

EUR Research Information Portal

Adaptive deontic logics

Published in:

IfCoLoG Journal of Logics and their Applications

Publication status and date:

Published: 01/05/2019

Document Version

Publisher's PDF, also known as Version of record

Document License/Available under:

Article 25fa Dutch Copyright Act

Citation for the published version (APA):

Van De Putte, F., Beirlaen, M., & Meheus, J. (2019). Adaptive deontic logics: A survey. *IfCoLoG Journal of Logics and their Applications*, 6(3), 523-608.

[Link to publication on the EUR Research Information Portal](#)

Terms and Conditions of Use

Except as permitted by the applicable copyright law, you may not reproduce or make this material available to any third party without the prior written permission from the copyright holder(s). Copyright law allows the following uses of this material without prior permission:

- you may download, save and print a copy of this material for your personal use only;
- you may share the EUR portal link to this material.

In case the material is published with an open access license (e.g. a Creative Commons (CC) license), other uses may be allowed. Please check the terms and conditions of the specific license.

Take-down policy

If you believe that this material infringes your copyright and/or any other intellectual property rights, you may request its removal by contacting us at the following email address: openaccess.library@eur.nl. Please provide us with all the relevant information, including the reasons why you believe any of your rights have been infringed. In case of a legitimate complaint, we will make the material inaccessible and/or remove it from the website.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/325755997>

Adaptive Deontic Logics: a Survey

Preprint · June 2018

CITATIONS

0

READS

157

1 author:



Frederik Van De Putte

University of Bayreuth and Ghent University

38 PUBLICATIONS 156 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project

What should we expect? Actions, assumptions, and norms [View project](#)

Project

DYCODE: The Dynamics of Constructive Deliberation [View project](#)

ADAPTIVE DEONTIC LOGICS: A SURVEY

FREDERIK VAN DE PUTTE
Ghent University
frederik.vandeputte@ugent.be

MATHIEU BEIRLAEN
Ghent University
mathieubeirlaen@gmail.com

JOKE MEHEUS
Ghent University
joke.meheus@ugent.be

Abstract

Adaptive Logics (ALs) are a viable and useful formal tool to handle various issues in deontic logic. In this paper, we motivate, explain, illustrate, and discuss the use of ALs in deontic logic. Published work on deontic ALs focusses mainly on conflict-tolerant deontic logics (logics that can accommodate conflicting obligations) and – to a lesser extent – on problems concerning factual and deontic detachment. So does the present paper. Near the end of the paper, however, we also indicate some of the possibilities that the adaptive logic framework creates for tackling other types of problems within deontic logic.

Keywords: Deontic Logic, Adaptive Logics, Conflict-tolerance, Non-monotonic Reasoning, Benchmark Examples

Preludium: Nathan’s predicament

One Friday evening, Nathan promises his mother that he will look after his little brother, Ben, on Saturday afternoon so that she can visit her sister. A couple of hours later, Nathan’s girlfriend Lisa calls. Being a typical teenager and hopelessly in love, he completely forgets about the promise he made earlier to his mother and

We thank Jesse Heyninck, Stef Frijters, for valuable comments on previous drafts of the paper. We are also indebted to Lou Goble for his incisive remarks and useful suggestions.

agrees with Lisa to go with her to the cinema on Saturday afternoon (to see this cool movie – children under the age of 13 not allowed!) and to go for a veggie burger in the evening. On Saturday, Lisa rings at the door. Almost simultaneously, his mother puts on her coat, meanwhile saying “So, I’ll be back by five. Don’t forget...”. Hearing this, Nathan remembers about *both* promises and immediately realizes what kind of situation he is in. Given his promises, there are several things he ought to do and it is clear that he cannot do them all. Keeping his promise to go for a veggie burger in the evening still seems feasible, but he cannot look after six year old Ben and at the same time take Lisa to this particular movie!

1 Introduction

Logical principles may fail to apply under certain conditions, and logical principles involving normative concepts are no exception. Even if we restrict our focus to the modalities “it is obligatory that” and “it is permitted that”, there are circumstances in which we cannot apply certain plausible rules of inference (unrestrictedly) on pain of highly undesirable outcomes or even plain triviality.

The example from the *preludium* provides one kind of illustration of this phenomenon. It concerns a context in which an agent, in this case Nathan, faces several obligations that cannot be jointly fulfilled. In such contexts, several clusters of otherwise plausible principles involving obligations and permissions are problematic. Let us look at two instances of such clusters.

Consider first the combination of the principle that whatever is obligatory is also permissible (OIP), and the principle of the interdefinability of obligation and permission (ID):

(OIP) If A is obligatory, then A is also permitted: $OA \supset PA$

(ID) A is obligatory iff $\neg A$ is not permitted: $OA \equiv \neg P\neg A$

If both A and its negation $\neg A$ are obligatory ($OA \wedge O\neg A$), then by (ID) and the first conjunct, $\neg P\neg A$. However, by (OIP) and the second conjunct, $P\neg A$. So we obtain a plain contradiction: $\neg P\neg A \wedge P\neg A$. Even if one is willing to accept that contradictions are not absurd, it seems hard to accept that conflicting obligations entail them. Opinions may differ on which of these two principles is the most salient one. It is clear, however, that at least one of them has to be abandoned or adequately restricted if we want to avoid the outcome that conflicting obligations entail plain contradictions.

A second cluster of principles which is problematic in the face of conflicting obligations consists of the aggregation principle (Agg), the principle that “ought

implies can” (OIC), and the impossibility of contradictory states of affairs (CP):

- (Agg) If A and B are obligatory, then so is their conjunction: $(OA \wedge OB) \supset O(A \wedge B)$
- (OIC) If something is obligatory, then it is also possible: $OA \supset \diamond A$
- (CP) Contradictions are impossible: $\neg \diamond(A \wedge \neg A)$

If $OA \wedge O\neg A$, then, by (Agg), $O(A \wedge \neg A)$ and hence by (OIC), $\diamond(A \wedge \neg A)$. But this is in direct contradiction with (CP). Again, one of the principles from the cluster cannot be upheld (unrestrictedly) if we are to accommodate conflicting obligations, or at least if we want to avoid that such conflicts result in plain contradictions.

Besides conflicting obligations, there are other types of circumstances in which plausible logical principles may fail to apply. One that we want to consider here concerns the violation of conditional obligations, i.e. statements of the form “If A is the case, then B is obligatory” – formally, $O(B \mid A)$. Each of the rules of factual detachment (FD) and deontic detachment (DD) is intuitively appealing as a rule for detaching unconditional obligations from conditional ones:

- (FD) If it is obligatory that B given condition A , and if A is the case, then it is obligatory that B : $A, O(B \mid A) \vdash OB$
- (DD) If it is obligatory that B given condition A , and if A is obligatory, then it is obligatory that B : $OA, O(B \mid A) \vdash OB$.

The combination of (FD) and (DD) is known to cause trouble in so-called contrary-to-duty cases: cases in which a secondary obligation kicks in once a possibly conflicting primary obligation was violated. The following is an example of such a case.

Lisa and Nathan are a couple since eleven months. Lisa wants their first anniversary to be special and promises Nathan to take him to a “real” restaurant. One can only pay in cash at this restaurant, so if they are going to the restaurant, then Lisa ought to withdraw one hundred dollars at an ATM beforehand. However, on the day of the event, Lisa changes her mind and decides that she is not going to the restaurant after all – perhaps she is no longer sure she wants to be Nathan’s girlfriend in the first place. In view of her promise, she (still) has the obligation to take Nathan to the restaurant: OA . She also still has the conditional obligation that, if she takes Nathan there, she has to withdraw the money: $O(B \mid A)$. However, if she is not going to any restaurant, then she should not withdraw a hundred dollars, since carrying around that much money for no reason would be hazardous: $O(\neg B \mid \neg A)$. And as it happens to be, she is not going to the restaurant: $\neg A$.

Let us now see how the combination of (FD) and (DD) causes trouble for cases like this. If the obligation OA is violated, i.e. $\neg A$ is the case, then the primary

conditional obligation $O(B \mid A)$ leads to the unconditional obligation OB via (DD), while the secondary (contrary-to-duty) obligation $O(\neg B \mid \neg A)$ leads to the unconditional obligation $O\neg B$ via (FD). In order to resolve this conflict, we must block the application of (DD) or that of (FD).¹

We will have much more to say about conflicting obligations and about the detachment of conditional obligations in the remainder of this paper. For now, these examples merely serve to illustrate a general point. In the circumstances described above – conflicting obligations and contrary-to-duty cases – one cannot consider principles such as the ones just mentioned as unrestrictedly valid. This leaves the logician who wants to explicate our reasoning in such cases with various options. One is to simply reject those principles, and hence declare a number of intuitive inferences simply invalid. Our stance towards this option is perhaps best summarized by the following words of van Benthem [96, p. 95]:

This is like turning down the volume on your radio so as not to hear the bad news. You will not hear much good news either.

A more promising option is to look for restricted versions or alternative, more fine-grained formulations of those principles. For instance, for the case of conflicting obligations, one may argue that (Agg) should only be applicable in case the conjunction of A and B is possible. For contrary-to-duty cases, one may reformulate (FD) as a principle that concerns dynamic updates, rather than (mere) factual input – see e.g. [97] where this is proposed.

We will not pursue this second option here, even though occasionally we will show that some concrete instances of it fail to deliver an appropriate logic of normative reasoning, either on philosophical or on purely technical grounds. Instead, we will focus on a third option, i.e. to take (some of) these problematic principles to be only valid in a defeasible, context-sensitive way.

That this option seems well in line with our intuitions is easily demonstrated by returning to our examples. As soon as Nathan realizes that looking after Ben is incompatible with going to that particular movie with Lisa, it seems quite rational to reject the conclusion that he ought to do both. But, suppose that his mother also made him promise to walk the family dog on Saturday evening. Would it be rational that, in view of the conflict concerning his afternoon plans, he also rejects the conclusion that he ought to go with Lisa for a veggie burger (at 6pm) and take the dog for a walk (at 10pm)? It seems that the one should have no bearing on the

¹Alternatively, we could bite the bullet and accept the outcome that both B and $\neg B$ are obligatory. But then our first illustration shows that we must give up other logical principles on pain of contradiction.

other. What this comes to is that, even if it makes sense to withdraw applications of (Agg) upon realizing that A and B are mutually exclusive, this need not affect other applications of (Agg).

In a similar vein, it seems quite natural that certain applications of (DD) are upheld *unless and until* it turns out that the unconditional obligation in the premises is violated. That Lisa has the obligation to withdraw money, even if she is not going to the restaurant at all, feels contra-intuitive to non-logicians. Is there something wrong with their intuitions? Not necessarily, and maybe even to the contrary. It seems quite justified that in cases like this, (DD) is treated as a *defeasible* rule of inference: the obligation is detached from the conditional obligation *provided* the unconditional obligation is not violated.

Note the difference between the third option and the first one. In our approach, we do not invalidate *principles*, we invalidate certain *applications* of principles and this is done only *when and where* necessary. This at once illustrates what we mean by context-sensitivity: whether an application of a certain principle or rule is validated or not depends on the specific context (the premises at issue).

The aforementioned clusters of principles governing obligations and permissions were originally introduced to hold unconditionally. The circumstances in which these principles are not (jointly) applicable, such as conflicts and violations, are often considered anomalous or exceptional. Other principles were acknowledged to be applicable only in a defeasible, context-sensitive manner right from their very introduction. We give only one example. Consider the *nullum crimen sine lege* principle: “If A is not forbidden, then A is permitted”. This principle is best thought of as a kind of *default* rule: assume (or infer) PA , unless $O\neg A$ follows from the premises. This rule is defeasible by its very nature, in the sense that at least some of its instances are violated in every interesting application context.

In order to apply inference rules in a logic in a context-sensitive, defeasible manner, the consequence relation of this logic has to be *non-monotonic*: given a set of premises from which a conclusion A is derivable, it must be possible to revoke A in the light of additional premises.² *Adaptive logics* (henceforth, ALs) provide a natural way to explicate the premise-sensitive, defeasible application of certain inference rules in a formal logic.

ALs are built on top of a core logic, called the *lower limit logic*, the inference rules of which hold unconditionally and unrestrictedly. An AL strengthens its lower limit logic by allowing a number of additional inference rules to be applied relative to the specific premises at hand. The term “adaptive logic” originates from this

²Formally, a logic \mathbf{L} is non-monotonic iff (if and only if) there are two sets of formulas Γ and Δ and there is a formula A such that A is \mathbf{L} -derivable from Γ , while A is *not* \mathbf{L} -derivable from $\Gamma \cup \Delta$.

premise-sensitivity: ALs “adapt” themselves to the premises under consideration.

Beside ALs, many other formalisms for modelling defeasible reasoning have been applied in a deontic context: default logic [53], defeasible deontic logic [71], formal argumentation theories [33; 77; 92; 18; 105], input/output logic [76], etc. These different frameworks are all linked to one another and to ALs in various ways – see e.g. [47] for some recent comparisons.

There is, however, a distinctive feature of ALs that sets them apart from other approaches to non-monotonic reasoning, viz. their dynamic proof theory. The idea behind this proof theory is that the non-monotonicity of the logic’s consequence relation is pushed into the object-level proofs. This means that a given derivation in a proof can become rejected in the light of other derivations within that same proof.³

Another important difference between the existing work on ALs and other types of non-monotonic logics is the pivotal role that classical logic (henceforth **CL**) plays within the latter. ALs are, at least in origin, more pluralistic in spirit regarding the meaning of the classical connectives, thus opening up to new perspectives on defeasible reasoning that are hard to detect when one sticks to **CL** as one’s underlying monotonic logic.⁴

The current paper’s aim is to motivate, explain, illustrate, and discuss the use of ALs in deontic logic. Published work on deontic ALs focusses mainly on conflict-tolerant deontic logics (logics that can accommodate conflicting obligations) and – to a lesser extent – on problems concerning factual and deontic detachment. So does the present paper. Near the end of the paper, however, we also indicate some of the possibilities that the adaptive logic framework creates for tackling other types of problems within deontic logic.

The outline of this paper is as follows. For ease of reference, we start by recalling the basic definitions concerning Standard Deontic logic, henceforth **SDL** (Section 2). In Section 3 we provide an introduction to the framework of ALs. By way of illustration, we first present two very simple adaptive logics that can handle examples as the one from the prelude (Section 3.1).

In Sections 5–7 we present and discuss a variety of conflict-tolerant deontic ALs that move further away from the standard view: unlike the logics from Section 3.1, the logics from Sections 5–7 have lower limit logics that are inferentially weaker than **SDL**. Section 4 provides the conceptual and technical basis for this discussion. Whereas Sections 5 and 6 are mainly based on existing work, Section 7 presents mostly new material that we think improves on the existing work in a number of

³We will define and illustrate the dynamic proof theory of ALs in Section 3.

⁴This aspect of ALs is nicely illustrated by our Section 7, where we introduce and discuss (adaptive) paraconsistent deontic logics.

ways – we explain this in Section 7.4.

Section 8 summarizes the merits and demerits of the conflict-tolerant ALs presented throughout Sections 3–7. In that section we also show how the simple logics introduced in Section 3.1 can be further refined in various ways.

The other main application of existing deontic ALs concerns the problem of detaching conditional obligations. We distinguish between various approaches to this problem in Section 9, and discuss adaptive versions of each of them.

In Section 10 we show how the *nullum crimen sine lege* principle can be captured within the AL framework, and how this gives rise to various extensions of the logics defined in previous sections. This at once paves the way for our last section in which we give a short summary of the paper and point to ideas for future research.

Throughout this paper our focus is on the illustration and motivation of the core ideas we present, rather than on formal details and meta-theoretical results. Whenever relevant, we provide pointers to the literature, cf. the subsections “further reading and open ends”.

Much of what we will write in this paper builds on Lou Goble’s work on normative conflicts, which is nicely summarized in [42]. We will provide references to specific parts of this (and other) work in due course. In general, we try to avoid overlap as much as possible, but whenever this maxim conflicts with keeping the present paper self-contained, we give priority to the latter.

We end this section with some more general comments regarding the plurality and diversity of logics to be discussed in this paper. Our stance on the matter can be described as follows.

For a start, various logics present themselves as useful depending on the specific type of application context, and the associated logical grammar one wants to study. But even if we keep the grammar fixed, there are various reasons for occupying oneself with not one but many logics for this grammar. That logic – even the logic of our most basic connectives like conjunction – is not god-given, and that there are no absolute grounds for preferring one logic over another, seems hardly contested nowadays. So all one can do is give pragmatic arguments, referring to general desiderata for logics on the one hand, and the needs of a given application on the other.

In the context of conflict-tolerant deontic logics, one way to argue for diversity is by referring to various explosion principles, as discussed in Section 4.2. For instance, if one does not *need* to accommodate conflicts between obligations and permissions, or if one can safely assume within a given domain that norms are at least internally consistent, then this should translate to one’s preferred logic for that domain. Moreover, there are many different ways one can interpret the O of a given (conflict-tolerant or other) deontic logic, which will yield different formal semantics

and hence different logics in turn.

Going non-monotonic (or in our case, going adaptive) does not reduce this plurality – quite to the contrary. To use Makinson’s words [61, p. 14]:

Leaving technical details aside, the essential message is as follows. Don’t expect to find *the* nonmonotonic consequence relation that will always, in all contexts, be the right one to use. Rather, expect to find several *families* of such relations, interesting *syntactic conditions* that they sometimes satisfy but sometimes fail, and principal *ways of generating* them mathematically from underlying structures.

Indeed, it will become clear throughout this paper that there are usually several interesting and sensible ways of going adaptive, starting from a given lower limit logic. In the absence of further philosophical arguments against the resulting logics, one needs to keep an open mind and study all of them.

2 Some formal preliminaries

Languages Throughout this paper, we use A, B, \dots as metavariables for formulas of a given formal language, and Γ, Δ, \dots as metavariables for sets of such formulas.

Let henceforth **CL** stand for the propositional fragment of classical logic, as based on a set of propositional variables (also called sentential letters) $\mathcal{S} = \{p, q, \dots\}$, the connectives $\neg, \vee, \wedge, \supset, \equiv$, and the logical constants \perp, \top . We use \mathcal{W} to denote the set of well-formed formulas in this language.

