Designing a Deontic Logic of Deadlines

Jan Broersen, Frank Dignum, Virginia Dignum, John-Jules Meyer

Institute of Information and Computing Sciences
Utrecht University,
P.O.Box 80.089, 3508 TB Utrecht, The Netherlands
broersen/dignum/virginia/jj@cs.uu.nl

Abstract. This paper studies the logic of a dyadic modal operator for being obliged to meet a condition ρ before a condition δ becomes true. Starting from basic intuitions we arrive at a simple semantics for deadline obligations in terms of branching time models. We show that this notion of deadline obligation can be characterized in the branching time logic CTL. The defined operator obeys intuitive logic properties, like monotony w.r.t. ρ and anti-monotony w.r.t. δ, and avoids some counter-intuitive properties like agglomeration w.r.t ρ and 'weak agglomeration' w.r.t. δ. However, obligations of this type are implied by the actual achievement of ρ before the deadline. We argue that this problem is caused by the fact that we model the obligation only from the point of view of its violation conditions. We show that the property might be eliminated by considering success conditions also.

1 Introduction

This paper studies the logic of a dyadic modal operator, denoted $O(\rho \leq \delta)$, for being obliged to meet a condition ρ before a condition δ becomes true. To satisfy the obligation, it suffices to satisfy the condition ρ only once, at a time of ones choosing, as long as it is *before* (or, ultimately, *at*) the point where the condition δ occurs. We refer to the operator $O(\rho \leq \delta)$ as a 'deontic deadline' operator. We do not claim that all deadlines have a deontic aspect. For instance, in the field of 'scheduling' [1], deadlines are hard constraints that have to be satisfied under all circumstances. However, in more realistic situations, where agents may choose to violate deadlines imposed on them by other agents, it is much harder to deny the deontic aspect[1].

Conceptually, deontic deadlines are interactions between two dimensions: a deontic (normative) dimension and a temporal dimension. So, to study deadlines, it makes sense to take a standard temporal logic, say CTL [2–4], and a standard deontic logic, say SDL [5], and combine the two in one system. This type of approach is for instance taken in [6]. The problem then is how to account for the interactions. Conceptually, we have to keep in mind that in the combined

[1] If deadlines are not due to commitments towards other agents, but the result of personal decisions based on personal desires, it is more adequate to talk about 'deadline intentions'.

system we can express that the normative content of the deontic operators can be temporal (e.g, being obliged to be polite *always*), but also that obligations can have some (non-)dynamical behavior over time (e.g., *always* being obliged to be polite). It is easy to mix up these essentially different propositions. The same kind of confusion threatens the study of deadlines. Is a deadline (1) an obligation *at* a certain point in time to achieve something *before* another point in time, or (2) is a deadline simply an obligation that persists in time until a deadline is reached, or (3) is it both? In natural language it is actually quite hard to be precise about this distinction. Therefore, for now, we rely on in informal understanding of the branching time temporal logic CTL, and the standard deontic logic SDL, to discuss the distinction using formulas. In CTL, the symbols E and A denote an existential and a universal quantifier respectively, ranging over possible future courses (branches) of time. Within the scope of these quantifiers, CTL uses the linear time operators (LTL [3]) $\varphi U \psi$ (strong Until, i.e., ψ will occur, and φ holds up until then), $\varphi U_w \psi$ (weak Until, i.e., if ψ will occur, then φ holds up until then) to talk about individual future courses of time (from now on simply called 'possible futures'). From SDL we use the operator O, for obligation.

In a language that combines CTL and SDL we can talk about both the temporal and deontic dimensions independently. For instance, we can talk about a certain obligation being preserved over time: $A(O\rho U_w \delta)$, which says 'the obligation to achieve ρ persists until δ, and if δ does not occur, it persists forever'. Or we can talk about the obligation that a certain condition ρ has to be achieved before a condition δ occurs: $O(\neg E(\neg\rho U \delta))$. This says 'it is obliged that on no possible future ρ is avoided until δ becomes true' (alternatively we may read this as 'it is forbidden that on some possible future ρ is avoided up until δ becomes true').

However, in this paper we do not use a language where we can talk about both dimensions independently. We see the deontic dimension as 'embedded' in the temporal dimension[2], the only difference being that it is considered exclusively with certain violation [7] and ideality [8] constants that hold or do not hold at certain points in time.

The advantage of this approach is that we study deadlines entirely in CTL supplemented with violation[3] and ideality constants. The disadvantage is that the language is not expressive enough to talk about the deontic and temporal dimensions independently. In particular, we cannot talk about the dynamics of obligations. So, a background assumption of our study will be that agents do not get new obligations, or are explicitly discharged of some of their obligations, when time evolves. In yet other words: there are no explicit 'deontic updates'. This implies that if in a next state the deadline δ or the achievement ρ is not realized, the deadline obligation persists. In sections 4 and 5, we present formulas

[2] Technically this corresponds with the deontic accessibility relation being enclosed by the temporal accessibility relation.

[3] The idea of expressing the semantics of deontic deadlines by characterizing violation conditions in CTL supplemented with violation constants [7], was first explored in [9].

that correspond to this background assumption for two different version of the deadline operator.

