

‘What I fail to do today, I have to do tomorrow’: a logical study of the propagation of obligations

Jan Broersen¹ and Julien Brunel²
broersen@cs.uu.nl brunel@irit.fr

¹ Institute of Information and Computing Sciences
Utrecht University, The Netherlands

² Institut de Recherche en Informatique de Toulouse
University of Toulouse, France

Abstract. We study a logical property that concerns the preservation of future directed obligations that have not been fulfilled yet. We call this property ‘propagation property’. The goal is to define a combination of temporal and deontic logics which satisfies this property. Our starting point is the product of temporal and deontic logics. We investigate some modifications of the semantics of the product in order to satisfy the propagation property, without losing too much of the basic properties of the product. We arrive at a semantics in which we only consider ideal histories that share the same past as the current one, and that enables an interesting characterization of the states in which obligations propagate: these are the states where there are no violations of present directed obligations.

1 Introduction

A strong intuition concerning the interaction of deontic and temporal modalities is that an obligation to achieve something in the future should propagate to future moments if it is not met presently. This is particularly true for deadline obligations; if I have to finish my paper before the end of the week, and I do not finish it today, tomorrow I still have to finish it before the end of the week. But, the propagation property also pertains to future directed obligations without a deadline: if today I need to give a party someday, and I do not give the party today, then tomorrow I still have to give the party someday.

We want to emphasize that such properties are only valid if we assume that the ‘deontic realm’ is not changed by an explicit update of the norms that construct it. What we call the ‘deontic realm’ is relative to the actual situation and to an external body of norms that determines what the obligations of the agent in the given situation are (e.g. a lawbook, or some other set of rules the agents have to comply to). If we allowed the body of norms to vary, the propagation property would not hold. For instance, to come back to the above mentioned examples, there may be an unexpected extension of the deadline for the paper, or my friends have waited in vain for too long and no longer demand that I give the party. This means that I no longer have to finish the paper

before the end of the week, and, that I no longer have to organize the party. In such cases, the preservation of the original obligations is not prolonged, due to a change of the norms, and accordingly, a change of the deontic realm. In this paper we will not be concerned with these explicit updates of the norms; we only consider logical properties for the case where the norms are settled, and where it makes sense to reason about the preservation and propagation of the resulting obligations on the basis of what actually happens. So, the only changes of the deontic realm we consider are the ones due to changes of the *situation*.

The problem of propagation of obligations is an instance of the more general problem of the interaction of 'what is obligatory' with 'what is actually the case'. In deontic logic such interactions are only considered sporadically. For instance, in SDL [18], we do not have the interaction property $\mathbf{O}(\varphi \vee \psi) \wedge \neg\varphi \Rightarrow \mathbf{O}(\psi)$, although it might be considered quite reasonable: if I have to be smart or strong, and in fact I am not strong, I have to be smart. A possible ground for not wanting this property is that in combination with 'weakening', that is $\mathbf{O}(\varphi) \Rightarrow \mathbf{O}(\varphi \vee \psi)$, which is valid in SDL, we get that $\mathbf{O}(\varphi) \wedge \neg\varphi \Rightarrow \mathbf{O}(\psi)$. Thus by the combination of weakening and the proposed interaction of obligations with conditions being settled, we get that when there is a violation, everything is obligatory. Then, two reactions are possible: (1) indeed this property is bad, and we have to see what we can do to avoid it (while keeping the interaction with facts, in which we are interested), or (2) maybe this property is not as bad as it seems. Below we elaborate on both reactions.

If we want to avoid $\mathbf{O}(\varphi) \wedge \neg\varphi \Rightarrow \mathbf{O}(\psi)$ the obvious choice would be to attack weakening. Ross' famous paradox [17] indeed questions this property. Ross' paradox just says that $\mathbf{O}(\varphi) \Rightarrow \mathbf{O}(\varphi \vee \psi)$ is not intuitive under some readings. For instance, being obliged to post a letter does not imply being obliged to post or burn it. But many deontic logicians (see e.g., [11]) have argued that the paradox is due to a naive interpretation of the formula. If we read $\mathbf{O}(\varphi)$ properly as 'ϕ is a necessary condition of any state that is optimal according to ones obligations', then the property poses no problems. Another way to say this is that $\mathbf{O}(\varphi)$ expresses an 'at least' reading of what is obligatory: it is obligatory to at least satisfy ϕ, but maybe other properties also have to be obligatory at the same time.

Another way to avoid $\mathbf{O}(\varphi) \wedge \neg\varphi \Rightarrow \mathbf{O}(\psi)$ is to refine the interaction between 'what is obligatory' with 'what is actually the case', i.e., to specialise the interaction property $\mathbf{O}(\varphi \vee \psi) \wedge \neg\varphi \Rightarrow \mathbf{O}(\psi)$. Our point is to consider that what happens today can only have an effect on what will be obligatory tomorrow, and not on what is obligatory today. We follow the idea that one time step is needed so that the deontic realm takes into account what happens. According to this reading, if I have to be smart or strong today, then the fact I am in one of the four possible situations (strong and smart, strong and not smart, etc.) does not change my obligation to be either smart or strong. In one of these situations (neither strong nor smart) the obligation is violated, while it is fulfilled in the other ones. On the other hand, things change in a temporal context. Indeed, if I have the obligation to be in Paris today or in Amsterdam tomorrow, then the

obligation I will have tomorrow will depend on what I do today: if I am not in Paris today, then tomorrow I will have the obligation to be in Amsterdam, otherwise I will not. (Whether I am in Paris or not today, today's obligation does not change, only tomorrow's obligation does.) It is closely related to the fact that an obligation only concerns the present or the future: an obligation that yesterday I was in Paris does not make sense. Thus, in case I am not in Paris today, tomorrow's obligation will not be 'to be in Amsterdam or to have been in Paris yesterday' but simply 'to be in Amsterdam'. So in this paper the interaction property we will consider is not $\mathbf{O}(\varphi \vee \psi) \wedge \neg\varphi \Rightarrow \mathbf{O}(\psi)$ but $\mathbf{O}(\varphi \vee X\psi) \wedge \neg\varphi \Rightarrow X\mathbf{O}(\psi)$, and we will call it propagation property.

