The Manchester Twins: Conflicts Between Directed Obligations

Stef Frijters [1]

Ghent University

Thijs De Coninck [2]

Ghent university

Abstract

Term-modal logic uses modal operators that are indexed with terms of the language, which allows for quantification over these operators. Term-modal *deontic* logics (TMDL) can capture reasoning with rules, directed, and undirected obligations. Using the rich language of TMDL, we identify different types of deontic conflicts between directed obligations and describe reasoning in the face of these conflicts. We develop several monotonic logics in the TMDL family and show that none is capable of capturing all plausible deontic principles, while also being conflict-tolerant. To remedy this we develop several non-monotonic extensions in the format of adaptive logics. We end by isolating one of these, $\mathbf{TMDL^m}$, and commenting on it.

Keywords: Conflict-tolerant deontic logic, term-modal logic, first-order, undirected obligations, directed obligations.

1 Introduction

In deontic reasoning, one often encounters conflicting obligations. These conflicting obligations do not always result from conflicting moral theories or legal systems. Take, for example, the commonly accepted general rule: 'Doctors have an obligation to their patients to benefit the health of these patients'. [3] Taken on its own, this rule is perfectly consistent. However, in certain specific

[1] stef.frijters@ugent.be Stef Frijters holds a PhD grant of the Research Foundation - Flanders on the research project "Towards a more integrated formal account of actual ethical reasoning, with applications in medical ethics." (G0D2716N).

[2] thijs.deconinck@ugent.be Thijs De Coninck holds a PhD grant fundamental research of the Research Foundation - Flanders (1167619N).

[3] We have a distributive reading of this rule, instead of a collective one. Thus we interpret it as "Every individual doctor has an obligation to each of their patients to benefit the health of that patient." and not as "The group of all doctors have an obligation . . . " or "Each doctor has an obligation to the group of all of his/her patients . . . ". This sentence is also not meant to be interpreted as a generic sentence.

situations it can lead to deontic conflicts. Let us illustrate this with an example loosely based on the Manchester twins case [14,11], summarised by Kaveny: [4]

> A pair of conjoined twins, known by the pseudonyms of "Jodie" and "Mary," were born in Manchester, England, hospital in August 2000. Mary's heart and lungs were essentially non-functioning; she was entirely dependent upon her connection with her stronger sister for survival. But Jodie's cardiovascular system could not continue to do the work necessary to support both babies indefinitely. Physicians predicted that without an operation to separate the twins, both babies soon would die, probably before their first birthday. Unfortunately, however, the surgical separation would be able to save only Jodie. Although likely to need several reconstructive operations, she was predicted to live a long and virtually normal life once her body was liberated from the burden of providing life support to her sister. Mary's fate would be very different; she was predicted to die in the course of the procedure. [11, p. 115]

In this specific situation, benefitting Jodie's health implies performing the operation, while benefitting Mary's health implies refraining from it. Both Jodie and Mary are patients of the same physician. [5] Thus, this physician has an obligation to Jody to perform the operation, and an obligation to Mary not to do so: a genuine *deontic conflict* [7].

We define a deontic conflict as a situation in which multiple obligations hold that are individually, but not jointly fulfillable. In our example, the physician can perform the surgery, or she can refrain from it, but she cannot do both. Thus, these two obligations are individually fulfillable, but not jointly. This differs from a situation in which one is faced with multiple obligations none of which is fulfillable. These are excluded by our definition of a deontic conflict.

We can be more precise about the kind of deontic conflict with which the physician is faced. This is a conflict between *directed obligations*. A directed obligation is characterized by the fact that it has both a *bearer* and a *counterparty*. The bearer of an obligation is the person who is (in principle) blamed if the obligation is not fulfilled. In the Manchester Twins case, the physician is the bearer of both conflicting obligations. A counterparty is the person *to whom* the bearer has the obligation [10,5]. In the Manchester twins case, Jodie is the counterparty to the directed obligation that the physician has to operate. Mary is the counterparty to the directed obligation that the physician has to not operate.

Under normal circumstances, i.e. at least when there are no conflicts, it is plausible that directed obligations imply *undirected obligations*. With undirected obligations, we mean obligations that are only tied to a bearer and not to a counterparty [10,5]. In this paper we consider undirected obligations to be

[4] We say that this example is 'loosely based on' the case, as the actual case was much more complicated than this summary suggests [14,11].

[5] In reality there was a team of physicians, all responsible for both Jodie and Mary, but we make abstraction of this.

action guiding in the sense that they should not offer contradictory demands [20]. Suppose that a has an obligation toward b to tutor b's daughter c (as a has promised b to do so). This directed obligation normally implies the undirected obligation on the part of a to tutor c. Such an implication is, however, not so straightforward in cases with a deontic conflict between directed obligations.

In this paper we develop several logics with the aim of capturing reasoning with possibly conflicting directed obligations. The logics should enable us to derive conflicts from general premise sets, while at the same time being weak enough not to trivialize these conflicts. Specifically, we will develop term-modal deontic logics (TMDL) in the vein of [5], based on the more general framework of term-modal logics [4].

Term-modal logics are first-order modal logics with modal operators that are indexed by terms of the language (variables and constants). This allows one to quantify over (the indexes of) modal operators. In [5], these term-modal operators are given a deontic interpretation, to allow for the formalisation of general deontic rules, directed, and undirected obligations. However, the logic presented in [5] is not conflict-tolerant. To develop conflict-tolerant TMDL, we will use the neighborhood semantics for term-modal logics developed in [6], instead of the relational semantics of [4] and [5].

The paper is organised as follows. We begin in Section 2 by setting out **DE**, a very weak term-modal deontic logic. In the same section, we also discuss a number of monotonic extensions of **DE**. These logics all allow us to derive directed obligations from more general premises and to capture different principles of reasoning with both directed and undirected obligations. The next section is devoted to deontic conflicts. We distinguish two kinds of conflicts between directed obligations and then describe reasoning in the face of these conflicts. We show that the monotonic logics of Section 2 cannot at the same time capture all plausible principles, while also tolerating conflicts. To remedy this, Section 4 is devoted to defeasible versions of two principles of deontic logic. We show how we can use these to extend the monotonic logics to non-monotonic adaptive logics [1,2,3,19]. We end the paper by presenting some avenues of future research (Section 5).

2 A family of monotonic term-modal deontic logics

This section is divided into four subsections. The first of these presents the formal language that will be used in all of the logics in this article. Section 2.2 is dedicated to a semantic characterization of the weakest logic that we present: **DE**. A sound and complete axiomatisation of **DE** is given in Section 2.3. After this we discuss some other plausible principles of deontic logic and the ways in which we can extend **DE** to obtain these.

2.1 The formal language and its interpretation

Let $C = \{a, b, \ldots\}$ be a countable set of constants and $V = \{x, y, \ldots\}$ a countable set of variables. We let $\alpha, \beta, \alpha_1, \ldots$ range over C and $\nu, \xi, \nu_1, \ldots$ over V. Let $T = C \cup V$ be the set of terms and let $\theta, \kappa, \theta_1, \ldots$ be the metavariables ranging over it. For each $n \in \mathbb{N}$, let $\mathcal{P}^n$ be a countable set of n-ary predicate symbols and let $\mathcal{P}$ denote the union of all $\mathcal{P}^n$. Note that our language includes propositional variables, i.c. the 0-ary predicate symbols.

