

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/jj/virginia@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 for the first time. 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’ [6], 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 aspect¹.

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 LTL [8], and a standard deontic logic, say SDL [15], and combine the two in one system. This type of approach is for instance taken in [2]. However, this gives us a complex language that is hard to handle, both conceptually and computationally. Conceptually,

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’.

we have to keep in mind that 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 assertions. 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* a certain point in time, or (2) is a deadline an obligation to achieve something, that persists in time until a deadline is reached, or (3) is it both? In natural language it is actually quite hard to be more precise about this distinction. Therefore, for now, we rely on an informal understanding of the branching time temporal logic CTL [3, 8, 4], and the standard deontic logic SDL [15], to discuss the distinction using formulas. The CTL-operators E and A are an existential and a universal quantifier respectively, ranging over possible future courses (branches) of time. The CTL-operators $\varphi U \psi$ (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), $X\varphi$ (φ holds neXt), $G\varphi$ (φ holds Globally in the future), and $F\varphi$ (at some point in the Future φ), are linear time operators ranging over moments of individual future courses of time (from now one simply called ‘possible futures’). From SDL we use the operator O , for obligation.

The first of the above three concepts corresponds with the formula $O(\neg E(\neg \rho U \delta))$, which says ‘it is obliged that on no possible future ρ is avoided up until δ becomes true’. Alternatively we could write $F(E(\neg \rho U \delta))$, which says ‘it is forbidden that on some possible future ρ is avoided up until δ becomes true’. The second of the above notions corresponds with the formula $A(O\rho U_w \delta)$, which says ‘the obligation to achieve ρ persists until δ , and if δ does not occur, it persists forever’. However, we think it is intuitive that obligations are not only discharged by the deadline condition, but also by the realization of the condition that is obliged. So, we assume that the obligation is preserved until the deadline *or* the point of its achievement: $A(O\rho U_w(\rho \vee \delta))$ (see also [2]). The third alternative, where we have a deadline obligation that itself (the obligation to achieve something before a certain point in time) is preserved until it is achieved or until the deadline is met, is then characterized by the formula $A(O(\neg E(\neg \rho U \delta))U_w(\rho \vee \delta))$.

We are interested in deadlines that correspond to the third point (and formula) above. However, the above discussion makes it clear that designing a logic of deontic deadlines is not a trivial matter. In particular it is easy to get confused by the distinction between the temporal *content* of norms and the temporal *dynamics* of norms. Therefore, we will make an assumption that will enable us to reduce the conceptual complexity, and also give a semantics using violation constants in a standard temporal language. Our assumption is that the second aspect of deadlines discussed above, is a general property of our logic. This means that our logic concerns the reasoning of an agent who considers the logical implications of deadline obligations that can only be discharged by complying before the deadline, or by failing to meet the deadline. So, it is assumed 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 nor the achievement is realized, the deadline obligation persists. This background assumption of the theory should appear as a theorem of our logic. Thus, we should have that $O_i(\rho \leq \delta) \rightarrow X(\neg(\rho \vee \delta) \rightarrow O_i(\rho \leq \delta))$. Or, more precise: $O_i(\rho \leq \delta) \rightarrow A(O_i(\rho \leq \delta)U_w(\rho \vee \delta))$.

The above discussed assumption makes it possible to study deadlines in a standard temporal logic extended with violation predicates for violations of individual agents. In this paper we choose for the branching time temporal logic CTL, because of its nice properties (P-complete complexity of the model checking problem for CTL, versus PSPACE-complete complexity for LTL [14]), and its popularity in agent theory [13]. The idea of expressing the semantics of deontic deadlines by characterizing violation conditions in CTL supplemented with violation constants, was first explored in [7].

The general idea is that a violation condition holds at exactly those moments where the agent violates a deontic deadline. This enables us to reason about violations explicitly, and about what to do if they occur, which is a distinctive feature of deontic reasoning. 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 true at some point in time. This abstract view contributes to the relevance of the present research for other logical systems. For instance, Rao and Georgeffs commitment strategies [12] are actually a sort of deadlines: an agent commits to an intention until the action is performed or believed not be feasible any longer.