As a further motivation for this work we will first point to some related problems. Preservation properties have been studied for intentions. However, intentions are preserved for a different reason. As Bratman [3, 4] explains, intentions serve to stabilize an agent's deliberations. An agent cannot continuously reconsider his decisions, simply because usually there is no time to do that. It is usually more rational to stick to reached decisions (thereby turning them into intentions), and to only let achievement of that what is intended be a cause for discharging the obligation. In AI (that is, AI as studied in computer science), the best-known formalizations of rationality postulates for the phenomenon of intention preservation are Rao and Georgeff's 'commitment strategies' [15]. For obligations, the reason to preserve them is different: they are preserved simply because an agent has to meet up to his obligations at some point, unless, of course, he is explicitly relieved from his obligations. But, as said, we do not consider that issue in this paper. Another motivating example is the preservation of goals in the mechanisms underlying agent programming languages like AgentSpeak [16] and 3APL [12, 8]. Such agent programming languages usually comprise programming rules that work on data structures for beliefs, plans and goals. The goals in the 'goal bases' of agent programs are often sets of propositional formulas denoting properties the agent aims at making true. The operational semantics of the programming languages typically treats the goals in the goal bases as information that has to be preserved unless a state is reached where the agent believes the goal is achieved. Of course, goals are not identical to obligations. But it is clear that they at least share the phenomenon of propagation. One of the motivations is thus to provide a logical basis for verification of agent programs against logically formulated rationality postulates about the propagation of goals and obligations.

The present paper is organised as follows. Section 2 presents the product of Linear Temporal Logic (*LTL*) [14] and Standard Deontic Logic (*SDL*) [18], which is an appropriate starting point for our investigation. Section 3 shows that the propagation property is not compatible with the genuine product and presents modifications of the product semantics which guarantee the propagation of obligations. Section 5 concludes the paper.

2 Product of temporal and deontic logic

We present here standard deontic logic *SDL* [18], linear temporal logic *LTL* [14], and their product logic $LTL \times SDL$. We choose the product as the starting point because the commutativity properties, which are specific for products, correspond to a setting without norm updates. As we explained in the introduction we are not interested in explicit updates here.

2.1 Deontic and temporal logics

Deontic logic is the modal logic of obligation, permission, and prohibition.

Definition 1 (Deontic language). *The deontic language \mathcal{DL} is defined as*

$$\mathcal{DL} ::= P \mid \perp \mid \mathcal{DL} \Rightarrow \mathcal{DL} \mid \mathbf{O}(\mathcal{DL})$$

The necessity operator \mathbf{O} is read it is obligatory that. The possibility operator $\mathbf{P} \stackrel{def}{=} \neg \mathbf{O} \neg$ is read it is permitted that. We can read $\mathbf{O}(\neg\varphi)$ as it is prohibited that φ .

The boolean operators are defined as usual:

$$\neg\varphi \stackrel{def}{=} \varphi \Rightarrow \perp \quad \top \stackrel{def}{=} \neg\perp \quad \varphi_1 \vee \varphi_2 \stackrel{def}{=} \neg\varphi_1 \Rightarrow \varphi_2 \quad \varphi_1 \wedge \varphi_2 \stackrel{def}{=} \neg(\varphi_1 \Rightarrow \neg\varphi_2)$$

The truth-relation \models between a world of a Kripke model and a formula is given through the usual possible world semantics, with which we consider the reader is familiar. Standard Deontic Logic *SDL* can be defined as the logic over the language \mathcal{DL} determined by the Kripke frames (W, R) such that R is serial, i.e. as the set of the \mathcal{DL} -formulas that are valid in every frame (W, R) such that R is serial. In a deontic frame, W is the set of the worlds, and the intuitive reading of R is that it associates each world with a set of ideal worlds, in which every obligation is fulfilled.

Definition 2 (Temporal language). *We consider a monadic operator X called next, and a dyadic operator U called until. Given a set P of atomic propositions, the temporal language \mathcal{TL} is defined as*

$$\mathcal{TL} ::= P \mid \perp \mid \mathcal{TL} \Rightarrow \mathcal{TL} \mid X \mathcal{TL} \mid \mathcal{TL} U \mathcal{TL}$$

The informal meaning of the temporal operators X and U are as follows:

$X\varphi$: “at the next moment, φ will hold.”

$\varphi_1 U \varphi_2$: “ φ_2 will eventually hold at some moment m , while φ_1 holds from now until the moment before m .”

The temporal operators F (finally, or eventually), G (globally, or always) and $F_{\leq k}$ (before k time units) are defined as the following abbreviations (the boolean operators are defined as for \mathcal{DL}):

$$F \varphi \stackrel{def}{=} \top U \varphi \quad G \varphi \stackrel{def}{=} \neg F \neg \varphi \quad F_{\leq k} \varphi \stackrel{def}{=} \begin{cases} \varphi & \text{if } k = 0 \\ \varphi \vee X F_{\leq k-1} \varphi & \text{else} \end{cases}$$

Definition 3 (Linear temporal structure and model). We consider here a linear, infinite, and discrete time. The unique temporal frame is $(\mathbb{N}, <)$ where

- \mathbb{N} is the set of the natural numbers.
- $<$ is the usual strict order on the natural numbers.

Given a set of atomic propositions P , a temporal valuation V is a function $V : \mathbb{N} \rightarrow 2^P$ which associates each state with a set of atomic propositions.

Let us define the satisfaction relation between a state of a model and a temporal formula.

Definition 4 (Satisfaction). Given a set of atomic propositions P , a temporal valuation V , a moment $i \in \mathbb{N}$, and a formula φ of \mathcal{TL} , the satisfaction relation \models is defined by induction on φ as follows:

$$\begin{aligned}
 i \models p & \quad \text{iff} \quad p \in V(i) \quad \text{where } p \in P \\
 i \not\models \perp & \\
 i \models \varphi_1 \Rightarrow \varphi_2 & \quad \text{iff} \quad \text{if } i \models \varphi_1 \text{ then } i \models \varphi_2 \\
 i \models X\varphi & \quad \text{iff} \quad i + 1 \models \varphi \\
 i \models \varphi_1 U \varphi_2 & \quad \text{iff} \quad \exists i' \geq i \text{ such that } i' \models \varphi_2 \text{ and} \\
 & \quad \forall i'' \in \mathbb{N} \text{ if } i \leq i'' < i' \text{ then } i'' \models \varphi_1
 \end{aligned}$$

The logic formulated in the language \mathcal{TL} which is determined by the unique frame $(\mathbb{N}, <)$ is called *LTL* (*Linear Temporal Logic*)[14].

