

Actions Made Explicit in BDI

Vineet Padmanabhan¹ Guido Governatori² Abdul Sattar¹

¹ Knowledge Representation & Reasoning Unit (KRRU)
School of Information Technology
Griffith University, Gold Coast Campus, Queensland, Australia.
{vineet,sattar}@cit.gu.edu.au
² Cooperative Information System Research Centre
Queensland University of Technology, Queensland, Australia
g.governatori@qut.edu.au

Abstract. The Belief, Desire, Intention (*BDI*) architecture is increasingly being used in a wide range of complex applications for agents. Many theories and models exist which support this architecture and the recent version is that of *Capability* being added as an additional construct. In all these models the concept of action is seen in an endogenous manner. We argue that the *Result* of an action performed by an agent is extremely important when dealing with composite actions and hence the need for an explicit representation of them. The Capability factor is supported using a RES construct and it is shown how the components of a composite action is supported using these two. Further, we introduce an OPP (opportunity) operator which in alliance with *Result* and *Capability* provides better semantics for practical reasoning in BDI.

1 Introduction

A paradigm shift is happening in both Artificial Intelligence and mainstream computer science with the advent of agents and agent-oriented approaches to developing systems, both on a theoretical and practical level. One such approach called *BDI* takes mental attitudes like *Belief*, *Desire* and *Intention* as the primitives and has given rise to a set of systems called *Intentional Agent Systems* [2, 5, 7, 9]. Of these the one by Rao and Georgeff [13] has been widely investigated due to its strong links with theoretical work. Many modifications have been made since the initial work, the most recent being the addition of a *Capability* [11] construct along with the three primitive modalities. In all these systems the concept of action is seen in an endogenous manner. Though it is possible to come up with accounts of action without representing them explicitly, many problems that plague endogenous formalisations can be avoided in exogenous ones. The later work by Rao [12] makes this shift but then it is restricted to the planning domain.

This paper can be viewed as a further extension of the existing BDI theory whereby we reason about the mental state of an agent during the execution of an action in an exogenous way. We investigate the close connection between the *result* of an action performed by a BDI agent and its capability of achieving that result. We argue that though the agent might have a capability to perform an action it need not be the case that the *opportunity* should always accompany it. This view gets importance when we

take into consideration composite actions where one action follows the other ($\phi_1; \phi_2$), which means an agent performs ϕ_1 followed by ϕ_2 . In such cases the *result* of the component parts of the action is needed for the overall success of the action. It also seems reasonable to declare that the agent has the relevant *opportunity* to perform the component actions in such a way that the execution leads to an appropriate state of affairs. By making actions explicit in BDI we try to avoid some of the problems that plague the endogenous systems when dealing with composite actions. We describe a formal relationship between the *Result*, *Opportunity*, *Belief*, *Desire* and *Intention* modalities. It is important to note the close connection between *Intention* and *Result*. For instance, if an agent intends to perform a plan, we can infer under certain conditions he intends the result of the plan. Similar is the case with Goals and Results.

This work is partially motivated by the KARO architecture of Van Linder [10] whereby we indicate how *Result* and *Opportunity* can be integrated to the existing BDI framework. Such an addition definitely paves way for a better understanding of the dynamics of BDI Systems. The rest of the paper is organised as follows. In section 2 we make a distinction between *intentional action* (actions with a pre-defined intention) and *intending an action* (actions with future Intention) and claim that composite actions fit well under *actions with future intention*. Section 3 gives a brief summary about the original BDI logic as developed by Rao [13] and the recent version of it with the *Capability* construct [11]. Sections 4 and 5 integrate two new operators RES and OPP with the existing BDI architecture. Section 6 gives the full picture of the new semantics. We have purposefully avoided the use of any temporal operators as it remains part of the future work. In section 7 we formalise the commitment axioms according to the new semantics and the conclusion and future work is depicted in section 8.

2 Intentional Action & Intending an Action

When one takes into account the compositional nature of actions $\phi_1; \phi_2$ (ϕ_1 followed by ϕ_2), it seems contradictory to believe that endogenous logics alone can account for the mental state of an agent during the execution of such actions. The problem with the current formalisms is in their failure to differentiate *Intentional Action* (Predefined Intention) from *Intending to do an Action* (Future Intention). Most of the work in BDI represent actions in the former manner. In the work of Rao [13] formulas like $BEL(\Phi)$, $GOAL(\Phi)$ etc. are used to denote the belief and goal of an agent performing an action ϕ . The formalism remains true for single actions, but when it comes to composite actions like $(\phi_1; \phi_2)$ it fails to do justice as it is taken for granted that the execution of the first action necessarily leads to the second without mentioning anything about the *result* of the first action on the second. Based on the existing BDI architecture the concept of composite actions could be formalised as

$$INT(does(\phi_1; \phi_2)) \Rightarrow does(\phi_1; \phi_2).$$

This need not be the case as the performance of ϕ_1 could result in a counterfactual state of affairs. It seems crucial to consider the *result* of the first action for the overall success of the composite action. In the same manner the formulas like

$$GOAL(\phi_1; \phi_2) \Rightarrow CAP(GOAL(\phi_1; \phi_2))$$

seem to be problematic as the formulation doesn't tell anything about the ability of the agent if the first action results in a counterfactual state of affairs. It doesn't mention anything regarding the *Opportunity* the agent has in performing the second action.