We model the deadlines themselves as propositions. This seems a reasonable choice given that we do not want to model a deadline in a logic of explicit time (real time). Our view is more abstract, and a deadline is simply a condition δ true at some point in time. A consequence of this abstract view is that we have to deal with the possibility that δ actually never occurs. Note that for a theory of deadlines that uses an explicit notion of time, this would never be a problem. In particular, the point 'two hours from now' will always occur, while meeting a condition 'δ' may be impossible or extremely unlikely. However, our abstract view contributes to the relevance of the present research for other logical systems. For instance, Rao and Georgeff's commitment strategies [10] are actually a sort of deadlines: an agent commits to an intention until the action is performed or believed not to be feasible any longer.

The choice in this paper for the temporal logic CTL is a pragmatic one. We believe the theory applies equally well, and maybe better, to linear time temporal logic (LTL [3]). However, CTL has nice properties (P-complete complexity of the model checking problem for CTL, versus PSPACE-complete complexity for LTL [11]), and is popular in agent theory [12].

2 Preliminaries: CTL

A well-formed formula φ of the temporal language $\mathcal{L}_{\mathrm{CTL}}$ is defined by:

$$\varphi, \psi, \ldots := p \mid \neg\varphi \mid \varphi \wedge \psi \mid E(\varphi U^{ee} \psi) \mid A(\varphi U^{ee} \psi)$$

where φ, ψ represent arbitrary well-formed formulas, and where the p are elements from an infinite set of propositional symbols $\mathcal{P}$. We use the superscript 'ee' to denote that this is the version of the 'until' where φ is not required to hold for the present, nor for the point where ψ, i.e., the present and the point where ϕ are *both* excluded. This gives us the following informal meanings of the until operators:

$E(\varphi U^{ee} \psi)$: there is a future for which eventually, at some point m, the condition ψ will hold, while φ holds from the next moment until the moment before m

$A(\varphi U^{ee} \psi)$: for all futures, eventually, at some point m, the condition ψ will hold, while φ holds from the next moment until the moment before m

We define all other CTL-operators as abbreviations[4]. Although we do not use all of the LTL operators X, F, and G in this paper, we give their abbreviations

[4] Often, the CTL-operators $EG\varphi$ and $E(\psi U \varphi)$ are taken as the basic ones, and other operators are defined in terms of them. The advantage of that approach is that we do not have to use the notion of 'full path', that is crucial for the truth condition of $A(\varphi U^{ee} \psi)$. However, that approach is not applicable here, since we cannot define the

(in combination with the path quantifiers E and A) in terms of the defined operators for the sake of completeness. We also assume the standard propositional abbreviations.

$EX\varphi \equiv_{def} E(\bot U^{ee}\varphi)$	$AX\varphi \equiv_{def} \neg EX\neg\varphi$
$EF\varphi \equiv_{def} \varphi \vee E(\top U^{ee}\varphi)$	$AG\varphi \equiv_{def} \neg EF\neg\varphi$
$AF\varphi \equiv_{def} \varphi \vee A(\top U^{ee}\varphi)$	$EG\varphi \equiv_{def} \neg AF\neg\varphi$
$A(\varphi U^{e}\psi) \equiv_{def} \varphi \wedge A(\varphi U^{ee}\psi)$	$E(\varphi U^{e}\psi) \equiv_{def} \varphi \wedge E(\varphi U^{ee}\psi)$
$A(\varphi U\psi) \equiv_{def} A(\varphi U^{e}(\varphi \wedge \psi))$	$E(\varphi U\psi) \equiv_{def} E(\varphi U^{e}(\varphi \wedge \psi))$
$A(\varphi U_{w}\psi) \equiv_{def} \neg E(\neg\psi U\neg\varphi)$	$E(\varphi U_{w}\psi) \equiv_{def} \neg A(\neg\psi U\neg\varphi)$

The informal meanings of the formulas with a universal path quantifier are as follows (the informal meanings for the versions with an existential path quantifier follow trivially):

$A(\varphi U^{e}\psi)$	: for all futures, eventually, at some point m, the condition ψ will hold, while φ holds from now until the moment before m
$A(\varphi U\psi)$	: for all futures, eventually, at some point the condition ψ will hold, while φ holds from now until then
$A(\varphi U_{w}\psi)$	: for all possible futures, if eventually ψ will hold, then φ holds from now until then, or forever otherwise
$AX\varphi$	: at any next moment φ will hold
$AF\varphi$	: for all futures, eventually φ will hold
$AG\varphi$	: for all possible futures φ holds globally