2.2 Temporal and deontic product

We define here the product of temporal and deontic logics. The product frames correspond to the usual product definition [10] (see figure 1 for an illustration).

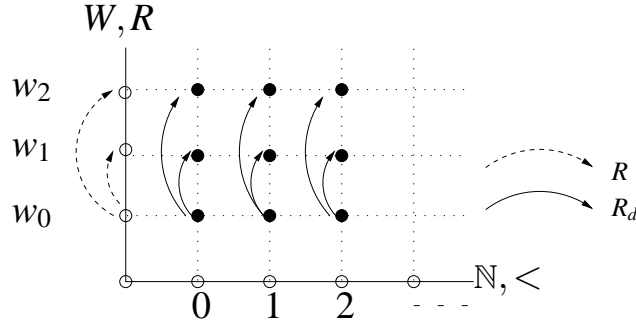


Fig. 1. Illustration of the product $(\mathbb{N}, <) \times (W, R)$

Definition 5 (Product frame, product model). Let $T = (\mathbb{N}, <)$ and $D = (W, R)$ be respectively a temporal frame and a deontic frame. Then the product frame $T \times D$ is a triple $(S, <_t, R_d)$ where

- $S = \mathbb{N} \times W$ (the set of the states) is the Cartesian product of the set \mathbb{N} of the natural numbers, viewed as a set of moments, and the set W of the worlds,
- $<_t \subseteq S \times S$ is the temporal relation on states such that $(i, w) <_t (i', w')$ if and only if $i < i'$ and $w = w'$,
- $R_d \subseteq S \times S$ is the deontic relation on the states such that $(i, w)R_d(i', w')$ if and only if wRw' and $i = i'$. We then say that w' is an ideal world of w at moment i , or that (i, w') is an ideal state of (i, w) .

Given a set P of atomic propositions, a valuation V for $T \times D$ is a function $V : S \rightarrow 2^P$ that associates each state with a set of atomic propositions. The pair $(T \times D, V)$ is then called a product model based on $T \times D$.

The language of the product logic $LTL \times SDL$ combines LTL operators and SDL operator.

Definition 6 (Syntax of $LTL \times SDL$). Given a countable set P of atomic propositions, the temporal deontic language \mathcal{TDL} of $LTL \times SDL$ is defined by:

$$\mathcal{TDL} ::= P \mid \perp \mid \mathcal{TDL} \Rightarrow \mathcal{TDL} \mid X(\mathcal{TDL}) \mid \mathcal{TDL} U \mathcal{TDL} \mid \mathbf{O}(\mathcal{TDL})$$

Usual boolean and temporal operators defined as abbreviations in the definition of temporal and deontic languages (definitions 1 and 2 respectively) are also available.

We can now define the satisfaction relation for the deontic and temporal product logic.

Definition 7 (Satisfaction). A formula φ of \mathcal{TDL} is interpreted on a state of a product model. Given a product model $((S, <_t, R_d), V)$, a state $s = (i, w) \in S$, and a formula φ , we can define the satisfaction relation \models by induction on φ :

$$\begin{aligned} s \models X\varphi & \quad \text{iff} \quad (i+1, w) \models \varphi \quad \text{where} \quad s = (i, w) \\ s \models \varphi_1 U \varphi_2 & \quad \text{iff} \quad \exists s' \geq_t s \quad \text{such that} \quad s' \models \varphi_2 \quad \text{and} \\ & \quad \forall s'' \in S \quad \text{if} \quad s \leq_t s'' <_t s' \quad \text{then} \quad s'' \models \varphi_1 \end{aligned}$$

where “ \leq_t ” is defined by $s \leq_t s'$ iff $s <_t s'$ or $s = s'$

$$s \models \mathbf{O}\varphi \quad \text{iff} \quad \forall s' \in S \quad \text{if} \quad sR_d s' \quad \text{then} \quad s' \models \varphi$$

A product model $((W, <_t, R_d), V)$ satisfies a formula φ if every state of the product satisfies it.

A product frame $F = (W, <_t, R_d)$ validates a formula φ if every model based on F satisfies it.

A formula φ is valid if every product frame validates it.

Let us elaborate on the interaction of the temporal and deontic dimensions. For instance, there is no difference between “it is permitted that φ holds tomorrow”, and “tomorrow, φ will be permitted”. This corresponds to the validity of $\mathbf{P}(X\varphi) \Leftrightarrow X\mathbf{P}\varphi$. Indeed, let $s = (i, w) \in W$ be a state. Suppose that $s \models \mathbf{P}(X\varphi)$. Then there is a state $s' = (i', w')$ such that $sR_d s'$ and $s' \models X\varphi$. So $(i+1, w') \models \varphi$. And thus $(i+1, w) \models \mathbf{P}\varphi$. So we can deduce $s \models X\mathbf{P}\varphi$. In the same way, we can show that $\models X\mathbf{P}\varphi \Rightarrow \mathbf{P}X\varphi$. Then,

$$\models \mathbf{P}X\varphi \Leftrightarrow X\mathbf{P}\varphi$$

This property can also be formulated as follows

$$\models \mathbf{O}X\varphi \Leftrightarrow X\mathbf{O}\varphi \tag{1}$$

The above commutativity properties are typical for product logics. The properties reflect the fact the deontic realm is not updated, as we said in the introduction. So, if it is obligatory to go to Paris tomorrow, then tomorrow it will be obligatory to go to Paris immediately, and vice versa. Now the question of the next section is whether or not we can add propagation properties to the temporal deontic product while leaving the product intact: is there a non-empty (and interesting) subset of the product frames which validate the propagation property? Intuitively, this should not be the case: propagation means that obligations are ‘created’ for future moments. The *trigger* for this creation is the circumstance that the obligations are not met presently.

3 Adding a propagation property

We want to consider a propagation property as general as possible. For instance we want to capture the obligation with deadline, or the obligation to meet something eventually (without deadline). The obligation to satisfy φ now, or ψ next seems to be the most general kind of obligation for which we want to study the propagation. Indeed, the obligation with deadline $\mathbf{O}(F_{\leq k}(\varphi))$ can be re-written $\mathbf{O}(\varphi \vee XF_{\leq k-1}(\varphi))$, and the obligation to satisfy φ eventually $\mathbf{O}(F\varphi)$ can be re-written $\mathbf{O}(\varphi \vee XF(\varphi))$.

