BDI architecture in the framework of Situation Calculus

Robert Demolombe
ONERA-Toulouse *
Robert.Demolombe@cert.fr

Pilar Pozos Parra
Benemérita Universidad Autnoma de Puebla †
ppozos@fcfm.buap.mx

Abstract

The BDI architecture (Beliefs, Desires and Intentions) has been accepted by an important part of the scientific community for agent modeling. Its positive feature is to propose a simple representation of agents' rational behaviour. In this paper a formalisation of the BDI concepts is presented in the framework of Situation Calculus. It is shown how the most important properties of the concept of intention are formalised, and this formalisation is compared with other formalisations of the BDI architecture, in particular with the formalisation proposed by Cohen and Levesque.

An interesting and original feature of the proposed framework is that it can be implemented using the same method as for implementing Reiter's basic action theories. The scenario presented in this paper has been implemented in Prolog.

1 Introduction

Several authors have proposed a logical formalisation of the concepts involved in the BDI architecture (see [Singh, 1994; Wooldridge, 2000; 2002; 1992; Singh *et al.*, 1998; Lespérance *et al.*, 1999; van Linder, 1996; Rao and Georgeff, 1991]). Most of them use modal logics to formalize cognitive concepts. Although technically intuitive and elegant, modal approches over estimate the reasoning capabilities of agents. For exemple, an agent who intendes p is assumed to intend all logical consequences of p (for the knowledge this is called logical omniscience). Real life cannot be logical omniscience. Moreover modal logics have been extensively studied and their implementation complexity is well known. In the other hand, there are other approches to solve this kind of problem but they do not support any inferences among the cognitive concepts.

A new proposal is presented in this paper in the perspective of finding a tradeoff between the expressive power of the formalism and the possibility to design a realistic implementation. That is why we have selected the Situation Calculus.

Intention is the more complex concept in the BDI architecture. Intentions are determined by a rational analysis of the current situation and of the future possible situations, and by the analysis of the actions that allow to change the situation. Cognitive agents have the ability to predict, to some extend, the consequences of the actions they have chosen to perform. They select the actions in function of the causal links between the actions and the situations. In other words they can foresee the future.

To design a rational agent the BDI architecture defines the roles played by the beliefs, the desires and the intentions. Bratman treats intentions as sets of actions that an agent has committed to perform in order to fulfill her goals. If an agent takes the decision to perform a sequence of actions that is because they allow to reach her goal. Then, a theory of intention requires a well defined theory of action, that is a theory of the evolution of the world. This has been formalised in the Situation Calculus, and that is one of the reasons why we have selected this formal framework.

It is possible in the Situation Calculus to represent in a simple way the evolution of the world [Reiter, 1991] and the evolution of beliefs [Demolombe and Pozos-Parra, 2000]. The later has been restricted to beliefs about the present situation. For instance, it can be represented that it is believed that it rains, but it cannot be represented that it will rain or that it has been raining. In the BDI architecture agents must to reason about the future. A significant contribution of this work is an extension to beliefs about the past and about the future. However, in the context of intention formalisation we only have to consider beliefs about the future.

Another contribution is the representation of evolution of goals (consistent with achievable desires) in the same style as the evolution of beliefs which has been presented in [Demolombe and Pozos-Parra, 2000].

The paper is organised as follows. We start with a brief introduction to the Situation Calculus and to the corre-

*2 Avenue E. Belin BP 4025, 31055 Toulouse Cedex 4, France

†Avenida San Claudio y Ro Verde, C.P.72570, C.U. Puebla, México

sponding representation of the evolution of the world, and of the agents' beliefs. Then, an extension of beliefs about the past and about the future is presented, The next step is to present the evolution of goals and of intentions. Finally, the proposed formalisation is compared with other ones.

2 Situation Calculus

Reiter in [Reiter, 1991] has proposed to represent the evolution of the world in the framework of Situation Calculus which is a classical first order logic with equality [1]. In this logic predicates whose truth value can dynamically change are called "fluents". They have one argument (the last argument) which is of the type situation. It is assumed that every change is caused by an action which is represented by a term (e.g. $advance$, $move(x, y)$).

