How Knowledge Triggers Obligation

Grossi, Davide; Kooi, Barteld; Su, Xingchi; Verbrugge, Rineke

Published in:
Logic, Rationality, and Interaction - 8th International Workshop, LORI 2021, Proceedings

DOI:
10.1007/978-3-030-88708-7_17

How Knowledge Triggers Obligation

A Dynamic Logic of Epistemic Conditional Obligation

Davide Grossi[1,2], Barteld Kooi[1], Xingchi Su[1(✉)], and Rineke Verbrugge[1]

[1] University of Groningen, Groningen, The Netherlands
{d.grossi,b.p.kooi,x.su,l.c.verbrugge}@rug.nl
[2] University of Amsterdam, Amsterdam, The Netherlands

Abstract. Obligations can be affected by knowledge. Several approaches exist to formalize knowledge-based obligations, but no formalism has been developed yet to capture the *dynamic* interaction between knowledge and obligations. We introduce the dynamic extension of an existing logic for knowledge-based obligations here. We motivate the logic by analyzing several scenarios and by showing how it can capture in an original manner several fundamental deontic notions such as absolute, *prima facie* and all-things-considered obligations. Finally, in the dynamic epistemic logic tradition, we provide reduction axioms for the dynamic operator of the new logic.

Keywords: Epistemic conditional obligation · Priority structure · Action model · Reduction axiom · Kangerian-Andersonian reduction

1 Introduction

Epistemic conditional obligations are a type of conditional obligations where the consequent is triggered by the knowledge of the antecedent. For example, a doctor has an obligation to treat a patient only if she *knows* that the man is ill. The classical approach to conditional obligation is based on Hansson's preference-based models [1], where the semantics of $\bigcirc(\varphi|\psi)$ is given as: the best ψ-states are φ-states. A large literature has developed out of this approach (e.g., [2–4]). More recently, a static logic of epistemic conditional obligations ($\mathbb{KCDL}$) has been presented in [5], which defines operators $\odot_i(\varphi|\psi)$ as: the best ψ-states that are *epistemically indistinguishable for agent* i also satisfy φ.

In this paper, priority structures are introduced as norms that remain *static* throughout. Accordingly, we build on the logic of epistemic conditional obligations to study the *dynamic* process whereby the acquisition of new information, or the change of factual circumstances, triggers changes of obligations. To do so we introduce a dynamic operator formalizing obligation change, and show how the new logic can systematize some fundamental deontic notions. The proposed logic is motivated by the following scenarios, among which Scenario 1 is taken from [6] and Scenario 2 is a variant of an example from [7].

S. Ghosh and T. Icard (Eds.): LORI 2021, LNCS 13039, pp. 201–215, 2021.
https://doi.org/10.1007/978-3-030-88708-7_17

Scenario 1. Uma is a doctor whose neighbour Sam is ill. And Sam is a patient at Uma's practice. But Uma does not know that Sam is ill. We intuitively think that Uma has no obligation to treat her neighbour. Then Sam's daughter Ann shouts loudly on the street that "My dad is ill, any help please?" Now Uma knows that Sam is ill and has an obligation to treat Sam.

Scenario 2. One coin is tossed and covered by a cup. Fumio and Chiyo have an obligation to bet correctly (if the coin lands heads up and they bet on heads, or if the coin lands tails up and they bet on tails). Chiyo then lifts the cup, looks at the coin and ensures that the coin is heads up by some sleight of hand. Fumio observes Chiyo looking at the coin, and he considers it possible that Chiyo has flipped the coin. So before Chiyo looks at the coin, they do not have obligations to bet on heads (or on tails). After Chiyo flips the coin, Chiyo has an obligation to bet on heads but Fumio still does not have an obligation to bet on heads.

Scenario 3. Driss promised to his friend that he will go to the party on time. But when he is on the way to the party, he sees a car accident happening. Now, Driss ought to call an ambulance and help the people in the car, although it could make himself be late for the party. Driss has a new obligation to call an ambulance which overrides the obligation to keep the promise.

Paper Outline. The technical background is introduced in Sect. 2. Section 3 gives the language and semantics for the dynamic epistemic conditional obligation, which is then used for modelling the scenarios mentioned above. Section 4 uses our logic to provide novel formalizations of important deontic notions such as absolute, *prima facie* and all-things-considered obligations, as well as of a new type of obligation, which we call safe obligations. In developing these notions we highlight how the existing body of theory on conditional belief dynamics in Dynamic Epistemic Logic [8] bears significance for the understanding of deontic conditionals and their dynamics. Finally, Sect. 5 provides a sound and complete axiom system for logic $\mathbb{DKCDL}$, based on standard reduction axioms and the Kangerian-Andersonian reduction of deontic operators [9,10]. One proof is omitted and one is only sketched for space reasons.

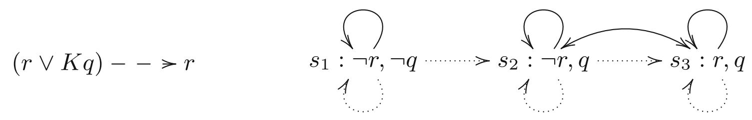
2 Preliminaries

We will introduce the static logic of epistemic conditional obligations $\mathbb{KCDL}$ from [5]. The language is $\mathcal{L}_{\mathrm{KCDL}}$. The semantic apparatus comprises epistemic betterness structures which contain a betterness relation, comparing states by deontic ideality, and epistemic relations for each agent. Let $\mathbf{P}$ be a countable set of propositional atoms and let $G = \{1, \cdots, n\}$ be a finite set of agents.

Definition 1. *The language $\mathcal{L}_{\mathrm{KCDL}}$ is given by the following BNF:*

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid \bigodot_i(\varphi|\varphi)$$

where $p \in \mathbf{P}$ and $i \in G$.



(a) $\mathcal{G}_k$ (b) $M_{\mathcal{G}k}$

Fig. 1. Two examples

Intuitively, $K_i\varphi$ stands for "agent i knows that φ"; $\bigodot_i(\varphi|\psi)$ stands for "if i knows that ψ, then φ ought to be the case". The language of epistemic logic is $\mathcal{L}_{\mathrm{EL}}$, which is identical to $\mathcal{L}_{\mathrm{KCDL}}$ but without the dyadic operator $\bigodot_i(\_|\_)$.

