Yet another semantics of goals and goal priorities

Jacques Wainer[1]

Abstract. This paper presents a semantics of goals that deals with the problem of closeness under believed implications. The paper develops the semantics of wishes, its relation to beliefs, and define goals in relation to wishes. Also, the semantics of a priority or importance relation among propositions (and goals) is developed. Finally the paper advances in extending this semantics to the multi-agent case.

1 Introduction

Goals are an important aspect of an agent's mental state. The "standard" semantics for goals (for example [2, 13]) is based on an extension of the possible world semantic of knowledge: besides an accessibility relation among the possible worlds that defines the knowledge of an agent, one would also define a desirability relation that is the base of the definition of goals. Like the accessibility relation, the desirability relation has some constraints that define further properties of the goal operator.

One of the main problems with that approach is that the goal operator will have the unreasonable property of being closed under believed implications, as it was pointed out in [9]. That is, if **G** is the goal operator and **B** is the knowledge or belief operator, then the expression below is true in a logic with the "standard" goal definition.

$$\mathbf{G}\alpha \land \mathbf{B}(\alpha \rightarrow \beta) \rightarrow \mathbf{G}\beta$$

The problem with this property is not that it presupposes unreasonable abilities for the agent, like the axiom K for belief that states that the agent knows all logical consequences of its knowledge, but that it just does not correspond to any definition of goal. Let us take an example from [9]. If an agent has the goal of having his tooth filled, and the agent believes the treatment will cause him pain, then one would not want to conclude that the agent has the goal of feeling pain. It is not a question of stating that in "some way" the agent wants to feel pain, because besides this tooth filling treatment, the agent will not seek to get pain by other means. What happens is that the agent considers it more important to have the tooth filled than to avoid pain.

A second problem of the standard semantic for goals is that it does not model the important characteristic that goals have different priorities. Agents continually have to decide which action to take in a situation when goals are potentially in conflict, and they decide based on the priority of each of the conflicting goals. For example, it is reasonable to assume that the agent above has the goal of *avoiding* pain (for example, he demands anesthetics in major surgery, and so on). So, not only not getting pain is not one of the agent's goals, but in fact on of his goals is not to get pain. The agent was faced with a conflict of goals: getting the tooth filled was in conflict to avoiding pain, and the agent decided to pursue the first goal because it had a higher priority than the second.

This paper will propose a semantic model of goals that addresses the problem of closure under believed implication and which provides a very intuitive definition of priority of goals. We will proceed by defining wishes, and its relation to beliefs. Then we will discuss the semantics of the priority relation, and define top level goals in relation to wishes. Finally the paper will extend the logic from a single agent to many agents.

2 Formal Aspects

We will develop the semantics of a logic $\mathcal{G}$ that deals with wishes, goals, belief and priorities. This logic will be based on an underlying modal temporal logic, which we will call LTL, since goals and time are strongly connected. For examples, only propositions in the future can be goal: one cannot have a goal of having been born in a different year because that cannot be changed (although one may wish to having been born in a different year).

We will first develop the basic underlying temporal logic, then define wishes, discuss its relation to beliefs, define goals and finally priorities.

2.1 The underlying temporal logic

We will use a modal linear-time, discrete temporal logic [3], which we will call LTL, on top of which we will define the semantics of belief and goals. We will define a **time-line**, or a **world**, as an infinite set of states, where each state is a function that attributes true or false to all propositional symbols in the language; and a mapping from its states to the integers. We will denote a world as the sequence $w = (\ldots x_{-k} \ldots x_{-1} x_0 x_1 \ldots x_j \ldots)$ where x_i is the state that corresponds to instant i.

In this logic one can define the temporal operators: $\overrightarrow{\Diamond}\,\alpha$ which states that α will eventually hold in a future state; $\overrightarrow{\Box}\,\alpha$ which states that α will always hold in future states; $\overleftarrow{\Diamond}\,\alpha$ which states that α did hold sometime in the past; and $\overleftarrow{\Box}$

[1] Department of Computer Science, State University of Campinas, Campinas, SP 13081-970, BRAZIL. E-mail: wainer@dcc.unicamp.br. This research was partially funded by CNPq grant 521464/93-5.

