BDI^{ATL}: An Alternating-time BDI Logic for Multiagent Systems

Roberto Montagna Giorgio Delzanno Maurizio Martelli
Viviana Mascardi

Dipartimento di Informatica e Scienze dell'Informazione – DISI,
Università di Genova, Via Dodecaneso 35, 16146, Genova, Italy.
`{montagna,giorgio,martelli,mascardi}@disi.unige.it`

Abstract

Many logics for modelling beliefs, desires and intentions of agents, such as Rao and Georgeff's BDI logic and Wooldridge's $\mathcal{LORA}$, are based on temporal logics like CTL/CTL* (Computational Tree Logic) in which the structure of time is branching in the future and linear in the past. Recently, many attempts have been made to define logics for BDI agents by using extensions of CTL. In this paper, we discuss BDI^{ATL} that is obtained by substituting ATL* (Alternating-Time Temporal Logic) to CTL* in Rao and Georgeff's logic. One of the main advantages of our approach is that in BDI^{ATL} we can express new commitment strategies that are more realistic than those proposed by Rao and Georgeff (and that could not be defined in their logic), since they take collaboration among agents into account. In particular, in this paper we discuss three variants of Rao and Georgeff's "open minded" commitment: "independent open minded", "optimistic open minded", and "pessimistic open minded", whose definition exploits the new features that ATL* adds to CTL*.

1 Introduction

Many logics for modelling beliefs, desires and intentions of agents, such as Rao and Georgeff's BDI logic [16, 14, 15] and Wooldridge's $\mathcal{LORA}$ [20], are based on temporal logics in which the structure of time is branching in the future and linear in the past like CTL (Computational Tree Logic) [7, 5]. CTL temporal connectives can be used to specify properties on the tree-like structure of the corresponding models. Rao and Georgeff's logic provides temporal modalities, like *next*, *eventually* and *always* as well as modalities to describe beliefs, goals (desires), and intentions. The semantics of the logic is given through rich Kripke models in which each world has an internal branching structure to represent the evolution of a given state of mind. The rational balance among these states of mind is guaranteed by means of a set of relations (e.g. *Strong Realism*: $\mathsf{Des}(i, \varphi) \supset \mathsf{Bel}(i, \varphi)$, or, equivalently, $\neg\mathsf{Bel}(i, \varphi) \supset \neg\mathsf{Des}(i, \varphi)$)

$\mathcal{LORA}$ is a refinement of both Rao and Georgeff's BDI logic and Cohen and Levesque's intention logic [6]; from the former it inherits the use of intentions as a distinct state of mind, whereas from the latter it inherits a complete set of operators to manage actions. Moreover, differently from both the previous logics aimed at specifying the mental state of *individual agents*[1], $\mathcal{LORA}$ is a logic for *multiagent systems*, and thus it includes operators suitable for modelling communication and cooperation among agents. In recent years several attempts of defining an alternative approach to modelling multiagent systems have been made (see for example the recent survey by Goranko and Jamroga, [8]). A first step in this direction is represented by the encoding of Rao and Georgeff's BDI logic in the μ-calculus presented in [17]. In [19] the authors propose to adopt ATL (Alternating-Time Temporal Logic, [13]) as the reference logic for reasoning about multiagent systems and epistemic knowledge.

ATL provides a form of selective quantification over paths of a tree-like Kripke model (*cooperation modalities*), where, instead, CTL permits only the existential and universal quantification

[1]Extensions of the Rao and Georgeff's logic to cope with multiagent systems exist, but they do not include operators to manage the cooperation and team work.

over paths. As an example, ATL can express properties like *there exists a strategy for a subset of agents to reach a given goal*.

ATL has been developed in the context of hardware verification [12] for modelling open systems. Here the term *open* means that the model takes also care of the environment, that can influence the behaviour of the overall system directly. Usually the environment is represented as the other components. ATL* extends ATL by allowing a more flexible way of composing modalities. In [10], besides presenting a complete study of many variants of the ATL logic and its relationships with BDI logics, Jamroga states that

> *ATL and coalition games can provide* BDI *models with a finer-grained structure of action (simultaneous choices). Furthermore, the cooperation modalities can be "imported" into the* BDI *framework to enable modeling, specifying and verifying agents' strategic abilities in the context of their beliefs, desires and intentions.*

Following this line of thoughts, in this paper we present an ATL-based logic for modelling cooperative BDI-agents, we will refer to as BDI^{ATL}. Our work on BDI^{ATL} represents a preliminary step toward the realisation of Jamroga's intuition.

