# A Paraconsistent Multi-Agent Framework for Dealing with Normative Conflicts

Mathieu Beirlaen and Christian Straßer

Centre for Logic and Philosophy of Science
University of Ghent, Belgium
{Mathieu.Beirlaen,Christian.Strasser}@UGent.be

# A Paraconsistent Multi-Agent Framework for Dealing with Normative Conflicts[⋆]

Mathieu Beirlaen and Christian Straßer

Centre of Logic and Philosophy of Science, Ghent University
{Mathieu.Beirlaen,Christian.Strasser}@UGent.be

**Abstract.** In a multi-agent deontic setting, normative conflicts can take a variety of different logical forms. In this paper, we present a very general characterization of such conflicts, including both intra- and inter-agent normative conflicts, conflicts between groups of agents, conflicts between obligations and permissions, and conflicts between contradictory norms. In order to account for the consistent possibility of this wide variety of conflict-types, we present a paraconsistent deontic logic, i.e. a logic that invalidates the classical principle of non-contradiction. Next, we strengthen this logic within the adaptive logics framework for defeasible reasoning. The resulting inconsistency-adaptive deontic logic interprets a given set of norms 'as consistently as possible'.

## 1 Introduction

The development of systems capable of tolerating conflicting norms is considered an important challenge in the fields of deontic logic [14] and normative multi-agent systems [7]. In this paper, we try to meet this challenge by consistently allowing for various types of normative conflicts within a non-classical multi-agent framework, i.e. a multi-agent framework that invalidates some rules and theorems of Standard Deontic Logic (**SDL**).

For reasons of presentation we will first introduce a classical variant of the framework (Section 2), and illustrate how the resulting logic **MDC** treats individual and collective obligations. Next, we present a subdivision of various types of normative conflicts (Section 3), and show that **MDC** cannot consistently allow for the possibility of such conflicts.

In order to prevent instances of normative conflicts from giving rise to deontic explosion, we introduce a paraconsistent variant of the logic **MDC**: the logic **MDP** (Section 4). As opposed to **MDC**, **MDP** can consistently deal with any type of normative conflict. However, the conflict-tolerance of **MDP** comes at a price. Since **MDP** gives up some of the rules validated by **SDL** (and hence by **MDC**), it loses much of the latter system's inferential power. This drawback is

[⋆] Research for this paper was supported by subventions from the Research Foundation – Flanders (FWO Vlaanderen). The authors are indebted to the anonymous referees for valuable comments and suggestions.

common to any monotonic paraconsistent deontic logic presented so far (Section 5).

The solution to this problem presented here consists of extending **MDP** within the adaptive logics framework [4]. In the resulting logic (called $\mathbf{MDP^m}$), some **MDC**-inferences are made conditional upon the behavior of the premises: $\mathbf{MDP^m}$ verifies only those inferences which rely on premises that can safely be assumed to behave 'normally'. The technically precise sense in which $\mathbf{MDP^m}$ does so is spelled out in Section 6. $\mathbf{MDP^m}$ has the nice property that for premise sets all members of which behave 'normally' in this sense, $\mathbf{MDP^m}$ delivers the same consequences as **MDC**.

## 2 A simple classical multi-agent framework

### 2.1 Language

We use a denumerable set $\mathcal{W}^a$ of propositional constants (atoms) $p, q, r, \ldots$. The $\langle \neg, \wedge, \vee, \supset, \equiv \rangle$-closure of $\mathcal{W}^a$ is denoted by $\mathcal{W}$. We call formulas in $\mathcal{W}$ (purely) propositional formulas.

Next to propositional formulas, we use a finite set $I = \{i_1, \ldots, i_n\}$ of agents. We will in the remainder often refer to groups of agents $J$ in $I$, i.e. non-empty subsets of $I$. The following notation is useful for this: $J \subseteq_\emptyset I$ iff $J \neq \emptyset$ and $J \subseteq I$. The set $\mathcal{W}_I = \{\langle A, J \rangle \mid A \in \mathcal{W}, J \subseteq_\emptyset I\}$ denotes the set of agent-proposition pairs. Throughout the paper, we will use "$A_J$" as a shortcut for "$\langle A, J \rangle$." Where $i \in I$, we will in the remainder of the paper abbreviate $A_{\{i\}}$ by $A_i$. A formula $A_J \in \mathcal{W}_I$ is translated as "group $J$ brings about $A$ by a joint effort". We will discuss and distinguish this notion from another, weaker reading of group obligations in Section 2.3. $\mathcal{W}_I^c$ is the set of all formulas in the $\langle \neg, \wedge, \vee, \supset, \equiv \rangle$-closure of $\mathcal{W}_I$. Where $\mathcal{W}^l$ denotes the set of literals (i.e. the set of atoms in $\mathcal{W}^a$ and their negations), we also define the set $\mathcal{W}_I^l = \{A_J \mid A \in \mathcal{W}^l, J \subseteq_\emptyset I\}$ of agent-literal complexes. Finally, the set $\mathcal{W}^c$ of well-defined formulas for the classical multi-agent framework is defined recursively as follows:

$$\mathcal{W}^c := \langle \mathcal{W} \cup \mathcal{W}_I \rangle \mid \mathsf{O}\langle \mathcal{W}_I^c \rangle \mid \mathsf{P}\langle \mathcal{W}_I^c \rangle \mid \neg \langle \mathcal{W}^c \rangle \mid \langle \mathcal{W}^c \rangle \wedge \langle \mathcal{W}^c \rangle \mid \\ \langle \mathcal{W}^c \rangle \vee \langle \mathcal{W}^c \rangle \mid \langle \mathcal{W}^c \rangle \supset \langle \mathcal{W}^c \rangle \mid \langle \mathcal{W}^c \rangle \equiv \langle \mathcal{W}^c \rangle$$

Where $A \in \mathcal{W}_I^c$, a formula $\mathsf{O}A$ [$\mathsf{P}A$] is interpreted as "it ought to be [is permitted] that $A$". Hence, $\mathsf{O}A_i$ is read as "It ought to be that agent $i$ brings about $A$". Similarly, $\mathsf{O}A_J$ is read as "The group of agents $J$ ought to bring about $A$ by a joint effort". We do not allow for formulas $\mathsf{O}A$ and $\mathsf{P}A$ where $A \in \mathcal{W}^c \setminus \mathcal{W}_I^c$ such as $\mathsf{O}B$ where $B$ is a propositional atom. This is because we are only interested in obligations that are addressed directly to (groups of) agents. Note that we do allow for formulas such as $\mathsf{O}(A_i \vee B_j)$ and $\mathsf{O}A_i \vee \mathsf{O}B_j$. While the former expresses that it ought to be the case that either $i$ brings about $A$ or that $j$ brings about $B$, the latter expresses that either $i$ ought to bring about $A$ or $j$ ought to bring about $B$. This difference corresponds to the distinction in **SDL** between the formulas $\mathsf{O}(A \vee B)$ and $\mathsf{O}A \vee \mathsf{O}B$.

There is another subtlety worth pointing out, namely the difference between $\mathsf{O}\neg(A_i)$ and $\mathsf{O}(\neg A)_i$. While the latter indicates $i$'s obligation to bring about $\neg A$, the former is literally read as "It ought to be that it is not the case that $i$ brings about $A$". This can be understood as $i$'s obligation to refrain from bringing about $A$.

### 2.2 The logic MDC

In this section we present a classical system for modeling normative reasoning. We presuppose that (i) norms dealt with by this system arise from the same source, and (ii) agents have epistemic access to *all* norms issued by this source.

Let us demonstrate how to adjust the Kripke-frames that are usually used in order to semantically characterize **SDL** to the multi-agent setting of **MDC**. We shortly outline some of the basic ideas. An **SDL**-model is a tuple $M = \langle W, R, v, w^0 \rangle$. $W$ is a set of worlds where each world is associated with a set of atoms by the assignment function $v : \mathcal{W}^a \to \wp(W)$. A propositional atom $A$ is said to hold in a world $w$ iff it is assigned to the world by $v$, i.e. $w \in v(A)$. The validity of complex formulas is then recursively defined as usual. $R \subseteq W \times W$ is a serial accessibility relation. A formula $A$ is obliged in a world $w$ iff it is valid in all the accessible worlds of $w$. Moreover, $w^0 \in W$ is the so-called actual world.

Let us now step-by-step generalize these frames for the multi-agent setting. First we need to introduce agents. We represent them by a finite non-empty set $I = \{i_1, \ldots, i_n\}$. An **MDC**-model is a tuple $M = \langle W, I, R, v, v_I, w^0 \rangle$ where as before $W$ is a set of worlds, $R \subseteq W \times W$ is a serial accessibility relation, $v : \mathcal{W}^a \to \wp(W)$ is an assignment function, and $w^0 \in W$ is the actual world. Just as before, the idea is that the propositional atom $A$ is the case in $w$ iff $w \in v(A)$.

We are not only interested in what is the case in our worlds, but also in causation, more precisely the question which agents cause certain events. In order to express this, our worlds are not just points, such as in the case of the **SDL**-semantics, but they are structured. Every world $w \in W$ is associated with tuples $\langle w, J \rangle$, for all $J \subseteq_{\emptyset} I$. We use $w_J$ as a shortcut for $\langle w, J \rangle$.

While in **SDL** the assignment $v$ associates a world $w \in W$ with atoms in order to express what atoms hold in $w$, we add now an additional assignment $v_I$ that associates each $w_J$ with literals in order to express what literals are brought about by the group of agents $J$. The idea is that a literal $A$ is brought about in $w$ by a group of agents $J$ iff $w_J \in v(A)$. Hence, $v_I : \mathcal{W}^l \to \wp(\{w_J \mid w \in W, J \subseteq_{\emptyset} I\})$.