Definition 2 (Epistemic Betterness Structures [5]). $M = \langle S, \sim_1, \cdots, \sim_n, \leqslant, V\rangle$ *is an epistemic betterness structure where* S *is the set of states,* $\sim_i$: $S \times S$ *is the epistemic relation for agent i (equivalence relation),* $\leqslant$: $S \times S$ *is a betterness relation (total pre-order) and* $V : \mathbf{P} \rightarrow \mathcal{P}(\mathbf{S})$ *is the valuation function over* S*. Let* $[s]^{\sim_i}$ *denote the set of states accessible from* s *by the epistemic relation* $\sim_i$*.*

Definition 3 (Semantics of $\mathcal{L}_{\mathrm{KCDL}}$ [5]). *The truth conditions of formulas can be defined over* M *as follows (only the non-trivial cases are shown):*

- $M, s \models K_i\varphi$ *iff* $[s]^{\sim_i} \subseteq \|\varphi\|_M$*;*
- $M, s \models \bigodot_i(\varphi|\psi)$ *iff* $\max_{\leqslant}([s]^{\sim_i} \cap \|\psi\|_M) \subseteq \|\varphi\|_M$

where

- $\|\varphi\|_M = \{s \in M \mid M, s \models \varphi\}$,
- $\max_{\leqslant} S = \{s \in S \mid \forall t \in S\ (s \leqslant t \Rightarrow t \leqslant s)\}$.

Observe that for non-empty $\|\psi\|_M$, the set $\max_{\leqslant}([s]^{\sim_i} \cap \|\psi\|_M)$ is non-empty, which makes the semantics of $\bigodot_i(\_|\_)$ well-defined.

Definition 4 (Epistemic Models [11]). *An epistemic model* $M_E = \langle S, \sim_1, \sim_2, \cdots, \sim_n, V\rangle$ *is an epistemic betterness structure without the betterness relation.*

An epistemic betterness structure is also an epistemic model extended with a betterness relation $\leqslant$ over the set of states. In the study on $\mathbb{KCDL}$ [5], betterness relations between states are given *a priori*. In this paper, *priority structures* will be introduced for ordering states. Priority structures were originally introduced in [12]. The domain of a priority structure is a finite set of relevant formulas. In this paper, we define it within $\mathcal{L}_{\mathrm{EL}}$-formulas. Priority structures enable us to define the betterness relations between states according to the $\mathcal{L}_{\mathrm{EL}}$-formulas that are satisfied on the states.

Definition 5 ($\mathcal{L}_{\mathrm{EL}}$-Priority Structures). *Given the language of the classical epistemic logic* $\mathcal{L}_{\mathrm{EL}}$*, an* $\mathcal{L}_{\mathrm{EL}}$*-priority structure is a tuple* $\mathcal{G} = \langle \Phi, \prec\rangle$ *such that:*

– *$\Phi \subset \mathcal{L}_{\text{EL}}$ and Φ is finite;*
– *$\prec$ is a strict order on Φ such that for all formulas $\varphi, \psi \in \Phi$, it holds that: if $\varphi \prec \psi$, then ψ logically implies φ.*

An example of an $\mathcal{L}_{\text{EL}}$-priority structure is shown in Fig. 1a, where a one-way dashed arrow from φ to ψ denotes $\varphi \prec \psi$.

A priority structure supplies a criterion for assessing the relative ideality of states. Given an $\mathcal{L}_{\text{EL}}$-priority structure, a betterness relation can be derived from a domain of an epistemic model. In this way, priority structures serve a similar purpose to norms in [13]. In this paper, we follow the approach of [4] to obtain betterness relations from priority structures.

Definition 6 (Epistemic Betterness Structures Based on Priority Structures). *Given an $\mathcal{L}_{\text{EL}}$-priority structure $\mathcal{G} = \langle \Phi, \prec \rangle$ and an epistemic model $M_E = \langle S, \sim_1, \cdots, \sim_n, V \rangle$, the structure $M = \langle S, \sim_1, \cdots, \sim_n, \leqslant_{\mathcal{G}}, V \rangle$ is an epistemic betterness structure based on $\mathcal{G}$ if M is M_E extended with the betterness relation $\leqslant_{\mathcal{G}}$, where $\leqslant_{\mathcal{G}}$ is defined as follows, for any two states s, $s' \in S$:*

$$s \leqslant_{\mathcal{G}} s' \iff \forall \varphi \in \Phi : s \in \|\varphi\|_{M_E} \Rightarrow s' \in \|\varphi\|_{M_E}$$

In other words, when an epistemic model M_E and an $\mathcal{L}_{\text{EL}}$-priority structure $\mathcal{G}$ are provided, we can construct an epistemic betterness structure by adding a betterness relation based on $\mathcal{G}$ to M_E. An example of an epistemic betterness structure based on $\mathcal{G}_k$ is shown as $M_{\mathcal{G}_k}$ in Fig. 1b, where a directed dotted arrow from s_i to s_j denotes $s_i \leqslant s_j$. According to $\mathcal{G}_k$, the state satisfying r is the best. So s_3 is the best. The state satisfying $r \vee Kq$ is better than those not satisfying it. So s_2 is better than s_1. Since an *epistemic betterness structure based on a priority structure* is an epistemic betterness structure, we will also call them just *epistemic betterness structures* in the following parts.

3 Dynamic Epistemic Conditional Obligation

In this section, we intend to establish a dynamic extension to $\mathbb{KCDL}$. *Action models*, originally introduced in dynamic epistemic logic (DEL), can characterize not only the information changes, but also the factual changes (truth value of the propositions).

Definition 7 (Action Models [14]). *An action model for a language $\mathcal{L}$ is a structure $U = \langle E, R_1, R_2, \cdots, R_n, pre, post \rangle$ where*

– *E is a finite non-empty set of events;*
– *for each $i \in G$, $R_i : E \times E$ is i's indistinguishability relation between events;*
– *$pre : E \rightarrow \mathcal{L}$ assigns to each event a precondition;*
– *$post : E \rightarrow (\mathbf{P} \rightarrow \mathcal{L})$ assigns to each event a postcondition for each atom. Each $post(e)$ function is required to change truth values of only finitely many propositions.*

For each $e \in E$, (U, e) is called a pointed action model.

In dynamic epistemic logic [15], action models operate on epistemic models, leading to new epistemic models.