It is important to make a division between the two action constructs of *Intentional* and *Intending* for our framework. The former relates to a predefined intention, where the *Result* of an action is taken for granted, whereas the latter concerns a future intention, where further deliberation is done as to what the result would be before an action is performed. Davidson [6] oversees such a division and extends the concept of *intentionally doing* to that of *intending to*. Though Bratman [1] points out this disparity the current formalisms does not allow for sound representation using the existing modal operators. Hence the need for additional constructs like RES and OPP. In intentional action, there is no temporal interval between what Davidson terms as *all-out evaluation and action*. So there is no room for further practical reasoning in which that all-out evaluation can play a significant role as input. The BDI framework gives primary importance to practical reasoning and hence to means-end reasoning which is important to avoid further deliberation at the time of action. Therefore it seems appropriate to categorise composite actions under future intentions as they play a crucial role in our practical thinking. More importantly, we form future intentions as part of larger plans whose role is to aid co-ordination of our activities over time. As elements in these plans, future intentions force the formation of yet further intentions and constrain the formation of other intentions and plans.

3 The BDI Logic

The logic developed by Rao and Georgeff [13] is based on Computational Tree Logic (*CTL**) [4] extended with a first order variant for the basic logic and a possible-worlds framework for the Belief, Goal and Intention operators. The world is modelled using a temporal structure with a branching time future and a single past called a *time-tree*. A situation refers to a particular time point in a particular world. Situations are mapped to one another by occurrence of events. The branches in a time tree can be viewed as representing the choices available to the agent at each moment in time. There are two kinds of formulae in the logic called the *state formulae* and *path formulae*. The former are evaluated at a specified time point in a time tree and the latter over a specified path in a time tree. Two modal operators *optional* and *inevitable* are used for path formulas. *optional* is said to be true of a path formula Φ at a particular point in a time-tree if Φ is true of at least one path emanating from that point. *inevitable* is said to be true of a path formula Φ at a particular point in a time-tree if Φ is true of all paths emanating from that point. The standard temporal operators \diamond (eventually), \square (always), \bigcirc (next) and \bigcup (until), operate over state and path formulae. These modalities can be combined to describe the options of an agent.

Beliefs, Goals and Intentions are modelled as a set of belief-, goal- and intention accessible worlds associated to an agent in each situation. An agent x has a belief Φ , at a time point t ($BEL(\Phi)$), if Φ is true in all belief-accessible worlds. It is the same case for *goals* ($GOAL(\Phi)$) and *intentions* ($INT(\Phi)$). The logic is based on the concept of strong realism which requires the goals to be compatible with beliefs, and intentions

with goals. This is done by requiring that for every belief-accessible world w at time-point t , there is a desire-accessible world w' at that time point which is a sub-world for w . The converse does not hold as there can be desire-accessible worlds that do not have corresponding belief-accessible worlds. There are similar relationships between goal-accessible and intention-accessible worlds. The axiomatization of beliefs is the standard weak-S5 (or KD45) modal system [8]. The D and K axioms are adopted for goals and intentions, which means that goals and intentions have to be closed under implication and have to be consistent. We are concerned with the semantics of the mental attitudes and the details concerning the possible worlds semantics for various state and path formulae, is given in Appendix A. The set of belief-accessible worlds of an agent x from world w at time t , is denoted by $B_t^w(x)$. Similarly we use $G_t^w(x)$ and $I_t^w(x)$ to denote set of Goal and Intention-accessible worlds of agent x in world w at time t , respectively. When we state the rules and axioms the world w is taken for granted and the formalism is based on the agent, action and time. The semantics for beliefs, goals and intentions can be defined formally as follows

Definition 1 For an interpretation M , with a variable assignment v , a possible world w and a temporal variable t , the semantics for the mental attitudes can be given as:

- $M, v, w_t \models \text{BEL}(\Phi)$ iff $\forall w' \in B_t^w(x), \langle M, v, w'_t \rangle \models \Phi$;
- $M, v, w_t \models \text{GOAL}(\Phi)$ iff $\forall w' \in G_t^w(x), \langle M, v, w'_t \rangle \models \Phi$;
- $M, v, w_t \models \text{INT}(\Phi)$ iff $\forall w' \in I_t^w(x), \langle M, v, w'_t \rangle \models \Phi$.