2 Preliminaries: CTL

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

$$\varphi, \psi, \dots := p \mid \neg\varphi \mid \varphi \wedge \psi \mid E(\varphi U \psi) \mid A(\varphi U \psi)$$

where φ, ψ represent arbitrary well-formed formulas, and where the p are elements from an infinite set of propositional symbols \mathcal{P} . The standard propositional abbreviations are assumed. Additionally the following abbreviations are applied:

$$\begin{array}{ll} EX\varphi \equiv_{def} E(\perp U \varphi) & AX\varphi \equiv_{def} \neg EX\neg\varphi \\ EF\varphi \equiv_{def} E(\top U \varphi) & AG\varphi \equiv_{def} \neg EF\neg\varphi \\ AF\varphi \equiv_{def} A(\top U \varphi) & EG\varphi \equiv_{def} \neg AF\neg\varphi \\ A(\varphi U_w \psi) \equiv_{def} \neg E(\neg\psi U \neg\varphi) & E(\varphi U_w \psi) \equiv_{def} \neg A(\neg\psi U \neg\varphi) \end{array}$$

The CTL-operators have the following informal meanings:

- $E(\varphi U \psi)$: there is a future for which eventually ψ will hold,
while φ holds until then
- $E(\varphi U_w \psi)$: there is a future such that, if eventually ψ will hold,
then φ holds until then, or forever otherwise
- $A(\varphi U \psi)$: for all possible futures eventually ψ will hold,
while φ holds until then
- $A(\varphi U_w \psi)$: for all possible futures, if eventually ψ will hold,
then φ holds until then, or forever otherwise
- $EX\varphi$: there is a next moment for which φ will hold
- $AX\varphi$: at any next moment φ will hold
- $EF\varphi$: for some future, eventually φ will hold
- $AG\varphi$: for all possible futures φ will be preserved
- $AF\varphi$: for all futures, eventually φ will hold
- $EG\varphi$: there is a future that preserves φ

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 \mathcal{M} is a sequence $\sigma = s_0, s_1, s_2, \dots$ 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 a full path starts at s if and only if $s_0 = s$. We denote the state s_i of a full path $\sigma = s_0, s_1, s_2, \dots$ in \mathcal{M} by σ_i .

Validity $\mathcal{M}, s \models \varphi$, of a CTL-formula φ in a world s of a modal action model \mathcal{M} is defined as (propositional symbols are interpreted as usual):

- $\mathcal{M}, s \models E(\varphi U \psi) \Leftrightarrow \exists \sigma$ in \mathcal{M} with $\sigma_0 = s$, and $\exists n$ such that:
 - (1) $\mathcal{M}, \sigma_n \models \psi$ and
 - (2) $\forall i$ with $0 \leq i \leq n$ it holds that $\mathcal{M}, \sigma_i \models \varphi$
- $\mathcal{M}, s \models A(\varphi U \psi) \Leftrightarrow \forall \sigma$ in \mathcal{M} such that $\sigma_0 = s$ it holds that $\exists n$ such that
 - (1) $\mathcal{M}, \sigma_n \models \psi$ and
 - (2) $\forall i$ with $0 \leq i \leq n$ it holds that $\mathcal{M}, \sigma_i \models \varphi$

Validity on a CTL model \mathcal{M} is defined as validity in all worlds 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}_{CTL} over the class of CTL models.

3 A dyadic deontic deadline operator

We extend the language \mathcal{L}_{CTL} by extending the set of propositional atoms with an infinite set of violation constants of the form $Viol(a)$, where $a \in \mathcal{A}$, and \mathcal{A} , an infinite set of agent identifiers. Note that this extends the language only minimally. furthermore, the formal interpretation of atoms $Viol(a)$ is treated like that of all other atomic propositions. So, we can view the set of propositional

constants of the form $Viol(a)$ also as a subset of \mathcal{P} : an infinite subset of ‘special purpose’ propositions for the application of CTL to the normative context.

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

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

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*:

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

This formula simply ‘negates’ the situation that should be excluded when a deontic deadline is in force. 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 in some state if and only if the the CTL formula is true in that state.

4 Logic Properties

What logic properties of the operator $O_a^V(\rho \leq \delta)$ does this bring us? First of all we get monotonicity with respect to ρ (other terminology: validity of the operator is closed under weakening of ρ)². Monotonicity says that we have as validities³:

$$\begin{aligned} &\models O_a^V((\rho \wedge \chi) \leq \delta) \rightarrow O_a^V(\rho \leq \delta) \\ &\models O_a^V(\rho \leq \delta) \rightarrow O_a^V((\rho \vee \chi) \leq \delta) \end{aligned}$$

This is perfectly 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

² All verifications are left to the reader.

³ 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.