Definition 8 (Updated Epistemic Models [14]). *Given an epistemic model* $M_E = \langle S, \sim_1, \cdots, \sim_n, V\rangle$ *and an action model* $U = \langle E, R_1, \cdots, R_n, pre, post\rangle$, *the result of executing* U *in* M_E *is the model* $M_E \otimes U = \langle S', \sim'_1, \cdots, \sim'_n, V'\rangle$:

- $S' = \{(s, e) \mid s \in S, e \in E$ *and* $M_E, s \models pre(e)\}$;
- *for each* $i \in G$, $\sim'_i = \{((s, e), (t, f)) \mid (s, e), (t, f) \in S', (s, t) \in \sim_i, (e, f) \in R_i\}$;
- $V'(p) = \{(s, e) \mid M_E, s \models post(e)(p)\}$.

3.1 Language and Semantics of $\mathcal{L}_{\mathrm{DKCDL}}$

Definition 9. *The language* $\mathcal{L}_{\mathrm{DKCDL}}$ *is given by the following BNF:*

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_i\varphi \mid \bigodot_i(\varphi|\varphi) \mid [(U, e)]\varphi,$$

where $p \in \mathbf{P}$, $i \in G$, *and* (U, e) *is a pointed action model.*

The language of dynamic epistemic logic is $\mathcal{L}_{\mathrm{DEL}}$, which is identical to $\mathcal{L}_{\mathrm{DKCDL}}$ but without the dyadic operator $\bigodot_i(\_|\_)$. Subsequently, we provide the semantics of the formula $[(U, e)]\varphi$ over epistemic betterness structures based on priority structures. Firstly, we need to define epistemic betterness structures updated by action models.

Definition 10 (Updated Epistemic Betterness Structures). *Given an epistemic model* $M_E = \langle S, \sim_1, \cdots, \sim_n, V\rangle$ *and a priority structure* $\mathcal{G} = \langle \Phi, \prec\rangle$, *the structure* $M = \langle S, \sim_1, \cdots, \sim_n, \leqslant_{\mathcal{G}}, V\rangle$ *is the epistemic betterness structure based on* $\mathcal{G}$. *Letting* $U = \langle E, R_1, \cdots, R_n, pre, post\rangle$ *be an action model, the result of executing* U *in* M *is the model* $M \otimes U = \langle S', \sim'_1, \cdots, \sim'_n, \leqslant', V'\rangle$ *where:*

- $\langle S', \sim'_1, \cdots, \sim'_n, V'\rangle = M_E \otimes U$;
- $\leqslant' = \{((s, e), (t, f)) \in S' \times S' \mid \forall\varphi \in \Phi : (s, e) \in \|\varphi\|_{M_E \otimes U} \Rightarrow (t, f) \in \|\varphi\|_{M_E \otimes U}\}$.

An updated epistemic betterness structure consists of its corresponding updated epistemic model and an updated betterness relation. The new betterness relation *re-orders* these new states based on the priority structure. Now we can give the truth condition of the formula $[(U, e)]\varphi$.

Definition 11. *The truth conditions of atoms, Boolean formulas, epistemic formulas and dyadic conditional obligations are identical to* $\mathbb{KCDL}$. *Let* M *be an arbitrary epistemic betterness structure based on priority structure* $\mathcal{G}$.

- $M, s \models [(U, e)]\varphi$ *iff* $M, s \models pre(e)$ *implies* $M \otimes U, (s, e) \models \varphi$.

3.2 Analysis of Scenarios 1, 2 and 3

We are now in a position to formalize how obligations change in response to information and factual changes. For each scenario, a priority structure is given in advance, which remains unchanged throughout the story. It specifies the betterness relations in both the initial and updated epistemic betterness structures. The information changes and factual changes are characterized by action models. After performing an action, the updated epistemic betterness structure will determine the agents' new obligations. In the following models, the transitive and reflexive closures of all types of relations are omitted in the figures.

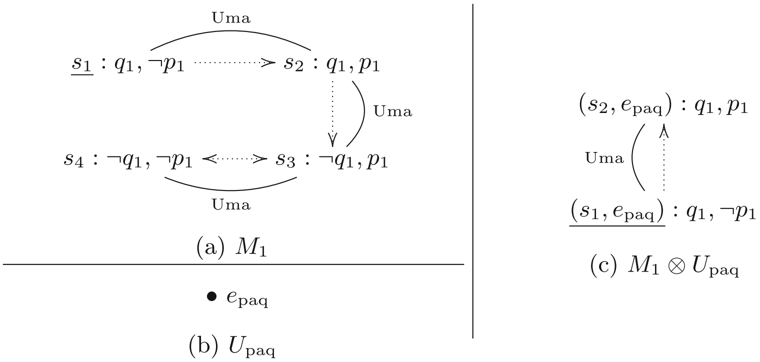


Fig. 2. Scenario 1: Ann is shouting loudly

Scenario 1: New Information Triggers Obligations. In Fig. 2, q_1 refers to 'Sam is ill' and p_1 refers to 'Sam is treated'. We first give the priority structure $\mathcal{G}_1$ for scenario 1. 'Sam is not ill' ($\neg q_1$) is always the best state of affairs and 'if Sam is ill, then Sam is treated' ($\neg q_1 \vee p_1$) is better than those cases where 'Sam is ill but Sam is not treated' ($q_1 \wedge \neg p_1$).

Accordingly, the initial epistemic betterness structure based on $\mathcal{G}_1$ is M_1 (see Fig. 2a). Over M_1, we have $M_1, s_1 \models \neg \bigodot_{\text{Uma}}(p_1|\top) \wedge \neg K_{\text{Uma}} q_1$, which means that Uma does not know that Sam is ill and does not have an obligation to see to it that Sam is treated. Then, Sam's daughter shouts loudly outside that her dad is ill. This action can be modeled by an action model of truthful public announcements, i.e., $(U_{\text{paq}}, e_{\text{paq}})$ (see Fig. 2b, 'paq' refers to 'public announcement that q_1'). An action model of truthful public announcement that φ is a singleton action model where the precondition equals to φ and the postconditions for all propositions are *id*. It consequently eliminates all $\neg\varphi$-states and keeps φ-states. So $pre(e_{\text{paq}}) = q_1$ and postconditions for all propositions on e_{paq} are *id*.