The rules and axioms depicting the semantic conditions is given as in [13]. The temporal variable t stands for a constant. We do not make any explicit representation of time as it remains part of future work.

Definition 2 Let Φ be a formula, BEL, INT and GOAL be the modal operators for the mental constructs, done, does be the operators for event types, and inevitable be the modal operator for a path formulae; then we have the following axioms:

- A1** $\text{GOAL}(\Phi) \Rightarrow \text{BEL}(\Phi)$
- A2** $\text{INT}(\Phi) \Rightarrow \text{GOAL}(\Phi)$
- A3** $\text{INT}(\text{does}(e)) \Rightarrow \text{does}(e)$
- A4** $\text{INT}(\Phi) \Rightarrow \text{BEL}(\text{INT}(\Phi))$
- A5** $\text{GOAL}(\Phi) \Rightarrow \text{BEL}(\text{GOAL}(\Phi))$
- A6** $\text{INT}(\Phi) \Rightarrow \text{GOAL}(\text{INT}(\Phi))$
- A7** $\text{done}(e) \Rightarrow \text{BEL}(\text{done}(e))$
- A8** $\text{INT}(\Phi) \Rightarrow \text{inevitable} \diamond (\neg \text{INT}(\Phi))$

Axiom A3 seems to be problematic because of the fact that the event e need not be necessarily restricted to a single action. If the agent has a choice of actions at the current time point, he/she would be incapable of acting *intentionally* until she deliberates and chooses one of them. It is the same case when the particular event is a composite action. The agent needs to deliberate on the result of the first action for the successful execution of the second one. It might also be the case that the agent lacks the relevant opportunity at that particular time point of doing the specific action. It becomes more relevant with the addition of the capability construct as given below.

The basic axioms with the capability construct are the same as those given in [11]. The temporal variable has been added in the semantics.

- C1** $CAP(\Phi) \Rightarrow BEL(\Phi)$
- C2** $GOAL(\Phi) \Rightarrow CAP(\Phi)$
- C3** $CAP(\Phi) \Rightarrow BEL(CAP(\Phi))$
- C4** $GOAL(\Phi) \Rightarrow CAP(GOAL(\Phi))$
- C5** $INT(\Phi) \Rightarrow CAP(INT(\Phi))$

The semantic condition of C2 and C3 can be given as follows

Definition 3 Let $C_t^w(x)$ be the set of capability-accessible worlds of agent x in world w at time t .

- $\forall w' \in C_t^w(x), \exists w'' \in C_t^{w'}(x)$ such that $w'' \sqsubseteq w'$;
- $\forall w' \in B_t^w, \forall w'' \in C_t^{w'}(x)$ we have $w'' \in C_t^w(x)$

The first constraint means that for every capability-accessible world w' at time-point t , there is a goal-accessible world w'' at that time-point which is sub-world of w' . The converse doesn't hold as there can be goal-accessible worlds that do not have corresponding capability-accessible worlds. The second constraint is more complicated and deviates from the original interpretation as in [13]. It means that for every belief-accessible world w' at time-point t , all the capability-accessible worlds w'' which is a member of the belief (capability)-accessible worlds w' at time-point t is a member of the capability-accessible worlds w at time-point t , i.e., if the agent has a capability to achieve Φ , then the agent believes that she has such a capability.

4 Integrating Results

The BDI logic and the semantic conditions stated in the previous section shows that the compositional behaviour of actions has not been dealt within the BDI architecture. With the recent addition of the *Capability* construct we believe that it is worthwhile exploring this concept. Whereas the BDI framework is concerned with finding out what it means exactly to have the ability to perform some action, we try to focus on the compositional behaviour of actions. In other terms we are concerned with finding a relation between the capability to perform a composite action and relate it with the capability for the components of that action. Not all actions are treated equally in our approach but instead the result of each action is determined individually and then the conclusion is made whether the agent succeeds in performing that action. Three types of actions are dealt with $(\phi_1; \phi_2)$ (ϕ_1 followed by ϕ_2), *(while Φ do ψ)* (ψ as long as Φ holds) and *(if Φ then ϕ_1 else ϕ_2)* (ϕ_1 if Φ holds and ϕ_2 otherwise). The composite action $(\phi_1; \phi_2)$ is discussed in detail. An additional operator RES (*result*) is introduced to show the success/failure of the component actions. The RES operator functions as a *practition* operator which indicates the sequence of actions being performed, i.e., which action is performed next. The existing BDI architecture doesn't mention anything about the actual execution of actions. Since the transition caused by the execution of the action $(\phi_1; \phi_2)$ equals the *sum* of the transition caused by ϕ_1 and the one caused by ϕ_2 in the state brought about

by execution of ϕ_1 , the RES operator helps in acting as a filter which checks whether the first action results in a counterfactual state or not. Such a filtering helps in avoiding further deliberation at the time of action as would otherwise be in situations arising from counterfactual states. For example, the success of the printer command (*lpr*), in a unix environment, depends on the result of the execution of the command in the spooler phase followed by the recognition of the command by the printer in the communication phase. Here the action needs to be broken down into compartments and the success of each action should be validated for the overall success. In such circumstances the RES operator helps in providing the necessary specification. This goes in alliance with our view of categorising composite actions under future intentions, where the scope of practical reasoning is more.