As a first attempt for formalizing a propagation property to be added to the product logic, we consider:

$$\mathbf{O}(\varphi \vee X\psi) \wedge \neg\varphi \Rightarrow X\mathbf{O}(\psi) \tag{2}$$

If it is obligatory to meet φ now, or ψ next, and φ is not satisfied now, then it will be obligatory next to meet ψ .

As argued in the introduction, we would not want that from the propagation property and the properties of the temporal deontic logic it follows that $\mathbf{O}\varphi \wedge \neg\varphi \Rightarrow X\mathbf{O}(\psi)$. Yet this property does follow from 2 in combination with (a temporal variant of) weakening of obligations: $\mathbf{O}\varphi \Rightarrow \mathbf{O}(\varphi \vee X\psi)$. To solve this problem, we re-formalize the propagation property, in order to prevent

that in combination with temporal weakening it can be used to derive this unwanted property. To achieve this, in the propagation property, we exclude that $\mathbf{O}(\varphi \vee X\psi)$ holds *only* because $\mathbf{O}(\varphi)$ holds³, and we thus arrive at the following property instead of (2):

$$\mathbf{O}(\varphi \vee X\psi) \wedge \neg \mathbf{O}\varphi \wedge \neg\varphi \Rightarrow X\mathbf{O}(\psi) \quad (3)$$

Similarly, we may explicitly exclude that $\mathbf{O}(\varphi \vee X\psi)$ holds *only* because $\mathbf{O}(X\psi)$ holds. So we may formulate the propagation formula as follows:

$$\mathbf{O}(\varphi \vee X\psi) \wedge \neg \mathbf{O}\varphi \wedge \neg \mathbf{O}X\psi \wedge \neg\varphi \Rightarrow X\mathbf{O}(\psi) \quad (4)$$

However, with the product property $X\mathbf{O}\varphi \leftrightarrow \mathbf{O}X\varphi$, property 4 is equivalent with $\mathbf{O}(\varphi \vee X\psi) \wedge \neg \mathbf{O}\varphi \wedge \neg \mathbf{O}X\psi \wedge \neg\varphi \Rightarrow \mathbf{O}X\psi$. This, in turn, is logically equivalent with $\mathbf{O}(\varphi \vee X\psi) \wedge \neg \mathbf{O}\varphi \wedge \neg\varphi \Rightarrow \mathbf{O}X\psi$, which is exactly the same property as 3. So, in the product setting, properties 3 and 4 are equivalent.

But, now we have to conclude that the propagation property is not compatible with a genuine product: we can consistently add property 3 to the product logic, but we will never have a case where $X\mathbf{O}(\psi)$ is really a *consequence* of $\mathbf{O}(\varphi \vee X\psi) \wedge \neg \mathbf{O}\varphi \wedge \neg\varphi$ being true. In fact, a product model satisfies property 4 only if it does not satisfy the hypothesis $\mathbf{O}(\varphi \vee X\psi) \wedge \neg \mathbf{O}(\varphi) \wedge \neg \mathbf{O}(X\psi) \wedge \neg\varphi$. (Indeed, if a product model satisfied the hypothesis $\mathbf{O}(\varphi \vee X\psi) \wedge \neg \mathbf{O}(\varphi) \wedge \neg \mathbf{O}(X\psi) \wedge \neg\varphi$, in some state s , for some φ and ψ , then we could deduce $\neg X\mathbf{O}(\psi)$ in s .) This corresponds to a product model where all the ideal states of a given state have the same valuation, which is clearly not interesting to work with.

The only way to preserve the propagation property is then to drop the ‘no learning’ property $X\mathbf{O}\varphi \Rightarrow \mathbf{O}X\varphi$. So, we will no longer have a genuine product. But this is in accordance with intuitions. Obligations may now be transferred to future states. The above discussion shows that this is incompatible with a product; we have to allow some dynamics in the deontic dimensions because obligations may be inherited from earlier states. We do however preserve the ‘perfect recall’ property $\mathbf{O}X\varphi \Rightarrow X\mathbf{O}\varphi$ that expresses that no obligations are ‘forgotten’ over time. Note that this last property is sufficient to ensure that properties 3 and 4 are equivalent. So, in the rest of the paper, we will study property 3, which is shorter.

3.1 Restricting the ideal states

Our goal in this section will be to define a semantics that satisfies the propagation property and the perfect recall property. To account for propagation, in the semantics we have to introduce a stronger interaction between what happens and what is obligatory, i.e., between what is true in the current world and what is true in the (next) ideal worlds. If we want to satisfy the perfect recall property, the set of ideal worlds in the next state is a *subset* of the set of the ideal worlds in the current state. The principle of propagation then should point us to what

³ Another strategy might be to attack the temporal weakening property directly.

subset to take. Our idea is that for ideal worlds at a next moment we should only take into account the worlds that share the same past as the current world until the present moment. This reflects the idea that what is deontically ideal at the next moment depends on what actually occurs presently.

Below, we first define the predicate $SamePast(s, s')$ which says that the states s and s' of a temporal deontic model share the same past :

$$SamePast((i, w), (i', w')) \stackrel{def}{=} i = i' \wedge \forall j < i \ V(j, w) = V(j, w')$$

When interpreting an obligation in a state s , we only consider the states s' which satisfy $sR_d s'$ and $SamePast(s, s')$.

Definition 8 (Semantics of the obligation (2)). *Given a product model $(S, <_t, R_d)$, a state s , and a formula φ , we now consider the following semantics for obligation:*

$$s \models \mathbf{O}\varphi \text{ iff } \forall s' \in S \text{ if } (SamePast(s', s) \text{ and } sR_d s') \text{ then } s' \models \varphi$$

With this new semantics, the deontic realm is described by fewer and fewer worlds (which means that more and more formulas are obligatory) when time passes. This is conform the fact that we keep $\mathbf{O}(X\varphi) \Rightarrow X\mathbf{O}(\varphi)$, and avoid $X\mathbf{O}(\varphi) \Rightarrow \mathbf{O}(X\varphi)$; no obligations are forgotten, but some obligations may appear (in particular when they are propagated from a more general obligation in the previous state).