A CTL model $\mathcal{M} = (S, \mathcal{R}, \pi)$, consists of a non-empty set S of states, an accessibility relation $\mathcal{R}$, and an interpretation function π for propositional atoms. A full path σ in M is a sequence $\sigma = s_0, s_1, s_2, \ldots$ such that for every $i \geq 0$, s_i is an element of S and $s_i\mathcal{R}s_{i+1}$, and if σ is finite with s_n its final situation, then there is no situation s_{n+1} in S such that $s_n\mathcal{R}s_{n+1}$. We say that the full path σ starts at s if and only if $s_0 = s$. We denote the state s_i of a full path $\sigma = s_0, s_1, s_2, \ldots$ in $\mathcal{M}$ by σ_i. Validity $\mathcal{M}, s \models \varphi$, of a CTL-formula φ in a world s of a model $\mathcal{M} = (S, \mathcal{R}, \pi)$ is defined as:

$$
\begin{array}{lll}
\mathcal{M}, s \models p & \Leftrightarrow & s \in \pi(p) \\
\mathcal{M}, s \models \neg\varphi & \Leftrightarrow & \text{not } \mathcal{M}, s \models \varphi \\
\mathcal{M}, s \models \varphi \wedge \psi & \Leftrightarrow & \mathcal{M}, s \models \varphi \text{ and } \mathcal{M}, s \models \psi \\
\mathcal{M}, s \models E(\varphi U^{ee}\psi) & \Leftrightarrow & \exists\sigma \text{ in } \mathcal{M} \text{ with } \sigma_0 = s, \text{ and } \exists n > 0 \text{ such that:} \\
 & & (1)\ \mathcal{M}, \sigma_n \models \psi \text{ and} \\
 & & (2)\ \forall i \text{ with } 0 < i < n \text{ it holds that } \mathcal{M}, \sigma_i \models \varphi \\
\mathcal{M}, s \models A(\varphi U^{ee}\psi) & \Leftrightarrow & \forall\sigma \text{ in } \mathcal{M} \text{ such that } \sigma_0 = s, \text{ it holds that } \exists n > 0 \text{ such that} \\
 & & (1)\ \mathcal{M}, \sigma_n \models \psi \text{ and} \\
 & & (2)\ \forall i \text{ with } 0 < i < n \text{ it holds that } \mathcal{M}, \sigma_i \models \varphi
\end{array}
$$

'exclusive' versions of the operators in terms of them. And, even if we take $EG\varphi$ and $E(\psi U^{ee}\varphi)$ as basic, we can still not define the for our purposes important operator $A(\psi U^{e}\varphi)$ as an abbreviation.

Validity on a CTL model $\mathcal{M}$ is defined as validity in all states of the model. If φ is valid on a CTL model $\mathcal{M}$, we say that $\mathcal{M}$ is a model for φ. General validity of a formula φ is defined as validity on all CTL models. The logic CTL is the set of all general validities of $\mathcal{L}_{\mathrm{CTL}}$ over the class of CTL models.

3 A dyadic deontic deadline operator

We minimally extend the language $\mathcal{L}_{\mathrm{CTL}}$ by extending the set of propositional atoms with a violation constant of the form $Viol$[5]. Furthermore, the formal interpretation of the atom $Viol$ is treated like that of all other atomic propositions. So, we can view the propositional constant $Viol$ also as a special element of $\mathcal{P}$: a 'special purpose' proposition solely used to interpret deontic formulas in a temporal dimension.

Let $\mathcal{M}$ be a CTL model, s a state, and σ a full path starting at s. A straightforward modal semantics for the operator $O^V(\rho \leq \delta)$, where the V is only a label to emphasize that this operator is defined in terms of Violations, is then defined as follows:

$$\begin{aligned} \mathcal{M}, s \models O^V(\rho \leq \delta) \Leftrightarrow\ & \forall \sigma \text{ with } \sigma_0 = s, \forall j : \\ & \text{if} \\ & \mathcal{M}, \sigma_j \models \delta \text{ and } \forall 0 \leq i \leq j : \mathcal{M}, \sigma_i \models \neg\rho \\ & \text{then} \\ & \mathcal{M}, \sigma_j \models Viol \end{aligned}$$

This says: if at some future point the deadline occurs, and until then the result has not yet been achieved, then we have a violation at that point. This semantic definition is equivalent to the following definition as a reduction to CTL:

$$O^V(\rho \leq \delta) \equiv_{def} \neg E(\neg\rho U(\delta \wedge \neg Viol))$$

This formula simply 'negates' the situation that should be excluded when a deontic deadline is in force[6]. In natural language this *negative* situation is: 'δ becomes true at a certain point, the achievement has not been met until then, and there is *no* violation at δ'. Therefore this CTL formula exactly characterizes the truth condition for the above defined deontic deadline operator: the semantic conditions are true is some state if and only if the the CTL formula is true in that state.

[5] For reasoning in a multi-agent context we may provide violation constants of the form $Viol(a)$ where $a \in \mathcal{A}$, and $\mathcal{A}$ an infinite set of agent identifiers.

[6] Alternatively this definition can be given using the weak until: $O^V(\rho \leq \delta) \equiv_{def} A((\neg\delta \vee Viol)U_w\rho)$. But for the version with the strong until it is much easier to see that it corresponds with the semantic truth conditions defined above.