The formal language $\mathcal{L}$ is defined by the following Backus-Naur form, where $\Pi \in \mathcal{P}^n$, $\theta, \kappa \in T$ and $\nu \in V$:

$$\varphi ::= \Pi(\theta_1, \ldots, \theta_n) \mid \theta = \kappa \mid \neg\varphi \mid \varphi \vee \varphi \mid \mathsf{O}_\theta\varphi \mid \mathsf{O}^\theta_\kappa\varphi \mid (\forall\nu)\varphi \mid [\mathsf{U}]\varphi$$

The other Boolean connectives are defined in the standard way. Additionally, $(\exists\nu)\varphi =_{df} \neg(\forall\nu)\neg\varphi$, $\mathsf{P}_\theta\varphi =_{df} \neg\mathsf{O}_\theta\neg\varphi$, $\mathsf{P}^\theta_\kappa\varphi =_{df} \neg\mathsf{O}^\theta_\kappa\neg\varphi$ and $\langle\mathsf{U}\rangle\varphi =_{df} \neg[\mathsf{U}]\neg\varphi$. We will write $\theta \neq \kappa$ instead of $\neg(\theta = \kappa)$. [6]

The notions of free and bound variables are as usual, with two additions (cf. Fitting et al. [4]): (1) The free occurrences of variables in $\mathsf{O}_\theta\varphi$ are all free occurrences of variables in φ and in addition θ if θ is a variable, and (2) the free occurrences of variables in $\mathsf{O}^\kappa_\theta\varphi$ are θ, if θ is a variable, κ, if κ is a variable, and all free occurrences of variables in φ. A wff φ is a sentence iff all the variables in φ are bound. Let $\mathcal{S}$ be the set of sentences of $\mathcal{L}$.

We interpret $\mathsf{O}^b_a\varphi$ as the directed obligation 'a has an obligation towards b that φ' and $\mathsf{O}_a\varphi$ as the undirected obligation 'a has an obligation that φ'. We will only use terms to refer to agents, and not to other objects, such as apples. In this way we can avoid being able to express sentences such as 'this apple has an obligation'.

$[\mathsf{U}]$ is a universal modal operator and we interpret $[\mathsf{U}]\varphi$ as 'φ is settled true'. This operator allows us to express more conflicts. As an example, we can look back at the tutoring case. Here, a had promised b to tutor c, say at three in the afternoon. As a result, a has an obligation towards b that a tutors c at three in the afternoon. Suppose that a has also promised their friend d to meet for an afternoon of playing computer games. The resulting (directed) obligation conflicts with the obligation that a has towards c, but only because it is impossible to fulfill both obligations. This is not a logical impossibility, but for all intents and purposes it is *settled true* that b does not both tutor c at three and also meets d for an afternoon of playing computer games. We can express this with the $[\mathsf{U}]$-operator.

$\mathcal{L}$ allows for a great deal of precision. Let Sx be interpreted as 'x performs the surgery'. In $\mathcal{L}$ we can express that it is obligatory for our physician (a), to perform the surgery, O_aSa, or that she has this obligation towards Jodie (j), O^j_aSa. $\mathcal{L}$ also has the expressive power to formalise sentences where the agent of the obligatory action is not the bearer of the obligation, such as in 'it is obligatory for the head of the hospital, b, that someone else performs the surgery': $\mathsf{O}_b(\exists x)(x \neq b \wedge Sx)$. [7] It is also possible to distinguish 'there is someone for whom it is obligatory to perform the surgery', $(\exists x)\mathsf{O}_xSx$, from 'it

[6] Note that the brackets around $(\theta = \kappa)$ are strictly speaking unnecessary.

[7] The sentence 'it is obligatory for the head of the hospital, b, that someone else performs the surgery' should not be confused with 'it is obligatory for the head of the hospital, b, that b *brings it about that* someone else performs the surgery'. In the second sentence the agent of the obligatory action is also the bearer of the obligation, whereas that is not the

is obligatory for someone that someone performs the surgery', $(\exists x)\mathsf{O}_x(\exists y)Sy$. Finally, we can express general rules such as the one from the introduction, i.e. that if x is a patient of y (Pxy), then y has an obligation towards x to benefit the health of x (Byx): $(\forall x)(\forall y)(Pxy \to \mathsf{O}^x_y Byx)$.

2.2 DE, the weakest logic

We now present a semantic characterization of **DE**, the weakest logic in the TMDL-family. These semantics are based on the neighborhood semantics for term-modal logics in [6]. A **DE**-model is a tuple $M = \langle W, \mathcal{A}, N^P, N^D, I, w_a \rangle$. W is a state domain, consisting of possible worlds $w, w_1, \ldots$ and $\mathcal{A}$ is an agent-domain, consisting of agents $p, p_1, p_2, \ldots$. Both are non-empty and are allowed to be at most countably infinite. I is an interpretation function. The actual world w_a is used to determine validity in the model (Definition 2.6, this becomes important in Section 4).

Definition 2.1 A **DE**-model is a tuple $M = \langle W, \mathcal{A}, N^P, N^D, I, w_a \rangle$, where:
1. $W \neq \emptyset$
2. $\mathcal{A} \neq \emptyset$
3. $N^P : W \times \mathcal{A} \to \wp(\wp(W))$ is a *neighborhood function* of M
3.1 for all $w \in W$ and $p \in \mathcal{A}$: if $X \in N^P(w,p)$ and $X \subseteq Y \subseteq W$, then $Y \in N^P(w,p)$
3.2 for all $w \in W$ and $p \in \mathcal{A}$: $W \in N^P(w,p)$
3.3 for all $w \in W$ and $p \in \mathcal{A}$: $\emptyset \notin N^P(w,p)$
3.4 for all $w \in W$ and $p \in \mathcal{A}$: if $X, Y \in N^P(w,p)$, then $X \cap Y \in N^P(w,p)$
4. $N^D : W \times \mathcal{A} \times \mathcal{A} \to \wp(\wp(W))$ is a *neighborhood function* of M
4.1. For all $w \in W$ and $p_1, p_2 \in \mathcal{A}$: $\emptyset \notin N^D(w, p_1, p_2)$
5. I is an *interpretation* function such that:
5.1. $I : T \to \mathcal{A}$
5.2. $I : \mathcal{P}^n \times W \to \wp(\mathcal{A}^n)$ for every natural number $n \in \mathbb{N}$ such that $1 \leq n$
5.3. $I : \mathcal{P}^0 \to \wp(W)$
6. $w_a \in W$.