$v$ associates only atoms (and not literals) with worlds because this provides enough information to uniquely define whether a complex propositional formula representing factual information holds in a world. We for instance do not need to assign worlds to negated propositional atoms such as $\neg A$, since by means of a semantic clause such as the following it can be determined whether $\neg A$ holds in a world $w$: (†) "$\neg A$ holds in $w$ in a model $M$ iff $A$ does not hold in $w$ in $M$". Note that, in order to determine whether $J$ brings about $\neg A$ in $w$, we cannot rely on the fact that $J$ does not bring about $A$. After all, from the fact that $J$ refrains from bringing about $A$ we cannot infer that $J$ brings about $\neg A$. The fact that $A$ or $\neg A$ holds in a world may be independent of actions by $J$. Hence, we need to specify for each literal by what group of agents it is brought about (if any).

In the **SDL**-semantics the clause (†) ensures that the worlds are *consistent* in the sense that it is not the case that for an atom $A$, $A$ holds in a world $w$ and at the same time $\neg A$ holds in the world $w$. Since $v_I$ associates worlds with both atoms and their negations we need to ensure the consistency by a frame-condition:

**F-Con** For all $A \in \mathcal{W}^a$, for all $w \in W$, and for all $J, K \subseteq_\emptyset I$: (i) if $w_J \in v_I(A)$ then $w_K \notin v_I(\neg A)$ and (ii) if $w_J \in v_I(\neg A)$ then $w_K \notin v_I(A)$.

Moreover, we want to ensure that whenever an agent or group brings about $A$, then $A$ is also the case (factually). This is guaranteed by adding the following frame condition:

**F-Fac** For all $A \in \mathcal{W}^a$ and all $w \in W$, (i) if $w_J \in v_I(A)$ then $w \in v(A)$ and (ii) if $w_J \in v_I(\neg A)$ then $w \notin v(A)$.

The valuation $v_M : \mathcal{W}^c \to W$ associated with the model $M$ is defined by:

$\mathsf{C}_I^l$ where $A_J \in \mathcal{W}_I^l : M, w \models A_J$ iff $w_J \in v_I(A)$

$\mathsf{C}_I\wedge$ where $A, B \in \mathcal{W} : M, w \models (A \wedge B)_J$ iff $M, w \models A_J$ and $M, w \models B_J$

$\mathsf{C}_I\vee$ where $A, B \in \mathcal{W} : M, w \models (A \vee B)_J$ iff $M, w \models A_J$ or $M, w \models B_J$

$\mathsf{C}_I\supset$ where $A, B \in \mathcal{W} : M, w \models (A \supset B)_J$ iff $M, w \not\models A_J$ or $M, w \models B_J$

$\mathsf{C}_I\equiv$ where $A, B \in \mathcal{W} : M, w \models (A \equiv B)_J$ iff $(M, w \models A_J$ iff $M, w \models B_J)$

$\mathsf{C}_I\neg\neg$ where $A \in \mathcal{W} : M, w \models (\neg\neg A)_J$ iff $M, w \models A_J$

$\mathsf{C}_I\neg\vee$ where $A, B \in \mathcal{W} : M, w \models (\neg(A \vee B))_J$ iff $M, w \models (\neg A \wedge \neg B)_J$

$\mathsf{C}_I\neg\wedge$ where $A, B \in \mathcal{W} : M, w \models (\neg(A \wedge B))_J$ iff $M, w \models (\neg A \vee \neg B)_J$

$\mathsf{C}_I\neg\supset$ where $A, B \in \mathcal{W} : M, w \models (\neg(A \supset B))_J$ iff $M, w \models (A \wedge \neg B)_J$

$\mathsf{C}_I\neg\equiv$ where $A, B \in \mathcal{W} : M, w \models (\neg(A \equiv B))_J$ iff $M, w \models ((A \vee B) \wedge (\neg A \vee \neg B))_J$

$\mathsf{C}^a$ where $A \in \mathcal{W}^a : M, w \models A$ iff $w \in v(A)$

$\mathsf{C}\neg$ where $A \in \mathcal{W}^c : M, w \models \neg A$ iff $M, w \not\models A$

$\mathsf{C}\wedge$ where $A, B \in \mathcal{W}^c : M, w \models A \wedge B$ iff $M, w \models A$ and $M, w \models B$

$\mathsf{C}\vee$ where $A, B \in \mathcal{W}^c : M, w \models A \vee B$ iff $M, w \models A$ or $M, w \models B$

$\mathsf{C}\supset$ where $A, B \in \mathcal{W}^c : M, w \models A \supset B$ iff $M, w \not\models A$ or $M, w \models B$

$\mathsf{C}\equiv$ where $A, B \in \mathcal{W}^c : M, w \models A \equiv B$ iff $(M, w \models A$ iff $M, w \models B)$

$\mathsf{CO}$ where $A \in \mathcal{W}_I^c : M, w \models \mathsf{O}A$ iff $M, w' \models A$ for all $w'$ such that $Rww'$

$\mathsf{CP}$ where $A \in \mathcal{W}_I^c : M, w \models \mathsf{P}A$ iff $M, w' \models A$ for some $w'$ such that $Rww'$

An **MDC**-model $M$ verifies $A \in \mathcal{W}^c$ ($M \Vdash_{\mathbf{MDC}} A$) iff $M, w^0 \models A$. Where $\Gamma \subseteq \mathcal{W}^c$, $M$ is an **MDC**-model of $\Gamma$ iff $M$ is an **MDC**-model and $M \Vdash_{\mathbf{MDC}} A$ for all $A \in \Gamma$. Moreover, $\models_{\mathbf{MDC}} A$ iff all **MDC**-models verify $A$, and $\Gamma \models_{\mathbf{MDC}} A$ iff all **MDC**-models of $\Gamma$ verify $A$.

All of the following inferences are valid in **MDC** (where $A, B \in \mathcal{W}_I^c$):

$$\begin{aligned} \mathsf{O}A, \mathsf{O}B &\models_{\mathbf{MDC}} \mathsf{O}(A \wedge B) \\ \mathsf{O}A &\models_{\mathbf{MDC}} \neg\mathsf{O}\neg A \\ \mathsf{O}(A \vee B), \mathsf{O}\neg A &\models_{\mathbf{MDC}} \mathsf{O}B \end{aligned}$$

## 2.3 More on group obligations

Where $i, j \in I$, the formula $\mathsf{O}A_{\{i,j\}}$ abbreviates a collective obligation for group $\{i, j\}$ to bring about $A$. Note that none of $\mathsf{O}A_i$, $\mathsf{O}A_j$, $\mathsf{O}A_i \vee \mathsf{O}A_j$, and $\mathsf{O}(A_i \vee A_j)$ is **MDC**-derivable from $\mathsf{O}A_{\{i,j\}}$. This is due to the fact that $\mathsf{O}A_{\{i,j\}}$ expresses that $i$ and $j$ should bring about $A$ by a joint effort. Collective obligations of this kind are called *strict collective obligations* by Dignum & Royakkers [11]. A strict collective obligation to bring about $A$ is satisfied only if *all* agents in the collective bring about $A$ *together*.

Not all collective obligations are strict collective obligations. Suppose, for instance, that a mother of three children orders her offspring to do the dishes. In order to satisfy this obligation, it might not matter if only one or two of the children actually do the dishes. All that matters is that, in the end, the dishes are clean. The obligation issued by this agent is hence not a strict collective obligation. It is what Dignum & Royakkers call a *weak collective obligation.* A weak collective obligation to bring about $A$ is satisfied as soon as any subset of the collective brings about $A$.

Although the formula $\mathsf{O}A_J$ is in **MDC** interpreted as a strict collective obligation, we can also define an obligation operator $\mathsf{O}^{\mathsf{w}}$ in order to express weak collective obligations:

$\mathsf{O}^{\mathsf{w}}A_J =_{\mathrm{df}} \mathsf{O}(\bigvee_{K \subseteq_{\emptyset} J} A_K)$

The weak collective obligation operator $\mathsf{O}^{\mathsf{w}}$ captures the intended meaning that if a group of agents ought to bring about a certain state of affairs, then this state of affairs ought to be brought about by some subset of the group.[1] It follows immediately by the definition and $\mathrm{C}_I\vee$ that $\models_{\mathbf{MP}} \mathsf{O}A_J \supset \mathsf{O}^{\mathsf{w}}A_J$. Obviously, if the group $J$ faces the strict collective obligation to bring about $A$, then some subgroup of $J$ –namely $J$ itself– has to bring about $A$. Note that $\mathsf{O}^{\mathsf{w}}A_i = \mathsf{O}A_i$.

The disambiguation of the notion of collective obligation by means of the distinction between strict and weak collective obligations allows us to further illustrate some subtle differences in **MDC**. Suppose that some agent $i$ ought to bring about $\neg A$, whereas agents $i$ and $j$ ought to bring about $A \vee B$. If the latter obligation is interpreted as a strict collective obligation, then it is **MDC**-derivable that $i$ and $j$ share the strict collective obligation to bring about $B$:

(1) $\mathsf{O}(\neg A)_i, \mathsf{O}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDC}} \mathsf{O}B_{\{i,j\}}$

[1] The $\mathsf{O}^{\mathsf{w}}$-operator as defined here is slightly different from the one defined by Dignum & Royakkers in [11]. We write the latter operator as $\mathsf{O}_{\mathsf{w}}$. Then $\mathsf{O}_{\mathsf{w}}A_J =_{\mathrm{df}} \bigvee_{K \subseteq_{\emptyset} J} \mathsf{O}A_K$. Note that $\mathsf{O}^{\mathsf{w}}A_{\{a,b\}} = \mathsf{O}(A_a \vee A_b \vee A_{\{a,b\}})$, while $\mathsf{O}_{\mathsf{w}}A_{\{a,b\}} = \mathsf{O}A_a \vee \mathsf{O}A_b \vee \mathsf{O}A_{\{a,b\}}$. We prefer to define weak obligation in terms of $\mathsf{O}^{\mathsf{w}}$ instead of $\mathsf{O}_{\mathsf{w}}$ because we take a weak (collective) obligation to be a single norm rather than a disjunction of norms.