For instance, the fact that a robot is at the position x in the situation s is represented by $position(x, s)$. Situations are denoted either by constants (for instance, S_0 for the initial siutation), or by variables (they are usually denoted by s, s', $s'' \ldots$), or by terms of the form $do(a, s)$, where the first argument is of the type action and the second argument is of the type situation. If $advance$ and $reverse$ have the type action, the following terms have also the type situation: $do(advance, S_0)$, $do(reverse, do(advance, S))$, ...

The intuitive meaning of the formula $position(x_1, do(advance, S_0))$ is that in the situation that results from the performance of the action $advance$ from the situation S_0, the robot is at the position x_1. The variables of the type action and situation can be quantified like in: $\exists a \exists s position(x_1, do(a, s))$.

The formalisation of the evolution of the world will be presented first through an example. It is assumed that in every situation a robot is at the position x after performance of the action $advance$ from the position $x - 1$, or after performance of the action $reverse$ from the position $x + 1$. Moreover, the robot is no more at the position x after performance of one of the actions $adavance$ or $reverse$ from the position x. In formal terms we have:
$\forall s \forall a \forall x[(a = advance \wedge position(x - 1, s)) \vee$
$(a = reverse \wedge position(x+1, s)) \rightarrow position(x, do(a, s))]$

$\forall s \forall a \forall x((a = advance \vee a = reverse) \wedge position(x, s)$
$\rightarrow \neg position(x, do(a, s)))$.

It is also assumed that the actions in the antecedent of the first formula (respectively the second formula) are the only actions that have the effect that the robot in the situation $do(a, s)$ is at the position x (respectively is not at the position x). Form these completeness assumptions we can infer the following successor state axiom for the fluent $position(x, s)$ [Reiter, 2001]:

$\forall s \forall a \forall x(position(x, do(a, s)) \leftrightarrow [(a = advance \wedge position(x - 1, s)) \vee (a = reverse \wedge position(x + 1, s))] \vee (position(x, s) \wedge [\neg a = advance \wedge \neg a = reverse]))$

This axiom allows to know the truth value of $position(x, do(a, s))$ for every action a and every situation s, provided we know in the situation s the truth values of the conditions in the right hand side in the above equivalence.

For instance, from $position(5, S_0)$ we can infer $\neg position(5, do(advance, S_0))$ and $position(6, do(advance, S_0))$. This means that after to advance the robot is at the position 6 and it is no more at the position 5.

In general, to represent the evolution of the world, for each fluent p we need a successor state axiom like (S_p). To make the notations simpler we have only explicited the arguments of the type situation and action:

(S_p) $\forall s \forall a(p(do(a, s)) \leftrightarrow \Upsilon_p^+(a, s) \vee p(s) \wedge \neg \Upsilon_p^-(a, s))$

The conditions $\Upsilon_p^+(a, s)$ and $\Upsilon_p^-(a, s)$ mention the action a and the situation s, but they do not mention the situation $do(a, s)$. In addition it is assumed that there is no situation and no action that can cause p to be both true and false. This assumption is formally represented by: $\neg \exists s \exists a(\Upsilon_p^+(a, s) \wedge \Upsilon_p^-(a, s))$.

The set of axioms of the form (S_p) for all the fluents defines the truth values of the atomic formulas in any circumstances, and it indirectly defines the truth value of any formula.

In [Scherl and Levesque, 1993] the authors Scherl and Levesque have extended the formalism to represent beliefs and their evolution. The notation $Bel(p, s)$ means that in the situation s an agent believes p. This notation denotes the following formula: $\forall s'(K(s', s) \rightarrow p[s'])$, which expresses that in every epistemic variant s' of s the property p holds. In their definition there is no restriction about the formula p. In particular p may also include beliefs. The definition of belief evolution is rather complex and we have no room to present it here. This approach though it is very general raises theoretical problems in the case of belief revision, and practical problems if one intents to implement it.

