{ACT}}}$ is such that for each $i \in Agt$ and $w_I, v_I \in W_c$, if $w_I R_{c,i} v_I$ then $E_c(i, w_I) = E_c(i, v_I)$.*
- *$B_c : Agt \times W_c \longrightarrow \mathbb{N}$ is such that for each $i \in Agt$ and $w_I, v_I \in W_c$, if $w_I R_{c,i} v_I$ then $B_c(i, w_I) = B_c(i, v_I)$.*
- *$C_c : Agt \times \mathcal{L}_{\mathsf{ACT}} \times W_c \longrightarrow \mathbb{N}$ is such that for each $i \in Agt$, $\alpha \in \mathcal{L}_{\mathsf{ACT}}$, and $w_I, v_I \in W_c$, if $w_I R_{c,i} v_I$ then $C_c(i, \alpha, w_I) = C_c(i, \alpha, v_I)$.*
- *$A_c : Agt \times W_c \longrightarrow 2^{Atm_A}$ is such that for each $i \in Agt$ and $w_I, v_I \in W_c$, if $w_I R_{c,i} v_I$ then $A_c(i, w_I) = A_c(i, v_I)$.*
- *$P_c : Agt \times W_c \times Atm_A \longrightarrow \mathbb{N}$ is such that for each physical action $\phi_A$ and $i \in Agt$ and $w_I, v_I \in W$, if $w_I R_{c,i} v_I$ then $P_c(i, w_I, \phi_A) = P_c(i, v_I, \phi_A)$.*
- *$U_c : Agt \times W_c \longrightarrow 2^{Atm_A}$ is such that, for each $i \in Agt$ and $w_I, v_I \in W$, if $w_I R_{c,i} v_I$ then $U_c(i, w_I) = U_c(i, v_I)$;*
- *$L_c : Agt \times Grp \times Grp \times Atm_A \times W_c \longrightarrow \{\mathsf{true}, \mathsf{false}\}$ is such that for each $i \in Agt$, $G, H \in Grp$, $\phi_A$ and $w_I, v_I \in W$, if $w_I R_{c,i} v_I$ then $L_c(i, H, G, \phi_A, w_I) = L_c(i, H, G, \phi_A, v_I)$;*
- *$T_c$ is the time function, as defined in Sect. 2.1;*
- *$V_c : W_c \longrightarrow 2^{Atm}$ is such that $V_c(w_I) = Atm \cap w_I$.*
- *$S_c : W_c \longrightarrow 2^{\{do_G(\phi_A; I), | \phi_A \in Atm_A, i \in Agt, G \in Grp, I \subseteq \mathbb{N}\}}$ such that $S_c(w_I) \subseteq w_I$.*

Where, analogously to what we have done before, $R_{c,i}(w_I) = \{v_I \in W_c \mid w_I R_{c,i} v_I\}$.

It is easy to verify that $M_c$ is an *L-INF* model as specified in Definition 1, since it satisfies conditions (**C1**)–(**C9**). Hence, it models the axioms and the inference rules introduced before. The following properties hold for each $w_I \in W_c$ and $\varphi \in \mathcal{L}_{\textit{L-INF}}$:

- $\mathbf{K}_i\, \varphi \in w_I$ iff $\forall v_I \in W_c$ such that $w_I R_{c,i} v_I$ we have $\varphi \in v_I$;
- if $\mathbf{B}_{i,j}\, \varphi \in w_I$ and $w_I R_{c,i} v_I$ then $\mathbf{B}_{i,j}\, \varphi \in v_I$.

Thus, $R_{c,i}$-related worlds have the same knowledge and $N_c$-related worlds have the same beliefs, i.e. there can be $R_{c,i}$-related worlds with different beliefs.