ECAI 94. *11th European Conference on Artificial Intelligence* Edited by A. Cohn
Published in 1994 by John Wiley & Sons, Ltd.

which states that α has always hold in the past. One can also define other temporal operators like **next**, **until**, **before**, in t moments in the future $\overleftarrow{\Diamond}_t$ and so on [3]. The definition of some of these operators is as follows, where $\models$ should be interpreted as $\models_{LTL}$

$$w,i \models \alpha \quad \text{iff} \quad x_i \models \alpha \text{ and } \alpha \text{ has no temporal operator}$$
$$w,i \models \overrightarrow{\Box}\, \alpha \quad \text{iff} \quad \forall j > i \quad w,j \models \alpha$$
$$w,i \models \overrightarrow{\Diamond}\, \alpha \quad \text{iff} \quad \exists j > i \quad w,j \models \alpha$$
$$w,i \models \overrightarrow{\Diamond}_t\, \alpha \quad \text{iff} \quad w,i+t \models \alpha$$

A formula α is satisfiable if there is a world such that $w, 0 \models \alpha$, and a formula is valid if for all worlds w, $w, 0 \models \alpha$.

2.2 Relations among worlds

In order to define the logic $\mathcal{G}$ that deals with wishes, beliefs, goals and priorities, we need to consider the set of all possible worlds (time-lines) given a certain set of basic propositional symbols, which we call W, and to define two relations among the element of this set. The relation A is the standard accessibility relation for belief, that is, A is serial, transitive and Euclidean, which corresponds to a logic of belief that satisfy the modal logic KD45 [8].

The $>$ relation is a preference relation between worlds. "$a > b$" states that the agent would rather be in world a than in world b, or in other words, the agent considers world a more desirable than world b. The preference relation is transitive and assymetric.

Let us consider now the case of two worlds that cannot be compared in terms of the relation $>$, that is, neither $a > b$ nor $b > a$ holds, which we will denote by $a \approx b$. In such a case, from the agent's point of view, both a and b are equally good worlds, and there is nothing "that matters for the agent" that distinguish them. Thus if $a > c$ then it must be the case that also $b > c$ since whatever made a better than c must be also present in b since they are indistinguishable from the point of view of "things that matters to the agent." The same reasoning holds if $c > a$. Therefore, besides transitivity and assymmetry, the $>$ relation must have the property that if $a \approx b$ and $a > c$ then $b > c$, and if $d > a$ then $d > b$. Finally we will require the $>$ relation to have an upper bound, that is, there is at least one world z such that there is no other world z' and $z' > z$. This last requirement will be necessary for the definition of wishes, but it imposes a strong constraint on the preference relation since there are uncountably many possible worlds. For example the preference relation cannot account for things like "I would prefer to live as much as possible," since such preference relation would have no upper bound.[2]

Due to these properties, the preference relation $>$ can be seen as a linear order of clusters of equally preferred worlds (figure 1). One can divide the set of all possible worlds W into disjoint partitions or clusters $P_0, P_1, \ldots P_k, \ldots$, where each P_i is a maximal set of equally preferred worlds. P_0 is the set of all maximally preferred worlds, P_1 is the set of all next preferred worlds, and so on. There may or may not be a least preferred partition. Two worlds in the same cluster are are equally preferred, and they are preferred to all worlds in all sets below them.

Furthermore, we can also see the relation A as a set of mutually accessible possible worlds, or better, a set of epistemically possible worlds (that is, worlds that the agent considers equally possible descriptions of the reality). Let us call this set $\overline{A}$.

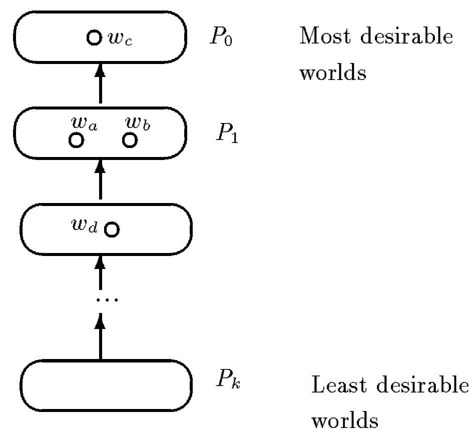

Figure 1.

2.3 Wishes

We define the agent's wishes (represented by the modal operator **W**) as the formulas that are satisfied in all maximally preferred worlds. The intuition is that wishes correspond to partial descriptions of all the worlds the agent thinks are the most desirable ones.