Thereafter, shown as Fig. 2c, the updated epistemic betterness structure $M_1 \otimes U_{\text{paq}}$ only contains the two states (s_1, e_{paq}) and (s_2, e_{paq}). We have $M_1 \otimes U_{\text{paq}}, (s_1, e_{\text{paq}}) \models \bigodot_{\text{Uma}}(p_1|\top)$, which means that Uma has an obligation

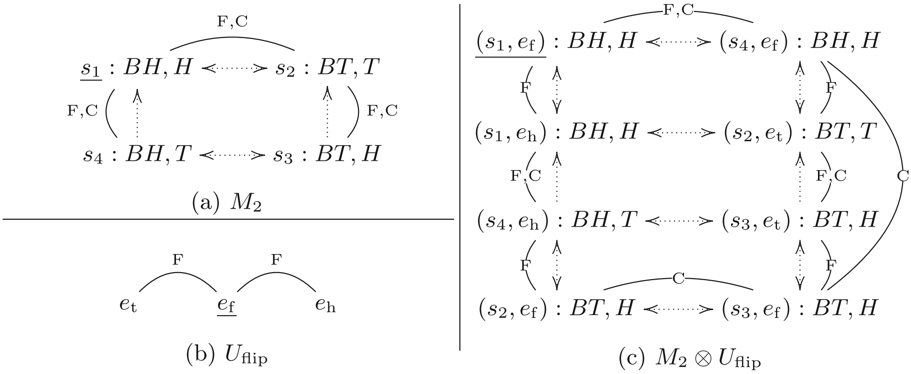


Fig. 3. Scenario 2: Chiyo ensures that the coin lands heads up

to see to it that Sam is treated. Therefore, we have $M_1, s_1 \models \bigodot_{\text{Uma}}(p_1|q_1) \rightarrow [(U_{\text{paq}}, e_{\text{paq}})] \bigodot_{\text{Uma}}(p_1|\top)$.

Scenario 2: Factual Change Triggers Obligations. In Figure 3, T refers to 'the coin lands tails up', H refers to 'the coin lands heads up', BT refers to 'betting on tails', and BH refers to 'betting on heads'. First, we give the priority structure $\mathcal{G}_2$ for Scenario 2. The best state of affairs is betting correctly $((BH \wedge H) \vee (BT \wedge T))$. Any other cases are worse. Let F denote Fumio and let C denote Chiyo.

The initial epistemic betterness structure based on $\mathcal{G}_2$ is M_2 shown as Fig. 3a. Over M_2, we have $M_2, s_1 \models \neg \bigodot_{\text{F}}(BT|\top) \wedge \neg \bigodot_{\text{F}}(BH|\top) \wedge \neg \bigodot_{\text{C}}(BT|\top) \wedge \neg \bigodot_{\text{C}}(BH|\top)$ since they cannot see the coin.

In Fig. 3b, the action model U_{flip} describes the case where Chiyo sees the coin and ensures that the coin lands heads up but Fumio cannot see Chiyo's action. Event e_{t} represents that Chiyo sees the coin is tails up but does not flip. Event e_{h} represents that Chiyo sees the coin is heads up but does not flip. Event e_{f} represents that Chiyo ensures that the coin lands heads up no matter whether it was heads up or tails up. The preconditions are $pre(e_{\text{t}}) = T$, $pre(e_{\text{h}}) = H$, and $pre(e_{\text{f}}) = \top$. The postconditions of e_{t} and e_{h} are id. The postconditions of e_{f} are $post(e_{\text{f}})(H) = \top$ and $post(e_{\text{f}})(T) = \bot$.

After Chiyo performs the action, Chiyo has an obligation to bet on heads. Since Fumio does not know whether Chiyo flips the coin, Fumio still does not have an obligation to bet on heads (or on tails). These new obligations can be shown over the updated epistemic betterness structure $M_2 \otimes U_{\text{flip}}$ (see Fig. 3c). We have $M_2 \otimes U_{\text{flip}}, (s_1, e_{\text{f}}) \models \bigodot_{\text{C}}(BH|\top) \wedge \neg \bigodot_{\text{F}}(BH|\top) \wedge \neg K_{\text{F}} \bigodot_{\text{C}}(BH|\top)$.

Scenario 3: Unconditional Obligations Are Defeasible. In Fig. 4, k refers to 'Driss keeps promise', a refers to 'a car accident happens', and s refers to 'Driss saves the people involved in the accident'. The priority structure $\mathcal{G}_3$ for Scenario 3 would have the best state of affairs to be those where there is no car accident and Driss keeps the promise $(\neg a \wedge k)$. The second best case is that no

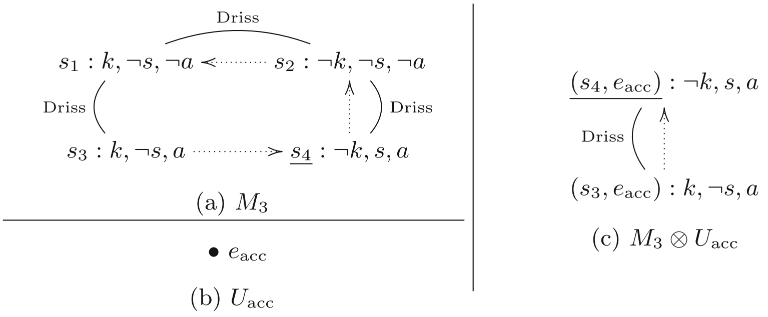


Fig. 4. Scenario 3: A car accident is happening

accident happens ($\neg a$). The third best case is that if an accident happens, then Driss saves the people ($\neg a \vee s$). Other cases are the worst.

Accordingly, we assume that there are only four possible situations in the initial epistemic betterness structure based on $\mathcal{G}_3$ (shown as M_3, Fig. 4a). We have $M_3, s_4 \models \bigodot_{\text{Driss}}(k|\top) \wedge \neg \bigodot_{\text{Driss}}(s|\top)$, which means that Driss ought to keep his promise unconditionally at that moment, but does not have an unconditional obligation to save the people.

The action model $(U_{\text{acc}}, e_{\text{acc}})$ (Fig. 4b) represents the event of the car accident, where $pre(e_{\text{acc}}) = \neg a$ and $post(e_{\text{acc}})(a) = \top$. In the updated epistemic betterness structure $M_3 \otimes U_{\text{acc}}$ (Fig. 4c), we have $M_3 \otimes U_{\text{acc}}, (s_4, e_{\text{acc}}) \models \bigodot_{\text{Driss}}(s|\top) \wedge \neg \bigodot_{\text{Driss}}(k|\top)$, which means that, after seeing the car accident, Driss's unconditional obligation to keep his promise is *overridden* by another unconditional obligation, namely, saving people.