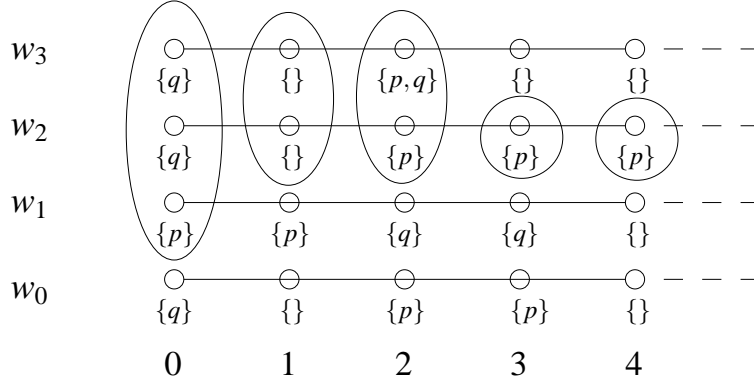


Fig. 2. Semantics of obligation

Let us illustrate, by the way of an example, how an obligation may propagate. Consider the product model illustrated in Figure 2, where, in state $(0, w_0)$, histories w_1 , w_2 , and w_3 , are ideal. Then, we have for instance $0, w_0 \models \mathbf{O}(p \vee XXp) \wedge \neg p$. Since w_0 does not satisfy p at instant 0, the history w_1 which satisfies p at instant 0 is not ideal anymore at the next instant.

So, only w_2 and w_3 (which satisfy XXp at instant 0) remain ideal at instant 1. Thus, the propagation applies, and we have $0, w_0 \models X\mathbf{O}(Xp)$. Let us now state the propagation property and propose a proof in the general case where φ is a propositional formula and ψ can be any formula.

Property 1 (Propagation property). Let M be a temporal deontic product model. Then it satisfies the propagation property for the obligation operator of definition 8:

$$M \models \mathbf{O}(\varphi \vee X\psi) \wedge \neg\mathbf{O}(\varphi) \wedge \neg\varphi \Rightarrow X\mathbf{O}(\psi)$$

for φ propositional formula, and ψ any formula.

Proof. Let $M = ((S, <_t, R_d), V)$ be a temporal deontic model, and $s = (i, w) \in S$ a state such that $s \models \mathbf{O}(\varphi \vee X\psi) \wedge \neg\mathbf{O}(\varphi) \wedge \neg\varphi$. Every s' such that $\text{SamePast}(s, s')$ and $sR_d s'$ satisfies $\varphi \vee X\psi$, and it is not the case that every such s' satisfies φ . If some of these states $s' = (i, w')$ have the same valuation as s , then they satisfy $\neg\varphi$ (since $s \models \neg\varphi$), and $\varphi \vee X\psi$. So, they satisfy $X\psi$. Thus, for every state $(i+1, w')$ which is ideal from $(i+1, w)$ and has the same past as $(i+1, w)$ satisfies ψ , i.e., $(i+1, w) \models \mathbf{O}\psi$. Otherwise (if none of the states s' have the same valuation as s), there is no ideal state having the same past as $(i+1, w)$. So, every formula is obligatory in $(i+1, w)$. In particular, $(i+1, w) \models \mathbf{O}\psi$. \square

The fact that φ is a propositional formula in accordance with intuition. Indeed, the propagation property expresses that what will be obligatory at the next step may depend on what happens now (φ not being true). So it is natural to consider that φ is a formula which only concerns the present moment, i.e. a propositional formula. Otherwise, if φ contained future operators, what will be obligatory at the next step would depend on something which has not happened yet.

So we have that some of the obligations that may appear at a next state are due to the propagation property. In fact, the following property claims that the propagation property completely characterizes the new obligations that appear.

Property 2 (Characterization of new obligations). For any formula ψ , if in a state s both the formulas $X\mathbf{O}(\psi)$ and $\neg\mathbf{O}(X\psi)$ hold, then there exists a propositional formula φ such that

$$s \models \mathbf{O}(\varphi \vee X\psi) \wedge \neg\varphi$$

So, if there will be next an obligation to satisfy ψ and if this obligation is new (i.e., now, there is no obligation to meet ψ next), then it is due to a current obligation to satisfy $\varphi \vee X\psi$ where φ is propositional and not fulfilled.

Proof. Let ψ a formula and s a state such that $s \models X\mathbf{O}(\psi) \wedge \neg\mathbf{O}(X\psi)$. Let E the set of the ideal states of s which do not satisfy $X\psi$:

$$E \stackrel{def}{=} \{s' \in S / sR_d s' \text{ and } s' \models \neg X\psi\}$$

We now define the set $V(E)$ of all the valuations of states in E . This set is finite (even if E is infinite) because it belongs to 2^{2^P} . $V(E) \stackrel{def}{=} \{V(s) / s \in E\}$. Then we define the propositional formula

$$\varphi \stackrel{def}{=} \bigvee_{v \in V(E)} \left(\bigwedge_{p \in v} p \wedge \bigwedge_{p \notin v} \neg p \right)$$

Then every ideal state of s either satisfies $X\psi$ or is in E and satisfies φ . So $s \models \mathbf{O}(\varphi \vee X\psi)$.

Moreover, since $s \models X\mathbf{O}(\psi)$, the states in E - which do not satisfy $X\psi$ - become not ideal at the next step. So they do not share the same atomic propositions with s . Thus $s \models \neg\varphi$. \square

Unfortunately, not everything is fine. In particular, the deontically ideal worlds may shrink to the empty set when time passes, as we saw in the proof of property 1. This conflicts with our desire to stay in accordance with *SDL* where obligations are always consistent: $\neg\mathbf{O}\perp$. Another formulation is the D axiom: $\mathbf{O}\varphi \Rightarrow \mathbf{P}\varphi$. Then, if from a state s , there is no ideal state with the same past, these properties cannot be satisfied, and every formula is obligatory in s , including $s \models \mathbf{O}\perp$. In particular this occurs if there is a violation of a proposition p in a state $s = (i, w)$, i.e., if $s \models p \wedge \mathbf{O}(\neg p)$. In this case no ideal state is associated with $(i + 1, w)$.

Property 3. With the new semantics of the obligation, the D axiom is not valid:

$$\not\models \neg\mathbf{O}(\perp) \quad \text{and} \quad \not\models \mathbf{O}\varphi \Rightarrow \mathbf{P}\varphi \quad \text{for any formula } \varphi$$

Proof. The problem is due to the fact that there may be states without any ideal states with the same past. Indeed, let $((S, <_d, R_d), V)$ be a product model, and $s \in S$ a state such that no other state has the same past as s . (It is easy to build such a model.) Then, according to definition 8, $s \models \perp$, which invalidates the D axiom. \square

As a solution to this problem, we might consider to add a constraint on the models expressing that from every state there exists an ideal state with the same past.

