Avoiding Deontic Explosion by Contextually Restricting Aggregation

Joke Meheus, Mathieu Beirlaen, and Frederik Van De Putte

Centre for Logic and Philosophy of Science
University of Ghent, Belgium
{Joke.Meheus,Mathieu.Beirlaen,frvdeput.VanDePutte}@UGent.be

Avoiding Deontic Explosion by Contextually Restricting Aggregation[⋆]

Joke Meheus, Mathieu Beirlaen, and Frederik Van De Putte

Ghent University, Centre for Logic and Philosophy of Science, Blandijnberg 2,
B-9000 Ghent, Belgium
{Joke.Meheus,Mathieu.Beirlaen,frvdeput.vandeputte}@UGent.be

Abstract. In this paper, we present an adaptive logic for deontic conflicts, called $\mathbf{P2.1}^r$, that is based on Goble's logic $\mathbf{SDL}a\mathbf{P}e$—a bimodal extension of Goble's logic **P** that invalidates aggregation for all *prima facie* obligations. The logic $\mathbf{P2.1}^r$ has several advantages with respect to $\mathbf{SDL}a\mathbf{P}e$. For consistent sets of obligations it yields the same results as Standard Deontic Logic and for inconsistent sets of obligations, it validates aggregation "as much as possible". It thus leads to a richer consequence set than $\mathbf{SDL}a\mathbf{P}e$. The logic $\mathbf{P2.1}^r$ avoids Goble's criticisms against other non-adjunctive systems of deontic logic. Moreover, it can handle all the 'toy examples' from the literature as well as more complex ones.

Key words: conflict-tolerant deontic logic, non-adjunctive deontic logic, deontic explosion, defeasible deontic reasoning, adaptive logic

1 Introduction

Over the last two decades a plethora of deontic logics have been proposed for which deontic explosion (to derive OB from OA and $O\neg A$) is not valid. A large number of these systems are obtained by rejecting or restricting the aggregation principle—from OA and OB to derive $O(A \wedge B)$.[1]

Given the way in which humans build up their norms, it seems realistic to suppose that they adhere to norms deriving from 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 normative systems are in conflict, should not lead to the conclusion that there is a normative system that forces one to do the impossible. Thus, the intuitive principle "Ought implies Can" can be preserved.

In [6, p. 466], Lou Goble stated that giving up aggregation is "perhaps the most natural suggestion for avoiding deontic explosion". In several papers Goble

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

[1] See, for instance, [1], [2], [3], [4], [5].

advocated the use of one particular such logic, namely the logic **P** [2], [3], [4]. **P** is a very well-behaved system and has a natural interpretation in a Kripke-like semantics.[2] It has moreover a nice axiomatization and avoids any kind of explosion when applied to conflicting obligations.[3]

Still, the logic **P** has a serious drawback: it is too weak, especially when applied to obligations that are mutually compatible. For instance, in the famous Horty example [7, 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 is not permitted to fight in the army, or what comes to the same, 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) does not enable one to do so. In other words, simply invalidating aggregation results in a logic that is too weak.

In [2], Goble extends his logic **P** to the bimodal logic **SDL**a**P**e. 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. (The duals P_e and P_a are defined in the usual way.) 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* standard. 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 conflicted by other obligations.

The logic **SDL**a**P**e behaves exactly like Standard Deontic Logic (**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 consistent '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. 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, formalizing the premises in the wrong way (because some conflicts were not detected), may still lead to explosion. One could, of course, play it safe and formalize all obligations with the O_e-operator, but then the inferential power reduces to that of **P** and we are back to square one.

In this paper, we shall present a logic that leads to exactly the same consequence set as **SDL** for sets of premises that are conflict-free. For sets of premises

[2] The idea behind this semantics will be spelled out in Section 4. Goble also proposed a preferential semantics for **P** in [3], and a neighbourhood semantics in [2].

[3] As we shall shortly see, other kinds of deontic conflicts, such as those between obligations and permissions, *do* lead to explosion in **P**.

that are not conflict-free (but that have **P**-models), the set of consequences is "as rich as possible", without any form of deontic explosion being validated.

One of the basic ideas behind $\mathbf{P2.1}^r$ is that 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 stops 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.

The logic $\mathbf{P2.1}^r$ is an adaptive logic[4] and is based on Gobles $\mathbf{SDL}a\mathbf{P}e$. The logic $\mathbf{P2.1}^r$ is non-monotonic and its proof theory is dynamical (conclusions derived at some stage of a proof may be rejected at a later stage),[5] but is sound and complete with respect to a (static) semantics. We shall argue that Goble's objections from [6] against non-adjunctive approaches do not apply to $\mathbf{P2.1}^r$ and show that $\mathbf{P2.1}^r$ leads to the desired results for the Smith example and for more complex ones than that.