The neighborhood function N^P assigns to each world-agent pair a set of propositions that are obligatory for this agent (each proposition being a set of worlds). This will be used to interpret the undirected obligation operator. N^P has a number of conditions. The first of these ensures inheritance: if a proposition is obligatory, then what necessarily follows from this proposition will also be obligatory. The second condition ensures that what is necessary is obligatory, and the third ensures that what is impossible cannot be obligatory. The final condition corresponds to aggregation: if two propositions are obligatory, then their conjunction is obligatory as well. Taken together, this

---

case in the first sentence. That obligations exist where the bearer is not the agent of the obligatory action has been argued in [5,12,10]. To properly express the second sentence, we could extend our language with a term-modal 'bring it about'-operator. The technical results in [6] combined with the neighborhood semantics of [9] allow one to give a sound and complete logic for this extended language. However, since this extension is not essential for what follows, we leave a development of this approach for future work.

means that the undirected obligation operator is as as the obligation operator of standard deontic logic.

The neighborhood function N^D assigns to every triple consisting of a world and two agents a set of propositions that are obligatory for the first agent towards the second agent. Condition 4.1. ensures that what is obligatory, is also possible. The reason for this condition is that we do not want the logic to model unfulfillable directed obligations. We defined a conflict as a situation in which multiple (directed) obligations hold that can each be individually fulfilled, but which are not jointly fulfillable. The ought-implies-can principle for directed obligations that is expressed by condition 4.1. ensures that all directed obligations can indeed be individually fulfilled.

To interpret quantifiers, we define ν-alternatives, before we give the semantic clauses. As usual, for any $\varphi \in \mathcal{L}$ and **DE**-model $M = \langle W, \mathcal{A}, N^P, N^D, I, w_a \rangle$, $[\![\varphi]\!]_M =_{\mathsf{df}} \{w \in W \mid M, w \vDash \varphi\}$.

Definition 2.2 [ν-alternative] For any $\nu \in V$, $M' = \langle W, \mathcal{A}, N^P, N^D, I', w_a \rangle$ is a ν-alternative to $M = \langle W, \mathcal{A}, N^P, N^D, I, w_a \rangle$ iff I' differs at most from I in the member of $\mathcal{A}$ that I' assigns to ν.

Definition 2.3 [Semantic Clauses] For any **DE**-model $M = \langle W, \mathcal{A}, N^P, N^D, I, w_a \rangle$:
SC1 $M, w \models P(\theta_1, \ldots, \theta_n)$ iff $\langle I(\theta_1), \ldots, I(\theta_n) \rangle \in I(P, w)$
SC1' $M, w \models P$ iff $w \in I(P)$
SC2 $M, w \models \neg\varphi$ iff $M, w \not\models \varphi$
SC3 $M, w \models \varphi \vee \psi$ iff $M, w \models \varphi$ or $M, w \models \psi$
SC4 $M, w \models \theta = \kappa$ iff $I(\theta) = I(\kappa)$
SC5 $M, w \models \mathsf{O}_\theta \varphi$ iff $[\![\varphi]\!]_M \in N^P(w, I(\theta))$
SC6 $M, w \models \mathsf{O}^\kappa_\theta \varphi$ iff $[\![\varphi]\!]_M \in N^D(w, I(\theta), I(\kappa))$
SC7 $M, w \models (\forall \nu)\varphi$ iff for every ν-alternative M': $M', w \models \varphi$
SC8 $M, w \models [\mathsf{U}]\varphi$ iff $M, w' \models \varphi$ for all $w' \in W$.

In the following three definitions we define semantic consequence, validity and validity in a model. In this last definition, we use the actual world.

Definition 2.4 Where $\Gamma \subseteq \mathcal{S}$ and $\varphi \in \mathcal{S}$, φ *is a semantic consequence of* Γ, $\Gamma \Vdash \varphi$, iff for every **DE**-model $M = \langle W, \mathcal{A}, N^P, N^D, I \rangle$ and $w \in W$: if $M, w \models \psi$ for all $\psi \in \Gamma$, then $M, w \models \varphi$.

Definition 2.5 Where $\Gamma \subseteq \mathcal{S}$ and $\varphi \in \mathcal{S}$, **DE** *validates* φ iff for every **DE**-model $M = \langle W, \mathcal{A}, N^P, N^D, I, w_a \rangle$ and $w \in W$: $M, w \models \varphi$.

Definition 2.6 Where $\varphi \in \mathcal{S}$, φ *is valid in a model* M, $M \models \varphi$, iff $M, w_a \models \varphi$

2.3 Axiomatisation of DE

A sound and strongly complete axiomatisation of **DE** is obtained by closing a complete axiomatisation of classical propositional logic (**CL**) with all instances of the axiom schemata in Table 1 under the rules of Table 2.[8] $\varphi(\theta/\kappa)$ is

[8] Soundness and completeness follow from previous results in [6]. See also [4,17,5].

the result of replacing all free occurrences of κ in φ by θ, relettering bound variables if necessary to avoid rendering new occurrences of θ bound in $\varphi(\theta/\kappa)$. $\varphi(\theta//\kappa)$ is the result of replacing various (not necessarily all or even any) free occurrences of θ in φ by occurrences of κ, again relettering if necessary [18, p. 57].

(UK)	$[\mathsf{U}](\varphi \to \psi) \to ([\mathsf{U}]\varphi \to [\mathsf{U}]\psi)$	(UI)	$(\forall\nu)\varphi \to \varphi(\alpha/\nu)$
(UT)	$[\mathsf{U}]\varphi \to \varphi$	(REF)	$\alpha = \alpha$
(U5)	$\langle\mathsf{U}\rangle\varphi \to [\mathsf{U}]\langle\mathsf{U}\rangle\varphi$	(SUB)	$(\alpha = \beta) \to (\varphi \to \varphi(\alpha//\beta))$
(UBF)	$(\forall\nu)[\mathsf{U}]\varphi \to [\mathsf{U}](\forall\nu)\varphi$	(ND)	$\alpha \neq \beta \to [\mathsf{U}]\alpha \neq \beta$
(DREU)	$(\mathsf{O}_\alpha^\beta\varphi \wedge [\mathsf{U}](\varphi \leftrightarrow \psi)) \to \mathsf{O}_\alpha^\beta\psi$		
(DIC)	$\mathsf{O}_\alpha^\beta\varphi \to \langle\mathsf{U}\rangle\varphi$		
(PK)	$\mathsf{O}_\alpha(\varphi \to \psi) \to (\mathsf{O}_\alpha\varphi \to \mathsf{O}_\alpha\psi)$		
(PIC)	$\mathsf{O}_\alpha\varphi \to \langle\mathsf{U}\rangle\varphi$		
(PN)	$[\mathsf{U}]\varphi \to \mathsf{O}_\alpha\varphi$		

Table 1
Axiom schemata

(MP)	if $\varphi \to \psi$ and φ, then ψ
(UG)	if $\vdash \varphi \to \psi(\alpha/\nu)$ and α not in φ or ψ, then $\vdash \varphi \to (\forall\nu)\psi$.
(UNEC)	if $\vdash \varphi$, then $\vdash [\mathsf{U}]\varphi$

Table 2
Rules

There is little that is surprising in this axiomatisation. $[\mathsf{U}]$ is an **S5**-operator, O_α is a normal modal operator satisfying the ought-implies-can principle and O_α^β is a classical modal operator with the ought-implies-can principle. The other schemes are familiar from first-order modal logic. What might be surprising is that we do not have the Barcan formula for the obligation-operators even though we work with a constant domain semantics. This is a result of using neighborhood semantics instead of relational semantics [6].

