

On Non-normal Modal Description Logics

Tiziano Dalmonte¹, Andrea Mazzullo², and Ana Ozaki²

¹ Aix Marseille Univ, Université de Toulon, CNRS, LIS, Marseille, France
tiziano.dalmonte@etu.univ-amu.fr

² Free University of Bozen-Bolzano
surname@inf.unibz.it

Abstract. Non-normal modal logics based on neighbourhood semantics can be used to formalise normative, epistemic and coalitional reasoning in autonomous and multi-agent systems, since they do not validate principles known to be problematic in applications. These principles, satisfied by all modal logics interpreted over relational frames, also affect several modal description logics (MDLs) used in knowledge representation. We study *non-normal MDLs*, obtained by extending \mathcal{ALC} -based languages with non-normal modal operators. These logics increase the expressive power of their propositional counterparts, and allow for complex modelling of obligations, beliefs, abilities and strategies. On the computational side, standard reasoning tasks are not more difficult than in basic normal MDLs, with a NEXPTIME upper bound for satisfiability that can be lowered further in fragments with modal operators only over axioms.

1 Introduction

Several approaches to the formal study of normative, epistemic and action-based notions are based on modal logic (ML) operators [9, 14, 16]. In the normative setting, for instance, the so-called *standard deontic logic* (SDL) extends propositional logic with unary operators, intuitively interpreted as ‘it is obligatory’ and ‘it is permitted’. First-order extensions have been considered as well [16]. Research on autonomous systems [11], machine ethics [2], and normative multi-agent systems [21] is drawing attention to challenging application scenarios for deontic logics in computer science. Other motivations come from knowledge management in legal domains (e.g. legal ontologies and expert systems [8, 13]), Semantic Web applications (e.g. legislative XML and RuleML [5, 18]), as well as verification of normative systems, and modelling of the normative behaviour of organisations (e.g. company policies specifications or contracting [19]).

The semantics of MLs, and of SDL in particular, is traditionally based on *relational frames*, consisting of a set of possible worlds endowed with a binary accessibility relation [9, 16]. These structures, used to interpret modal operators (e.g. deontic, epistemic, dynamic, etc.), represent the connections between possible situations. For instance, in SDL, a proposition is said to be obligatory in some possible world w , if it holds in all worlds related to w , interpreted as morally ideal alternatives to w . However, all the so-called *normal* MLs, based on this

semantics, face the problem of validating principles that, in several applications, can be hardly associated with an acceptable meaning. In SDL, these principles lead to several counter-intuitive conclusions, often presented in the form of *deontic paradoxes*. For instance, if it is obligatory to perform an action, and if this action always implies a negative consequence, then we are forced to conclude that also the latter is obligatory. Problematic arguments like this one represent a strong limitation to the applicability of SDL to normative reasoning [16].

To overcome these problems, a different semantics, based on *neighbourhood* (or *minimal*) *models*, has been proposed [9]. Instead of using a set of worlds endowed with an accessibility relation, this approach associates to each situation w a family of sets of worlds. These sets intuitively represent the propositions that are obligatory (or believed, brought about, etc.) in w . MLs based on this semantics can satisfy weaker principles, without validating those axioms and rules that are common to all normal MLs. For this reason, they are called *non-normal MLs*. At the propositional level, non-normal MLs based on neighbourhood semantics have received considerable attention [20, 10], with results reducing validity in propositional non-normal MLs to validity in normal ones [17, 15]. To increase the expressive power of these formalisms, first-order non-normal MLs based on neighbourhood semantics have been considered as well [27, 3, 7].

Not much has been done yet in applications of non-normal MLs to knowledge representation, in particular, to normative automated reasoning. Most of the modal description logics (MDLs) considered in the literature are based on the standard relational semantics [14]. Modal extensions of \mathcal{ALC} with neighbourhood semantics have been introduced as a basis of coalition logic [22, 24] and agent communication [12] languages for reasoning over structured domains. However, in normative settings, these MDLs still share several problems of propositional normal MLs. Failing to address this issue can lead to serious drawbacks to normative reasoning in knowledge-based systems. In this paper we study non-normal MDLs, interpreted over neighbourhood models, satisfying only minimal requirements on the modal operators. With these formalisms, counter-intuitive inferences in normative scenarios can be blocked, while still retaining the expressive power needed in knowledge representation.

In Section 2 we present MDLs, both recalling the standard relational semantics, and introducing the neighbourhood models used for non-normal MDLs. In Section 3 we model with MDLs a scenario involving normative notions, discussing deontic paradoxes due to relational semantics, and how they can be blocked using neighbourhood models. Then, in Section 4, we study the complexity of the formula satisfiability problem for non-normal MDLs. We prove NEXPTIME-upper bounds for the complexity of the satisfiability problem, showing that reasoning is not harder than in basic (normal) modal DLs with the relational semantics [14]. Directions for future work are discussed in Section 5.

2 Preliminaries

We start by introducing the required notation for normal and non-normal MDLs.

Syntax. Let \mathbf{N}_C , \mathbf{N}_R and \mathbf{N}_I be countably infinite and pairwise disjoint sets of *concept names*, *role names*, and *individual names*, respectively. An $ML_{\mathcal{ALC}}^n$ *concept* is an expression of the form $C, D ::= A \mid \neg C \mid C \sqcap D \mid \exists r.C \mid \Box_i C$, where $A \in \mathbf{N}_C$, $r \in \mathbf{N}_R$, and \Box_i , with $1 \leq i \leq n$, are modal operators called *boxes*. An $ML_{\mathcal{ALC}}^n$ *atom* is either a *concept inclusion* (CI) of the form $C \sqsubseteq D$, or an *assertion* of the form $A(a)$ or $r(a, b)$, where C, D are $ML_{\mathcal{ALC}}^n$ concepts, $A \in \mathbf{N}_C$, $r \in \mathbf{N}_R$, and $a, b \in \mathbf{N}_I$. An $ML_{\mathcal{ALC}}^n$ *formula* is an expression of the form $\varphi, \psi ::= \pi \mid \neg\varphi \mid \varphi \wedge \psi \mid \Box_i \varphi$, where π is an $ML_{\mathcal{ALC}}^n$ atom, and $1 \leq i \leq n$. We will use the following standard definitions for concepts: $\perp \equiv A \sqcap \neg A$, $\top \equiv \neg\perp$; $\forall r.C \equiv \neg\exists r.\neg C$; $(C \sqcup D) \equiv \neg(\neg C \sqcap \neg D)$; $\Diamond_i C \equiv \neg\Box_i \neg C$ (\Diamond_i are called *diamonds*). Concepts of the form $\Box_i C$, $\Diamond_i C$ are called *modalised concepts*. Analogous conventions also hold for formulas. In particular, we write $C \doteq D$ for $C \sqsubseteq D \wedge D \sqsubseteq C$.