In [Demolombe and Pozos-Parra, 2000] we have presented another approach which is less general but it can be rather easily implemented [2]. In this approach a notion similar to modal operator is introduced. In fact a statement $B_i(p(s))$, where B_i is a modal operator which represents what the agent i believes and $p(s)$ is an atomic formula, is represented by the fluent $Bip(s)$. Nevertheless, for convenience, we adopt the notation $B_i(p(s))$ to denote $Bip(s)$. In the same way $B_i(\neg p(s))$ is used to denote $Binotp(s)$. In particular the fluent $Biposs(a, s)$ (or $B_i(Poss(a, s))$) represents the fact that the agent i believes in s that it is possible to perform the action a.

In general an agent may have four different mental attitudes with respect to his belief about a given property. Let us consider, for example, that the pilot of the robot can observe in a screen the position of the robot. If it displayed that the robot is at the position x , the pilot

[1] Some limited fragments of Situation Calculus require to deal with second order logic and with three types of terms: situation, action and object.

[2] A comparison of the two approaches can be found in [Petrick and Levesque, 2002].

believes that it is at the position x and he does not believe that it is not at the position x. In formal terms we have:
$\forall s \forall a \forall x(a = obs.pos \wedge position(x, s) \rightarrow B_p(position(x, do(a, s))))$

$\forall s \forall a \forall x(a = obs.pos \wedge position(x, s) \rightarrow \neg B_p(\neg position(x, do(a, s))))$

If the pilot observes that the robot is not at the position x, then he does not believe that it is in the position x and he believes that it is not in the position x. In formal terms we have:
$\forall s \forall a \forall x(a = obs.pos \wedge \neg position(x, s) \rightarrow \neg B_p(position(x, do(a, s))))$

$\forall s \forall a \forall x(a = obs.pos \wedge \neg position(x, s) \rightarrow B_p(\neg position(x, do(a, s))))$

In general the four possible formal attitudes of an agent i are represented by: $B_i(p(s))$, $\neg B_i(\neg p(s))$, $\neg B_i(p(s))$ and $B_i(\neg p(s))$. If it assumed, in the same way as we did for the evolution of the world, that the conditions that cause each attitude are complete conditions, therefore we have the two following successor belief state axioms that define the evolution of pilot's belief about $position(x, s)$ and about $\neg position(x, s)$:
$\forall s \forall a \forall x(B_p(position(x, do(a, s))) \leftrightarrow$
$a = obs.pos \wedge position(x, s) \vee$
$B_p(position(x, s)) \wedge \neg(a = obs.pos \wedge \neg position(x, s)))$

$\forall s \forall a \forall x(B_p(\neg position(x, do(a, s))) \leftrightarrow$
$a = obs.pos \wedge \neg position(x, s) \vee$
$B_p(\neg position(x, s)) \wedge \neg(a = obs.pos \wedge position(x, s)))$.

In general, to represent the evolution of agents' beliefs for each agent i and for each fluent p we have two successor belief state axioms of the following form:
$(S_p^{B_i})$ $\forall s \forall a(B_i(p(do(a, s))) \leftrightarrow \Upsilon^+_{B_i p}(a, s) \vee B_i(p(s)) \wedge \neg \Upsilon^-_{B_i p}(a, s))$

$(S_{\neg p}^{B_i})$ $\forall s \forall a(B_i(\neg p(do(a, s))) \leftrightarrow \Upsilon^+_{B_i \neg p}(a, s) \vee B_i(\neg p(s)) \wedge \neg \Upsilon^-_{B_i \neg p}(a, s))$

Like for successor state axioms we have to impose constraints to prevent the derivation of inconsistent beliefs (see [Demolombe and Pozos-Parra, 2000]). Moreover it has been shown that if the initial representation of the world and of agents' beliefs is consitent, then it is consistent after performance of any sequence of actions.

3 Evolution of general beliefs

In the axiomatisation presented in the previous section the situation s in sentences of the form $B_i(p(s))$ refers both to the situation where p holds and the situation where the belief holds. That means that beliefs are about the present.

As far as we know there is no proposal in the Situation Calculus to represent beliefs about the past and about the future. A natural extension of [Scherl and Levesque, 1993] in this direction might be defined as follows.