The language of **SDL** is obtained by adding to the grammar of **CL** the modal operators **O** for “it is obligatory that” and **P** for “it is permitted that”. We take both **O** and **P** (and the classical connectives) to be primitive by default in this paper; i.e. whenever one is defined in terms of the others in one logic or another, we will indicate so. For the sake of simplicity, we will focus on the fragment of this language in which no nested occurrences of **O** and **P** are allowed. This means that the set of well-formed formulas for **SDL** is defined as follows:

$$\mathcal{W}^d := \mathcal{W} \mid \neg\langle\mathcal{W}^d\rangle \mid \langle\mathcal{W}^d\rangle \vee \langle\mathcal{W}^d\rangle \mid \langle\mathcal{W}^d\rangle \wedge \langle\mathcal{W}^d\rangle \mid \langle\mathcal{W}^d\rangle \supset \langle\mathcal{W}^d\rangle \mid \langle\mathcal{W}^d\rangle \equiv \langle\mathcal{W}^d\rangle \mid \mathbf{O}\langle\mathcal{W}\rangle \mid \mathbf{P}\langle\mathcal{W}\rangle$$

Axiomatization The logic **SDL** is obtained by adding to **CL** the following axioms, rule, and definition:

- (K) $\mathbf{O}(A \supset B) \supset (\mathbf{O}A \supset \mathbf{O}B)$
- (D) $\mathbf{O}A \supset \neg\mathbf{O}\neg A$

- (N) if $\vdash A$, then $\vdash \mathbf{O}A$
 (Def_P) $\mathbf{P}A =_{df} \neg\mathbf{O}\neg A$

It is well-known that in the presence of (N), (K) can equivalently be expressed as the combination of the axiom of aggregation (Agg) and the rule of inheritance (Inh):

- (Agg) $(\mathbf{O}A \wedge \mathbf{O}B) \supset \mathbf{O}(A \wedge B)$
 (Inh) if $\vdash A \supset B$, then $\vdash \mathbf{O}A \supset \mathbf{O}B$

whence **SDL** can be equivalently characterized by adding (N), (Agg), (Inh), (D), and (Def_P) to **CL**. Note also that in the presence of (Agg), (D) is equivalent to the following principle:

- (P) $\neg\mathbf{O}(A \wedge \neg A)$

For ease of reference, we note some more derivable principles of **SDL**. The first is the axiom of distributivity (of **O** over \wedge):

- (Dist) $\mathbf{O}(A \wedge B) \supset (\mathbf{O}A \wedge \mathbf{O}B)$

Second, the replacement of equivalents rule (RE) is an immediate consequence of the behavior of \supset and \equiv in **CL** and (Inh):

- (RE) if $\vdash A \equiv B$, then $\vdash \mathbf{O}A \equiv \mathbf{O}B$

Third and last, in view of (Agg), (Inh), and the validity of disjunctive syllogism (DS) in **CL**, we have:

- (DDS) $(\mathbf{O}A \wedge \mathbf{O}(\neg A \vee B)) \supset \mathbf{O}B$

Semantics We work with the traditional Kripke-semantics for **SDL**, but to prepare for the semantics of other logics to be presented below, we work with a designated “actual” world. An **SDL**-model M is a quadruple $\langle W, w_0, R, v \rangle$, where W is a non-empty set of worlds, $w_0 \in W$ is the actual world, $R \subseteq W \times W$ is a serial⁵ accessibility relation and $v : W \rightarrow \mathcal{S}$ is a valuation function. $R(w)$ (the image of w under R) is the set of worlds that are accessible from the viewpoint of w , $R(w) = \{w' \mid (w, w') \in R\}$.

The semantic clauses for the sentential variables and the connectives are as usual; those for **O** and **P** are as follows:

- (SC1) $M, w \models \mathbf{O}A$ iff $M, w' \models A$ for all $w' \in R(w)$

⁵ R is serial iff for every $w \in W$, there is a $w' \in W$ such that $(w, w') \in R$.

(SC2) $M, w \models PA$ iff $M, w' \models A$ for some $w' \in R(w)$

Truth of a formula A at a world w is given by the relation \models . Truth in a model $M = \langle W, w_0, R, v \rangle$ is simply truth at w_0 . We say that M is a model of Γ iff all the members of Γ are true in M , i.e. iff for all $B \in \Gamma$, $M, w_0 \models B$. Semantic consequence is then defined as the preservation of truth in all models: $\Gamma \Vdash A$ iff A is true in all models of Γ .

Following customary notation, let $|A|_M =_{\text{df}} \{w \mid M, w \models A\}$. $|A|_M$ is also called the *truth set* (intension) of A . Note that the semantic clause for \mathbf{O} can be equivalently rewritten as follows: $M, w \models \mathbf{O}A$ iff $R(w) \subseteq |A|_M$.

3 Adaptive logics

Adaptive logics were originally introduced by Diderik Batens around the 1980s, and have since been applied to various forms of defeasible reasoning.⁶ The aim of this section is to highlight the basic features of ALs by means of a running example, viz. the logics \mathbf{SDL}_p^r and \mathbf{SDL}_p^m . These logics can handle simple cases of conflicting obligations such as the running example from the beginning of this paper. We explain the idea behind both logics in Section 3.1. Generic definitions for all ALs in the standard format from [10] are given in Section 3.2. We mention the most salient properties of all logics that are defined within this format in Section 3.3. Finally, we discuss some variants of the standard format that will turn out useful in the remainder of this paper (Section 3.4).

3.1 The basics

Before introducing the logics \mathbf{SDL}_p^r and \mathbf{SDL}_p^m , we present another predicament from Nathan’s life. The example will be used to illustrate the proof theory of \mathbf{SDL}_p^r and \mathbf{SDL}_p^m .

One evening, Nathan comes home from school. As soon as he enters the kitchen, he hears his father: “Remember, Nathan, it’s your turn to do the dishes tonight. Do them this time!” His mother immediately adds: “And forget about playing with Ben tonight. Before supper, you will do nothing but your homework. Your grades are terrible lately!” Not too enthusiastically, Nathan heads towards his room to do his homework. As soon as he wants to enter it, his twin sister Olivia leaves hers, in great despair: “Nathan, you have to help me. I am on “Ben watch” tonight, but he is driving me crazy and I am expecting this really, really important phone call! Play with him until supper, will you? I’ll do anything for you in return!” Nathan

⁶See Section 3.5 for references to the literature on ALs.

finds himself again in a difficult situation. He can obey his father and do the dishes. No problem there. But what should he do until supper? Olivia helps him out on a quite regular basis and he feels he ought to return the favor this time. But if he plays with Ben, he will not be able to do his homework.

This example and the one from the *preludium* have three important characteristics in common. The first is that they both concern a situation in which an agent faces several obligations, not all of which can be fulfilled. The second is that, for each of the separate obligations, there is some *prima facie* reason. In the example from the *preludium*, Nathan's specific obligations hold in view of the general rule "One ought to keep one's promises". In this last example, the obligation that Nathan ought to do the dishes holds in view of his father's command. The third characteristic is that, although not all obligations can be met, some of them can. Nathan cannot look after Ben and take Lisa to that particular movie, but he *can* go for a veggie burger in the evening. Similarly, Nathan cannot do his homework and at the same time play with Ben, but he *can* do the dishes.

In this paper, we will use the term *prima facie obligations* for any obligation for which there is some *prima facie* reason (some general rule, a command, ...). As the examples show (and as we all know from daily life), there are situations in which not all *prima facie* obligations can be *binding*. Nathan cannot go to that particular movie with Lisa (in view of his promise to her) and at the same time *not* go there (in view of his promise to his mother and the fact that six year olds are not allowed for this particular movie). We will use the term *actual obligations* for obligations that are binding and that should be acted upon.

Examples in which not all *prima facie* obligations can be met raise the following question: how do we decide, in a given situation, which *prima facie* obligations are actual obligations and which are not? A first answer to this question seems to be that at least those *prima facie* obligations should be considered as actual obligations that are not in conflict with any other *prima facie* obligation. This seems to capture nicely our intuitions behind the examples. The fact that Nathan made conflicting promises with respect to what he will do in the afternoon should not prevent him from going for a veggie burger in the evening. The fact that he cannot help out his twin sister as well as obey his mother should not rule out that he at least obeys his father.

This is exactly the idea behind the logics \mathbf{SDL}_p^r and \mathbf{SDL}_p^m presented in this section: *prima facie* obligations are considered as actual obligations *unless and until* it turns out that they are in conflict with some other *prima facie* obligation. Or, put in a somewhat different form, the logics \mathbf{SDL}_p^r and \mathbf{SDL}_p^m validate the inference of actual obligations from *prima facie* obligations *as much as possible*. The exact meaning of this "as much as possible" will become clear below.

The logics have two further characteristics in common: they allow us to (a) accommodate conflicts at the level of *prima facie* obligations, and (b) reason about actual obligations in the standard way (i.e., applying all axioms of **SDL**).⁷ What (a) comes to is that both logics are conflict-tolerant: they do not lead to unwanted conclusions in the face of conflicting *prima facie* obligations.

We will now show, step by step, how the logics **SDL_p^r** and **SDL_p^m** are obtained.

The lower limit logic In order to make the distinction between *prima facie* obligations and actual obligations, we will use a bi-modal language that contains two obligation operators: \mathbf{O}^p and \mathbf{O} . The first is used for *prima facie* obligations, the second for actual obligations. The language is defined as follows:

$$\mathcal{W}^p := \mathcal{W} \mid \mathbf{O}\langle\mathcal{W}\rangle \mid \mathbf{O}^p\langle\mathcal{W}\rangle \mid \neg\langle\mathcal{W}^p\rangle \mid \langle\mathcal{W}^p\rangle \vee \langle\mathcal{W}^p\rangle \mid \langle\mathcal{W}^p\rangle \supset \langle\mathcal{W}^p\rangle \mid \langle\mathcal{W}^p\rangle \wedge \langle\mathcal{W}^p\rangle \mid \langle\mathcal{W}^p\rangle \equiv \langle\mathcal{W}^p\rangle$$

Note that we exclude nesting; i.e. none of the two operators occurs within the scope of another operator.

To obtain a logic that is tolerant with respect to conflicting *prima facie* obligations (characteristic (a) above), \mathbf{O}^p is treated as a propertyless operator, a “dummy”. This means that e.g. *prima facie* obligations cannot be derived from other *prima facie* obligations. Characteristic (b) is realized by assuming that \mathbf{O} is the ought-operator of **SDL**.

Let us call the resulting logic **SDL_p** – it is just **SDL** extended with the dummy-operator \mathbf{O}^p . In AL terminology, what we have done so far is define the *lower limit logic* of our AL. This logic constitutes the monotonic core of the AL. In other words, it consists of all the principles (rules, axioms) that are unconditionally valid within the logic.⁸

In order to obtain a logic that validates the inference from *prima facie* obligations to actual obligations as much as possible, **SDL_p** needs to be strengthened. One option that does *not* work is to simply add the axiom

$$(A) \quad \mathbf{O}^p A \supset \mathbf{O} A$$

to **SDL_p**. Let us call the resulting logic **SDL_p⁺**. In this stronger logic, conflicts at the level of *prima facie* obligations will be trivialized: if $\vdash_{\mathbf{CL}} \neg(A_1 \wedge \dots \wedge A_n)$,

⁷Our characteristics (a) and (b) correspond to Goble’s criteria of adequacy a) and b) for *prima facie* oughts versus all-things-considered oughts [42, p. 257].

⁸The lower limit logic of every AL has to satisfy certain general desiderata, which will be spelled out in Section 3.2.

then $\text{OP}A_1, \dots, \text{OP}A_n \vdash_{\text{SDL}_p^+} B$ for any B .⁹ Of course, we could weaken the logic of O , but then we would lose characteristic (b). This shows that we need a more refined way to fulfill our aim. We will now show how this can be realized within the framework of adaptive logics.

Going adaptive What we need is a way to steer between SDL_p and SDL_p^+ , avoiding the weakness of the former but also the explosive character of the latter. More precisely, we need a defeasible, context-sensitive version of (A). This can be done by assuming that formulas like $\text{OP}p \wedge \neg \text{Op}$, $\text{OP}q \wedge \neg \text{Oq}$, etc. are false *unless and until* proven otherwise.

In AL terminology, such formulas – the negations of defeasible assumptions – are called *abnormalities*.¹⁰ We will use Ω_p to refer to the set of all those abnormalities, i.e. all formulas of the form $\text{OP}A \wedge \neg \text{OA}$.

In an adaptive proof, we can derive formulas on the assumption that certain abnormalities are false. This is most easily illustrated with an example. Let d stand for “Nathan washes the dishes”, b for “Nathan plays with Ben” and h for “Nathan does his homework”. The *prima facie* obligations that Nathan faces in our second running example may then be formalized as $\text{OP}d$, $\text{OP}b$ and $\text{OP}(\neg b \wedge h)$. An adaptive proof from $\Gamma = \{\text{OP}d, \text{OP}b, \text{OP}(\neg b \wedge h)\}$ in which we try to derive the actual obligation for Nathan to wash the dishes (Od) may then look as follows:

1	$\text{OP}d$	Prem	\emptyset
2	$\text{OP}b$	Prem	\emptyset
3	$\text{OP}(\neg b \wedge h)$	Prem	\emptyset
4	$\text{Od} \vee \neg \text{Od}$	SDL	\emptyset
5	$\text{Od} \vee (\text{OP}d \wedge \neg \text{Od})$	1,4; SDL	\emptyset
6	Od	5; RC	$\{\text{OP}d \wedge \neg \text{Od}\}$

The fourth element of each line in this proof represents the condition of that line. This condition is always a (possibly empty) set of abnormalities. After introducing the premises on lines 1-3, we have used excluded middle to derive a new formula at line 4, and then derived line 5 using lines 1 and 4. We use “SDL” as a generic name for all rules and axioms of **SDL**. At line 6, Od is derived on the condition that the

⁹To see why, note that in SDL_p , conflicting actual obligations are trivialized just as in **SDL**. If we moreover allow for the unrestricted application of (A), this means that also conflicts at the level of *prima facie* obligations are trivialized.

¹⁰Our terminology here and below suggests a link with Makinson’s *Default Assumption Consequence Relations* [61]. Indeed, as shown in [99], one can establish an exact correspondence between Makinson’s construction and ALs that use the minimal abnormality strategy.

abnormality $\text{O}^p d \wedge \neg \text{O} d$ is false. This is done by means of the rule RC (shorthand for *conditional rule*) which allows us to push abnormalities to the condition within an adaptive proof.

Here are two other applications of RC:

:	:	:	:
7	$\text{O} b \vee (\text{O}^p b \wedge \neg \text{O} b)$	2; SDL	\emptyset
8	$\text{O} b$	7; RC	$\{\text{O}^p b \wedge \neg \text{O} b\}$
9	$\text{O}(\neg b \wedge h) \vee (\text{O}^p(\neg b \wedge h) \wedge \neg \text{O}(\neg b \wedge h))$	3; SDL	\emptyset
10	$\text{O}(\neg b \wedge h)$	9; RC	$\{\text{O}^p(\neg b \wedge h) \wedge \neg \text{O}(\neg b \wedge h)\}$

At this point, the reader may become suspicious. Clearly, $\text{O} b$ and $\text{O}(\neg b \wedge h)$ cannot both be true. By means of well-known **SDL**-principles, we can derive from our premises that at least one of the two corresponding abnormalities is true:

$$11 \quad (\text{O}^p b \wedge \neg \text{O} b) \vee (\text{O}^p(\neg b \wedge h) \vee \neg \text{O}(\neg b \wedge h)) \quad 2,3; \text{SDL} \quad \emptyset$$

Formulas like the one at line 11 are called *Dab-formulas* (**Dab** is shorthand for “disjunction of abnormalities”). Note that this **Dab**-formula is derived on the empty condition. Hence, it is an unconditional consequence of the premises – it cannot be false, if the premises are true. Moreover, it is *minimal*: neither of its disjuncts $\text{O}^p b \wedge \neg \text{O} b$ or $\text{O}^p(\neg b \wedge h) \vee \neg \text{O}(\neg b \wedge h)$ is derived on the empty condition in the above proof.¹¹

At lines 8 and 10 respectively, we relied on the assumption that the first, respectively the second of these abnormalities is false. But line 11 clearly indicates that those two assumptions cannot be jointly true. So a mechanism is needed to *retract* the inferences at lines 8 and 10.

Formally, this is taken care of by a *marking definition*, which stipulates which lines are marked, and hence considered “out” at a given stage of an adaptive proof. How the marking proceeds depends on the so-called *adaptive strategy*. The logics **SDL**_p^r and **SDL**_p^m are based respectively based on the Reliability strategy and the Minimal Abnormality strategy. Let us look at these in turn.