Relational Semantics. A *relational frame* (or *R-frame*) is a structure $F = (W, \{R_i\}_{i \in [1, n]})$, where W is a non-empty set and each R_i is a binary relation on W . An $ML_{\mathcal{ALC}}^n$ *relational model* (or *R-model*) based on an R-frame F is a structure $M = (F, \Delta, I)$, where Δ is a non-empty set, called the *domain* of M , and I is a function associating with every $w \in W$ an \mathcal{ALC} *interpretation* (or *model*) $I(w) = (\Delta, \cdot^{I(w)})$ having domain Δ , and where $\cdot^{I(w)}$ is a function such that: for all $A \in \mathbf{N}_C$, $A^{I(w)} \subseteq \Delta$; for all $r \in \mathbf{N}_R$, $r^{I(w)} \subseteq \Delta \times \Delta$; for all $a \in \mathbf{N}_I$, $a^{I(w)} \in \Delta$, and for all $u, v \in W$, $a^{I(u)} = a^{I(v)}$ (denoted by a^I). Given an R-model $M = (F, \Delta, I)$ and a world w in F , the *interpretation of a concept C in w* , written $C^{I(w)}$, is defined by taking:

$$\begin{aligned} (\neg C)^{I(w)} &= \Delta \setminus C^{I(w)}, & (C \sqcap D)^{I(w)} &= C^{I(w)} \cap D^{I(w)}, \\ (\exists r.C)^{I(w)} &= \{d \in \Delta \mid \exists e \in C^{I(w)} : (d, e) \in r^{I(w)}\}, \\ (\Box_i C)^{I(w)} &= \{d \in \Delta \mid \forall v \in W : wR_i v \Rightarrow d \in C^{I(v)}\}, \end{aligned}$$

A concept C is *satisfied in M* if there is w in F s.t. $C^{I(w)} \neq \emptyset$, and that C is *satisfiable* (over R-models) if there is an R-model in which it is satisfied. The *satisfaction of a $ML_{\mathcal{ALC}}^n$ formula φ in w of M* , written $M, w \models \varphi$, is defined as:

$$\begin{aligned} M, w \models C \sqsubseteq D &\text{ iff } C^{I(w)} \subseteq D^{I(w)}, \\ M, w \models A(a) &\text{ iff } a^I \in A^{I(w)}, & M, w \models r(a, b) &\text{ iff } (a^I, b^I) \in r^{I(w)}, \\ M, w \models \neg\varphi &\text{ iff } M, w \not\models \varphi, & M, w \models \varphi \wedge \psi &\text{ iff } M, w \models \varphi \text{ and } M, w \models \psi, \\ M, w \models \Box_i \varphi &\text{ iff for all } v \in W: \text{ if } wR_i v, \text{ then } M, v \models \varphi. \end{aligned}$$

Given an R-frame $F = (W, \{R_i\}_{i \in [1, n]})$ and an R-model $M = (F, \Delta, I)$, we say that φ is *satisfied in M* if there is $w \in W$ s.t. $M, w \models \varphi$, and that φ is *satisfiable* (over R-models) if it is satisfied in some R-model. Also, φ is said to be *valid in M* , $M \models \varphi$, if it is satisfied in all w of M , and it is *valid on F* if, for all M based on F , φ is valid in M , writing $F \models \varphi$. Moreover, φ *logically implies* a formula ψ , writing $\varphi \models \psi$, if $M, w \models \varphi$ implies $M, w \models \psi$, for every M and every w in M . Recall that the concept satisfiability problem can be reduced to the formula satisfiability problem, since C is satisfiable iff $\neg(C \sqsubseteq \perp)$ is satisfiable.

Neighbourhood Semantics. A *neighbourhood frame* (or *N-frame*) is a pair $\mathcal{F} = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$, where \mathcal{W} is a non-empty set and, for each $1 \leq i \leq n$, $\mathcal{N}_i : \mathcal{W} \rightarrow \mathcal{P}(\mathcal{P}(\mathcal{W}))$ is called a *neighbourhood function*. A frame is *supplemented* if for all $\alpha, \beta \subseteq \mathcal{W}$, $\alpha \in \mathcal{N}_i(w)$ and $\alpha \subseteq \beta$ implies $\beta \in \mathcal{N}_i(w)$; it is *closed under intersection* if $\alpha \in \mathcal{N}_i(w)$ and $\beta \in \mathcal{N}_i(w)$ implies $\alpha \cap \beta \in \mathcal{N}_i(w)$; and it *contains the unit* if for all $w \in \mathcal{W}$, $\mathcal{W} \in \mathcal{N}_i(w)$. An $ML_{\mathcal{ALC}}^n$ *neighbourhood model* (or *N-model*) based on an N-frame \mathcal{F} is a triple $\mathcal{M} = (\mathcal{F}, \Delta, \mathcal{I})$, where $\mathcal{F} = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$ is a neighbourhood frame, Δ is a non-empty set called the *domain* of \mathcal{M} , and \mathcal{I} is a function associating with every $w \in \mathcal{W}$ an \mathcal{ALC} interpretation $\mathcal{I}(w) = (\Delta, \cdot^{\mathcal{I}(w)})$, defined as above. Given a model $\mathcal{M} = (\mathcal{F}, \Delta, \mathcal{I})$ and a world w in \mathcal{F} , the *interpretation of a concept C in w* , written $C^{\mathcal{I}(w)}$, is defined as for the relational semantics, except from:

$$(\Box_i C)^{\mathcal{I}(w)} = \{d \in \Delta \mid [C]_d^{\mathcal{M}} \in \mathcal{N}_i(w)\},$$

where, for all $d \in \Delta$, the set $[C]_d^{\mathcal{M}} = \{v \in \mathcal{W} \mid d \in C^{\mathcal{I}(v)}\}$ is called the *truth set of C with respect to d* . We say that a concept C is *satisfied in \mathcal{M}* if there is w in \mathcal{F} s.t. $C^{\mathcal{I}(w)} \neq \emptyset$, and that C is *satisfiable* (over N-models) if there is an N-model in which it is satisfied. The *satisfaction of an $ML_{\mathcal{ALC}}$ formula φ in w of \mathcal{M}* , written $\mathcal{M}, w \models \varphi$, is defined analogously to relational semantics, and as follows for modalised formulas:

$$\mathcal{M}, w \models \Box_i \varphi \quad \text{iff} \quad [\varphi]^{\mathcal{M}} \in \mathcal{N}_i(w),$$

where $[\psi]^{\mathcal{M}}$ denotes the set $\{v \in \mathcal{W} \mid \mathcal{M}, v \models \psi\}$ of the worlds v that satisfy ψ , also called the *truth set of ψ* . As a consequence of the above definition, we obtain the following condition for diamond formulas: $\mathcal{M}, w \models \Diamond_i \varphi$ iff $[\neg \varphi]^{\mathcal{M}} \notin \mathcal{N}_i(w)$. Given an N-frame $\mathcal{F} = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$ and an N-model $\mathcal{M} = (\mathcal{F}, \Delta, \mathcal{I})$, we say that φ is *satisfied in \mathcal{M}* if there is $w \in \mathcal{W}$ s.t. $\mathcal{M}, w \models \varphi$, and that φ is *satisfiable* (over N-models) if it is satisfied in some N-model. Other semantical definitions can be easily adapted from the relational semantics case.

Frames and Satisfiability Problems. In the following, we use \mathfrak{F} to stand either for an N- or R-frame, and \mathfrak{M} for a N- or R-model. To define the $ML_{\mathcal{ALC}}^n$ formula satisfiability problems studied in this paper, we consider the principles listed in Table 1 (where C, D and φ, ψ are $ML_{\mathcal{ALC}}^n$ concepts and formulas, respectively). Here, S is either a frame \mathfrak{F} , or a model \mathfrak{M} . For a principle P , if $S = \mathfrak{F}$ (respectively, $S = \mathfrak{M}$), we say that P *holds on \mathfrak{F}* (respectively, *in \mathfrak{M}*). On the correspondence between the principles in Table 1 and conditions over frames and models, we have the following result (see e.g. [20] for the propositional case).

Theorem 1. *Given an N-frame \mathcal{F} , we have that: (i) congruence holds on \mathcal{F} ; (ii) monotonicity holds on \mathcal{F} iff \mathcal{F} is supplemented; (iii) agglomeration holds on \mathcal{F} iff \mathcal{F} is closed under intersection; (iv) necessitation holds on \mathcal{F} iff \mathcal{F} contains the unit. Given an R-frame F , congruence, monotonicity, agglomeration, and necessitation hold on F ; moreover, for every R-model M , they hold in M .*

(Congruence)	$S \models C \doteq D$ implies $S \models \Box_i C \doteq \Box_i D$.
(Monotonicity)	$S \models \varphi \leftrightarrow \psi$ implies $S \models \Box_i \varphi \leftrightarrow \Box_i \psi$.
(Agglomeration)	$S \models C \sqsubseteq D$ implies $S \models \Box_i C \sqsubseteq \Box_i D$.
(Necessitation)	$S \models \varphi \rightarrow \psi$ implies $S \models \Box_i \varphi \rightarrow \Box_i \psi$.
	$S \models \Box_i C \sqcap \Box_i D \sqsubseteq \Box_i (C \sqcap D)$.
	$S \models \Box_i \varphi \wedge \Box_i \psi \rightarrow \Box_i (\varphi \wedge \psi)$.
	$S \models \top \sqsubseteq C$ implies $S \models \top \sqsubseteq \Box_i C$.
	$S \models \varphi$ implies $S \models \Box_i \varphi$.

Table 1. Principles over frames and models.

By the $ML_{\mathcal{A}\mathcal{L}\mathcal{C}}^n$ *formula satisfiability problem in a class of (respectively, N- or R-) frames* \mathcal{C} we mean the problem of deciding whether an $ML_{\mathcal{A}\mathcal{L}\mathcal{C}}^n$ formula is satisfied in a (respectively, N- or R-) model based on a frame in \mathcal{C} . The *formula satisfiability problem for $\mathbf{E}_{\mathcal{A}\mathcal{L}\mathcal{C}}^n$, $\mathbf{M}_{\mathcal{A}\mathcal{L}\mathcal{C}}^n$, and $\mathbf{K}_{\mathcal{A}\mathcal{L}\mathcal{C}}^n$* is the $ML_{\mathcal{A}\mathcal{L}\mathcal{C}}^n$ formula satisfiability problem in the class of N-frames, supplemented N-frames, and R-frames, respectively.

3 Modelling

In this section we model well-known paradoxes that normal MLs with relational semantics have to face in normative applications [16]. Firstly, the MDLs language introduced in Section 2 is used to provide a running example, that also illustrates more expressive (with respect to the propositional case) features of the language. Then, we show how principles validated by all normal MLs, and thus also by the standard MDLs based on relational models, can affect reasoning in deontic settings. We focus on the problems associated with necessitation, agglomeration, and monotonicity in normal MDLs, claiming that the flexibility of neighbourhood semantics represents an advantage in blocking problematic inferences.

Modelling Scenario. Consider the following variant of the classical *trolley problem* scenario [25]. A tram is moving towards a toddler lying on the tracks. Although it is not possible to stop the trolley, an agent (called *signaller*), possibly an autonomous control system, can activate a switch that would divert it on a side track, saving the toddler’s life. However, on the side track lies an elderly that would be killed with the activation of the switch. Therefore, the switching system has to decide among two (horrible) alternatives: (1) do not activate the switch, allowing the tram to kill the toddler; (2) activate the switch, saving the toddler’s life and allowing an elderly to be killed instead.