By our analysis on Scenario 3, we would say that epistemic conditional obligations are defeasible. The defeasibility of $\bigodot_i(\_|\_)$ appears in two different aspects. In the static logic of epistemic conditional obligation $\mathbb{KCDL}$ [5], it invalidates the formula $\bigodot_i(\varphi|\psi) \to \bigodot_i(\varphi|\psi \wedge \chi)$, which means that a stronger condition could override the old obligation [16,17]. In the dynamic extension shown in the current paper, the formula $\bigodot_i(\varphi|\top) \to [(U,e)]\neg \bigodot_i(\varphi|\top)$ is satisfiable, which means that even an unconditional obligation could be released after taking some action. All the unconditional obligations formalized in Scenario 3 can be denoted by *prima facie* obligations, a notion that is strongly related to defeasibility. We will discuss these notions in the following section.

4 Information and Knowledge-Based Obligation

In the context of conditional beliefs, van Benthem comments that "conditional beliefs *pre-encode* beliefs that we would have if we learnt certain things" [18]. Baltag and Smets state that "conditional beliefs give descriptions of the agent's *plan* about what he will believe ... after receiving new information" [8]. Similarly, we take the view that conditional obligations pre-encode what states of

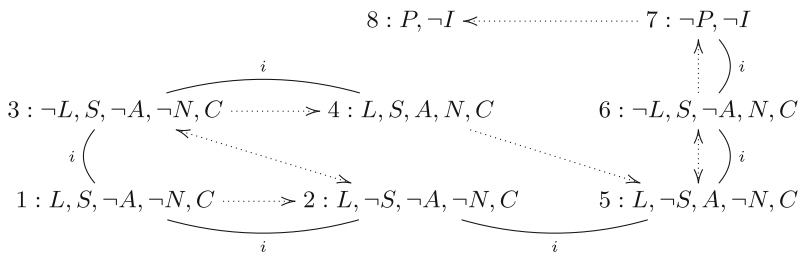


Fig. 5. M (a president is facing a world war)

affairs would be the best if specific facts were to hold. In deontic logic terminology, conditional obligations pre-encode the so-called factual detachment of obligations [2,19]. And, continuing the above analogy, an epistemic conditional obligation pre-encodes what can be referred to as *epistemic detachment*: $\mathbb{KCDL} \models (\bigodot_i(\varphi|\psi) \wedge K_i\psi) \rightarrow \bigodot_i(\varphi|\top)$. An unconditional obligation follows from an epistemic conditional obligation and the knowledge of the antecedent [5].

We will be using a running example to show how the above intuitions lead to natural formalizations, within language $\mathcal{L}_{\mathrm{DKCDL}}$, of several philosophical notions concerning obligations. In Fig. 5, $M = \langle W, \sim_i, \leqslant, V\rangle$ where $W = \{n \in \mathbb{N} \mid 1 \leqslant n \leqslant 8\}$; relations $\sim_i$ and $\leqslant$ are as depicted in the figure; $V(P) = \{8\}$, $V(I) = \{n \mid 1 \leqslant n \leqslant 6\}$, $V(L) = \{1, 2, 4, 5\}$, $V(S) = \{1, 3, 4, 6\}$, $V(A) = \{4, 5\}$, $V(N) = \{4, 6\}$, and $V(C) = W$. Model M describes a scenario where there is a world war and i is the president of a country. i has already come to know that a world war happens but she does not know whether her country is involved in the war. In model M, proposition P refers to 'the world is peaceful', I refers to 'i's country is involved in the war', C refers to 'i protects her civilians', L refers to 'the territorial land is invaded', S refers to 'the territorial sea is invaded', A refers to 'i sends the army', and N refers to 'i sends the navy'. In order to capture different notions concerning obligation, we need to define information sets.

Definition 12 (Information Set). *Given a pointed epistemic betterness structure (M, s) and a finite set of literals $Q = \{p_m, \neg p_m \mid 1 \leq m \leq n$ for some $n \in \mathbb{N}$ and $p_m \in \mathbf{P}\}$, let $I \subset Q$ and for each $p_m \in Q$ (or $\neg p_m \in Q$), if $p_m \in I$ ($\neg p_m \in I$), then $\neg p_m \in I$ ($p_m \in I$). Then the information set of state s is $I_s = \{\varphi \in I \mid M, s \models \varphi\}$.*

I_s consists of all true facts that have happened when s is the actual world. For the set $Q \setminus I$, it represents the state of affairs that would occur in the future. In the current example, let $I_5 = \{\neg P, I, L, \neg S\}$ be the information set in world 5 representing all states of affairs that have happened, thereby can be learnt by i. The remaining propositions $\{A, \neg N, C\}$ represent the states of affairs that would occur as a result of i's action.

Ideal Conditional Obligation $\bigcirc(\varphi|\psi)$: Hansson's conditional obligations $\bigcirc(\varphi|\psi)$ are defined over betterness structures, i.e., $M = \langle S, \leqslant, V\rangle$, where epistemic relations are absent. We call them ideal conditional obligations here to indicate that they describe the obligations regardless of agents' epistemic information. The term 'ideal' is borrowed from Jones and Pörn's [20]. Formula $\bigcirc(\varphi|\psi)$ can be read as: φ is ideally good given the situation ψ. The semantics of $\bigcirc(\varphi|\psi)$ is: all best ψ-states also satisfy φ, which considers all ontically possible states. Moreover, $\bigcirc(\_|\_)$ is a global operator, which implies that it does not depend on the state at which you evaluate it. Formula $\bigcirc(\varphi|\top)$ is a special type of ideal conditional obligations. It describes that φ is the ideal state of affairs over all ontically possible states.

In M, we have $M, 5 \models \bigcirc(P|\top) \wedge K_i\neg P$, which means that the president i has an ideal obligation to guarantee a peaceful world regardless of the information set I_5, although she knows that peace is no longer possible.

Epistemic Unconditional Obligation $\odot_i(\varphi|\top)$: Formula $\odot_i(\varphi|\top)$ tells the agent what ought to be the case given her current information. In M, over information set I_5, i only knows that $\neg P$. Based on her current information, we have $M, 5 \models \odot_i(\neg I|\top)$. Intuitively, she ought to guarantee that her country is not involved in the war. Arguably, epistemic unconditional obligations correspond to the notion of *absolute obligation* used by McCloskey to denote those obligations that an agent ought to comply with at a specific moment or under specific information [21]. To be subject to an absolute obligation is "to be in a moral situation with moral commitment" [21]. These properties are reflected in the intuition of $\odot_i(\varphi|\top)$. In M, we have $M, 5 \models \odot_i(A|L) \wedge \neg K_i L$, which means that i has an obligation to send an army when she knows that their territorial land is invaded, but she does not know that they are invaded. So it means that i does not have an absolute obligation to send the army. But, due to $M, 5 \models \odot_i(\neg I|\top)$, i has an absolute obligation not to have her country involved in the war.