4 Logical Properties

What logical properties of the operator $O^V(\rho \leq \delta)$ does this bring us? We first discus the property that corresponds to out background assumption that there are no deontic updates. It holds that:

$$\models O^V(\rho \leq \delta) \rightarrow A(O^V(\rho \leq \delta) U_w \rho)$$

To see that this holds[7], it is easiest to fall back on the semantics of the operator. The semantics says that on futures (branches of time) where δ occurs at some point t, while until then ρ has not been done once, there is a violation at t. Now, if we follow such a branch for some time-steps in the future, and we do not meet a ρ, then, the deadline conditions do still apply: still it holds that if δ will occur later on, we get a violation if we do not meet a ρ until then. An important observation is that even if we have passed one or more δ-states, the obligation still applies; only if we meet a ρ, the conditions are no longer guaranteed. Thus, the defined notion of deadline persists, even if we have passed a deadline. This might be considered counter-intuitive, since it seems correct to assume that the deadline obligation itself is discharged by passing the deadline. Therefore, in section 5 we show how to define a version that does not have this property. However, we consider the present notion of deadline not as counter-intuitive, but merely as a variant. Persistence of the obligation at the passing of a deadline is not a priori counter-intuitive. An example is the following: you are obliged to repair the roof of your house before it will rain (or otherwise you and your interior get wet). This obligation is only discarded by the act of repairing the roof, and not by the event of raining.

We now turn to other properties of the operator of section 3. First of all we get monotonicity with respect to ρ (other terminology: validity of the operator is closed under weakening of ρ)[8]. Monotonicity says that we have as validities[9]:

$$\models O^V((\rho \wedge \chi) \leq \delta) \rightarrow O^V(\rho \leq \delta)$$
$$\models O^V(\rho \leq \delta) \rightarrow O^V((\rho \vee \chi) \leq \delta)$$

This is in accordance with intuition: if ρ is made logically weaker, it is easier to satisfy. So, if the stronger condition has to be accomplished before δ occurs, then certainly also the weaker condition has to be accomplished before δ occurs.

A property we do not have is agglomeration with respect to ρ, i.e.:

---

[7] Alternatively we may write this as $\models O(\rho \leq \delta) \rightarrow \neg E(\neg\rho U \neg O(\rho \leq \delta))$. But in our opinion, here the version with the weak until is easier to understand.

[8] In section 6 we will see a simple way to prove weakening and strengthening for the defined operators. However, all other verifications are left to the reader.

[9] To express the property we call 'monotonicity' it suffices to give just one of these theorems, because they can be derived from each other using only the rules of uniform substitution and substitution by logical equivalents. However, to check the intuitiveness of monotony, especially for deontic operators, it is wise to consider both these 'appearances' of monotony.

$$\not\models O^V(\rho \leq \delta) \wedge O^V(\chi \leq \delta) \rightarrow O^V((\rho \wedge \chi) \leq \delta)$$

This shows that $O^V(\rho \leq \delta)$, is monotonic with respect to ρ, but is *not* a normal modal operator with respect to ρ. This means that it is a *strictly* monotonic modal operator with respect to ρ. Exactly this same logic behavior is known from intentions [13]: an intention for p and an intention for q do not necessarily give an intention for $p \wedge q$, because we may intend p for another point in the future then the point for which we intend q. That the behavior of deadline obligations is similar to that of future directed intentions is not unlikely, given the intuition that intentions can be seen as a kind of obligations to oneself. A consequence of the absence of agglomeration is that it is consistent to have $O^V(p \leq \delta) \wedge O^V(\neg p \leq \delta)$. Consistency of obligations of the form Op and $O\neg p$ is heavily debated in deontic logic. Here we have consistency simply because we are free to choose our time of compliance, as long as it is before the deadline.

Also we get that the operator is anti-monotonic with respect to δ (other terminology: validity of the operator is closed under strengthening of δ):

$$\models O^V(\rho \leq \delta) \rightarrow O^V(\rho \leq (\delta \wedge \gamma))$$
$$\models O^V(\rho \leq (\delta \vee \gamma)) \rightarrow O^V(\rho \leq \delta)$$

For this version of the operator, this is in accordance with intuition: if δ is made logically stronger, it is harder to satisfy. And if ρ already has to be accomplished before the weaker condition occurs, it will certainly have to be accomplished before the stronger condition occurs. This property does not go through for the version of the deadline we discuss in section 5. As said, in that variant, the obligation is discharged by the first condition δ we meet. Then, by strengthening δ, it is not necessarily the case that we preserve the obligation.

A property we do not have for $O^V(\rho \leq \delta)$ is 'weak agglomeration' with respect to ρ, i.e.:

$$\not\models O^V(\rho \leq \delta) \wedge O^V(\rho \leq \gamma) \rightarrow O^V(\rho \leq (\delta \vee \gamma))$$

This means that the deontic deadline operator $O^V(\rho \leq \delta)$, is strictly anti-monotonic with respect to δ. If it would also obey weak agglomeration, it would have been a window operator [14, 15], which means that it would have been a normal modal operator with respect to $\neg\delta$ [16, 17]. However, the operator is *strictly* anti-monotonic. This is intuitive. Weak agglomeration should not hold, because having to achieve something before tomorrow and having to achieve the same thing before the end of the day does not imply that I have the choice to do it before tomorrow *or* before the end of the day: it simply gives me no other choice than to do it before the end of the day.