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 $A \vee B, A \wedge B, A \supset B, A \equiv B \in \mathcal{W}^M$

[4] See [8] for an introduction to adaptive logics and [9] for an overview of their metatheoretic properties.

[5] The stage of a proof refers to the set of lines that occur in the proof up to a certain line. Thus, "stage 14 of the proof" refers to the first fourteen lines in the proof.

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).

For reasons of readability, and because we prefer to name the adaptive logic based on **SDL***a***P***e* in such a way that the connection is clear, we shall from now on use the name **P2** to refer to **SDL***a***P***e*. The 1 in $\mathbf{P2.1}^r$ indicates that $\mathbf{P2.1}^r$ is only one member of a larger family of logics based on **P2**, and the "r" refers to the adaptive strategy.[6]

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 A More General Approach to Normative Conflicts

There is a strong tendency in the literature to restrict deontic conflicts to moral conflicts or moral dilemmas. In [10], [11] and [12], for instance, systems of deontic logic are presented that focus exclusively on *moral* conflicts. In this paper, we shall not require that the normative statements are moral (they can come from legal codes, moral codes, traffic regulations, promises, ...).

Most authors moreover focus almost exclusively on cases where, for some A, both OA and $O\neg A$ hold. However, a non-explosive deontic logic should be able to account for other types of conflicts, such as, for instance, conflicts between an obligation and a permission. Suppose, for instance, that Yilmaz ought not to drink alcohol according to his religious beliefs. However, according to the laws of his country, he is permitted to drink alcohol. If one would simply formalize this conflict in the language of **P** as $O\neg A \wedge PA$ and apply **P** to it, explosion would follow, just as in **SDL**. However, a more natural formalization can be obtained in the language of **P2**. What the first premise comes to is that Yilmaz has a *prima facie* obligation not to drink alcohol ($O_e\neg A$). The second premise entails that it is not an actual obligation not to drink alcohol: there are normative systems that allow one to drink alcohol. This is why we would formalize the second premise as $\neg O_a\neg A$, or what comes to the same, as P_aA. Applying **P2** or $\mathbf{P2.1}^r$ to $O_e\neg A \wedge P_aA$ does not result in explosion.

We shall also not presuppose that all conflicts between obligations are 'direct conflicts' (that is, conflicts of the form $OA \wedge O\neg A$). One should, for instance, also allow for deontic conflicts that involve more than two obligations, such as OA, OB, and $O(\neg A \vee \neg B)$ (see below for a discussion of this kind of conflicts). Remarkably, hardly any attention was paid to such conflicts (some exceptions

[6] The adaptive strategy used for $\mathbf{P2.1}^r$, is the Reliability Strategy. What an adaptive strategy comes to will become clear below.

where the possibility of conflicts between more than two obligations is considered are [13], [14], and [15]).

Furthermore, we shall not presuppose that all normative conflicts can be *reduced* to direct conflicts. 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 the following assumptions (see [6, p. 462]):

RM If $\vdash A \supset B$, then $\vdash OA \supset OB$
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.[7] 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 our view, this kind of reduction is not the most natural formalization of the situation and causes the loss of crucial information. One loses, for instance, the information that there is a *link* between saving or not saving the first twin and saving or not saving the second one. One also loses the information that there are no possible worlds in which both twins are saved. This is why, where T_1, respectively T_2, stands for saving the first twin, respectively the second twin, we would formalize the twin example as $\Gamma_1 = \{O_e T_1,\ O_e T_2,\ O_a \neg(T_1 \wedge T_2)\}$. The all-things-considered obligation in Γ_1 is intended to capture the idea that it is impossible to save both twins (that is, that there is no accessible world in which both are saved).

Admittedly, the two direct conflicts that Goble starts from are derivable by $\mathbf{P2.1}^r$ from Γ_1, but our formalization is stronger (one cannot derive the members of Γ_1 from Goble's formalization), and it retains the information that is otherwise lost.

In our view, there are also examples of normative conflicts that cannot be reduced to direct conflicts. As an example 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 to quarrel over almost anything and Bob's girlfriend is afraid that this may put a damper on the party). In this case, Bob is facing three *prima facie* obligations

[7] 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.

(1) he has a *prima facie* obligation to invite John — $O_e I_j$
(2) he has a *prima facie* obligation to invite Peter — $O_e I_p$
(3) he has a *prima facie* obligation not to invite both Peter and John — $O_e \neg(I_j \wedge I_p)$

Note that (1) and (2) are jointly compatible and jointly possible (analogously for (1) and (3) and for (2) and (3)). Nothing prevents Bob from inviting both his friends, except for the promise that he made to his girlfriend. As (1) and (2) are jointly possible, no direct conflict of the form $OA \wedge O\neg A$ is derivable by $\mathbf{P2.1}^r$ from (1)–(3), even if one adheres to NM.

