Non-Adjunctive Deontic Logics That Validate Aggregation as Much as Possible*

Joke Meheus, Mathieu Beirlaen,
Frederik Van De Putte, and Christian Straßer

Centre for Logic and Philosophy of Science
University of Ghent, Belgium
Joke.Meheus@UGent.be
Mathieu.Beirlaen@UGent.be
frvdeput.vandeputte@UGent.be
Christian.Strasser@UGent.be

October 8, 2010

1 Introduction

The aim of this paper is to present two new deontic logics that enable one to reason sensibly in the presence of normative conflicts. As is well-known, Standard Deontic Logic (henceforth **SDL**) presupposes that there are no conflicts between obligations. This is clearly seen from the fact that **SDL** validates the following principles

D $\vdash \neg(OA \wedge O\neg A)$
DEX $\vdash (OA \wedge O\neg A) \supset OB$

The explosion principle DEX is contained in any logic that contains all of

EFQ $\vdash (A \wedge \neg A) \supset B$
RM If $\vdash A \supset B$ then $OA \supset OB$
AND $(OA \wedge OB) \supset O(A \wedge B)$

Hence, any conflict-tolerant deontic logic has to invalidate at least one of these principles. One of the most common solutions is to invalidate or restrict the aggregation principle AND. Deontic logics that invalidate AND (or restrict it in some way) will be called non-adjunctive. Examples can be found in [4], [6], [7], [8], [13], [14].

*Research for this paper was supported by subventions from Ghent University and from the Research Foundation – Flanders (FWO - Vlaanderen). The authors are highly indebted to Lou Goble for his notes and comments on a previous paper. These made it possible to present the logics in this paper in a way that is intuitively more justified.

Given the way in which humans build up their norms, it seems realistic to suppose that they adhere to norms belonging to conflicting normative systems. Non-adjunctive deontic logics are especially suited to handle such cases, since they do not allow to derive $O(A \wedge \neg A)$ from OA and $O\neg A$. This makes good sense: the observation that two norms are in conflict, should not lead to the conclusion that there is a norm that forces one to do the impossible. Thus, the intuitive principle "Ought implies Can" can be preserved.

In [9, p. 466], Lou Goble stated that giving up aggregation is "perhaps the most natural suggestion for avoiding deontic explosion". In several papers Goble advocated the use of one particular such logic, namely the logic **P** [6], [7], [8]. **P** is a very well-behaved system and has a natural interpretation in a Kripke-like semantics.[1] It has moreover a nice axiomatization and avoids any kind of explosion when applied to conflicting obligations.

Still, the logic **P** has a serious drawback: it is too weak, especially when applied to obligations that are mutually compatible. For instance, in Horty's famous Smith example [10, p. 37], Smith is confronted with two obligations: (i) he ought to fight in the army or perform alternative service to his country ($O(F \vee S)$) and (ii) he ought not the fight in the army ($O\neg F$). As there is no conflict among these obligations, it seems reasonable to infer OS. Nevertheless, the logic **P**, as well as other non-adjunctive deontic logics, do not enable one to do so. In other words, simply invalidating aggregation results in a logic that is too weak.

At some point, Goble no longer considered **P** as the best solution for a conflict-tolerant deontic logic. However, given the attractiveness of a non-adjunctive approach to deontic conflicts, he and others made several attempts to restrict aggregation rather than to invalidate it. However, as Goble has shown in [9], all of them are inadequate (for instance, because all of them validate some form of explosion or lead to some otherwise counter-intuitive consequences). Moreover none of them can handle the Jones example, which is a simple variant of the Smith example. The obligations Jones is confronted with are: (i) he ought to pay taxes and fight in the army or perform alternative service ($O(T \wedge (F \vee S))$), and (ii) he ought not to pay taxes and not to fight in the army ($O(\neg T \wedge \neg F)$). In this case, Jones clearly faces a conflict concerning whether or not to pay his taxes. However, there is no problem concerning the 'Smith part' of his obligations. Hence, also in this case we would like to infer that Jones should perform alternative service.

Several solutions have been proposed for the Jones problem. One is to reformulate the premises. If the Jones example is reformulated as four separate obligations, then most non-adjunctive deontic logics can deal with it. But that seems like putting the cart before the horse. In complex cases, it requires *reasoning* to localize the conflicts and *this* reasoning now seems to be outside the scope of logic. Moreover, in such cases, formalizing the premises in the wrong way (because some conflicts were not detected), may still lead to explosion.

This led Goble to the following observation (personal communication):

> If the Jones Argument really should be taken to be valid in the form originally given, or at least if it really does appear to be valid in this form, then the challenge posed by normative conflicts seems

[1] The idea behind this semantics will be spelled out in Section 4. Goble also proposed a preferential semantics for **P** in [7] and in [6].

> to remain for the logic of ought. *We do not yet have an account that provides for all the inferences it should, if that includes the Jones argument, while at the same time avoiding the catastrophe of deontic explosion in the face of normative conflicts.* If this is correct, and none of the strategies considered so far yields a fully adequate logic for normative conflicts, I would, nevertheless, not draw the pessimistic conclusion that no such logic is possible, nor the optimistic conclusion that no such conflicts are possible. *It does seem, however, that if the Jones Argument stands, then more work must be done to determine the fundamental principles of the logic of ought.* (our emphasis)

In this paper, we shall present two logics, $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$, that can handle the Jones example (in its original form) as well as more complex ones. The only price to be paid is that one has to go adaptive (see below).

Both $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ are based on Goble's logic $\mathbf{SDL_aP_e}$ from [6], which we shall henceforth call **P2**. The system **P2** is a bimodal extension of the logic **P**. The language of the latter contains two sets of deontic operators: the operator O_e, which is the one from **P**, and the new operator O_a.[2] Goble's motivation for this additional ought-operator is that O_eA expresses that, under *some* set of norms, A ought to be case, but cannot express that A holds under *any* set of norms. The O_a-operator gives one exactly this. This results in a greater expressive power and also in different ways for formalizing conflicts (see Section 3). Another reading of the operators is that O_eA stands for the *prima facie* obligation to do A and that O_aA stands for the *actual* ("all-things-considered") obligation to do A. In line with what is common, we shall accept the idea that a *prima facie* obligation functions as an actual obligation, in case it is not incompatible with other obligations.