For modelling purposes, it is crucial to have a formalism that allows to specify both the factual features of the setting, and the ethical or legal aspects involved. We assume that the domain consists of *objects* (e.g. a switch, a signaller, a toddler, etc.), performing some *actions* (e.g. activating a switch) and bringing about some *consequences* (e.g. to save, to kill) on each other. We represent classes

of objects with (unary) concepts, such as `Switch`, `Signaller` and `Toddler`, while actions and consequences are formalised using (binary) roles, e.g., `activates`, `saves` and `kills`. Obligations and permissions are indicated using \Box and \Diamond , respectively. For instance, the concept $\exists\text{activates.Switch}$ intuitively denotes the set of objects that activate some switch, whereas $\Box\exists\text{activates.Switch}$ is the set of entities that are obliged to do so. The following example shows a simple N-model interpreting these latter concepts according to the definitions given in Section 2.

Example 1. Let $\mathcal{M} = (\mathcal{F}, \Delta, \mathcal{I})$ be a $ML_{\mathcal{ALC}}$ N-model, where $\mathcal{F} = (\mathcal{W}, \mathcal{N})$ is such that $\mathcal{W} = \{w, v\}$ and $\mathcal{N}(w) = \{\{v\}\}$, $\mathcal{N}(v) = \{\{w\}\}$. Moreover, $\Delta = \{d_1, d_2, d_3\}$, and let $\text{Switch}^{\mathcal{I}(w)} = \text{Switch}^{\mathcal{I}(v)} = \{d_2\}$, $\text{activates}^{\mathcal{I}(w)} = \emptyset$, and $\text{activates}^{\mathcal{I}(v)} = \{(d_1, d_2)\}$. We have $(\exists\text{activates.Switch})^{\mathcal{I}(w)} = \emptyset$, $(\exists\text{activates.Switch})^{\mathcal{I}(v)} = \{d_1\}$. Moreover, $[\exists\text{activates.Switch}]_{d_1}^{\mathcal{M}} = \{v\}$ and $[\exists\text{activates.Switch}]_{d_i}^{\mathcal{M}} = \emptyset$, for $i \in \{2, 3\}$. Thus, $(\Box\exists\text{activates.Switch})^{\mathcal{I}(w)} = \{d_1\}$ and $(\Box\exists\text{activates.Switch})^{\mathcal{I}(v)} = \emptyset$.

Non-normal MDLs allow also for a meaningful distinction between *de re* (applied to concepts) and *de dicto* (applied to formulas) modalities. For instance, a signaller can be defined as an agent with the permission to activate a switch, and a guard as an agent having the duty to check the rails, i.e.,

$$\text{Signaller} \doteq \text{Agent} \Box \Diamond \exists \text{activates.Switch}, \quad \text{Guard} \doteq \text{Agent} \Box \Box \exists \text{checks.Rail}.$$

Using modal operators over formulas, it is possible to express additional normative specifications. For example, stating that it is obligatory that a guard who detects some dangerous situation must alert a station agent, which in turn has the duty to alert an emergency service:

$$\Box (\text{Guard} \Box \exists \text{detects.DangerousSituation} \sqsubseteq \Box \exists \text{alerts.}(\text{StationAgent} \Box \Box \exists \text{alerts.EmergencyService})).$$

This flexibility in the application of modal operators allows us to assign different sets of duties to the agents involved, while still enforcing these statements as obligatory. To see this, compare the above definition of `Guard` with the case of a bystander that happens to detect a situation of danger while checking the rails. We could expect that it is obligatory for them to alert an emergency service, without requiring that they ought to alert a station manager to do so. Namely,

$$\Box (\exists \text{checks.Rail} \Box \exists \text{detects.DangerousSituation} \sqsubseteq \Box \exists \text{alerts.EmergencyService}).$$

Problems for Necessitation. Consider the CI, valid on every frame, that signallers either save a toddler, or they do not, i.e.,

$$\text{Signaller} \sqsubseteq \exists \text{saves.Toddler} \sqcup \neg \exists \text{saves.Toddler}.$$

By Theorem 1, we have that on R-frames this formula is also obligatory:

$$\Box (\text{Signaller} \sqsubseteq \exists \text{saves.Toddler} \sqcup \neg \exists \text{saves.Toddler}).$$

This conclusion violates what is known in the literature as the *principle of deontic contingency* [16], according to which what ought to be the case cannot be enforced within a deontic system by virtue of logical validity alone.

More in general, given a formula φ that is valid in a R-model, we always have that some logically implied formula ψ is obligatory, even when φ contains only factual statements describing the domain of interest. For instance, suppose that the following formula χ , stating that a toddler is a person, and that no person is a trolley, is valid in a R-model M :

$$\text{Toddler} \sqsubseteq \text{Person} \wedge \text{Person} \sqsubseteq \neg \text{Trolley}.$$

The formula χ specifies only some of the factual features of the trolley problem scenario, without any reference to normative notions, and it logically implies $\text{Toddler} \sqsubseteq \neg \text{Trolley}$. However, since χ is assumed to be valid in M , we are forced to conclude that also the latter is valid in M , and thus, due to Theorem 1, we have that $\Box(\text{Toddler} \sqsubseteq \neg \text{Trolley})$ is valid in M . Although true as a factual statement, it is not clear why it should be inferred that this ought to be the case.

Problems for Agglomeration. Given the following concept D describing a *moral dilemma*,

$$\Box \exists \text{activates.Switch} \sqcap \Box \neg \exists \text{activates.Switch},$$

(that is, the obligation to activate a switch and the obligation not to do it), by Theorem 1 we have that the following CI is valid on all R-frames:

$$D \sqsubseteq \Box(\exists \text{activates.Switch} \sqcap \neg \exists \text{activates.Switch}).$$

In other words, all objects incurring in a moral dilemma are also objects that are obligated to do something inconsistent. This issue is sometimes presented in the literature as the problem of *self-inconsistent obligations* for deontic agents [16].