2.4 Some further principles for the directed obligation operator

In this section we discuss four more logical principles for the directed obligation operator: (DP), necessitation, inheritance and aggregation. Standard deontic logic (**SDL**) satisfies the last three, but each of these can also be given up (see for example [19]). By adding any combination of the four conditions to Definition 2.1, we can define extensions of **DE**. We do so in Table 3 on page 10. The first column gives the name of the logic, the next four columns the conditions that it satisfies.

As we stated in the introduction, under normal circumstances directed and undirected obligations are related to each other in a natural way. If a has towards b a directed obligation to tutor c, then a has an undirected obligation to tutor c. This principle, stating that directed obligations imply undirected obligations, will be called (DP): $\mathsf{O}_\alpha^\beta\varphi \to \mathsf{O}_\alpha\varphi$. We can validate it easily by adding the following condition (that we call (dp)): for all $w \in W$, $p_1, p_2 \in \mathcal{A}$ and $X \subseteq W$: if $X \in N^D(w, p_1, p_2)$ then $X \in N^P(w, p_1)$.

Necessitation is the principle that anything that is settled true, is also obligatory: $[\mathsf{U}]\varphi \to \mathsf{O}^{\beta}_{\alpha}\varphi$. We can validate it by adding the condition (n) to our models: for all $p_1, p_2 \in \mathcal{A}$ and $w \in W$: $W \in N^D(w, p_1, p_2)$.

Inheritance is the principle: $\mathsf{O}^{\beta}_{\alpha}(\varphi \wedge \psi) \to (\mathsf{O}^{\beta}_{\alpha}\varphi \wedge \mathsf{O}^{\beta}_{\alpha}\psi)$. It is validated by models satisfying the condition (m): for all $w \in W$ and $p_1, p_2 \in \mathcal{A}$, if $X \in N^D(w, p_1, p_2)$ and $X \subseteq Y \subseteq W$, then $Y \in N^D(w, p_1, p_2)$. Note that any models satisfying condition (m) also validates the principle that we will call inheritance*: $(\mathsf{O}^{\beta}_{\alpha}\varphi \wedge [\mathsf{U}](\varphi \to \psi)) \to \mathsf{O}^{\beta}_{\alpha}\psi$.

Finally, aggregation (between directed obligations with the same bearer) says that if φ and ψ are obligatory, then their conjunction is also obligatory: $(\mathsf{O}^{\beta}_{\alpha}\varphi \wedge \mathsf{O}^{\beta}_{\alpha}\psi) \to \mathsf{O}^{\beta}_{\alpha}(\varphi \wedge \psi)$. It corresponds to the condition (c): for all $w \in W$, $p_1, p_2 \in \mathcal{A}$ and $X, Y \in N^D(w, p_1, p_2)$, $X \cap Y \in N^D(w, p_1, p_2)$.

3 Deontic conflicts

We distinguish two different types of conflicts between directed obligations, before discussing the kind of reasoning that is employed when encountering such conflicts.

3.1 Types of deontic conflict

In the introduction we distinguished deontic conflicts from situations in which an impossible proposition is obligatory. We see a *deontic conflict* as a situation in which two or more obligations hold that are not jointly fulfillable, but neither of which is impossible to fulfill on its own. In the Manchester twins case, the doctor has an obligation towards Jodie to perform the surgery, and another obligation towards Mary not to perform the surgery. In this article, we focus on such conflicts between directed obligations with the same bearer.

We distinguish two kinds of deontic conflicts between directed obligations with the same bearer: bilateral and multilateral conflicts. Multilateral conflicts are conflicts between directed obligations with distinct counterparties (for example $\{\mathsf{O}^b_a Qa, \mathsf{O}^c_a \neg Qa\}$ or $\{\mathsf{O}^b_a Pa, \mathsf{O}^b_a Qa, \mathsf{O}^c_a \neg(Qa \wedge Pa)\}$ in a context where $b \neq c$). In the Manchester twins case there is such a multilateral conflict: Mary is the counterparty of one obligation, and Jodie of the other obligation. Bilateral conflicts are conflicts where all the obligations involved are directed and where the counterparty is the same for all those obligations (for example $\{\mathsf{O}^b_a Qa, \mathsf{O}^b_a \neg Qa\}$). [9]

Consider the following case of a bilateral conflict: A patient w with cystic fibrosis is in need of a life-saving blood transfusion by doctor b. However, w is a Jehovah's witness, and refuses the transfusion on religious grounds [15, pp. 34-35]. The same general rule holds as in the Manchester twins case: 'Doctors have an obligation to their patients to benefit the health of these

[9] In this paper we will not consider conflicts between directed obligations with different bearers, but it is possible to make analogous constructions for these.

patients'. From this rule and the information at hand it follows that 'Doctor b has an obligation towards w to administer a blood transfusion to w'. However, this time there is also a second rule in play: 'Doctors have an obligation to their patients to respect the autonomy of these patients'. Since patient w refuses a blood transfusion, respecting the autonomy of w necessarily implies not administering a blood transfusion to w. Hence, b is faced with a bilateral conflict: b has an obligation towards w to administer the blood transfusion, and b has another obligation towards w not to administer the blood transfusion.

Not every conflict is a conflict between the obligatoriness of a proposition and its negation. Sometimes, as in the tutoring and gaming example above, the incompatibility of obligatory propositions is not due to logical impossibility, but due to contingent circumstances. We can use the $[\mathsf{U}]$-operator to express that two propositions are mutually incompatible. At other times, we will have conflicts between three or more obligations, e.g. $\{\mathsf{O}^a_d(P \vee Q), \mathsf{O}^b_d\neg P, \mathsf{O}^c_d\neg Q\}$. Finally, it is also possible to have existentially quantified formulas as part of a conflict. Thus, we consider $(\exists x)(\mathsf{O}^a_x Px \wedge \mathsf{O}^b_x\neg Px)$ to be a multilateral conflict as well.

Premise sets will usually not explicitly contain formulas that fit neatly into the definition of deontic conflicts above. Instead, we have to deduce these by means of deontic reasoning. In the Manchester twins case, the premises are: (1) all doctors have an obligation to their patients to benefit the health of these patients, (2) Jodie and Mary are patients of physician a, (3) it is necessary that if physician a acts to benefit Jodie, then she does perform the surgery and (4) if physician a acts to benefit Marie, then she does not perform the surgery. We can express these premises in the language as follows:

(i) $(\forall x)(\forall y)(Pxy \to \mathsf{O}^x_y Byx)$

(ii) $Pja \wedge Pma$

(iii) $[\mathsf{U}](Baj \to Sa)$

(iv) $[\mathsf{U}](Bam \to \neg Sa)$

No combination of these formulas fits the definition of a deontic conflict. We need a logic that is strong enough to derive a conflict from such a premise set, but does not lead to triviality once it does so. **DE** allows us to derive the conflict $\{\mathsf{O}^j_a Baj, \mathsf{O}^m_a\neg Bam, \neg\langle\mathsf{U}\rangle(Baj \wedge Bam)\}$ (by (UI), (MP), (UK) and (UNEC)). [10]

3.2 Reasoning in the face of deontic conflicts

When we are faced with a deontic conflict, we do not throw our hands up in the air and forego any further reasoning. We also do not conclude that everything is suddenly obligatory. This leads us to our first desideratum: deontic conflicts should not be trivialized. This means that if we have a deontic conflict in our premises, we should not be able to derive $\perp$.