Definition 9 (Ideal existence constraint on models). Let $M = ((S, <_d, R_d), V)$ be a temporal deontic product model. We say that M satisfies the ideal existence constraint if

$$\forall s \in S \quad \exists s' \in S \quad \text{such that} \quad sR_d s' \text{ and } \text{SamePast}(s, s')$$

This constraint now guarantees validity of the D-axiom.

Property 4 (D axiom). Let M be a temporal deontic product model that satisfies the ideal existence constraint. Then

$$M \models \neg\mathbf{O}\perp \quad \text{or equivalently} \quad M \models \mathbf{O}\varphi \Rightarrow \mathbf{P}\varphi$$

for any formula φ .

Proof. Let $M = ((S, <_d, R_d), V)$ be a temporal deontic product model that satisfies the ideal existence constraint, and $s \in S$ a state. From definition 9 we have that there exists an ideal state s' with the same past as s . So, from the definition of obligation (definition 8), $s \models \neg \mathbf{O}(\perp)$. \square

However, again we have to face a problem: the ideal existence constraint interacts with the identical past criterion in an undesirable way. In particular, if there is a simple obligation $\mathbf{O}p$ that is violated in the current world, the identical past criterion demands that all ideal next worlds satisfy $\neg p$ in their previous state. However, of these there can be none, since otherwise we would not have had $\mathbf{O}p$. But then, under the identical past criterion, there can be no ideal next worlds as soon as there is a simple obligation $\mathbf{O}p$ that is violated presently. But then this directly conflicts with the ideal existence constraint. So, if in our logic, we impose both properties, we actually get that obligations can never be violated.

Property 5 (No violation). Let M be a model satisfying the ideal existence constraint and φ a formula. Then $M \models \neg(\varphi \wedge \mathbf{O}(\neg\varphi))$.

The conclusion has to be that we still have to refine the semantics: the violation of obligations should be possible, without losing the interaction between what happens and the deontic realm.

3.2 Levels of deontic ideality

Another way to view the problem of the previous section is to say that the semantics should be able to deal with ‘contrary to duty’ (CTD) situations. In states where there is a violation, something happens that is contrary to what is obligatory for that state. It should not be the case that such situations cause the deontic realm to collapse. So when there is a violation, it should still be possible to point out what is obligatory and what not, despite of the violation in the present state.

We look for a solution to the problem by switching to *levels* of ideality. Rather than an accessibility relation which gives the ideal states, we consider a preference relation \leq_d , where $s \leq_d s'$ means that the state s' is “better” than the state s . This allows us to have several “levels of ideality”. The ideal states will be the best states among those which share the same past as the current state. The idea is now that if a state (i, w) violates an obligation of a propositional formula then the ideal states of $(i + 1, w)$ are states which were not ideal for (i, w) : the deontic realm thus switches to a *lower level* of ideality. This contrasts with the setting of the previous section, where in this case there would be no ideal states left.

Definition 10 (Temporal deontic frame and model). A temporal deontic frame

$(S, <_t, \leq_d)$ is defined as the product $(\mathbb{N}, <) \times (W, \leq_{pref})$ of a temporal frame $(\mathbb{N}, <)$ and a deontic frame (W, \leq_{pref}) , where \leq_{pref} , considered as a preference relation, is a total quasi-order (total and transitive relation) on W .

A temporal deontic model is defined as a product model based on a temporal deontic frame.

For the temporal and boolean operators the satisfaction relation is defined as above. For the obligation operator it is defined as follows.

Definition 11 (Semantics of the obligation (3)). Given a temporal deontic model

$((S, <_t, \leq_d), V)$, and a state $s \in S$, φ is obligatory if there is a state with the same past as s such that every “better” state with the same past satisfies φ .

$s \models \mathbf{O}\varphi$ iff $\exists s' \in S$ such that $\text{SamePast}(s, s')$ and
and $\forall s'' \in S$ if $(\text{SamePast}(s, s'') \wedge s' \leq_d s'')$ then $s'' \models \varphi$

Remark 1. If every set of states has at least one maximum element for the quasi-order \leq_d (i.e., the relation \geq_d , defined by $s \geq_d s'$ iff $s' \leq_d s$, is a well-quasi-order), then we can define the set of the best states among those having the same past:

$\text{BestSamePast}(s) \stackrel{def}{=} \{s' \in S / \text{SamePast}(s, s') \text{ and}$
 $\forall s'' \in S \text{ if } \text{SamePast}(s, s'') \text{ then } s'' \leq_d s'\}$

And the semantic definition of $\mathbf{O}(\varphi)$ becomes more simple:
 $s \models \mathbf{O}\varphi$ iff $\forall s' \in \text{BestSamePast}(s) \quad s' \models \varphi$

For the newly defined models (definition 10) with levels of ideality, there is no need for a constraint to guarantee the validity of the D axiom. Indeed, the existence of a state with the same past is guaranteed by the current state itself. Thus the existence of ideal states is also guaranteed. (Recall that the ideal states are the best among those which share the same past as the current state.) So the D axiom is valid and violations can be satisfied.

However, there still is a phenomenon that has to be considered more closely. When an obligation of a proposition p is violated in a state (i, w) , then the ideal states at the step $i+1$ are completely disjoint from the ideal states at the step i . This is easy to see: if $(i, w) \models \neg p \wedge \mathbf{O}p$, then all the ideal states of (i, w) satisfy p . On the other hand, the ideal states of $(i+1, w)$ have the same past as $(i+1, w)$, and thus they are states $(i+1, w')$ such that (i, w') does not satisfy p . So none of the ideal worlds of (i, w) are ideal for $(i+1, w)$ and vice versa. The problem is now that in such states, the propagation property is not guaranteed anymore because of the change to a completely different set of lower level ideal worlds.

Actually, the condition that makes the set of ideal worlds change between (i, w) and $(i+1, w)$ is a little more general than suggested by the example with the violation of an atomic proposition. More in general, the condition concerns the violation of an obligation for any propositional formula which can be seen as an immediate obligation, that is, any propositional formula concerning the present moment. So, if such an obligation is violated, the current ideal worlds

will not be considered as ideal in the future. The current norms become obsolete, and we switch to the norms of a lower level. If not, we have a strong link between what is obligatory now and next, and the propagation property holds.

To characterize these two kinds of states, we define the condition $IdealSameProp(s)$ on a state s which expresses that for every state with the same past as s , there is better state which still has the same past at the next step. This condition ensures that some of the current ideal worlds are still ideal at the next step.