BDI^{ATL} is born from our effort to find a BDI logic suitable for modelling the behaviour of agents structured according to the "CooBDI" architecture [1].
CooBDI is characterised by a built-in mechanism for retrieving plans from cooperative agents, for example when no local plans suitable for achieving a certain desire are available. In particular, the cooperation strategy of an agent includes the set of agents with which is expected to cooperate (its partner agents, or its "friends"). BDI^{ATL} allows us to express new commitment strategies that are more realistic than those proposed by Rao and Georgeff (and that could not be defined in their logic), since they take collaboration among agents into account. In particular, in this paper we discuss three variants of Rao and Georgeff's "open minded" commitment: "independent open minded", "optimistic open minded", and "pessimistic open minded". In these commitment strategies we exploit the new feature that ATL* adds to CTL*, namely *cooperation modalities*, to express the way of thinking of rational agents.

Thus, we are currently tackling the problem of a well-founded modelling and implementation of cooperative BDI-agents from both the theorical side (the logical foundations given by means of BDI^{ATL}), and the practical one (the CooBDI architecture and its implementations "Coo-AgentSpeak" [2] and "CooWS" [3]). It is well known that in the BDI context, filling the gap between theory and practice is still an open problem and we do not claim that we are going to solve it. However, although the gap between CooBDI and BDI^{ATL} is far from being filled, we hope that adopting a mixed top-down and bottom-up approach to the problem will help us to converge towards a (partial) solution, or at least towards a better comprehension of what "cooperative BDI agents" are.

To the best of our knowledge, there are no previous attempts to extend Rao and Georgeff's BDI logic by integrating ATL into it, although lots of work has been done on similar issues. For example, in [9] Jamroga, van der Hoek and Wooldridge present an intention logic (inspired to the work of Cohen and Leveque) based on ATL, and the same authors, in [18], discuss CATL, an extension of ATL which supports reasoning about the abilities of agents and their coalitions in game-like multi-agent systems.
The paper is structured in the following way: Section 2 describes the syntax and semantics of BDI^{ATL}, Section 3 discusses some of the advantages of using ATL* instead of CTL* inside a BDI logic, and Section 4 concludes the paper. The ideas behind the proof that Rao and Georgeff's multiagent logic (BDI^{CTL}, from now on) can be embedded into BDI^{ATL} are summarised in the Appendix.

2 BDI^{ATL}: Syntax and Semantics

BDI^{ATL} is based on ATL and includes three modalities to represent the mental attitudes of agents. Its syntax is nearly identical to the syntax of BDI^{CTL}, apart from quantification over paths.

2.1 Syntax

As in the normal CTL based logic there are two types of formulae: *state formulae* and *path formulae*. *State formulae* are evaluated in a given situation. They are defined as follows:

- any first-order formula is a state formula;
- if φ_1 and φ_2 are state formulae and x is a variable, then also $\neg\varphi_1$, $\varphi_1 \vee \varphi_2$, $\exists x\, \varphi_1$ are state formulae;
- if φ is a state formula and i is an agent then also $\mathsf{Bel}(i, \varphi)$, $\mathsf{Des}(i, \varphi)$ and $\mathsf{Int}(i, \varphi)$ are state formulae;
- **[New w.r.t. BDI^{CTL}]** if ψ is a path formula and A is a group (set) of agents, then $\langle\langle A\rangle\rangle\psi$ is a state formula.

Path formulae are evaluated along a given path in a given world. They are defined as follows:

- any state formula is also a path formula;
- if ψ_1 and ψ_2 are path formulae then also $\neg\psi_1$, $\psi_1 \vee \psi_2$, $\exists x\, \psi_1$, $\psi_1 \mathcal{U} \psi_2$, and $\bigcirc\psi_1$ are path formulae.

2.2 Semantics

Since the semantics of ATL formulae is given using concurrent game structures, the semantics of BDI^{ATL} has to be given using concurrent game structures too. While in Kripke structures a state transition represents a step of a closed system, in a concurrent game structure (CGS), a state transition results from choices made by the system components and by the environment and represents simultaneous steps by the components and environment. A *concurrent game structure* for BDI^{ATL} can be defined as

$$S = \langle D_{Ag}, W, T, D_U, D_{Ac}, \Phi, d, \delta, \mathcal{B}, \mathcal{D}, \mathcal{I}\rangle$$

where:

- D_{Ag} is the set of agents;
- W is the set of worlds, $W = \langle T', \delta'\rangle$ where $T' \subseteq T$ and δ' is the same of δ but it is restricted to the time points in T';
- T is the set of time points;
- D_U is the universe of discourse;
- D_{Ac} is the set of actions;
- Φ is a mapping of first order entities to elements of D_U for any given world and any given time point;
- d is a mapping $D_{Ag} \times T \to \wp(D_{Ac})$ such that for each player $a \in K$ and each time point t, $d_a(t)$ defines the set of moves available to agent a in time point t.
  For each time point t a *move vector* is a tuple $\langle j_1, \ldots, j_k\rangle$ such that $j_a \in d_a(t)$ for each agent a. The *move function* $D : T \to (\prod_{a\in D_{Ag}} D_{Ac})$ is such that each $t \in T$ is mapped to $(\prod_{a\in D_{Ag}} d_a(t))$;
- δ is the transition function $\delta : T \times \prod_{a\in D_{Ag}} D_{Ac} \to T$;
- $\mathcal{B}$ is the belief-accessibility relation ($\mathcal{B} : D_{Ag} \to \wp(W \times T \times W)$).
- $\mathcal{D}$ is the desire-accessibility relation ($\mathcal{D} : D_{Ag} \to \wp(W \times T \times W)$).
- $\mathcal{I}$ is the intention-accessibility relation ($\mathcal{I} : D_{Ag} \to \wp(W \times T \times W)$).