The logic **P2** behaves exactly like **SDL** for the O_a-operator and like **P** for the O_e-operator. This seems to give the logic some advantages over **P**. Given the proper formalization, one can make sure that for all non-conflicting 'parts' of the premises, the same results are obtained as with **SDL**. For instance, in the Smith example, formalizing the premises as $O_a(F \vee S)$ and $O_a\neg F$ ensures that O_aS is derivable. This solution presupposes, however, that one knows in advance which premises can be safely formalized with the O_a-operator. This in turn presupposes that one knows in advance which 'parts' of the premises are problematic.

The systems $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ are not the first adaptive logics that are based on **P**. In [12], we presented the logic $\mathbf{P2.1}^r$. At first sight, $\mathbf{P2.1}^r$ satisfied all desiderata. It has all the nice properties of **P**, it leads to the same consequence set as **SDL** for conflict-free premise sets, and it allows one to deal with both the Smith example and the Jones example, *in their original formulation.*

However, it turns out that $\mathbf{P2.1}^r$ does not entirely live up to its expectations. For simple examples, like the Jones example, it works fine. However, it breaks down for specific sets of more complex premises. Consider, for instance, the following premise set

(1) $O_e(p \vee q)$
(2) $O_e(r \vee s)$

[2] The duals P_e and P_a are defined in the usual way.

(3) $\neg O_a((p \vee q) \wedge (r \vee s))$
(4) $O_e t$

There is clearly an incompatibility between (1) and (2) in view of (3). However, there is nothing wrong with (4). Hence, one expects to be able to derive $O_a t$ from this premise set, but $\mathbf{P2.1}^r$ does not allow for this inference. The logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ solve this problem, while retaining all the nice properties of $\mathbf{P2.1}^r$.

The basic idea behind the two new logics is the same as that behind $\mathbf{P2.1}^r$: O_e-obligations are interpreted "as much as possible" as O_a-obligations (that is, unless and until the premises explicitly prevent this). Thus, *prima facie* obligations are interpreted as actual obligations, unless and until the context and the logic stop one from doing so. As is clear from the above, all classical operations can be applied to actual obligations (aggregation, disjunctive syllogism, ...). Which *prima facie* obligations are interpreted as actual obligations and which not is solely dependent on formal grounds. Note also that the logic adapts *itself* to the set of premises and localizes *itself* the conflicts. No interference of the user is required for this.

As mentioned above, the logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ are adaptive logics.[3] Both logics are non-monotonic and their proof theory is dynamical (conclusions derived at some stage of a proof may be rejected at a later stage),[4] but is sound and complete with respect to a (static) semantics.

2 Some Preliminaries

We shall use $\mathcal{L}$ to refer to the standard language of classical propositional logic and $\mathcal{S}$ to refer to the set of schematic letters. $\mathcal{L}^M$ is obtained from $\mathcal{L}$ by extending it with the modal operators O_e, O_a, P_e and P_a. Let "$\neg$", "$\vee$", "O_e" and "O_a" be primitive, the other logical constants being defined by

D1 $A \supset B =_{df} \neg A \vee B$
D2 $A \wedge B =_{df} \neg(\neg A \vee \neg B)$
D3 $A \equiv B =_{df} (A \supset B) \wedge (B \supset A)$
D4 $P_e A =_{df} \neg O_e \neg A$
D5 $P_a A =_{df} \neg O_a \neg A$

Where $\mathcal{W}$ is the set of all well-formed formulas of $\mathcal{L}$, the set of well-formed formulas of $\mathcal{L}^M$ is defined as the smallest set $\mathcal{W}^M$ that satisfies the following conditions:

(i) If $A \in \mathcal{W}$ then $A \in \mathcal{W}^M$
(ii) If $A \in \mathcal{W}$ then $O_e A$, $O_a A$, $P_e A$, $P_a A \in \mathcal{W}^M$
(iii) If $A \in \mathcal{W}^M$ then $\neg A \in \mathcal{W}^M$
(iv) If $A, B \in \mathcal{W}^M$ then $A \vee B, A \wedge B, A \supset B, A \equiv B \in \mathcal{W}^M$

[3]See [2] for an introduction to adaptive logics and [1] for an overview of their metatheoretic properties.

[4]A *stage* of a proof can be seen as a sequence of lines and a proof can be seen as a chain of stages. Every proof starts off with stage 1. Adding a line to a proof by applying one of the rules of inference brings the proof to its next stage, which is the sequence of all lines written down so far.

We shall use $\mathcal{W}^a$ to refer to the set of atoms (schematic letters and their negations). As is clear from the definition of $\mathcal{W}^M$, we restrict ourselves to first degree modalities. This simplifies the characterization of the logic and does not cause too much harm—nearly all papers on deontic conflicts constrain themselves to first degree modalities (either explicitly of implicitly).

There is one aspect in which we shall differ in our presentation of the logic **P2**. Goble is only interested in the theorems of his logic, not in a semantic consequence relation. As we are mainly interested in the consequence relation, we shall modify his semantics in such a way that we introduce a real world in the models. This will also make it easier to explain how the adaptive logic works.

3 Incompatible Obligations

In [12], we defended a view on normative conflicts that is more general than what is common in the literature. On our view, not all normative conflicts are *moral* conflicts (they can come from legal codes, moral codes, traffic regulations, promises, ...). We also do not restrict ourselves to cases where, for some A, both OA and $O\neg A$ hold. We also allow, for instance, for conflicts of the form $O\neg A \wedge PA$ and of the form $OA,\ OB,\ O(\neg A \vee \neg B)$. Finally we do not assume that all normative conflicts can be *reduced* to direct conflicts (that is, conflicts of the form $OA \wedge O\neg A$).

The latter seems to be the position of Goble. On the one hand, he allows for situations where two obligations are jointly incompatible (OA, OB and $\vdash \neg(A \wedge B)$ hold). On the other hand, he considers situations where two obligations are jointly impossible (OA, OB and $\neg\Diamond(A \wedge B)$ hold). However, he argues that both cases are reducible to situations where two direct conflicts hold: $OA \wedge O\neg A$ and $OB \wedge O\neg B$. For this reduction, he relies on RM as well as on the following assumption (see [9, p. 462]):

NM $\ \vdash \neg\Diamond(A \wedge \neg B) \supset (OA \supset OB)$

The reduction Goble has in mind is typically applied to cases where two obligations are jointly incompatible or jointly impossible. In the case of the drowning twins, for example, one has to imagine a situation where two identical twins are drowning and the situation is such that one can save either of them, but one cannot save both of them.[5] In Goble's view, the impossibility to save both, reduces the normative conflict to two direct conflicts: "one ought to save the first twin and one ought not to save the first twin" and analogously for the second twin.

In the rest of this section, we shall concentrate on incompatible obligations. We shall distinguish two types.