In general, if some group faces a strict collective obligation, then it should try to satisfy this obligation in a way that conflicting obligations are avoided whenever possible. This is exactly what happens in the above example.

If we interpret $i$ and $j$'s obligation to bring about $A \vee B$ as a weak collective obligation, then $\mathsf{O}B_{\{i,j\}}$ is no longer **MDC**-derivable, but the weaker obligation $\mathsf{O}^{\mathsf{w}}B_{\{i,j\}}$ is:

(2) $\mathsf{O}(\neg A)_i, \mathsf{O}^{\mathsf{w}}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDC}} \mathsf{O}^{\mathsf{w}}B_{\{i,j\}}$

Again, conflicting obligations are neatly avoided: $i$ and $j$'s weak obligation to bring about $A \vee B$ is satisfied in a consistent way whenever $i$, $j$, or $i$ and $j$ together bring about $B$.

If instead of supposing that $i$ has the obligation to bring about $\neg A$, we suppose that $i$ merely has the obligation to refrain from bringing about $A$, the above reasoning no longer applies:

(3) $\mathsf{O}\neg(A_i), \mathsf{O}(A \vee B)_{\{i,j\}} \not\models_{\mathbf{MDC}} \mathsf{O}B_{\{i,j\}}$

That $i$ ought to refrain from bringing about $A$, does not entail that the group $\{i,j\}$ ought to do so.[2] Hence there is no strict obligation for $\{i,j\}$ to bring about $B$. In the variant for weak collective obligation, a similar reasoning applies:

(4) $\mathsf{O}\neg(A_i), \mathsf{O}^{\mathsf{w}}(A \vee B)_{\{i,j\}} \not\models_{\mathbf{MDC}} \mathsf{O}^{\mathsf{w}}B_{\{i,j\}}$

That $i$ should refrain from bringing about $A$ does not allow us to derive a weak collective obligation for $i$ and $j$ to bring about $B$, because $\mathsf{O}^{\mathsf{w}}(A \vee B)_{\{i,j\}}$ is also satisfied if, for instance, $j$ brings about $A$ or if $i$ and $j$ together (in the strict sense) bring about $A$.

## 3 Normative conflicts

In single-agent settings, normative conflicts (moral conflicts, deontic conflicts) are usually conceived as situations in which an agent has two (or more) conflicting obligations. In the language of **MDC**, such *intra-agent* conflicts between obligations can have two logical forms. Where the agent in question is represented by the subscript $i$, we say that $i$ faces an obligation-obligation conflict (in short, an $\mathsf{OO}$-*conflict*) if, for some $A$, either $\mathsf{O}A_i \wedge \mathsf{O}(\neg A)_i$ or $\mathsf{O}A_i \wedge \mathsf{O}\neg(A_i)$. In the first case, $i$ has both an obligation to bring about $A$ and an obligation to bring about $\neg A$. In the second case, $i$ has both an obligation to bring about $A$ and an obligation to refrain from bringing about $A$. Similarly, a group of agents $J$ faces an $\mathsf{OO}$-conflict if $\mathsf{O}A_J \wedge \mathsf{O}(\neg A)_J$ or if $\mathsf{O}A_J \wedge \mathsf{O}\neg(A_J)$.

In a multi-agent setting, we have to allow for the possibility of *inter-agent* conflicts next to intra-agent conflicts. Conflicts of obligations between *different* (groups of) agents can arise only in case one of the agents or groups, say $J$, has to bring about a state of the world inconsistent with a state of the world that should be brought about by another agent or group, say $K$, i.e. if $\mathsf{O}A_J \wedge \mathsf{O}(\neg A)_K$.

[2] Suppose, for example, that $i$ has to refrain from lifting a heavy cupboard (because, for instance, $i$ has chronic back pain). From this it does not follow that $i$ should still refrain from doing so if she is assisted by $j$.

Note that if $J \neq K$, a formula $\mathsf{O}A_J \wedge \mathsf{O}\neg(A_K)$ no longer guarantees a conflict of obligations in the multi-agent setting: it is perfectly possible that agent or group $J$ brings about $A$, while another agent or group $K$ refrains from bringing about $A$. Altogether, in a multi-agent framework an $\mathsf{OO}$-conflict has one of the following two logical forms: $\mathsf{O}A_J \wedge \mathsf{O}\neg(A_J)$ or $\mathsf{O}A_J \wedge \mathsf{O}(\neg A)_K$ (where possibly $J = K$).

Logicians often limit their study of normative conflicts to conflicts between two or more obligations, e.g. [13, 15, 17, 18, 25]. However, other types of normative conflicts can occur. It might, for instance, be the case that an agent or group $J$ ought to bring about $A$, while $J$ is also permitted to refrain from doing so, i.e. $\mathsf{O}A_J \wedge \mathsf{P}\neg(A_J)$. Moreover, $J$ might have the obligation to bring about $A$ while a possibly different group or agent $K$ is permitted to bring about $\neg A$, i.e. $\mathsf{O}A_J \wedge \mathsf{P}(\neg A)_K$. In what follows such conflicts will be called obligation-permission conflicts or $\mathsf{OP}$-*conflicts*. For some examples of $\mathsf{OP}$-conflicts in a single-agent setting, see [6, 30]. The possibility of $\mathsf{OP}$-conflicts was also defended in [1, 2, 8, 35].

In [6, 24] examples were given of *contradicting norms*. Suppose, for instance, that in some country the constitution contains the following clauses concerning parliamentary elections: (i) it is not the case that women are permitted to vote, and (ii) property holders are permitted to vote. Suppose further that (possibly due to a recent revision of the property law) women are allowed to hold property. Then the law is inconsistent: any female property holder $i$ is both permitted and not permitted to vote: $\mathsf{P}V_i \wedge \neg\mathsf{P}V_i$ (example from [24, pp. 184-185]).

The same reasoning holds, of course, for formulas of the form $\mathsf{O}A \wedge \neg\mathsf{O}A$ (where $A \in \mathcal{W}_I^c$). As hinted at above, normative conflicts of the type $\mathsf{P}A \wedge \neg\mathsf{P}A$ or $\mathsf{O}A \wedge \neg\mathsf{O}A$ are called contradicting norms.

Next to contradicting norms, i.e. different norms that contradict *each other*, one might also face a *contradictory norm*, i.e. a norm that contradicts *itself*. A contradictory norm is of the form $\mathsf{O}(A_J \wedge \neg(A_J))$, $\mathsf{P}(A_J \wedge \neg(A_J))$, $\mathsf{O}(A_J \wedge (\neg A)_K)$, $\mathsf{P}(A_J \wedge (\neg A)_K)$, $\mathsf{O}(A \wedge \neg A)_J$, or $\mathsf{P}(A \wedge \neg A)_J$. For a defense of contradictory norms, we refer to [24].

Unfortunately, none of the types of normative conflicts presented above can be dealt with consistently by the logic **MDC**. **MDC** trivializes all instances of all types of normative conflicts. This gives rise to what is usually called *deontic explosion*: the fact that from a deontic conflict any obligation follows. See [13, 30] for a more detailed discussion of this phenomenon in deontic logic. An oversight of the various types of normative conflicts and their accompanying principles of deontic explosion is provided in the table below. Where $A \in \mathcal{W}$ and $B, C \in \mathcal{W}_I^c$:

| | | |
|---|---|---|
| OO-conflicts: | $\mathsf{O}A_J \wedge \mathsf{O}\neg(A_J) \models \mathsf{O}C$, | $\mathsf{O}A_J \wedge \mathsf{O}(\neg A)_K \models \mathsf{O}C$ |
| OP-conflicts: | $\mathsf{O}A_J \wedge \mathsf{P}\neg(A_J) \models \mathsf{O}C$, | $\mathsf{O}A_J \wedge \mathsf{P}(\neg A)_K \models \mathsf{O}C$ |
| Contradicting norms: | $\mathsf{O}B \wedge \neg\mathsf{O}B \models \mathsf{O}C$, | $\mathsf{P}B \wedge \neg\mathsf{P}B \models \mathsf{O}C$ |
| Contradictory norms: | $\mathsf{O}(A_J \wedge \neg(A_J)) \models \mathsf{O}C$, | $\mathsf{P}(A_J \wedge \neg(A_J)) \models \mathsf{O}C$, |
| | $\mathsf{O}(A_J \wedge (\neg A)_K) \models \mathsf{O}C$, | $\mathsf{P}(A_J \wedge (\neg A)_K) \models \mathsf{O}C$, |
| | $\mathsf{O}(A \wedge \neg A)_J \models \mathsf{O}C$, | $\mathsf{P}(A \wedge \neg A)_J \models \mathsf{O}C$ |

## 4 Avoiding deontic explosion: the logic MDP

Since **MDC** causes explosion when faced with a normative conflict, and since we want to allow for the consistent possibility of normative conflicts, we need a logic that is weaker than **MDC**.[3] The situation is analogous in non-agentive settings. There too, **SDL** gives rise to explosion in view of formulas of the form $\mathsf{O}A \wedge \mathsf{O}\neg A$, $\mathsf{O}A \wedge \mathsf{P}\neg A$, etc. And there too, authors have suggested weakening the logic in order to tolerate normative conflicts; for some examples, see [12, 21, 25, 28, 31, 32]. A good oversight can be found in [13].

The solution presented here is to replace the classical negation operator by a weaker negation operator that renders invalid the *Ex Contradictione Quodlibet* principle (ECQ), i.e. $A \wedge \neg A \models B$. One of the main reasons for invalidating ECQ in deontic logic is that it is the only possible solution for consistently allowing for contradicting norms.

Logics that invalidate ECQ are usually called *paraconsistent* logics. In a single-agent deontic setting, paraconsistent deontic logics have been presented in [6, 10, 24]. To the best of our knowledge, this solution was never before used in a multi-agent deontic setting.