Definition 1 $\langle W, w_0, A, > \rangle \models \mathbf{W}\alpha$ *iff* $\langle W, w, A, > \rangle \models \alpha$, *for all* w *such that there is no* w', $w' > w$. *Or using the partitions notation:* $\langle W, w_0, A, > \rangle \models \mathbf{W}\alpha$ *iff* $\langle W, w, A, > \rangle \models \alpha$, *for all* $w \in P_0$

We must point out that the temporal component of the logic does not include the wish operator (or the other modal operators to be defined below). Thus, in the logic $\mathcal{G}$ one can express a wish of having been rich ($\mathbf{W} \overleftarrow{\Diamond} r$), one cannot express having had the wish of being rich ($\overleftarrow{\Diamond}\, \mathbf{W} r$).

Wishes have the following properties:

$$\mathbf{W}\alpha \text{ if } \alpha \text{ is a LTL tautology} \quad (1)$$
$$\mathbf{W}\alpha \rightarrow \neg\mathbf{W}\neg\alpha \quad (2)$$
$$\mathbf{W}\alpha \wedge \mathbf{W}(\alpha \rightarrow \beta) \rightarrow \mathbf{W}\beta \quad (3)$$
$$\mathbf{W}\alpha \leftrightarrow \mathbf{W}\mathbf{W}\alpha \quad (4)$$
$$\neg\mathbf{W}\alpha \leftrightarrow \mathbf{W}\neg\mathbf{W}\alpha \quad (5)$$

(1) states that something that is necessarily true is a wish. (2) states that wishes are consistent, and (3) that wishes are closed under implication. (4) needs some more further elaboration. The direction that seems interesting is $\mathbf{W}\mathbf{W}\alpha \rightarrow \mathbf{W}\alpha$,

[2] A technical and possibly unsatisfying solution for this problem would be to limit the diversity of possible worlds by defining a possible world to be a finite set of states. In this case the temporal aspect of the logic would be bounded, and one could only refer say to a million years into the future and into the past, and not to the whole eternity.

which states that if one has a wish of having a wish, then one have this wish. The intuition for this property follows from reading the $\rightarrow$ as "causes," and from seeing a wish as a "cognitive action" that is very easy to achieve: a wish can appear in one's mind just by wishing that wish. The other direction of the biconditional is less intuitive but it is probably harmless. The same holds for (5), but in this case referring to not having a wish.

2.4 Beliefs and Wishes

The agent's beliefs ($\mathbf{B}$) is defined as usual based on the relation A:

Definition 2 *$\langle W, w_0, A, > \rangle \models \mathbf{B}\alpha$ iff $\langle W, w, A, > \rangle \models \alpha$, for all w such that $(w_0, w) \in A$. Or using the set $\overline{A}$ approach: $\langle W, w_0, A, > \rangle \models \mathbf{B}\alpha$ iff $\langle W, w, A, > \rangle \models \alpha$, for all $w \in \overline{A}$.*

The definitions above imply that the knowledge operator has the usual KD45 properties [8].

The relation of wishes and beliefs is captured by the following axioms:

$$\mathbf{W}\alpha \leftrightarrow \mathbf{B}\mathbf{W}\alpha \tag{6}$$
$$\neg\mathbf{W}\alpha \leftrightarrow \mathbf{B}\neg\mathbf{W}\alpha \tag{7}$$

That is, wishes and lack of wishes are available for introspection.

It is important to point out that wishes are not closed under believed implication. That is:

$$\mathbf{W}\alpha \wedge \mathbf{B}(\alpha \rightarrow \beta) \not\models_{\mathcal{G}} \mathbf{W}\beta$$

On the other hand, wishes are closed under logical implication. That is, if $\alpha \rightarrow \beta$ is a LTL tautology, then if α is a wish, β will also be a wish. For example, if $\mathbf{W}(\alpha \wedge \beta)$ then both $\mathbf{W}\alpha$ and $\mathbf{W}\beta$.