The first type consists of conflicts of the form $OA_1,\ \ldots,\ OA_n,\ O\neg(A_1 \wedge \ldots \wedge A_n)$. The idea is that $A_1,\ \ldots,\ A_n$ can be jointly fulfilled, but there is an additional obligation not to fulfill them all. As a simple example of this type, consider the situation where Bob, at different moments in time, promised his two best friends, John and Peter, to invite them to his birthday party. However, he also promised his girlfriend not to invite them both. (John and Peter are known

[5] As a more realistic example, one may think of the kind of heartbreaking decision some parents have to make in the case of Siamese twins.

to quarrel over almost anything and Bob's girlfriend is afraid that this may put a damper on the party.) As there is no reason in this case to prefer one obligation above the other (all three stem from promises made at different moments in time), we formalize all obligations involved as *prima facie* obligations:

(5) Bob has a *prima facie* obligation to invite John — $O_e I_j$
(6) Bob has a *prima facie* obligation to invite Peter — $O_e I_p$
(7) Bob has a *prima facie* obligation not to invite both Peter and John — $O_e \neg(I_j \wedge I_p)$

The second type consists of conflicts where the joint fulfillment of a certain number of obligations is not merely forbidden, but simply impossible. As an example, one may again think of the drowning twins. Another example is that of the marrying daughters. Charlotte promised both her daughters to be present at their wedding, but unfortunately they planned their wedding on the same day in two different continents, so that Charlotte cannot possibly attend both weddings.

Given the language of **P2**, there are different ways to formalize incompatible obligations of the second type. We shall concentrate on two of them. A formalization that immediately comes to mind is to express the impossibility to fulfill a certain number of obligations by the universal obligation not to fulfill them all. This would give us the following formalization in the drowning twin case:

(8) I have a *prima facie* obligation to save the first twin — $O_e T_1$
(9) I have a *prima facie* obligation to save the second twin — $O_e T_2$
(10) I have the *universal* obligation not to save both — $O_a \neg(T_1 \wedge T_2)$

At first sight, this formalization seems appealing: the universal obligation seems to capture the idea that it is impossible to save both twins (that is, that there is no accessible world in which both twins are saved).

However, there are several objections possible. The first is that it leads to the same kind of reduction that we discussed above. Given (8)–(10), it follows that I have the obligation to save the first twin and also the obligation not to save him. The second concerns the notion of a "deontically perfect world". The above formalization leads to a very strong restriction on what counts as a deontically perfect alternative for the actual world. One not only has to assume that a deontically perfect world has at least the same natural laws as the actual world (which is a reasonable requirement), but also that its history is exactly as our world's history *up to the point* where at least one of the twins can no longer be saved.

Here lies the difficulty. It is a reasonable requirement that a deontically perfect world has the same past, but a different future than our world. But where shall we draw the line? After all, falling in the water and drowning is not an instantaneous process. If we allow that the histories of the accessible worlds diverge from one another at an earlier point in time than the moment where at least one of the twins is actually dying, things are different. In that case, there are accessible worlds in which both twins are saved (for instance, the world where at the crucial moment one of my friends passes by and we each save one of the twins).

In view of this, we favour a weaker formalization of the twin example: we only require that it is not an actual obligation to save both. Thus, instead of (10), we obtain

(11) I do *not* have the universal obligation to save both — $\neg O_a(T_1 \wedge T_2)$

This formalization has several advantages. One is that the link between the two incompatible obligations is preserved: there is no reduction to a number of direct conflicts. As we shall see below, this allows us to follow different 'strategies' when dealing with incompatible obligations of the second type. It also nicely agrees with a certain interpretation of the "Ought implies Can" principle. A state of affairs that is impossible to realize (in our world) should not be a universal obligation. This is captured by (11).

4 Rejecting Aggregation: The Logic P2

Let us now turn to the logic that will form the basis of our adaptive logic. The idea behind **P2** is actually very simple: in a Kripke-like semantics, aggregation is invalidated by considering a *set* of accessibility relations instead of only one. Intuitively, each accessibility relation can be thought of as corresponding to one of the normative systems an agent adheres to.

A **P2**-model M is a quadruple $\langle W, \mathcal{R}, v, w_0 \rangle$ where W is a set of possible worlds, $\mathcal{R}$ is a non-empty set of serial accessibility relations R on W, $v : \mathcal{S} \times W \rightarrow \{0, 1\}$ is an assignment function, and $w_0 \in W$ is the real world. The valuation v_M defined by the model M is characterized by:

C1 where $A \in \mathcal{S}$, $v_M(A, w) = v(A, w)$
C2 $v_M(\neg A, w) = 1$ iff $v_M(A, w) = 0$
C3 $v_M(A \vee B, w) = 1$ iff $v_M(A, w) = 1$ or $v_M(B, w) = 1$
C4 $v_M(O_e A, w) = 1$ iff, for some $R \in \mathcal{R}$, $v_M(A, w') = 1$ for all w' such that Rww'
C5 $v_M(O_a A, w) = 1$ iff, for every $R \in \mathcal{R}$, $v_M(A, w') = 1$ for all w' such that Rww'

A **P2**-model M verifies A ($M \Vdash A$) iff $v_M(A, w_0) = 1$. M is a **P2**-model of Γ iff $M \Vdash A$ for all $A \in \Gamma$, and $\Gamma \vDash_{\mathbf{P2}} A$ iff all **P2**-models of Γ verify A.

P2 is axiomatized by extending an axiomatization of classical propositional logic with the following axioms and rules:

K_a	$O_a(A \supset B) \supset (O_a A \supset O_a B)$
D_a	$O_a A \supset \neg O_a \neg A$
RN_a	if $\vdash A$ then $\vdash O_a A$
RM_e	if $\vdash A \supset B$ then $\vdash O_e A \supset O_e B$
N_e	if $\vdash A$ then $\vdash O_e A$
P_e	if $\vdash A$ then $\vdash \neg O_e \neg A$
K_{ae}	$O_a(A \supset B) \supset (O_e A \supset O_e B)$

The first three postulates deliver **SDL** for O_a and the next three deliver **P** for O_e.[6] The last axiom links the two operators.

[6] **P** is as **P2**, except that there is only one O-operator and that obviously C5 does not hold in it.

We define $\Gamma \vdash_{\mathbf{P2}} A$ iff A is derivable from Γ by the axioms and rules of **P2**. In [6], Goble proved soundness and (weak) completeness for **P2**.