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.:

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

This shows that $O_a^V(\rho \leq \delta)$, is monotonic with respect to ρ , but is *not* a normal modal operator with respect to ρ . So, it is a strictly monotonic modal operator with respect to ρ . Exactly this same logic behavior is known from intentions [5]: 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 $O_a^V(p \leq \delta) \wedge O_a^V(\neg p \leq \delta)$. Consistency of obligations of the form O_p 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):

$$\begin{aligned} &\models O_a^V(\rho \leq \delta) \rightarrow O_a^V(\rho \leq (\delta \wedge \gamma)) \\ &\models O_a^V(\rho \leq (\delta \vee \gamma)) \rightarrow O_a^V(\rho \leq \delta) \end{aligned}$$

This is perfectly 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.

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

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

This means that the deontic deadline operator $O_a^V(\rho \leq \delta)$, is strictly anti-monotonic with respect to δ . If it would also obey weak agglomeration, it would have been a window operator [1, 11], which means that it would have been a normal modal operator with respect to $\neg\delta$ [10, 9]. 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 $O_a^V(\rho \leq \delta)$ ⁴:

⁴ Taking advantage of the definability in CTL, it can be shown that we actually obey a stronger version of this property: $\models O_a^V(\rho \leq \delta) \wedge AG(\delta \rightarrow \chi) \wedge O_a^V(\chi \leq \gamma) \rightarrow O_a^V(\rho \leq \gamma)$

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

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:

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

Another property we do obey is reflexivity:

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

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 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.

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:

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

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 \perp (from the anti-monotony with respect to δ).

$$\models O_a^V(\rho \leq \perp)$$

Clearly, \perp is a condition that will be never met. So, an obligation to perform something before the (absent) point that \perp , 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 \perp or δ equals \top . These conditions deserve extra attention. First we discuss the case where ρ equals \perp . 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 $O_a(\perp \leq \delta)$ 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 $\neg O_a^V(\perp \leq \delta)$, which is similar to standard deontic logic's D-axiom $\neg O\perp$. However, it is clear that we do not validate $\neg O_a^V(\perp \leq \delta)$. In our semantics, this would mean that we validate $EF(\delta \wedge \neg Viol(a))$, 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 $\neg O_a^V(\perp \leq \delta)$. We believe 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 $O_a^V(p \leq \delta) \wedge O_a^V(\neg p \leq \delta)$. 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 we can also satisfy $O_a^V(\perp \leq \delta)$. 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 $O_a^V(\perp \leq \delta)$ by choosing \perp for δ . We get $O_a^V(\perp \leq \perp)$, which is not only satisfiable, but also valid. So, obligations of the form $O_a^V(\perp \leq \delta)$ 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 $O_a^V(\rho \leq \top)$ is satisfiable. Similar to the above case, we may even choose ρ to be \top to get the valid formula $O_a^V(\top \leq \top)$, 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: $O_a^V(\perp \leq \top)$: agents cannot achieve the impossible now, since, by definition, the present state is not an impossibility. And indeed, we have the following property:

$$\models \neg O_a^V(\perp \leq \top)$$

There is one property of the logic we did not discuss yet, namely, the background assumption mentioned in the introduction. Indeed, we satisfy

$$\models O_i(\rho \leq \delta) \rightarrow A(O_i(\rho \leq \delta)U_w(\rho \vee \delta))$$

To see that this holds, 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 δ , nor do we 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. But, if we do meet ρ or δ , the conditions are no longer guaranteed.

5 A counter-intuitive logic property

The operator defined in the previous section obeys 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 definition 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:

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

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

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

This is easy to verify, by substituting the CTL characterization of the deadline obligation: $\neg E(\neg \rho U \delta) \rightarrow \neg E(\neg \rho U (\delta \wedge \neg Viol(a)))$. This must hold, since, if $\neg E(\neg \rho U (\delta \wedge \neg Viol(a)))$ is anti-monotonic with respect to δ , then also $\neg E(\neg \rho U \delta)$ is. So, by strengthening δ to $\delta \wedge \neg Viol(a)$ we preserve truth.

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

However, we consider the issue whether or not $\rho \rightarrow O_a^V(\rho \leq \delta)$ to be slightly different from, for instance, the issue whether or not $O_a^V(\top \leq \delta)$. 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.

6 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 $O_a^S(\rho \leq \delta)$ using success conditions (propositional ‘ideality’ constants) only. We show that if we look at the operator from this more positive angle, we arrive at similar logic properties. However, also this approach has a (quite obvious) counter-intuitive consequence. We show that to eliminate all counter-intuitive properties we have to combine both failure and success conditions.