Let us adopt the notation $do^{-1}(a, s)$ to denote the situation s' such that $s = do(a, s')$. Intuitively $do^{-1}(a, s)$ represents the situation where we are before to perform a. Then, it can be noticed that we have the following properties:

$p(do(a, s)) \leftrightarrow \forall s'(s' = do(a, s) \rightarrow p(s'))$
$p(do^{-1}(a, s)) \leftrightarrow \forall s'(s = do(a, s') \rightarrow p(s'))$

Now we can represent the fact that it is believed in the situation s that p holds in $do(a, s)$ (respectively in $do^{-1}(a, s)$) by the following formulas:

$Bel(p(do(a, s), s) \stackrel{\text{def}}{=} \forall s'(K(s', s) \rightarrow p[do(a, s')])$

$Bel(p(do^{-1}(a, s), s) \stackrel{\text{def}}{=} \forall s'(K(s', s) \rightarrow p[do^{-1}(a, s')])$

These definitions could be easily extended in the case where instead of the action a we have any sequence of actions. Even if these definitions are acceptable from a theoretical point of view, they would be too complex to be implemented, and we prefered to go in a simpler direction which is presented below.

To represent beliefs about the past, the present and the future we use the notation $B_i(p(s_1), s_2)$ to denote atomic formulas of the form $Bip(s_1, s_2)$ ($Binotp(s_1, s_2)$ has similar intuitive meaning). This notation can be intuitively read: in the situation s_2 the agent i believes that p is true in the situation s_1. This extension allows to represent beliefs about the future when we have $s_2 < s_1$, beliefs about the past when $s_1 < s_2$, and beliefs about the present when $s_1 = s_2$. [3]

The successor belief state axioms $(S_p^{B_i})$ and $(S_{\neg p}^{B_i})$ are extended as follows:

(B_i^p) $\forall s \forall a(B_i(p(s_1), do(a, s)) \leftrightarrow \Upsilon^+_{B_i p}(a, s_1, s) \vee B_i(p(s_1), s) \wedge \neg \Upsilon^-_{B_i p}(a, s_1, s))$

$(B_i^{\neg p})$ $\forall s \forall a(B_i(\neg p(s_1), do(a, s)) \leftrightarrow \Upsilon^+_{B_i \neg p}(a, s_1, s) \vee B_i(\neg p(s_1), s) \wedge \neg \Upsilon^-_{B_i \neg p}(a, s_1, s))$

The conditions Υs may contain communication actions which are a generalisation of sensing actions that cause belief change in [Scherl and Levesque, 1993]. For instance, if in 1990 an agent believes that the soccer cupworld in 1986 happened in Italy, which is represented by $B_i(cupworld(Italy, 1986), 1990)$ [4], and in 2002 this agent learns from a newspaper that it happened in Mexico, then his new belief is $B_i(cupworld(Mexico, 1986), 2002)$. Then, if until 2003 he gets no more information about this event we have $B_i(cupworld(Mexico, 1986), 2003)$.

Since in the following we shall use mainly beliefs about the future, we adopt the notation:

$Bf_i(p(s_1), s_2) \stackrel{\text{def}}{=} s_2 < s_1 \wedge B_i(p(s_1), s_2)$

4 Evolution of goals

To represent the evolution of goals the language is extended with predicates of the form $Gip(s)$ which are de-

[3] The predicate $s < s'$ represents the fact that the situation s' is obtained from s after performance of one or several actions.

[4] Here, we have identified the dates as the situations to make the example easier to understand.

noted by $G_i(p(s))$, and whose intuitive meaning is that in the situation s the agent i has the goal that p became true in the future (similarly for $Ginotp(s)$). Here p is retricted to be an atomic formula or the negation of an atomic formula.

In the same way as for beliefs an agent may have four different attitudes in terms of goals with regard to a proposition. For example, we may have $G_i(position(5, s))$: in the situation s the agent i has the goal to be at the position 5, $G_i(\neg position(5, s))$: in the situation s the agent i has the goal not to be at the position 5, $\neg G_i(position(5, s))$: in the situation s the agent i does not have the goal to be at the position 5 and $\neg G_i(\neg position(5, s))$: in the situation s the agent i does not have the goal not to be at the position 5.