2.5 Priorities

Based on the $>$ relation it is possible to define an ordering of preference among propositions for the agent. We will use a new binary modal operator $\sqsupset$ (called priority relation) to state that a proposition has higher preference or priority than another proposition. Informally, $\alpha \sqsupset \beta$ states that the agent prefer worlds where β is false and α is true to worlds where α is false and β is true. Formally:

Definition 3 *$\langle W, w_0, A, P \rangle \models \alpha \sqsupset \beta$ for α and β LTL formulas iff for all worlds w such that $w, 0 \models_{LTL} \neg\alpha \wedge \beta$ there is a world w' such that $w', 0 \models_{LTL} \alpha \wedge \neg\beta$ and $w' > w$.*

The "$\sqsupset$" operator has the following properties:

$$\text{if } \alpha \sqsupset \beta \text{ and } \beta \sqsupset \gamma \text{ then } \alpha \sqsupset \gamma \tag{8}$$
$$\text{if } \models_{LTL} \alpha \leftrightarrow \gamma \text{ and } \alpha \sqsupset \beta \text{ then } \gamma \sqsupset \beta \tag{9}$$
$$\text{if } \models_{LTL} \beta \leftrightarrow \gamma \text{ and } \alpha \sqsupset \beta \text{ then } \alpha \sqsupset \gamma \tag{10}$$
$$\text{if } \alpha \sqsupset \beta \text{ then } \mathbf{B}(\alpha \sqsupset \beta) \tag{11}$$

Priorities are connected with choices. If a agent comes to believe that if p were true, then eventually α would be true, and β would be false, and if p were false then eventually α would be false, and β would be true, and if $\alpha \sqsupset \beta$, then the agent should choose to pursue p. Unfortunately the language of $\mathcal{G}$ is not expressive enough to represent choice (but see the discussion of goals below), and thus we cannot make use of such intuition in a formal definition. Nevertheless, we believe that the semantic definition of priorities is intuitive enough and it can serve as a stepping stone for further investigation.

3 Goals

Defining goals is a harder task than defining wishes since the intuitions about goals are less clear. People seem to use the term goal as meaning propositions that would bring reality closer to one's wishes. It follows that there is no clear distinction between goals and subgoals since if one thinks that α would bring reality closer to one's wishes, and β (among other propositions) would bring α to be true, then β will bring reality closer to the wishes, and thus β is a goal. A complete semantic definition of of goals would probably include aspects of a causal theory [11, 5], and autoepistemic theory [10] so it can model the following reasonable recursive definition of goals (or subgoals): if the agent believes that *only β causes α*, and α is a goal, then β is also a goal. For example, β could a conjunction $\beta_1 \wedge \beta_2 \ldots$, which is the plan to achieve α, or a disjunction $\beta_1 \vee \beta_2$, which would represent alternative plans to achieve α.

We will not further develop these ideas beyond this point. Intead, we will concentrate on the relation between goals and wishes, or informally, the terminating case of the recursive definition of goals above. There are goals that are not explained as being part of a plan to achieve other goals, but refer directly to the agent's wishes. They can be seen as *top level goals*, but we will refer to them in the rest of the paper as simply *goals*.

Other complications about goals is that they are always about propositions in the future: one cannot have a goal of having been born in a different year, but one can have a goal of publishing seven articles before the end of the year.

3.1 Goals and wishes

There are two possible definitions for goals. The first one states that goals are wishes about the future that the agent considers possible. Formally:

Definition 4 *We will define* **goals**, *represented by the modal operator* $\mathbf{G}$, *as wishes that the agent believe are possible. That is:*

$$\mathbf{G}(\vec{\Diamond}\, \alpha) \quad \text{iff} \quad \mathbf{W}(\vec{\Diamond}\, \alpha) \wedge \neg\mathbf{B}\neg(\vec{\Diamond}\, \alpha) \text{ and}$$
$$\mathbf{G}(\vec{\Box}\, \alpha) \quad \text{iff} \quad \mathbf{W}(\vec{\Box}\, \alpha) \wedge \neg\mathbf{B}\neg(\vec{\Box}\, \alpha)$$

We will call goals of the form $\mathbf{G}\,\vec{\Diamond}\,\alpha$ as **achievement** goals, and goals of the form $\mathbf{G}\,\vec{\Box}\,\alpha$, **permanent** goals.[3] In this paper we will consider only achievement goals.