Problems for Monotonicity. Since $\perp \sqsubseteq C$, for any ML_{ACC}^n concept C , is valid on R-frames, by Theorem 1 we have for instance that on R-frames it is valid

$$\Box(\exists \text{activates.Switch} \sqcap \neg \exists \text{activates.Switch}) \sqsubseteq \Box \exists \text{kills.Toddler}$$

Together with the CI discussed in the previous paragraph, we obtain that an object in the extension of a moral dilemma (such as the one described by D) is an object for which anything is obligatory, hence the names of *universal obligatoriness problem* [16], or *deontic explosion* [7].

Another problematic inference is known in the literature as the *Ross's paradox*. We have for instance that the following CI is valid on all R-frames:

$$\exists \text{saves.Toddler} \sqsubseteq \exists \text{saves.Toddler} \sqcup \exists \text{kills.Toddler}.$$

If the concept $\Box \exists \text{saves.Toddler}$, denoting the set of objects for which it is obligatory to save a toddler, is satisfiable, by Theorem 1 we obtain that the following concept is satisfiable as well:

$$\Box(\exists \text{saves.Toddler} \sqcup \exists \text{kills.Toddler}).$$

In other words, there can be individuals for which it is obligatory to save some toddlers or to kill them. However, it is not easy to explain why the obligation to save some toddlers should imply another obligation that can be fulfilled by killing some of them. Indeed, if it is not possible for an agent to fulfil the obligation to save some toddlers, it is still possible that they could partially attend to their duties by respecting other normative constraints. An obligation that can be fulfilled by killing some toddlers is a highly undesirable consequence that could not be used as a partial justification in case the former goal (to save a toddler) is not reachable. Therefore, it should not be derived in the normative system [16].

Another similar difficulty related to monotonicity is known as the *Good Samaritan paradox* [16]. Suppose that the deontic concept $\Box\exists\text{activates.Switch}$, denoting the set of entities for which it is obligatory to activate the switch, is satisfied in a R-model M , and that $\exists\text{activates.Switch} \sqsubseteq \exists\text{kills.Elderly}$, (meaning that if someone activates the switch, then they kill some elderly) is valid in that model. By Theorem 1, we have that also the following is valid in M :

$$\Box\exists\text{activates.Switch} \sqsubseteq \Box\exists\text{kills.Elderly}$$

Thus, the concept $\Box\exists\text{kills.Elderly}$ is satisfied in M , i.e., there is an object for which it is obligatory to kill an elderly. Although the killing of an elderly is a consequence implied by the activation of the switch, the obligation to do so is a consequence that a trustworthy moral agent should not be able to draw.

Example 2. In the N-model of Example 1, let $\text{kills}^{\mathcal{I}(w)} = \text{kills}^{\mathcal{I}(v)} = \{(d_1, d_3)\}$ and $\text{Elderly}^{\mathcal{I}(w)} = \text{Elderly}^{\mathcal{I}(v)} = \{d_3\}$. For all u of \mathcal{M} , $\mathcal{M}, u \models \exists\text{activates.Switch} \sqsubseteq \exists\text{kills.Elderly}$. For the concept $\Box\exists\text{kills.Elderly}$, we have $[\exists\text{kills.Elderly}]_{d_1}^{\mathcal{M}} = \{w, v\}$, and $[\exists\text{kills.Elderly}]_{d_3}^{\mathcal{M}} = \emptyset$. Since $\{w, v\} \notin \mathcal{N}(w)$, $(\Box\exists\text{kills.Elderly})^{\mathcal{I}(w)} = \emptyset$. Similarly, $\emptyset \notin \mathcal{N}(v)$, and so $(\Box\exists\text{kills.Elderly})^{\mathcal{I}(v)} = \emptyset$. Therefore, in particular, the concept $\Box\exists\text{activates.Switch}$ is satisfied in \mathcal{M} , and the formula $\exists\text{activates.Switch} \sqsubseteq \exists\text{kills.Elderly}$ is satisfied in all worlds of \mathcal{M} . However, the concept $\Box\exists\text{kills.Elderly}$ is *not* satisfied in \mathcal{M} . Hence, the Good Samaritan paradox does not apply.

4 Satisfiability in Non-normal Modal Description Logics

At the propositional level, logics \mathbf{E}^n and \mathbf{M}^n have both been used as a basis for weak deontic systems [1, 9] (although \mathbf{M}^n suffers from several problems discussed in Section 3), as well as to interpret praxeological operators, such as ‘agent i has the ability to bring about φ ’ [6, 20]. Moreover, \mathbf{M}^n has been combined with \mathcal{ALC} , as a basis for further coalition logic extensions of description logic languages [24, 23], and \mathbf{E}^n modal operators have been applied over \mathcal{ALC} axioms to formalise reasoning about agents’ intentions [12] (however, without establishing tight complexity results). In this section we study the complexity of the formula satisfiability problem in $\mathbf{E}_{\mathcal{ALC}}^n$ and $\mathbf{M}_{\mathcal{ALC}}^n$. This result is then lowered for fragments, denoted by $\mathbf{E}_{\mathcal{ALC}}^{n|g}$ and $\mathbf{M}_{\mathcal{ALC}}^{n|g}$, in which the modal operators are applied *globally*, i.e., over \mathcal{ALC} axioms only.

Satisfiability in $\mathbf{E}_{\mathcal{ALC}}^n$ and $\mathbf{M}_{\mathcal{ALC}}^n$. We provide a NEXPTIME upper bound by using a reduction, lifted from the propositional case, to multi-modal $\mathbf{K}_{\mathcal{ALC}}^m$. The translation \cdot^\dagger from $ML_{\mathcal{ALC}}^n$ to $ML_{\mathcal{ALC}}^{3n}$ is defined as follows [17, 15]:

$$\begin{aligned} A^\dagger &= A, & (\exists r.C)^\dagger &= \exists r.C^\dagger, & (C \sqsubseteq D)^\dagger &= C^\dagger \sqsubseteq D^\dagger, & (\vartheta)^\dagger &= \vartheta, \\ (\neg\gamma)^\dagger &= \neg\gamma^\dagger, & (\gamma \circ \delta)^\dagger &= \gamma^\dagger \circ \delta^\dagger, & (\Box_i\gamma)^\dagger &= \Diamond_{i_1}(\Box_{i_2}\gamma^\dagger \circ \Box_{i_3}\neg\gamma^\dagger) \end{aligned}$$

where $A \in \mathbf{N}_{\mathbf{C}}$, $r \in \mathbf{N}_{\mathbf{R}}$, ϑ is an assertion, γ and δ are both either $ML_{\mathcal{ALC}}^n$ concepts or formulas, and $\circ \in \{\sqcap, \wedge\}$. Using this translation, one can show that the satisfiability problem in N-frames is reducible to the formula satisfiability in the relational case [17, 15]. Since satisfiability in $\mathbf{K}_{\mathcal{ALC}}^{3n}$ is known to be NEXPTIME-complete [14, Theorem 15.15], we obtain the following complexity result.