The semantics of path formulae is given as in BDI^{CTL}. We use the path formula satisfaction relation $\models_{\mathcal{P}}$; a path formula interpretation is a tuple $\langle M, v, w, \lambda\rangle$, where M is a model, v is a variable assignment, w is a world, and λ is a computation.
Also the semantics of the state formulae is given as in BDI^{CTL}, apart from the semantics of the path quantification. In order to define it, we use the state formula satisfaction relation $\models_{\mathcal{S}}$, a state formula interpretation, a tuple $\langle M, v, w, t\rangle$, where M is a model (CGS), v is a variable assignment, w is a world, and t is a time point, and some auxiliary definitions given below.

successor Let t_0 and t_1 be time points, then t_1 is the successor of t_0 if $\exists j \in D(t_0).\, \delta(t_0, j) = t_1$.

computation An infinite sequence of time points $\lambda = t_0, t_1, t_2, t_3, \ldots$ is a computation if $\forall i \geq 0\; t_{i+1}$ is a successor of t_i. A computation starting from the time point t is called t-computation. Let $\lambda = t_0, t_1, t_2, t_3, \ldots$ be a computation, then $\lambda[i]$ is the i_th time point of λ, $\lambda[0, i]$ is the finite prefix $t_0, \ldots, t_i$, and $\lambda[i, \infty]$ is the infinite suffix $t_i, t_{i+1}, t_{i+2}, \ldots$.

strategy A strategy is a function $f_a(\lambda t) \in d_a(t)$ where $a \in D_{Ag}$, and $\lambda \in T^*$; a strategy for a particular agent represents the action chosen by the agent in a particular situation. It is important to note that agent's decisions depend on the history of the computation.

set of strategies A set of strategies is defined as $F_A = \{f_a \mid a \in A\}$ where $A \subseteq D_{Ag}$; a set of strategies F_A is the set of the strategies of a group of agents A.

outcomes The outcomes of F_A from t ($out(t, F_A)$) is the set of t-computations that the agents in A can enforce. Formally, $\lambda = t_0, t_1, t_2, t_3, \ldots \in out(t, F_A)$ if and only if $t = t_0$ and for all $i \geq 0$ there is a move vector $j = \langle j_0, \ldots, j_{|D_{Ag}|}\rangle \in D(t_i)$ such that for all $a \in A$ $j_a = f_a(\lambda[0, i])$ and $\delta(t_i, j) = t_{i+1}$.

Semantics of the path quantification. We define the semantics of path quantification by means of the $\models_{\mathcal{S}}$ relation that states when $\langle\langle A\rangle\rangle\psi$ is satisfied in $\langle M, v, w, t\rangle$.

- $\langle M, v, w, t\rangle \models_{\mathcal{S}} \langle\langle A\rangle\rangle\psi$ iff there exists a set F_A of strategies, such that for all t-computations $\lambda \in out(t, F_A) \quad \langle M, v, w, \lambda\rangle \models_{\mathcal{P}} \psi$.

As usual, many useful operators can be derived from the initial set such as the path quantifier $[\![\ ]\!]$, that is the dual operator of $\langle\langle\ \rangle\rangle$ and expresses the impossibility of avoidance. If the agents in A can enforce a set Λ of computations, then the agents in $D_{Ag} \setminus A$ cannot avoid Λ. Therefore, $\langle M, v, w, t\rangle \models_{\mathcal{S}} \langle\langle A\rangle\rangle\psi \Rightarrow \langle M, v, w, t\rangle \models_{\mathcal{S}} [\![D_{Ag} \setminus A]\!]\psi$. The converse of this statement is not necessarily true.
To encode BDI^{CTL} in BDI^{ATL} we follow the well-know encoding of CTL in ATL, and we translate the path existential quantification E into the operator for the cooperation of all the agents ($\mathsf{E}\psi \rightarrow \langle\langle D_{Ag}\rangle\rangle\psi$) and the path universal quantification A into the operator for the cooperation of the empty set of agents ($\mathsf{A}\psi \rightarrow \langle\langle\emptyset\rangle\rangle\psi$)[2]. To complete the encoding it is necessary to deal with some other differences: one is that CTL is conceived to represent asynchronous turn based games, instead ATL is thought to manage synchronous multiplayer games. An overview of the ideas behind the encoding can be found in the Appendix.