The combination of monotony for ρ and anti-monotony for δ gives us the following transitivity property for the deontic deadline operator $O^V(\rho \leq \delta)$ [10]:

[10] Taking advantage of the definability in CTL, it can be shown that we actually obey a stronger version of this property: $\models O^V(\rho \leq \delta) \wedge AG(\delta \rightarrow \chi) \wedge O^V(\chi \leq \gamma) \rightarrow O^V(\rho \leq \gamma)$

$$\models O^V(\rho \leq \delta) \wedge O^V(\delta \leq \gamma) \rightarrow O^V(\rho \leq \gamma)$$

Also this property is intuitive: if an agent is obliged to brush his teeth before going to bed, and take a medicine before he brushes his teeth, then he is certainly obliged to take has medicine before going to bed.

Clearly, the deadline operator should not be symmetric. Indeed we have:

$$\not\models O^V(\rho \leq \delta) \rightarrow O^V(\delta \leq \rho)$$

Another property we do obey is reflexivity:

$$\models O^V(\gamma \leq \gamma)$$

This is exactly the reason why we use the symbol '$\leq$' and not the symbol '$<$', in the denotation for the operator. If we achieve the obliged condition *at* the point of the deadline, we are still in time. In particular, if the deadline condition itself coincides with the condition we are obliged to achieve, whatever this condition is, we are always 'just in time' to meet the deadline. However, some would say that it is counter-intuitive to actually always be *obliged* to achieve any γ up and until γ.

We might argue that the situation is comparable to the axiom $O\top$ of standard deontic logic. The common denominator of these properties is that they concern an obligation for something that actually can never be violated. The point is that although it seems strange that our logic validates obligations for things that cannot be violated, it is not harmful either. No agent will ever let his decision making be influenced by obligations for things that are true inevitably and always. In other words, such obligations are void. However, we will see in section 7 a solution to a related, but more serious problem will discard this property, which means that we no longer have to defend it by saying that it is counter-intuitive but harmless.

Let us now consider the related issue of having a tautology or contradiction as deadline, or as a condition to achieve. We first consider the case where ρ equals $\top$. We have that:

$$\models O^V(\top \leq \delta)$$

This is related to the monotony with respect to ρ; we can weaken ρ up until it coincides with $\top$. This situation is similar to standard deontic logic's $O\top$, which we already discussed.

Just like we can weaken ρ up until $\top$, we can strengthen δ up until $\bot$ (from the anti-monotony with respect to δ).

$$\models O^V(\rho \leq \bot)$$

Clearly, $\bot$ is a condition that will be never met. So, an obligation to perform something before the (absent) point that $\bot$, can never be violated. We can postpone the obligation forever, without ever falling pray to a violation. In our view, such obligations are void. Therefore, also this case is similar to standard deontic logic's $O\top$.

Another issues is the case where ρ equals $\bot$ or δ equals $\top$. These conditions deserve extra attention. First we discuss the case where ρ equals $\bot$. This concerns the question whether something general holds for obligations for conditions that under no circumstance can be achieved. One view is that obligations of the form $O(\bot \leq \delta)$ are impossible or inconsistent. After all, it seems reasonable to take the position that one can never be obliged to achieve the impossible. This view would demand that we validate $\neg O^V(\bot \leq \delta)$, which is similar to standard deontic logic's D-axiom $\neg O\bot$. However, it is clear that we do not validate $\neg O^V(\bot \leq \delta)$. In our semantics, this would mean that we validate $EF(\delta \wedge \neg Viol)$, which directly contradicts our intuitions: it is not the case that any condition δ will be met eventually. But this does not answer the question whether we *should* obey $\neg O^V(\bot \leq \delta)$. We belief we should not. Note first that our setting is weaker than that of standard deontic logic. In particular, since we do not have agglomeration, we can satisfy $O^V(p \leq \delta) \wedge O^V(\neg p \leq \delta)$. This simply says that before δ, we have to satisfy p at some point, and we have to satisfy $\neg p$ at some point. That this cannot be the same point does not exclude the conjunction. However, this does not yet explain why it is not excluded that we satisfy $O^V(\bot \leq \delta)$. This is because this is no ordinary obligation but a deadline obligation. As we already discussed, we can have that the deadline itself is a condition that can never occur. And we argued that for that situation, the obligation is trivially met. But then we can also satisfy the formula $O^V(\bot \leq \delta)$ by choosing $\bot$ for δ. We get $O^V(\bot \leq \bot)$, which is not only satisfiable, but also valid. So, obligations of the form $O^V(\bot \leq \delta)$ are not inconsistent; in particular they can be met if δ never occurs. Intuitively: an agent can consistently meet up to the obligation to do something impossible before δ just in case that δ will never occur. Analogously, we can discuss the case where δ equals $\top$. Now the agent is obliged to achieve ρ *now*. In our semantics this is possible. Therefore $O^V(\rho \leq \top)$ is satisfiable. Similar to the above case, we may even choose ρ to be $\top$ to get the valid formula $O^V(\top \leq \top)$, which says: an agent is obliged to obey a tautology now.