The evolution of goals is determined by actions of the kind *select*, whose effect is to adopt a goal, or of the kind *abandon*, whose effect is to give up a goal. Then, in the case of our robot example we have:

$$\forall s \forall a \forall x (a = select.pos(x) \rightarrow G_r(position(x, do(a, s))))$$

$$\forall s \forall a \forall x (a = select.not.pos(x) \rightarrow G_r(\neg position(x, do(a, s))))$$

$$\forall s \forall a \forall x (a = abandon.pos(x) \rightarrow \neg G_r(position(x, do(a, s))))$$

$$\forall s \forall a \forall x (a = abandon.not.pos(x) \rightarrow \neg G_r(\neg position(x, do(a, s))))$$

These axioms define the four possible attitudes of the robot. If, in addition, it is assumed that the conditions in the antecedents of each axiom represent all the circumstances that cause each attitude, then with the same reasoning as for beliefs we have the successor goal state axioms:

$$\forall s \forall a \forall x (G_r(position(x, do(a, s))) \leftrightarrow a = select.pos(x) \vee G_r(position(x, s)) \wedge \neg(a = abandon.pos(x)))$$

$$\forall s \forall a \forall x (G_r(\neg position(x, do(a, s))) \leftrightarrow a = select.not.pos(x) \vee G_r(\neg position(x, s)) \wedge \neg(a = abandon.not.pos(x)))$$

In general for each agent i and for each fluent p we have two successor goal state axioms of the form:

$(S_p^{G_i})$ $\forall s \forall a (G_i(p(do(a, s))) \leftrightarrow \Upsilon^+_{G_i p}(a, s) \vee G_i(p(s)) \wedge \neg \Upsilon^-_{G_i p}(a, s))$

$(S_{\neg p}^{G_i})$ $\forall s \forall a (G_i(\neg p(do(a, s))) \leftrightarrow \Upsilon^+_{G_i \neg p}(a, s) \vee G_i(\neg p(s)) \wedge \neg \Upsilon^-_{G_i \neg p}(a, s))$

Notice that the fact that an agent has no goal with respect to the fact that it is at the position x can be represented by $\neg G_r(position(x, s)) \wedge \neg G_r(\neg position(x, s))$. To prevent situations where an agent has inconsistent goals, like in $G_r(position(x, s)) \wedge G_r(\neg position(x, s))$, we have to impose constraints to the conditions Υs as we did for beliefs in [Demolombe and Pozos-Parra, 2000].

5 Intentions

In the context of Multi Agent Systems, intention is a concept that allows to relate goals with beliefs and commitments. Also, we can see intention as the concept that motivates agents' actions. Most of the works in the field of intention are based on the proposal by Cohen and Levesque [Cohen and Levesque, 1990] which itself is based on Bratman's proposal [Bratman, 1987]. In these approaches intention is oriented to the future and it generates subintentions to reach to satisfy the initial goal.

In fact, intention is implicitly based on a plan generation technique, or on a prediction of future situations, which is one of the characteristics of rational agents.

Bratman [Bratman, 1987] says that intentional actions have the following properties: intentions are originally part of the problems that an agent has to solve, and intentions must be consistent.

The fact that an agent has the intention to perform the sequence of action $T = [a_1, a_2, \ldots, a_n]$ in the situation s in order to satisfy the goal p (resp. $\neg p$) is represented by $Iip(T, s)$ (resp. $Iinotp(T, s)$) and it is denoted by $I_i(T, p(s))$ (resp. $I_i(T, \neg p(s))$). Here p is restricted to be an atomic formula. This fact presupposes that the following conditions are satisfied:

1)the agent i does not believe that p (resp. $\neg p$) is the case now, i.e. we have $\neg B_i(p(s))$ (resp. $\neg B_i(\neg p(s))$),

2) the agent i believes that after performance of T p (resp. $\neg p$) will be the case, i.e. $Bf_i(p(do(T, s)), s)$ (resp. $Bf_i(\neg p(do(T, s)), s)$),[5]

3) the agent i believes that it is possible to perform T in s, i.e. $B_i(Poss(T, s), s)$; this last sentence is taken as an abreviation for the belief that it is possible to perform the action a_j in the situation s_j, i.e. $B_i(Poss(a_j, s_j), s_i)$, where $s_1 = s$ and for j in [1,n] $s_j = do(a_{j-1}, s_{j-1})$,

4) the agent i believes that some agent (it may be himself) has the ability to perform each action a_j in T.