The second possible definition for goals is that goals are the formulas (about the future) that are satisfied by all maximally preferred worlds *that are considered possible.* Let us see

[3] There seems to be another kind of goals, the cyclic goals, a recurent goal whenever a situation repeats itself. For example, the goal of arriving early home after the working day. Cyclic goals seems to be of the form $\vec{\Box}\,(\beta \rightarrow \vec{\Diamond}\, \alpha)$, or $\vec{\Box}\,(\beta \rightarrow \vec{\Diamond}_t\, \alpha)$.

one example where the two definitions differ. Thus, instead of considering the intersection of most preferred worlds with the set of worlds considered possible, this second approach would look at the maximally preferred worlds in the set $\overline{A}$. In this paper we will concentrate on the first definition, since it has simpler formal properties. We are currently exploring the consequences of the second definition.

Under the first definition goals exhibit the following properties:

$$\mathbf{G} \vec{\Box}\, \alpha \quad \text{if } \alpha \text{ is a LTL tautology} \tag{12}$$

$$\mathbf{G}\alpha \rightarrow \neg\mathbf{G}\neg\alpha \tag{13}$$

$$\mathbf{G}\alpha \wedge \mathbf{G}(\alpha \rightarrow \beta) \quad \rightarrow \quad \mathbf{G}\beta \tag{14}$$

$$\mathbf{G}\alpha \leftrightarrow \mathbf{G}\mathbf{G}\alpha \tag{15}$$

$$\neg\mathbf{G}\alpha \leftrightarrow \mathbf{G}\neg\mathbf{G}\alpha \tag{16}$$

The proofs are based on the equivalence $\mathbf{G}\alpha \equiv \mathbf{W}\alpha \wedge \neg\mathbf{B}\neg\alpha$.

Like wishes, goals are not closed under believed implications, but they are closed under tautological implication, so for example, $\mathbf{G}(\alpha \wedge \beta)$ implies both $\mathbf{G}\alpha$ and $\mathbf{G}\beta$.

4 Tentative Extension for Multi-agents

The logic of goals above does not easily expand to a multi-agent logic. The obvious expansion, creating an accessibility relation A_i and a preference relation P_i for each agent i, will not yield a logic with reasonable properties. In this logic, which we call n-$\mathcal{G}$, one can define belief, wishes, goals and goal priority for each agent.

The main problem of the logic n-$\mathcal{G}$ is the relation of one agent's belief and wishes with another agent's wishes. Under such semantics, the formulas below would be are valid in the logic n-$\mathcal{G}$.

$$\mathbf{W}_i\alpha \rightarrow \mathbf{B}_j\mathbf{W}_i\alpha$$
$$\mathbf{W}_i\alpha \rightarrow \mathbf{W}_j\mathbf{W}_i\alpha$$

In particular, the first formula states that an agent's wishes are known by all other agents, and indeed one can prove that they are common knowledge, which is clearly an unreasonable property.

These properties do not carry over to goals, because of the belief (or lack of belief) component of the definition of goals. Therefore $\mathbf{G}_i\alpha$ would not imply $\mathbf{B}_j\mathbf{G}_i\alpha$ or $\mathbf{G}_j\mathbf{G}_i\alpha$ in general. Thus, although I know that you wish α, I may not think that you believe α to be possible, and thus I would not believe that α is one of your goals. Of course, although the result is reasonable, that is, I do not know what are your goals, the reason for that result is the wrong one. Goals should not be mutually known because goals, like beliefs, are private mental aspects of each agent.

4.1 Recursive Kripke models for goals

In order to accommodate many agents without the unreasonable properties of the logic n-$\mathcal{G}$ we will use the device of recursive Kripke structures [7]. The intuition is simple: we believe that the model of wishes, goals and beliefs described for the logic $\mathcal{G}$ is very intuitive when it refers to just one agent. What needs to be reworked is the model of what does it means for an agent to have a metal model of another agent. If the pair of relations A and $>$ capture so well the intuition of the mental state of a single agent, then for an agent to have a mental model of another agent corresponds to assigning a pair $\langle A, >\rangle$ to that agent. The pair $\langle A, >\rangle$ would describe the wishes and beliefs of that agent. Thus, each possible world the agent considers it possible (that is, in the A relation) should not only be a possible time-line, but should also include an assignment of pairs $\langle A_i, >_i\rangle$ for each agent i (the other agents and itself). Thus, a world will be accessible for agent 1 if it is a possible description of reality as agent 1 sees it, and also if that world attributes pairs $\langle A, >\rangle$ in such a way that they describe agent 1's view of the other agents's (and itself) mental states. Similarly for the preference relation: the agent does not only rank two worlds in terms of the reality they describe, but also in terms of what mental state (pairs $\langle A, >\rangle$) each world assigns to all agents.