The logic obtained by replacing the classical negation of **MDC** by a weaker, paraconsistent negation is called **MDP**.[4]

Since we want **MDP** to invalidate all explosion principles from the table in Section 3, frame condition **F-Con** must be given up. In **MDC**, **F-Con** excludes accessible worlds which validate both $A_J$ and $(\neg A)_K$ for some $J$ and $K$. Hence this condition immediately trivializes e.g. normative conflicts of the form $\mathsf{O}A_J \wedge \mathsf{O}(\neg A)_K$ or $\mathsf{O}A_J \wedge \mathsf{P}(\neg A)_K$. Thus if we want to consistently allow for all types of normative conflicts, **F-Con** must be rejected.

Giving up **F-Con** takes us one step closer towards a conflict-tolerant deontic logic. However, even if **F-Con** is rejected, triviality still ensues in view of e.g. conflicts of the form $\mathsf{O}A_J \wedge \mathsf{O}\neg(A_J)$ or $\mathsf{O}A_J \wedge \mathsf{P}\neg(A_J)$. Hence more work is needed in order to make the new logic fully conflict-tolerant, i.e. in order to invalidate all explosion principles stated for **MDC** in the table in Section 3.

Analogous to **MDC**-models, **MDP**-models are tuples $\langle W, I, R, v, v_I, w^0 \rangle$. The only difference is that the factual assignment $v$ is now defined more broadly, i.e. $v : \mathcal{W}^l \cup \{\neg(A_J) \mid A \in \mathcal{W}^l\} \rightarrow \wp(W)$. Moreover we remove the **MDC**-frame condition **F-Con**, and replace **F-Fac** with **F-Fac′**:

[3] Some authors circumnavigate the problems posed by normative conflicts by making their formal system more expressive rather than by weakening its axioms or rules. For instance, in [18] Kooi & Tamminga add super- and subscripts to the deontic operators in order to express the source and the interest group in view of which a norm holds. However, in their system explosion still ensues when faced with conflicting norms that hold for the same source and interest group. Such 'hardcore' normative conflicts are sometimes called *symmetrical* conflicts [20, 27]. In order to consistently allow for the possibility of these conflicts in deontic logic, we need a non-standard formalism, i.e. a formalism that invalidates one or more of the theorems and rules of **SDL**.

[4] The negation of **MDP** is that of the paraconsistent logic **CLuNs** as found in e.g. [3, 5].

**F-Fac′** For all $A \in \mathcal{W}^a$, all $w \in W$ and all $J \subseteq_\emptyset I$, (i) if $w_J \in v_I(A)$ then $w \in v(A)$ and (ii) if $w_J \in v_I(\neg A)$ then $w \notin v(A)$ or $w \in v(\neg A)$.

The valuation $v_M : \mathcal{W}^c \to W$ is defined by $\mathrm{C}^l_I$, $\mathrm{C}_I\wedge$, $\mathrm{C}_I\vee$, $\mathrm{C}_I\supset$, $\mathrm{C}_I\equiv$, $\mathrm{C}_I\neg\neg$, $\mathrm{C}_I\neg\vee$, $\mathrm{C}_I\neg\wedge$, $\mathrm{C}_I\neg\supset$, $\mathrm{C}_I\neg\equiv$, $\mathrm{C}^a$, $\mathrm{C}\wedge$, $\mathrm{C}\vee$, $\mathrm{C}\supset$, $\mathrm{C}\equiv$, $\mathrm{C}\bot$, $\mathrm{C}\mathsf{O}$, $\mathrm{C}\mathsf{P}$, and the following:

$C\neg'$ where $A \in \mathcal{W}^l \cup \mathcal{W}^l_I : M, w \models \neg A$ iff $M, w \not\models A$ or $w \in v(\neg A)$
$C\neg\neg$ where $A \in \mathcal{W}^c : M, w \models \neg\neg A$ iff $M, w \models A$
$C\neg\vee$ where $A, B \in \mathcal{W}^c : M, w \models \neg(A \vee B)$ iff $M, w \models \neg A \wedge \neg B$
$C\neg\wedge$ where $A, B \in \mathcal{W}^c : M, w \models \neg(A \wedge B)$ iff $M, w \models \neg A \vee \neg B$
$C\neg\supset$ where $A, B \in \mathcal{W}^c : M, w \models \neg(A \supset B)$ iff $M, w \models A \wedge \neg B$
$C\neg\equiv$ where $A, B \in \mathcal{W}^c : M, w \models \neg(A \equiv B)$ iff $M, w \models (A \vee B) \wedge (\neg A \vee \neg B)$

As before, an **MDP**-model $M$ verifies $A$ ($M \Vdash_{\mathbf{MDP}} A$) iff $M, w^0 \models A$. Where $\Gamma \subseteq \mathcal{W}^c$, $M$ is an **MDP**-model of $\Gamma$ iff $M$ is an **MDP**-model and $M \Vdash_{\mathbf{MDP}} A$ for all $A \in \Gamma$. Moreover, $\models_{\mathbf{MDP}} A$ iff all **MDP**-models verify $A$, and $\Gamma \models_{\mathbf{MDP}} A$ iff all **MDP**-models of $\Gamma$ verify $A$.

$C\neg\neg, C\neg\wedge, C\neg\vee, C\neg\supset$, and $C\neg\equiv$ ensure that de Morgan laws and the double-negation rule are valid for complex formulas in $\mathcal{W}^c$. Due to the weakened negation of **MDP** this does not follow anymore from the other semantic clauses.

**MDP** allows for both $A_J$ and $\neg(A_J)$ to be true at one and the same accessible world. Consequently, this logic can consistently model situations in which for an agent or group $J$ it ought to be that $J$ brings about $A$ and that $J$ refrains from bringing about $A$. In general, for any $A \in \mathcal{W}^c$, **MDP** allows for both $A$ and $\neg A$ to be true at one and the same accessible world. This is exactly what we need if we also want to consistently allow for the possibility of contradicting norms.

Altogether, the paraconsistent multi-agent deontic logic **MDP** is fully conflict-tolerant (where $A \in \mathcal{W}$, and $B, C \in \mathcal{W}^c_I$):

| | | |
|---|---|---|
| OO-conflicts: | $\mathsf{O}A_J \wedge \mathsf{O}\neg(A_J) \not\models_{\mathbf{MDP}} \mathsf{O}C$, | $\mathsf{O}A_J \wedge \mathsf{O}(\neg A)_K \not\models_{\mathbf{MDP}} \mathsf{O}C$ |
| OP-conflicts: | $\mathsf{O}A_J \wedge \mathsf{P}\neg(A_J) \not\models_{\mathbf{MDP}} \mathsf{O}C$, | $\mathsf{O}A_J \wedge \mathsf{P}(\neg A)_K \not\models_{\mathbf{MDP}} \mathsf{O}C$ |
| Contradicting norms: | $\mathsf{O}B \wedge \neg\mathsf{O}B \not\models_{\mathbf{MDP}} \mathsf{O}C$, | $\mathsf{P}B \wedge \neg\mathsf{P}B \not\models_{\mathbf{MDP}} \mathsf{O}C$ |
| Contradictory norms: | $\mathsf{O}(A_J \wedge \neg(A_J)) \not\models_{\mathbf{MDP}} \mathsf{O}C$, | $\mathsf{P}(A_J \wedge \neg(A_J)) \not\models_{\mathbf{MDP}} \mathsf{O}C$, |
| | $\mathsf{O}(A_J \wedge (\neg A)_K) \not\models_{\mathbf{MDP}} \mathsf{O}C$, | $\mathsf{P}(A_J \wedge (\neg A)_K) \not\models_{\mathbf{MDP}} \mathsf{O}C$, |
| | $\mathsf{O}(A \wedge \neg A)_J \not\models_{\mathbf{MDP}} \mathsf{O}C$, | $\mathsf{P}(A \wedge \neg A)_J \not\models_{\mathbf{MDP}} \mathsf{O}C$ |

## 5 Drawbacks of MDP

In an **MDP**-model, accessible worlds can consistently verify contradictions. This is what causes **MDP** to avoid deontic explosion when faced with a normative conflict. However, this property comes at a cost. We illustrate this by means of an example. Suppose that Frank has baked cookies and that it's hot in his kitchen. In order to let some fresh air in, Frank ought to open the door or open the window ($\mathsf{O}(D \vee W)_f$). However, if someone opens the door, the neighbour's dog might smell Frank's cookies and try to steal them from the table. Hence Frank should take care that the door remains closed ($\mathsf{O}(\neg D)_f$). In this situation Frank can consistently satisfy his obligations by simply opening the window.

Yet $\mathsf{O}W_f$ is not **MDP**-derivable from $\mathsf{O}(D \vee W)_f$ and $\mathsf{O}(\neg D)_f$. Note that there are **MDP**-models of the premise set $\Gamma_1 = \{\mathsf{O}(D \vee W)_f, \mathsf{O}(\neg D)_f\}$ in which inconsistent worlds are accessible from the actual world, i.e. worlds in which both $D_f$ and $(\neg D)_f$ are true (and, consequently, in which both $D$ and $\neg D$ are true). In these worlds, $W_f$ may be false while the premises are true. In contrast, all the **MDC**-models of our premise set $\Gamma_1$ are such that all the accessible worlds are consistent and verify $(D \vee W)_f$, $(\neg D)_f$ and hence $W_f$. This is the reason why $\Gamma_1 \models_{\mathbf{MDC}} \mathsf{O}W_f$ while $\Gamma_1 \not\models_{\mathbf{MDP}} \mathsf{O}W_f$. Obviously our premise set is not conflicting. In such cases we would ideally expect from any deontic logic that its models do not verify normative conflicts. Hence, in our case we are interested in **MDP**-models that –just like the **MDC**-models– do not validate $D_f \wedge (\neg D)_f$ in any of the accessible worlds, i.e. models $M$ for which $M \nVdash_{\mathbf{MDP}} P(D_f \wedge (\neg D)_f)$. It is easy to see that all these models validate $W_f$ in all the accessible worlds, just like the **MDC**-models. In other words, since $\Gamma_1 \models_{\mathbf{MDP}} \mathsf{O}W_f \vee P(D_f \wedge (\neg D)_f)$ we get $\mathsf{O}W_f$ by selecting models that do not validate any normative conflicts.