Theorem 1 *For any finite* $\Gamma \subset \mathcal{W}$, $\Gamma \vdash_{\mathbf{P2}} A$ *iff* $\Gamma \vDash_{\mathbf{P2}} A$.

5 Informal Presentation of the Logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$

5.1 Informal Presentation of the Logic $\mathbf{P2.2}^r$

The logic $\mathbf{P2.2}^r$ is an adaptive extension of the logic **P2**. The logic **P2** constitutes the stable part of $\mathbf{P2.2}^r$: anything that is **P2**-derivable from a premise set is derivable in $\mathbf{P2.2}^r$. In addition to this, it is allowed that O_e-obligations are interpreted as O_a-obligations whenever they behave "normally". A first approximation of this idea is that it is allowed that O_aA is derived from O_eA *unless* $O_eA \wedge \neg O_aA$ is **P2**-derivable from the premises. A formula of the form $O_eA \wedge \neg O_aA$ will be called an *abnormality*—it is a formula that blocks a desired inference (in this case the transition from O_eA to O_aA). We shall see below that several restrictions are needed with respect to the abnormalities and that we also need a more sophisticated notion of "normal" behavior. But let us first illustrate the main ideas by means of an example.

Suppose that Johnson faces the following three *prima facie* obligations:

O1 he ought to pay taxes, and fight in the army or perform alternative service to his country — $O_e(T \wedge (F \vee S))$
O2 he ought not to pay taxes and not to fight in the army — $O_e(\neg T \wedge \neg F)$
O3 he ought to pay taxes or donate to charity — $O_e(T \vee C)$

In order to localize the conflicts and to see what follows, we start a $\mathbf{P2.2}^r$-proof by introducing first the premises:

1	$O_e(T \wedge (F \vee S))$	PREM	$\emptyset$
2	$O_e(\neg T \wedge \neg F)$	PREM	$\emptyset$
3	$O_e(T \vee C)$	PREM	$\emptyset$

The only unusual element in this proof is in the last column. This element is called the *condition* of the line at issue and is always empty in the case of premises. Its function will become clear below.

Suppose that we continue the proof as follows:

4	$O_e(F \vee S)$	1; RU	$\emptyset$
5	$O_e\neg F$	2; RU	$\emptyset$
6	O_eT	1; RU	$\emptyset$
7	$O_e\neg T$	2; RU	$\emptyset$

Each of these formulas follows by **P2** from the premises and hence can be unconditionally derived in the proof. The rule RU is a generic rule that allows one to derive any formula that is **P2**-derivable.

In view of these formulas, it seems intuitively clear that we want to derive O_eS and even O_aS from $O_e(F \vee S)$ and $O_e\neg F$, but that we do not want to derive O_eC or O_aC from $O_e(T \vee C)$ and $O_e\neg T$. The reason is that there is a conflict in the second case (see lines 6 and 7), but not in the first case. We shall see below that $\mathbf{P2.2}^r$ gives us precisely this outcome. But first we need to discuss some small complications.

A first complication is that from some sets of conflicting deontic statements no formula of the form $O_eA \wedge \neg O_aA$ is **P2**-derivable. For instance, from the set of premises $\{O_ep,\ O_eq,\ O_e\neg(p \wedge q)\}$, no single formula of the form $O_eA \wedge \neg O_aA$ is derivable, but $(O_ep \wedge \neg O_ap) \vee (O_eq \wedge \neg O_aq)$ is. It is in view of such cases that the expression "to interpret a set of premises as normally as possible" becomes ambiguous. It is disambiguated by the *adaptive strategy*. In the case of $\mathbf{P2.2}^r$, the strategy is Reliability.[7] To explain this strategy, we first need some definitions.

Where Δ is a non-empty, finite set of abnormalities, the disjunction $\bigvee(\Delta)$ will be called a *Dab*-formula and will be written as $Dab(\Delta)$.[8] A *Dab*-formula $Dab(\Delta)$ will be called a *minimal Dab*-formula at stage s of a proof, if, at that stage of the proof, $Dab(\Delta)$ is derived on the condition $\emptyset$, and no $Dab(\Delta')$ such that $\Delta' \subset \Delta$ is derived on the condition $\emptyset$.

What the Reliability Strategy comes to is that, whenever a *minimal Dab*-formula is unconditionally derived in a proof at a certain stage, then all disjuncts that occur in that *Dab*-formula are considered as behaving abnormally and are therefore considered unreliable. As we shall see below, the unreliable formulas at a stage s determine which lines (if any) should be *marked*. Intuitively, a line is marked if its condition is violated. According to the Reliability Strategy, a condition is violated at a certain stage if at that stage its condition contains an unreliable formula. Formulas that occur on marked lines are not considered as derived in the proof.

The second complication is that we need some fine-tuning on the form of the abnormalities. By simply using $O_eA \wedge \neg O_aA$ as the logical form for the abnormalities, we would obtain an extremely weak logic. The reason for this is easily demonstrated by means of the following example. Consider $\Gamma = \{O_ep,\ O_e\neg p,\ O_eq\}$. As there is clearly no conflict with respect to O_eq, O_aq should be $\mathbf{P2.2}^r$-derivable from Γ. However, the *Dab*-formula $(O_eq \wedge \neg O_aq) \vee (O_e(\neg p \vee \neg q) \wedge \neg O_a(\neg p \vee \neg q))$ is **P2**-derivable from Γ, whereas neither of its disjuncts is. Hence, this *Dab*-formula is minimal and, in view of the Reliability Strategy, $O_eq \wedge \neg O_aq$ is unreliable. This blocks the desired inference from O_eq to O_aq. That is why we cannot simply take all formulas of the form $O_eA \wedge \neg O_aA$ to be abnormalities.

A solution to this particular problem would be to restrict the abnormalities to the set $\{O_eA \wedge \neg O_aA \mid A \in \mathcal{W}^a\}$. However, in doing so, we would obtain an adaptive logic that is too poor. It would, for instance, not be possible to infer $O_a(p \vee q)$ from $O_e(p \vee q)$. This brings us to the question when we should consider it as an abnormality that an obligation of the form $O_e(A_1 \vee \ldots \vee A_n)$ (with $A_1, \ldots, A_n \in \mathcal{W}^a$ and $n \geq 2$) cannot be generalized to $O_a(A_1 \vee \ldots \vee A_n)$. A natural answer to this question is that we have an abnormality when $O_e(A_1 \vee \ldots \vee A_n)$ is true whereas $O_a(A_1 \vee \ldots \vee A_n)$ is false *or* there is a

[7] The two most common strategies for adaptive logics are the Reliability Strategy and the Minimal Abnormality Strategy—the former is a bit more cautious than the latter—see [1]. In Section 5.2, we shall illustrate the Minimal Abnormality strategy by means of an example.