Definition 4 Let ϕ_1, ϕ_2 be actions, then the axioms for the operator RES are:

$$\mathbf{R1} \text{ CAP}(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} \text{BEL}(\text{does}(\phi_i)) \wedge \text{BEL}(\text{RES}(\text{does}(\phi_1))) \neq \perp$$

$$\mathbf{R2} \text{ GOAL}(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} \text{CAP}(\text{does}(\phi_i)) \wedge \text{RES}(\text{does}(\phi_1)) \neq \perp$$

$$\mathbf{R3} \text{ CAP}(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} \text{BEL}(\text{CAP}(\text{does}(\phi_i))) \wedge \text{RES}(\text{does}(\phi_1)) \neq \perp$$

$$\mathbf{R4} \text{ GOAL}(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} \text{CAP}(\text{GOAL}(\text{does}(\phi_i))) \wedge \text{RES}(\text{does}(\phi_1)) \neq \perp$$

$$\mathbf{R5} \text{ INT}(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} \text{CAP}(\text{INT}(\text{does}(\phi_i))) \wedge \text{RES}(\text{does}(\phi_1)) \neq \perp$$

The first axiom states that an agent has the capability of performing a composite action $\phi_1; \phi_2$ then at some point of time the agent believes in doing ϕ_1 and ϕ_2 and believes that the performance of ϕ_1 does not end in counterfactual state of affairs (i.e, it does not end in falsity). Similarly the third axiom states that an agent has the capability of performing a composite action $\phi_1; \phi_2$, then at some point of time, the agent believes that it has the capability of doing ϕ_1 and believes in the capability of doing ϕ_2 and the result of ϕ_1 does not end up in a counterfactual state of affairs.

The semantic conditions for RES are similar to those given in Definition 3. For instance it can be shown that the semantic condition for R2 is

$$\forall w' \in C_t^w(x), \exists w'' \in G_t^w(x), \exists w''' \in R_t^w(x) \text{ such that } w'' \sqsubseteq w' \text{ and } w''' \sqsubseteq w'$$

where $R_t^w(x)$ is the set of result-accessible worlds of agent x in world w at time t .

This constraint means that for every capability-accessible world w' at time-point t , there is a goal-accessible world w'' at that time-point which is a sub-world of w' and a result-accessible world w''' which is a sub-world of w' . The converse doesn't hold as there can be Goal-accessible worlds that do not have corresponding capability as well as result-accessible worlds that do not have corresponding capability but only has the opportunity. We shall deal with the opportunity construct in the next section.

The action constructors dealing with *while Φ do ψ* (which means that ψ as long as Φ holds) and *if Φ then ϕ_1 else ϕ_2* (ϕ_1 if Φ holds and ϕ_2 otherwise) is crucial from computational point of view. For an agent to be able to perform an action *while Φ do ψ* it

has to have the ability to perform some finite actions constituting the body of the while-loop as well as the opportunity to perform all the steps. Agents should not be able to perform an action that goes indefinitely. These specifications are formally represented by the following two axioms.

R6 $CAP(\text{while } \Phi \text{ do } \psi) \Rightarrow [\neg\Phi \vee (\Phi \wedge BEL(CAP(\text{does}(\psi)))) \wedge RES(\text{done}(\phi)) \neq \perp]$

R7 $CAP(\text{if } \Phi \text{ then } \phi_1 \text{ else } \phi_2) \Rightarrow$
 $[\Phi \wedge BEL(CAP(\text{does}(\phi_1))) \wedge RES(\text{done}(\phi_1)) \neq \perp] \vee$
 $[\neg\Phi \wedge BEL(CAP(\text{does}(\phi_2))) \wedge RES(\text{done}(\phi_2)) \neq \perp].$

The first proposition states that an agent is capable of performing an action *while* Φ *do* ψ , as long as Φ holds and the agent believes that it has the capability of ψ and result of ψ does not end in falsity. Similarly R7 can be read as, an agent has the capability of performing an action *if* Φ *then* ϕ_1 *else* ϕ_2 , if Φ holds and the agent believes that it has the capability of ϕ_1 and the result of ϕ_1 is true, or it is the case that, Φ does not hold and the agent believes that it has the capability of ϕ_2 and result of ϕ_2 does not end in a counterfactual state of affairs.