Reliability For **SDL**_p^r, a line is marked whenever its condition contains an abnormality that is a disjunct of a minimal **Dab**-formula that has been derived in the

¹¹In fact, neither of them *can* be derived in this proof on the empty condition, since they simply do not follow from Γ by **SDL**_p.

same proof. For instance, in the above example, lines 8 and 10 are marked, whereas all other lines are not marked. This is indicated by a \checkmark -symbol at the end of the line:

1	$\text{O}^{\text{P}}d$	Prem	\emptyset
2	$\text{O}^{\text{P}}b$	Prem	\emptyset
3	$\text{O}^{\text{P}}(\neg b \wedge h)$	Prem	\emptyset
4	$\text{O}d \vee \neg \text{O}d$	SDL	\emptyset
5	$\text{O}d \vee (\text{O}^{\text{P}}d \wedge \neg \text{O}d)$	1; SDL	\emptyset
6	$\text{O}d$	5; RC	$\{\text{O}^{\text{P}}d \wedge \neg \text{O}d\}$
7	$\text{O}b \vee (\text{O}^{\text{P}}b \wedge \neg \text{O}b)$	2; SDL	\emptyset
8	$\text{O}b$	7; RC	$\{\text{O}^{\text{P}}b \wedge \neg \text{O}b\} \checkmark$
9	$\text{O}(\neg b \wedge h) \vee (\text{O}^{\text{P}}(\neg b \wedge h) \wedge \neg \text{O}(\neg b \wedge h))$	3; SDL	\emptyset
10	$\text{O}(\neg b \wedge h)$	9; RC	$\{\text{O}^{\text{P}}(\neg b \wedge h) \wedge \neg \text{O}(\neg b \wedge h)\} \checkmark$
11	$(\text{O}^{\text{P}}b \wedge \neg \text{O}b) \vee (\text{O}^{\text{P}}(\neg b \wedge h) \vee \neg \text{O}(\neg b \wedge h))$	2,3; SDL	\emptyset

In general, lines with an empty condition are never marked. But also those lines whose condition is not problematic in view of the minimal **Dab**-formulas in the proof remain unmarked (witness line 6 in the example). So at the end of the day, some instances of (A) are trustworthy in the light of the premises, while other instances of (A) are not. This illustrates the premise-sensitivity of adaptive logics that was mentioned in Section 1.¹²

The fact that lines can become marked in a proof means that we cannot simply define logical consequence in terms of being derivable in a proof. We need a more robust notion of derivability; this is called *final derivability*. The basic idea is that something is finally derivable if and only if it can be derived in a “stable” way. Spelling out this intuition is not as straightforward as it may seem, as it requires quantification over extensions of proofs. We refer to Definitions 3.3 and 3.4 in the next section for the exact details.

¹²Some may argue that, in light of the premise set, the inferences at lines 8 and 10 were never rational in the first place. Admittedly, in cases like Γ above, it can easily be seen which *prima facie* obligations can make it into actual obligations, and which cannot on pain of triviality. But then again, such cases are not the only ones we may encounter in practice. Conflicts may exist between many different *prima facie* obligations, and they may be very hard to trace. Once we move to the predicate level, it may even be undecidable whether a certain set of *prima facie* obligations is consistent. One may well be calculating up to eternity before ever knowing for sure whether a certain inference is safe.

Minimal Abnormality The logic \mathbf{SDL}_p^m works in exactly the same way as \mathbf{SDL}_p^r , except that the marking in both logics is slightly different. Consider the following extension of our proof:

1	$\mathcal{O}^P d$	Prem	\emptyset
2	$\mathcal{O}^P b$	Prem	\emptyset
3	$\mathcal{O}^P(\neg b \wedge h)$	Prem	\emptyset
4	$\mathcal{O}d \vee \neg \mathcal{O}d$	SDL	\emptyset
5	$\mathcal{O}d \vee (\mathcal{O}^P d \wedge \neg \mathcal{O}d)$	1; SDL	\emptyset
6	$\mathcal{O}d$	5; RC	$\{\mathcal{O}^P d \wedge \neg \mathcal{O}d\}$
7	$\mathcal{O}b \vee (\mathcal{O}^P b \wedge \neg \mathcal{O}b)$	2; SDL	\emptyset
8	$\mathcal{O}b$	7; RC	$\{\mathcal{O}^P \neg r \wedge \neg \mathcal{O} \neg r\} \checkmark$
9	$\mathcal{O}(\neg b \wedge h) \vee (\mathcal{O}^P(\neg b \wedge h) \wedge \neg \mathcal{O}(\neg b \wedge h))$	3; SDL	\emptyset
10	$\mathcal{O}(\neg b \wedge h)$	9; RC	$\{\mathcal{O}^P(\neg b \wedge h) \wedge \neg \mathcal{O}(\neg b \wedge h)\} \checkmark$
11	$(\mathcal{O}^P b \wedge \neg \mathcal{O}b) \vee (\mathcal{O}^P(\neg b \wedge h) \vee \neg \mathcal{O}(\neg b \wedge h))$	2,3; SDL	\emptyset
12	$\mathcal{O}(b \vee h)$	8; (Inh)	$\{\mathcal{O}^P b \wedge \neg \mathcal{O}b\} ?$
13	$\mathcal{O}(b \vee h)$	10; (Inh)	$\{\mathcal{O}^P(\neg b \wedge h) \wedge \neg \mathcal{O}(\neg b \wedge h)\} ?$

Note first that, since we used the formula at line 8 to derive the one at line 12, the latter inherits the condition of the former. Likewise, line 13 is derived on the same condition as line 10. Taken together, lines 12 and 13 indicate that $\mathcal{O}(b \vee h)$ is true if either of the abnormalities in the \mathbf{Dab} -formula at line 11 is false.

Should lines 12 and 13 in this proof be marked? Clearly, there is a problem with at least one of the two involved abnormalities. Since there is no reason to prefer the falsehood of one over that of the other, that means both abnormalities are “unreliable” at this proof stage. However, if we assume that as few abnormalities as possible are false – until and unless proven otherwise –, then in cases like these we will assume that only one of both abnormalities is true. And in that case, $\mathcal{O}(b \vee h)$ does follow.

To turn this idea into a general method for marking lines in an adaptive proof, we need the concept of a (\subset -minimal) choice set. Suppose that the \mathbf{Dab} -formulas at the current stage of our proof are $\mathbf{Dab}(\Delta_1), \mathbf{Dab}(\Delta_2), \dots$. A choice set of $\{\Delta_1, \Delta_2, \dots\}$ is a set φ that contains at least one member of each Δ_i . In view of our proof, we know that (at least) the members of one choice set of $\{\Delta_1, \Delta_2, \dots\}$ should be true in view of the premises. However, we are still free to assume that *only* the members of a \subset -minimal choice set of $\{\Delta_1, \Delta_2, \dots\}$ are true. Suppose now moreover that, for every such minimal choice set φ , we can derive A on a condition Θ that does not

overlap with φ . This means that we have sufficient reasons to infer A – since every minimally abnormal way of interpreting the current proof stage will make A true. Following this general line of reasoning, lines 12 and 13 will not be marked, but lines 8 and 10 will be marked just as before.

To summarize: one can be cautious to different degrees when reasoning defeasibly; this difference is modeled by the adaptive strategy. According to the *reliability strategy* (usually indicated with a superscript r), both lines 12 and 13 are marked. According to *minimal abnormality*, they are both unmarked. In general, reliability is slightly weaker (more cautious) than minimal abnormality – see Theorem 3.15.

We now turn to the general characterization of ALs. A critical discussion of the logics \mathbf{SDL}_p^r and \mathbf{SDL}_p^m is postponed until Section 8. There we evaluate \mathbf{SDL}_p^r and \mathbf{SDL}_p^m by various criteria that are introduced in Section 4.

3.2 The standard format

The locus classicus for the standard format is Batens' [10]; an earlier version of it appeared in [9]. Here, we will follow the more recent presentation from [12], indicating minor differences where they occur. We will only explain the general characteristics, and refer to the works just cited for more details.

Standardly, a logic is defined as a function $\mathbf{L} : \wp(\mathcal{W}_L) \rightarrow \wp(\mathcal{W}_L)$, where \mathcal{W}_L is the set of formulas in the formal language of \mathbf{L} . This also holds for adaptive logics. For adaptive logics in standard format, the language should at least contain the classical disjunction \vee .¹³ For reasons of convenience, we will in this paper assume that the language also contains the classical negation \neg .

Every logic \mathbf{AL}^x is defined by a triple:

1. A lower limit logic \mathbf{LLL} . This is a reflexive, transitive, monotonic and compact logic¹⁴ that has a characteristic semantics and for which at least the disjunction \vee behaves classically.
2. A set of abnormalities $\Omega \subseteq \mathcal{W}_{\mathbf{LLL}}$ that is specified in terms of one or several logical forms.
3. An adaptive strategy: Reliability (when $x = r$) or Minimal Abnormality (when $x = m$).

¹³The assumption that the language contains a classical disjunction can be questioned on philosophical grounds. In [72; 12] it is shown that one can do without this assumption, if one rephrases everything in terms of multi-conclusion sequents.

¹⁴Let \mathbf{Cn} be the consequence operation of a logic \mathbf{L} . \mathbf{L} is *reflexive* iff for all Γ , $\Gamma \subseteq \mathbf{Cn}_L(\Gamma)$. \mathbf{L} is *transitive* iff for all Γ, Γ' : if $\Gamma' \subseteq \mathbf{Cn}_L(\Gamma)$, then $\mathbf{Cn}_L(\Gamma \cup \Gamma') \subseteq \mathbf{Cn}_L(\Gamma)$. \mathbf{L} is *monotonic* iff for all Γ, Γ' , $\mathbf{Cn}_L(\Gamma) \subseteq \mathbf{Cn}_L(\Gamma \cup \Gamma')$. \mathbf{L} is *compact* iff for all Γ, A , if $A \in \mathbf{Cn}_L(\Gamma)$, then there is a finite $\Gamma' \subseteq \Gamma$ with $A \in \mathbf{Cn}_L(\Gamma')$.

For instance, the adaptive logic \mathbf{SDL}_p^r from Section 3.1 is defined by the triple $\langle \mathbf{SDL}_p, \Omega_p, r \rangle$; the logic \mathbf{SDL}_p^m is defined by $\langle \mathbf{SDL}_p, \Omega_p, m \rangle$. The logical form that specifies Ω_p is $\mathbf{O}^p A \wedge \neg \mathbf{O} A$. In general, it is required that only countably many logical forms specify the set of abnormalities.

In the remainder of this section, we presuppose a fixed \mathbf{LLL} , Ω , and strategy $x \in \{r, m\}$. We use $\mathbf{Dab}(\Delta)$ to denote the (classical) disjunction of the members of Δ , where it is presupposed that Δ is a finite subset of Ω .

Proof theory The core idea behind the adaptive proof theory is to take all the inference rules of the lower limit logic for granted and to allow in addition for defeasible applications of some rules. Defeasible inferences in adaptive proofs are conditional. Hence, the usual way in which lines in proofs are presented – by a line number, a formula, and a justification – is enriched by a fourth element: a condition. A condition in turn is a set of abnormalities.

Suppose some formula A is derived on the condition $\{B_1, B_2, \dots, B_n\} \subseteq \Omega$. The intended reading is that A is derived on the assumption that all the abnormalities B_1, \dots, B_n are false.

Adaptive proofs are characterized by three generic rules and a marking definition. Let us first discuss the generic rules. In what follows we skip the line numbers and justification of lines.

$$\begin{array}{ll}
 \text{Prem} & \text{If } A \in \Gamma: \\
 & \frac{\vdots \quad \vdots}{A \quad \emptyset} \\
 \\
 \text{RU} & \text{If } A_1, \dots, A_n \vdash_{\mathbf{LLL}} B: \\
 & \frac{\begin{array}{l} A_1 \quad \Delta_1 \\ \vdots \quad \vdots \\ A_n \quad \Delta_n \end{array}}{\begin{array}{l} B \quad \Delta_1 \cup \dots \cup \Delta_n \end{array}} \\
 \\
 \text{RC} & \text{If } A_1, \dots, A_n \vdash_{\mathbf{LLL}} B \vee \mathbf{Dab}(\Theta): \\
 & \frac{\begin{array}{l} \vdots \quad \vdots \\ A_n \quad \Delta_n \end{array}}{\begin{array}{l} B \quad \Delta_1 \cup \dots \cup \Delta_n \cup \Theta \end{array}}
 \end{array}$$

By means of Prem, any premise may be introduced on the empty condition. Of course, we do not need any defeasible assumptions in order to state premises. The unconditional rule (RU) makes it possible to apply any inference rule of \mathbf{LLL} in an adaptive proof. Note that RU may also be applied to lines that were derived on defeasible assumptions, i.e. where $\Delta_i \neq \emptyset$ for some $i \in \{1, \dots, n\}$. The assumptions

under which the A_i 's were derived thus carry forward to the line at which B is derived. In virtue of Prem and RU, ALs inherit all the inferential power of **LLL**: any **LLL**-proof can be rephrased as an **AL**-proof just by adding the empty condition in the fourth column and by replacing the respective **LLL**-rules by Prem or RU.

In Section 3.1, we sometimes referred explicitly to the axiom that was used to derive a specific line in an adaptive proof. In the remainder we use RU as a metavariable for all axioms and (derivable) rules of the **LLL**; whenever useful we will indicate in footnotes which exact axioms were applied in order to derive a new line.

The rule that permits the introduction of new conditions in an adaptive proof is RC, the conditional rule. Suppose that we can derive $B \vee \text{Dab}(\Theta)$ by means of **LLL**, i.e. that either B is the case or some of the abnormalities in Θ . Then RC allows us to derive B on the assumption that none of the abnormalities in Θ is true. Making this assumption amounts to adding all members of Θ to the condition by means of RC. Similarly as for RU, in case some of the lines that are used for the inference step are conditional inferences, we carry forward their conditions as well.

Apart from the possibility to make conditional derivations via RC, a second distinctive aspect of adaptive proofs is the marking definition, which is applied at each *stage* of a proof. A stage is simply a sequence of lines, obtained by the application of the above rules. For concrete examples, we will identify stages with their last line. So for example the last stage of the last proof displayed in Section 3.1 is referred to as stage 13.

$\text{Dab}(\Delta)$ is a **Dab**-formula at stage s of a proof, iff it is the second element of a line of the proof with an empty condition, and derived by means of RU.¹⁵ $\text{Dab}(\Delta)$ is a *minimal* **Dab**-formula at stage s iff there is no other **Dab**-formula $\text{Dab}(\Delta')$ at stage s such that $\Delta' \subset \Delta$. Where $\text{Dab}(\Delta_1), \text{Dab}(\Delta_2), \dots$ are the minimal **Dab**-formulas at stage s of a proof, let $\Sigma_s(\Gamma) = \{\Delta_1, \Delta_2, \dots\}$. Finally, let $U_s(\Gamma) = \bigcup \Sigma_s(\Gamma)$.

Definition 3.1 (Marking for **AL^r**). *A line l is marked at stage s iff, where Δ is its condition, $\Delta \cap U_s(\Gamma) \neq \emptyset$.*

In terms of assumptions, this means that according to the reliability strategy, an assumption is “safe” at stage s iff the corresponding abnormality is not a member of $U_s(\Gamma)$, and an inference is “safe” at s iff it only relies on assumptions that are safe at s .

Returning to our example of page 15, we can see that $\Sigma_{11}(\Gamma) = \{\{\text{O}^{\text{P}}b \wedge \neg \text{O}b, \text{O}^{\text{P}}(\neg b \wedge h) \wedge \neg \text{O}(\neg b \wedge h)\}\}$ and hence $U_{11}(\Gamma) = \{\text{O}^{\text{P}}b \wedge \neg \text{O}b, \text{O}^{\text{P}}(\neg b \wedge h) \wedge \neg \text{O}(\neg b \wedge h)\}$.

¹⁵Here, our terminology differs slightly from that in [12]. Batens uses the term “**Dab**-formula at stage s ” for any disjunction of abnormalities derived at s , whereas we preserve it for those that have been derived by means of RU. Batens calls the latter “*inferred Dab*-formulas”.

This explains why lines 8 and 10 are marked at stage 11 of the proof.

The marking definition for minimal abnormality requires some more terminology. Recall that, where Σ is a set of sets, φ is a *choice set* of Σ iff for every $\Delta \in \Sigma$, $\varphi \cap \Delta \neq \emptyset$. φ is a *minimal choice set* of Σ iff there is no choice set ψ of Σ such that $\psi \subset \varphi$. Let $\Phi_s(\Gamma)$ be the set of \subset -minimal choice sets of $\Sigma_s(\Gamma)$. Marking for minimal abnormality proceeds as follows:

Definition 3.2 (Marking for \mathbf{AL}^m). *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 at which A is derived on a condition Θ for which $\Theta \cap \varphi = \emptyset$.*

In our simple example on page 16, $\Phi_{13}(\Gamma) = \{\{\mathbf{O}^p b \wedge \neg \mathbf{O} b\}, \{\mathbf{O}^p(\neg b \wedge h) \wedge \neg \mathbf{O}(\neg b \wedge h)\}\}$. In view of condition (ii) in Definition 3.2, lines 8 and 10 are marked for minimal abnormality at stage 13, but lines 12 and 13 are not. Note that all of these lines are marked for reliability.

If a line that has A as its second element is marked at stage s , this indicates that according to our best insights at this stage, A cannot be considered derivable. If the line is unmarked at stage s , we say that A is derivable at stage s of the proof. Since marks may come and go as a proof proceeds, we also need to define a stable notion of derivability. This definition is the same for both strategies.

Where s is a proof stage, an *extension* of s is every stage s' that contains the lines occurring in s in the same order. Hence putting lines in front of s , inserting them somewhere in between lines of s , or simply adding them at the end of s may all result in an extension of s .

Definition 3.3. *A is finally derived from Γ at line l of a stage s iff (i) A is the second element of line l , (ii) line l is unmarked at s , and (iii) every extension of s in which line l is marked may be further extended in such a way that line l is unmarked again.*

Definition 3.4. $\Gamma \vdash_{\mathbf{AL}^x} A$ ($A \in \mathbf{Cn}_{\mathbf{AL}^x}(\Gamma)$) iff A is finally derived at a line of a stage in an \mathbf{AL}^x -proof from Γ .

Note that in order to be finally derivable, A must be derived at a line l , where $l \in \mathbb{N}$. This means that every formula that is finally derivable from Γ can be finally derived in a *finite* proof from Γ . However, we need a meta-level argument to show that clauses (ii) and (iii) in Definition 3.3 are satisfied, and hence that $\Gamma \vdash_{\mathbf{AL}^x} A$.

Semantics On the supposition that \mathbf{LLL} is characterized by a model theoretic semantics (with the semantic consequence relation $\Vdash_{\mathbf{LLL}}$), one can also give a semantics for \mathbf{AL}^x . The rough idea is as follows: from the set of \mathbf{LLL} -models of a

given premise set, \mathbf{AL}^x selects a subset of “preferred” models. Whatever holds in those preferred models, follows by \mathbf{AL}^x .¹⁶

What counts as a preferred model depends on the strategy used. For minimal abnormality, only those models of the premise set are selected which verify a \subset -minimal set of abnormalities. For reliability, a threshold of *unreliable* abnormalities (with respect to a given premise set Γ) is defined, and only the models that do not verify any abnormalities other than the unreliable ones, are selected.