Some readers might argue that there *is* an incompatibility: it is impossible to obey to all three obligations at the same time: $\neg\Diamond((I_j \wedge I_p) \wedge \neg(I_j \wedge I_p))$. However, in order to derive the direct conflict $O_e(I_j \wedge I_p) \wedge O_e\neg(I_j \wedge I_p)$ (in view of NM), one first needs $O_e(I_j \wedge I_p)$, and this is not derivable by **P2** from (1) and (2). Critical readers might continue that we can easily reformulate our premises in such a way that a direct conflict becomes derivable (for instance, by replacing (1) and (2) by $O_e(I_j \wedge I_p)$. However, this is exactly the kind of move that (throughout the paper), we want to avoid. We want to take sets of premises at face value and let the context and the logic decide what follows from what, without interferences (or 'preparations') from the part of the reasoner.

When discussing the formalization of normative conflicts, one may also think of Horty's *visiting parents* example [16, p. 581]. Suppose that you have an obligation to visit both your own and your spouse's parents during the holiday season, and that, because they live in separate parts of the country, it is not possible to visit both pairs. Whichever pair you eventually decide on, you should notify them of your visit. Let V_1 be that you visit your own parents, N_1 that you notify your own parents of your visit, and let N_2 and V_2 be the respective propositions for your spouse's parents. Horty represents these obligations as $\Gamma_2 = \{O_e(V_1 \wedge N_1),\ O_e(V_2 \wedge N_2),\ O_e\neg(V_1 \wedge V_2)\}$.[8]

Horty uses this example to make a case against consistent aggregation in favour of consistent consequent aggregation. According to the former, but not the latter $O_e(N_1 \wedge N_2)$ is derivable from the premises. According to Horty the derivability of $O_e(N_1 \wedge N_2)$ is incorrect: although it is possible to notify both pairs of parents that you are planning to visit them, this is not what you *ought* to do in this situation.

As will become clear below, $O_a(N_1 \wedge N_2)$ is derivable by $\mathbf{P2.1}^r$ from Γ_2. Is this a problem for our logic? We believe it is not. In our view, there is nothing wrong with the outcome, it is the formalization that is mistaken (for this particular example). This requires some explanation.

Horty uses a *truth-functional conjunction* to formalize the first two premises, even though it is clear that there is a *connection* between visiting one specific pair of parents and notifying *that* pair of parents of your visit. Even Horty, in his intuitive reading, seems to acknowledge that there is such a connection. In his words:

[8] Horty takes these to be your *prima facie* obligations, hence we have formalized them with the O_e-operator.

> Your prima facie obligations can then be represented through the two imperatives [. . .] telling us that you should notify and *then* visit your parents, but also that you should notify and *then* visit your spouse's parents. [16, p. 581] (our italics)

So, Horty clearly sees at least a *temporal* connection between the two conjuncts, but this is not captured by the truth-functional conjunction. (To see that Horty has something more in mind than is captured by his formalization, note that the sentence "you should notify your parents and then visit them" has a different meaning than the sentence "you should visit your parents and then notify them").

But there is something worse, there is also a *conditional* connection between the two conjuncts: the obligation to notify someone holds only *in view of* the obligation to visit this person. Also this is lost in Horty's formalization. A formalization that retains this information is $\{O_eV_1,\ O_eV_2,\ O_e(V_1 \supset N_1),\ O_e(V_2 \supset N_2),\ O_e\neg(V_1 \wedge V_2)\}$. or, if you prefer, $\{O_eV_1,\ O_eV_2,\ V_1 \supset O_1N_1,\ V_2 \supset O_eN_2,\ O_e\neg(V_1 \wedge V_2)\}$.[9] If Horty's example is formalized in either of these ways, one no longer obtains the unwanted consequence that one has an actual obligation to notify both pairs of parents.

There is another way to plead for our case. Horty's way of formalizing is adequate only for actions that are independent from one another. Suppose, for instance, that, on the one hand, you have the obligation to attend your daughter's wedding and to buy her a suitable wedding gift and, on the other hand, you have the obligation to attend your son's wedding and to buy him a suitable wedding gift. As both your children decided to marry on the same day in different parts of the country you cannot attend both weddings. Still, you can buy a wedding gift for both of them, and this seems the proper thing to do. So, where A_d, respectively A_s, stands for attending your daughters wedding, respectively your son's wedding, and G_d, respectively G_s, stands for buying a wedding gift for your daughter, respectively for your son, there is nothing wrong with the Horty-like formalization $\{O_e(A_d \wedge G_d),\ O_e(A_s \wedge G_s),\ O_e\neg(A_d \wedge A_s)\}$, and there is also nothing wrong with the fact that the actual obligation $O_a(G_d \wedge G_s)$ is $\mathbf{P2.1}^r$-derivable in this case.

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.