5 Integrating Opportunity

Though in many cases it seems reasonable to assume that Capability implies Opportunity, when it comes to practical reasoning Opportunity seem to play a significant role. Van Linder [10] explains opportunity in terms of the correctness of action. An action is correct for some agent to bring about some proposition iff (*if and only if*) the agent has the opportunity to perform the action in such a way that its performance results in the proposition being true. Integrating opportunity lays further constraint on the part of the agent to think about an action before getting committed. Consider the example of a lion in a cage, which is perfectly well capable of eating a zebra, but ideally never has the opportunity to do so.¹ Using the BDI formalism we would have to conclude that the lion is capable of performing the sequential composition *eat zebra ; fly to the moon* which hardly seems to be intuitive. In such situations it is very important to know the combination of Capability and Opportunity so that no unwarranted conclusions can be drawn. We introduce an operator OPP whose intuitive meaning is agent x has the opportunity. The axioms for the OPP operator together with the Capability construct can be given as follows

Definition 5 *Let ϕ_1, ϕ_2 be actions, then we have*

O1 $CAP(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} BEL(OPP(\text{does}(\phi_i)))$

O2 $GOAL(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} CAP(\text{does}(\phi_i)) \wedge OPP(\text{does}(\phi_i))$

O3 $CAP(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} BEL(CAP(\text{does}(\phi_i))) \wedge OPP(\text{does}(\phi_i))$

¹ The example is taken from [10].

- O4** $\text{GOAL}(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} \text{CAP}(\text{GOAL}(\text{does}(\phi_i))) \wedge \text{OPP}(\text{does}(\phi_i))$
- O5** $\text{INT}(\text{does}(\phi_1; \phi_2)) \Rightarrow \bigwedge_{i=1,2} \text{CAP}(\text{INT}(\text{does}(\phi_i))) \wedge \text{OPP}(\text{does}(\phi_i))$
- O6** $\text{CAP}(\text{while } \Phi \text{ do } \psi) \Rightarrow [\neg\Phi \vee (\Phi \wedge \text{BEL}(\text{CAP}(\text{does}(\psi)))) \wedge \text{OPP}(\text{does}(\psi))]$
- O7** $\text{CAP}(\text{if } \Phi \text{ then } \phi_1 \text{ else } \phi_2) \Rightarrow [\Phi \wedge \text{BEL}(\text{CAP}(\text{does}(\phi_1))) \wedge \text{OPP}(\text{does}(\phi_1))] \vee [\neg\Phi \wedge \text{BEL}(\text{CAP}(\text{does}(\phi_2))) \wedge \text{OPP}(\text{does}(\phi_2))]$

The third axiom states that an agent has the capability of performing $\phi_1; \phi_2$ then the agent believes that he has the capability of ϕ_1 , if he has the opportunity of ϕ_1 , and, he has the capability of ϕ_2 , if he has the opportunity of ϕ_2 . Similarly O7 can be interpreted as an agent has the capability of doing the action (if Φ then ϕ_1 else ϕ_2) then either Φ holds and the agent believes that he/she has the capability of ϕ_1 provided the opportunity exists or $\neg\Phi$ holds and the agent has the capability of ϕ_2 provided the opportunity exists. The other axioms can be interpreted in a similar manner.

6 Opportunity + Results

In [10] a division is made between *optimistic agents* and *pessimistic agents* and the interpretation of the OPP formulae is done accordingly. They make use of two dynamic operators $\langle do_i(\alpha) \rangle \varphi$ and $[do_i(\alpha)]\varphi$. The first one denotes that an agent i has to have the opportunity to perform the action α in such a way that φ will result from the performance (*Pessimistic Approach*): A pessimistic agent needs certainty. The second one is the dual of the first and states that if the opportunity to do α is present then φ would be among the results of $do_i(\alpha)$ (*Optimistic Approach*). The formula $[do_i(\alpha)]\varphi$ is noncommittal about the opportunity of the agent i to perform the action α . We do not go for such a division and interpret the OPP formulae in a realistic manner linked with the RES operator. Such a formalism helps in avoiding unwarranted results as were seen in the earlier examples. In what follows we present the axioms capturing this intuition.