The agglutination of a possible world (that is, a time line) and n pairs of relations $\langle A_i, >_i\rangle$ is called an augmented possible world. Therefore, the relations A and $>$ should be relations among augmented possible worlds.

Definition 5 *An* **augmented possible world** a^j *is a tuple:*

$$\langle w^j, \langle A_1^j, >_1^j\rangle, \langle A_2^j, >_2^j\rangle, \ldots \langle A_n^j, >_n^j\rangle\rangle$$

where the first entry is a possible world (time-line), and each of the pairs $\langle A_i^j, >_i^j\rangle$ *corresponds to an accessibility and a preference relation on* all *augmented possible worlds for agent* i.

Formally, let the set W_ω be the set of all augmented possible worlds, then both A_i^j and $>_i^j$ are subsets of $W_\omega \times W_\omega$. The relations A_i^j and $>_i^j$ should also exhibit the properties discussed in the previous section, that is A_i^j is serial, transitive, and Euclidean; and $>_i^j$ should partition the set W_ω into sets of equally preferred augmented possible worlds, with an maximally preferred partition.

The definition of augmented possible worlds is recursive and depends on relations defined over the set of all augmented possible worlds. For the moment we do not have a proof that this set is well defined, but we will proceed in this paper with such assumption.

Satisfiability is defined below:

- $a^j \models p$ iff $w^j, 0 \models_{LTL} p$ where p is a LTL formula
- $a^j \models \mathbf{B}_i\alpha$ iff $a^j \models \alpha$ for all augmented possible worlds a^k such that $(a^j, a^k) \in A_i^j$
- $a^j \models \mathbf{W}_i\alpha$ iff $a^j \models \alpha$ for all augmented possible worlds a^j such that a^j is maximally preferred in $>_i^k$,

where $a^j = \langle w^j, \langle A_1^j, >_1^j\rangle, \ldots \langle A_n^j, >_n^j\rangle\rangle$. Goals and the priority relation $\sqsupset$ are defined as before, but are indexed by the agent. Validity is defined as the satisfiability by all augmented possible worlds.

Given the definitions above, goals and wishes satisfy the properties (1) to (3), for each of the operators $\mathbf{W}_i$, and the similar properties for goals. But to satisfy formulas that contain subformulas of different depth (roughly, the number of nested modal operators) like $\mathbf{G}_i\alpha \rightarrow \mathbf{B}_i\mathbf{G}_i\alpha$, one needs further constraints on the relations A_i^j and $>_i^j$. To describe these constraints it is convenient to see the relation A_i^j as a set of augmented possible worlds, as we did for the single agent case, written as $\overline{A}_i^j$. It will be also convenient to refer to the set of maximally preferred worlds in a relation $>_i^j$, which will be written as P_{0i}^j, that is, the most preferred partition of $>_i^j$. If

$$a^a = \langle w^a, \langle \overline{A}_1^a, >_1^a\rangle, \ldots \langle \overline{A}_n^a, >_n^a\rangle\rangle$$
$$\text{and } a^b = \langle w^b, \langle \overline{A}_1^b, >_1^b\rangle, \ldots \langle \overline{A}_n^b, >_n^b\rangle\rangle.$$

then we should have the following constraints:

$$\text{for all } a^b \in \overline{A}_i^a, \text{ then } \overline{A}_i^b = \overline{A}_i^a$$

$$\text{for all } a^b \in P_{0i}^a, \text{ then } P_{0i}^a = P_{0i}^b$$

$$\text{for all } a^b \in \overline{A}_i^a, \text{ then } >_i^a = >_i^b$$

The first constraint forces the introspection properties of the belief operator to be true for each agent. That is $\mathbf{B}_i\alpha \rightarrow \mathbf{B}_i\mathbf{B}_i\alpha$ and $\neg\mathbf{B}_i\alpha \rightarrow \mathbf{B}_i\neg\mathbf{B}_i\alpha$ are both valid.

The second constraint forces the equivalent of introspection for the wish operator for each agent, that is, properties (4) and (5). This constraint can be weaken to, for example, get rid of the non-intuitive property that $\mathbf{W}_i\alpha \rightarrow \mathbf{W}_i\mathbf{W}_i\alpha$, while keeping the property that $\mathbf{W}_i\mathbf{W}_i\alpha \rightarrow \mathbf{W}_i\alpha$. This weaker version of (4) is the result of changing the second constraint above to:

$$\text{for all } a^b \in P_{0i}^a, \text{ then } P_{0i}^a \subseteq P_{0i}^b$$

Finally, the third constraint is responsible for properties (6), (7), and (11), that mix wishes and the priority operator $\sqsupset_i$ with the belief operator.