[9] A similar formalization is found in the literature on the Chisholm paradox, where sentences of the form "It ought to be that if X visits Y then X tells Y (s)he is coming" are typically formalized as either $O(V_y \supset N_y)$ or $V_y \supset ON_y$ (see, for instance, [13]).

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, $v : \mathcal{S} \times W \to \{0, 1\}$ 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 iff $v_M(A, w_0) = 1$, $\models_{\mathbf{P2}} A$ iff all **P2**-models verify A, and $\Gamma \models_{\mathbf{P2}} A$ 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$_a e$ $O_a(A \supset B) \supset (O_e A \supset O_e B)$

The first three postulates deliver **SDL** for O_a and the next three deliver **P** for O_e.[10] The last axiom links the two operators.

As was mentioned in the introduction, **P2** has several nice properties. But the fact remains that it is a very poor logic, and that its inferential strength can only be increased by making the correct 'guesses' on what the unproblematic premises are. At some point, Goble no longer considered **P2** 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.

5 Restricting Aggregation

A detailed discussion of the main attempts to restrict aggregation can be found in [6, pp. 467-473]. Here, we shall only briefly mention some of them.

The first one is that of *consistent aggregation* or

CAND If $\nvdash A \supset \neg B$ then $\vdash (OA \wedge OB) \supset O(A \wedge B)$

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

Although this suggestion appears natural, it is much too strong. In the presence of a normative conflict $OA \wedge O\neg A$ and some random formula B such that $\nvdash \neg B$, CAND allows one to derive $O(A \wedge (\neg A \vee B))$, from which follows OB. Hence CAND validates the following contra-intuitive and explosion-like principle:

DEX-1 If $\nvdash \neg B$ then $\vdash (OA \wedge O\neg A) \supset OB$

The second form of restricted aggregation is that of *permitted aggregation* or

PAND $\vdash P(A \wedge B) \supset ((OA \wedge OB) \supset O(A \wedge B))$

Instead of allowing aggregation for obligations that are *jointly compatible* (as is the case for CAND), this alternative allows aggregation for obligations that are *jointly permissible*. Unfortunately, PAND suffers from similar problems as CAND. Whereas in the case of CAND a variant of deontic explosion follows from a normative conflict $OA \wedge O\neg A$ in the presence of some formula $\neg B$ that is logically contingent, in the case of PAND a variant of explosion follows from $OA \wedge O\neg A$ in the presence of some formula B such that B is permitted. To see why, note that $PB \equiv P((A \vee B) \wedge (\neg A \vee B))$, that from OA follows $O(A \vee B)$ and that from $O\neg A$ follows $O(\neg A \vee B)$. Hence, by PAND, we obtain $O((A \vee B) \wedge (\neg A \vee B))$ (which is equivalent to OB) from $OA \wedge O\neg A$. This yields

DEX-2 $\vdash (OA \wedge O\neg A) \supset (PB \supset OB)$

More generally, Goble argues that any restricted aggregation rule of the form:

RAND If $Cond(A \wedge B)$ then $\vdash (OA \wedge OB) \supset O(A \wedge B)$

where $Cond(A \wedge B)$ denotes a certain restriction over A and B, is bound to lead to a related type of deontic explosion:

DEX-gen $\vdash Cond(B) \supset ((OA \wedge O\neg A) \supset OB)$

That is, all propositions that satisfy *Cond* will become obligatory as soon as a deontic conflict arises. Hence we should look for some way to restrict aggregation that is not based solely on the behaviour of the two obligations that one wants to aggregate, but on the behaviour of the whole set of obligations *and* their consequences. Our proposal meets this requirement, and thus overcomes the problems faced by restricted aggregation.

Apart from the fact that constrained aggregation leads to some form of explosion, we can also provide a more philosophical argument against these systems. It seems strange to assume that the reasoner should add formulas of the form $Cond(A \wedge B)$, in order to obtain $O(A \wedge B)$ from OA and OB. If we know beforehand which obligations may be aggregated safely, there is no genuine problem. An important feature of the logic $\mathbf{P2.1}^r$ is that it does not presuppose such knowledge: it is the logic *itself* that localizes which obligations can be aggregated, without the need to add any new premise.

Goble discusses two more classes of solutions to the aggregation-problem [6, pp. 469-471]. The first is "constrained consistent aggregation", where the aggregation is restricted to consistent subsets of the premise set Γ (see [16] for an update of this proposal). A major drawback of this system is that the set of derivable obligations depends largely on the way the premises are formalized: equivalent premise sets can yield different results. Moreover, this approach cannot handle more complex situations like the Johnson example, that we shall present in Section 7.

The second class of solutions consists of bimodal systems, where aggregation is restricted to obligations for one of the two ought-operators—see, for instance, [17] and [15]. Goble argues convincingly that, although these logics avoid explosion, they seem far-fetched from the viewpoint of everyday deontic reasoning.