[8] Note that it is allowed that Δ is a singleton. In that case, $Dab(\Delta)$ is simply the formula that is contained in Δ.

'shorter' abnormality from which $O_e(A_1 \vee \ldots \vee A_n)$ is obtained. Thus, where $\Gamma = \{O_e p,\ O_e \neg p,\ O_e q\}$, it would not count as an abnormality that $O_e(p \vee r)$ (which is **P2**-derivable from Γ) cannot be generalized to $O_a(p \vee r)$ (in view of the conflict between $O_e p$ and $O_e \neg p$), but it would count as an abnormality that $O_e(q \vee r)$ cannot be generalized to $O_a(q \vee r)$ (since there is no conflict regarding q or r).

This brings us to a more refined type of abnormalities.[9] Where Θ is a finite and non-empty set and $\Theta \subseteq \mathcal{W}^a$, and where $\sigma(\Theta) = \{O_e(\bigvee \Theta') \wedge \neg O_a(\bigvee \Theta') \mid \Theta' \subseteq \Theta \text{ and } \Theta' \neq \emptyset\}$, the form of the abnormalities is $\bigvee(\sigma(\Theta))$. As an example, consider the set $\Theta = \{p,\ q,\ \neg r\}$. In that case, $\bigvee(\sigma(\Theta))$ stands for the formula[10] $(O_e(p \vee q \vee \neg r) \wedge \neg O_a(p \vee q \vee \neg r)) \vee (O_e(p \vee q) \wedge \neg O_a(p \vee q)) \vee (O_e(p \vee \neg r) \wedge \neg O_a(p \vee \neg r)) \vee (O_e(q \vee \neg r) \wedge \neg O_a(q \vee \neg r)) \vee (O_e p \wedge \neg O_a p) \vee (O_e q \wedge \neg O_a q) \vee (O_e \neg r \wedge \neg O_a \neg r)$. For reasons of transparency, we shall in the remainder use $\dagger(p \vee q \vee \neg r)$ instead of $\bigvee(\sigma(\{p, q, \neg r\}))$ More generally, we shall use $\dagger(A_1 \vee \ldots \vee A_n)$ (where $n \geq 1$) instead of $\bigvee(\sigma(\{A_1, \ldots, A_n\}))$.

We can now return to our Johnson example. One way to continue the proof is as follows:

8	$O_a \neg F$	5; RC	$\{\dagger(\neg F)\}$
9	$O_a(F \vee S)$	4; RC	$\{\dagger(F \vee S)\}$
10	$O_a S$	8, 9; RU	$\{\dagger(\neg F), \dagger(F \vee S)\}$

Lines 8 and 9 are applications of the conditional rule RC. This is a rule that leads to the introduction of a new condition. Note that $(O_e \neg F \supset O_a \neg F) \vee (O_e \neg F \wedge \neg O_a \neg F)$ is **P2**-derivable from the premises. A way to read this is: $O_e \neg F$ entails $O_a \neg F$, or $O_e \neg F \wedge \neg O_a \neg F$ is true. This is the motor behind the adaptive proof theory: abnormalities are assumed to be false unless and until proven otherwise. Thus we derive $O_a \neg F$ from $O_e \neg F$ on the condition that $\dagger(\neg F)$ is false. If at some point in the proof the condition of line 8 is no longer fulfilled, then this line is marked, indicating that the formula at that line is no longer considered as derived in the proof. An analogous reasoning holds for line 9. In this case, $(O_e(F \vee S) \supset O_a(F \vee S)) \vee \dagger(F \vee S)$ is **P2**-derivable from the premises, and also here, the abnormalities are assumed to be false unless and until proven otherwise.

Line 10 features an application of the unconditional rule RU. Note that when the unconditional rule is applied, no new abnormalities are added to the condition, but any formula that occurs in a non-empty condition is 'carried' over to the condition of the conclusion of the application. The reason for this is easy to understand. If, at some point, line 8 or line 9 is marked (because its condition is no longer satisfied), then evidently any line that depends on it, should also be marked.

The following continuations of the proof are meant to illustrate that neither $O_a T$ nor $O_a C$ are $\mathbf{P2.2}^r$-derivable. Analogous to line 8, the conditional rule allows one to add a line to the proof on which $O_a \neg T$ is derived on the appropriate condition:

11	$O_a \neg T$	7; RC	$\{\dagger(\neg T)\}$

[9] The form of the abnormalities is inspired by an idea of Lou Goble (personal communication).

[10] For reasons of readability, we use continuous disjunctions.

At this stage of the proof, the formula $O_a\neg T$ is considered as derived. We may also continue the proof as follows

12	$O_a(T \vee C)$	3; RC	$\{\dagger(T \vee C)\}$
13	O_aC	11, 12; RU	$\{\dagger(\neg T), \dagger(T \vee C)\}$

Things change, however, as soon as the following lines are added:

14	$\dagger(\neg T)$	6, 7; RU	$\emptyset$
15	$\dagger(T \vee C)$	6, 7; RU	$\emptyset$

These lines make it clear that the conditions of lines 11–13 are not fulfilled, and hence, they are marked:

11	$O_a\neg T$	7; RC	$\{\dagger(\neg T)\}\checkmark^{14}$
12	$O_a(T \vee C)$	3; RC	$\{\dagger(T \vee C)\}\checkmark^{15}$
13	O_aC	11, 12; RU	$\{\dagger(\neg T), \dagger(T \vee C)\}\checkmark^{14}$
14	$\dagger(\neg T)$	6, 7; RU	$\emptyset$
15	$\dagger(T \vee C)$	6, 7; RU	$\emptyset$

From stage 15 on, O_aT, $O_a(T \vee C)$ and O_aC are no longer considered to be derived in the proof. Since the *Dab*-formulas on lines 14 and 15 cannot be 'shortened', lines 11–13 will remain marked in any extension of the proof. For this simple example, it is also easy to see that lines 8–10 will not be marked in any extension of the proof. This is why we say that the formulas on these lines are *finally* derived with respect to the premises 1–3.[11] Note especially that O_aS is $\mathbf{P2.2}^r$-derivable from the premises, even though it is 'connected' to a problematic obligation.