Theorem 2. *Satisfiability in $\mathbf{E}_{\mathcal{ALC}}^n$ is in NEXPTIME.*

Proof (Sketch). Let φ be an $ML_{\mathcal{ALC}}^n$ formula s.t. $\mathcal{M}, w \models \varphi$, for some N-model $\mathcal{M} = (\mathcal{F}, \Delta, \mathcal{I})$ and some w in $\mathcal{F} = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$. We define an R-frame $F = (W, \{R_{i_j}\}_{i \in [1, n], j \in [1, 3]})$ and an $ML_{\mathcal{ALC}}^{3n}$ R-model $M = (F, \Delta, I)$ s.t.:

- $W = \{(w, 0) \mid w \in \mathcal{W}\} \cup \{(\alpha, 1) \mid \alpha \in \bigcup_{v \in \mathcal{W}} \mathcal{N}_i(v)\}$
- $R_{i_1} = \{((w, 0), (\alpha, 1)) \mid \alpha \in \mathcal{N}_i(w)\}$;
- $R_{i_2} = \{((\alpha, 1), (w, 0)) \mid w \in \alpha\}$
- $R_{i_3} = \{((\alpha, 1), (w, 0)) \mid w \notin \alpha\}$
- for every $(w, 0) \in W$, $I(w, 0) = \mathcal{I}(w)$; for every $(\alpha, 1) \in W$, $X^{I(\alpha, 1)} = \emptyset$, for all $X \in \mathbf{N}_{\mathbf{C}} \cup \mathbf{N}_{\mathbf{R}}$, and $a^{I(\alpha, 1)} = a^{\mathcal{I}}$, for all $a \in \mathbf{N}_{\mathbf{I}}$.

The pairs $(w, 0), (\alpha, 1)$ are used to ensure that W is the disjoint union of the sets of worlds w and subsets α of \mathcal{W} . By induction on concept and formulas occurring in φ , one can show that $M, (w, 0) \models \varphi^\dagger$. Conversely, given a $ML_{\mathcal{ALC}}^{3n}$ formula φ^\dagger s.t. $M, w \models \varphi^\dagger$, for some $ML_{\mathcal{ALC}}^{3n}$ R-model $M = (F, \Delta, I)$ based on $F = (W, \{R_{i_j}\}_{i \in [1, n], j \in [1, 3]})$, and some $w \in W$, we define a $ML_{\mathcal{ALC}}^n$ N-model $\mathcal{M} = (\mathcal{F}, \Delta, \mathcal{I})$ based on $\mathcal{F} = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$ s.t. $\mathcal{W} = W$, and for all $w \in W$:

- $\alpha \in \mathcal{N}_i(w)$ iff there is $v \in W$ s.t. $wR_{i_1}v$ and: (i) for all $u \in W$, $vR_{i_2}u \Rightarrow u \in \alpha$, and (ii) for all $u \in W$, $vR_{i_3}u \Rightarrow u \notin \alpha$;
- $\mathcal{I}(w) = I(w)$.

Again, by induction, we obtain that $\mathcal{M}, w \models \varphi$. □

The translation \cdot^\ddagger from $ML_{\mathcal{ALC}}^n$ to $ML_{\mathcal{ALC}}^{2n}$ is defined as \cdot^\dagger on all concepts and formulas, except from the modalised concepts or formulas γ [17, 15]:

$$(\Box_i\gamma)^\ddagger = \Diamond_{i_1}\Box_{i_2}\gamma^\ddagger.$$

We obtain an upper bound analogous to the one for $\mathbf{E}_{\mathcal{ALC}}^n$ by a reduction of the formula satisfiability problem for $\mathbf{M}_{\mathcal{ALC}}^n$ to the same problem for $\mathbf{K}_{\mathcal{ALC}}^{2n}$ [17, 15, 14].

Theorem 3. *Satisfiability in $\mathbf{M}_{\mathcal{ALC}}^n$ is in NEXPTIME.*

Proof (Sketch). The proof is similar to the one of Theorem 2. Given an N-model based on a supplemented N-frame satisfying an $ML_{\mathcal{ALC}}^n$ formula φ , we define an $ML_{\mathcal{ALC}}^{2n}$ R-model satisfying φ^\ddagger as above, by using relations R_{i_1} and R_{i_2} only. To prove the inductive step for modalised formulas $\Box_i\psi$ occurring in φ , we use the fact that, in N-models \mathcal{M} based on supplemented N-frames $\mathcal{F} = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$, $\mathcal{M}, w \models \Box_i\psi$ is equivalent to: there is $\alpha \in \mathcal{N}_i(w)$ s.t. $\alpha \subseteq [\psi]^\mathcal{M}$. Conversely, given a $ML_{\mathcal{ALC}}^{2n}$ R-model $M = (F, \Delta, I)$ based on $F = (W, \{R_{i_j}\}_{i \in [1, n], j \in [1, 2]})$ and satisfying φ^\ddagger , we define a $ML_{\mathcal{ALC}}^n$ N-model $\mathcal{M} = (\mathcal{F}, \Delta, \mathcal{I})$ based on $\mathcal{F} = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$ s.t. $\mathcal{W} = W$ and, for all $w \in W$: $\mathcal{I}(w) = I(w)$; $\alpha \in \mathcal{N}_i(w)$ iff there is $v \in W$ s.t. $wR_{i_1}v$ and for all $u \in W$, $vR_{i_2}u \Rightarrow u \in \alpha$. The N-frame \mathcal{F} so defined is supplemented: for all $w \in W$, if $\alpha \in \mathcal{N}_i(w)$ and $\alpha \subseteq \beta \subseteq W$, then there is $v \in W$ s.t. $wR_{i_1}v$ and for all $u \in W$, $vR_{i_2}u \Rightarrow u \in \beta$, i.e., $\beta \in \mathcal{N}_i(w)$. Moreover, by induction, we have that \mathcal{M} satisfies φ . \square