## 5. An example

In this section, we propose an example to explain the usefulness of this kind of logic and of the new extensions. Consider a group $G_{AI}$ of three agents, who constitute a research group, say in $AI$. They are the co-authors of a paper that has to be submitted to a conference: the first author, *bob*, deals with the drafting of the introduction and finding the references, the second one, $alice$, deals with the experiments, and the third one, $patrick$, deals with the formalization part. The second one, $alice$, is the only one who can perform the experiments because she has the required certifications; the other two agents are enabled to perform different tasks, such as write the abstract, searching references, checking the correctness of the formal part, and so on. However, none of the agents can perform the checking of the English. Luckily, there is another research group, say in Software Engineering, $G_{SE}$, that is willing (as we assume is dictated by the function $L$ in Definition 1) to provide help, and has among its members an agent, $mary$, that is able, is allowed and is the most willing (of the group) to check the English. Hence, according to point (ix) of Definition 2, this agent will be lent from $G_{SE}$ to $G_{AI}$ in the required time interval.

The group $G_{AI}$ receives notification of a deadline for a paper, so they decide to organize themselves for submitting it. The group will reason, and devise the intention/goal $\mathbf{K}_i(\Box_I(\mathit{submit\_fullpaper}(t_0, t_2))$: the group intends to submit their paper within the indicated time interval $I$. Here $t_0$ is the time instant when the group begins to organize to write the paper after the call was issued at $t_{call}$ ($t_0 \geq t_{call}$), so $I = [t_{call}, t_1]$ where $t_1$ is the deadline and $t_2$ is the time instant when they really submit the paper and $t_2 \leq t_1$.

Among the physical actions that agents in the group can perform are for instance the following: $\mathit{submit\_abstract}$, $\mathit{do\_experiment}$, $\mathit{write\_introduction}$, $\mathit{write\_formal\_part}$ and $\mathit{write\_experiment\_results}$. They are instead not able to perform the action $\mathit{check\_the\_english}$.

The group will now be required to perform a planning activity. Assume that, as a result of the planning phase, the knowledge base of each agent $i$ contains the following rule, that specifies how to reach the intended goal in terms of actions to perform and sub-goals to achieve (listed after the " $\rightarrow$ "):

$$\begin{aligned}
\mathbf{K}_i\big(&\Box_I(\mathit{submit_fullpaper}(t_0, t_2)) \rightarrow \\
&\Box_{I_1}(\mathit{submit_abstract}(t_0, t_3)) \wedge \Box_{I_2}(\mathit{do_experiment}(t_0, t_4)) \wedge \\
&\Box_I(\mathit{write_formal_part}(t_0, t_5)) \wedge \Box_{I_3}(\mathit{check_the_english}(t_5, t_6))\big)
\end{aligned}$$

where $I_1, I_2, I_3 \subseteq I$, $t_3$ is the time instant when the authors submit the abstract, and $t_3 \leq t_1$, $t_4$ is the time instant when the author $alice$ has finished the experiments and has written the results at $t_4 \leq t_1$, finally $t_5$ is the time instant when the other agent has finished to write the formal part. At this point, someone can start checking the English before submitting. By $t_6 \leq t_2$ the English check should have been completed, so that the paper can be submitted.

Assume now that the knowledge base of the group $G_{AI}$, shared by every agent $i \in G_{AI}$ contains also facts, stating the abilities, permissions, and preferences of agents of the group, which turn out to be able to perform each of the necessary actions, but one. Which agent will in particular perform each action $\phi_A$? According to items (x) and (xiv) and (xi) in Definition 2, this agent will be chosen as the one that best prefers to perform this action, among those that can do it. Formally, $maxpref\_do_G(i, \phi_A, I)$ identifies the agent $i$ in the group with the highest degree of preference on performing $\phi_A$, and $can\_do_G(\phi_A, I)$ is true if, in the time interval $I$, there is some agent $i$ in the group which is able and allowed to perform $\phi_A$.

Since no member of the group is capable of checking the English, and supposing that in the present situation, $G_{SE}$ is willing to help by lending an agent $j$, i.e., that $L(j, G_{SE}, check\_the\_english, G_{AI}, w_{I_3}) = \mathsf{true}$ then by point (xiv) of Definition 2 the best willing agent in $G_{SE}$, say agent $j$ is $mary$, will be lent to $G_{AI}$ for interval $I_3$, and will thus perform the check of the English. Thus we will have:

$$
\begin{aligned}
&\mathbf{K}_i\big(\Box_{I_1}(submit\_abstract(t_0, t_3)) \wedge can\_do_{G_{AI}}(submit\_abstract(t_0, t_3), I_1) \wedge \\
&\quad maxpref\_do_{G_{AI}}(bob, submit\_abstract(t_0, t_3), I_1) \rightarrow do_{bob}(submit\_abstract(t_0, t_3), I_1)\big) \\
&\mathbf{K}_i\big(\Box_{I_2}(do\_experiment(t_0, t_4)) \wedge can\_do_{G_{AI}}(do\_experiment(t_0, t_4), I_2) \wedge \\
&\quad maxpref\_do_{G_{AI}}(alice, do\_experiment(t_0, t_4), I_2) \rightarrow do_{alice}(do\_experiment(t_0, t_4), I_2)\big) \\
&\mathbf{K}_i\big(\Box_{I}(write\_formal\_part(t_0, t_5)) \wedge can\_do_{G_{AI}}(write\_formal\_part(t_0, t_5), I) \wedge \\
&\quad maxpref\_do_{G_{AI}}(partick, write\_formal\_part(t_0, t_5), I) \rightarrow do_{partick}(write\_formal\_part(t_0, t_5), I)\big) \\
&\mathbf{K}_i\big(\Box_{I_3}(check\_the\_english(t_5, t_6)) \wedge can\_do_{G_{AI}}(check\_the\_english(t_5, t_6), I_3) \wedge \\
&\quad lend_{G_{AI}}(mary, G_{SE}, check\_the\_english, I_3) \rightarrow do_{mary}(check\_the\_english(t_5, t_6), I_3)\big)
\end{aligned}
$$

Hence, the $G_{AI}$ group will be able to complete the plan and achieve the goal.

## 6. Conclusion

In this paper we extended an epistemic logical framework previously introduced in [11, 12].

The extension enriches the framework with: (a) The possibility of specifying time intervals to express the time periods in which agents' beliefs hold. Clearly, this affects the executability of actions and effects of the mental actions that agents use to modify their belief sets; (b) The possibility of exchanging agents between groups, i.e., an agent can be lent by a group $H$ to another group $G$; this possibility is relevant to model situations where in $G$ there is no agent that can perform a specific action and is conditioned to the willingness of group $H$ to lend (established by the outcome of the lending function $L$ included in the definition of *L-INF* model), and to the availability of an agent able and authorized to perform the required action. Then, a group can obtain help from another group if the latter includes (at least) one agent that can perform the required action. This greatly enhances a group's capabilities to reach its goals.

## References

[1] A. I. Goldman, Theory of mind, in: E. Margolis, R. Samuels, S. P. Stich (Eds.), The Oxford Handbook of Philosophy of Cognitive Science, volume 1, Oxford University Press, 2012, pp. 402–424.

[2] S. Costantini, A. Formisano, V. Pitoni, An epistemic logic for modular development of multi-agent systems, in: N. Alechina, M. Baldoni, B. Logan (Eds.), EMAS 2021: 9th International Workshop on Engineering Multi-Agent Systems, Proceedings, volume 13190 of *LNCS*, Springer, 2021.

[3] S. Costantini, A. Formisano, V. Pitoni, Modelling agents roles in the epistemic logic L-DINF, in: O. Arieli, G. Casini, L. Giordano (Eds.), Proc. of the 20th International Workshop on Non-Monotonic Reasoning, NMR 2022, Part of the Federated Logic Conference (FLoC 2022), Haifa, Israel, August 7-9, 2022, volume 3197 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2022, pp. 70–79. URL: https://ceur-ws.org/Vol-3197/paper7.pdf.

[4] S. Costantini, A. Formisano, V. Pitoni, An epistemic logic for formalizing group dynamics of agents, Interaction Studies 23 (2022) 391–426. doi:10.1075/is.22019.cos.

[5] S. Costantini, A. Formisano, V. Pitoni, Timed memory in resource-bounded agents, in: C. Ghidini, B. Magnini, A. Passerini, P. Traverso (Eds.), Proc. of AI*IA 2018, volume 11298 of *LNCS*, Springer, 2018, pp. 15–29.

[6] S. Costantini, A. Formisano, V. Pitoni, Temporalizing epistemic logic L-DINF, in: R. Calegari, G. Ciatto, A. Omicini (Eds.), Proc. of the 37th Italian Conference on Computational Logic, Bologna, Italy, June 29 - July 1, 2022, volume 3204 of *CEUR Workshop Proc.*, 2022.