It was hinted at in the introduction that the proof theory of $\mathbf{P2.2}^r$ is dynamical. What this comes to is that lines may be unmarked at some stage in the proof, marked at a later stage and sometimes again unmarked at a still later stage. As is usual for adaptive logics, a distinction can be made between an internal dynamics and an external dynamics. The internal dynamics occurs when lines are marked (respectively, unmarked) because of new insights in the premises (for instance, the marking of line 11 in view of line 14 in the above proof). The external dynamics occurs when the markings change due to the addition of new premises. In the remainder of this section, we shall illustrate the external dynamics.

Suppose that Johnson, after a reasoning process that is explicated by the above proof, discusses the matter with his girlfriend and that she convinces him that he ought not to perform alternative service to his country. This new premise brings us in a new situation: whereas lines 8–10 are finally derivable with respect to the premises 1–3, they are no longer finally derivable when this new premise is added at line 16.

8	$O_a\neg F$	5; RC	$\{\dagger(\neg F)\}\checkmark^{17}$
9	$O_a(F \vee S)$	4; RC	$\{\dagger(F \vee S)\}\checkmark^{18}$
10	O_aS	8, 9; RU	$\{\dagger(\neg F), \dagger(F \vee S)\}\checkmark^{17}$
11	$O_a\neg T$	7; RC	$\{\dagger(\neg T)\}\checkmark^{14}$

[11] The precise definition of *final derivability* follows in the next section.

12	$O_a(T \vee C)$	3; RC	$\{\dagger(T \vee C)\}\checkmark^{15}$
13	$O_a C$	11, 12; RU	$\{\dagger(\neg T), \dagger(T \vee C)\}\checkmark^{14}$
14	$\dagger(\neg T)$	6, 7; RU	$\emptyset$
15	$\dagger(T \vee C)$	6, 7; RU	$\emptyset$
16	$O_e \neg S$	PREM	$\emptyset$
17	$\dagger(\neg F) \vee \dagger(\neg S)$	1, 2, 16, RU	$\emptyset$
18	$\dagger(F \vee S)$	1, 2, 16, RU	$\emptyset$

What this illustrates is that the formulas on lines 8–10 are finally derivable with respect to the premises on lines 1–3, but not with respect to the premises on lines 1–3 and 16.

Like the logic $\mathbf{P2.1^r}$ presented in [12], the logic $\mathbf{P2.2}^r$ can handle the Johnson example in its formalization above, without adding any allegedly 'hidden' or 'tacit' premises. To the best of our knowledge, these are the first non-adjunctive deontic logics that have this property. There are other logics that can handle the Smith case, but they can only deal with the Johnson example by reformulating the premises as $\{O_eT,\ O_e(F \vee S),\ O_e\neg T,\ O_e\neg F,\ O_e(T \vee C)\}$, or by extending the premise set with 'hidden' assumptions. In the former case, crucial information may be lost. In the latter case it may be far from evident—especially in very complex situations—which premises can be added safely, and in making the wrong 'guesses' one may cause explosion after all. This is why we consider it important to start from the original premises and leave their analysis to the logic itself (see also the discussion in [12], p. 160).

5.2 Informal Presentation of the Logic $\mathbf{P2.2}^m$

The logic $\mathbf{P2.2}^m$, like $\mathbf{P2.2}^r$, is an adaptive extension of **P2** that allows for the interpretation of O_e-obligations as O_a-obligations whenever the former "behave normally". The difference between the two adaptive systems is that they do not disambiguate the latter expression in the same way. Whereas $\mathbf{P2.2}^r$ uses the Reliability strategy in order to formally specify the idea of behaving "normally", $\mathbf{P2.2}^m$ uses the Minimal Abnormality strategy.

For now, we shall simply illustrate the difference between the logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ by means of an example. This illustration will be quite informal. The technical details are spelled out in the next section.

As an illustration of the difference between the two logics, we use a formalization of the drowning twins example that we already presented in Section 3. Where T_1, respectively T_2, stands for the obligation to save the first, respectively the second twin, we face the following three normative statements:

1	O_eT_1	PREM	$\emptyset$
2	O_eT_2	PREM	$\emptyset$
3	$\neg O_a(T_1 \wedge T_2)$	PREM	$\emptyset$

Using the conditional rule, we can add lines to the proof at which O_aT_1 and O_aT_2 are derived on the appropriate condition:

4	O_aT_1	1; RC	$\{\dagger(T_1)\}$
5	O_aT_2	2; RC	$\{\dagger(T_2)\}$

Since the minimal *Dab*-formula $\dagger(T_1)\vee\dagger(T_2)$ is a **P2**-consequence of our premises, lines 4 and 5 above are marked once we have derived the formula on line 6:

4	$O_a T_1$	1; RC	$\{\dagger(T_1)\}\checkmark^6$
5	$O_a T_2$	2; RC	$\{\dagger(T_2)\}\checkmark^6$
6	$\dagger(T_1) \vee \dagger(T_2)$	1-3; RU	$\emptyset$

So far, the above proof might as well be a $\mathbf{P2.2}^r$-proof: in view of line 6, lines 4 and 5 are marked according to both the Reliability and the Minimal Abnormality strategy. Things change, however, when we make the following move:

7	$O_a(T_1 \vee T_2)$	4; RU	$\{\dagger(T_1)\}$
8	$O_a(T_1 \vee T_2)$	5; RU	$\{\dagger(T_2)\}$

The disjunctive universal obligation $O_a(T_1 \vee T_2)$ is derivable on the condition $\{\dagger(T_1)\}$ and on the condition $\{\dagger(T_2)\}$. According to the Reliability strategy, both of these conditions are considered unreliable. This causes the marking of lines 7 and 8 (in view of line 6) in a $\mathbf{P2.2}^r$-proof.

However, not so in a $\mathbf{P2.2}^m$ proof. According to the Minimal Abnormality strategy, we need not assume that *both* obligations in the minimal *Dab*-formula on line 6 behave abnormally. As only one of the disjuncts has to be true (in order for the premises to be true), we can assume that one of the obligations behaves normally. So, if, on the one hand, the formula $\dagger(T_1)$ in the condition of line 7 is considered as true, we can assume that the formula in the condition of line 8 is false. If, on the other hand, the formula $\dagger(T_2)$ in the condition of line 8 is considered as true, we can assume that the formula in the condition of line 7 is false. What this comes to is that, on either assumption, $O_a(T_1 \vee T_2)$ is true. Hence, $O_a(T_1 \vee T_2)$ is a $\mathbf{P2.2}^m$-consequence of our premises. In Section 6, we shall discuss in a formally precise way why, with Minimal Abnormality, lines 7 and 8 are unmarked.