To define the \mathbf{AL}^x -semantics in exact terms, we need some more notation. Validity of a formula A in a model M will be written as $M \models A$. M is an \mathbf{LLL} -model of Γ iff $M \models A$ for all $A \in \Gamma$. $\mathcal{M}_{\mathbf{LLL}}(\Gamma)$ denotes the set of \mathbf{LLL} -models of Γ . Where M is an \mathbf{LLL} -model, its *abnormal part* is given by $\text{Ab}(M) =_{\text{df}} \{B \mid B \in \Omega, M \models B\}$.

For reliability, the selection of preferred models is in some sense analogous to the marking definition. $\text{Dab}(\Delta)$ is a *minimal Dab-consequence* of Γ iff $\Gamma \Vdash_{\mathbf{LLL}} \text{Dab}(\Delta)$ and there is no $\Delta' \subset \Delta$ for which $\Gamma \Vdash_{\mathbf{LLL}} \text{Dab}(\Delta')$. Where $\text{Dab}(\Delta_1), \text{Dab}(\Delta_2), \dots$ are the minimal Dab-consequences of Γ , let $\Sigma(\Gamma) = \{\Delta_1, \Delta_2, \dots\}$. Let $U(\Gamma) = \bigcup \Sigma(\Gamma)$. We say that $U(\Gamma)$ is the set of *unreliable* formulas with respect to Γ .

Definition 3.5. An \mathbf{LLL} -model M of Γ is *reliable* iff $\text{Ab}(M) \subseteq U(\Gamma)$.

Definition 3.6. $\Gamma \Vdash_{\mathbf{AL}^r} A$ iff A is verified by all reliable models of Γ .

For minimal abnormality, the semantics’ simplicity stands in sharp contrast to the intricate marking definition:

Definition 3.7. An \mathbf{LLL} -model M of Γ is *minimally abnormal* iff there is no \mathbf{LLL} -model M' of Γ such that $\text{Ab}(M') \subset \text{Ab}(M)$.

Definition 3.8. $\Gamma \Vdash_{\mathbf{AL}^m} A$ iff A is verified by all minimally abnormal models of Γ .

In the remainder, we will denote the set of \mathbf{AL}^x -models of a set Γ by $\mathcal{M}_{\mathbf{AL}^x}(\Gamma)$.

Upper Limit Logic The so-called *upper limit logic* of \mathbf{AL}^x is defined as the Tarski-logic¹⁷ obtained by adding all negations of abnormalities as axioms to \mathbf{LLL} . That is, where $\Omega^\neg = \{\neg A \mid A \in \Omega\}$, $\Gamma \vdash_{\mathbf{ULL}} A$ iff $\Gamma \cup \Omega^\neg \vdash_{\mathbf{LLL}} A$. By the compactness of \mathbf{LLL} , $\Gamma \vdash_{\mathbf{ULL}} A$ iff there are $B_1, \dots, B_n \in \Omega$ such that

¹⁶Note that this is similar to the semantics of circumscription (where models are selected in which the abnormal predicates have a minimal extension) and Shoham-style preferential semantics (where all the \prec -minimal models are selected, for a given order \prec on the models of a premise set). However, in ALs, the selection depends on purely syntactic properties of the models, viz. the formulas (more specifically, the abnormalities) that they verify. This in turn gives ALs fairly strong meta-theoretic properties – see Section 3.3.

¹⁷A Tarski-logic is a logic whose consequence relation is reflexive, monotonic, and transitive.

$\Gamma \cup \{\neg B_1, \dots, \neg B_n\} \vdash_{\mathbf{LLL}} A$. \mathbf{AL}^x can be seen as steering a middle course between \mathbf{LLL} and \mathbf{ULL} (see Theorem 3.15 below).

In our running example, \mathbf{SDL}_p^+ is the upper limit logic of both \mathbf{SDL}_p^r and \mathbf{SDL}_p^m . Note that in general, \mathbf{ULL} does not depend on the strategy of \mathbf{AL}^x .

3.3 Some meta-properties of ALs in standard format

Once defined within the standard format, it is guaranteed that an AL satisfies a number of meta-properties. We only mention some of them here for the ease of reference. Their proofs can be found in [10].

First of all, the dynamic proof theory is sound and complete with respect to the semantics of \mathbf{AL}^x :

Theorem 3.9 (Soundness and Completeness). $\Gamma \vdash_{\mathbf{AL}^x} A$ iff $\Gamma \Vdash_{\mathbf{AL}^x} A$.

It follows from this result that one can rely on semantic considerations in order to prove that a formula A is finally derivable from a given Γ . We will in the remainder rely freely on Theorem 3.9, switching between the semantic and proof theoretic consequence relation where suitable.

Recall that the semantics of an AL consists in selecting a subset of the \mathbf{LLL} -models of Γ . Now, when a model M is not selected, we should be able to justify this in terms of another model M' that *is* selected, and is more normal than M . This is what the following theorem gives us:

Theorem 3.10 (Strong Reassurance). *If $M \in \mathcal{M}_{\mathbf{LLL}}(\Gamma) - \mathcal{M}_{\mathbf{AL}^x}(\Gamma)$, then there is an $M' \in \mathcal{M}_{\mathbf{AL}^x}(\Gamma)$ such that $\text{Ab}(M') \subset \text{Ab}(M)$.*

In other words, the preference relation defined in terms of \subset and the abnormal part relation is smooth with respect to every set $\mathcal{M}_{\mathbf{LLL}}(\Gamma)$.¹⁸ It is well-known that a selection semantics based on such a smooth preference relation warrants the following properties in turn:¹⁹

Theorem 3.11 (Consistency Preservation). *If Γ has \mathbf{LLL} -models, then $\mathcal{M}_{\mathbf{AL}^x}(\Gamma) \neq \emptyset$. Hence, Γ is \mathbf{AL}^x -trivial iff Γ is \mathbf{LLL} -trivial.*

Theorem 3.12 (Cumulative Indifference). *If $\Gamma' \subseteq \text{Cn}_{\mathbf{AL}^x}(\Gamma)$, then $\text{Cn}_{\mathbf{AL}^x}(\Gamma) = \text{Cn}_{\mathbf{AL}^x}(\Gamma \cup \Gamma')$.*

¹⁸A partial order \prec is *smooth* with respect to a set X iff for all $x \in X$, either x is \prec -minimal in X , or there is some \prec -minimal $y \in X$ such that $y \prec x$.

¹⁹See e.g. [60].

In the literature on non-monotonic logics, cumulative indifference is often divided into two properties: cumulative transitivity or cut (if $\Gamma' \subseteq \text{Cn}_{\mathbf{AL}^\times}(\Gamma)$, then $\text{Cn}_{\mathbf{AL}^\times}(\Gamma \cup \Gamma') \subseteq \text{Cn}_{\mathbf{AL}^\times}(\Gamma)$) and cumulative or cautious monotonicity (if $\Gamma' \subseteq \text{Cn}_{\mathbf{AL}^\times}(\Gamma)$, then $\text{Cn}_{\mathbf{AL}^\times}(\Gamma) \subseteq \text{Cn}_{\mathbf{AL}^\times}(\Gamma \cup \Gamma')$).

Strong reassurance, consistency preservation, and cumulative indifference are generally considered desirable for non-monotonic consequence relations, see e.g. [61]. It speaks in favor of ALs (in standard format) that they satisfy each of these properties. In particular, cautious monotonicity is a very intuitive property: if a formula follows from a premise set Γ , then it ought to follow from any Γ' that is obtained by extending Γ with some logical consequences of Γ . The extended premise set Γ' contains no genuinely new information, as the additions are in a sense already contained in Γ .

Suppose that Γ and Γ' are **LLL**-equivalent, i.e. $\text{Cn}_{\mathbf{LLL}}(\Gamma) = \text{Cn}_{\mathbf{LLL}}(\Gamma')$. It follows that they have the same set of **LLL**-models and that $U(\Gamma) = U(\Gamma')$. Hence in view of the semantics, they will also have the same **AL^x**-models, and hence be **AL^x**-equivalent. So we have a fairly straightforward criterion to decide when two premise sets are equivalent according to **AL^x**:²⁰

Theorem 3.13. *If $\text{Cn}_{\mathbf{LLL}}(\Gamma) = \text{Cn}_{\mathbf{LLL}}(\Gamma')$, then $\text{Cn}_{\mathbf{AL}^\times}(\Gamma) = \text{Cn}_{\mathbf{AL}^\times}(\Gamma')$.*

The next property on the list is specific to ALs, as it concerns the notion of an abnormality. It will be of particular use in Sections 5-7.

Say a premise set Γ is *normal* iff $\Gamma \cup \{\neg A \mid A \in \Omega\}$ is not **LLL**-trivial; in other words, iff it is **ULL**-consistent. The theorem states that every adaptive logic is as powerful as its upper limit logic when normal premise sets are concerned:²¹

Theorem 3.14 (ULL-recapture). *Γ is a normal premise set iff $\text{Cn}_{\mathbf{AL}^\times}(\Gamma) = \text{Cn}_{\mathbf{ULL}}(\Gamma)$.*

The last theorem simply recalls the relation between **LLL**, **AL^r**, **AL^m** and **ULL**, which was illustrated in Section 3.1:

Theorem 3.15. $\text{Cn}_{\mathbf{LLL}}(\Gamma) \subseteq \text{Cn}_{\mathbf{AL}^r}(\Gamma) \subseteq \text{Cn}_{\mathbf{AL}^m}(\Gamma) \subseteq \text{Cn}_{\mathbf{ULL}}(\Gamma)$.

3.4 Variants and extensions of the standard format

In this section, we briefly consider two variants of the standard format that are useful in the context of deontic reasoning; we will occasionally refer back to both

²⁰Similar criteria for equivalence are discussed in [14]; an extended and updated version of this paper can be found in [87, Chapter 4].

²¹Our name for the theorem is inspired by discussions in paraconsistent logic, where a similar property is called “classical recapture” [81].

variants in the remainder of this paper. We focus on the essential ideas in both cases; the metatheory of these (and many other) variants of the standard format is studied at length in [87, Chapter 5].

Normal Selections The minimal abnormality strategy corresponds to what is called the *skeptical* solution to the problem of multiple extensions in default logic.²² That is, A is finally \mathbf{AL}^m -derivable from Γ if and only if, for *every* maximal set $\Delta \subseteq \Omega^\neg$ such that $\Gamma \cup \Delta$ is \mathbf{LLL} -satisfiable, $\Gamma \cup \Delta \vdash_{\mathbf{LLL}} A$.²³

Rather than taking the universal quantification over such maximal sets, one may also quantify existentially over them. That is, say $\Gamma \vdash_{\mathbf{AL}^n} A$ iff *there is a* maximal set $\Delta \subseteq \Omega^\neg$ such that $\Gamma \cup \Delta \vdash_{\mathbf{LLL}} A$. The superscript n refers to “normal selections”, which is the name of the adaptive strategy of the resulting logics. Proof-theoretically, such logics are characterized in exactly the same way as ALs in standard format, with the only exception that the marking definition is simplified:

Definition 3.16 (Marking for Normal Selections). *A line l in a proof with condition Δ is marked at stage s iff $\mathbf{Dab}(\Delta)$ is derived on the empty condition at s .*

The consequence relation $\vdash_{\mathbf{AL}^n}$ is usually very strong, and yet does not trivialize premise sets as long as they are \mathbf{LLL} -consistent. However, it will not in general be closed under \mathbf{LLL} . More generally, many of the nice properties we discussed in Section 3.3 can fail for $\vdash_{\mathbf{AL}^n}$.

To understand this, consider the logic \mathbf{SDL}_p^n , defined by the triple

$$\langle \mathbf{SDL}_p, \Omega_p, \text{normal selections} \rangle \tag{1}$$

Let $\Gamma = \{\mathbf{O}^p p, \mathbf{O}^p q, \neg \mathbf{O}(p \wedge q)\}$. Note that this premise set has the following minimal \mathbf{Dab} -consequence:

$$(\mathbf{O}^p p \wedge \neg \mathbf{O} p) \vee (\mathbf{O}^p q \wedge \neg \mathbf{O} q) \tag{2}$$

Since this is a minimal \mathbf{Dab} -consequence of Γ , both $\mathbf{O} p$ and $\mathbf{O} q$ are individually compatible with Γ . Hence, both $\mathbf{O} p$ and $\mathbf{O} q$ are finally \mathbf{SDL}_p^n -derivable from Γ , on the respective conditions $\{\mathbf{O}^p p \wedge \neg \mathbf{O} p\}$ and $\{\mathbf{O}^p q \wedge \neg \mathbf{O} q\}$. However, $\mathbf{O} p \wedge \mathbf{O} q$ is not finally \mathbf{SDL}_p^n -derivable from Γ , since one needs to rely on the falsity of both

²²Analogous problems arise in Input/Output-logic, inheritance networks, and abstract argumentation, giving rise to similar distinctions between less and more cautious “modes of reasoning” – see [87, Sect. 2.8] for more discussion.

²³This is a well-known property that is often used in the metatheory of ALs; see e.g. [99] for a proof of it.

abnormalities in order to obtain this conclusion. This shows that the consequence relation of \mathbf{SDL}_p^n is not closed under the rule of conjunction, even if \wedge behaves classically in the lower limit logic.

In the context of deontic logic, normal selections has been used to characterize one variant of Horty’s approach to conflicting obligations [89]. Likewise, it has been applied to characterize constrained Input/Output-logics that are defined in terms of the join of the maximal unconflicted sets of generators [88]. We will shortly return to the latter systems in Section 9.2.

Prioritized adaptive logics Another useful variation of the standard format is obtained by distinguishing between various types of abnormalities, and by giving priority to some of these when minimizing abnormality. This can be done in at least three clearly distinct ways – see [98] for a detailed study of these. Here we will only discuss one of these three, viz. the so-called *lexicographic adaptive logics* first presented in [102]; we moreover confine ourselves to the minimal abnormality-variant of these systems. Although these logics can be fully characterized in terms of a dynamic proof theory, we focus on their semantics, which is a straightforward generalization of the \mathbf{AL}^m -semantics.

Let $\langle \Omega_i \rangle_{i \in I}$ (for $I \subseteq \mathbb{N}$) be a sequence of sets of abnormalities. Intuitively, the idea is that we consider the members of Ω_1 to be the “worst” abnormalities; those of Ω_2 as “slightly less problematic (yet still abnormal)”, etc. Thus, we want to make sure when selecting models, that we first minimize with respect to Ω_1 , next with respect to Ω_2 , etc. This is done in terms of a lexicographic order \sqsubset on the abnormal parts of the models:

Definition 3.17. *Where $\Delta, \Delta' \subseteq \bigcup_{i \in I} \Omega_i$: $\Delta \sqsubset \Delta'$ iff there is a $j \in I$ such that (1) for all $k < j$ (if any), $\Delta \cap \Omega_k = \Delta' \cap \Omega_k$ and (2) $\Delta \cap \Omega_j \subset \Delta' \cap \Omega_j$.*

The preference relation \sqsubset on abnormal parts of models yields a smooth preference relation on every set $\mathcal{M}_{\mathbf{LLL}}(\Gamma)$ [102]. Hence, just as for minimal abnormality, we can select the \sqsubset -minimal models of a premise set and define semantic consequence in terms of those models. Then it is again a matter of routine to show that this consequence relation satisfies all the nice properties of the standard format.

For an illustration of this format of ALs, let us suppose that *prima facie* obligations come in various degrees $i \in \mathbb{N}$ of importance, where degree 1 is most important, degree 2 is slightly less important, etc. Let $O_i^p A$ denote that A is *prima facie* obligatory, with degree i . Then intuitively, we expect that from $\{O_1^p p, O_2^p q, O_2^p r, \neg O(p \wedge q)\}$ we can derive $O p$ but not $O q$. Moreover, we also expect $O r$ to be derivable, since r is not involved in the conflict. This is exactly the result we obtain if we define our sequence of sets of abnormalities as $\langle \{O_i^p A \wedge \neg O A\} \rangle_{i \in \mathbb{N}}$.

The format of lexicographic ALs is relatively new; the first ideas for it date back to 2010. It has been applied to deontic logic in [103], where a lexicographic variant of the logic from [67] is proposed.

3.5 Further reading

The first ALs were developed a little before 1980 by Diderik Batens, as a new, “dialectical” approach to (non-explosive) reasoning with inconsistent theories.²⁴ Nowadays these logics are called “inconsistency-adaptive logics” – more on them in Section 7.²⁵

From its first days, this research was pluralist in the sense that various (monotonic) paraconsistent logics were used to define ALs. Around the mid 1990s, the idea emerged that besides inconsistency, various other types of “abnormality” with respect to classical (propositional or first order) logic can be used as a basis to define ALs – see e.g. [6]. The resulting logics are nowadays called “corrective ALs”, in contradistinction to “ampliative” ALs, which only saw light around 2000.²⁶ The latter are, roughly, ALs that characterize a given type of inference which goes *beyond* one’s chosen standard of deduction (usually first order **CL**), such as compatibility [13], inductive generalization [11], abduction [69; 19], etc.

The notion of an adaptive strategy was only fully developed in the 1990s – see in particular [7]. Before that, only the proof theory of *reliability* and the semantics of *minimal abnormality* were known.

The standard format as presented in this section, was introduced in [10]. Its further development in turn facilitated applications in various new areas during the last decade, ranging from foundations of set theory [108], over causal discovery [56; 20], to deontic logic.

For a recent and compact introduction into ALs (with a focus on their application to paraconsistent reasoning), we refer to [12]. A thorough discussion of the standard format and several of its generalizations can be found in Part I of [87]. Slightly older papers that present the basics of ALs are [9] and [10].