[10] With any of the logics in Table 3 that validate inheritance, we can derive $\{\mathsf{O}^j_a Sa, \mathsf{O}^m_a\neg Sa\}$ as well.

Any extension of **DE** that validates aggregation trivialises bilateral conflicts. Consider, for example, the premise set $\{\mathsf{O}^b_a Qa, \mathsf{O}^b_a \neg Qa\}$. By aggregation, we can derive $\mathsf{O}^b_a(Qa \wedge \neg Qa)$. By the ought-implies-can principle, we can derive $\langle\mathsf{U}\rangle(Qa \wedge \neg Qa)$. By **CL** and the S5 properties of $[\mathsf{U}]$, we derive $\perp$.

When an extension of **DE** validates (DP), then it tolerates none of the conflicts identified above. From a conflict between directed obligations, we can derive a conflict between undirected obligations. Since O_α is a normal modal operator, **DE** trivialises such conflicts.

Name	(m)	(c)	(n)	(dp)	bilateral	multilateral
DE					✓	✓
DM	x				✓	✓
DC		x				✓
DR	x	x				✓
DN			x		✓	✓
DMN	x		x		✓	✓
DCN		x	x			✓
DK	x	x	x			✓
DE + DP				x		
DM + DP	x			x		
DC + DP		x		x		
DR + DP	x	x		x		
DN + DP			x	x		
DMN + DP	x		x	x		
DCN + DP		x	x	x		
DK + DP	x	x	x	x		

Table 3
The different monotonic logics

However, there are other possible desiderata than conflict-tolerance that we must take into account. For obligations that are not tainted by conflicts we want to be able to apply all the principles from Section 2.4 that we deem to be plausible. However, since (DP) and aggregation are incompatible with a logic that is conflict-tolerant, this means that we need defeasible versions of these principles. [11] If we find the principle (DP) plausible, then we should, for example, be able to derive $\mathsf{O}_a Qa$ from $\{\mathsf{O}^b_a Qa\}$ or from $\{\mathsf{O}^b_a Qa, \mathsf{O}^c_a Pa, \mathsf{O}^d_a \neg Pa\}$, but not from $\{\mathsf{O}^b_a Qa, \mathsf{O}^b_a \neg Qa\}$. Similarly, if one finds aggregation of directed obligations plausible, but also wants the logic to be conflict-tolerant, then one would want to be able to derive $\mathsf{O}^b_a(Qa \wedge Pa)$ from $\{\mathsf{O}^b_a Qa, \mathsf{O}^b_a Pa\}$, but not $\mathsf{O}^b_a(Qa \wedge \neg Qa)$ from $\{\mathsf{O}^b_a Qa, \mathsf{O}^b_a \neg Qa\}$. This is the second desideratum.

[11] This kind of problem is typical for the type of solution to normative conflicts that we propose here. Goble notes the same problem for conflict-tolerant variants of **SDL** [7, p. 297]: if the logic is weak enough to be conflict-tolerant, then it does not validate all principles of **SDL**, and if the logic does validate all principles of **SDL**, then it is not conflict-tolerant.

Finally, we should note that we do not want conflicts between undirected obligations to be derivable from conflicting directed obligations. In this paper we are only concerned with conflicts between directed obligations. Our undirected obligations should be action guiding in the sense that they should not offer contradictory demands [20]. All undirected obligations of an agent should be jointly fulfillable. Therefore we are only concerned with logics that satisfy ought-implies-can and aggregation for undirected obligations.

4 Adaptive extensions

The principles (DP) and aggregation seem incompatible with tolerating the conflicts between directed obligations. Nevertheless, one can argue that these principles are plausible for obligations not involved in a conflict (cf. the second desideratum). In this section we develop adaptive logics that take this idea into account. These logics allow us to apply (DP) or aggregation in unproblematic cases, but block the application of the same principles for obligations that are in conflict. [12]

In Section 4.1 we explain the basic ideas of adaptive logics, using a toy logic. Section 4.2 sets out two problems with this toy logic. The last two sections, 4.3 and 4.4, develop adaptive logics based on the conflict-tolerant logics that were presented in Section 3, while taking the two problems of the toy logic into account.

4.1 Adaptive logic, a toy example

To explain adaptive logics, we will use a running toy example of such a logic. The motivating idea behind this toy logic is that we would like to have a logic where the principle (DP) is blocked only in case the obligations involved are in conflict with other obligations. For example, we should be able to derive $\mathsf{O}_a\varphi$ from $\{\mathsf{O}_a^b\varphi\}$, but not from $\{\mathsf{O}_a^b\varphi, \mathsf{O}_a^c\neg\varphi\}$. In this last premise set, $\mathsf{O}_a^b\varphi$ is part of a conflict and so applying (DP) would lead to triviality.

Our toy logic uses the standard format of adaptive logics [19]. Every logic in the standard format is defined by a lower limit logic (LLL), a set of abnormalities and an adaptive strategy. For our present purposes the LLL can be any of the logics in Table 3 that does not validate (DP). The adaptive logic validates all of the valid formulas of the LLL, so taking a logic that validates (DP) as LLL will result in an adaptive logic that does not block (DP) in any circumstance. For this toy example, we will use **DK** as the LLL.

Abnormalities are those formulas that we want the logic to falsify as often as possible. How this 'as often as possible' is interpreted exactly is determined by the adaptive strategy. In our case we want all negations of instances of (DP) to be falsified as long as this does not lead to triviality. So we use $\Omega = \{\mathsf{O}_\alpha^\beta\varphi \wedge \neg\mathsf{O}_\alpha\varphi \mid \alpha, \beta \in C \text{ and } \varphi \in \mathcal{L}\}$ as the set of abnormalities.

[12] The main advantage of adaptive logic over other non-monotonic formalisms is that adaptive logics (in the standard format) give us a dynamic proof theory [19, p. 528]. This dynamic proof theory "explicates the dynamics of defeasible reasoning" [16, p. 9]. (Another advantage is the transparent handling of premise sets [16, pp. 87-88].)

For all logics in this section, including our toy logic, we will use the adaptive strategy known as 'minimal abnormality' [19]. To explain this strategy, we first need some preliminary definitions. We say that a model M is a model of Γ iff for all $\varphi \in \Gamma$, $M \models \varphi$. For any model M, $\mathrm{Ab}(M) =_{\mathsf{df}} \{\varphi \mid \varphi \in \Omega \text{ and } M \models \varphi\}$.

Definition 4.1 An LLL-model M of Γ is *minimally abnormal* iff there is no LLL-model M' of Γ such that $\mathrm{Ab}(M') \subset \mathrm{Ab}(M)$.