However, from the above discussion, it follows that there is a deadline obligation that really should be inconsistent: $O^V(\bot \leq \top)$: agents cannot achieve the impossible now, since, by definition, the present state is not an impossibility. And indeed, we have the following property:

$$\models \neg O^V(\bot \leq \top)$$

5 A variant without strengthening of the deadline condition

The deadlines as discussed in section 3 are not discarded by meeting the deadline: as long as the condition ρ is not yet achieved, we have a violation at every point where the deadline condition δ holds. In other words: the obligation is not discarded by having failed a deadline. Here we drop this property. Thus the obligation is dropped the first time we meet the deadline condition, irrespective of whether we have achieved the goal or not. In the definition of section 3, we need to add that only the first δ occurring, is relevant.

$$\begin{aligned}\mathcal{M}, s \models O'^V(\rho \leq \delta) \Leftrightarrow\ & \forall\sigma \text{ with } \sigma_0 = s, \forall j : \\ & \text{if} \\ & \mathcal{M}, \sigma_j \models \neg\rho \wedge \delta \text{ and } \forall 0 \leq i < j : \mathcal{M}, \sigma_i \models \neg\rho \wedge \neg\delta \\ & \text{then} \\ & \mathcal{M}, \sigma_j \models Viol\end{aligned}$$

This says: if at some future point the deadline occurs for the first time, and until then the result has not yet been achieved, then we have a violation at that point. For this notion of deadline it is a slightly harder to give a CTL characterization. We need to use the notion of until that talks about the states until the last state before φ (i.e., $\psi U^e \varphi$).

$$O'^V(\rho \leq \delta) \equiv_{def} \neg E((\neg\rho \wedge \neg\delta) U^e (\delta \wedge \neg\rho \wedge \neg Viol))$$

The main point of this variant is thus that it has a different dynamical behavior. In particular, it is discarded by the first δ, even if the achievement has not been met. Therefore, the following preservation property holds:

$$\models O'^V(\rho \leq \delta) \rightarrow A(O'^V(\rho \leq \delta) U_w (\rho \vee \delta))$$

For this variant all logical properties of the variant of section 3 hold, except strengthening. Thus:

$$\not\models O'^V(\rho \leq \delta) \rightarrow O'^V(\rho \leq (\delta \wedge \gamma))$$
$$\not\models O'^V(\rho \leq (\delta \vee \gamma)) \rightarrow O'^V(\rho \leq \delta)$$

It is clear that the following holds for the relation between the two variants:

$$\models O^V(\rho \leq \delta) \rightarrow O'^V(\rho \leq \delta)$$

6 A counter-intuitive logical property

The operators defined in sections 3 and 5 obey intuitive properties. However, there is a property, or more precise, a class of properties, which are satisfied by it, but whose intuitiveness is disputable. These possibly counter-intuitive properties are caused by the definition of a deadline from the viewpoint of its violation conditions only. The idea behind the definitions was 'give an exact temporal characterization of the conditions under which the deadline is *violated*'. This idea is correct as long as we are interested in the temporal conditions *implied by* a deontic deadline. But what about the temporal conditions that *give rise to* a deontic deadline? It turns out that here something might be missing. For instance, we have the following property (From now on we will only consider the first version of the operator. The discussion for the other version is analogous.):

$$\models \rho \rightarrow O^V(\rho \leq \delta)$$

It says that the deadline obligation of section 3 is implied by the actual achievement of ρ in the current state. Moreover, this property is only an instance of a more general, stronger property that holds for the deontic deadline operator of section 3. The obligation is valid in any state where it is sure that the deadline will be met. In particular:

$$\models \neg E(\neg\rho U \delta) \rightarrow O^V(\rho \leq \delta)$$

This can be verified by substituting the CTL characterization of the deadline obligation: $\neg E(\neg\rho U \delta) \rightarrow \neg E(\neg\rho U(\delta \wedge \neg Viol))$. We may see this as the strengthening of δ to $\delta \wedge \neg Viol$ in the schema $\neg E(\neg\rho U \delta)$. It is quite easy to see that this strengthening property holds. We start with the fact that validity of the the schema $E(\varphi U \psi)$ is closed under *weakening* with respect to φ and with respect to ψ, that is, if at some point in a model we satisfy $E(\varphi U \psi)$, we also satisfy both $E((\varphi \vee \gamma)U\psi)$ and $E(\varphi U(\psi \vee \gamma))$. But this means[11] that the schema $\neg E(\varphi U \psi)$ is closed under *strengthening* with respect to ψ, which is what we needed to show (with $\neg\rho$ substituted for φ, and δ for ψ).

Now the question rises whether we cannot defend intuitiveness of this property in the same way as we defended intuitiveness of, for instance $\models O^V(\gamma \leq \gamma)$ and $\models O^V(\top \leq \delta)$ and $\models O^V(\rho \leq \bot)$. We might argue that if ρ is unavoidable, in particular, if it is true now, then the deadline $O^V(\rho \leq \delta)$ is void, because it concerns an achievement that is met anyway.

However, we consider the issue whether or not $\rho \rightarrow O^V(\rho \leq \delta)$ to be different from, for instance, the issue whether or not $O^V(\top \leq \delta)$. Whereas the second obligation is void because the obligation concerns a tautology, i.e., something that is considered to be true inevitably and always, the first obligation results from a condition that can be considered to be only *occasionally* true. Therefore, we would like to have a mechanism that enables us to avoid this property while retaining the good properties.