6 Desiderata for a Conflict-Tolerant Deontic Logic

The discussion in the previous sections gives us several requirements for an adequate logic for deontic conflicts. Evidently, no form of deontic explosion should be validated. For sets of premises that are conflict-free, the logic should lead to the same results as **SDL** and for sets of premises that contain conflicts it should be "as rich as possible". The logic should also not presuppose that one knows in advance which obligations behave consistently and it should not be sensitive to the accidental formulation of the premises. Finally, it should be able to handle other normative conflicts than $OA \wedge O\neg A$.

The logic that we shall present in the next two sections satisfies all these requirements.

7 Intuitive Characterization of $\mathbf{P2.1}^r$

The logic $\mathbf{P2.1}^r$ is an adaptive extension of the logic **P2**. The logic **P2** constitutes the stable part of $\mathbf{P2.1}^r$: anything that is **P2**-derivable from a premise set is unconditionally derivable in $\mathbf{P2.1}^r$. In addition to this, it is allowed that O_e-obligations are interpreted "as much as possible" as O_a-obligations. 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 "as much as possible". But let us first illustrate the main ideas by means of an example.

Suppose that Johnson faces the following three 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 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 $\mathbf{P2.1}^r$-proof by entering 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 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 now 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 clearly a conflict in the second case (see lines 6 and 7), but not in the first case. We shall see below that $\mathbf{P2.1}^r$ 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 $\mathbf{P2.1}^r$, the strategy is Reliability.[11] To explain this strategy, we first need some definitions.

Where Δ is a finite set of abnormalities, the disjunction $\bigvee(\Delta)$ will be called a *Dab*-formula and will be written as $Dab(\Delta)$. 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, no $Dab(\Delta')$ is derived, such that $\Delta' \subset \Delta$.

What the Reliability Strategy comes to is that, whenever a *minimal Dab*-formula is unconditionally derived in the proof at a certain stage, then all disjuncts that occur in that *Dab*-formula are considered as behaving abnormally (or as 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. A condition is violated at a certain stage if at that stage its

[11] The two most common strategies in adaptive logics are the Reliability Strategy and the Minimal Abnormality Strategy—the former is a bit more cautious than the latter—see [9].

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 restriction on the form of the abnormalities. Without such restriction, we would obtain a so-called flip-flop logic: a logic that behaves exactly like **SDL** for consistent sets of premises, but like **P** for inconsistent sets of premises. The reason for this is easily demonstrated by means of the following example. Consider $\Gamma_3 = \{O_e p,\ O_e\neg p,\ O_e q\}$. As there is clearly no conflict with respect to $O_e q$, $O_a q$ should be $\mathbf{P2.1}^r$-derivable from Γ_3. However, the disjunction $(O_e q \wedge \neg O_a q) \vee (O_e(\neg p \vee \neg q) \wedge \neg O_a(\neg p \vee \neg q))$ is **P2**-derivable from Γ_3, whereas neither of its disjuncts is. Hence, in view of the Reliability Strategy, $O_e q \wedge \neg O_a q$ would be considered as unreliable and this would block the desired inference from $O_e q$ to $O_a q$. This is why we shall only consider those formulas of the form $O_e A \wedge \neg O_a A$ as abnormalities where A is an atom. This brings us to the third and last complication.

If we would restrict the abnormalities to the set $\{O_e A \wedge \neg O_a A \mid A \in \mathcal{W}^a\}$, 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 $O_e(A_1 \vee \ldots \vee A_n)$ (for $n \geq 2$) cannot be generalized to $O_a(A_1 \vee \ldots \vee A_n)$. A natural answer to this question is that it counts as 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, *unless* $O_e(A_1 \vee \ldots \vee A_n)$ is obtained from a 'shorter' obligation that behaves abnormally. Thus, in the case of Γ_3, 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 $O_a(q \vee r)$. This brings us to a second type of abnormalities. Let $\dagger A$ abbreviate $(O_e A \wedge \neg O_a A)$. Where $A_1, \ldots, A_n \in \mathcal{W}^a$, and $n \geq 2$, the general form of this second type of abnormalities is $O_e(A_1 \vee \ldots \vee A_n) \wedge \neg\dagger A_1 \wedge \ldots \wedge \neg\dagger A_n \wedge \neg O_a(A_1 \vee \ldots \vee A_n)$.