3 New commitment strategies

Commitment strategies are a central concept in the rationality of an agent, because the agent's behaviour depends on his intentions, and his intentions depend on his commitment strategies. In Rao and Georgeff's work, three commitment strategies have been proposed: blind, single minded, open minded.

[2] $\mathsf{E}\psi$ and $\mathsf{A}\psi$ will hereafter be used (inside of the ATL formulae) as a shorthand for $\langle\langle D_{Ag}\rangle\rangle\psi$ and $\langle\langle\emptyset\rangle\rangle\psi$ respectively.

Blind commitment. A blindly committed agent is an agent who maintains his intentions until he believes that he has achieved them.

$$\mathsf{Int}(\mathsf{A}\Diamond\varphi) \supset \mathsf{A}(\mathsf{Int}(\mathsf{A}\Diamond\varphi)\,\mathcal{U}\,\mathsf{Bel}(\varphi))$$

The blind commitment strategy is obviously too strong, since it does not permit to an agent to give up from one of his intentions. By modifying the "give up" requirement, it is possible to define the single minded commitment strategy, that can be further refined resulting into the open minded commitment strategy.

Single minded commitment. A single minded committed agent is an agent who maintains his intentions as long as he believes that they are still options.

$$\mathsf{Int}(\mathsf{A}\Diamond\varphi) \supset \mathsf{A}(\mathsf{Int}(\mathsf{A}\Diamond\varphi)\,\mathcal{U}\,(\mathsf{Bel}(\varphi) \vee \neg\mathsf{Bel}(\mathsf{E}\Diamond\varphi)))$$

If we analyse the single minded commitment strategy, we can identify some problems in the path quantification E of the last disjunction. In fact, it may seem irrational that an agent pursues an intention that he believes achievable only with the collaboration of all agents. In this case an agent can discard an intention only if he doesn't believe that there is at least one path in which his intention will become true. The path quantification E forces the agent to maintain his intention as long as he believes that there is at least one possibility to satisfy it, doesn't matter how much complex the system is or how many agents have to cooperate to achieve it.

Open minded commitment. An open minded committed agent is an agent who maintains his intentions as long as they are still goals.

$$\mathsf{Int}(\mathsf{A}\Diamond\varphi) \supset \mathsf{A}(\mathsf{Int}(\mathsf{A}\Diamond\varphi)\,\mathcal{U}\,(\mathsf{Bel}(\varphi) \vee \neg\mathsf{Des}(\mathsf{E}\Diamond\varphi)))$$

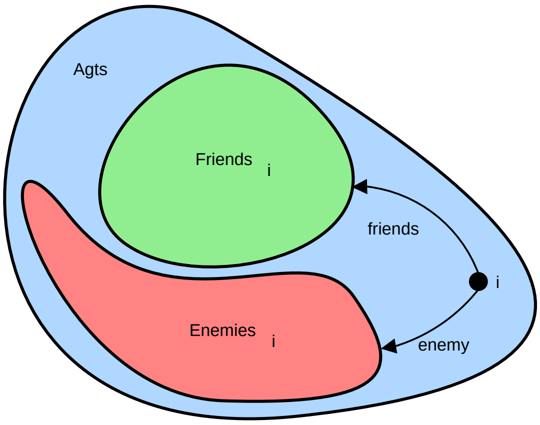
If we analyse the open minded commitment strategy, we can notice that this new kind of commitment overcomes the problem which incurs the previous one, but unfortunately the way used to solve the problem doesn't really remove it. The use of Des in place of Bel permits to an agent to discard an intention if it doesn't desire to achieve it, and this allows agents to perform reconsideration [4] and then represents a step forward towards rationality. However, if we assume that a strong realism relation holds between desires and intentions, it may happen that an open minded agent behaves exactly as a single minded one, thus making the effort of defining a more rational commitment strategy vain.

Trying to avoid the strong realism assumption is not the right direction in order to define an agent with a rational behaviour. On the one hand, imposing a relation between desires and beliefs stronger than the strong realism one leads to an irrational agent's behaviour, as demonstrated by [20]. On the other hand, adopting a weaker relation between mental modalities (for example, the *Weak Realism with Option* defined as $(\mathsf{Bel}(i, \mathsf{E}\varphi) \supset \neg\mathsf{Des}(i, \neg\mathsf{E}\varphi)) \wedge (\mathsf{Des}(i, \mathsf{E}\varphi) \supset \neg\mathsf{Bel}(i, \neg\mathsf{E}\varphi))$), together with an open minded commitment strategy, may allow an agent to behave in a definitely unrealistic way, such as going on to desire to reach a state where φ is true, while believing that φ will never become true.