Prima Facie Obligation $\odot_i^{\mathrm{P}}\varphi$: The formalization of *prima facie* obligations via unconditional obligation (ideal obligation in this paper) was first advanced in [22]. We expand on this tradition here, showing how our formalism also accommodates a natural formalization of this type of obligations. We use ideas from McCloskey's analysis [21] to justify our approach. We take as starting point McCloskey's observation that "an actual obligation does not differ 'qualitatively' from a *prima facie* obligation ..." and hence it could be captured by a formula $\odot_i(\varphi|\top)$. However, we still need to distinguish unconditional obligations that can be overridden (*prima facie*) from those that cannot. The overriding phenomenon, we argue, has to do with the acquisition of new information that brings about new *prima facie* obligations that override previous ones. This is a dynamic phenomenon, and our framework is well-suited to capture it.

Suppose that we only consider the single-agent case and the actions of truthful public announcements (see Sect. 3.2). Any public announcement introduces some true information. Taking the notation in public announcement logic, given an epistemic betterness structure $M = \langle W, \sim_1, \leqslant, V\rangle$, $M|_\varphi = \langle W \cap \|\varphi\|_M, \sim_1'$ $, \leqslant', V'\rangle$ where $\sim_1'$, $\leqslant'$, and V' are $\sim_1$, $\leqslant$, and V restricted to the set $W \cap \|\varphi\|_M$, respectively. We use a new operator $\bigodot_1^{\mathrm{P}} \varphi$ to denote 1's *prima facie* obligation to ensure φ. So, *prima facie* obligations can be defined as follows.

Definition 13. *Given a pointed epistemic betterness structure* (M, s)*,*

$$M, s \models \bigodot_1^{\mathrm{P}} \varphi \textit{ iff there exists } \psi \in I_s \textit{ such that } M|_\psi, s \models \bigodot_1(\varphi|\top).$$

The semantics of $\bigodot_1^{\mathrm{P}} \varphi$ means that it is *prima facie* obligatory that φ for 1 if and only if after receiving *some* true information, 1 has an epistemic unconditional obligation to ensure φ. In our example, we have $M, 5 \models \bigodot_i^{\mathrm{P}} A$ since $M|_L, 5 \models \bigodot_i(A|\top)$. But $M, 5 \models \neg \bigodot_i(A|\top)$. These mean that the president has a *prima facie* obligation to send the army once she knows that the territorial land is invaded. But this *prima facie* obligation is currently not an absolute obligation. Similarly, i also has a *prima facie* obligation to send both their army and navy, $\bigodot_i^{\mathrm{P}}(A \wedge N)$, once she knows that the territorial land and sea are invaded. But this *prima facie* obligation will never become an absolute obligation since S is not true in 5. We argue that the above definition of *prima facie* obligation succeeds in addressing the reservations moved by [2] to the approach to *prima facie* obligations based on conditional obligations.

All-Things-Considered Obligation $\bigodot_i^{\mathrm{A}} \varphi$: All-things-considered obligations are usually compared with *prima facie* obligations. Prakken and Sergot state that "To find out what one's duty proper is, one should consider all things, ··· [it] can be based on any aspect of the factual circumstances and find which one is more incumbent" [2]. The statement suggests that an all-things-considered obligation should be the most ideal state of affairs when introducing all true information. It is also strongly related to van der Torre's *exact factual detachment* in the context of objective conditional obligation when all factual premises are given (see Chap. 4.1 in [3]). We define it as follows:

Definition 14. *Given a pointed epistemic betterness structure* (M, s)*,*

$$M, s \models \bigodot_i^{\mathrm{A}} \varphi \textit{ iff } M|_{\bigwedge I_s}, s \models \bigodot_i(\varphi|\top).$$

$M|_{\bigwedge I_s}$ is the model updated by introducing all information on s. In our example, $M, 5 \models \bigodot_i^{\mathrm{A}}(A \wedge \neg N)$, which means that the president has an all-things-considered obligation to send the army rather than the navy. However, since she does not know that their territorial land has been invaded at the moment, this all-things-considered obligation is not an absolute obligation *yet*.

Safe Knowledge-Based Obligation $\bigodot_i^{\mathrm{S}} \varphi$: We introduce a type of obligation that, to the best of our knowledge, has not yet been discussed in the literature, but which arises naturally in our framework. We have mentioned that absolute obligations are defeasible given different information. But it is still possible to find some obligations that cannot be defeated by the acquisition of new information. In the study on conditional beliefs, Baltag and Smets define *safe beliefs*, where 'safe' means that they are persistent under revision with any true information [8]. Although their definition is founded on a connected plausibility relation, we can follow their idea and define *safe knowledge-based obligations* as follows:

Definition 15. *Given an epistemic betterness structure* $M = \langle S, \sim_1, \leqslant, V\rangle$, $M, s \models \bigodot_1^{\mathrm{S}} \varphi$ *if and only if the following two conditions are satisfied:*

1. $M, s \models \varphi$;
2. *for each* $t, r \in [s]^{\sim_1}$, *if* $t \in \|\varphi\|_M$ *and* $t \leqslant r$, *then* $r \in \|\varphi\|_M$.

Intuitively, if $M, s \models \bigodot_1^{\mathrm{S}} \varphi$, then φ is satisfied in the actual state and $\|\varphi\|_M \cap [s]^{\sim_1}$ is $\leqslant$-upward-closed. As a consequence, it is easy to check that $M, s \models \bigodot_1^{\mathrm{S}} \varphi \rightarrow \bigodot_1(\varphi|\top)$. Moreover, for any $\psi \in I_s$, we have $M|_{\psi}, s \models \bigodot_1(\varphi|\top)$. Thus, $M|_{\bigwedge I_s}, s \models \bigodot_1(\varphi|\top)$. This means that a safe obligation to ensure φ will never be defeated by introducing new true information. It will always be an absolute obligation as well as a *prima facie* obligation.

In our example, $\|C\|_M \cap [5]^{\sim_i}$ is $\leqslant$-upward-closed since C is satisfied over the whole set. Thus, $M, 5 \models \bigodot_i^{\mathrm{S}} C$. This means that i has a safe knowledge-based obligation to protect her civilians, no matter what information she received.