We can now return to our Johnson example. We shall use $\ddagger(A_1 \vee \ldots \vee A_n)$ to abbreviate $O_e(A_1 \vee \ldots \vee A_n) \wedge \neg\dagger A_1 \wedge \ldots \wedge \neg\dagger A_n \wedge \neg O_a(A_1 \vee \ldots \vee A_n)$. 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	$\{\dagger F,\ \dagger S,\ \ddagger(F \vee S)\}$
10	$O_a S$	8, 9; RU	$\{\dagger\neg F,\ \dagger F,\ \dagger S,\ \ddagger(F \vee S)\}$

Lines 8 and 9 are applications of the conditional rule. 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. One way to read this is: $O_a\neg F$ is derivable from $O_e\neg F$ or $O_e\neg F \wedge \neg O_a\neg F$ is true. This is the motor behind the proof theory: abnormalities are assumed to be false unless and until proven otherwise. 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 on that line is no longer considered as derived in the proof. An analogous reasoning holds for line 9. In this case, $(O_e(F \vee S) \supset O_a(F \vee S)) \vee ((O_e F \wedge \neg O_a F) \vee (O_e S \wedge \neg O_a S) \vee \ddagger(F \vee S))$

is **P2**-derivable from the premises, and also here, the abnormalities are assumed to be false unless and until proven otherwise.

Line 10 is an application of the unconditional rule. Note that when the unconditional rule is applied, no new formulas are added to the condition, but any formula that occurs in a non-empty condition is 'carried' over to the conclusion of the application. The reason for this is easy to understand. If, at some point, line 8 or line 9 has to be marked (because the 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 O_aT is not $\mathbf{P2.1}^r$-derivable. Analogous to line 8, the conditional rule allows one to add a line to the proof on which O_aT is derived on the appropriate condition:

11	O_aT	6; RC	$\{O_eT \wedge \neg O_aT\}$

At this stage of the proof, the formula O_aT is considered as derived. Things change, however, as soon as the following line is added:

12	$O_eT \wedge \neg O_aT$	6, 7; RU	$\emptyset$

This line makes it clear that the condition of line 11 is not fulfilled, and hence, line 11 is marked in view of line 12:

11	O_aT	6; RC	$\{O_eT \wedge \neg O_aT\}\checkmark^{12}$
12	$O_eT \wedge \neg O_aT$	6, 7; RU	$\emptyset$

From stage 12 on, O_aT is no longer considered to be derived in the proof. It is easy to check that line 11 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 from the premises 1–3. (The precise definition of *final derivability* follows in the next section.) Note especially that O_aS is $\mathbf{P2.1}^r$-derivable from the premises, even though it is 'connected' to a problematic obligation. Although approaches like Horty's constrained consistent aggregation can handle the Smith case, they can only deal with the Johnson example by reformulating the premises as $\{O_eT,\ O_e(F \vee S),\ O_e\neg T,\ O_e\neg F,\ O_e(T \vee C)\}$—see also below.

It was hinted at in the introduction that the proof theory of $\mathbf{P2.1}^r$ 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 because of new insights in the premises (for instance, when an abnormality that was originally not noticed is derived at a later stage). The external dynamics occurs when new premises are added. 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 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.

8	$O_a\neg F$	5; RC	$\{\dagger\neg F\}\checkmark^{14}$
9	$O_a(F \vee S)$	4; RC	$\{\dagger F, \dagger S, \ddagger(F \vee S)\}\checkmark^{15}$
10	O_aS	8, 9; RU	$\{\dagger\neg F, \dagger F, \dagger S, \ddagger(F \vee S)\}\checkmark^{14}$
11	O_aT	6; RC	$\{O_eT \wedge \neg O_aT\}\checkmark^{12}$
12	$O_eT \wedge \neg O_aT$	6, 7; RU	$\emptyset$
13	$O_e\neg S$	PREM	$\emptyset$
14	$\dagger\neg F \vee \dagger\neg S$	1, 2, 13, RU	$\emptyset$
15	$\dagger F \vee \dagger S \vee \ddagger(F \vee S)$	1, 2, 13, 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 13.

To the best of our knowledge, the logic $\mathbf{P2.1}^r$ is the first system that can handle the Johnson example in its *actual* form and without adding any allegedly 'hidden' or 'tacit' premises—for instance, that it is permitted not to fight *and* to fight of perform civil service). That we insist on the *actual* form of the Johnson example is no accident. There are approaches available that can handle the Johnson example, but only at the expense of reducing it to an instance of the Smith example (with some extra premises). Some readers may argue, for instance, that the Johnson example is better formalized as

O1' he ought to pay taxes — O_eT
O2' he ought to fight in the army or perform alternative service to his country — $O_e(F \vee S)$
O3' he ought not to pay taxes — $O_e\neg T$
O4' he ought not to fight in the army — $O_e\neg F$
O5' he ought to pay taxes or donate to charity — $O_e(T \vee C)$

Given this reformulation, there are indeed alternative approaches available that lead to the desired result (that OS is derivable but OC is not). This, however, misses the whole point. In *our* approach we do not need this reformulation to obtain the desired results. Any reformulation that is not logically equivalent to the original one may lead to the loss of crucial information (for instance, the person who formulated O1-O3 may have wanted to express that Johnson is dealing with norms from three different sources. This information is lost in the reformulation O1'-O5'). This is why we consider it important to start from the original premises and leave their analysis to the logic itself.