- OR1** $\text{CAP}(\text{does}(\phi_1; \phi_2)) \Rightarrow$
- $$\left[\begin{array}{c} \text{BEL}(\text{does}(\phi_1)) \wedge \text{OPP}(\text{does}(\phi_1)) \wedge \text{RES}(\text{done}(\phi_1)) \neq \perp \\ \wedge \text{BEL}(\text{does}(\phi_2)) \wedge \text{OPP}(\text{does}(\phi_2)) \end{array} \right]$$
- OR2** $\text{GOAL}(\text{does}(\phi_1; \phi_2)) \Rightarrow$
- $$\left[\begin{array}{c} \text{CAP}(\text{does}(\phi_1)) \wedge \text{OPP}(\text{does}(\phi_1)) \wedge \text{RES}(\text{done}(\phi_1)) \neq \perp \\ \wedge \text{CAP}(\text{does}(\phi_2)) \wedge \text{OPP}(\text{does}(\phi_2)) \end{array} \right]$$
- OR3** $\text{CAP}(\text{does}(\phi_1; \phi_2)) \Rightarrow$
- $$\left[\begin{array}{c} \text{BEL}(\text{CAP}(\text{does}(\phi_1))) \wedge \text{OPP}(\text{does}(\phi_1)) \wedge \text{RES}(\text{done}(\phi_1)) \neq \perp \\ \wedge \text{BEL}(\text{CAP}(\text{does}(\phi_2))) \wedge \text{OPP}(\text{does}(\phi_2)) \end{array} \right]$$
- OR4** $\text{GOAL}(\text{does}(\phi_1; \phi_2)) \Rightarrow$
- $$\left[\begin{array}{c} \text{CAP}(\text{GOAL}(\text{does}(\phi_1))) \wedge \text{OPP}(\text{does}(\phi_1)) \wedge \text{RES}(\text{done}(\phi_1)) \neq \perp \\ \wedge (\text{CAP}(\text{GOAL}(\text{does}(\phi_2))) \wedge \text{OPP}(\text{does}(\phi_2))) \end{array} \right]$$

OR5 $\text{INT}(\text{does}(\phi_1; \phi_2)) \Rightarrow$

$$\left[\begin{array}{l} \text{CAP}(\text{INT}(\text{does}(\phi_1)) \wedge \text{OPP}(\text{does}(\phi_1)) \wedge \text{RES}(\text{done}(\phi_1)) \neq \perp) \\ \wedge (\text{CAP}(\text{INT}(\text{does}(\phi_2)) \wedge \text{OPP}(\text{does}(\phi_2))) \end{array} \right]$$

OR6 $\text{CAP}(\text{while } \Phi \text{ do } \psi) \Rightarrow$

$$\left[\begin{array}{l} (\Phi \wedge \text{BEL}(\text{CAP}(\text{does}(\psi)) \wedge \text{OPP}(\text{does}(\psi))) \\ \wedge \text{RES}(\text{done}(\psi)) \neq \perp) \vee (\neg\Phi) \end{array} \right]$$

OR7 $\text{CAP}(\text{if } \Phi \text{ then } \phi_1 \text{ else } \phi_2) \Rightarrow$

$$\left[\begin{array}{l} \Phi \wedge \text{BEL}(\text{CAP}(\text{does}(\phi_1)) \wedge \text{OPP}(\text{does}(\phi_1)) \wedge \\ \text{RES}(\text{done}(\phi_1)) \neq \perp) \vee \\ (\neg\Phi \wedge \text{BEL}(\text{CAP}(\text{does}(\phi_2)) \wedge \text{OPP}(\text{does}(\phi_2)) \wedge \\ \text{RES}(\text{done}(\phi_1)) \neq \perp)) \end{array} \right]$$

Axioms OR1–OR7 are a formalisation of the results and opportunities together with the capability operator for composite actions. OR3 states that agents have the capability of doing a composite action $(\phi_1; \phi_2)$ to achieve Φ then the agent believes that it has the capability, provided the right opportunity, in each of the atomic states and the resulting condition is in alliance with its beliefs, i.e., it does not result in counterfactual situations. The actual execution of actions is made explicit through such a formalisation. Similarly OR6 states that if an agent has the Capability and Opportunity to perform a while-loop then it keeps this opportunity under execution of the body of the loop as long as the condition holds, i.e., as long as the result is true.

7 Commitment Axioms Revisited

In [13] a division is made in the commitment strategies of an agent, categorising an agent as *blindly committed agent*, *single minded agent*, and *open-minded agent*. A *blindly committed agent* maintains her intentions until she *actually* believes that she has achieved them; the *single minded agent* maintains her intentions as long as she believes that they are still options; finally an *open-minded agent* maintains her intentions as long as the intentions are still her goals. Based on the semantics given in the previous section the formalisation can be given as follows

CA1 $\text{INT}(\text{inevitable} \diamond \Phi) \Rightarrow$

$$\text{inevitable}(\text{INT}(\text{inevitable} \diamond \Phi) \cup \text{BEL}(\text{RES}(\Phi)))$$