The solution offered above is obviously not working as soon as we have to deal with conflicting premise sets. Suppose Frank invited his aunt Maggie for a cup of coffee and cookies in the afternoon. However, his other aunt Beth is an awfully jealous person: she would be deeply insulted if she's not also invited. Hence Frank has the obligation to also invite Beth ($\mathsf{O}B_f$). On the other hand, Maggie cannot stand Beth (she's a rather difficult person) and whenever they are together all hell breaks loose. Thus, Frank should make sure that Beth is not invited ($\mathsf{O}(\neg B)_f$). Let $\Gamma_2 = \Gamma_1 \cup \{\mathsf{O}B_f, \mathsf{O}(\neg B)_f\}$. While **MDC** trivializes $\Gamma_2$, **MDP** does not trivialize $\Gamma_2$ but is again too weak. For the same reason as above, $\Gamma_2 \not\models_{\mathbf{MDP}} \mathsf{O}W_f$. However, in contrast to above we cannot now simply select models whose worlds are consistent since there are no such models. Indeed, all models of $\Gamma_2$ are such that in all accessible worlds $B_f$ and $(\neg B)_f$ are valid. In other words, all models validate $\mathsf{O}(B_f \wedge (\neg B)_f)$. But, similar to above, the idea is to not take into consideration models that validate $\mathsf{P}(D_f \wedge (\neg D)_f)$.

In a nutshell the idea is to strengthen **MDP** by selecting models whose accessible worlds are "as non-conflicting as possible". This idea will be realized by means of the adaptive logic $\mathbf{MDP^m}$.

Before we introduce this logic in Section 6, let us focus on some other weaknesses of **MDP**. For instance, all of the following inferences are invalid in **MDP**, for the same reason why $\mathsf{O}W_f$ is not **MDP**-derivable from $\mathsf{O}(D \vee W)_f$ and $\mathsf{O}(\neg D)_f$: because of the possibility of contradictions being true in accessible worlds in **MDP**-models.[5] Where $A, B \in \mathcal{W}$, and $C, D \in \mathcal{W}^c$:

(1) $\mathsf{O}(A \vee B)_J, \mathsf{O}(\neg A)_J \not\models_{\mathbf{MDP}} \mathsf{O}B_J$

[5] These problems are common to monotonic logics with a paraconsistent negation. In [6], it was argued that the paraconsistent deontic logics presented in [10, 24, 25] are too weak to account for deontic reasoning.

(2) $\mathsf{O}(A \vee B)_J, \mathsf{O}\neg(A_J) \not\models_{\mathbf{MDP}} \mathsf{O}B_J$
(3) $\mathsf{O}A_J \not\models_{\mathbf{MDP}} \neg\mathsf{P}\neg(A_J)$
(4) $\neg\mathsf{O}A_J \not\models_{\mathbf{MDP}} \mathsf{P}\neg(A_J)$
(5) $\mathsf{P}A_J \not\models_{\mathbf{MDP}} \neg\mathsf{O}\neg(A_J)$
(6) $\neg\mathsf{P}A_J \not\models_{\mathbf{MDP}} \mathsf{O}\neg(A_J)$
(7) $C \vee D, \neg C \not\models_{\mathbf{MDP}} D$
(8) $C \supset D \not\models_{\mathbf{MDP}} \neg D \supset \neg C$

Items (1) and (2) represent deontic variants of Disjunctive Syllogism, (3)–(6) represent variants of the interdefinability between obligations and permissions, (7) is the propositional version of Disjunctive Syllogism, and (8) is Contraposition. In contrast, the following inferences *are* valid in **MDP**.

(1') $\mathsf{O}(A \vee B)_J, \mathsf{O}(\neg A)_J \models_{\mathbf{MDP}} \mathsf{O}B_J \vee \mathsf{P}(A_J \wedge (\neg A)_J)$
(2') $\mathsf{O}(A \vee B)_J, \mathsf{O}\neg(A_J) \models_{\mathbf{MDP}} \mathsf{O}B_J \vee \mathsf{P}(A_J \wedge \neg(A_J))$
(3') $\mathsf{O}A_J \models_{\mathbf{MDP}} \neg\mathsf{P}\neg(A_J) \vee \mathsf{P}(A_J \wedge \neg(A_J))$
(4') $\neg\mathsf{O}A_J \models_{\mathbf{MDP}} \mathsf{P}\neg(A_J) \vee (\mathsf{O}A_J \wedge \neg\mathsf{O}A_J)$
(5') $\mathsf{P}A_J \models_{\mathbf{MDP}} \neg\mathsf{O}\neg(A_J) \vee \mathsf{P}(A_J \wedge \neg(A_J))$
(6') $\neg\mathsf{P}A_J \models_{\mathbf{MDP}} \mathsf{O}\neg(A_J) \vee (\mathsf{P}A_J \wedge \neg\mathsf{P}A_J)$
(7') $C \vee D, \neg C \models_{\mathbf{MDP}} D \vee (C \wedge \neg C)$
(8') $C \supset D \models_{\mathbf{MDP}} (\neg D \supset \neg C) \vee (D \wedge \neg D)$

In items (1')–(8'), all formulas on the right-hand side of the "$\vee$"-sign represent normative conflicts. As in our example above, interpreting premise sets as non-conflicting as possible will validate the deontic and propositional versions of Disjunctive Syllogism, the interdefinability between obligations and permissions, and Contraposition as much as possible. Indeed, given that the normative conflicts on the right-hand side of "$\vee$" are false in (1')–(8'), the left-hand disjuncts must be true.

## 6 The adaptive logic $\mathbf{MDP^m}$

Adaptive logics are characterized by means of a triple consisting of a lower limit logic (henceforth LLL), a set of abnormalities $\Omega$, and an adaptive strategy.[6] The LLL constitutes the stable part of an adaptive logic: everything that is LLL-derivable from a given set of premises, is still derivable by means of the adaptive logic. Formulating adaptive logics in the standard format has the advantage that a rich meta-theory is immediately available for this format [4]. Although adaptive logics come with an attractive dynamic proof theory we will for the sake of conciseness focus in this paper exclusively on the semantics.

Typically, an adaptive logic enables one to derive, for most premise sets, some extra consequences on top of those that are LLL-derivable. These supple-

[6] For an introduction to adaptive logics, see [4]. However, familiarity with this framework for non-monotonic reasoning is not necessary for understanding the workings of the logic $\mathbf{MDP^m}$.

mentary consequences are obtained by interpreting a premise set "as normally as possible". The exact interpretation of this idea depends on the adaptive strategy which defines which models of the LLL are selected.[7] For our present purposes, we shall use the Minimal Abnormality strategy. The logic $\mathbf{MDP^m}$ is characterized by:

(1) LLL: **MDP**

(2) Set of abnormalities: $\mathit{\Omega} = \mathit{\Omega}_1 \cup \mathit{\Omega}_2 \cup \mathit{\Omega}_3$, where
$\mathit{\Omega}_1 = \{A \wedge \neg A \mid A \in \mathcal{W}^c\}$
$\mathit{\Omega}_2 = \{\mathsf{P}(A_J \wedge \neg(A_J)) \mid A \in \mathcal{W}, J \subseteq_\emptyset I\}$
$\mathit{\Omega}_3 = \{\mathsf{P}(A_J \wedge (\neg A)_K) \mid A \in \mathcal{W}; J, K \subseteq_\emptyset I\}$

(3) Adaptive strategy: Minimal Abnormality

By (1) we make sure that we select **MDP**-models. This ensures that $\mathbf{MDP^m}$ inherits the conflict-tolerance from **MDP**.

As mentioned, adaptive logics interpret premise sets in a way that as few abnormalities as possible are verified. The attentive reader will have noticed that not all conflict-types that were listed in the table in Section 3 occur in $\mathit{\Omega}$. This is justified due to the fact that all other conflict-types give rise to abnormalities in $\mathit{\Omega}$, as the following table shows (where $A \in \mathcal{W}$, and $B, C \in \mathcal{W}_I^c$)[8]:

$$\begin{array}{rlrl}
\mathsf{O}A_J, \mathsf{O}\neg(A_J) \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge \neg(A_J)), & \mathsf{O}A_J, \mathsf{O}(\neg A)_K \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge (\neg A)_K) \\
\mathsf{O}A_J, \mathsf{P}\neg(A_J) \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge \neg(A_J)), & \mathsf{O}A_J, \mathsf{P}(\neg A)_K \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge (\neg A)_K) \\
\mathsf{O}(A_J \wedge \neg(A_J)) \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge \neg(A_J)), & \mathsf{O}(A_J \wedge (\neg A)_K) \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge (\neg A)_K) \\
\mathsf{O}(A \wedge \neg A)_J \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge (\neg A)_J), & \mathsf{P}(A \wedge \neg A)_J \models_{\mathbf{MDP}} & \mathsf{P}(A_J \wedge (\neg A)_J)
\end{array}$$

For our semantic selection we will make use of the notion of the *abnormal part* of an **MDP**-model, i.e. the set of all abnormalities verified by it: $Ab(M) = \{A \in \mathit{\Omega} \mid M \Vdash_{\mathbf{MDP}} A\}$. The Minimal Abnormality strategy selects all **MDP**-models of a premise set $\mathit{\Gamma}$ which have a *minimal* abnormal part (w.r.t. set-inclusion).