5 Conclusions

This paper presents a first step in a formalization of goals that, in terms of the power of the logic, stands between the standard formalization (for example [2] and the minimal formalization of [9]. Furthermore, the formalization proposed here allows one to define a very clear semantics for the concept of priorities among propositions in general or among goals in particular.

The first incompleteness of the research reported in this paper is the extension to multiagents: the application of the concepts of recursive Kripke-structures [7] to model goals and beliefs of many agents. We believe that the use of recursive Kripke structures provides a clear intuition of modeling the mental state of agents reasoning about other agents' mental states, but the technical parts of such approach are incomplete. In particular, as we mentioned above, it is not proven that augmented possible worlds are well defined. It is possible that the ideas of recursive Kripke structures can be casted in the form of the knowledge structures (with an added dimension for the preference relation) of [4], which are incrementally constructed, and thus are well defined.

The research also have to advance on a broader definition of goals, one that includes causal and auto-epistemic aspects, and further understand the different kinds of goals: permanent, cyclic and achievement goals and how they interact with the concept of priorities.

Finally, the semantics presented here have some interesting and yet unexplored connections with two forms of nonmonotonic reasoning: conditional logics [6, 1] and preference logics [12, 14]. The structure of the semantics of the logic $\mathcal{G}$ is equivalent to the semantics CO* for conditional logics presented in [1], and the definition of the priority operator have some similarities with the conditional operator defined therein. On the other hand, if one sees the preference relation > not as a preference among worlds, but as a preference among models (of LTL), the second definition of goals mentioned in this paper would be very similar the definition preference entailment.

REFERENCES

[1] Craig Boutilier, 'Conditional logics for default reasoning and belief revision', Technical Report KRR–TR–92–1, University of Toronto, Computer Science Department, (1992).

[2] Philip R. Cohen and Hector Levesque, 'Intention is choice with commitment', *Artificial Intelligence*, **42**(3), (1990).

[3] E. A. Emerson, 'Temporal and modal logics', in *Handbook of Theoretical Computer Science*, ed., J. van Leeuwen, volume Volume B, Formal Methods and Semantics, 995–1072, Elsevier Science Publishers and The MIT Press, (1990).

[4] R. Fagin, J. Y. Halpern, and M. Y. Vardi, 'A model-theoretic analysis of knowledge', *Journal of the ACM*, **2**, 382–428, (April 1991).

[5] H. Geffner, *Default Reasoning: Causal and Conditional Theories*, MIT Press, 1991.

[6] M. L. Ginsberg, 'Counterfactuals', *Artificial Intelligence*, **30**(1), 35–79, (1986).

[7] P. J. Gmytrasiewics and E. H. Durfee, 'A logic of knowledge and belief for recursive modeling: Preliminary report', in *Proc. of the Tenth National Conference on Artificial Intelligence*, pp. 628–634. AAAI Press/The MIT Press, (1992).

[8] J. Y. Halpern and Y. Moses, 'A guide to completeness and complexity for modal logics of knowledge and belief', *Artificial Intelligence*, **54**(3), 319–379, (1992).

[9] K. Konolige and M. Pollack, 'A representationalist theory of intention', in *Proc. of the 13th International Joint Conference on Artificial Intelligence*, pp. 390–395, (1993).

[10] Hector J. Levesque, 'All I know: A study in autoepistemic logic', *Artificial Intelligence*, **42**, 263–310, (1990).

[11] D. Lewis, 'Causation', *Journal of Philosophy*, **70**, 556–567, (1973).

[12] John McCarthy, 'Applications of circumscription to formalizing common-sense reasoning', *Artificial Intelligence*, **28**, 89–116, (1986).

[13] A. S. Rao and M. P. Georgeff, 'Modelling rational agents within a bdi-architeture', in *KR91*, Cambridge, MA, (1991).

[14] Y. Shoham, 'A semantical approach to non-monotonic logics', in *Proc. of the Tenth International Joint Conference on Artificial Intelligence*, pp. 388–392, (1987).