7 A solution

We argue that this problem is caused by the fact that we model the obligation only from the point of view of its violation conditions. We show that the undesired property is eliminated by considering success conditions also. The solution we arrive at, preserves the good properties. First we investigate how we can define a deadline operator $O^S(\rho \leq \delta)$ using success conditions (propositional 'ideality'

[11] We actually use some background theory here about how logical properties of defined operators can be determined by looking at the way they are constructed from simpler operators. In particular, a negation in the definition flips closure under strengthening to closure under weakening and vice versa. This is why any modal operator $M\varphi$ is closed under weakening (strengthening) if and only if its dual $\neg M \neg \varphi$ is closed under weakening (strengthening).

constants) only. We show that if we look at the operator from this more positive angle, we arrive at similar logical properties. However, also this approach has a (quite obvious) counter-intuitive consequence. We show that to eliminate all counter-intuitive properties we may combine both failure and success conditions.

We extend the language $\mathcal{L}_{\text{CTL}}$ with an ideality constant[8] Idl. Let $\mathcal{M}$ be a CTL model, s a state, and σ a full path starting at s. We can now define a success condition based semantics for a deontic deadline operator $O^S(\rho \leq \delta)$, where the S stands for Success, as follows:

$$\begin{aligned}\mathcal{M}, s \models O^S(\rho \leq \delta) \Leftrightarrow\ & \forall \sigma \text{ with } \sigma_0 = s, \forall j : \\ & \text{if} \\ & \mathcal{M}, \sigma_j \models \delta \\ & \text{then} \\ & \exists 0 \leq i \leq j : \mathcal{M}, \sigma_i \models \rho \wedge Idl\end{aligned}$$

This says: for all possible futures it holds that if at some point the deadline occurs, then until then, there has at least been one ideal state where ρ has been achieved. Note that it would not be correct to define that *all* ρ-states before δ are ideal; if a ρ is met, the obligation is discharged, and no ideal states should occur anymore[12].

The above semantic definition is equivalent to the following definition as a reduction to CTL:

$$O^S(\rho \leq \delta) \equiv_{def} \neg E(\neg(\rho \wedge Idl) U \delta)$$

Note that due to its form, this definition also obeys all the logical properties discussed in section 4. To be more precise, also the operator $O^S(\rho \leq \delta)$ is a monotonic operator with respect to ρ (i.e., closed under weakening with respect to ρ), and an anti-monotonic operator with respect to δ (i.e., closed under strengthening with respect to δ). And, in addition, it does *not* obey the counter-intuitive $\rho \rightarrow O^S(\rho \leq \delta)$, because now it requires the presence of an ideal state to have an obligation of the form $O^S(\rho \leq \delta)$. To be more precise, we have that:

$$\not\models \neg E(\neg \rho U \delta) \rightarrow O^S(\rho \leq \delta)$$

This follows, because, as we argued in section 4, the construct $\neg E(\neg \rho U \delta)$, is not closed under agglomeration with respect to ρ, which implies that it is certainly not anti-monotonic (closed under strengthening) with respect to ρ. So ρ cannot be strengthened to $\rho \wedge Idl$ while preserving truth.

However, obviously, also with this operator something is wrong. We have that:

$$\models O^S(\rho \leq \delta) \rightarrow \neg E(\neg \rho U \delta)$$

[12] This actually implies that an ideal state can only be the *first* ρ-state encountered before the deadline δ. The consequences of introducing this stronger condition will be investigated on another occasion.

That is, deadline obligations $O^S(\rho \leq \delta)$ cannot be violated; success is guaranteed. Before giving the remedy, let us first explain why the above property is valid for the success based deadline definition. Validity of the schema $\neg E(\neg(\rho \wedge Idl) U \delta)$ is closed under weakening with respect to $\rho \wedge Idl$, so weakening $\rho \wedge Idl$ to ρ, does not destroy truth.

Now how can we combine the intuitions from the present section with the ones of the previous sections, to arrive at a deadline operator that excludes all counterintuitive properties? We will not give the semantic truth-conditions of this final operator we define, and leave it to a characterization as a CTL formula (the semantic truth-conditions can easily be obtained by combining the conditions for the earlier defined operators):

$$O(\rho \leq \delta) \equiv_{def} \neg E(\neg(\rho \wedge Idl) U(\delta \wedge \neg Viol))$$

First of all, it is clear that this operator preserves the good properties. Due to its form we have monotonicity with respect to ρ, anti-monotonicity with respect to δ, etc. But we also avoid the counter-intuitive property $\neg E(\neg \rho U \delta) \rightarrow O(\rho \leq \delta)$, because we have strengthened ρ to $\rho \wedge Idl$. And we avoid the counter-intuitive $O(\rho \leq \delta) \rightarrow \neg E(\neg \rho U \delta)$, because we have strengthened δ to $\delta \wedge \neg Viol$ (which means that δ is weaker than $\delta \wedge \neg Viol$). Informally, the formula says that there is a deadline obligation only if there is a violation if the achievement is not met at the deadline, or there is success if the achievement is accomplished before the deadline.