We also consider it important that, in our approach, it is not necessary to extend the original set of premises with 'hidden' assumptions. The reason is that, in complex cases, it may be far from evident which premises can be safely added and that making the wrong 'guesses' may lead to explosion after all.

To end this section, we address a different kind of objection. Some readers may wonder why we chose **P2** as our underlying logic, instead of the simpler **P**, and also why we not simply chose $OA \wedge O\neg A$ as the form of our abnormalities. The reason is simple: starting from **P** and taking $OA \wedge \neg A$ as the form of the abnormalities does not result in an adaptive logic that has the desired properties. For instance, it would not enable one to handle the Smith example. This is because $O\neg F$, $O(F \vee S) \nvdash_{\mathbf{P}} OS \vee (OF \wedge O\neg F)$. So, even if one assumes that $OF \wedge O\neg F$ is false *unless and until* proven otherwise, this will not yield OS as a conditional conclusion. Put more generally, the problem is that aggregation would not be validated contextually, in view of $\{OA \wedge OB\} \nvdash_{\mathbf{P}} O(A \wedge B) \vee (OA \wedge O\neg A) \vee (OB \wedge O\neg B)$. Also, combining **P** with all formulas of the form $(OA \wedge OB) \wedge \neg O(A \wedge B)$ as abnormalities will not do, since this leads to a flip-flop logic. The upshot is that we could not find any form for the abnormalities (expressible in the language of **P**) that would lead to the desired adaptive logic. We do not exclude, however, that by varying on **P** (for instance, by giving up interdefinability) one might obtain a system that is simpler than $\mathbf{P2.1}^r$. Still, the purpose of the present paper was to stay as close as possible to Goble's **P**-systems.

8 The Adaptive Logic $\mathbf{P2.1}^r$

In this section, we present the logic $\mathbf{P2.1}^r$ in a formally precise way. As any other adaptive logic in standard format, the logic $\mathbf{P2.1}^r$ is characterized by a triple: a lower limit logic (a reflexive, transitive, monotonic, uniform, and compact logic for which there is a positive test)[12], a set of abnormalities Ω (characterized by a, possibly restricted, logical form) and a strategy. The lower limit logic is the logic that determines the stable part of the adaptive logic, and that also determines the unconditional rule. In the case of $\mathbf{P2.1}^r$, the lower limit logic is **P2** and the strategy is Reliability.

As we have seen, the abnormalities in $\mathbf{P2.1}^r$ are characterized by two different forms of formulas. In this case, however, there is no problem to take the union Ω of the two separate sets Ω_1 and Ω_2, and to define both the semantics and the derivability relation with respect to this unified set. The set of abnormalities are defined as follows:

$$\Omega_1 = \{\dagger A \mid A \in \mathcal{W}^a\}$$
$$\Omega_2 = \{\ddagger(A_1 \vee \ldots \vee A_n) \mid A_1, \ldots, A_n \in \mathcal{W}^a; n \geq 2\}$$
$$\Omega = \Omega_1 \cup \Omega_2$$

In order to define the semantics, we need some further definitions. We first define the abnormal part of a **P2**-model:

[12] A property for objects of a given kind is *decidable* iff there is a mechanical procedure that leads to the answer YES if the property holds and to the answer NO if the property does not hold. 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.

Definition 1. $Ab(M) = \{A \in \Omega \mid M \Vdash A\}$

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 2. *Where $Dab(\Delta_1)$, $Dab(\Delta_2)$, $\ldots$ 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.*

In view of these definitions, the semantic consequence relation of $\mathbf{P2.1}^r$ is given by:

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

Definition 4. *$\Gamma \vDash_{\mathbf{P2.1}^r} A$ iff A is verified by all reliable models of Γ.*

As is common for all adaptive logics in standard format, the proof theory of $\mathbf{P2.1}^r$ 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 line 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 definition proceeds in terms of the *minimal Dab-formulas* derived at a stage of the proof:

Definition 5. *$Dab(\Delta)$ is a* minimal *Dab*-formula *at stage s iff, at stage s, $Dab(\Delta)$ is derived on condition $\emptyset$, and no $Dab(\Delta')$ with $\Delta' \subset \Delta$ is derived on condition $\emptyset$.*

Definition 6. *Where $Dab(\Delta_1)$, ..., $Dab(\Delta_n)$ are the minimal Dab-formulas derived on condition $\emptyset$ at stage s, $U_s(\Gamma) = \Delta_1 \cup \ldots \cup \Delta_n$.*

Definition 7. *Where Δ is the condition of line i, line i is marked at stage s iff $\Delta \cap U_s(\Gamma) \neq \emptyset$.*

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 8. *A is* finally derived *from Γ at line i of a proof at stage s iff A is derived at a 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.*