5 Reduction and Axiomatization

In this section, we will show that each $\mathcal{L}_{\mathrm{DKCDL}}$-formula in the form of $\bigodot_i(\varphi|\psi)$ can be reduced to some $\mathcal{L}_{\mathrm{DEL}}$-formula by a Kangerian-Andersonian reduction (KA-reduction) [9,10]. In its classical form, the reduction treats deontic operators $\bigcirc\varphi$ as $\Box(Q \rightarrow \varphi)$ where Q denotes a propositional ideality constant standing for ‘all obligations are met’. This reduction approach has been explored extensively in the literature, in a variety of settings (e.g. [4,23,24]).

Given a priority structure $\mathcal{G} = \langle \Phi, \prec\rangle$ and an arbitrary formula $\chi \in \Phi \cup \{\top\}$, define $\Phi_\chi = \{\chi' \in \Phi \mid \chi' \succ \chi\}$[1]. Thus, Φ_χ consists of all the formulas in $\mathcal{G}$ better than χ. The KA-reduction of $\bigodot_i(\varphi|\psi)$ relies on the formula:

$$\lambda_\psi^i : \bigvee_{\chi \in \Phi \cup \{\top\}}((\chi \wedge \psi) \wedge K_i(\bigvee \Phi_\chi \rightarrow \neg\psi))$$

Formula λ_ψ^i says “ψ is consistent with some χ in the priority structure and agent i knows that any state of affairs that is better than χ (i.e., $\bigvee \Phi_\chi$) must falsify ψ”.

Lemma 1. *Given a priority structure* $\mathcal{G} = \langle \Phi, \prec\rangle$ *and an arbitrary epistemic betterness structure* (M, s) *based on* $\mathcal{G}$, $M, s \models \lambda_\psi^i$ *iff* $s \in \max_{\leqslant}([s]^{\sim_i} \cap \|\psi\|_M)$.

Proof. ($\Rightarrow$) Suppose, to reach a contradiction, that $s \notin \max_{\leqslant}([s]^{\sim_i} \cap \|\psi\|_M)$. We split the proof into two cases: • Case 1: If $s \notin \|\psi\|_M$, then $M, s \not\models \lambda_\psi^i$. Contradiction. • Case 2: If $s \in \|\psi\|_M$ and s is not the best ψ-state, then there exists $t \in [s]^{\sim_i}$ such that $M, t \models \psi$ and $t > s$. Since there must exist $\chi \in \Phi \cup \{\top\}$ such that $M, s \models \chi \wedge \psi$, we have for any $r > s$ that there exists $\chi' \succ \chi$ (if χ is $\top$, then $\chi' \succ \top$ for each $\chi' \in \Phi$) such that $M, r \models \chi'$. Since $t \in [s]^{\sim_i}$,

[1] If $\chi \in \Phi$ and there is no $\chi' \in \Phi$ such that $\chi' \succ \chi$, let $\Phi_\chi = \varnothing$ and $\bigvee \Phi_\chi = \bot$. If $\chi = \top$, $\Phi_\chi = \Phi$.

$M, t \models \bigvee \Phi_\chi \to \neg\psi$. By $t > s$, we have that $M, t \models \chi'$, which implies that $M, t \models \bigvee \Phi_\chi$. So $M, t \models \neg\psi$. Contradiction. Therefore, $s \in \max_\leqslant([s]^{\sim_i} \cap \|\psi\|_M)$.

($\Leftarrow$) Suppose that $s \in \max_\leqslant([s]^{\sim_i} \cap \|\psi\|_M)$. There must exist $\chi \in \Phi \cup \{\top\}$ such that $M, s \models (\chi \wedge \psi) \wedge \neg \bigvee \Phi_\chi$. We then prove by two cases: • Case 1: If there is no $t \in [s]^{\sim_i}$ such that $t > s$, this implies that $\bigvee \Phi_\chi = \bot$. It is trivial that for each $r \sim_i s$, $M, r \models \bigvee \Phi_\chi \to \neg\psi$. So $M, s \models K_i(\bigvee \Phi_\chi \to \neg\psi)$. • Case 2: If there is $t \in [s]^{\sim_i}$ such that $t > s$, there must exist $\chi' \succ \chi$ in Φ such that $M, t \models \chi' \wedge \neg\psi$, which implies that $M, t \models \bigvee \Phi_\chi \to \neg\psi$. As for each $r \sim_i s$ such that $r \not> s$, $M, r \models \neg \bigvee \Phi_\chi$, this also implies that $M, r \models \bigvee \Phi_\chi \to \neg\psi$. So for all $u \sim_i s$, $M, u \models \bigvee \Phi_\chi \to \neg\varphi$. Therefore, $M, s \models K_i(\bigvee \Phi_\chi \to \neg\psi)$.

Therefore, λ^i_ψ captures the best ψ-state among the set of epistemically indistinguishable states for agent i. The outermost operator of λ^i_ψ is not $\bigodot_i(\_|\_)$.

Proposition 1 (KA-reduction). *Given an epistemic betterness structure (M, s) based on the priority structure $\mathcal{G}$,*

$$M, s \models \bigodot_i(\varphi|\psi) \leftrightarrow K_i(\lambda^i_\psi \to \varphi)$$

The proof involves a routine argument. The equivalence above helps to reduce each formula in the form of $\bigodot_i(\varphi|\psi)$ to a $\mathcal{L}_{\mathrm{DEL}}$-formula without any dyadic deontic operator.

5.1 Axiomatization $\mathbb{DKCDL}$

It should be noted that λ^i_ψ is defined by some ψ, some $i \in G$ and a certain priority structure. Therefore, our proof system is to be established based on a fixed priority structure $\mathcal{G}$.