[7] R. Koymans, Specifying real-time properties with metric temporal logic, Real-Time Systems 2 (1990) 255–299.

[8] A. S. Rao, M. Georgeff, Modeling rational agents within a BDI architecture, in: Proc. of the Second Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'91), Morgan Kaufmann, 1991, pp. 473–484.

[9] H. V. Ditmarsch, J. Y. Halpern, W. V. D. Hoek, B. Kooi, Handbook of Epistemic Logic, College Publications, 2015. Editors.

[10] R. W. Weyhrauch, Prolegomena to a theory of mechanized formal reasoning, Artif. Intell. 13 (1980) 133–170.

[11] S. Costantini, V. Pitoni, Towards a logic of "inferable" for self-aware transparent logical agents, in: C. Musto, D. Magazzeni, S. Ruggieri, G. Semeraro (Eds.), Proc. of the Italian Workshop on Explainable AI co-located with AI*IA 2020, volume 2742 of *CEUR Workshop Proc.*, CEUR-WS.org, 2020, pp. 68–79.

[12] S. Costantini, A. Formisano, V. Pitoni, An epistemic logic for multi-agent systems with budget and costs, in: W. Faber, G. Friedrich, M. Gebser, M. Morak (Eds.), Logics in Artificial Intelligence, Proc. of 17th European Conference, JELIA'21, volume 12678 of *LNCS*, 2021.


# A. Proofs of Strong Completeness

By exploiting the notion of canonical model (Definition 3), we can develop a proof of strong completeness for the axiomatic system defined in Sect. 3. Let us start with some preliminary results:

**Lemma 1.** *For all $w_I \in W_c$ and $\mathbf{B}_{i,j}\,\varphi, \mathbf{B}_{i,j}\,\psi \in \mathcal{L}_{\textit{L-INF}}$, if $\mathbf{B}_{i,j}\,\varphi \in w_I$ but $\mathbf{B}_{i,j}\,\psi \notin w_I$, it follows that there exists $v_I \in R_{c,i}(w_I)$ such that $\varphi \in v_I \leftrightarrow \psi \notin v_I$.*

*Proof.* Let $w_I \in W_c$ and $\mathbf{B}_{i,j}\,\varphi \in w_I$ and $\mathbf{B}_{i,j}\,\psi \notin w_I$. Assume now that for every $v_I \in R_{c,i}(w_I)$ we have $\varphi \in v_I \wedge \psi \in v_I$ or $\varphi \notin v_I \wedge \psi \notin v_I$; then, it follows that $\mathbf{K}_i(\varphi \leftrightarrow \psi) \in w_I$, and, by axiom (6), that $\mathbf{B}_{i,j}\,\psi \in w_I$, which is a contradiction. □

**Lemma 2.** *For all $\varphi \in \mathcal{L}_{L\text{-}INF}$ and $w_I \in W_c$ it holds that $\varphi \in w_I$ if and only if $M_c, w_I \models \varphi$.*