The models of our adaptive logic are those models of the LLL that are minimally abnormal. Take our toy logic and the premise set $\{\mathsf{O}_a^b Q\}$. There are **DK** models M of this premise set such that $\mathrm{Ab}(M) = \emptyset$. In all of these models $M \models \mathsf{O}_a Q$ (otherwise $\mathsf{O}_a^b Q \wedge \neg \mathsf{O}_a Q \in \mathrm{Ab}(M)$ and thus $\mathrm{Ab}(M) \neq \emptyset$). Hence, $\mathsf{O}_a Q$ is a semantic consequence of $\{\mathsf{O}_a^b Q\}$ in our toy logic. However, if we take the premise set $\{\mathsf{O}_a^b Q, \mathsf{O}_a^b \neg Q\}$, then for all LLL-models M: $M \models \mathsf{O}_a^b Q \wedge \neg \mathsf{O}_a Q$ or $M \models \mathsf{O}_a^b \neg Q \wedge \neg \mathsf{O}_a \neg Q$. So $\mathsf{O}_a Q$ is not a consequence of this premise set.

The standard format of adaptive logic also gives us a proof theory, and soundness and completeness for our adaptive logic. Due to space constraints we will not elaborate on this here. Instead, we refer interested readers to [19].

4.2 Two problems with the toy logic

In the previous section, we presented a toy logic that gives us an adaptive version of (DP). In this section, we present two problems of this toy logic. Then, in the next section, we will present an adaptive logic that solves these problems.

The first problem with the toy logic is that it is what adaptive logicians call a *flip-flop*. An adaptive logic is a flip-flop iff, for all premise sets from which an abnormality is derivable, the formulas that are derivable with the adaptive logic are the same as those derivable with the LLL [19]. In other words, in the presence of an abnormality, the adaptive logic collapses into the LLL.

To illustrate this, let us take the premise set $\{\mathsf{O}_a^b Pa, \mathsf{O}_a^c \neg Pa, \mathsf{O}_a^d Qa\}$ as an example. Here there is a conflict between the obligations that a has towards b and c. So we would want to block the derivation of $\mathsf{O}_a Pa$ and $\mathsf{O}_a \neg Pa$. However, $\mathsf{O}_a^d Qa$ is unproblematic and so we do want $\mathsf{O}_a Qa$ to be derivable.

Sadly, this is impossible in the toy logic. To see this, consider the following three abnormalities: $\mathsf{O}_a^b(Pa \vee \neg Qa) \wedge \neg \mathsf{O}_a(Pa \vee \neg Qa)$, $\mathsf{O}_a^c(\neg Pa \vee \neg Qa) \wedge \neg \mathsf{O}_a(\neg Pa \vee \neg Qa)$ and $\mathsf{O}_a^d Qa \wedge \neg \mathsf{O}_a Qa$. Each minimally abnormal model validates at least one of these abnormalities and every one of these formulas is validated in at least one of the minimally abnormal models. Since the last formula is validated in some minimally abnormal models, one cannot derive $\mathsf{O}_a Qa$ in the toy logic.

A second problem with the abnormalities of the toy logic can be illustrated by taking as a premise set $\{(\exists x)\mathsf{O}_x^a Pb\}$. From this, we would want to be able to derive $(\exists x)\mathsf{O}_x Pb$. However, there are minimally abnormal models of $\{(\exists x)\mathsf{O}_x^a Pb\}$ that do not validate $\{(\exists x)\mathsf{O}_x Pb\}$.

Take, for instance, a **DK**-model $M = \langle W, \mathcal{A}, N^P, N^D, I, w_a \rangle$ where $W = \{w_a, w_b, w_c\}$ and $\mathcal{A} = \{p_1, p_2\}$. For every $w \in W$ and $p \in \mathcal{A}$, let $N^P(w, p) = \{w_c\}$. Let $N^D(w_a, p_1, p_2) = \{w_b\}$ and for every $\langle w, p, p' \rangle \in \{\langle w, p, p' \rangle \mid w \in$

W and $p, p' \in \mathcal{A}\} \setminus \langle w_a, p_1, p_2\rangle$, let $N^D(w, p, p') = \{w_c\}$. Let I be such that for all $\theta \in T$, $I(\theta) = p_2$, $I(P, w_b) = \{p_2\}$ and $I(P, w_1) = I(P, w_2) = \emptyset$. M validates $(\exists x)\mathsf{O}^b_x Pb \wedge \neg(\exists x)\mathsf{O}_x Pb$, but this formula is not an abnormality. M is a minimally abnormal model of $\{(\exists x)\mathsf{O}^a_x Pb\}$ and does not validate $(\exists x)\mathsf{O}_x Pb$, thus we cannot derive this in our toy logic.

4.3 Adaptive (DP)

The problem of flip-flops is well-known in the study of adaptive logics. We can use the following solution, taken from [13]. [13]

Let $\mathcal{L}^a$ be the literals in $\mathcal{L}$. Where $\Theta \subseteq \mathcal{L}^a$ is finite and non-empty, we define $\sigma^\kappa_\theta(\Theta)$ as follows:

$$\sigma^\kappa_\theta(\Theta) =_{\mathsf{df}} \{\mathsf{O}^\kappa_\theta(\bigvee \Theta') \wedge \neg\mathsf{O}_\theta(\bigvee \Theta') \mid \Theta' \subseteq \Theta \text{ and } \Theta' \neq \emptyset\}$$

We define the set of abnormalities, Ω^I, as follows:

$$\Omega^I =_{\mathsf{df}} \{\bigvee(\sigma^\beta_\alpha(\Theta)) \mid \Theta \subseteq \mathcal{L}^a, \Theta \neq \emptyset, \Theta \text{ is finite and } \alpha, \beta \in C\}$$

This approach gets rid of our flip-flop problem. Recall our example premise set from above: $\{\mathsf{O}^b_a Pa, \mathsf{O}^c_a \neg Pa, \mathsf{O}^d_a Qa\}$. We could not derive $\mathsf{O}_a Qa$, since there were minimally abnormal models validating $\mathsf{O}^b_a Qa \wedge \neg\mathsf{O}_a Qa$, as every model validated at least one of $\mathsf{O}^b_a(Pa \vee \neg Qa) \wedge \neg\mathsf{O}_a(Pa \vee \neg Qa)$, $\mathsf{O}^c_a(\neg Pa \vee \neg Qa) \wedge \neg\mathsf{O}_a(\neg Pa \vee \neg Qa)$ and $\mathsf{O}^d_a Qa \wedge \neg\mathsf{O}_a Qa$. However, with the new definition of abnormalities, the first two of these three formulas are no longer abnormalities, while the last still is. Thus, models that validate $\mathsf{O}^b_a Qa \wedge \neg\mathsf{O}_a Qa$ are no longer minimally abnormal. Hence, we can derive $\mathsf{O}_a Qa$.