**Definition 1.** *An* **MDP***-model $M$ of $\mathit{\Gamma}$ is* minimally abnormal *iff there is no* **MDP***-model $M'$ of $\mathit{\Gamma}$ such that $Ab(M') \subset Ab(M)$.*

The semantic consequence relation of the logic $\mathbf{MDP^m}$ is defined by selecting the minimally abnormal **MDP**-models:

**Definition 2.** *$\mathit{\Gamma} \models_{\mathbf{MDP^m}} A$ iff $A$ is verified by all minimally abnormal* **MDP***-models of $\mathit{\Gamma}$.*

The fact that the set of $\mathbf{MDP^m}$-models of $\mathit{\Gamma}$ is a subset of the set of **MDP**-models of $\mathit{\Gamma}$ immediately ensures that $\mathbf{MDP^m}$ strengthens **MDP**.

**Theorem 1.** *If $\mathit{\Gamma} \models_{\mathbf{MDP}} A$, then $\mathit{\Gamma} \models_{\mathbf{MDP^m}} A$.*

[7] Besides adaptive logics many other formal frameworks make use of semantic selections, e.g. [19, 26].

[8] conflicts of the form $\mathsf{O}B \wedge \neg\mathsf{O}B$, $\mathsf{P}B \wedge \neg\mathsf{P}B$, $\mathsf{P}(A_J \wedge \neg(A_J))$, or $\mathsf{P}(A_J \wedge (\neg A)_K)$ are not listed in the table, since these conflicts already have the form of an abnormality.

For an illustration of the logic, let's return to the example presented in Section 5. Remember that $\Gamma_1 \not\models_{\mathbf{MDP}} \mathsf{O}W_f$. However, $\Gamma_1 \models_{\mathbf{MDP}} \mathsf{O}W_f \vee \mathsf{P}(D_f \wedge (\neg D)_f)$. By C$\vee$, we know that (†) for all **MDP**-models $M$ of $\Gamma_1$: if $M \not\Vdash_{\mathbf{MDP}} \mathsf{P}(D_f \wedge (\neg D)_f)$ then $M \Vdash_{\mathbf{MDP}} \mathsf{O}W_f$.

No abnormality $A \in \Omega$ is an **MDP**-consequence of $\Gamma_1$, hence there are **MDP**-models $M$ of $\Gamma_1$ such that $Ab(M) = \emptyset$. By Definition 1, these and only these are the minimal abnormal models of $\Gamma_1$. It follows that, for all minimal abnormal models $M$ of $\Gamma_1$, $M \not\Vdash_{\mathbf{MDP}} \mathsf{P}(D_f \wedge (\neg D)_f)$. By (†), it follows that $M \Vdash_{\mathbf{MDP}} \mathsf{O}W_f$ for all minimal abnormal models $M$ of $\Gamma_1$. Hence, by Definition 2, $\Gamma_1 \models_{\mathbf{MDP^m}} \mathsf{O}W_f$.

By the same reasoning as applied in the example above, we can show that all of (1")-(8") below are **MDP$^{\mathbf{m}}$**-valid in view of the **MDP**-validity of (1')-(8') as displayed in Section 5:

(1") $\mathsf{O}(A \vee B)_J, \mathsf{O}(\neg A)_J \models_{\mathbf{MDP^m}} \mathsf{O}B_J$
(2") $\mathsf{O}(A \vee B)_J, \mathsf{O}\neg(A_J) \models_{\mathbf{MDP^m}} \mathsf{O}B_J$
(3") $\mathsf{O}A_J \models_{\mathbf{MDP^m}} \neg\mathsf{P}\neg(A_J)$
(4") $\neg\mathsf{O}A_J \models_{\mathbf{MDP^m}} \mathsf{P}\neg(A_J)$
(5") $\mathsf{P}A_J \models_{\mathbf{MDP^m}} \neg\mathsf{O}\neg(A_J)$
(6") $\neg\mathsf{P}A_J \models_{\mathbf{MDP^m}} \mathsf{O}\neg(A_J)$
(7") $C \vee D, \neg C \models_{\mathbf{MDP^m}} D$
(8") $C \supset D \models_{\mathbf{MDP^m}} \neg D \supset \neg C$

In a similar fashion, we can show that other intuitive **MDC**-inferences are also **MDP$^{\mathbf{m}}$**-valid in the absence of normative conflicts. Remember from Section 2.3 that $\mathsf{O}(\neg A)_i, \mathsf{O}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDC}} \mathsf{O}B_{\{i,j\}}$ and $\mathsf{O}(\neg A)_i, \mathsf{O}^{\mathsf{w}}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDC}} \mathsf{O}^{\mathsf{w}}B_{\{i,j\}}$. Both of these inferences are invalidated by **MDP**. However, $\mathsf{O}(\neg A)_i, \mathsf{O}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDP}} \mathsf{O}B_{\{i,j\}} \vee \mathsf{P}(A_{\{i,j\}} \wedge (\neg A)_i)$, and $\mathsf{O}(\neg A)_i, \mathsf{O}^{\mathsf{w}}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDP}} \mathsf{O}^{\mathsf{w}}B_{\{i,j\}} \vee \mathsf{P}(A_i \wedge (\neg A)_i) \vee \mathsf{P}(A_j \wedge (\neg A)_i) \vee \mathsf{P}(A_{\{i,j\}} \wedge (\neg A)_i)$. Note that none of the minimal abnormal **MDP**-models of $\{\mathsf{O}(\neg A)_i, \mathsf{O}(A \vee B)_{\{i,j\}}\}$ and $\{\mathsf{O}(\neg A)_i, \mathsf{O}^{\mathsf{w}}(A \vee B)_{\{i,j\}}\}$ validate one of the abnormalities $\mathsf{P}(A_{\{i,j\}} \wedge (\neg A)_i), \mathsf{P}(A_i \wedge (\neg A)_i)$, or $\mathsf{P}(A_j \wedge (\neg A)_i)$. Hence $\mathsf{O}(\neg A)_i, \mathsf{O}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDP^m}} \mathsf{O}B_{\{i,j\}}$ and $\mathsf{O}(\neg A)_i, \mathsf{O}^{\mathsf{w}}(A \vee B)_{\{i,j\}} \models_{\mathbf{MDP^m}} \mathsf{O}^{\mathsf{w}}B_{\{i,j\}}$.

The following theorem shows that for any **MDC**-consistent premise set the **MDP$^{\mathbf{m}}$**-consequences are identical to the **MDC**-consequences:

**Theorem 2.** *For all* **MDC**-*consistent* $\Gamma$, $\Gamma \models_{\mathbf{MDP^m}} A$ *iff* $\Gamma \models_{\mathbf{MDC}} A$.

A proof of Theorem 2 is contained in the Appendix. Note that (1")-(8") immediately follow as a corollary to Theorem 2.

If all **MDP**-models of given a premise set verify at least one abnormality, then **MDP$^{\mathbf{m}}$** is still considerably stronger than **MDP**. Consider the premise set $\Gamma_2$ from Section 5, where we enriched $\Gamma_1$ with the conflicting obligations concerning the invitation of aunt Beth, $\mathsf{O}B_f$ and $\mathsf{O}(\neg B)_f$. Here too, Frank's obligation to open the window is an **MDP$^{\mathbf{m}}$**-consequence: $\Gamma_2 \models_{\mathbf{MDP^m}} \mathsf{O}W_f$. Although there are no models of $\Gamma_2$ that have an empty abnormal part since all models validate the abnormality $\mathsf{P}(B_f \wedge (\neg B)_f)$, the minimal abnormal models do not validate $\mathsf{P}(D_f \wedge (\neg D)_f)$ (as the reader can easily verify).

Imagine now that we add to $\Gamma_2$ the premise $\mathsf{O}(\neg W)_f$, which abbreviates Frank's obligation to take care that the window remains closed (e.g. because it was painted recently and the paint is not dry yet). Let us call this extended premise set $\Gamma_3$. Then $\Gamma_3 \models_{\mathbf{MDP}} \mathsf{P}(D_f \wedge (\neg D)_f) \vee \mathsf{P}(W_f \wedge (\neg W)_f)$. Consequently, all minimally abnormal **MDP**-models $M$ of $\Gamma_3$ verify at least one of $\mathsf{P}(D_f \wedge (\neg D)_f)$ and $\mathsf{P}(W_f \wedge (\neg W)_f)$. $\Gamma_3$ has minimally abnormal **MDP**-models which verify $\mathsf{P}(D_f \wedge (\neg D)_f)$. Since it is no longer the case that, for all minimally abnormal **MDP**-models $M$ of $\Gamma_3$, $M \not\Vdash_{\mathbf{MDP}} \mathsf{P}(D_f \wedge (\neg D)_f)$, for these models it no longer follows that $M \Vdash_{\mathbf{MDP}} \mathsf{O}W_f$. Hence $\Gamma_3 \not\models_{\mathbf{MDP^m}} \mathsf{O}W_f$. Since $\Gamma_1 \subset \Gamma_3$ and $\Gamma_1 \models_{\mathbf{MDP^m}} \mathsf{O}W_f$, this shows that the logic $\mathbf{MDP^m}$ is non-monotonic.

The following theorems state some further meta-theoretical properties of $\mathbf{MDP^m}$. Let $\mathcal{M}_{\Gamma}^{\mathbf{MDP}}$ $[\mathcal{M}_{\Gamma}^{\mathbf{MDP^m}}]$ abbreviate the set of **MDP**- [$\mathbf{MDP^m}$-] models of $\Gamma$.