Satisfiability in $\mathbf{E}_{\mathcal{ALC}}^{n|g}$ and $\mathbf{M}_{\mathcal{ALC}}^{n|g}$. We now show tight complexity results for $\mathbf{E}_{\mathcal{ALC}}^{n|g}$ and $\mathbf{M}_{\mathcal{ALC}}^{n|g}$ using the notion of a propositional abstraction of a formula (as in, e.g., [4]). Here, one can separate the satisfiability test into two parts, one for the description logic dimension and one for the dimension of the neighbourhood frame. The *propositional abstraction* φ_{prop} of an $\mathbf{E}_{\mathcal{ALC}}^{n|g}$ formula φ is the result of replacing each \mathcal{ALC} atom in φ by a propositional variable, so that there is a 1 : 1 relationship between the \mathcal{ALC} atoms π occurring in φ and the propositional letters p_π used for the abstraction. The semantics of φ_{prop} is given in terms of propositional N-models $(\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]}, \mathcal{V})$ for \mathbf{E}^n , where $(\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]})$ is a N-frame and $\mathcal{V} : \mathbf{N}_P \rightarrow \mathcal{P}(\mathcal{W})$ is a function mapping propositional variables in \mathbf{N}_P to sets of worlds (see [9, 26]). We denote by $\mathbf{N}_P(\varphi)$ the set $\{p_\pi \in \mathbf{N}_P \mid \pi \text{ is an } \mathcal{ALC} \text{ atom in } \varphi\}$. Given an $\mathbf{E}_{\mathcal{ALC}}^{n|g}$ formula φ , we say that a propositional N-model $\mathcal{M}^P = (\mathcal{W}, \{\mathcal{N}_i\}_{i \in [1, n]}, \mathcal{V})$ of φ_{prop} is φ -consistent if, for all $w \in \mathcal{W}$, the following \mathcal{ALC} formula is satisfiable

$$\bigwedge_{p_\pi \in \mathbf{N}_P(w)} \pi \wedge \bigwedge_{p_\pi \in \overline{\mathbf{N}_P(w)}} \neg\pi$$

where $\mathbf{N}_P(w) = \{p_\pi \in \mathbf{N}_P(\varphi) \mid w \in \mathcal{V}(p_\pi)\}$ and $\overline{\mathbf{N}_P(w)} = \mathbf{N}_P(\varphi) \setminus \mathbf{N}_P(w)$. We now formalise the connection between $\mathbf{E}_{\mathcal{ALC}}^{n|g}$ formulas and their propositional abstractions with the following lemma.

Lemma 1. *A formula φ is $\mathbf{E}_{\mathcal{ALC}}^{n|g}$ satisfiable iff φ_{prop} is satisfied in a φ -consistent \mathbf{E}^n model.*

We assume that the primitive connectives used to build propositional formulas are \neg and \wedge (\vee is expressed using \neg and \wedge), and we denote by $\text{sub}(\varphi_{\text{prop}})$ the set of subformulas of φ_{prop} closed under single negation. A *valuation* for a propositional ML formula φ_{prop} is a function $\nu : \text{sub}(\varphi_{\text{prop}}) \rightarrow \{0, 1\}$ that satisfies the following conditions: (1) for all $\neg\psi \in \text{sub}(\varphi_{\text{prop}})$, $\nu(\psi) = 1$ iff $\nu(\neg\psi) = 0$; (2) for all $\psi_1 \wedge \psi_2 \in \text{sub}(\varphi_{\text{prop}})$, $\nu(\psi_1 \wedge \psi_2) = 1$ iff $\nu(\psi_1) = 1$ and $\nu(\psi_2) = 1$; and (3) $\nu(\varphi_{\text{prop}}) = 1$. We say that a valuation for φ_{prop} is φ -consistent if any N-model of the form $(\{w\}, \{\mathcal{N}_i\}_{i \in [1, n]}, \mathcal{V})$ satisfying $w \in \mathcal{V}(p_\pi)$ iff $\nu(p_\pi) = 1$, for

all $p_\pi \in \mathbf{Np}(\varphi)$, is φ -consistent. We now establish that satisfiability of φ_{prop} in a φ -consistent model is characterized by the existence of a φ -consistent valuation satisfying the property described in Lemma 2.

Lemma 2. *A formula φ_{prop} is satisfied in a φ -consistent \mathbf{E}^n model iff there is a φ -consistent valuation ν for φ_{prop} such that if $\Box_i\psi_1$ and $\Box_i\psi_2$ are in $\text{sub}(\varphi_{\text{prop}})$, $\nu(\Box_i\psi_1) = 1$, and $\nu(\Box_i\psi_2) = 0$, then $(\psi_1 \wedge \neg\psi_2) \vee (\neg\psi_1 \wedge \psi_2)$ is satisfied in a φ -consistent \mathbf{E}^n model.*

To determine satisfiability of φ_{prop} in a φ -consistent model, we use Lemma 2 and the fact that there are only quadratically many formulas of the form $\psi_1 \wedge \neg\psi_2$, where ψ_1 and ψ_2 are subformulas of φ_{prop} . We observe that satisfiability in \mathcal{ALC} is EXPTIME-complete [14] and so, one can determine in exponential time whether a valuation is φ -consistent. For an EXPTIME upper bound, one can deterministically compute all possible φ -consistent valuations for $\psi_1 \wedge \neg\psi_2$ and decide satisfiability of φ_{prop} by a φ -consistent model using a bottom-up strategy (as in [26]). As satisfiability in \mathcal{ALC} is EXPTIME-hard our upper bound is tight.