In the twin example, this outcome is desirable: even though we cannot save both twins, we still face the actual obligation to save at least one of them.[12] In other contexts, it may be undesirable to derive the universal obligation to fulfill at least one of the incompatible obligations. In the wedding example, for instance, one may argue that Charlotte should go to neither of the weddings, because otherwise one of the daughters may feel distressed. This outcome can be obtained by applying $\mathbf{P2.2}^r$ instead of $\mathbf{P2.2}^m$.

6 Formal Characterization of the Logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$

In this section, we present the logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ in a formally precise way. As any other adaptive logic in standard format, they are characterized by a triple: a lower limit logic (a reflexive, transitive, monotonic, uniform, and compact logic for which there is a positive test)[13], a set of abnormalities Ω

[12] Several authors have argued that, in case of a conflict between two *prima facie* obligations $O_e A$ and $O_e B$, the all-things-considered obligation $O_a(A \vee B)$ should be derivable. See, for instance, [3], [5], [11].

[13] There is a *positive test* for objects of a given kind iff there is a mechanical procedure that leads to the answer YES if the property holds. If the property does not hold the procedure may lead to the answer NO, but may continue forever.

(characterized by a, possibly restricted, logical form) and an adaptive strategy. In the case of $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$, the lower limit logic is **P2**.

Where Θ is a finite and non-empty set of atoms, let $\sigma(\Theta)$ be defined as in Section 5. The set of abnormalities is defined as follows:

$$\Omega = \left\{ \bigvee(\sigma(\Theta)) \mid \Theta \subseteq \mathcal{W}^a, \Theta \neq \emptyset, \Theta \text{ is finite} \right\}$$

As explained before, the adaptive logics interpret sets of premises as normally as possible, in a sense specified by the adaptive strategy. This strategy is Reliability for $\mathbf{P2.2}^r$ and Minimal Abnormality for $\mathbf{P2.2}^m$.

In order to define the semantics of these logics, we need some further definitions. We shall say that a *Dab*-formula $Dab(\Delta)$ is a *Dab*-consequence of Γ if it is **P2**-derivable from Γ and that it is a *minimal Dab*-consequence if there is no $\Delta' \subset \Delta$ such that $Dab(\Delta')$ is also a *Dab*-consequence of Γ. The set of formulas that are *unreliable* with respect to Γ, denoted by $U(\Gamma)$, is defined by

Definition 1 *Where $Dab(\Delta_1)$, $Dab(\Delta_2)$, ... are the minimal Dab-consequences of Γ, $U(\Gamma) = \Delta_1 \cup \Delta_2 \cup \ldots$ is the set of formulas that are* unreliable *with respect to Γ.*

Next we define the abnormal part of a **P2**-model M:

Definition 2 $Ab(M) = \{A \in \Omega \mid M \Vdash A\}$

Together, these definitions allow us to select the reliable models and to define the semantic consequence relation for $\mathbf{P2.2}^r$:

Definition 3 *A **P2**-model M of Γ is* reliable *iff $Ab(M) \subseteq U(\Gamma)$.*

Definition 4 *$\Gamma \vDash_{\mathbf{P2.2}^r} A$ iff A is verified by all reliable models of Γ*

The definition of a minimally abnormal model only refers to the abnormal part of the **P2**-models:

Definition 5 *A **P2**-model M of Γ is* minimally abnormal *iff there is no **P2**-model M' of Γ such that $Ab(M') \subset Ab(M)$.*

The semantic consequence relation for $\mathbf{P2.2}^m$ is defined in terms of the minimally abnormal models:

Definition 6 *$\Gamma \vDash_{\mathbf{P2.2}^m} A$ iff A is verified by all minimally abnormal models of Γ.*

As is common for all adaptive logics in standard format, the proof theory of **P2.2** is characterized by three generic inference rules and a marking definition. The inference rules only refer to the lower limit logic, in our case **P2**. Where Γ is the set of premises, the inference rules are given by

PREM If $A \in \Gamma$:

$$\begin{array}{ll} \ldots & \ldots \\ \hline A & \emptyset \end{array}$$

RU If $A_1, \ldots, A_n \vdash_{\mathbf{P2}} B$:

$$\begin{array}{ll} A_1 & \Delta_1 \\ \ldots & \ldots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \end{array}$$

RC If $A_1, \ldots, A_n \vdash_{\mathbf{P2}} B \vee Dab(\Theta)$

$$\begin{array}{ll} A_1 & \Delta_1 \\ \ldots & \ldots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \cup \Theta \end{array}$$

The premise rule PREM simply states that, at any stage of a proof, a premise may be introduced on the empty condition. What the unconditional rule RU comes to is that whenever $A_1, \ldots, A_n \vdash_{\mathbf{P2}} B$ and $A_1, \ldots, A_n$ occur in the proof on the conditions $\Delta_1, \ldots, \Delta_n$, then B may be added to the proof on the condition $\Delta_1 \cup \ldots \cup \Delta_n$. The conditional rule RC is analogous, except that here a new condition is introduced.

The marking definitions proceed in terms of the *minimal Dab-formulas* derived at a stage of the proof:

Definition 7 *$Dab(\Delta)$ is a* minimal *Dab*-formula *at stage s iff, at stage s, $Dab(\Delta)$ is derived on the condition $\emptyset$, and no $Dab(\Delta')$ with $\Delta' \subset \Delta$ is derived on condition $\emptyset$.*

In order to define the marking for Minimal Abnormality, we also need the notion of a choice set. A *choice set* of a set of sets Σ is a set that contains a member of each set in Σ, and a *minimal choice set* of Σ is a choice set of Σ of which no proper subset is a choice set of Σ.

Definition 8 *Where $Dab(\Delta_1)$, $\ldots$, $Dab(\Delta_n)$ are the minimal Dab-formulas at stage s, $U_s(\Gamma) = \Delta_1 \cup \ldots \cup \Delta_n$ and $\Phi_s(\Gamma)$ is the set of minimal choice sets of $\Sigma = \{\Delta_1, \ldots, \Delta_n\}$.*

In view of these definitions, the marking definitions for both adaptive strategies are given by:

Definition 9 (Marking for Reliability) *Where Δ is the condition of line l, line l is marked at stage s iff $\Delta \cap U_s(\Gamma) \neq \emptyset$.*

Definition 10 (Marking for Minimal Abnormality) *A line l with formula A is marked at stage s iff, where its condition is Δ: (i) there is no $\varphi \in \Phi_s(\Gamma)$ such that $\varphi \cap \Delta = \emptyset$, or (ii) for a $\varphi \in \Phi_s(\Gamma)$, there is no line on which A is derived on a condition Θ for which $\Theta \cap \varphi = \emptyset$.*

The last definition enables us to spell out more clearly how the marking works in the example of the drowning twins (see Section 5.2). Consider first line 7 of the proof. Where $\Gamma = \{O_e T_1,\ O_e T_2,\ \neg O_a(T_1 \wedge T_2)\}$, $\Phi_8(\Gamma) = \{\{\dagger(T_1)\}, \{\dagger(T_2)\}\}$. As $\{\dagger(T_2)\}$ does not overlap with the condition of line 7, the first criterion to mark line 7 is not met. Moreover, as $O_a(T_1 \vee T_2)$ is derived on the condition

$\{\dagger(T_1)\}$ as well as on the condition $\{\dagger(T_2)\}$, also the second criterion for marking line 7 is not met. An analogous reasoning applies to line 8 of the proof.