To express the condition 3) we need to solve the qualification problem in the context of beliefs. Our proposal is an extension of the formalisation of the qualification problem presented by Reiter in [Reiter, 2001]. That is the action precondition belief axiom here has the form:

$$B_i(Poss(a, s)) \leftrightarrow \pi_a^{B_i}(s)$$

To express the conditions 2) and 3), we have introduced predicates of the form $Bfpossip(do(T, s), s)$ (that is denoted by $Bfposs_i(p(do(T, s)), s)$) whose meaning is represented by the formulas $Bf_i(p(do(T, s)), s) \wedge B_i(Poss(T, s))$. Their intuitive meaning is that the agent i believes that p holds in the situation $do(T, s)$ and that it is possible to perform T.

The condition 4) is implicitly satisfied by the successor belief state axioms and by the above precondition axiom, because if an action occurs in a successor belief state axiom the agent believes that it is possible to perform it.

Another important property of intention is persistence, which means that an intention persists as long as the above conditions are satisfied.

This definition of intention is supported by the fact that an agent has to find a sequence of actions which is executable and which allows to satisfy his goal. But the

[5] $do(T, s)$ is an abreviation for $do(a_n, do(a_{n-1}, do(\ldots, do(a_1, s))\ldots))$.

world may change because other agents perform actions or because the environment changes, and the preconditions to perform actions may be invalidated independently of the planification performed by the agent. Then, the agent is not guaranteed to realise his intention.

To solve this problem it might be possible to take into account the information the agent has about the other agents' intentions. A naive solution to this problem might be to characterise the evolution of facts of the form $B_i(I_j(T,p(s)))$ which denote predicates of the form $BiIjp(T,s)$, and whose meaning is that the agent i believes that the agent j has the intention to perform T in order to reach the goal p.

It has been seen that the definition of intention is based on the possibility to generate plans, but right now we are not guaranteed that the agent will perform the planed actions. Indeed, to have the intention to perform some action is not just to have the ability to plan this action but also to commit himself to do this action.

Finally, our proposal to define the evolution of intentions is defined by axioms of the form:

$(S_p^{I_i})$ $\forall s \forall a (I_i(T,p(do(a,s))) \leftrightarrow G_i(p(do(a,s))) \wedge [(a = commit \wedge Bfposs_i(p(do(T,s)),s)) \vee I_i([a,a_1,\ldots,a_n],p(s)) \vee \phi^+_{Ip} \vee I_i(T,p(s)) \wedge \neg\phi^-_{Ip}])$

$(S_{\neg p}^{I_i})$ $\forall s \forall a (I_i(T,\neg p(do(a,s))) \leftrightarrow G_i(\neg p(do(a,s))) \wedge [(a = commit \wedge Bfposs_i(\neg p(do(T,s)),s)) \vee I_i([a,a_1,\ldots,a_n],\neg p(s)) \vee \phi^+_{I\neg p} \vee I_i(T,s) \wedge \neg\phi^-_{I\neg p}])$

The intuitive meaning of these axioms is that the agent has the intention to perform the sequence of action $T = [a_1,\ldots,a_n]$ in order to have p (resp. $\neg p$), and one of the follwing condition holds:

1) The action which has been performed is a commitment action and the agent believes that after performance of T p (resp. $\neg p$) holds. This condition means that if the agent has not performed the commitment action, even if he has a goal he has not the intention to satisfy his goal.

2) In the previous situation he had the intention to perform $[a,a_1,\ldots,a_n]$ and his intention does not include any more a because a has just happened.