Definition 16. *The proof system $\mathbb{DKCDL}$ consists of the following axiom schemas and inference rules:*

(TAUT)	*All instances of tautologies*
(K)	$K_i(\varphi \to \psi) \to (K_i\varphi \to K_i\psi)$
(T)	$K_i\varphi \to \varphi$
(4)	$K_i\varphi \to K_iK_i\varphi$
(5)	$\neg K_i\varphi \to K_i\neg K_i\varphi$
(U – A)	$[(U,e)]p \leftrightarrow (pre(e) \to post(e)(p))$
(U – N)	$[(U,e)]\neg\varphi \leftrightarrow (pre(e) \to \neg[(U,e)]\varphi)$
(U – C)	$[(U,e)](\varphi \wedge \psi) \leftrightarrow ([(U,e)]\varphi \wedge [(U,e)]\psi)$
(U – K)	$[(U,e)]K_i\varphi \leftrightarrow (pre(e) \to \bigwedge_{e' \sim_i e} K_i[(U,e')]\varphi)$
(KA)	$\bigodot_i(\varphi|\psi) \leftrightarrow K_i(\lambda^i_\psi \to \varphi)$
(MP)	*From φ and $\varphi \to \psi$, infer ψ*
(N)	*From φ, infer $K_i\varphi$*
(RE)	*From $\varphi \leftrightarrow \psi$, infer $\chi \leftrightarrow \chi[\varphi/\psi]$*

$\mathbb{DKCDL}$ is given based on the proof system for dynamic epistemic logic with postconditions **UM** [14], except (KA), which is given so as to reduce the dyadic deontic operators. (RE) is the inference rule *replacement (substitution) of equivalents*, which is admissible in $\mathbb{DKCDL}$. The notation $\chi[\varphi/\psi]$ denotes any formula obtained by replacing one or more occurrences of ψ in χ with φ.

Theorem 1. $\mathbb{DKCDL}$ *is sound and strongly complete with respect to the class of epistemic betterness structures.*

Proof (Sketch of proof). Soundness can be obtained straightforwardly from soundness of DEL and validity of (KA) (Proposition 1). Completeness can be proved by translating $\mathcal{L}_{\mathrm{DKCDL}}$-formulas to $\mathcal{L}_{\mathrm{EL}}$-formulas via KA-reduction, reduction axioms for dynamic operators, and induction on the complexity of the formulas (see Chap. 7.4 in [15] and Theorem 11 in [25]).

6 Conclusion

We extended the static logic of epistemic conditional obligation $\mathbb{KCDL}$ with a dynamic operator. We introduced priority structures as linguistic resources for referring to the betterness ordering on states of affairs. Accordingly, when an agent's epistemic conditional obligation is triggered (to an unconditional obligation) by getting new information or coming to know that some facts changed, the updated epistemic betterness structure shows the new information and the new betterness relation. Therefore, $\mathbb{DKCDL}$ can explicitly capture obligation change. We showed how this logic can naturally accommodate, in an original way, several key deontic notions such as ideal obligation, absolute obligation, *prima facie* obligation, all-things-considered obligation, and safe knowledge-based obligation. Furthermore, we established the sound and strongly complete axiom system $\mathbb{DKCDL}$ with respect to epistemic betterness structures, by giving a Kangerian-Andersonian reduction for the deontic operator and reduction axioms for the dynamic operator.

References

1. Hansson, B.: An analysis of some deontic logics. Nous, 373–398 (1969)
2. Prakken, H., Sergot, M.: Dyadic deontic logic and contrary-to-duty obligations. In: Nute, D. (eds.) Defeasible Deontic Logic. Synthese Library (Studies in Epistemology, Logic, Methodology, and Philosophy of Science), vol. 263. pp. 223–262. Springer, Dordrecht (1997)https://doi.org/10.1007/978-94-015-8851-5_10
3. van de Torre, L.: Reasoning about obligations: defeasibility in preference-based deontic logic. Tinbergen Institute Research Series 140. Amsterdam: Thesis (1997)
4. van Benthem, J., Grossi, D., Liu, F.: Priority structures in deontic logic. Theoria **80**(2), 116–152 (2014)
5. Su, X.: Knowledge-based conditional obligation. Short Papers Adv. Modal Logic AiML **2020**, 112–116 (2020)
6. Pacuit, E., Parikh, R., Cogan, E.: The logic of knowledge based obligation. Synthese **149**(2), 311–341 (2006)

7. Horty, J.: Epistemic oughts in stit semantics. Ergo J. Philos. **6** (2019)
8. Baltag, A., Smets, S.: Dynamic belief revision over multi-agent plausibility models. In: Proceedings of LOFT, pp. 11–24, University of Liverpool (2006)
9. Anderson, A.R.: A reduction of deontic logic to alethic modal logic. Mind **67**(265), 100–103 (1958)
10. Kanger, S.: New foundations for ethical theory. In: Hilpinen, R. (ed.) Deontic Logic: Introductory and Systematic Readings. SYLI, vol. 33, pp. 36–58. Springer, Dordrecht (1970). https://doi.org/10.1007/978-94-010-3146-2_2
11. Fagin, R., Moses, Y., Halpern, J.Y., Vardi, M.Y.: Reasoning about Knowledge. MIT Press (2003)
12. Liu, F.: Changing for the better: Preference dynamics and agent diversity. Ph.D. thesis, University of Amsterdam (2008)
13. van der Torre, L., Tan, Y.-H.: An update semantics for deontic reasoning. In: Norms, Logics and Information Systems, pp. 73–90 (1998)
14. van Ditmarsch, H., Kooi, B.: Semantic results for ontic and epistemic change. In: Proceedings LOFT, vol. 7, pp. 87–117, Amsterdam UP (2008)
15. van Ditmarsch, H., van der Hoek, W., Kooi, B.: Dynamic Epistemic Logic. Springer Science & Business Media (2007). https://doi.org/10.1007/978-1-4020-5839-4
16. Kraus, S., Lehmann, D., Magidor, M.: Nonmonotonic reasoning, preferential models and cumulative logics. Artif. Intell. **44**, 167–207 (1990)
17. Governatori, G., Rotolo, A., Sartor, G.: Temporalised normative positions in defeasible logic. Proc. ICAIL **10**, 25–34 (2005)
18. Benthem, J.v., Liu, F.: Dynamic logic of preference upgrade. J. Appl. Non-Class. Log, **17**(2), 157–182 (2007)
19. Greenspan, P.S.: Conditional oughts and hypothetical imperatives. J Philos. **72**(10), 259–276 (1975)
20. Jones, A.J., Pörn, I.: Ideality, sub-ideality and deontic logic. Synthese **65**(2), 275–290 (1985)
21. McCloskey, H.: Ross and the concept of a prima facie duty. Aust. J. Philos. **41**(3), 336–345 (1963)
22. Alchourrón, C.E.: Detachment and defeasibility in deontic logic. Stud. Log. **57**(1), 5–18 (1996)
23. Meyer, J.-J.C.: A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic. Notre Dame J. Formal Log. **29**, 109–136 (1988)
24. De Lima, T., Royakkers, L., Dignum, F.: A logic for reasoning about responsibility. Logic J. IGPL **18**(1), 99–117 (2010)
25. Kooi, B.: Expressivity and completeness for public update logics via reduction axioms. J. Appl. Non-Class. Logics **17**(2), 231–253 (2007)