A formula A is said to be derived at stage s of a proof if it occurs on a line in the proof that is unmarked at stage s. As the marking proceeds in terms of the minimal *Dab*-formulas that are derived at a certain stage, it is clear that marking is a dynamic matter: a line may be unmarked at a stage s, marked at a later stage s' and again unmarked at an even later stage s''. This is why a more stable notion of derivability is needed:

Definition 11 *A is* finally derived *from Γ at line i of a proof at stage s iff A is derived on line i at stage s and every extension of the proof in which line i is marked has an extension in which i is unmarked.*

Let in the remainder **P2.2** be a generic term referring to either of the two adaptive systems $\mathbf{P2.2^r}$ or $\mathbf{P2.2^m}$. As may be expected, the derivability relation of **P2.2** is defined with respect to the notion of final derivability

Definition 12 *$\Gamma \vdash_{\mathbf{P2.2}} A$ (A is finally derivable from Γ) iff A is finally derived in an **P2.2**-proof from Γ.*

For all adaptive logics in standard format, soundness and completeness are warranted in view of the soundness and completeness of the lower limit logic—see [1] for the proofs. Hence in view of Theorem 1 we obtain:

Theorem 2 *For every finite $\Gamma \subset \mathcal{W}$: $\Gamma \vdash_{\mathbf{P2.2}} A$ iff $\Gamma \vDash_{\mathbf{P2.2}} A$.*

The fact that **P2.2** is in standard format moreover warrants that it has a number of other meta-theoretic properties, such as proof invariance:[14]

Theorem 3 *If $\Gamma \vdash_{\mathbf{P2.2}} A$, then every **P2.2**-proof from Γ can be extended in such a way that A is finally derived in it.*

Whenever there is a **P2**-model of some Γ then there is a reliable, respectively minimally abnormal, model of Γ. Hence, the adaptive logic **P2.2** only trivializes a premise set if its lower limit logic does. Moreover, every **P2**-model M of Γ is either reliable or there is a reliable model M' of Γ such that $Ab(M') \subset Ab(M)$. This immediately entails that every minimally abnormal model is reliable. Hence, the following theorem holds.

Theorem 4 *For every Γ, $\mathrm{Cn}_{\mathbf{P2}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{P2.2^r}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{P2.2^m}}(\Gamma)$.*

7 In Conclusion

In this paper, we presented the logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$. Both are adaptive logics based on Goble's $\mathbf{SDL_aP_e}$. The difference between the two logics is that, in the case of two or more incompatible obligations $O_eA_1, \ldots, O_eA_n$, where the incompatibility is expressed by $\neg O_a(A_1, \ldots, A_n)$, $\mathbf{P2.2}^r$ does not allow one to derive the universal obligation $O_a(A_1 \vee \ldots \vee A_n)$, whereas $\mathbf{P2.2}^m$ does. We

[14] We refer to [2] for an overview of the meta-theoretic properties and the proofs that hold for all adaptive logics in standard format.

have argued that both outcomes are sensible for different kinds of application contexts.

The logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ are not the first adaptive logics based on Goble's $\mathbf{SDL_aP_e}$. The first one was the logic $\mathbf{P2.1}^r$ as presented in [12]. We have shown, however, that $\mathbf{P2.1}^r$ breaks down in the case of disjunctive obligations that are incompatible. The logics presented here solve that problem. They are also more elegant and easier to handle than $\mathbf{P2.1}^r$, because only one set of abnormalities is required.

As compared to other conflict-tolerant deontic logics, the logics $\mathbf{P2.2}^r$ and $\mathbf{P2.2}^m$ have several strengths. They preserve all nice properties of non-adjunctive deontic logics (for instance, that $O(A \wedge \neg A)$ is never derivable), but are much stronger (and less sensitive to the formulation of the premises) than any other system we know. Also, they do not presuppose that one knows in advance which obligations behave abnormally.

An important open problem concerns the design of logics in the same family that can handle cases where not all premises are equally preferred. Another open problem concerns the design of an adaptive version of the unimodal logic **P**. Preliminary results for both problems are coming forward.

References

[1] Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.

[2] Diderik Batens. *Adaptive Logics and Dynamic Proofs. Mastering the Dynamics of Reasoning, with Special Attention to Handling Inconsistency.* Forthcoming.

[3] David Brink. Moral conflict and its structure. *Philosophical Review*, 103:215–247, 1994.

[4] Brian F. Chellas. *Modal Logic: An Introduction.* Cambridge University Press, 1980.

[5] Alan Donagan. Consistency in rationalist moral systems. *Journal of Philosophy*, 81:291–309, 1984.

[6] Lou Goble. Multiplex semantics for deontic logic. *Nordic Journal of Philosophical Logic*, 5:113–134, 2000.

[7] Lou Goble. Preference semantics for deontic logic. Part I: Simple models. *Logique et Analyse*, 183–184:383–418, 2003.

[8] Lou Goble. Preference semantics for deontic logic. Part II: Multiplex models. *Logique et Analyse*, 185–188:335–363, 2004.

[9] Lou Goble. A logic for deontic dilemmas. *Journal of Applied Logic*, 3:461–483, 2005.

[10] John F. Horty. Moral dilemmas and nonmonotonic logic. *Journal of Philosophical Logic*, 23:35–65, 1994.

[11] John F. Horty. Reasoning with moral conflicts. *Nous*, 37:557–605, 2003.

[12] 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.

[13] Peter K. Schotch and Raymond E. Jennings. Non-kripkean deontic logic. In Risto Hilpinen, editor, *New Studies in Deontic logic*, pages 149–162. Reidel, Dordrecht, 1981.

[14] Bas Van Fraassen. Values and the heart's command. *Journal of Philosophy*, 70:5–19, 1973.