*Proof.* The proof is by induction on the structure of formulas $\varphi \in \mathcal{L}_{L\text{-}INF}$. For example, if $\varphi = p(t_1, t_2)$, $w_I \in W_c$, and $T(p(t_1, t_2)) \subseteq I$, then $p(t_1, t_2) \in w_I$ iff $p(t'_1, t'_2) \in V_c(w_I)$ for all $t'_1, t'_2$ such that $t_1 \leq t'_1 \leq t'_2 \leq t_2$. This means that $M_c, w_I \models p(t_1, t_2)$ by Definition 2. Other cases follows similarly from the definition of $W_c$ and the semantics of *L-INF*. For formulas of the form $\mathbf{B}_{i,j}\,\varphi$ we have to consider the neighborhood function $N$: assume $\mathbf{B}_{i,j}\,\varphi \in w_I$ for $w_I \in W_c$ and $T(\mathbf{B}_{i,j}\,\varphi) \subseteq I$. We have that $D_\varphi(i, w_I) = \{v_I \in R_{c,i}(w_I) \mid \varphi \in v_I\}$. By the definition of $\parallel \cdot \parallel^{M}_{i,w_I}$, and by the inductive hypothesis, we have $D_\varphi(i, w_I) = \parallel \varphi \parallel^{M_c}_{i,w_I} \cap R_{c,i}(w_I)$. Hence, $M_c, w_I \models \mathbf{B}_{i,j}\,\varphi$. Suppose $\mathbf{B}_{i,j}\,\varphi \notin w_I$, and then $\neg\mathbf{B}_{i,j}\,\varphi \in w_I$; we have to prove $\parallel \varphi \parallel^{M_c}_{w_I} \cap R_{c,i}(w_I) \notin N_c(i, j, w_I)$. Choose $D \in N_c(i, j, w_I)$. By definition, we know that $D = D_\psi(i, w_I)$ for some $\psi$, with $\mathbf{B}_{i,j}\,\psi \in w_I$. By Lemma 1 there is some $v_I \in R_{c,i}(w_I)$ such that $\varphi \in v_I \leftrightarrow \psi \notin v_I$. By the induction hypothesis, we obtain that either $v_I \in (\parallel \varphi \parallel^{M_c}_{w_I} \cap R_{c,i}(w_I)) \setminus D_\psi(i, w_I)$ or $v_I \in D_\psi(i, w_I) \setminus (\parallel \varphi \parallel^{M_c}_{i,w_I} \cap R_{c,i}(w_I))$ holds. Consequently, in both cases, $D_\psi(i, w_I) \neq \parallel \varphi \parallel^{M_c}_{i,w_I} \cap R_{c,i}(w_I)$. Thanks to the arbitrariness in the choice of $D$, we conclude that $\parallel \varphi \parallel^{M_c}_{i,w_I} \cap R_{c,i}(w_I) \notin N_c(i, j, w_I)$. Hence $M_c, w_I \not\models \mathbf{B}_{i,j}\,\varphi$. □

A crucial result states that for any *L-DINF* formula we can find an equivalent *L-INF* formula:

**Lemma 3.** *For all $\varphi \in \mathcal{L}_{L\text{-}DINF}$ there exists $\tilde{\varphi} \in \mathcal{L}_{L\text{-}INF}$ such that L-DINF $\vdash \varphi \leftrightarrow \tilde{\varphi}$.*

*Proof.* The proof is by induction on the structure of formulas. Let us show the case of $\varphi{=}p(t_1, t_2)$, the others cases are shown analogously. By axiom (12) we have $[G : \alpha]p(t_1, t_2) \leftrightarrow p(t_1, t_2)$ with $T(p(t_1, t_2)) \subseteq I$and by axiom (26) we have $\frac{[G:\alpha]p(t_1,t_2)\leftrightarrow p(t_1,t_2)}{\varphi\leftrightarrow\varphi[[G:\alpha]p(t_1,t_2)/p(t_1,t_2)]}$ with $T([G : \alpha]p(t_1, t_2)) \subseteq I$; then, we can obtain $\tilde{\varphi}$ by replacing $[G{:}\alpha]p(t_1, t_2)$ with $p(t_1, t_2)$ in $\varphi$. □

The previous lemmas allow us to prove the following:

**Theorem 1.** *L-INF is strongly complete for the class of L-INF models.*

*Proof.* Any consistent set $\Phi$ may be extended to a maximal consistent set $w_i{}^\star \in W_c$ and $M_c, w_I{}^\star \models \Phi$ by Lemma 2. Then, *L-INF* is strongly complete for the class of *L-INF* models. □

**Theorem 2.** *L-DINF is strongly complete for the class of L-INF models.*

*Proof.* If $\Phi$ is a consistent set of $\mathcal{L}_{L\text{-}DINF}$ formulas then by applying Lemma 3 we can obtain the set $\tilde{\Phi} = \{\tilde{\varphi} \mid \varphi \in \Phi\}$, which is a consistent set of $\mathcal{L}_{L\text{-}INF}$ formulas. By Theorem 1 $M_c, w_I \models \tilde{\Phi}$. But since *L-DINF* is sound and for each $\varphi \in \Phi$, *L-DINF* $\vdash \varphi \leftrightarrow \tilde{\varphi}$, it follows $M_c, w_I \models \Phi$ then *L-DINF* is strongly complete for the class of *L-INF* models. □