CA2 $\text{INT}(\text{inevitable} \diamond \Phi) \Rightarrow$

$$\text{inevitable}(\text{INT}(\text{inevitable} \diamond \Phi) \cup \text{BEL}(\text{CAP}(\Phi)) \vee \neg \text{BEL}(\text{OPP}(\text{optional} \diamond \Phi)))$$

CA3 $\text{INT}(\text{inevitable} \diamond \Phi) \Rightarrow$

$$\text{inevitable}(\text{INT}(\text{inevitable} \diamond \Phi) \cup \text{BEL}(\text{GOAL}(\Phi)) \vee \neg \text{CAP}(\text{optional} \diamond \Phi))$$

The *self-aware agent* mentioned in [11] can be added to the above set of commitment strategies directly. It seems that the formalisation depicted above is much more intuitive than the one given by Rao and Georgeff [13]. For instance the axiom of blind commitment states that, if an agent intends that inevitably Φ be eventually true, then the agent

will inevitably maintain its intentions until she believes in the *result* of Φ . The addition of *result* is important in the sense that the blindly committed agent maintains the intentions until the agent *actually* believes that she has achieved them, i.e., until the agent has a justified true belief. This condition is needed for an agent blindly-committed to her means to inevitably eventually *believe* that she has achieved her means or ends. It also seems to be in alliance with the philosophical theories concerning the nature of belief. Similarly a single-minded agent maintains her intentions as long as she believes that she has got the capability for it. Since we do not say anything about an agent *optionally* achieving particular means or ends, even if the opportunity is present, the agent does not believe that optionally Φ be eventually true which is captured by the $\neg\text{BEL}(\text{OPP}(\text{optional}\diamond\Phi))$. Finally, an open-minded agent maintains her intentions as long these intentions are still her goals or as long as she lacks the ability of optionally achieving them.

8 Conclusion and Future Work

The representation and reasoning about composite actions in a BDI environment forms the primary contribution of this work. Our work is motivated by the fact that many BDI systems provide no clue as to the actual execution of actions, and are only able to perform actions in an endogenous manner. When dealing with composite actions the actual execution of actions need to be represented and reasoned about for the overall success of the action. The addition of the two operators RES and OPP strengthens the semantics and functions as a filter in avoiding counterfactual situations. Though some mention has been done in [3] about the composite action construct $(\phi_1; \phi_2)$, it has been restricted to the Intention domain and nothing has been mentioned regarding the result of the actions. The only other comparable work is given by [10].

An explicit representation of temporal constructs can be seen as a further extension to this work. We have used the temporal operator as a static variable. When it comes to composite actions it is important to mention explicitly the time of each action and the temporal duration of the commitment an agent has towards each action. The interpretation of the \bigcirc (next) operator in the original logic needs to be verified. For example when it comes to composite actions like $(\phi; \psi)$ the temporal operator \bigcirc can be interpreted either as $\diamond(\phi \Rightarrow \bigcirc\psi)$ or $(\phi \Rightarrow \bigcirc\psi)$. The temporal notion as to whether the action is performed now or eventually needs to be clarified. It would also be worthwhile to investigate *does* $(\phi; \psi)$ in terms of $(\text{done}\phi; \text{does}\psi)$, i.e., to find whether *does* $(\phi; \psi)$ is concurrent or sequential.

9 Acknowledgments

This work was partially supported by the Australian Research Council (ARC) **grant A49601783**. We would like to thank the members of the Knowledge Representation and Reasoning Unit (KRRU) for valuable suggestions.

A Possible World Semantics

A structure M is a tuple $M = \langle W, \{S_w\}, \{R_w\}, \mathbf{B}, \mathbf{D}, \mathbf{I}, \mathbf{L} \rangle$ where W is a set of possible worlds, S_w is a set of time points in world w ; $R_w \subseteq S_w \times S_w$ is a total binary *temporal accessibility* relation; \mathbf{L} is a truth assignment function that assigns to each atomic formula the set of world-time pairs at which it holds. \mathbf{B} is a *belief-accessibility* relation that maps a time-point in a world to a set of worlds that are belief accessible to it; and \mathbf{D} and \mathbf{I} are *desire* and *intention* accessibility relations, respectively, that are defined in the same way as \mathbf{B} .

There are two types of formulas: *state formulas* (which are evaluated at a state in a time-tree) and *path formulas* (which are evaluated against a path in a time-tree). They are defined as follows.

- any propositional formula is a state formula; if $\Phi \wedge \Phi'$ are state formulas then so too are $\Phi \vee \Phi'$ and $\neg\Phi$
- if Φ is a state formula then so too are $\text{BEL}(\Phi)$ and $\text{INT}(\Phi)$
- if Ψ is a path formula then *optional*(Ψ) and *inevitable*(Ψ) are state formulas.
- Any state formula is also a path formula
- if Ψ and Ψ' are path formulas then so too are $\Psi \vee \Psi'$, $\neg\Psi$, $\bigcirc\Psi$, $\diamond\Psi$ and $\square\Psi$

A *full path* in w is an infinite sequence of time points such that $(w_i, w_{i+1}) \in R_w$ for all i . Satisfaction of a state formula Φ is defined with respect to a structure M , a world w and a time point t , denoted by $M, w_t \models \Phi$.