**Theorem 3.** *If $M \in \mathcal{M}_{\Gamma}^{\mathbf{MDP}} - \mathcal{M}_{\Gamma}^{\mathbf{MDP^m}}$, then there is a $M' \in \mathcal{M}_{\Gamma}^{\mathbf{MDP^m}}$ such that $Ab(M') \subset Ab(M)$ (Strong reassurance).*

For the proof of Theorem 3, we refer to [4].[9]

**Theorem 4.** *If $\Gamma \models_{\mathbf{MDP^m}} A$ for all $A \in \Gamma'$, then $\mathcal{M}_{\Gamma}^{\mathbf{MDP^m}} = \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP^m}}$.*

*Proof.* Suppose (†) $\Gamma \models_{\mathbf{MDP^m}} A$ for all $A \in \Gamma'$. Consider a $M \in \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP^m}}$. Then $M \in \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP}}$ and whence $M \in \mathcal{M}_{\Gamma}^{\mathbf{MDP}}$. Assume $M \notin \mathcal{M}_{\Gamma}^{\mathbf{MDP^m}}$. By the strong reassurance there is a $M' \in \mathcal{M}_{\Gamma}^{\mathbf{MDP^m}}$ such that $Ab(M') \subset Ab(M)$. In view of (†), $M' \Vdash_{\mathbf{MDP}} A$ for every $A \in \Gamma'$. Hence, $M' \in \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP}}$. But then $M \notin \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP^m}}$,— a contradiction.

Consider a $M \in \mathcal{M}_{\Gamma}^{\mathbf{MDP^m}}$. By (†), $M \Vdash_{\mathbf{MDP}} A$ for every $A \in \Gamma'$. By definition also $M \in \mathcal{M}_{\Gamma}^{\mathbf{MDP}}$. Hence $M \in \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP}}$. Assume $M \notin \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP^m}}$. Hence, there is a $M' \in \mathcal{M}_{\Gamma\cup\Gamma'}^{\mathbf{MDP}}$ for which $Ab(M') \subset Ab(M)$. By definition, $M \in \mathcal{M}_{\Gamma}^{\mathbf{MDP}}$. But then $M \notin \mathcal{M}_{\Gamma}^{\mathbf{MDP^m}}$,— a contradiction.

**Corollary 1.** *If $\Gamma \models_{\mathbf{MDP^m}} A$ for all $A \in \Gamma'$, then*

*(i) if $\Gamma \models_{\mathbf{MDP^m}} A$ then $\Gamma \cup \Gamma' \models_{\mathbf{MDP^m}} A$ (Cautious Monotonicity);*
*(ii) if $\Gamma \cup \Gamma' \models_{\mathbf{MDP^m}} A$ then $\Gamma \models_{\mathbf{MDP^m}} A$ (Cautious Cut).*

**Theorem 5.** $\mathcal{M}_{\Gamma}^{\mathbf{MDP^m}} = \mathcal{M}_{\{B\mid\Gamma\models_{\mathbf{MDP^m}}B\}}^{\mathbf{MDP^m}}$*, and whence $\Gamma \models_{\mathbf{MDP^m}} A$ iff $\{B \mid \Gamma \models_{\mathbf{MDP^m}} B\} \models_{\mathbf{MDP^m}} A$ (Fixed point).*

*Proof.* Since obviously $\{B \mid \Gamma \models_{\mathbf{MDP^m}} B\} \models_{\mathbf{MDP^m}} A$ for all $A \in \Gamma$, this is an immediate consequence of Theorem 4 (where $\Gamma' = \{B \mid \Gamma \models_{\mathbf{MDP^m}} B\}$).

[9] In [4], the strong reassurance property is proven for logics that fit the so-called standard format for adaptive logics. In order for the proof for strong reassurance from [4] to work, $\mathbf{MDP^m}$ needs to contain all classical connectives. $\mathbf{MDP^m}$ can easily be adjusted to do so by adding the constant false symbol "$\bot$" to its language, and by defining a classical negation connective "$\sim$" as $\sim A =_{\mathrm{df}} A \supset \bot$.

## 7 Outlook

The central problem tackled in this paper is the modeling of normative conflicts in multi-agent deontic logic. Because of this focus and reasons of conciseness, we have presented the logics **MDC**, **MDP** and $\mathbf{MDP^m}$ in a very basic form. In this section we will briefly demonstrate that they can be enhanced in various ways.

Some may wish to increase the expressiveness of our logics by alethic modalities. One way to technically realize this is to add another accessibility relation $R'$ to the **MDC**-models so that the models are tuples $\langle W, I, R, R', v, w^0\rangle$.[10] Validity for the $\Box$-operator is characterized as usual: $M, w \models \Box A$ iff for all $w' \in W$, if $R'ww'$ then $M, w' \models A$ (and the dual version for $\Diamond$). By requiring $R \subseteq R'$ it could be ensured that the Kantian "ought implies can" holds: $\mathsf{O}A \supset \Diamond A$.

Another extension could indicate the authority that issues a norm. For instance, $\mathsf{O}^a A_J$ reads "authority $a$ issues the norm that $J$ brings about $A$". Technically, introducing authorities is straightforward. First, we enhance our models by a set $A$ of authorities. This set may intersect with or even be identical to the set of agents $I$. Second, instead of one accessibility relation we introduce an accessibility relation $R^a$ for each authority $a \in A$. The semantic clauses are adjusted as expected: $M, w \models \mathsf{O}^a A$ iff for all $w' \in W$, if $R^a ww'$ then $M, w' \models A$ (and dually for $\mathsf{P}^a$).

In a way technically analogous to the representation of different authorities via superscripts to the deontic operators, we could add subscripts for distinguishing between various interest groups in view of which the norms hold (cfr. [18]). Moreover, the adaptive framework could be enhanced so as to allow for varying degrees of priority amongst norms and/or conditional norms [29].

The framework used in this paper is elementary not only in its limited expressive power, but also in its treatment of the notions of action and agency. At the moment, this paper is lacking a comparison with other frameworks for representing agency in deontic logic. Further research includes (i) the relation of the agentive setting applied here with other such settings, e.g. dynamic logic [9, 22], stit theory [16, 18], and their historical predecessors [23, 33, 34]; and (ii) the application of the inconsistency-adaptive approach for accommodating normative conflicts within these other frameworks for accounting for action in deontic logic.

## A Appendix: proof of Theorem 2

For every adaptive logic, there is a so-called *upper limit logic*. The upper limit logic **UMDP** of $\mathbf{MDP^m}$ is defined as follows: given a premise set $\Gamma$ we select all **MDP**-models $M$ of $\Gamma$ such that $Ab(M) = \emptyset$. **UMDP** is a monotonic logic that trivializes premise sets that give rise to abnormalities.

**Lemma 1.** For each **UMDP**-model $M$ of $\Gamma$, **F-Con** holds.

[10] We will exemplify all enhancements by means of **MDC**. The arguments are analogous for **MDP** and $\mathbf{MDP^m}$.

*Proof.* Let $M = \langle W, I, R, v, v_I, w^0 \rangle$. Suppose for some $w \in W$, some $A \in \mathcal{W}^a$, and some $J, K \subseteq_\emptyset I$, $w_J \in v_I(A)$ and $w_K \in v_I(\neg A)$. By **F-Fac**$'$, $w \in v(A)$ and $w \in v(\neg A)$. If $w = w^0$, by C$^a$, C$\neg'$ and C$\wedge$, $M, w^0 \models A \wedge \neg A$ and hence $A \wedge \neg A \in Ab(M)$,—a contradiction. If $w \neq w^0$, then by C$^a$, C$\neg'$, C$\wedge$ and C$\mathsf{P}$, $M, w^0 \models \mathsf{P}(A_J \wedge (\neg A)_K)$ and hence $\mathsf{P}(A_J \wedge (\neg A)_K) \in Ab(M)$,—a contradiction. Hence, **F-Con** holds. □

Let an **MDP**-model $\langle W, I, R, v, v_I, w^0 \rangle$ be **MDC**-*like* iff, (a) for all $A \in \mathcal{W}^a$, $w \in v(\neg A)$ iff $w \notin v(A)$; (b) for all $A_J \in \mathcal{W}^l_I$, $w \in v(\neg(A_J))$ iff $w_J \notin v_I(A)$; and (c) **F-Fac** holds. We say that two models *are equivalent* iff they validate the same formulas.

**Lemma 2.** For each **UMDP**-model $M = \langle W, I, R, v, v_I, w_0 \rangle$ there is an equivalent **MDC**-like **UMDP**-model $M' = \langle W, I, R, v', v_I, w_0 \rangle$.

*Proof.* Define $v'$ as follows: (1) where $A \in \mathcal{W}^l \cup \{\neg(B_J) \mid B \in \mathcal{W}^l\}$, $w^0 \in v'(A)$ iff $w^0 \in v(A)$; (2) where $w \in W \setminus \{w^0\}$ and $A \in \mathcal{W}^a$, $w \in v'(A)$ iff there is a $J \subseteq_\emptyset I$ for which $w_J \in v_I(A)$; (3) where $w \in W \setminus \{w^0\}$ and $\neg A \in \mathcal{W}^l$, $w \in v'(\neg A)$ iff $w \notin v'(A)$; and (4) where $A \in \mathcal{W}^l$, $w \in v'(\neg(A_J))$ iff $w_J \notin v_I(A)$.

**F-Con** only depends on $v_I$ and hence holds for $M'$ due to Lemma 1. **F-Fac**$'$(i) holds by (2) and (ii) by (3) and due to **F-Con**. Hence, $M'$ is a **MDP**-model.

**F-Fac** (i) holds due to **F-Fac**$'$ (i). Let $w_J \in v_I(\neg A)$. Suppose first that $w = w^0$. By **F-Fac**$'$, $w^0 \in v'(\neg A)$ or $w^0 \notin v'(A)$ and whence $w^0 \in v(\neg A)$ or $w^0 \notin v(A)$. Assume that $w^0 \in v(\neg A) \cap v(A)$. But then $A \wedge \neg A \in Ab(M)$,—a contradiction. Hence $w^0 \notin v(A)$ and whence $w^0 \notin v'(A)$. Let now $w \in W \setminus \{w^0\}$. By **F-Con**, there is no $K \subseteq_\emptyset I$ for which $w_K \in v_I(A)$. Hence, by (2), $w \notin v'(A)$. Thus, **F-Fac** holds for $M'$.