A positive side-effect of this operator is that we now have that $\not\models O(\gamma \leq \gamma)$ and $\not\models O(\top \leq \delta)$. So, some of the properties we considered to be intuitively unattractive, but harmless, are no longer valid. But, we do still have that $\models O(\rho \leq \bot)$.

8 Discussion and conclusion

Given that obligation concerns action, that action involves change, and that change presupposes time, deontic and temporal aspects have very strong conceptual connections. Therefore, any contribution to the study of such connections is welcome.

In this paper we discussed intuitions concerning the notion of ‘being obliged to obey a condition ρ before a condition δ occurs’. We made a simplifying assumption that enabled us to study this notion in the logic CTL, minimally extended with violation constants. We defined two dyadic modal operators for the mentioned notion, and showed that they obey several intuitive logical properties. Finally, to prevent the operators from obeying some counter-intuitive property also, we proposed to consider success conditions.

It would be interesting to test the logic by means of a CTL-theorem prover. There are no such implemented theorem provers available. However,they can be written by using the results in either [18] or [19]. We plan to do this in the near future.

There are many directions for future research. For instance, we want to investigate whether the semantics also applies to other temporal formalisms (in particular LTL). Another point concerns abandoning the background assumption that there are no deontic updates. How much of the theory can be preserved if we do allow updates? Also we want to study the notion of permission in this setting (a simple definition is $P(\rho \leq \delta) \equiv_{def} \neg O(\neg\rho \leq \delta)$).

Finally we note that the combination of failure and success conditions was used before in deontic formalisms [20]. However, to our knowledge, the idea to evaluate failure and success conditions at *different* points in time for defining the semantics of a deontic concept, is new.

We thank Leendert van der Torre, Joris Hulstijn, Mehdi Dastani and Henry Prakken for lively and illuminating discussions on this subject. We thank the anonymous referees for valuable suggestions and references.

References

1. Dean, T., Firby, R., Miller, D.: Hierarchical planning involving deadlines, travel time, and resources. Computational Intelligence **6:1** (1990) 381–398
2. Clarke, E., Emerson, E., Sistla, A.: Automatic verification of finite-state concurrent systems using temporal logic specifications. ACM Transactions on Programming Languages and Systems **8** (1986)
3. Emerson, E.: Temporal and modal logic. In Leeuwen, J.v., ed.: Handbook of Theoretical Computer Science, volume B: Formal Models and Semantics. Elsevier Science (1990) 996–1072
4. Clarke, E., Grumberg, O., Long, D.: Verification tools for finite-state concurrent systems. In: A decade of concurrency. Volume 803 of Lecture Notes in Computer Science. Springer (1993) 124–175
5. Wright, G.v.: Deontic logic. Mind **60** (1951) 1–15
6. Broersen, J., Dastani, M., Torre, L.v.d.: BDIO-CTL: Obligations and the specification of agent behavior. In: Proceedings of Eighteenth International Joint Conference on Artificial Intelligence (IJCAI2003). (2003)
7. Anderson, A.: A reduction of deontic logic to alethic modal logic. Mind **67** (1958) 100–103
8. Kanger, S.: New foundations for ethical theory. In Hilpinen, R., ed.: Deontic Logic: Introductory and Systematic Readings. D. Reidel Publishing Company (1971) 36–58
9. Dignum, F., Kuiper, R.: Combining dynamic deontic logic and temporal logic for the specification of deadlines. In Jr., R.S., ed.: Proceedings of thirtieth HICSS. (1997)
10. Rao, A., Georgeff, M.: Modeling rational agents within a BDI-architecture. In Allen, J., Fikes, R., Sandewall, E., eds.: Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR'91), Morgan Kaufmann Publishers (1991) 473–484
11. Schnoebelen, P.: The complexity of temporal logic model checking. In Balbiani, P., Suzuki, N.Y., Wolter, F., Zakharyaschev, M., eds.: Advances in Modal Logic. Volume 4. (2003) 393–436
12. Schild, K.: On the relationship between BDI-logics and standard logics of concurrency. Autonomous agents and multi-agent systems **3** (2000) 259–283

13. Cohen, P., Levesque, H.: Intention is choice with commitment. Artificial Intelligence **42** (1990) 213–261
14. Benthem, J.v.: Minimal deontic logics. Bulletin of the Section of Logic **8** (1979) 36–42
15. Lutz, C., Sattler, U.: The complexity of reasoning with boolean modal logic. In: Advances in Modal Logic. Volume 3., World Scientific (2002) 365–387
16. Gasquet, O., Herzig, A.: Translating non-normal modal logics into normal modal logics. In Jomes, A., Sergot, M., eds.: Proceedings International Workshop on Deontic Logic, TANO, Oslo (1993)
17. Gasquet, O., Herzig, A.: (From classical to normal modal logics)
18. Bolotov, A., Fisher, M.: A clausal resolution method for CTL branching-time temporal logic. Journal of Experimental and Theoretical Artificial Intelligence **11** (1999) 77–93
19. Shankar, N.: Machine-assisted verification using theorem proving and model checking. In Broy, M., ed.: Mathematical Methods in Program Development. Springer (1997)
20. Torre, L.v.d., Tan, Y.: Diagnosis and decision making in normative reasoning. Artificial Intelligence and Law **7** (1999) 51–67