ALs have been compared to various other generic frameworks for defeasible and/or non-monotonic reasoning in the past, including Makinson’s *default assumption consequence relations* [99], abstract argumentation [95], and modal logics [2]. There is also an interesting line of research on the relation between ALs and Rescher-Manor consequence relations for “contextualized” reasoning with inconsistent premises [82]. In fact, the logics \mathbf{SDL}_p^r , \mathbf{SDL}_p^m , and \mathbf{SDL}_p^n can be seen as adaptive

²⁴In [4], Batens refers to an (unpublished) manuscript from 1979, “Dynamische processen en dialectische logica’s”, as the first paper on this subject.

²⁵The term “adaptive” appears to be introduced in 1981 [4].

²⁶See e.g. [69] for a discussion of this distinction.

variants of the *Free*, the *Strong*, and the *Weak* Rescher-Manor consequence relation respectively [68].

4 Revisionist adaptive deontic logics

The logics \mathbf{SDL}_p^f and \mathbf{SDL}_p^m from Section 3 reserve the **SDL**-operator O for actual obligations, while they allow for the non-trivial formalization of conflicting (*prima facie*) obligations in terms of the new operator O^p . Via this grammatical enrichment, we obtain a conflict-tolerant adaptive logic, without having to revise any of the core principles of **SDL**. Indeed, \mathbf{SDL}_p^x is built on top of \mathbf{SDL}_p , which is in turn an extension of **SDL**.

Instead of extending the grammar of **SDL** while keeping its core principles intact, we may also accommodate conflicts by keeping the grammar of **SDL** intact while giving up some of its core principles. This means that we *revise* the underlying logic, to use the terminology from [42]. We therefore call the adaptive logics based on such “weak” deontic logics *revisionist adaptive deontic logics*. The aim of sections 5–7 is to present and discuss this branch of ALs.

We provide some general insight into the various types of revisionist (adaptive) deontic logics that are on the market in Section 4.1. Next, we will introduce some conceptual machinery that allows us to compare and evaluate such logics (Section 4.2).

4.1 **SDL**: three ways of giving it up (while keeping it)

If we are to reason non-trivially in the face of conflicting obligations, we need to give up at least some part of **SDL**. For the time being, let us focus on conflicts of the type $OA \wedge O\neg A$ (we will consider several other types below). First, if the logic of \neg is classical, then the (D)-axiom needs to be given up in order to avoid that everything follows from $OA \wedge O\neg A$. This means we are left with the minimal normal modal logic **K**, which is fully characterized by **CL**, the rule of necessitation (N) and the normality schema (K).

But giving up (D) alone will not do. As soon as (Agg), (Inh), and *Ex Contradictione Quodlibet* (ECQ) are valid, deontic conflicts result in deontic explosion, i.e. the conclusion that everything is obligatory:²⁷

$$OA, O\neg A \vdash OB \tag{DEX}$$

²⁷(ECQ) is the (classically valid) inference from $A, \neg A$ to arbitrary B .

Suppose OA and $O\neg A$. By (Agg), $O(A \wedge \neg A)$. By (ECQ) and (Inh), OB . Since all three of these principles are derivable within **K**, deontic conflicts imply deontic explosion also in this minimal logic.

So at least one of (Agg), (Inh), or (ECQ) has to go. It can be shown – and will be shown in the next three sections – that giving up either (Agg), or (Inh), or (ECQ) is sufficient in order to accommodate conflicts of the type $OA \wedge O\neg A$.²⁸ So in the remainder we will focus on these three principles, rather than on the “official” characterization of **SDL** in terms of (N), (K) and (D).

In Section 5, we will consider deontic logics that are obtained by giving up (Inh).²⁹ This means that e.g. $O(A \wedge B)$ does not imply OA , and OA does not imply $O(A \vee C)$ in these logics, absent further information about A , B , and C . As a result, $O(A \wedge B)$ can be true for conflicting (i.e., mutually incompatible) A and B , but this need not imply that OC is true for any arbitrary (non-contradictory) C .

Section 6 is concerned with conflict-tolerant deontic logics that invalidate (Agg). Thus, in these logics, OA and OB can be true without $O(A \wedge B)$ being true. As a result, the step from $OA \wedge O\neg A$ to $O(A \wedge \neg A)$ is blocked and we cannot get to the conclusion that any B is obligatory.

Finally, Section 7 focuses on alternative, weaker accounts of negation, which invalidate (ECQ). This allows us to keep (D).

So there are several, well-studied ways to avoid (DEX) and thus to accommodate deontic conflicts within a formal logic. However, giving up principles of **SDL** comes at a price. As we will show below, these principles are at the heart of intuitively plausible patterns of inference – see Section 4.2 for a number of examples. Giving up the principles means that one either has to deny head-on the validity of those inferences, or to explain them as enthymatic arguments, i.e. arguments with a number of tacit, hidden premises. Even if such a strategy is successful to some extent, it turns out very difficult to develop a general logical (and philosophically justifiable) procedure that allows one to obtain such tacit premises for a given case.

Going adaptive allows us to give up principles, whilst keeping them *as much as possible*, i.e., as long as they do not lead to deontic explosion. The core idea behind revisionist adaptive deontic logics is to start from a monotonic, conflict-tolerant deontic logic **L** and to try to apply the missing **SDL**-rule(s) in a premise-sensitive, defeasible way, thus steering a middle course between the excesses of **SDL** and the inferential weakness of **L**.

²⁸One may of course give up even more principles, but we will focus on the simple cases where only one of the three is given up. All that we write on revisionist deontic logics and their adaptive extensions applies *mutatis mutandis* to such weaker logics.

²⁹In some but not all of these logics, also (Agg) is restricted. In all of them, classical logic is preserved for the connectives and replacement of equivalents (RE) holds.

Before we continue, an important side-remark is in place. In [42, Sect. 5.4], Goble also develops two new, monotonic conflict-tolerant deontic logics that are inferentially very powerful, in the sense that they validate (a variant of) (Agg), (DDS), and (Dist). The basic idea behind these logics is to give up the principle of extensionality (RE), and to opt for a weaker notion of “analytic equivalence” instead. In recent (unpublished) work, Anglberger and Korbmacher have developed a semantics for the resulting logics, based on truthmaker semantics for hyperintensional logics [32]. We will not discuss these new systems in the present paper, since it is as yet unclear whether and how sensible adaptive logics based on them could be developed.

4.2 Criteria for comparison and evaluation

When discussing and comparing the ALs defined in the next three sections, we will look at two aspects in particular. First, we will consider various types of deontic conflicts, and compare the logics in terms of which of these types they can accommodate properly. Second, we look at how the logics behave with respect to specific benchmark examples known from the literature.

Explosion principles In the specific context of conflict-tolerant deontic logics, it is common to demand some additional consistency constraints on top of the consistency preservation property from Theorem 3.11. In particular, we want to take great care to avoid the validity of *explosion principles*, i.e. principles according to which a set of arbitrary formulas is derivable given a (specific type of) normative conflict. These can come in various types, as we now explain.

We already referred to the principle of deontic explosion (DEX) in Section 4.1. In [93], some more refined explosion principles are specified that serve as touchstones for measuring the conflict-tolerance of various deontic logics. Here are some examples:³⁰

$$OA, O\neg A \vdash OB \vee O\neg B \tag{3}$$

$$OA, O\neg A \vdash OB \vee PB \tag{4}$$

$$OA, O\neg A \vdash OB \vee \neg O\neg B \tag{5}$$

$$OA, O\neg A \vdash PB \tag{6}$$

Principles (3)-(6) weaken the right-hand side of (DEX). We can devise further – equally undesirable – explosion principles by strengthening its left-hand side via the addition of logically unrelated information. For instance, where γ is any subset of $\{OD, \neg O\neg D, PE, \neg O\neg E, \neg OF, \neg O\neg F, PG, P\neg G\}$,

$$\{OA, O\neg A\} \cup \gamma \vdash OB \vee PB \tag{7}$$

³⁰Recall that we treat O and P as primitive operators unless stated otherwise; cf. Section 2.

More fine-grained explosion principles may be obtained by stipulating that principles like (3)-(7) are avoided even for B that satisfy certain additional constraints. For instance, Goble showed that the following principle is valid in deontic logics which restrict (Agg) to conjunctions of jointly consistent obligations [41]:

$$\text{If } \not\vdash \neg B, \text{ then } OA, O\neg A \vdash OB \quad (8)$$

The above forms of explosion are all still limited in (at least) one sense, in that they are focused on binary conflicts between obligations, i.e. formulas of the form $OA \wedge O\neg A$. There seems to be no reason to us as to why one should focus solely on such types of conflicts between norms, ignoring all others. For instance, there seems to be no logical reason why self-contradictory norms should be excluded – if an authority can issue mutually incompatible commands, then why can't it issue (highly complex but) self-contradictory commands as well? Likewise, why not consider conflicts between obligations and permissions?

Consider the following variant of an example from [?, p. 305]: a couple you know is having a party. One of them leaves a message: “I am sorry, you cannot come – it’s close friends only.” The other also leaves a message: “you can surely come to the party if you like – there will anyway be plenty of food for everyone.” Absent further information, the resulting norms can best be formalized as $\neg Pp$ and Pp , where p stands for “go to the party”. Even if we assume that O and P are interdefinable, this does not result in a conflict of the form $OA \wedge O\neg A$, but rather in a direct contradiction, i.e. $OA \wedge \neg OA$.

So all in all, there seem to be reasons for taking into account explosion principles such as the following:

$$OA, P\neg A \vdash OB \quad (9)$$

$$O(A \wedge \neg A) \vdash OB \quad (10)$$

Candidate conflict-tolerant deontic logics should be tested not only for the validity of (DEX), but also for the validity of more refined principles like (3)-(8) above. In doing so, we do not consider it the task of any such logic to invalidate all forms of explosion; rather, we treat the explosion principles as a useful way to compare and classify given deontic logics.

In the next two sections, we will focus on the following explosion principles – apart from (DEX):

$$\begin{array}{ll}
 \text{O}(A \wedge \neg A) \vdash \text{OB} & (\text{DEX-O}\perp) \\
 \text{P}(A \wedge \neg A) \vdash \text{PB} & (\text{DEX-P}\perp) \\
 \text{OA} \wedge \text{P}\neg A \vdash B & (\text{DEX-OP}\neg) \\
 \text{OA} \wedge \neg \text{PA} \vdash B & (\text{DEX-O}\neg\text{P})
 \end{array}$$

We choose these five principles since they allow us to compare the (non)explosive behavior of the various logics discussed below in a succinct way. In Section 7 we will consider some additional forms of explosion that can be avoided by using paraconsistent deontic logics.

Benchmark examples. Research in the fields of deontic logic and non-monotonic logic is to a large extent driven by a relatively small set of benchmark examples aimed at testing the formal system in question (the reader may be familiar with Tweety the penguin, the good Samaritan, and the gentle murderer, just to name a few). When faced with such examples, counter-intuitive outcomes are taken to reflect badly on a formal system, so these benchmark examples provide a criterion for checking whether a formal system meets our informal intuitions.

A warning is in order here, however. The fact that a formal system provides intuitive outcomes for the relevant benchmark examples is not a sufficient condition for positively evaluating the system in question. For instance, the system may be devised in an *ad hoc* manner to deal specifically with a small set of examples, at the cost of violating one or more rationality postulates. Moreover, some of these examples may reflect intuitions on which not everyone agrees, leaving room for dispute. In some cases the fact that our logic does *not* give us the expected outcome for some concrete example may inform us that our intuitions are perhaps incoherent, whence this is not in itself a sufficient reason to reject the logic. So, as was the case with explosion principles, we will use our benchmark examples as means to classify given logics, not as absolute criteria for their usefulness.³¹

With this warning in mind, let us list a number of examples which have been used to evaluate conflict-tolerant deontic logics studied in the literature. For each of them, we indicate some of the basic **SDL**-principles which allow us to infer the conclusion from the given premises. We use (CL) as a generic name for all inferences that are **CL**-valid.

1. *The Smith Argument.* — (Agg), (Inh), (CL)

³¹For a critical discussion of the use of examples as intuition-pumps in the evaluation of logics for defeasible reasoning, see [78].

- (i) Smith ought to fight in the army or perform alternative service to his country ($O(f \vee s)$).
- (ii) Smith ought not to fight in the army ($O\neg f$).
- \therefore (iii) Smith ought to perform alternative service to his country (Os).

2. *The Jones Argument.* — (Inh), (CL)

- (i) Jones ought to tell a joke and sing a song ($O(j \wedge s)$).
- \therefore (ii) Jones ought to tell a joke (Oj).

3. *The Roberts Argument, version 1.* — (Inh), (CL)

- (i) Roberts ought to pay federal taxes and register for national service ($O(t \wedge r)$).
- (ii) Roberts ought not to pay federal taxes but volunteer to help the homeless in his community ($O(\neg t \wedge v)$).
- \therefore (iii) Roberts ought to register for national service and ought to volunteer to help the homeless ($Or \wedge Ov$).

4. *The Roberts Argument, version 2.* — (Inh), (CL), (Agg)

- (i) Roberts ought to pay federal taxes and register for national service ($O(t \wedge r)$).
- (ii) Roberts ought not to pay federal taxes but volunteer to help the homeless in his community ($O(\neg t \wedge v)$).
- \therefore (iii) Roberts ought to register for national service and volunteer to help the homeless ($O(r \wedge v)$).

5. *The Thomas Argument.* — (Inh), (Agg), (CL)

- (i) Thomas ought to pay federal taxes and either fight in the army or perform alternative service to his country ($O(t \wedge (f \vee s))$).
- (ii) Thomas ought neither to pay federal taxes nor fight in the army ($O(\neg t \wedge \neg f)$).
- \therefore (iii) Thomas ought to perform alternative service to his country (Os).

6. *The Natascha Argument, version 1.* — (K) / (Inh), (Agg), (CL)

- (i) Natascha ought to take Sarah to the concert ($\mathbf{O}s$).
 - (ii) Natascha ought to take Martin to the concert ($\mathbf{O}m$).
 - (iii) It is not the case that Natascha ought to take Sarah *and* Martin to the concert ($\neg\mathbf{O}(s \wedge m)$).
 - (iv) If she takes Sarah, she ought to buy an extra ticket ($\mathbf{O}(s \supset t)$).
 - (v) If she takes Martin, she ought to buy an extra ticket ($\mathbf{O}(m \supset t)$).
- \therefore (vi) Natascha ought to buy an extra ticket ($\mathbf{O}t$).

7. *The Natascha Argument, version 2.* — (K) / (Inh), (Agg), (CL)

- (i) Natascha ought to take Sarah to the concert ($\mathbf{O}s$).
 - (ii) Natascha ought to take Martin to the concert ($\mathbf{O}m$).
 - (iii) Natascha ought not to take Sarah *and* Martin to the concert ($\mathbf{O}\neg(s \wedge m)$).
 - (iv) If she takes Sarah, she ought to buy an extra ticket ($\mathbf{O}(s \supset t)$).
 - (v) If she takes Martin, she ought to buy an extra ticket ($\mathbf{O}(m \supset t)$).
- \therefore (vi) Natascha ought to buy an extra ticket ($\mathbf{O}t$).

The Smith argument was first presented by Horty [49; 50; 52; 53]; the name “Smith” is due to Goble [43; 42]. The Jones, Roberts, and Thomas arguments are variations on examples from [43; 42]. The Natascha argument is new.

The validity of these arguments is not undisputed. The Jones argument, for instance, which concerns the application of the inheritance principle (Inh), has been called into question [34; 45; 74]. The Natascha argument concerns the derivation of a so-called *floating conclusion*, a conclusion entailed by each of two mutually conflicting obligations. The status of such conclusions is debatable.³²

In both versions of the Natascha argument, the idea behind the third premise is that for some reason or another, Natascha cannot possibly take both Sarah and Martin to the concert — e.g. because there is only one additional ticket left at the counter. In the absence of alethic modalities, we translate information concerning what is (im)possible directly into the language of **SDL**. While the first version of

³²See [51; 57; 78] for arguments pro and contra the derivation of floating conclusions in non-monotonic logic. In a moral context, the derivability of floating conclusions has been defended by Brink [26].

this argument relies on the principle of “ought implies can” (OIC) and contraposition, the second relies on the stronger principle of “permitted implies can” (PIC), interdefinability of O and P , and contraposition. Both (OIC) and (PIC) are controversial.³³ However, here we focus merely on the formal premises as such, not on the question whether they represent the example in the most natural way.

5 Adaptive inheritance

The first type of conflict-tolerant deontic logics mentioned in Section 4.1 is obtained by giving up or weakening the rule of inheritance (Inh). In the present section, we discuss one specific subclass of such logics, showing how they can be strengthened by going adaptive.

5.1 Logics with unconflicted inheritance

Restricting inheritance In a number of papers, Goble presented the **LUM**-family of deontic logics.³⁴ The language of these logics is just that of **SDL**, with P defined as the dual of O . The logics in the **LUM**-family do not simply reject inheritance, but replace it with a weaker principle that accounts for a number of intuitive applications of (Inh). This requires some explanation.

Let $UA =_{df} \neg(OA \wedge O\neg A)$ denote that A is unconflicted. All **LUM**-systems extend **CL** with the necessitation rule (N), the replacement of equivalents rule (RE), as well as the following rule of “unconflicted” inheritance (RUM):

$$\text{If } A \vdash B, \text{ then } UA, OA \vdash OB \quad (\text{RUM})$$

(RUM) allows for those applications of the inheritance rule (Inh) which involve only unconflicted obligations. In terms of permission, the rule states that whenever A is both obligatory and permitted, then whatever is logically weaker than A is also obligatory. This rule is therefore also sometimes referred to as “permitted inheritance” (RPM).

In addition to (N), (RE), and (RUM), the systems in the **LUM**-family are defined in terms of (a selection among) (P), (Agg), and “consistent” and “permitted” aggregation rules (C-Agg) and (P-Agg):

$$\text{If } \not\vdash \neg(A \wedge B) \text{ then } OA, OB \vdash O(A \wedge B) \quad (\text{C-Agg})$$

$$PA, PB, OA, OB \vdash O(A \wedge B) \quad (\text{P-Agg})$$

³³See [110] for a comprehensive discussion of the first of these two principles.

³⁴We adopt the presentation and nomenclature from [43]. For more details and references, we refer to Section 5.3.