Given a temporal deontic model $((S, <_t, \leq_d), V)$ and a state $s \in S$,

$$IdealSameProp(s) \stackrel{def}{=} \forall s' \in S \text{ if } SamePast(s, s') \text{ then} \\ \exists s'' \in S \text{ such that } SamePast(s, s'') \text{ and } V(s) = V(s'') \text{ and } s' \leq_d s''$$

Remark 2. If (W, \geq_{pref}) is a well-quasi-order, then $IdealSameProp(s)$ is defined in a more simple way:

$$IdealSameProp(s) \stackrel{def}{=} \exists s' \in BestSamePast(s) \text{ such that } V(s) = V(s')$$

Property 6. Given a temporal deontic model $((S, <_t, \leq_d), V)$ and a state $s \in S$, the condition $IdealSameProp(s)$ holds iff there is no violation of a propositional formula in s , that is, iff for any propositional formula φ , $s \models \neg(\mathbf{O}(\varphi) \wedge \neg\varphi)$.

Proof. We first prove that if $IdealSameProp(s)$ does not hold, then there is some propositional formula φ such that $s \models \mathbf{O}(\varphi) \wedge \neg\varphi$. We then prove the other direction.

' \Leftarrow ': Suppose that $IdealSameProp(s)$ does not hold, i.e.,

$$\exists s' \in S \text{ such that } SamePast(s, s') \text{ and} \\ \forall s'' \geq_d s' \text{ if } SamePast(s, s'') \text{ then } V(s) \neq V(s'')$$

Then, we consider such a state s' and define the set $VAL(s')$ of all the valuations of the states which are at least as good as s' and share the same past.

$$VAL(s') \stackrel{def}{=} \{V(s'') / s' \leq_d s'' \text{ and } SamePast(s', s'')\}$$

$VAL(s')$ is finite since it is included in the set 2^{2^P} . Let us consider the propositional formula φ defined as follows:

$$\varphi \stackrel{def}{=} \bigvee_{v \in VAL(s')} \left(\bigwedge_{p \in v} p \wedge \bigwedge_{p \notin v} \neg p \right)$$

Since every such state s'' has a valuation which is distinct from the valuation of s , then $s \models \neg\varphi$. Besides, from the definition of obligation we have that $s \models \mathbf{O}(\varphi)$. Thus, $s \models \mathbf{O}(\varphi) \wedge \neg\varphi$.

' \Rightarrow ': Let us suppose now that there exists some propositional formula φ such that $s \models \mathbf{O}(\varphi) \wedge \neg\varphi$. Then,

$$\exists s' \in S \text{ such that } SamePast(s, s') \text{ and} \\ \forall s'' \geq_d s' \text{ if } SamePast(s, s'') \text{ then } s'' \models \varphi$$

Every such s'' has a valuation which differs from the valuation of s , i.e., $V(s'') \neq V(s)$, since $s \models \neg\varphi$ and $s'' \models \varphi$. Therefore, $IdealSameProp(s)$ does not hold. \square

In a state s that satisfies $IdealSameProp(s)$, the deontic realm that will be considered next is a subset of the current deontic realm. So we still have, as in section 3.1, that no obligations are forgotten, but some may appear. If $IdealSameProp(s)$, then $s \models \mathbf{O}(X\varphi) \Rightarrow X\mathbf{O}(\varphi)$, but $X\mathbf{O}(\varphi) \Rightarrow \mathbf{O}(X\varphi)$ does not hold necessarily.

Property 7 (Propagation). A state which does not satisfy any violation of a propositional formula satisfies the propagation property.

If $IdealSameProp(s)$ then

$$s \models \mathbf{O}(\varphi \vee X\psi) \wedge \neg\mathbf{O}\varphi \wedge \neg\varphi \Rightarrow X\mathbf{O}\psi$$

for φ propositional formula, and ψ any formula.

Proof. The proof is similar to the proof of property 1 in section 3.1, except that we have not the case where every formula is obligatory in the temporal successor of s . \square

We still have, as in section 3.1, property 2, a more precise characterization.

Property 8 (Characterization of new obligations). For any formula ψ , if in a state s , which satisfies $IdealSameProp(s)$, both the formulas $X\mathbf{O}(\psi)$ and $\neg\mathbf{O}(X\psi)$ hold, then there exists a propositional formula φ such that

$$s \models \mathbf{O}(\varphi \vee X\psi) \wedge \neg\mathbf{O}(\varphi) \wedge \neg\varphi$$

When an obligation appears, it is necessary due to the propagation of some more general obligation in the previous state. So the propagation property completely characterizes the new obligations that appear.

Proof. The proof follows the same idea as the proof of property 2 in section 3.1. \square

As said in the introduction of section 3, as a consequence of the general propagation property, if a state s satisfies $IdealSameProp(s)$, then it satisfies the following property of propagation for an obligation with deadline, since $F_{\leq k}\varphi \Leftrightarrow \varphi \vee XF_{\leq k-1}\varphi$, for $k > 0$:

$$s \models \mathbf{O}(F_{\leq k}\varphi) \wedge \neg\mathbf{O}(\varphi) \wedge \neg\varphi \Rightarrow X\mathbf{O}(F_{\leq k-1}\varphi)$$

for any deadline $k > 0$, and φ propositional formula. This property expresses that if it is obligatory to satisfy φ before a deadline k (and it is not obligatory to satisfy it now) then, if φ is not true now, the obligation is propagated.

In a state which does not satisfy $IdealSameProp$, that is, a state which violates an obligation of some propositional formula, the deontic realm of the next state switches to a lower level. We consider that when a state violates the present rules, then they become obsolete. In such a state, $\mathbf{O}(X\varphi) \Rightarrow X\mathbf{O}(\varphi)$ is not guaranteed, and neither is any link between what is satisfied in the current state, and what is obligatory next.

4 Branching time structures

Many proposals to combine temporal and deontic concepts use a branching time structure, where the ideal alternatives are subsets of the possible future worlds [13, 2, 1, 9]. This has several advantages. In particular, the principle “must implies can” is guaranteed. And our “identical past” criterion is automatically satisfied by branching time structures. However, some of the intuitive properties we have discussed in this paper do not necessarily hold in a branching time setting, and may be hard to implement. For instance, we consider that an atomic proposition p can be both true in the current state and false in some ideal state. In pure branching time approaches [1, 2] this is usually not possible. Branching time approaches thus have problems modelling immediate obligations. For example, $O(p) \wedge \neg p$ is not satisfiable in such logics.