3) Possibly other conditions ϕ^+_p holds which causes that he adopts the new intention (for instance the perception of the presence of an obstacle may cause the intention to change his path).

4) In the previous situation he has the same intention and it is not the case that some conditions represented by ϕ^-_p holds; these conditions have the effect to abandon his intention (for instance the perception of the presence of an obstacle may cause the agent to abandon his intention to advance).

6 Comparison with other works

In the theory of intentional actions proposed by Cohen and Levesque an action is considered as a sequence of events, where an event is a primitive concept, and a proposition can be satisfied after performance of an action.

Action performance is denoted by the operator $DONE$, and $DONE(a)$ means that performance of the action a has just happened. In the Situation Calculus this infromation is implicitly represented by the definition of situations. For instance, $position(2, do(advance, S_0))$ means that the robot is at the position 2 in the situation that immediately follows the performance of the action *advance*.

They introduce formulas of the form $\exists x(HAPPENS\ x; \alpha?)$, which are denoted by $\Diamond\alpha$, whose meaning is that there exists a sequence of events x such that after performance of x α holds. We have a similar notion in the Situation Calculus which is represented by $Bf_i(p(do(T,s)),s)$, whose meaning is that agent i believes that after performance of T p holds. Notice that the difference is that this property is explicitly part of the agent's set of beliefs. That is the agent can reason about this property.

Goals are a subset of beliefs that obey the logic (KD), in particular they are consistent. In our proposal we have presented an axiomatics of the evolution of goals, which are restricted to literals, and satisfy (D). That is, we have $G_i(p(s)) \rightarrow \neg G_i(\neg p(s))$, where p is an atom. If in addition we impose the axiom schemas $G_i(p(s)) \rightarrow \exists T Bf_i(p(do(T,s)),s)$ and $G_i(\neg p(s)) \rightarrow \exists T Bf_i(\neg p(do(T,s)),s)$ goals are restricted to what we believe that it can happen in the future.

Cohen and Levesque make use of the operator $LATER$ to represent properties that may hold in the future and do not hold at the present time. They have $LATER(p) = \neg p \wedge \Diamond p$. Then, they can define an achievement goal as a property that it believed not to hold now and can hold in the future. An achievement goal is denoted by $A - GOAL$ and we have $A - GOAL(p) = GOAL(LATER(p)) \wedge BEL(\neg p)$[6]. This notion is not necessary in our approach, given our definitions of G_i, $Bfposs_i$ and B_i.

An agent may eventually abandon his goal. This is formally represented by $\vdash \Diamond\neg(GOAL(LATER(p)))$. This idea to have the possibility to abandon a goal is represented in our approach by the action *abandon* in the succesor goal state axioms. In the situation $do(abandon, s)$ a goal has been abandoned.

Persistent goals capture the notion of fanatic commitment, that is a goal that persists until it is satified or it is believed that it will never be possible to satisfy it. In formal terms persistent goals are represented by $P - GOAL(p) = A - GOAL(LATER(p)) \wedge [BEFORE((BEL(p) \vee BEL(\Box\neg p))\ \neg GOAL(LATER(p))]$.

In other words a persistent goal is a goal to be satisfied or that it will never be satisfied. In our proposal the notion of persistence can be directly expressed in the successor goal state axioms. For that purpose, successor goal state axioms must take the form: $G_i(p(do(a,s))) \leftrightarrow \Gamma^+_p \vee G_i(p(s)) \wedge \neg[\Gamma^-_p \vee B_i(p(s)) \vee \forall s' Bf_i(\neg p(s'),s)]$.

[6] For simplicity we have omited the reference to the agent unlike in the original paper.

This form of the axioms expresses the idea that an agent abandon a goal p only if he believes that it is satisfied (i.e. $B_i(p(s))$) or if he believes that it will be impossible to satisfy it (i.e. $\forall s' Bf_i(\neg p(s'), s)$). Of course, they may be other circumstances that cause that a goal is abandoned. These circumstances are represented by Γ_p^-. For example, when the agent has received the order to abandon his goal.