Note that, since \mathbf{P} is the dual of \mathbf{O} , the antecedent of (P-Agg) just means that A and B are obligatory, and that neither of their negations are obligatory. The systems **LUM.a-LUM.c** extend **CL** by adding:

- LUM.a:** (N), (RE), (RUM), (Agg)
- LUM.b:** (N), (RE), (RUM), (P), (C-Agg)
- LUM.c:** (N), (RE), (RUM), (P), (P-Agg)

A semantics for these three logics can easily be obtained, following the well-known generalization of Kripke-semantics into neighbourhood semantics – cf. [29, Chapters 7 & 8] and [85]. Say a **LUM**-model is of the type $M = \langle W, w_0, n_{\mathbf{O}}, v \rangle$, where W is a non-empty set of worlds, $w_0 \in W$ is the actual world, $n_{\mathbf{O}} : W \rightarrow \wp(\wp(W))$ maps each world $w \in W$ to the set of *obligatory propositions at w* , and v is a valuation function. The semantic clause for \mathbf{O} in such models reads:

$$(\text{SC-O}) \quad M, w \models \mathbf{O}A \text{ iff } |A|_M \in n_{\mathbf{O}}(w)$$

Truth in a model is defined as usual, viz. as truth at w_0 ; semantic consequence is defined by quantifying over all models in which the premises are true.

This gives us the minimal classical modal logic **E**, which is characterized fully by adding (RE) to **CL**. Imposing a number of restrictions on such models, we obtain the additional axioms and rules listed above. These conditions are:

- (CO-RUM) if $X \in n_{\mathbf{O}}(w)$, $W \setminus X \notin n_{\mathbf{O}}(w)$, and $X \subseteq Y$, then $Y \in n_{\mathbf{O}}(w)$
- (CO-N) $W \in n_{\mathbf{O}}(w)$
- (CO-P) $\emptyset \notin n_{\mathbf{O}}(w)$
- (CO-Agg) if $X \in n_{\mathbf{O}}(w)$ and $Y \in n_{\mathbf{O}}(w)$, then $X \cap Y \in n_{\mathbf{O}}(w)$
- (CO-C-Agg) if $X \in n_{\mathbf{O}}(w)$, $Y \in n_{\mathbf{O}}(w)$, and $X \cap Y \neq \emptyset$, then $X \cap Y \in n_{\mathbf{O}}(w)$
- (CO-P-Agg) if $X \in n_{\mathbf{O}}(w)$, $Y \in n_{\mathbf{O}}(w)$, $W \setminus X \notin n_{\mathbf{O}}(w)$, and $W \setminus Y \notin n_{\mathbf{O}}(w)$, then $X \cap Y \in n_{\mathbf{O}}(w)$

For an extensive comparison and discussion of the various **LUM**-logics, we refer to [42, Sect. 5.3]. In the remainder, we will focus on ALs obtained from them.

Going adaptive To understand the specific motivation for going adaptive in the case of the **LUM**-logics, it will be useful to reconsider the benchmark examples from Section 4.2. The Smith and Jones arguments are invalid in all three of the **LUM**-logics, but valid once we add the premises $\mathbf{U}\neg f$ (for the Smith argument) and $\mathbf{U}(j \wedge s)$ (for the Jones argument). The Roberts and Thomas arguments are more problematic. In the Roberts argument, for instance, we cannot just add the

premises $\mathbf{U}(t \wedge r)$ and $\mathbf{U}(\neg t \wedge v)$ in order to render the argument valid, since doing so would trivialize the premise set.³⁵

More generally, it is problematic that in the **LUM**-systems we need to add the ‘tacit’ information that a formula is unconflicted before we can apply the restricted distribution rule. This worry was first raised in [94], and acknowledged by Goble:

For one thing, the additional non-conflict condition on the distribution rule seems rather *ad hoc*; there is little to recommend it except its success in disarming deontic explosion. For another, it seems risky to try to account for the plausibility of arguments by considering them enthymematic for straight-forwardly valid arguments. In context it may be all right to accept the alleged tacit premise, but we cannot rely on that. With more complicated arguments it might be quite uncertain what unspoken premises of non-conflict are implicitly present [43, pp. 210-211].

Both problems can be overcome by strengthening the **LUM**-systems within the adaptive logics framework. On the one hand, we can validate all those applications of distribution that do not lead to deontic explosion. On the other hand, it is the logic itself that fixes which applications of distribution are tolerable; no interference of any user is required for this. We explain how this works below, focusing on the adaptive extensions of the logic **LUM.a**. For the other logics in this family, the difficulties and properties are roughly analogous. We will point out salient differences as we go along.

The logics $\mathbf{LUM.a}^x$ A natural way of strengthening Goble’s **LUM**-systems is to work under the assumption that obligations are unconflicted, so that an obligation \mathbf{OA} behaves abnormally in case it is conflicted, i.e. in case $\neg\mathbf{UA}$ or, equivalently, $\mathbf{OA} \wedge \mathbf{O}\neg A$:

$$\Omega = \{\mathbf{OA} \wedge \mathbf{O}\neg A \mid A \in \mathcal{W}\}$$

The logic **ADPM.1^r** from [94] is the AL defined by the triple $\langle \mathbf{LUM.a}, \Omega, \text{reliability} \rangle$. In an **ADPM.1^r**-proof, (Inh) can be applied via the conditional rule RC, assuming that the obligations involved are not conflicted:

$$\begin{array}{llll} 1 & \mathbf{O}(p \wedge q) & \text{Prem} & \emptyset \\ 2 & \mathbf{O}p & 1; \text{RC} & \{\mathbf{O}(p \wedge q) \wedge \mathbf{O}\neg(p \wedge q)\} \end{array}$$

³⁵For Roberts, first note that $\vdash (t \wedge r) \supset \neg(\neg t \wedge v)$. By (RPM), $\neg\mathbf{O}\neg(t \wedge r) \supset (\mathbf{O}(t \wedge r) \supset \mathbf{O}\neg(\neg t \wedge v))$ or, equivalently, $\mathbf{O}\neg(t \wedge r) \vee (\mathbf{O}(t \wedge r) \supset \mathbf{O}\neg(\neg t \wedge v))$. By premises (i) and (ii) of the Roberts argument, we get $(\mathbf{O}(t \wedge r) \wedge \mathbf{O}\neg(t \wedge r)) \vee (\mathbf{O}(\neg t \wedge v) \wedge \mathbf{O}\neg(\neg t \wedge v))$ by **CL**. So adding $\mathbf{U}(t \wedge r)$ and $\mathbf{U}(\neg t \wedge v)$ would make the argument **CL**-inconsistent. For Thomas the argument is analogous.

The conditional derivation at line 2 is legitimate in view of the **LUM.a**-valid inference

$$\mathbf{O}(p \wedge q) \vdash \mathbf{O}p \vee (\mathbf{O}(p \wedge q) \wedge \mathbf{O}\neg(p \wedge q)) \quad (11)$$

Unfortunately, Goble pointed out that **ADPM.1^r** suffers from a problem [43, Sect. 4.3.1]. Although we can indeed apply distribution conditionally in **ADPM.1^r**, the corresponding application of RC in the proof is marked as soon as a (possibly unrelated) conflict follows from the premise set. The problem is best illustrated by means of a simple example.

1	$\mathbf{O}(p \wedge q)$	Prem	\emptyset
2	$\mathbf{O}r$	Prem	\emptyset
3	$\mathbf{O}\neg r$	Prem	\emptyset
4	$\mathbf{O}p$	1; RC	$\{\mathbf{O}(p \wedge q) \wedge \mathbf{O}\neg(p \wedge q)\} \checkmark$
5	$(\mathbf{O}(p \wedge q) \wedge \mathbf{O}\neg(p \wedge q)) \vee$ $(\mathbf{O}(p \wedge r) \wedge \mathbf{O}\neg(p \wedge r)) \vee$ $(\mathbf{O}(p \wedge \neg r) \wedge \mathbf{O}\neg(p \wedge \neg r))$	1-3; RU	\emptyset

The **Dab**-formula derived at line 5 is minimal at this stage of the proof, and causes the marking of line 4.³⁶ This **Dab**-formula is a minimal **Dab**-consequence of the premise set $\{\mathbf{O}(p \wedge q), \mathbf{O}r, \mathbf{O}\neg r\}$. Consequently, there is no extension of this proof in which line 4 is unmarked, and hence

$$\mathbf{O}(p \wedge q), \mathbf{O}r, \mathbf{O}\neg r \not\vdash_{\mathbf{ADPM.1}^r} \mathbf{O}p \quad (12)$$

The same holds if we use the minimal abnormality strategy instead of reliability (the reasoning is analogous):

$$\mathbf{O}(p \wedge q), \mathbf{O}r, \mathbf{O}\neg r \not\vdash_{\mathbf{ADPM.1}^m} \mathbf{O}p \quad (13)$$

This problem generalizes: in the presence of a conflict between two obligations, we can construct minimal **Dab**-formulas containing abnormalities pertaining to seemingly unrelated and unproblematic formulas, blocking unproblematic applications of RC. The logics **ADPM.1^r** and **ADPM.1^m** are therefore called *flip-flops* [10]. In the absence of conflicts, their consequence set is the same as their ULL, namely **SDL**.³⁷ As soon as one conflict is present, however, their consequence set collapses into that of their lower limit logic **LUM.a**.

³⁶By (11), $\mathbf{O}p \vee (\mathbf{O}(p \wedge q) \wedge \mathbf{O}\neg(p \wedge q))$. Suppose $\mathbf{O}p$. Then (i) by (Agg), $\mathbf{O}(p \wedge r)$ and, by (RUM) and **CL**, $\mathbf{O}\neg(p \wedge \neg r) \vee (\mathbf{O}(p \wedge r) \wedge \mathbf{O}\neg(p \wedge r))$; analogously (ii) by (Agg), $\mathbf{O}(p \wedge \neg r)$ and, by (RUM) and **CL**, $\mathbf{O}\neg(p \wedge r) \vee (\mathbf{O}(p \wedge \neg r) \wedge \mathbf{O}\neg(p \wedge \neg r))$. Altogether, by **CL**, $(\mathbf{O}(p \wedge q) \wedge \mathbf{O}\neg(p \wedge q)) \vee (\mathbf{O}(p \wedge r) \wedge \mathbf{O}\neg(p \wedge r)) \vee (\mathbf{O}(p \wedge \neg r) \wedge \mathbf{O}\neg(p \wedge \neg r))$.

³⁷It was shown in [94, Th. 7] that **SDL** is the ULL of **ADPM.1^r**.

There is a natural fix to this flip-flop problem, due to Goble [43]. Let $S(A)$ denote the set of all subformulas of A (including A itself). Where $S(A) = \{B_1, \dots, B_n\}$, we define³⁸

$$\sharp(A) = (\mathbf{O}B_1 \wedge \mathbf{O}\neg B_1) \vee \dots \vee (\mathbf{O}B_n \wedge \mathbf{O}\neg B_n)$$

Following Goble, we let $\mathbf{LUM.a}^r = \langle \mathbf{LUM.a}, \Omega^S, \text{reliability} \rangle$, where³⁹

$$\Omega^S = \{\sharp(A) \mid A \in \mathcal{W}\}$$

In an $\mathbf{LUM.a}^r$ -proof, the formula derived at line 5 of our proof above is no longer a **Dab**-formula. Rather, we obtain the following proof:

1	$\mathbf{O}(p \wedge q)$	Prem	\emptyset
2	$\mathbf{O}r$	Prem	\emptyset
3	$\mathbf{O}\neg r$	Prem	\emptyset
4	$\mathbf{O}p$	1; RC	$\{\sharp(p \wedge q)\}$
5	$\sharp(p \wedge q) \vee \sharp(p \wedge r) \vee \sharp(p \wedge \neg r)$	1-3; RU	\emptyset
6	$\sharp(p \wedge r)$	2,3; RU	\emptyset
7	$\sharp(p \wedge \neg r)$	2,3; RU	\emptyset

The abnormalities $\sharp(p \wedge q)$, $\sharp(p \wedge r)$, and $\sharp(p \wedge \neg r)$ denote the formulas (14), (15), and (16) respectively:

$$(\mathbf{O}(p \wedge q) \wedge \mathbf{O}\neg(p \wedge q)) \vee (\mathbf{O}p \wedge \mathbf{O}\neg p) \vee (\mathbf{O}q \wedge \mathbf{O}\neg q) \tag{14}$$

$$(\mathbf{O}(p \wedge r) \wedge \mathbf{O}\neg(p \wedge r)) \vee (\mathbf{O}p \wedge \mathbf{O}\neg p) \vee (\mathbf{O}r \wedge \mathbf{O}\neg r) \tag{15}$$

$$(\mathbf{O}(p \wedge \neg r) \wedge \mathbf{O}\neg(p \wedge \neg r)) \vee (\mathbf{O}p \wedge \mathbf{O}\neg p) \vee (\mathbf{O}\neg r \wedge \mathbf{O}\neg\neg r) \vee (\mathbf{O}r \wedge \mathbf{O}\neg r) \tag{16}$$

The inference made at line 4 is legitimate in view of the $\mathbf{LUM.a}$ -valid inference

$$\mathbf{O}(p \wedge q) \vdash \mathbf{O}p \vee \sharp(p \wedge q) \tag{17}$$

Since $\sharp(p \wedge r)$ and $\sharp(p \wedge \neg r)$ are $\mathbf{LUM.a}$ -derivable from the premises $\mathbf{O}r$ and $\mathbf{O}\neg r$, the **Dab**-formula derived at line 5 of the proof is not minimal at stage 7. Consequently, line 4 is unmarked at this stage. As opposed to $\mathbf{ADPM.1}^r$ and $\mathbf{ADPM.1}^m$, the logics $\mathbf{LUM.a}^r$ and $\mathbf{LUM.a}^m$ lead to the following desirable outcome:

$$\mathbf{O}(p \wedge q), \mathbf{O}r, \mathbf{O}\neg r \vdash_{\mathbf{LUM.a}^r} \mathbf{O}p \tag{18}$$

$$\mathbf{O}(p \wedge q), \mathbf{O}r, \mathbf{O}\neg r \vdash_{\mathbf{LUM.a}^m} \mathbf{O}p \tag{19}$$

³⁸Our expression $\sharp(A)$ is equivalent to the negation of Goble's expression $\mathcal{U}(A)$ in [43; 42]. Note that \sharp is not a (modal or other) operator but just a symbol that allows us to abbreviate a formula.

³⁹Goble uses the name \mathbf{ALUM}^r for the logic that we call $\mathbf{LUM.a}^r$.

5.2 Evaluating the logics

Explosion principles The adaptive logics based on the **LUM**-family are conflict-tolerant to the same extent as their respective lower limit logics. This means, for a start, that (DEX) is invalid in all of them. Since they are **CL**-based and in view of the interdefinability of **O** and **P**, they also accommodate conflicts of the form $OA \wedge \neg PA$, which simply reduce to conflicts between obligations.

However, the logics do not tolerate the other types of deontic conflicts that were discussed in Section 4.2. While $O(A \wedge \neg A)$ is consistent in **LUM.a** – and hence also in **LUM.a^x**, it is inconsistent in each of **LUM.b** and **LUM.c** in view of the (P)-axiom. It follows that ALs based on the latter two logics cannot make sense of self-contradictory obligations. Also, all the (adaptive) **LUM**-logics trivialize conflicts of the form $OA \wedge P\neg A$, as these reduce to plain contradictions in view of (Def_P) and (RE). Finally, $P(A \wedge \neg A)$ (which is equivalent to $\neg O(\neg A \vee A)$) is also trivial in these logics, in view of the necessitation rule (N).