Note that (a) holds for $v'$ due to (3), and (b) holds due to (4). Hence, $M'$ is **MDC**-like.

$M \Vdash_{\mathbf{MDP}} A$ iff $M' \Vdash_{\mathbf{MDP}} A$ is shown by an induction over the length of the formula $A$. The induction base is easily established. Where $A \in \mathcal{W}^a$ the equivalence holds by C$^a$ and (1). Where $A \in \mathcal{W}^l_I$ the equivalence holds due to C$^l_I$. For the induction step let first $A = \neg A'$. Suppose $A' \in \mathcal{W}^a \cup \mathcal{W}^l_I$. Note that by the induction hypothesis, C$\neg'$ and (1) we have the same valuation for $A$. Let now $A' \in \mathcal{W}_I \setminus \mathcal{W}^l_I$. Since both models have the same assignment $v_I$ the valuation is analogous due to C$^l_I$, C$_I\wedge$, C$_I\vee$, C$_I\supset$, C$_I\equiv$, C$_I\neg\neg$, C$_I\neg\vee$, C$_I\neg\wedge$, C$_I\neg\supset$, and C$_I\neg\equiv$. The similar cases for $A' = B \pi C$ where $B, C \in \mathcal{W}^c$ and $\pi \in \{\vee, \wedge, \supset, \equiv\}$ resp. for $A' = \neg B$ where $B \in \mathcal{W}^c$ are left to the reader. The induction proceeds in a similar way if $A \in \mathcal{W}_I \setminus \mathcal{W}^l_I$, $A = \mathsf{O}A'$ or $A = \mathsf{P}A'$ where $A' \in \mathcal{W}^c_I$, or $A = B \pi C$ where $B, C \in \mathcal{W}^c$ and $\pi \in \{\vee, \wedge, \supset, \equiv\}$. Since $M$ and $M'$ are equivalent, $Ab(M') = Ab(M) = \emptyset$ and whence $M'$ is an **UMDP**-model. □

**Corollary 1.** $\Gamma \models_{\mathbf{UMDP}} A$ iff each **MDC**-like **UMDP**-model $M$ of $\Gamma$ validates $A$.

Where $M = \langle W, I, R, v, v_I, w^0 \rangle$ is an **MDC**-like **UMDP**-model, let $M_c = \langle W, I, R, v_c, v_I, w^0 \rangle$ be an **MDC**-model where $v_c : \mathcal{W}^a \to \wp(W), A \mapsto v(A)$. Note that $M_c$ is indeed an **MDC**-model since $M$ satisfies **F-Con** and **F-Fac** and thus by the definition also $M_c$.

**Lemma 3.** $M$ and $M_c$ are equivalent.

The Lemma is proved by a similar induction over the length of $A$ as in the proof of Lemma 2. Due to space restrictions this is left to the reader.

**Lemma 4.** Where $M$ is an **MDC**-model of $\Gamma$, $Ab(M) = \emptyset$.

*Proof.* Suppose $A \in Ab(M)$. Let $A = B \wedge \neg B \in \Omega_1$. By C$\neg$, $M \models B$ and $M \not\models B$,—a contradiction. The other cases are similar and left to the reader. □

Where $M = \langle W, I, R, v, v_I, w^0 \rangle$ is an **MDC**-model, let $M_p = \langle W, I, R, v_p, v_I, w^0 \rangle$ be an **MDC**-like **MDP**-model where $v_p : \mathcal{W}^l \cup \{\neg(A_J) \mid A \in \mathcal{W}^l, J \subseteq_\emptyset I\} \to \wp(W)$ is defined by: $w \in v_p(A)$ iff $M, w \models A$. The reader can easily verify that $M_p$ is **MDC**-like.

**Lemma 5.** $M$ and $M_p$ are equivalent and $M_p$ is a **UMDP**-model.

Again, the proof of the equivalence proceeds by a similar induction over the length of $A$ as in the proof of Lemma 2. By Lemma 4, $Ab(M) = \emptyset$ and hence $Ab(M_p) = \emptyset$. Hence, $M_p$ is a **UMDP**-model.

Theorem 2 is an immediate consequence of Corollary 1, Lemma 3 and Lemma 5.

## References

1. Carlos E. Alchourrón. Logic of norms and logic of normative propositions. *Logique & Analyse*, 47:242–268, 1969.
2. Carlos E. Alchourrón and Eugenio Bulygin. The expressive conception of norms. In Risto Hilpinen, editor, *New Studies in Deontic Logic*, pages 95–124. D. Reidel Publishing Company, Dordrecht, 1981.
3. Diderik Batens. A survey of inconsistency-adaptive logics. In Diderik Batens, Graham Priest, and Jean-Paul Van Bendegem, editors, *Frontiers of Paraconsistent Logic*, pages 49–73. Baldock: Research Studies Press, Kings College Publication, 2000.
4. Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.
5. Diderik Batens and Joke Meheus. Recent results by the inconsistentcy-adaptive labourers. In Jean-Yves Béziau, Walter Carnielli, and Dov Gabbay, editors, *Handbook of Paraconsistency*, pages 81–99. College Publications, London, 2007.
6. Mathieu Beirlaen, Joke Meheus, and Christian Straßer. An inconsistency-adaptive deontic logic for normative conflicts. Under review.
7. Guido Boella, Leendert Van Der Torre, and Harko Verhagen. Introduction to the special issue on normative multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 17:1–10, August 2008.
8. Guido Boella and Leendert van der Torre. Permissions and obligations in hierarchical normative systems. In *Proceedings of the 9th international conference on Artificial intelligence and law*, ICAIL '03, pages 109–118, New York, NY, USA, 2003. ACM.
9. Jan Broersen. Action negation and alternative reductions for dynamic deontic logics. *Journal of Applied Logic*, 2:153–168, 2004.
10. Newton Da Costa and Walter Carnielli. On paraconsistent deontic logic. *Philosophia*, 16:293–305, 1986.
11. Frank Dignum and Lambèr Royakkers. Collective commitment and obligation. In C. Ciampi and E. Marinai, editors, *Proceedings of 5th Int. conference on Law in the Information Society, Firenze, Italy*, pages 1008–1021, 1998.
12. Lou Goble. Multiplex semantics for deontic logic. *Nordic Journal of Philosophical Logic*, 5(2):113–134, 2000.
13. Lou Goble. A logic for deontic dilemmas. *Journal of Applied Logic*, 3:461–483, 2005.
14. Jörg Hansen, Gabriella Pigozzi, and Leendert van der Torre. Ten philosophical problems in deontic logic. In Guido Boella, Leon van der Torre, and Harko Verhagen, editors, *Normative Multi-agent Systems*, number 07122 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.
15. John Francis Horty. Moral dilemmas and nonmonotonic logic. *Journal of Philosophical Logic*, 23(1):35–66, 1994.

16. John Francis Horty. *Agency and Deontic Logic.* Oxford University Press, 2001.
17. John Francis Horty. Reasoning with moral conflicts. *Noûs*, 37:557–605, 2003.
18. Barteld Kooi and Allard Tamminga. Moral conflicts between groups of agents. *Journal of Philosophical Logic*, 37:1–21, 2008.
19. Sarit Kraus, Daniel J. Lehmann, and Menachem Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.
20. Terrance McConnell. Moral dilemmas. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Summer 2010 edition, 2010.
21. Joke Meheus, Mathieu Beirlaen, and Frederik Van De Putte. Avoiding deontic explosion by contextually restricting aggregation. In Guido Governatori and Giovanni Sartor, editors, *DEON (10th International Conference on Deontic Logic in Computer Science)*, volume 6181 of *Lecture Notes in Artificial Intelligence*, pages 148–165. Springer, 2010.
22. John-Jules Meyer. A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29:109–136, 1988.
23. Hector-Neri Casta neda. The paradoxes of deontic logic: the simplest solution to all of them in one fell swoop. In Risto Hilpinen, editor, *New Studies in Deontic Logic*, pages 37–85. D. Reidel Publishing Company, Dordrecht, 1981.
24. Graham Priest. *In Contradiction: A Study of the Transconsistent (2nd Edition).* Oxford University Press, 2006.
25. Richard Routley and Val Plumwood. Moral dilemmas and the logic of deontic notions. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic. Essays on the Inconsistent*, pages 653–702. Philosophia Verlag, München, 1989.
26. Yoav Shoham. A semantical approach to nonmonotonic logics. In Matthew L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 227–250. Morgan Kaufmann Publishers, 1987.
27. Walter Sinnott-Armstrong. *Moral Dilemmas.* Basil Blackwell, Oxford/New York, 1988.
28. Christian Straßer. An adaptive logic framework for conditional obligations and deontic dilemmas. *Logic and Logical Philosophy*, 19(1–2):95–128, 2010.
29. Christian Straßer. A deontic logic framework allowing for factual detachment. *Journal of Applied Logic*, 9(1):61–80, 2011.
30. Christian Straßer and Mathieu Beirlaen. Towards more conflict-tolerant deontic logics by relaxing the interdefinability between obligations and permissions. Under review.
31. Christian Straßer, Joke Meheus, and Mathieu Beirlaen. Tolerating deontic conflicts by adaptively restricting inheritance. Under review.
32. Leendert van der Torre and Yao Hua Tan. Two-phase deontic logic. *Logique et Analyse*, 2(171-172):411–456, 2000.
33. Georg Henrik von Wright. *Norm and Action. A Logical Enquiry.* Routledge and Kegan Paul, London, 1963.
34. Georg Henrik von Wright. On the logic of norms and actions. In Risto Hilpinen, editor, *New Studies in Deontic Logic*, pages 3–35. D. Reidel Publishing Company, Dordrecht, 1981.
35. Georg Henrik von Wright. Deontic logic: a personal view. *Ratio Juris*, 12(1):26–38, 1999.