We extend the language \mathcal{L}_{CTL} with an infinite set of ideality constants of the form $Idl(a)$, where $a \in \mathcal{A}$, and \mathcal{A} , the infinite set of agent identifiers.

Let \mathcal{M} be 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 $O_a^S(\rho \leq \delta)$, where the S stands for *Success*, as follows:

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

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. This semantic definition is equivalent to the following definition as a reduction to CTL :

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

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

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

This follows, because, as we argued in the previous section, 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 $\rho \wedge Idl(a)$ while preserving truth.

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

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

That is, deadline obligations $O_a^S(\rho \leq \delta)$ cannot be violated; success is guaranteed. Before giving the remedy, let us first explain why the above property is valid. Again this follows from the properties we discussed in the previous section.

In particular, the construction $\neg E(\neg(\rho \wedge Idl(a)U\delta))$ is monotonic with respect to $\rho \wedge Idl(a)$, so weakening $\rho \wedge Idl(a)$ to ρ , does not destroy truth.

Now how can we combine the intuitions from the two definitions so far, to arrive at a deadline operator that excludes all counterintuitive properties? We will not give a the semantic truth-conditions of this third and 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 two defined operators):

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

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 $\neg E(\neg\rho U\delta) \rightarrow O_a(\rho \leq \delta)$, because we have strengthened ρ to $\rho \wedge Idl(a)$. And we avoid the counter-intuitive $O_a(\rho \leq \delta) \rightarrow \neg E(\neg\rho U\delta)$, because we have strengthened δ to $\delta \wedge \neg Viol(a)$ (which means that δ is weaker than $\delta \wedge \neg Viol(a)$). 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.

7 Discussion and conclusion

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

It would be interesting to test the logic, using a CTL-theorem prover. However, to our knowledge, there is still not such a prover available. The apparent absence of such a prover is quite surprising, given the popularity of CTL as a temporal specification language, and the relatively good computational complexity of the logic.

We thank Leendert van der Torre, Joris Hulstijn, Mehdi Dastani and Henry Prakken for lively and illuminating discussions on this subject.

References

1. J.F.A.K. van Benthem. Minimal deontic logics. *Bulletin of the Section of Logic*, 8(1):36–42, 1979.
2. J.M. Broersen, M. Dastani, and L.W.N. van der Torre. BDIO-CTL: Obligations and the specification of agent behavior. In *Proceedings of Eighteenth International Joint Conference on Artificial Intelligence (IJCAI2003)*, 2003.

3. E.M. Clarke, E.A. Emerson, and A.P. Sistla. Automatic verification of finite-state concurrent systems using temporal logic specifications. *ACM Transactions on Programming Languages and Systems*, 8(2), 1986.
4. E.M. Clarke, O. Grumberg, and D. Long. Verification tools for finite-state concurrent systems. In *A decade of concurrency*, volume 803 of *Lecture Notes in Computer Science*, pages 124–175. Springer, 1993.
5. P.R. Cohen and H.J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(3):213–261, 1990.
6. T. Dean, R.J. Firby, and D. Miller. Hierarchical planning involving deadlines, travel time, and resources. *Computational Intelligence*, 6:1:381–398, 1990.
7. F. Dignum and R. Kuiper. Combining dynamic deontic logic and temporal logic for the specification of deadlines. In R. Sprague Jr., editor, *Proceedings of thirtieth HICSS*, 1997.
8. E.A. Emerson. Temporal and modal logic. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science, volume B: Formal Models and Semantics*, chapter 14, pages 996–1072. Elsevier Science, 1990.
9. O. Gasquet and A. Herzig. From classical to normal modal logics.
10. O. Gasquet and A. Herzig. Translating non-normal modal logics into normal modal logics. In A.I.J. Jones and M. Sergot, editors, *Proceedings International Workshop on Deontic Logic*, TANO, Oslo, 1993.
11. C. Lutz and U. Sattler. The complexity of reasoning with boolean modal logic. In *Advances in Modal Logic*, volume 3, pages 365–387. World Scientific, 2002.
12. A.S. Rao and M.P. 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.
13. K. Schild. On the relationship between BDI-logics and standard logics of concurrency. *Autonomous agents and multi-agent systems*, 3:259–283, 2000.
14. Ph. Schnoebelen. The complexity of temporal logic model checking. In P. Balbiani, N-Y. Suzuki, F. Wolter, and M. Zakharyashev, editors, *Advances in Modal Logic*, volume 4, pages 393–436, 2003.
15. G.H. von Wright. Deontic logic. *Mind*, 60:1–15, 1951.