What happens is that assuming more sophisticated and realistic relations between the agent's states of mind of an open minded agent, leads to a less rational behaviour of the agent himself. This means that, in order to save both the open mindedness of the agent and his rational behaviour, working on the relations between the agent's states of mind is not enough. It is indeed necessary to define new commitment strategies that are based on the open minded one, that lead to rational behaviour, and that are not as sensible to the changes in the assumptions on the relations between mental modalities as the Rao and Georgeff's open minded commitment strategy.

One advantage of using BDI^{ATL} and, in particular, of using quantification over paths and groups of agents, is that we can express more realistic commitment strategies without requiring that *all* the agents in the system cooperate to the achievement of an agent's intention. By using the path quantification we can modify the "discard condition" that permits to an agent to reject an intention, and we can define commitment strategies that meet our requirements of leading to a rational behaviour based on what the agent thinks about the cooperation inside the system, no matter what agent desires about the cooperation itself.

As instance we present three variants of the "open minded commitment" strategy: "independent open minded", "optimistic open minded", and "pessimistic open minded". We suppose that, from the point of view of an agent i, the set of all the agents that belong to the MAS ($Agts$ in the drawing) contains both agents that i considers friends ($Friends_i$), and agents that i considers enemies ($Enemies_i$). It is rational to assume that these two sets are disjoint.



- $Agts \overset{def}{\equiv} D_{Ag}$
- $i \in D_{Ag}$
- $Friends_i \subseteq D_{Ag}$
- $Enemies_i \subseteq D_{Ag}$
- $Friends_i \cap Enemies_i = \emptyset$
- $i \notin Enemies_i$
- $i \notin Friends_i$

Independent open minded.

$$\mathsf{Int}(i,\ \mathsf{A}\Diamond\varphi) \rightarrow \mathsf{A}(\mathsf{Int}(i,\ \mathsf{A}\Diamond\varphi)\,\mathcal{U}\,(\mathsf{Bel}(i,\ \varphi) \vee \neg\mathsf{Des}(\mathsf{E}\Diamond\varphi) \vee \neg\mathsf{Bel}(i,\ \langle\langle\{i\}\rangle\rangle\Diamond\varphi)))$$

If an agent uses this commitment strategy he preserves only the intentions that he believes have not been achieved yet, that he continues to desire, and that he believes can be achieved even if all the other agents try to avoid his objective. By adopting this strategy, an agent pursues only the intentions that he thinks to be achievable, no matter what the other agents do.

Optimistic open minded.

$$\mathsf{Int}(i,\ \mathsf{A}\Diamond\varphi) \rightarrow \mathsf{A}(\mathsf{Int}(i,\ \mathsf{A}\Diamond\varphi)\,\mathcal{U}\,(\mathsf{Bel}(i,\ \varphi) \vee \neg\mathsf{Des}(\mathsf{E}\Diamond\varphi) \vee \neg\mathsf{Bel}(i,\ \langle\langle D_{Ag} \setminus Enemies_i\rangle\rangle\Diamond\varphi)))$$

In this strategy we use the quantification over groups of agents to express an optimistic way of thinking. An agent with this commitment strategy maintains an intention until he believes that this intention is achieved or he desires to achieve it or he doesn't believe that the collaboration of all the agents, but his enemies, is enough to achieve this intention.

Pessimistic open minded.

$$\mathsf{Int}(i,\ \mathsf{A}\Diamond\varphi) \rightarrow \mathsf{A}(\mathsf{Int}(i,\ \mathsf{A}\Diamond\varphi)\,\mathcal{U}\,(\mathsf{Bel}(i,\ \varphi) \vee \neg\mathsf{Des}(\mathsf{E}\Diamond\varphi) \vee \neg\mathsf{Bel}(i,\ \langle\langle\{i\} \cup Friends_i\rangle\rangle\Diamond\varphi)))$$

An agent with this commitment strategy maintains an intention until he believes that this intention is achieved or he desires to achieve it or he doesn't believe that he can force the achievement of the intention even with the help of all his friendly agents.

In all the three new commitment strategies, we have added a condition to the "until" part of the basic "open minded commitment" strategy; the added condition represents an information that cannot be derived by the *Strong Realism* applied to definition of the "single minded commitment" strategy[3]. In fact, the following formulae hold:

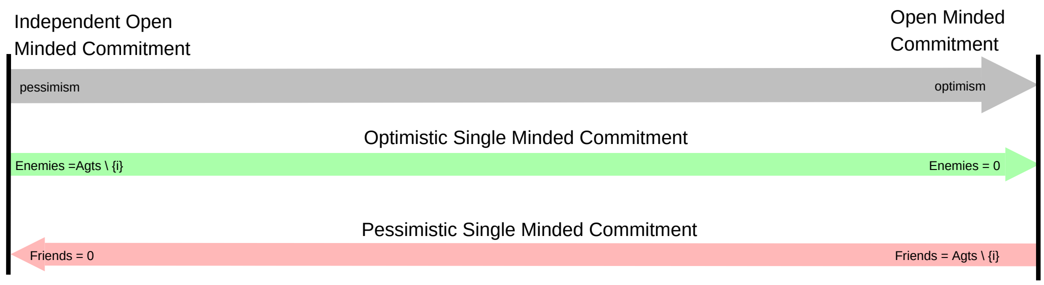
$$\begin{array}{rcll}
\neg\mathsf{Bel}(i,\ \mathsf{E}\Diamond\varphi) & \supset & \neg\mathsf{Des}(i,\ \mathsf{E}\Diamond\varphi) & \text{Strong Realism.}\\
\neg\mathsf{Bel}(i,\ \langle\langle\{i\}\rangle\rangle\Diamond\varphi) & \not\supset & \neg\mathsf{Des}(i,\ \mathsf{E}\Diamond\varphi) & \text{If there is more than one agent.}\\
\neg\mathsf{Bel}(i,\ \langle\langle D_{Ag} \setminus Enemies_i\rangle\rangle\Diamond\varphi) & \not\supset & \neg\mathsf{Des}(i,\ \mathsf{E}\Diamond\varphi) & \text{If } Enemies_i \neq \emptyset.\\
\neg\mathsf{Bel}(i,\ \langle\langle\{i\} \cup Friends_i\rangle\rangle\Diamond\varphi) & \not\supset & \neg\mathsf{Des}(i,\ \mathsf{E}\Diamond\varphi) & \text{If } \{i\} \cup Friends_i \neq D_{Ag}.
\end{array}$$

[3]On the contrary, the "open minded commitment" strategy is derived by the "single minded commitment" strategy with the assumption of a *Strong Realism* relation between mental modalities.

The added condition represents what the agent thinks about the cooperation inside the system, and then it is unrelated to what the agent desires.

At a first glance, it might seem that the "optimistic open minded" strategy is not that optimistic. If an agent adopts this commitment strategy he implicitly thinks that all the enemy agents will do everything in their possibility to inhibit the achievement of his intentions. On the other hand, the "pessimistic single minded" strategy might seem too optimistic, because if an agent adopts this commitment strategy, he implicitly thinks that there is the possibility that all the friendly agents will help his to achieve his intentions.

The balance between the optimistic and pessimistic approaches has to be found in the use of the cooperation modalities. The optimistic approach uses the set $Enemies_i$ to identify all the agents that could help to force a computation. The pessimistic approach uses the set $Friends_i$ to identify all the agents that could help to achieve the objective.



In the figure above we see how the new commitment strategies can be put into a strict relation with the "open minded" one. Previously we have stated that the original strategy seems, in some sense, too optimistic, now we can see that the "independent open minded" has exactly the opposite problem and finally that the other two strategies cover all the "rational space" between these two extremes. The "optimistic open minded" converges either to the "independent open minded" when the agent thinks that all the other agents are enemies or to the "open minded" when the agent thinks that he doesn't have any enemy. For the "pessimistic open minded" we can identify a very similar behaviour related to the $Friends$ set.

4 Conclusions and future work

In this paper we have introduced a new BDI logic, BDI^{ATL}, focussing our attention on its temporal aspects. The relationships among the mental modalities have not been analysed because all the work done by Wooldridge for his $\mathcal{LORA}$ logic is still valid for BDI^{ATL} too (see Appendix C of [11]).

Temporal logic comes in three varieties: LTL (Linear-time Temporal Logic), that assumes implicit universal quantification over all paths that are generated by the execution of the system; CTL (Computational Tree Logic), that allows explicit existential and universal quantification over all paths, and ATL (Alternating-time Temporal Logic), that offers selective quantification over those paths that are possible outcomes of games. All the formulae representable in LTL are also representable in CTL*, and the formulae representable in CTL* are also representable in ATL*.

We started from this well-known assumption to extend the Rao and Georgeff's BDI logic by substituting its CTL* foundation with a new ATL* foundation. This new logic BDI^{ATL} can be used to represent all the formulae that are representable in BDI^{CTL}, moreover it is possible to take advantage of the greater expressive power to express new formulae.

Although in this paper we have only analysed the advantages of our approach with respect to the definition of some realistic commitment strategies, BDI^{ATL} has many other interesting features that can be exploited, and that we are currently analysing. In fact, beyond the new possible paths quantification, the ATL foundation of BDI^{ATL} permits to manage the contemporaneous execution of more than one action. Taking advantage of this feature, it would be possible to represent the environment like any other agents. The result of the execution of an action by a *normal* agent would then become the combination of the execution of this action with the action chosen by the environment agent. The *environment* agent could be used to solve another chronic problem of the

BDI logic: in nearly all BDI logics time goes on only if an action/event occurs, consequently in a stationary phase (where all agents wait for something) the line-time is stopped. To overcome this situation, it would be sufficient that the environment agent executed periodically a sort of *no-op* action leading to a time point transition. This would be enough to force the time to go ahead.