Benchmark examples The Smith and Jones arguments are **LUM.a^x**-valid. The premises of these arguments are **SDL**-consistent and hence normal, which means that (by Theorem 3.14), the adaptive logics are just as strong as **SDL** for these cases.⁴⁰ The Roberts and Thomas arguments are not valid in **LUM.a^r** or **LUM.a^m**. Here is a proof illustrating why the Roberts arguments are not valid in **LUM.a^x**:

1	$O(t \wedge r)$	Prem	\emptyset
2	$O(\neg t \wedge v)$	Prem	\emptyset
3	Or	1; RC	$\{\#(t \wedge r)\} \checkmark$
4	Ov	2; RC	$\{\#(\neg t \wedge v)\} \checkmark$
5	$O(r \wedge v)$	3,4; RU	$\{\#(t \wedge r), \#(\neg t \wedge v)\} \checkmark$
6	$\#(t \wedge r) \vee \#(\neg t \wedge v)$	1,2; RU	\emptyset

In order to infer Or and Ov via RC we need to rely on the falsity of $\#(t \wedge r)$ and $\#(\neg t \wedge v)$. However, further inspection of the premises teaches us that the disjunction of these abnormalities is **LUM.a**-derivable from the premises. To see why, note that this disjunction is **LUM.a**-equivalent to the following formula, which

⁴⁰Goble showed that the upper limit logic of **LUM.a^x** is **SDL**, see [43, Observation 4.1].

is a **LUM.a**-consequence of the premises:⁴¹

$$\begin{aligned} & (\mathbf{O}(t \wedge r) \wedge \mathbf{O}\neg(t \wedge r)) \vee (\mathbf{O}(\neg t \wedge v) \wedge \mathbf{O}\neg(\neg t \wedge v)) \vee \\ & (\mathbf{O}t \wedge \mathbf{O}\neg t) \vee (\mathbf{O}r \wedge \mathbf{O}\neg r) \vee (\mathbf{O}v \wedge \mathbf{O}\neg v) \end{aligned} \quad (20)$$

The minimal **Dab**-formula derived at line 6 blocks the derivation of the formulas derived at lines 3-5, causing the invalidity of the Roberts arguments. The same mechanism blocks the derivation of the conclusion of the Thomas argument.⁴²

The Natascha argument, version 1, is **LUM.a**-valid (and hence **LUM.a^x**-valid), but only because its premise set is **LUM.a**-trivial: from premises (i) and (ii) we can derive the negation of premise (iii) by (Agg). In contrast, the second version of the Natascha argument is **LUM.a**-satisfiable. Here is an **LUM.a^m**-proof for this argument:

1	$\mathbf{O}s$	Prem	\emptyset
2	$\mathbf{O}m$	Prem	\emptyset
3	$\mathbf{O}\neg(s \wedge m)$	Prem	\emptyset
4	$\mathbf{O}(s \supset t)$	Prem	\emptyset
5	$\mathbf{O}(m \supset t)$	Prem	\emptyset
6	$\mathbf{O}(s \wedge t)$	1, 4; RU	\emptyset
7	$\mathbf{O}(m \wedge t)$	2, 5; RU	\emptyset
8	$\mathbf{O}t$	6; RC	$\{\sharp(s \wedge t)\}$
9	$\mathbf{O}t$	7; RC	$\{\sharp(m \wedge t)\}$
10	$\sharp(s \wedge t) \vee \sharp(m \wedge t)$	1-3; RU	\emptyset

The formulas derived at lines 6 and 7 are **LUM.a**-derivable from the premises via applications of (Agg) and (RE). From each of these formulas we can derive $\mathbf{O}t$ via RC. Since we are working with the minimal abnormality strategy, lines 8 and 9 are unmarked at stage 10. If we were to use reliability, however, both lines would be marked. Indeed, the modified Natascha argument is valid for **LUM.a^m**, while invalid for **LUM.a^r**:

$$\mathbf{O}s, \mathbf{O}m, \mathbf{O}\neg(s \wedge m), \mathbf{O}(s \supset t), \mathbf{O}(m \supset t) \not\vdash_{\mathbf{LUM.a}^r} \mathbf{O}t \quad (21)$$

$$\mathbf{O}s, \mathbf{O}m, \mathbf{O}\neg(s \wedge m), \mathbf{O}(s \supset t), \mathbf{O}(m \supset t) \vdash_{\mathbf{LUM.a}^m} \mathbf{O}t \quad (22)$$

⁴¹By **CL**, $(\mathbf{O}(t \wedge r) \wedge \mathbf{O}\neg(t \wedge r)) \vee \neg(\mathbf{O}(t \wedge r) \wedge \mathbf{O}\neg(t \wedge r))$. Since $\mathbf{O}(t \wedge r)$, $\neg(\mathbf{O}(t \wedge r) \wedge \mathbf{O}\neg(t \wedge r))$ entails $\mathbf{O}t$ by (RUM). Analogously, by **CL**, $(\mathbf{O}(\neg t \wedge v) \wedge \mathbf{O}\neg(\neg t \wedge v)) \vee \neg(\mathbf{O}(\neg t \wedge v) \wedge \mathbf{O}\neg(\neg t \wedge v))$. Since $\mathbf{O}(\neg t \wedge v)$, $\neg(\mathbf{O}(\neg t \wedge v) \wedge \mathbf{O}\neg(\neg t \wedge v))$ entails $\mathbf{O}\neg t$ by (RUM). Altogether, by **CL**, $(\mathbf{O}(t \wedge r) \wedge \mathbf{O}\neg(t \wedge r)) \vee (\mathbf{O}(\neg t \wedge v) \wedge \mathbf{O}\neg(\neg t \wedge v)) \vee (\mathbf{O}t \wedge \mathbf{O}\neg t)$. By **CL** again, (20) follows.

⁴²In the Thomas case, the culpable **Dab**-formula is the disjunction $\sharp(t \wedge (f \vee s)) \vee \sharp(\neg t \wedge \neg f)$. We leave the verification to the interested reader.

The behavior of the ALs based on **LUM.b** and **LUM.c** is roughly analogous to the preceding case, with one exception. The premises in version 1 of the Natascha argument are inconsistent in **LUM.a** and **LUM.b**, but consistent in **LUM.c**. That is, we cannot aggregate premises (i) and (ii) of this argument, in the absence of the permission statements P_s and P_m . Parallel to the situation for the modified Natascha argument in **LUM.a^x**, we obtain the conclusion O_t with **LUM.c^m** for the original Natascha argument, while we do not obtain it with **LUM.c^r**. The following proof illustrates that O_t is **LUM.c^m**-derivable:⁴³

1	O_s	Prem	\emptyset
2	O_m	Prem	\emptyset
3	$\neg O(s \wedge m)$	Prem	\emptyset
4	$O(s \supset t)$	Prem	\emptyset
5	$O(m \supset t)$	Prem	\emptyset
6	$O(s \wedge m)$	1,2; RC	$\{\#s, \#m\}\checkmark^{12}$
7	$O(s \wedge t)$	1, 4; RU	$\{\#s, \#(s \supset t)\}\checkmark^{12}$
8	$O(m \wedge t)$	2, 5; RU	$\{\#m, \#(m \supset t)\}\checkmark^{12}$
9	O_t	7; RC	$\{\#s, \#(s \supset t), \#(s \wedge t)\}$
10	O_t	8; RC	$\{\#m, \#(m \supset t), \#(m \wedge t)\}$
11	$\#(s \wedge t) \vee \#(m \wedge t)$	1-3; RU	\emptyset
12	$\#s \vee \#m$	1-3; RU	\emptyset
13	$\#s \vee \#(m \wedge t)$	12; RU	\emptyset
14	$\#(s \wedge t) \vee \#m$	12; RU	\emptyset

The inferences at lines 13 and 14 hold in view of the **CL**-validity of $\#s \supset \#(s \wedge t)$ and $\#m \supset \#(m \wedge t)$ respectively. Where $\Gamma_n = \{O_s, O_m, \neg O(s \wedge m), O(s \supset t), O(m \supset t)\}$:

$$\Phi_{14}(\Gamma_n) = \{\{\#(s \wedge t), \#s\}, \{\#(m \wedge t), \#m\}\} \quad (23)$$

It is easily verified that, in view of Definition 3.2, lines 9 and 10 are unmarked. If we were to use the reliability strategy instead, then by Definition 3.1 these lines would be marked in the proof above.

$$\Gamma_n \not\vdash_{\mathbf{LUM.c}^r} O_t \quad (24)$$

$$\Gamma_n \vdash_{\mathbf{LUM.c}^m} O_t \quad (25)$$

The formula O_t is a floating conclusion with respect to Γ_n . As pointed out in Section 4, it is a matter of debate whether or not floating conclusions are acceptable.

⁴³Lines 7 and 8 can be derived by means of (P-Agg). Note that this rule requires that the two formulas to be aggregated are themselves unconflicted. Hence we need RC to make these two derivations.

We do not add anything to this debate here. It suffices for us to point out that each stance can be formally represented within the AL framework.

5.3 Further reading and open ends

The **LUM**-systems were introduced by Goble in [35; 41; 36], where they were called ‘logics of permitted distribution’ or **DPM**. They were called ‘logics of unconflicted distribution’ or **LUM** in [42; 43]. Adaptive extensions of these systems were presented in [94; 87; 43]. Moreover, in [91], dyadic variants of the **LUM**-systems were also strengthened within the AL framework (see also Section 9.1 below).

There are many other types of deontic logics which invalidate (Inh). First, there is the general class of classical modal logics of which the logic **E** (cf. supra) is but one example. Second, Goble [34; 37] developed a very rich semantics for deontic logics, based on an idea from [54]. On this semantics, OA is true iff the closest A -worlds are all better than the closest $\neg A$ -worlds. Third and last, in more recent work, Cariani [27] proposed yet another semantics for “ought” which invalidates (Inh) in a principled way – see also [100; 101] for a formal investigation into this proposal. For each of these types of logics, one can ask whether it makes sense to strengthen them adaptively, and if so, which technical difficulties arise and what behavior the resulting logics will display. In particular, it would be interesting to learn whether some such variants perform better than the currently available logics, in dealing with the Roberts arguments and the Thomas argument.

6 Adaptive aggregation

A popular way to accommodate deontic conflicts in a formal system is by rejecting the aggregation principle (Agg), and with it the normality schema (K). In its simplest form, this proposal gives us the deontic logic **P**.⁴⁴ We will focus on two relatively basic ALs obtained from **P** in this section.

6.1 Adaptive aggregation: a basic example

Rejecting aggregation The language of **P** is the same as that of **SDL**, with **P** defined as the dual of **O**. As before, we will not consider nested occurrences of **O**. **P** is axiomatized by adding the axiom (D) to **CL** and closing the resulting set under modus ponens (MP), the necessitation rule (N), and the rule of inheritance (Inh). Each of the following are facts about the derivability relation of **P**:

⁴⁴Again, we follow Goble’s nomenclature. See the end of this section for pointers to the literature on this and related logics.

$$\vdash O(p \vee \neg p) \quad (26)$$

$$Op \vdash O(p \vee q) \quad (27)$$

$$O(p \wedge q) \vdash Op, Oq \quad (28)$$

$$Op, Oq \not\vdash O(p \wedge q) \quad (29)$$

$$O(p \wedge (\neg p \vee q)) \vdash Oq \quad (30)$$

$$Op, O(\neg p \vee q) \not\vdash Oq \quad (31)$$

In view of (Inh), Replacement of (Classical) Equivalents (RE) is valid in \mathbf{P} . So in Chellas' terms, \mathbf{P} is a non-normal but classical modal logic [29].

One way to motivate and understand the rejection of (Agg) in \mathbf{P} is in terms of multiple normative standards that ground our obligations, where OA is unspecific about the normative standard that grounds the obligation that A . Under such a reading, OA and OB may well be true even if there is no single standard that grounds the conjunction of both obligations, and hence $O(A \wedge B)$ can still fail.⁴⁵ For instance, varying on our Smith example, one's duty to fight in the army might be based on the laws of one's country, whereas one's personal pacifist ethics grounds the claim that one ought not to fight in the army. Still, it does not follow that one ought to do the logically impossible, viz. to fight in the army and not fight in the army.

A semantics for \mathbf{P} is obtained from the **SDL**-semantics (cf. Section 2) by generalizing the notion of an accessibility relation R . \mathbf{P} -models are then of the type $\langle W, w_0, \mathcal{R}, V \rangle$, where W , w_0 , and V are as before, but \mathcal{R} is a non-empty set of serial accessibility relations, rather than a single such relation. The semantic clause for O then reads as follows:

$$(SC-O) \quad M, w \models OA \text{ iff there is an } R \in \mathcal{R} \text{ such that } M, w' \models A \text{ for all } w' \text{ such that } Rww'$$

In other words, the single normative standard from **SDL** is replaced with a set of such standards, and we quantify (existentially) over such standards in order to determine the truth of OA . It is well-known that \mathbf{P} is sound and complete with respect to this semantics – see [38, Theorem 1]. Other semantics can also be given for \mathbf{P} . We refer the reader to [42, pp. 300-301] for an overview of these.

⁴⁵The idea that one can relativize deontic logic to a given “moral code”, and that what is obligatory under one such code may not be obligatory (or even forbidden) under another, is at least as old as Von Wright's *Deontic Logic* – see [109, p. 15]. The difference here is that in \mathbf{P} , the code that is at stake remains implicit, and OA only means that A is obligatory under at least *some* moral code.

Going adaptive Even if aggregation is invalid on the reading of **O** just presented, in practice we do often aggregate our obligations. One simple way to argue for this is by referring to the benchmark examples from Section 4. It can easily be verified that neither the Smith argument nor the second variant of the Roberts argument is valid in **P**.

More generally, it is one thing to say that we take into account various normative standards and treat them as independent grounds or reasons when trying to determine what our obligations are. It is quite another thing to argue that none of these obligations can themselves be aggregated when doing so; this seems to go against much of our intuition.⁴⁶ For instance, when deciding how to get to the office in the morning, I may apply norms concerning the environment, norms uttered by my boss, and norms concerning my own safety and that of others. There seems to be no *prima facie* reason why we cannot integrate these various norms when settling for a single way to get to the office – e.g. I may conclude that I ought to bike to the office, since that way I will be in time for a meeting without causing air-pollution. The presence of deontic conflicts in itself seems insufficient to warrant a full rejection of aggregation, and, as we will show below, there is no logical reason for doing so either.

One needs to be careful here though. We cannot just add (Agg) to **P**, as this would give us again full **SDL** and hence deontic explosion in the face of deontic conflicts.⁴⁷ Moreover, as shown in [41, Sect. 2], there is no obvious conditional variant of (Agg) that can do a similar job, without in turn yielding some variant of deontic explosion.⁴⁸ So some obligations can be aggregated, but not all. As we will show in the remainder of this section, going adaptive allows us to steer a middle course between the weakness of **P** and deontic explosion.

The logics \mathbf{P}^x The most straightforward way one might strengthen **P** adaptively, is by treating all formulas of the form $\mathbf{O}A \wedge \mathbf{O}B \wedge \neg\mathbf{O}(A \wedge B)$ as abnormalities. However, just as in the case of **ADPM1^r**, this will give us a flip-flop. To see why, consider $\Gamma = \{\mathbf{O}p, \mathbf{O}\neg p, \mathbf{O}q, \mathbf{O}r\}$. Intuitively speaking, there is no problem with q and r in this example, and hence we expect $\mathbf{O}(q \wedge r)$ to be derivable. Such an inference can indeed be made within a proof of the adaptive logic thus defined.

⁴⁶Compare [42, p. 253]: “Even if what one ought to do is often determined by different sources or authorities, insofar as propositions of what one ought to do serve as guides to action or as standards of evaluation of an agent’s overall actions, there must be a common ought derived from those separate sources”.

⁴⁷We safely leave it to the reader to check that adding (Agg) to **P** yields full **SDL**.

⁴⁸See also [42, Section 5.2]. In particular, Goble shows that adding the axiom (C-Agg) (cf. Section 5.1) to **P** will result in a variant of deontic explosion.

However, we can derive a disjunction of abnormalities (in that adaptive logic) from Γ which will block the derivation. This **Dab**-formula is a disjunction of the following three formulas:

$$\mathbf{O}q \wedge \mathbf{O}r \wedge \neg\mathbf{O}(q \wedge r) \quad (32)$$

$$\mathbf{O}(p \vee \neg(q \wedge r)) \wedge \mathbf{O}\neg p \wedge \neg\mathbf{O}((p \vee \neg(q \wedge r)) \wedge \neg p) \quad (33)$$

$$\mathbf{O}(q \wedge r) \wedge \mathbf{O}\neg(q \wedge r) \wedge \neg\mathbf{O}((q \wedge r) \wedge \neg(q \wedge r)) \quad (34)$$

Suppose that (32) is false but the premises are true. Then $\mathbf{O}(q \wedge r)$ is the case. Likewise, since $\mathbf{O}(p \vee \neg(q \wedge r))$ follows by (Inh) from $\mathbf{O}p$, (33) can only be false (in view of the premises) if its last conjunct is false, and hence $\mathbf{O}\neg(q \wedge r)$ is true. But then the third abnormality, (34) must be true.

It is not hard to see where the problem could be in cases like this. That is, since $\mathbf{O}p, \mathbf{O}\neg p \in \Gamma$, we should not use these obligations – nor weakenings of them – in order to apply aggregation. In other words, obligations that are themselves conflicted, or subformulas of which are conflicted, should be treated as abnormal.

This brings us to a slightly more complicated set of abnormalities, which is due to Goble [43]. As before, let $\sharp(A)$ denote the disjunction of all formulas $\mathbf{O}B \wedge \mathbf{O}\neg B$, where $B \in S(A)$ (B is a subformula of A). Let $\natural(A, B) = (\mathbf{O}A \wedge \mathbf{O}B \wedge \neg\mathbf{O}(A \wedge B)) \vee \sharp(A \wedge B)$. We now define

$$\Omega_{\mathbf{P}} = \{\natural(A, B) \mid A, B \in \mathcal{W}\}$$

In other words, we have an abnormality with respect to A and B iff they are both obligatory and their conjunction is not obligatory, *or* a proper subformula of them is conflicted. This means that as soon as e.g. $\mathbf{O}p, \mathbf{O}\neg p$ holds, all abnormalities $\natural(A, B)$ with $p \in S(A)$ are true. Under this definition, none of the formulas (32)-(34) are abnormalities. The corresponding disjunction of $\Omega_{\mathbf{P}}$ -abnormalities

$$\natural(q, r) \vee \natural(p \vee \neg(q \wedge r), \neg p) \vee \natural(q \wedge r, \neg(q \wedge r)) \quad (35)$$

is not a minimal **Dab**-consequence of Γ , since $\natural(p \vee \neg(q \wedge r), \neg p)$ alone follows from Γ .

Let the logics $\mathbf{P}^{\mathbf{r}}$ and $\mathbf{P}^{\mathbf{m}}$ be the adaptive logics defined by the triple $\langle \mathbf{P}, \Omega_{\mathbf{P}}, \times \rangle$, where $\times \in \{\mathbf{r}, \mathbf{m}\}$.⁴⁹ It can easily be checked that the upper limit logic of $\mathbf{P}^{\mathbf{r}}$ and $\mathbf{P}^{\mathbf{m}}$ is just **SDL**: adding the negation of all members of $\Omega_{\mathbf{P}}$ as axioms to \mathbf{P} , is equivalent

⁴⁹In Goble's work, the first of these two logics is known as **AP^r**. As before, we skip the initial "A" since the superscript suffices to mark the difference with the monotonic logic \mathbf{P} .

to adding (Agg) to \mathbf{P} .⁵⁰ This means that normal premise sets in the logics \mathbf{P}^x are just **SDL**-consistent premise sets (where ‘normal’ is understood in the technical sense specified on page 23). Hence by Theorem 3.14, whenever a premise set is **SDL**-consistent, its \mathbf{P}^x -consequence set will be identical to its **SDL**-consequence set:

Theorem 6.1. *If Γ is **SDL**-consistent, then $Cn_{\mathbf{P}^r}(\Gamma) = Cn_{\mathbf{P}^m}(\Gamma) = Cn_{\mathbf{SDL}}(\Gamma)$.*

6.2 Evaluating the logics

Explosion principles The logic \mathbf{P} , and with it \mathbf{P}^r and \mathbf{P}^m , clearly accommodates conflicts of the basic type $OA, O\neg A$. By (RE), (Def_P) and **CL**-properties, also conflicts of the type $OA, \neg PA$ are consistent in \mathbf{P} and its adaptive extensions.