The second problem with the abnormalities of the toy logic (that is not solved by taking Ω^I as abnormalities) was illustrated by the premise set $\{(\exists x)\mathsf{O}^a_x Pb\}$. From this, we would want to be able to derive $(\exists x)\mathsf{O}_x Pb$. However, there are minimally abnormal models of $\{(\exists x)\mathsf{O}^a_x Pb\}$ that do validate $\{(\exists x)\mathsf{O}_x Pb\}$. To solve both the first and second problem of the toy logic, we define the following two sets of abnormalities:

$$\Omega^1 =_{\mathsf{df}} \{(\exists\nu)\bigvee(\sigma^\nu_\alpha(\Theta)) \mid \Theta \subseteq \mathcal{L}^a, \Theta \neq \emptyset, \Theta \text{ is finite, } \alpha \in C \text{ and } \nu \in V\}$$

$$\Omega^2 =_{\mathsf{df}} \{(\exists\nu)(\exists\xi)\bigvee(\sigma^\xi_\nu(\Theta)) \mid \Theta \subseteq \mathcal{L}^a, \Theta \neq \emptyset, \Theta \text{ is finite and } \nu, \xi \in V\}$$

Let $\Omega^{DP} = \Omega^1 \cup \Omega^2$. Models that validate $\{(\exists x)\mathsf{O}^a_x Pb\}$, but not $\{(\exists x)\mathsf{O}_x Pb\}$ are no longer minimally abnormal with these new abnormalities. Hence, $\{(\exists x)\mathsf{O}_x Pb\}$ is derivable.

If we had taken only Ω^2 as our set of abnormalities, then we could not derive $\mathsf{O}_a\varphi$ from $\{\mathsf{O}^b_a\varphi\}$, but only $(\exists x)\mathsf{O}_x\varphi$. By taking only Ω^1, we run into the opposite problem: not being able to derive $(\exists x)\mathsf{O}_x\varphi$ from $\{(\exists x)\mathsf{O}^a_x\varphi\}$. Thus, we need the union of both.

[13] See also [8,19].

Taking the set Ω^{DP} as the abnormalities solves both problems. We can use this set to get a defeasible form of (DP) for any of the logics in Table 3 that do not validate (DP). We simply take the selected monotonic logic from Table 3 as LLL, Ω^{DP} as the set of abnormalities and minimal abnormality as the strategy. Each of these logics is tolerant to the same conflicts as its LLL, but is stronger than its LLL. In particular, it satisfies the second desideratum for (DP): when there are no conflicts, then we are able to apply (DP).

4.4 Adaptive aggregation

To satisfy desideratum 2 for aggregation of directed obligations, we need an adaptive form of this aggregation. A first suggestion might be to use as abnormalities the set of all formulas of the form $\mathsf{O}^{\beta}_{\alpha}\varphi \wedge \mathsf{O}^{\beta}_{\alpha}\psi \wedge \neg\mathsf{O}^{\beta}_{\alpha}(\varphi \wedge \psi)$. However, this leads to similar problems as the two we identified in Section 4.2. Luckily, the solution of these problems is analogous to those presented in Section 4.3 for the adaptive form of (DP).

Again let $\mathcal{L}^a$ be the literals in $\mathcal{L}$ and let $\Theta \subseteq \mathcal{L}^a$ be finite and non-empty.

$$\tau^{\kappa}_{\theta}(\Theta, K) =_{\mathsf{df}} \{\mathsf{O}^{\kappa}_{\theta}(\bigvee \Theta') \wedge \mathsf{O}^{\kappa}_{\theta}(\bigvee K') \wedge \neg\mathsf{O}^{\kappa}_{\theta}(\bigvee \Theta' \wedge \bigvee K') \mid \Theta' \subseteq \Theta, K' \subseteq K \text{ and } \Theta', K' \neq \emptyset\}$$

$$\Omega^1_C =_{\mathsf{df}} \{\bigvee(\tau^{\beta}_{\alpha}(\Theta)) \mid \Theta \subseteq \mathcal{L}^a, \Theta \neq \emptyset, \Theta \text{ is finite and } \alpha, \beta \in C\}$$

$$\Omega^2_C =_{\mathsf{df}} \{(\exists\nu)(\exists\xi)\bigvee(\tau^{\xi}_{\nu}(\Theta)) \mid \Theta \subseteq \mathcal{L}^a, \Theta \neq \emptyset, \Theta \text{ is finite and } \nu, \xi \in V\}$$

$$\Omega^3_C =_{\mathsf{df}} \{(\exists\nu)\bigvee(\tau^{\beta}_{\nu}(\Theta)) \mid \Theta \subseteq \mathcal{L}^a, \Theta \neq \emptyset, \Theta \text{ is finite, } \nu \in V \text{ and } \beta \in C\}$$

$$\Omega^4_C =_{\mathsf{df}} \{(\exists\nu)\bigvee(\tau^{\nu}_{\alpha}(\Theta)) \mid \Theta \subseteq \mathcal{L}^a, \Theta \neq \emptyset, \Theta \text{ is finite, } \alpha \in C \text{ and } \nu \in V\}$$

$$\Omega^C =_{\mathsf{df}} \Omega^1_C \cup \Omega^2_C \cup \Omega^3_C \cup \Omega^4_C$$

The flip-flop problem would already have been solved by only taking Ω^1_C as our set of abnormalities. This solution is analogous to the one in [13, p. 10] and can be seen as an adaptation to aggregation of the solution to the flip-flop problem in Section 4.3.

To solve the second problem we need all four of Ω^1_c-Ω^4_C. Without Ω^4_c we would not be able to derive $(\exists x)\mathsf{O}^x_a(Pa \wedge Qa)$ from $\{(\exists x)(\mathsf{O}^x_a Pa \wedge \mathsf{O}^x_a Qa\}$, i.e. we would not be able to apply aggregation to formulas with existential quantification over the counterparty. Similarly, without Ω^3_C we would not be able to apply aggregation to formulas with existential quantification over the bearer, without Ω^2_C we would have trouble with existential quantification over both the bearer and the counterparty, and without Ω^1_C we would have problems with formulas without existential quantification.

Now we can take any of the monotonic logics from Table 3 that do not validate aggregation for directed obligations and use this logic as the LLL of an adaptive logic. We take Ω^C as the set of abnormalities and minimal abnormality as the strategy. The resulting adaptive logic satisfies desideratum

2 for aggregation and does not suffer from the two problems from Section 4.2. In addition, it is tolerant to the same conflicts as its LLL.

Let us illustrate that desideratum 2 is satisfied by taking **DMN** as the LLL. If we take as premises the set $\{\mathsf{O}^b_a Qa, \mathsf{O}^b_a Pa\}$, then we can derive $\mathsf{O}^b_a(Qa \wedge Pa)$. Any models that do not validate $\mathsf{O}^b_a(Qa \wedge Pa)$ are not minimally abnormal, as they validate the abnormality $\mathsf{O}^b_a Qa, \wedge \mathsf{O}^b_a Pa \wedge \neg \mathsf{O}^b_a(Qa \wedge Pa)$. Similarly, from $\{\mathsf{O}^b_a Qa, \mathsf{O}^b_a Pa, \mathsf{O}^c_a \neg Pa\}$ we can derive $\mathsf{O}^b_a(Qa \wedge Pa)$. All models of the premise set that do not validate $\mathsf{O}^b_a(Qa \wedge Pa)$, do validate the abnormality $\mathsf{O}^b_a Qa \wedge \mathsf{O}^b_a Pa \wedge \neg \mathsf{O}^b_a(Qa \wedge Pa)$ and are therefore not minimally abnormal.[14]

4.5 Combining adaptive (DP) and aggregation

We can also combine adaptive aggregation and adaptive (DP). Take as LLL any logic from Table 3 that does not validate (DP) nor aggregation for directed obligations, take as abnormalities $\Omega^C \cup \Omega^{DP}$ and as a strategy minimal abnormality. The resulting logic is tolerant to both kinds of conflicts (as its LLL is tolerant to both) and satisfies desideratum 2 for both aggregation and (DP). We consider for a moment the strongest of these logics, the one with **DMN** as its LLL. For ease of reference, we will call it $\mathbf{TMDL^m}$.