Theorem 4. *Satisfiability in $\mathbf{E}_{\mathcal{ALC}}^{n|\mathbf{g}}$ is EXPTIME-complete.*

Regarding the proof for $\mathbf{M}_{\mathcal{ALC}}^{n|\mathbf{g}}$, we first point out that Lemma 1 can be easily adapted to $\mathbf{M}_{\mathcal{ALC}}^{n|\mathbf{g}}$. The proof for our EXPTIME result for $\mathbf{M}_{\mathcal{ALC}}^{n|\mathbf{g}}$ is analogous to the one given for $\mathbf{E}_{\mathcal{ALC}}^{n|\mathbf{g}}$, except that here we use a variant of Lemma 2 tailored for \mathbf{M}^n (see Proposition 3.8 in [26]). Thus, we obtain also the following result.

Theorem 5. *Satisfiability in $\mathbf{M}_{\mathcal{ALC}}^{n|\mathbf{g}}$ is EXPTIME-complete.*

5 Conclusion

We have studied non-normal MDLs based on neighbourhood models, showing how to express normative specifications in a deontic scenario, and highlighting the differences with relational semantics. We have established complexity results for the satisfiability problem in the non-normal MDLs $\mathbf{E}_{\mathcal{ALC}}^n$ and $\mathbf{M}_{\mathcal{ALC}}^n$, and their respective fragments $\mathbf{E}_{\mathcal{ALC}}^{n|\mathbf{g}}$ and $\mathbf{M}_{\mathcal{ALC}}^{n|\mathbf{g}}$, with modal operators applied only over the description logic axioms. These logics represent the basis for further extensions in deontic, epistemic, and dynamic contexts. As future work, we plan to study logics interpreted over suitably constrained neighbourhood frames, so to provide additional reasoning capabilities in multi-agent systems. MDLs extended with non-normal dyadic modal operators, to express conditional obligations or beliefs, such as ‘it is obligatory/believed that φ , given ψ ’, represent another direction for further research.

References

1. Anglberger, A.J., Gratzl, N., Roy, O.: Obligation, free choice, and the logic of weakest permissions. *Rev. Symb. Logic* **8**(4), 807–827 (2015)
2. Arkoudas, K., Bringsjord, S., Bello, P.: Toward ethical robots via mechanized deontic logic. In: *AAAI fall symposium on machine ethics*. pp. 17–23 (2005)
3. Arló-Costa, H.L., Pacuit, E.: First-order classical modal logic. *Stud. Logica* **84**(2), 171–210 (2006)
4. Baader, F., Ghilardi, S., Lutz, C.: LTL over description logic axioms. *ACM T. Comput. Logic* **13**(3), 21:1–21:32 (2012)
5. Boer, A., Winkels, R., Vitali, F.: Metalex XML and the legal knowledge interchange format. In: *Computable Models of the Law*, pp. 21–41 (2008)
6. Brown, M.A.: On the logic of ability. *J. Philos. Logic* **17**(1), 1–26 (1988)
7. Calardo, E., Rotolo, A.: Quantification in some non-normal modal logics. *J. Philos. Logic* **46**(5), 541–576 (2017)
8. Casellas, N.: *Legal Ontology Engineering*. Springer (2011)
9. Chellas, B.F.: *Modal logic: an introduction*. Cambridge University Press (1980)
10. Dalmonte, T., Olivetti, N., Negri, S.: Non-normal modal logics: Bi-neighbourhood semantics and its labelled calculi. In: *AiML*. pp. 159–178 (2018)
11. Dennis, L.A., Fisher, M., Slavkovik, M., Webster, M.: Formal verification of ethical choices in autonomous systems. *Robot. Auton. Syst.* **77**, 1–14 (2016)
12. Erdur, R.C., Seylan, I.: The design of a semantic web compatible content language for agent communication. *Expert Syst.* **25**(3), 268–294 (2008)
13. Francesconi, E.: A description logic framework for advanced accessing and reasoning over normative provisions. *Artif. Intell. Law* **22**(3), 291–311 (2014)
14. Gabbay, D.M., Kurucz, A., Wolter, F., Zakharyashev, M.: *Many-dimensional Modal Logics: Theory and Applications*. Elsevier (2003)
15. Gasquet, O., Herzig, A.: From classical to normal modal logics. In: *Proof theory of modal logic*, pp. 293–311. Springer (1996)
16. Hansson, S.O.: Alternative semantics for deontic logic. In: *Handbook of Deontic Logic and Normative Systems*, pp. 445–497. College Publication (2013)
17. Kracht, M., Wolter, F.: Normal monomodal logics can simulate all others. *J. Symbolic Logic* **64**(1), 99–138 (1999)
18. Lam, H., Hashmi, M.: Enabling reasoning with LegalRuleML. *TPLP* **19**(1), 1–26 (2019)
19. Meyer, J.C., Wieringa, R.J., Dignum, F.: The role of deontic logic in the specification of information systems. In: *Logics for Databases and Information Systems*. pp. 71–115. Kluwer (1998)
20. Pacuit, E.: *Neighborhood Semantics for Modal Logic*. Springer (2017)
21. Pigozzi, G., van der Torre, L.W.N.: Multiagent deontic logic and its challenges from a normative systems perspective. *FLAP* **4**(9), 2929–2993 (2017)
22. Seylan, I., Erdur, R.C.: A tableau decision procedure for \mathcal{ALC} with monotonic modal operators and constant domains. *ENTCS* **231**, 113–130 (2009)
23. Seylan, I., Jamroga, W.: Coalition description logic with individuals. *ENTCS* **262**, 231–248 (2010)
24. Seylan, I., Jamroga, W.: Description logic for coalitions. In: *AAMAS*. pp. 425–432 (2009)
25. Thomson, J.J.: The Trolley Problem. *Yale LJ* **94**(6), 1395–1415 (1985)
26. Vardi, M.Y.: On the complexity of epistemic reasoning. In: *LICS*. pp. 243–252 (1989)
27. Waagbø, G.: Quantified modal logic with neighborhood semantics. *Math. Logic Quart.* **38**(1), 491–499 (1992)