As may be expected, the derivability relation of $\mathbf{P2.1}^r$ is defined with respect to the notion of final derivability

Definition 9. *$\Gamma \vdash_{\mathbf{P2.1}^r} A$ (A is finally derivable from Γ) iff A is finally derived in an $\mathbf{P2.1}^r$-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 [9] for the proofs. The soundness and completeness of **P2** therefore yield:

Theorem 1. *$\Gamma \vdash_{\mathbf{P2.1}^r} A$ iff $\Gamma \vDash_{\mathbf{P2.1}^r} A$.*

The fact that $\mathbf{P2.1}^r$ is in standard format moreover warrants that it has a number of other meta-theoretic properties, such as proof invariance:[13]

Theorem 2. *If $\Gamma \vdash_{\mathbf{P2.1}^r} A$, then every $\mathbf{P2.1}^r$-proof from Γ can be extended in such a way that A is finally derived in it.*

9 In Conclusion

In this paper, we presented the logic $\mathbf{P2.1}^r$, which is only one logic from a family of logics that are based on Goble's **SDL***a***P***e*. A simple extension of $\mathbf{P2.1}^r$ ensures that in the case of incompatible obligations (formalized, for instance, as O_eA, O_eB, $O_e\neg(A \wedge B)$), the general obligation to do A or B ($O_a(A \vee B)$) is derivable. Other logics in the same family can handle cases where not all normative statements are equally preferred.

As compared to other conflict-tolerant deontic logics, the logic $\mathbf{P2.1}^r$ has several strengths (see also Section 6). It preserves all nice properties of non-adjunctive deontic logics (for instance, that $O(A \wedge \neg A)$ is never derivable), but

[13] We refer to [8] for an overview of the meta-theoretic properties and the proofs that hold for all adaptive logics in standard format.

is much stronger (and less sensitive to the formulation of the premises) than any other system we know. Still, none of the forms of deontic explosion that were discussed in Section 5 are validated. Evidently, this does not mean that $\mathbf{P2.1}^r$ is free from any kind of explosion. For instance, $\mathbf{P2.1}^r$ cannot handle plain contradictions, such $O_eA \wedge \neg O_eA$. In order to handle that kind of conflicts, one needs a system for which the negation is paraconsistent outside the scope of a modal operator.

Another result from the present paper is the broadening of the notion of a normative conflict. Our logic not only enables one to deal with other kinds of conflicts than the one that is usually studied, but also enables one to discern links between conflicting statements. Especially the latter has been completely ignored up to now.

References

1. Van Fraassen, B.: Values and the heart's command. Journal of Philosophy **70** (1973) 5–19
2. Goble, L.: Multiplex semantics for deontic logic. Nordic Journal of Philosophical Logic **5** (2000) 113–134
3. Goble, L.: Preference semantics for deontic logic. Part I: Simple models. Logique et Analyse **183–184** (2003) 383–418
4. Goble, L.: Preference semantics for deontic logic. Part II: Multiplex models. Logique et Analyse **185–188** (2004) 335–363
5. Schotch, P.K., Jennings, R.E.: Non-kripkean deontic logic. In Hilpinen, R., ed.: New Studies in Deontic logic. Reidel, Dordrecht (1981) 149–162
6. Goble, L.: A logic for deontic dilemmas. Journal of Applied Logic **3** (2005) 461–483
7. Horty, J.F.: Moral dilemmas and nonmonotonic logic. Journal of Philosophical Logic **23** (1994) 35–65
8. Batens, D.: Adaptive Logics and Dynamic Proofs. Mastering the Dynamics of Reasoning, with Special Attention to Handling Inconsistency. (Forthcoming)
9. Batens, D.: A universal logic approach to adaptive logics. Logica Universalis **1** (2007) 221–242
10. da Costa, N.C., Carnielli, W.: On paraconsistent deontic logic. Philosophia **16** (1986) 293–305
11. Holbo, J.: Moral dilemmas and the logic of obligation. American Philosophical Quarterly **39** (2002) 259–274
12. Routley, R., Plumwood, V.: Moral dilemmas and the logic of deontic notions. In Priest, G., Routley, R., Norman, J., eds.: Paraconsistent Logic. Essays on the Inconsistent. Philosophia Verlag, München (1989) 653–702
13. McConnell, T.: Moral dilemmas. Published online at http://plato.stanford.edu/entries/moral-dilemmas/ (2006)
14. Conee, E.: Against moral dilemmas. The Philosophical Review **91** (1982) 87–97
15. Hansen, J.: Problems and results for logics about imperatives. Journal of Applied Logic **2** (2004) 39–61
16. Horty, J.F.: Reasoning with moral conflicts. Nous **37** (2003) 557–605
17. Van der Torre, Leendert Tan, Y.H.: Two-phase deontic logic. Logique et Analyse **43** (2000) 411–456