Imagine an extension of the Manchester Twins case where it is necessary for performing the surgery to prepare surgical equipment, $[\mathsf{U}](Sa \to Ea)$. With $\mathbf{TMDL^m}$ we can derive from this and the premises (i)-(iv) from Section 3.1 that a has an obligation towards Jodie to prepare the surgical equipment: By (UI) and (MP), we can derive $\mathsf{O}^j_a Baj$ from (i). By two applications of inheritance* (see Section 2.4), we first derive $\mathsf{O}^j_a Sa$ and then $\mathsf{O}^j_a Ea$. This seems appropriate for cases of multilateral conflicts where it is not clear which obligation (if any) should be given up. When there is only a multilateral conflict, then we can still derive the obligations that the bearer has towards each counterparty. However, when we decide for some extra-logical reason that the obligation not to perform the surgery prevails, then we might no longer be willing to make this derivation.

5 Conclusion

In this article we distinguished bilateral and multilateral conflicts. We developed a number of monotonic extensions of the term-modal deontic logic **DE**, and showed which of these tolerate what kinds of conflicts. We then noted that none of the conflict-tolerant extensions validate aggregation of directed obligations with the same bearer, or the derivation of undirected obligations from directed obligations. They did not even validate these for obligations that were not involved in any conflict. Since these principles are arguably plausible, we developed defeasible versions of (DP) and aggregation. This allows us to construct non-monotonic logics that validate these principles as much as possible.

All of this gives us a broad range of logics that tolerate bilateral and multilateral conflicts. Whatever combination of the principles discussed in Section

[14] Naturally, any logics whose LLL validates (DP) will trivialise this last premise set.

2.4 one finds plausible can be used to construct a conflict-tolerant logic. If one finds inheritance, necessitation, both or neither plausible, then one can use the conflict-tolerant logic in Table 3 that validates exactly these principles. If, in addition, one finds aggregation, (DP) or both plausible, then one can add these as defeasible principles, as was described in Section 4. We see this as the main result of the present paper.

This opens the door to different avenues of future research. One can also consider conflicts between directed or undirected obligations with different bearers. Another option is to involve impersonal obligations, i.e. obligations not tied to any bearer or counterparty. One could ask whether, in a conflict between an impersonal obligation and an undirected obligation, one of the kinds prevails over the other.

It is also possible to consider more involved formalisations of general rules, based on a more in depth account of conditional obligations. For this article we have used the material implication to interpret both general rules and conditional obligations. However, this does not take into account the possibility of exceptions to general rules, nor the problem of contrary-to-duty obligations. Integrating a richer account of conditional obligations with term-modal deontic logics might open the way to different conflict-tolerant deontic logics. This is especially interesting in conjunction with a point made in Section 4.5, that in the case of a multilateral conflict, $\mathbf{TMDL^m}$ allows one to derive the obligations of every bearer-counterparty pair as if there was no conflict. [15]

References

[1] Batens, D., *Dialectical dynamics within formal logics*, Logique et Analyse **29** (1986), pp. 161–173.

[2] Batens, D., *Dynamic dialectical logics*, Paraconsistent Logic. Essays on the Inconsistent (1989), pp. 187–217.

[3] Batens, D., *A universal logic approach to adaptive logics*, Logica Universalis **1** (2007), pp. 221–242.

[4] Fitting, M., L. Thalmann and A. Voronkov, *Term-modal logics*, Studia Logica **69** (2001), pp. 133–169.
URL `https://doi.org/10.1023/A:1013842612702`

[5] Frijters, S., J. Meheus and F. Van De Putte, *Representing reasoning with rules and rights: Term-modal deontic logic*, in: S. Rahman, M. Armgardt and H. C. Nordtveit Kvernenes, editors, *New Systematic and Historic Studies in Legal Reasoning and Logic*, In Press .

[6] Frijters, S. and F. Van De Putte, *Classical term-modal logics* (Submitted).

[7] Goble, L., *Prima facie norms, normative conflicts, and dilemmas*, Handbook of deontic logic and normative systems **1** (2013), pp. 241–351.

[8] Goble, L., *Deontic logic (adapted) for normative conflicts*, Logic Journal of the IGPL **22** (2014), pp. 206–235.

[9] Governatori, G. and A. Rotolo, *On the axiomatisation of Elgesem's logic of agency and ability*, Journal of Philosophical Logic (2005), pp. 403–431.

[10] Herrestad, H. and C. Krogh, *Deontic logic relativised to bearers and counterparties*, Anniversary Anthology in Computers and Law (1995), pp. 453–522.

[15] The authors would like to thank Frederik Van De Putte, Joke Meheus, Federico Faroldi, Nathan Wood and two anonymous reviewers for their valuable feedback and suggestions.

[11] Kaveny, M. C., *Conjoined twins and catholic moral analysis: Extraordinary means and casuistical consistency*, Kennedy Institute of Ethics Journal **12** (2002), pp. 115–140.

[12] McNamara, P., *Agential obligation as non-agential personal obligation plus agency*, Journal of Applied Logic **2** (2004), pp. 117–152.

[13] Meheus, J., M. Beirlaen, F. Van De Putte and C. Straßer, *Non-adjunctive deontic logics that validate aggregation as much as possible* (2010), unpublished.

[14] Paris, J. J. and A. C. Elias-Jones, *"Do we murder Mary to save Jodie?" an ethical analysis of the separation of the Manchester conjoined twins*, Postgraduate Medical Journal **77** (2001), pp. 593–598.
URL `http://pmj.bmj.com/content/77/911/593`

[15] Parker, M. and D. Dickenson, "The Cambridge medical ethics workbook: Case studies, commentaries and activities," Cambridge University Press, 2001.

[16] Straßer, C., "Adaptive Logics for Defeasible Reasoning," Springer, 2014.

[17] Thalmann, L., "Term-modal logic and quantifier-free dynamic assignment logic," Ph.D. thesis, Acta Universitatis Upsaliensis (2000).

[18] Thomason, R. H., *Some completeness results for modal predicate calculi*, in: K. Lambert, editor, *Philosophical Problems in Logic*, Synthese Library **29**, Springer Netherlands, 1970 pp. 56 – 76.

[19] Van De Putte, F., M. Beirlaen and J. Meheus, *Adaptive deontic logics: a survey*, Journal of Applied Logics - IfCoLog Journal of Logics and their Applications (2019).

[20] Van De Putte, F. and C. Straßer, *A Logic for prioritized normative reasoning*, Journal of Logic and Computation **23** (2012), pp. 563–583.
URL `https://doi.org/10.1093/logcom/exs008`