Our main directions of work are then to analyse the possibility to define more and more refined versions of the commitment strategies, and to analyse how to explicitly deal with the environment agent, taking the outcomes of recent events on this topic (for example, the Environment Technical Forum held in Budapest in September 2005) as guidelines for our work. Since BDI^{ATL} seems also suitable for representing the team work and team formation, we also plan to investigate the possible interaction among "collective mental states" [20] and cooperative modalities.

References

[1] D. Ancona and V. Mascardi. Coo-BDI: Extending the BDI Model with Cooperativity. In J. A. Leite, A. Omicini, L. Sterling, and P. Torroni, editors, *Proceedings of the First Declarative Agent Languages and Technologies Workshop (DALT'03), Revised Selected and Invited Papers*, pages 109–134. Springer-Verlag, 2004. LNAI 2990.

[2] D. Ancona, V. Mascardi, J. F. Hübner, and R. H. Bordini. Coo-AgentSpeak: Cooperation in AgentSpeak through Plan Exchange. In N. R. Jennings, C. Sierra, L. Sonenberg, and M. Tambe, editors, *Proceedings of AAMAS 2004 (Int. Conf. on Autonomous Agents and Multiagent Systems)*, pages 698–705. ACM press, 2004.

[3] L. Bozzo, V. Mascardi, D. Ancona, and P. Busetta. CooWS: Adaptive BDI Agents Meet Service-Oriented Computing (Extended Version). In this volume.

[4] M. E. Bratman. *Intention, Plans and Practical Reason*. Harvard University Press, Nov 1987.

[5] E. M. Clarke and E. A. Emerson. Design and Synthesis of Synchronization Skeletons Using Branching-Time Temporal Logic. In *Logic of Programs*, pages 52–71, 1981.

[6] P. R. Cohen and H. J. Levesque. Intention is Choice with Commitment. *Artificial Intelligence*, 42(3):213–261, 1990.

[7] E. A. Emerson and J. Y. Halpern. "Sometimes" and "not never" revisited: on branching versus linear time temporal logic. *J. ACM*, 33(1):151–178, 1986.

[8] V. Goranko and W.J. Jamroga. Comparing semantics of logics for multi-agent systems. *Synthese*, 139(2):241–281, 2004.

[9] W. Jamroga, W. van der Hoek, and M. Wooldridge. Intentions and Strategies in Game-Like Scenarios. In *3rd Workshop on Multi-Agent Systems: Theory and Applications (MASTA 2005)*, 2005.

[10] W. J. Jamroga. *Using Multiple Models of Reality. On Agents who Know how to Play Safer*. PhD thesis, University of Twente, January 2004.

[11] R. Montagna. Teoria e modellazione logica degli agenti bdi cooperativi. Master's thesis, DISI – Università di Genova, 2004. In English.

[12] T. A. Henzinger R. Alur and O. Kupferman. Alternating-Time temporal Logic. In *Proceeding of FOCS*, pages 100–109. IEEE Computer Society Press, 1997.

[13] T. A. Henzinger R. Alur and O. Kupferman. Alternating-Time temporal Logic. In *Journal of the ACM 49*, pages 672–713. 2002.

[14] A. S. Rao and M. P. Georgeff. Asymmetry Thesis and Side-Effect Problems in Linear-Time and Branching-Time Intention Logics. In John Myopoulos and Ray Reiter, editors, *Proceedings of the 12th International Joint Conference on Artificial Intelligence (IJCAI-91)*, pages 498–505. Morgan Kaufmann publishers Inc. San Mateo, CA, USA, 1991.

[15] A. S. Rao and M. P. Georgeff. Deliberation and Intentions. In *Proceeding of 7th Conference on Uncertainity in Artificial Intelligence, Los Angeles*. Morgan Kaufmann publishers Inc. San Mateo, CA, USA, May 1991.

[16] A. S. Rao and M. P. Georgeff. Modelling Rational Agents within a BDI-Architecture. In *Proceeding of the Second International Conference of Principles of Knowledge Representation and Reasoning*. Morgan Kaufmann publishers Inc. San Mateo, CA, USA, February 1991.

[17] K. Schild. On the Relationship Between BDI Logics and Standard Logics of Concurrency. *Autonomous Agents and Multi-Agent Systems*, 3(3):259–283, 2000.

[18] W. van der Hoek, W. Jamroga, and M. Wooldridge. A logic for strategic reasoning. In *Proc. of the 4th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'05)*. ACM Press, 2005.

[19] W. van der Hoek and M. Wooldridge. Tractable multiagent planning for epistemic goals. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems*, pages 1167–1174. ACM Press, 2002.

[20] M. Wooldridge. *Reasoning about Rational Agents*. Mit press (Cambridge, Massachussetts/London, England), June 2000.