Finally, they define the intention to perform an action a in terms of persistent goals as follows: $INTEND_1(a) = P - GOAL[DONE(BEL(HAPPENS\ a))?; a]$. That means that the intention to perform a is the fanatic goal to perform this action when the agent believes that the action can be performed. This definition is captured in the proposal we have presented in the previous section. Moreover, in the definition of the intention to perform T, it must be explicit which is the associated achievement goal p. In that case it is clear why the agent have her intention.

Their approach include commitment in the core semantical definition of intentions. Another contrasting approach is [Rao and Georgeff, 1991] that consider intention as a basic attitude and treat it like beliefs and goals. This approach considers commitment as constraint for the intention revision and intention. Our approach consider too commitment as a constraint for the intention revision.

7 Conclusion

We have presented an extension of Situation Calculus in which are formalised the BDI concepts. The general idea was to avoid to deal with modal logic and rather to represent modal concepts in terms of fluents. That was possible thanks to strong restrictions about the expressive power of the formulas. In fact these formulas are restricted to literals. The benefit of this approach is that the evolution of beliefs, goals and intentions can be easily defined by successor axioms. That is, we have followed the initial simple idea proposed by Reiter for successor state axioms. By doing so we have avoided difficult theoretical problems, like belief, goal and intention revision. Moreover, it has been possible to implement this approach and to run simple examples to check the validity of our ideas.

References

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

[Cohen and Levesque, 1990] P.R. Cohen and H. Levesque. Intention Is Choice with Commitment. *Artificial Intelligence*, 42, 1990.

[Demolombe and Pozos-Parra, 2000] R. Demolombe and M. P. Pozos-Parra. A simple and tractable extension of situation calculus to epistemic logic. In Z. W. Ras and S. Ohsuga, editors, *Proc. of 12th Internatioonal Symposium ISMIS 2000.* Springer. LNAI 1932, 2000.

[Lespérance *et al.*, 1999] Y. Lespérance, H. Levesque, and R. Reiter. A situation calculus approach to modeling and programming agents. In A. Rao and M. Wooldbridge, editors, *Foundations and Theories of Rational Agents.* Kluwer, 1999.

[Petrick and Levesque, 2002] R. Petrick and H. Levesque. Knowledge equivalence in combined action theories. In *Proceeding 8th International Conference on Principles of Knowledge Representation and Reasoning*, 2002.

[Rao and Georgeff, 1991] A.N. Rao and M.P. Georgeff. Modeling Rational Agents within a BDI Architecture. In *Proc. 2nd Int. Conf. on Knowledge Representation and Reasoning.* Morgan Kaufmann, 1991.

[Reiter, 1991] R. Reiter. The frame problem in the situation calculus: a simple solution (sometimes) and a completeness result for goal regression. In V. Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, pages 359–380. Academic Press, 1991.

[Reiter, 2001] R. Reiter. *Knowledge in Action: Logical Foundations for Specifying and Implementing Dynamical Systems.* MIT Press, 2001.

[Scherl and Levesque, 1993] R. Scherl and H. Levesque. The Frame Problem and Knowledge Producing Actions. In *Proc. of the National Conference of Artificial Intelligence.* AAAI Press, 1993.

[Singh *et al.*, 1998] M. P. Singh, A. Rao, and M. Georgeff. Formal Method in DAI : Logic Based Representation and Reasoning. In G. Weis, editor, *Introduction to Distributed Artificial Intelligence*, New York, 1998. MIT Press.

[Singh, 1994] M. P. Singh. *Multiagent Systems. A Theoretical Framework for Intentions, Know-How, and Communications.* Springer LNAI 799, 1994.

[van Linder, 1996] B. van Linder. *Modal Logics for Rational Agents.* PhD thesis, University of Utrecht, 1996.

[Wooldridge, 1992] M. Wooldridge. *The Logical Modelling of Computational Multi-Agent Systems.* PhD thesis, University of Manchester, 1992.

[Wooldridge, 2000] M. Wooldridge. *Reasoning about Rational Agents.* The MIT Press, 2000.

[Wooldridge, 2002] M. Wooldridge. *Introduction to Multi Agent Systems.* J. Wiley and Sons, 2002.