Satisfaction of a path formula Ψ is defined with respect to a structure M , a world w , and a full path $(w_{t_0}, w_{t_1}, \dots)$ in world w .

- $M, w_t \models \Phi$ iff $(w, t) \in \mathbf{L}(\Phi)$, where Φ is an atomic formula.
- $M, w_t \models \neg\Phi$ iff $M, w_t \not\models \Phi$
- $M, w_t \models \Phi_1 \vee \Phi_2$ iff $M, w_t \models \Phi_1$ or $M, w_t \models \Phi_2$
- $M, (w_{t_0}, w_{t_1}, \dots) \models \Phi$ iff $M, w_{t_0} \models \Phi$, where Φ is a state formula.
- $M, (w_{t_0}, w_{t_1}, \dots) \models \bigcirc\Phi$ iff $M, (w_{t_1}, \dots) \models \Phi$
- $M, (w_{t_0}, w_{t_1}, \dots) \models \diamond\Psi$ iff $\exists w_{t_k} \in (w_{t_0}, w_{t_1}, \dots)$ s.t. $M, (w_{t_k}, w_{t_{k+1}}, \dots) \models \Psi$
- $M, w_{t_0} \models \square\Psi$ iff $M, (w_{t_0}, w_{t_1}, \dots) \models \Psi$, for all full paths $(w_{t_0}, w_{t_1}, \dots)$
- $M, (w_{t_0}, w_{t_1}, \dots) \models \Phi_1 \bigcup \Phi_2$ iff for some $i \geq 0$, $M, w_t \models \Phi_2$ and for all $0 \leq j < i$, $M, w_{t_j} \models \Phi_1$
- $M, w_{t_0} \models \text{optional}(\Psi)$ iff there exists a full path $(w_{t_0}, w_{t_1}, \dots)$ such that $M, (w_{t_0}, w_{t_1}, \dots) \models \Psi$.

References

1. M. E. Bratman. *Intentions, Plans and Practical Reason*. Harvard University Press, Cambridge, MA, 1987.
2. M.E. Bratman, D.J. Israel, and M.E Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4:349–355, 1988.
3. Lawrence Cavedon, Lin Padgham, Anand Rao, and Elizabeth Sonenberg. Revisiting rationality for agents with intentions. In *Eight Australian joint conference on Artificial Intelligence*, 1995.

4. E. A. Emerson. *Temporal and Modal Logic*. Elsevier, Cambridge, 1990.
5. M. Georgeff and F. Ingrand. Decision making in an embedded reasoning system. In *In proc. of the International Joint Conference on Artificial Intelligence- IJCAI*, pages 972–978, 1989.
6. Bruce Vermazen & Merrill Hintikka. *Essays on Davidson: Actions and Events*. Clarendon Press, Oxford, 1985.
7. M. Huber. Jam: A bdi-theoretic mobile agent architecture. In *In Proceedings of the Third International Conference on Autonomous Agents - Agents 99*, pages 236–243, Seattle, WA, 1999.
8. G.E. Hughes and M.J. Cresswell. *An Introduction to Modal Logic*. Routledge, London, 1968.
9. Fischer K. and Muller J. P. & Pischel M. A pragmatic bdi architecture. In Michael J. Woodridge, Muller J. P., and Milind Tambe, editors, *Intelligent Agents II, Agent theories, Architectures, and Languages, IJCAI'95 Workshop (ATAL), Montreal, Canada*, volume 1037 of *Lecture notes in Computer Science*, pages 203–218. Springer-Verlag, 1996. subseries of Lecture notes in Artificial Intelligence (LNAI).
10. Bernardus Van Linder. *Modal Logic for Rational Agents*. PhD thesis, Department of Computer Science, Utrecht University, 19th June 1996.
11. Lin Padgham and Patrick Lambrix. Agent capabilities: Extending bdi theory. In *Proceedings of Seventeenth National Conference on Artificial Intelligence(AAAI-2000)*, pages 68–73, Austin, Texas USA, July 30-August 3 2000. AAAI Press/The MIT Press.
12. Anand S. Rao. Means-end plan recognition- towards a theory of reactive recognition. Technical note - 49, Australian Artificial Intelligence Institute, 1994.
13. A.S. Rao and M.P. Georgeff. Modelling rational agents within a bdi-architecture. In Allen J. Fikes R. Sandewall E., editor, *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*, pages 473–484. Morgan Kaufmann, 1991.