Appendix

In this appendix we discuss the main ideas behind the proof that BDI^{CTL} can be embedded in BDI^{ATL}. The complete demonstration can be found in [11]. The demonstration is carried out according with this sequence of steps:

1. We demonstrate that for each formula of BDI^{CTL} logic there is a syntactical equivalent formula in BDI^{ATL} logic by means of the function *transl* : BDI^{CTL} $\rightarrow$ BDI^{ATL}
2. We demonstrate that for each model of BDI^{CTL} there is an equivalent model of BDI^{ATL}.
3. We present some formulae of BDI^{ATL} that are not representable in BDI^{CTL}.

The third point demonstrates that BDI^{ATL} is strictly more expressive than BDI^{CTL}.

Syntax. From a syntactic point of view, BDI^{CTL} and BDI^{ATL} formulae are almost identical, apart from some differences in the path quantifications. The path existential quantification in BDI^{CTL} corresponds to the collaboration of all the agents in BDI^{ATL}; in the general case, in fact, a single path can be enforced only by a common strategy of all the agents. Since the final outcome of the system depends on the interaction among the agents choices and the behaviour/reaction of the environment, we have also to take into account the agentification of the environment.
The path universal quantification in BDI^{CTL} corresponds to the "null" collaboration BDI^{ATL}; every possible computation is acceptable, no matter what the other agents and the environment do (or do not). Thus, every agent (and the environment too) can follow his strategy. Formally, the considerations above can be expressed by means of the *transl* function as follows:

$$\begin{aligned} transl(\mathsf{E}\psi) &\stackrel{def}{=} \langle\langle D_{Ag}\rangle\rangle transl(\psi) \\ transl(\mathsf{A}\psi) &\stackrel{def}{=} \langle\langle \emptyset \rangle\rangle transl(\psi) \end{aligned}$$

Semantics. In order to encode the semantics of BDI^{CTL} into BDI^{ATL}, we have to take into account the fact that BDI^{CTL} was thought to deal with asynchronous turn based games, while BDI^{ATL} was thought to manage synchronous multiplayer games.
Intuitively, we have to guarantee that at each time point, the system behaves as if only one agent executed an action (or an event occured). The idea is to use the functions d to enforce that at least one agent acts at each time point and to use the function δ to manage the situations in which more than one action/event occur.

To guarantee that these conditions hold, we need to assume a complete order inside of the set of agents (D_{Ag}) and the existence of the "null" action $\varepsilon \in D_{Ac}$. In the following we define both d and δ in an informal way based on natural language; their formal definitions can be found in [11].

The function d is defined as:

$a \in d(i,t)$	if in BDI^{CTL} the execution of the action a by the agent i forces a transition starting from t	(1)
$\varepsilon \in d(i,t)$	if in BDI^{CTL} there is an action a that can be executed by an agent j, with $j > i$ and this action force a transition starting from t	(2)
	$\vee$	
	if in BDI^{CTL} there is not any an action a that can be executed by the agent i at time point t	(3)

The combination of the first and third conditions guarantees that every agent at every time point has at least one action to execute. The second condition permits to all the agents that can execute an action (but the last one in the complete order) to choose to stand still.

The function δ^{ATL} is defined as:

$\delta^{ATL}(t, \langle j_1, \ldots, j_{|D_{Ag}|}\rangle) \stackrel{def}{=} t'$ iff there is an agent n s.t. $j_n = a$ and for all the agents m $(m < n)$ $j_m = \varepsilon$ and in BDI^{CTL} the execution of the action a by the agent n force the transition from t to t'

The transition function δ^{ATL} takes into account only the first non-null action executed (respect to the agent's order). Using d and δ^{ATL} it is possible to create from every possible BDI^{CTL} model M, a BDI^{ATL} model M' that is equivalent to M by choosing, for the BDI^{ATL} model, the same set of time points of the BDI^{CTL} model. This choice is necessary because in BDI^{CTL}, from each time point there is at least one transition that starts and then the d function returns a set of vector moves in which every element contains at least one non-null action.

Finally we can conclude that if we apply the δ function to a vector move returned by the d function, it returns the time point related to the execution of the first action of the vector move; then the transition encoded is exactly equal to a transition of the BDI^{CTL} model.

Expressiveness. We have demonstrated that all the BDI^{CTL} formulae are representable in BDI^{ATL}, then the new logic is expressive at least as the BDI^{CTL}. Moreover, there are many BDI^{ATL} formulae that represent situations that are not representable in BDI^{CTL}. For example, if $D_{Ag} = \{a, b, c\}$, the formula $\langle\langle\{a, b\}\rangle\rangle\varphi$ can neither be reduced to $\mathsf{A}\varphi$ nor to $\mathsf{E}\varphi$; in fact, it is less restrictive than the first, and more restrictive than the second.