An exception has to be made for deontic STIT formalisms [13], where moments are partitioned into choices, and where a proposition can be obligatory but not true at the same moment. However, it is quite unclear how to implement a principle of propagation in the conceptually rich setting of STIT models. Actually, we can represent our model in a tree like view as in STIT theory by putting together the histories sharing the same past. In a STIT framework, the moment/history pairs (m, h) would naturally correspond to our states (i, w) , and the histories h to our deontic worlds w . To have a branching time view, the states (i, w) that share the same past would be represented by a unique moment m . So two deontic worlds w and w' that share the same past until i would naturally be represented by histories that go through the moment m , where m represents both states (i, w) and (i, w') . But in the STIT framework, there is a valuation of atomic propositions for every moment/history pair. So two histories that go through the same moment m , also go through the same moments w' for every $w' < w$, while the valuation they have may differ. Our model can thus be easily translated to a STIT setting under the restriction that two histories sharing the same past moments also share the same valuation on these moments. More formally, for a moment m , and two histories h_1 and h_2 ,

$$(m \in h_1 \text{ and } m \in h_2) \text{ if and only if } (\forall m' < m \forall p \in P \ m', h_1 \models p \text{ iff } m', h_2 \models p)$$

The strict operator $<$ allows us to have different valuations for the same moment m depending on the history we consider, if the histories split after this moment. Our framework is close to a STIT framework with this restriction. To characterize a state (called moment in a STIT model) that satisfies *IdealSameProp* corresponds in such a STIT model to a moment/history pair (m, h) which has the same valuation as some pair (m, h') , where h' is ideal. Such pairs, which would need a formal characterization, would satisfy the propagation properties we have studied.

5 Conclusion

In this paper, we have studied properties concerning the propagation of obligations for the future that have not been fulfilled yet. Because we do not want

to consider explicit updates of the deontic realm, we started with a product of temporal and deontic logics, and concluded that to account for propagation properties we had to drop the property of ‘no learning’. Going through some “hybrid” semantics, we settled on a semantics with levels of ideality exposing interesting similarities with branching time STIT settings. We characterized the states where obligations are propagated: these are the states in which no immediate obligations are violated. In such states, obligations are not forgotten, and the new obligations are due to the propagation of some past obligations which are not fulfilled yet. On the other hand, in the states which violate immediate obligations, the ideal states switch to a lower level of ideality, and the obligations of the next state do not depend on what is true now.

The propagation of obligations has been studied in the restricted case of dedicated operators for obligations with a deadline, for instance in [5–7, 9]. But, to our knowledge, the more general propagation property we have focused on is new.

We only considered a linear time setting with one agent. We cannot express “must implies can”, nor “it is obligatory to make something possible”, nor “it is obligatory for an agent to do something”. Therefore it would be interesting to further develop the link with STIT models mentioned in section 4. We plan to formalize it and thus study the propagation of obligations in a framework which allows branching time and multi-agents reasoning.

Another issue is the decidability of our logic. A clue is that the genuine product $LTL \times SDL$ is decidable (see [10] for the decidability of $LTL \times K$), but the decision problem is non elementary.

References

1. L. Åqvist. Combinations of tense and deontic logic. *Journal of Applied Logic*, 3:421–460, 2005.
2. P. Bailhache. Canonical models for temporal deontic logic. *Logique et Analyse*, pages 3–21, 1995.
3. M. Bratman. *Intention, plans, and practical reason*. Harvard University Press, Cambridge Massachussetts, 1987.
4. M. Bratman. Planning and the stability of intentions. *Minds and Machines*, 2:1–16, 1992.
5. J. Broersen. Strategic deontic temporal logic as a reduction to ATL, with an application to Chisholm’s scenario. In L. Goble and J.-J. C. Meyer, editors, *Deontic Logic and Artificial Normative Systems, 8th International Workshop on Deontic Logic in Computer Science (DEON’06)*, volume 4048 of *Lecture Notes in Computer Science*, pages 53–68. Springer, 2006.
6. J. Brunel, J.-P. Bodeveix, and M. Filali. A state/event temporal deontic logic. In L. Goble and J.-J. C. Meyer, editors, *Deontic Logic and Artificial Normative Systems, 8th International Workshop on Deontic Logic in Computer Science (DEON’06)*, volume 4048 of *Lecture Notes in Computer Science*, pages 85–100. Springer, 2006.
7. F. Cuppens, N. Cuppens-Bouahia, and T. Sans. Nomad: a security model with non atomic actions and deadlines. In *Proceedings of the 18th IEEE Computer Security Foundations Workshop*, June 2005.

8. M. Dastani, B. van Riemsdijk, F. Dignum, and J. Meyer. A programming language for cognitive agents: Goal directed 3APL. In *Proceedings of the First Workshop on Programming Multiagent Systems: Languages, frameworks, techniques, and tools (ProMAS03)*. ACM Press, 2003.
9. R. Demolombe, P. Bretier, and V. Louis. Formalisation de l'obligation de faire avec délais. In *Proc. Journées Francophones sur la Modélisation Formelle de l'Interaction*, Caen, 2005.
10. D. Gabbay, A. Kurucz, F. Wolter, and M. Zakharyachev. *Many-Dimensional Modal Logics: Theory and Applications*. Elsevier, 2003.
11. R. Hilpinen. *New studies in deontic logic*. Reidel, 1981.
12. K. V. Hindriks, F. S. de Boer, W. van der Hoek, and J.-J. C. Meyer. Agent programming in 3apl. *International Journal of Autonomous Agents and Multi-Agent Systems*, 2(4):357–401, 1999.
13. J. Horty. *Agency and Deontic Logic*. Oxford University Press, 2001.
14. A. Pnueli. The temporal semantics of concurrent programs. *Theoretical Computer Science*, 13:45–60, 1981.
15. A. Rao and M. Georgeff. Modeling rational agents within a BDI-architecture. In J. Allen, R. Fikes, and E. Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91)*, pages 473–484. Morgan Kaufmann Publishers, 1991.
16. A. S. Rao. AgentSpeak(L): BDI agents speak out in a logical computable language. In R. van Hoe, editor, *Seventh European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, Eindhoven, The Netherlands, 1996.
17. A. Ross. Imperatives and logic. *Theoria*, 7:53–71, 1941.
18. G. v. Wright. Deontic logic. *Mind*, 60:1–15, 1951.