All other types of deontic conflicts listed in Section 4.2 will be trivialized within these logics. The reasons are similar to those for **LUM.b** and **LUM.c**: $O(A \wedge \neg A)$ is contradictory in view of (P), $OA \wedge P\neg A$ is contradictory in view of (Def_P) and (RE), and $P(A \wedge \neg A)$ is false in view of (N) and (Def_P). So the simplicity of \mathbf{P} comes at an important price, viz. that it can only handle conflicting obligations and does not allow us to reason about conflicting information concerning (obligations and) permissions.⁵¹

Benchmark examples The arguments for Jones and Roberts 1 are valid in both \mathbf{P}^r and \mathbf{P}^m . This is easy to verify since the arguments are already valid in \mathbf{P} in view of its validating (Inh), and since both \mathbf{P}^r and \mathbf{P}^m are extensions of \mathbf{P} . The premises of the Smith argument are normal: no **Dab**-formula can be derived from them. As a result, we can aggregate the obligations in the argument and derive O_s .

The second Roberts argument is also valid in \mathbf{P}^x , but here the reasoning is slightly more intricate. First, applying (Inh), we can derive Or and Ov from the premises. To apply aggregation to these two formulas, we need to assume that neither r nor v are conflicted, given the premise set. This is clearly the case: the only conflict that follows from the premises, is $Ot, O\neg t$. The following \mathbf{P}^r -proof illustrates how we can obtain the desired conclusion for Roberts 2, while avoiding the aggregation of conflicted obligations:

⁵⁰To see why this is so, note first that if we negate all formulas of the form $\mathfrak{h}(A, B)$, then *a fortiori* we negate all formulas of the form $OA \wedge OB \wedge \neg O(A \wedge B)$, and hence we affirm all instances of (Agg). In addition, we also negate all formulas of the form $OA \wedge O\neg A$, but these are anyway **SDL**-valid.

⁵¹Note also that simply rejecting (P) will not allow us to have a satisfactory account of conflicts of the type $O(A \wedge \neg A)$: due to (Inh) these conflicts will still lead to deontic explosion.

1	$O(t \wedge r)$	Prem	\emptyset
2	$O(\neg t \wedge v)$	Prem	\emptyset
3	Or	1; RU	\emptyset
4	Ov	2; RU	\emptyset
5	$O(r \wedge v)$	4,5; RC	$\{\natural(r, v)\}$
6	Ot	1; RU	\emptyset
7	$O\neg t$	2; RU	\emptyset
8	$O(t \wedge v)$	6,4; RU	$\{\natural(t, v)\}\checkmark^{11}$
9	$O(t \wedge \neg t)$	6,7; RC	$\{\natural(t, \neg t)\}\checkmark^{10}$
10	$\natural(t, \neg t)$	6,7; RU	\emptyset
11	$\natural(t, v)$	6,7; RU	\emptyset

Since $\natural(t, v)$ follows from the premises, we cannot finally derive $O(t \wedge v)$ from them. So even if there is no direct conflict between t and v , the fact that t is itself conflicted is sufficient to block its aggregation with other (unproblematic) obligations.⁵²

The reasoning for the Thomas argument is wholly analogous to the second Roberts case, with the difference that we apply (Inh) once more after aggregating $O(f \vee s)$ and $O\neg f$ to $O((f \vee s) \wedge \neg s)$. This gives us the desired conclusion Os .

For the Natascha arguments, it turns out that with the \mathbf{P} -based adaptive logics the strategies make no difference. The point is that, although we can obviously not apply aggregation to Os and Om , we can still aggregate Os and $O(s \supset t)$ (and likewise, Om and $O(m \supset t)$). The fact that the pair (m, s) behaves abnormally ($\natural(m, s)$ follows from the premises of the argument) does not imply that either of $(s, s \supset t)$ or $(m, m \supset t)$ behave abnormally. Hence we can finally derive Ot on two different conditions in both \mathbf{P}^r and \mathbf{P}^m . We illustrate this for the first variant of the Natascha argument:

⁵²As pointed out by Goble, allowing aggregation for all A, B such that $A \wedge B$ is consistent is simply a no-go in the context of \mathbf{P} , since it will lead to another form of deontic explosion. See [41, Sect. 2.4.1].

1	Os	Prem	\emptyset
2	Om	Prem	\emptyset
3	$\neg O(s \wedge m)$	Prem	\emptyset
4	$O(s \supset t)$	Prem	\emptyset
5	$O(m \supset t)$	Prem	\emptyset
6	$O(s \wedge (s \supset t))$	1,4; RC	$\{\natural(s, s \supset t)\}$
7	$O(m \wedge (m \supset t))$	2, 5; RC	$\{\natural(m, m \supset t)\}$
8	Ot	6; RU	$\{\natural(s, s \supset t)\}$
9	Ot	7; RU	$\{\natural(m, m \supset t)\}$

For none of the variants of the Natascha arguments, the disjunction of abnormalities $(Os \wedge O\neg s) \vee (Om \wedge O\neg m)$ is \mathbf{P} -derivable from the premises. Nor is there another **Dab**-formula which prevents lines 8 and 9 from being finally derivable. So, to sum up, all inferences from our benchmark examples are valid in the logics \mathbf{P}^r and \mathbf{P}^m .

This is not to say that there is no difference between \mathbf{P}^r and \mathbf{P}^m . Consider e.g. $\Gamma = \{Op, Oq, Or, \neg O(p \wedge r) \vee \neg O(q \wedge r)\}$. From this premise set, \mathbf{P}^r will not allow us to finally derive $O(p \wedge r) \vee O(q \wedge r)$, whereas \mathbf{P}^m will. To understand this, note that $\natural(p, r) \vee \natural(q, r)$ is a minimal **Dab**-consequence of Γ , whence both abnormalities are unreliable in view of Γ . However, since nothing prevents us to assume that either the first or the second abnormality is *false*, using *minimal abnormality* we can derive $O(p \wedge r) \vee O(q \wedge r)$.

6.3 Further reading and open ends

Bernard Williams [111] was the first to advocate a rejection of (Agg) on philosophical grounds; Marcus [62] is another important proponent of such a rejection. More formally worked out proposals can be found in [107; 29; 84]. Later, Goble developed the semantics and metatheory of \mathbf{P} and variants of it in detail – see in particular [38; 40; 39]. For a more complete overview of the literature on \mathbf{P} and close (monotonic) relatives, we refer to [42, Section 5.2].

The first adaptive logic that applies the idea of “adaptive aggregation” was published in [66], and later reworked in [67]. These logics are however based on a richer lower limit logic, viz. the logic $\mathbf{SDL}_a\mathbf{P}_e$ from [38]. In this system, one can express both an “existential” notion of obligation O_e (whose logic is \mathbf{P}) and a “universal” notion of obligation O_a , whose logic is \mathbf{SDL} . The two modalities are connected by the following bridging principle:

$$(B) \quad O_a(A \supset B) \wedge O_e A \vdash O_e B$$

which entails i.a. that every universal obligation is also an existential obligation, $O_a A \supset O_e A$. Alternatively, one can interpret the logics in terms of our distinction between *prima facie* obligations and actual obligations (cf. Section 3.1).

Adaptive logics that are based on \mathbf{P} itself are discussed in [43]; here we only discussed the second of the two. The other AL discussed by Goble appears to be slightly weaker. For instance, in this logic, the Natascha argument is only valid if we use *minimal abnormality*. More generally, in this logic any conflict of the type $OA \wedge OB \wedge \neg O(A \wedge B)$ “infects” all the subformulas of A and B . We leave the full inspection and proof of this claim for another occasion.

An interesting issue concerns the enrichment of the aforementioned ALs with operators that allow one to express (technical, physical, practical) impossibility at the object level. Indeed, in Williams’ famous essay, he argues that purely logical conflicts between oughts are only a special case of a much more common type of conflicts, viz. conflicts between two obligations whose joint fulfillment is impossible for *contingent* reasons – e.g. because of the particular physical situation we find ourselves in [111]. This raises a number of questions concerning the interplay between alethic and deontic modalities, which would take us well beyond the scope of the present paper – see however [15, Chapter 4] for a first attempt to combine alethic and deontic modalities.

7 Inconsistency-adaptive deontic logics

As noted in Section 3.5, the first adaptive logics were *inconsistency-adaptive*. These logics are members of the larger family of *paraconsistent logics*, i.e. logics which invalidate (ECQ).

Note that (ECQ) bears close affinity to (DEX). To obtain the latter from the former we only need to prefix the formulas involved with an O -operator. Besides the approaches we saw in Sections 5 and 6, a third natural way to invalidate (DEX) is by invalidating (ECQ).

Going paraconsistent has a couple of additional benefits in the context of deontic logic. A first is that it allows us to preserve the interdefinability of O and P , while invalidating (DEX- $OP\neg$). Assuming the interdefinability of O and P , the formula $OA \wedge P\neg A$ is equivalent to the contradictions $OA \wedge \neg OA$ and $\neg P\neg A \wedge P\neg A$. By (ECQ), these contradictions entail everything. To prevent such explosive behavior, it suffices to invalidate (ECQ).

A second advantage is that only a paraconsistent deontic logic can invalidate the explosion principles (DEX- $O\neg O$) and (DEX- $P\neg P$), for the obvious reason that these

principles are instances of (ECQ):

$$OA, \neg OA \vdash OB \quad (\text{DEX-O}\neg\text{O})$$

$$PA, \neg PA \vdash OB \quad (\text{DEX-P}\neg\text{P})$$

There are independent reasons as to why, in some contexts, we may want to tolerate *contradictory norms*, i.e. formulas of the form $OA \wedge \neg OA$ or $PA \wedge \neg PA$. Priest, for instance, gives the following example. Suppose that, in some country, women are not permitted to vote, while property holders are permitted to vote. Suppose further that, perhaps due to a recent revision of the property law, women are permitted to hold property. Then female property holders are both permitted and not permitted to vote ($Pv \wedge \neg Pv$) [81, pp. 184–185].

In this section, we present inconsistency-adaptive deontic logics. We will work stepwise, starting with the paraconsistent logic **CLuN**, its deontic extension **DCLuN**, and adaptive strengthenings **DCLuN^x** (Section 7.1). After that, we will consider several variants of **DCLuN** and their associated adaptive logics (sections 7.2 and 7.3).

7.1 Paraconsistent adaptive deontic logic

A paraconsistent core logic We use the paraconsistent logic **CLuN** as our starting point. **CLuN** is an acronym for ‘Classical Logic with gluts for Negation’. A truth-value *glut* for negation relative to a formula A occurs when both A and its negation are true; **CLuN** allows such gluts whereas **CL** disallows them. The deontic logics to be presented in this section are extensions of **CLuN**, but they are defined so that plenty of other paraconsistent logics may replace **CLuN** as their core logic. In Sections 7.2 and 7.3 we will mention some alternatives.

The set \mathcal{W}^\sim of well-formed **CLuN**-formulas is the following:

$$\mathcal{W}^\sim := \mathcal{S} \mid \sim \langle \mathcal{W}^\sim \rangle \mid \neg \langle \mathcal{W}^\sim \rangle \mid \langle \mathcal{W}^\sim \rangle \vee \langle \mathcal{W}^\sim \rangle \mid \langle \mathcal{W}^\sim \rangle \wedge \langle \mathcal{W}^\sim \rangle \mid \langle \mathcal{W}^\sim \rangle \supset \langle \mathcal{W}^\sim \rangle \mid \langle \mathcal{W}^\sim \rangle \equiv \langle \mathcal{W}^\sim \rangle$$

In the remainder, we will stick to \neg as the connective denoting classical negation. Beside \neg , \mathcal{W}^\sim contains the connective \sim which we will use as our paraconsistent negation sign. In fact, \sim is the only **CLuN**-connective which behaves differently from the classical connectives. We obtain **CLuN** by adding the following axiom schema to **CL**:

$$A \vee \sim A \quad (\text{EM}\sim)$$

We write $\Gamma \vdash_{\text{CLuN}} A$ to denote that A is **CLuN**-derivable from Γ .

The **CLuN**-semantics is defined as follows. To obtain a **CLuN**-model M , we extend the assignment function v_a of **CL** so that it assigns truth values not only to schematic letters, but also to formulas of the form $\sim A$, i.e. $v_a : \mathcal{S} \cup \{\sim A \mid A \in \mathcal{W}^\sim\} \rightarrow \{0, 1\}$. Next, we extend v_a to a valuation function v as follows:

(SC1) For formulas $A \in \mathcal{S} \cup \{\sim A \mid A \in \mathcal{W}^\sim\} : M \models A$ iff $v_a(A) = 1$.

(SC2) For $\neg, \vee, \wedge, \supset, \equiv$, the semantic clauses for **CLuN** are those of **CL**.

Finally, in order to validate the axiom (EM \sim), we require that all **CLuN**-models satisfy the following condition: for all $A \in \mathcal{W}^\sim$, $M \models A$ or $M \models \sim A$. A semantic consequence relation for **CLuN** is defined as follows: $\Gamma \Vdash_{\mathbf{CLuN}} A$ iff for all **CLuN**-models M : if $M \models B$ for all $B \in \Gamma$, then $M \models A$.

Before we move on to deontic extensions of **CLuN**, we point out a number of relevant properties of this logic for ease of reference:

(i) **CLuN** is paraconsistent, but not *paracomplete*: while (ECQ) is **CLuN**-invalid for \sim , the excluded middle principle (EM \sim) is **CLuN**-valid.

(ii) In contrast to well-known paraconsistent logics such as Priest's **LP**, **CLuN** validates modus ponens:

$$A, A \supset B \vdash B \quad (\text{MP})$$

Note that $A \supset B$ and $\sim A \vee B$ are not **CLuN**-equivalent: if $v(A) = v(\sim A) = v(\sim B) = 1$ and $v(B) = 0$, then $v(A \supset B) = 0$ while $v(\sim A \vee B) = 1$.

(iii) De Morgan's laws and the double negation laws are invalid for \sim in **CLuN**. This means that complex contradictions are not reducible to contradictions between elementary letters:

$$(p \wedge q) \wedge \sim(p \wedge q) \not\vdash (p \wedge \sim p) \vee (q \wedge \sim q) \quad (36)$$

$$(p \vee q) \wedge \sim(p \vee q) \not\vdash (p \wedge \sim p) \vee (q \wedge \sim q) \quad (37)$$

$$(p \supset q) \wedge \sim(p \supset q) \not\vdash (p \wedge \sim p) \vee (q \wedge \sim q) \quad (38)$$

$$\sim\sim(p \wedge \sim p) \not\vdash p \wedge \sim p \quad (39)$$

(iv) Contraposition, modus tollens, and disjunctive syllogism are invalid for \sim in **CLuN**:

$$A \supset B \not\vdash \sim B \supset \sim A \quad (40)$$

$$A \supset B, \sim B \not\vdash \sim A \quad (41)$$

$$A \vee B, \sim A \not\vdash B \quad (42)$$

A paraconsistent deontic logic A technically straightforward way to construct a deontic logic on the basis of **CLuN** is the following. First, we extend the language \mathcal{W}^\sim with the deontic operator \mathbf{O} , preventing nested occurrences of the deontic operator:

$$\mathcal{W}_\mathbf{O}^\sim := \mathcal{W}^\sim \mid \mathbf{O}\langle \mathcal{W}^\sim \rangle \mid \sim\langle \mathcal{W}_\mathbf{O}^\sim \rangle \mid \neg\langle \mathcal{W}_\mathbf{O}^\sim \rangle \mid \langle \mathcal{W}_\mathbf{O}^\sim \rangle \vee \langle \mathcal{W}_\mathbf{O}^\sim \rangle \mid \langle \mathcal{W}_\mathbf{O}^\sim \rangle \wedge \langle \mathcal{W}_\mathbf{O}^\sim \rangle \mid \langle \mathcal{W}_\mathbf{O}^\sim \rangle \supset \langle \mathcal{W}_\mathbf{O}^\sim \rangle \mid \langle \mathcal{W}_\mathbf{O}^\sim \rangle \equiv \langle \mathcal{W}_\mathbf{O}^\sim \rangle$$

The logic **DCLuN** is axiomatized by adding to **CLuN** the axioms (K), (D), and closing the resulting set under (N) and (MP). Note that for (D) we need the original version (cf. page 8), hence with classical negations (\neg) only.

The semantics for **DCLuN** looks as follows. A model is a quadruple $M = \langle W, w_0, R, v \rangle$ where W is a non-empty set, $w_0 \in W$, $R \subseteq W \times W$ is a serial accessibility relation, and $v : \mathcal{W}_\mathbf{O}^\sim \times W \rightarrow \{1, 0\}$ is a valuation function. As with **CLuN**, we first assign truth values to both schematic letters and formulas of the form $\sim A$: $v_a : \mathcal{S} \cup \{\sim A \mid A \in \mathcal{W}_\mathbf{O}^\sim\} \times W \rightarrow \{0, 1\}$. v_a is extended to v as follows:

(SC1') For formulas $A \in \mathcal{S} \cup \{\sim A \mid A \in \mathcal{W}_\mathbf{O}^\sim\}$: $M, w \models A$ iff $v_a(A, w) = 1$.

(SC2') For $\mathbf{O}, \neg, \vee, \wedge, \supset, \equiv$, the semantic clauses for **DCLuN** are exactly those of **SDL** (cf. Section 2).

A model M is a **DCLuN-model** iff it satisfies the following condition on v :

$$\text{for all } w \in W, \text{ for all } A : v(A, w) = 1 \text{ or } v(\sim A, w) = 1 \quad (C_u)$$

$\Gamma \Vdash_{\mathbf{DCLuN}} A$ iff for all **DCLuN-models** M : if $M, w_0 \models B$ for all $B \in \Gamma$, then $M, w_0 \models A$.

The proof of soundness for this logic is a matter of routine. For completeness, we can use the well-known technique of canonical models (see e.g. [24, Chapter 4]), adjusted to the setting with an actual world. Fix a maximal, \neg -consistent set $\Gamma \subseteq \mathcal{W}_\mathbf{O}^\sim$. We build the canonical model $M_\Gamma^c = \langle W^c, \Gamma, R^c, V^c \rangle$ for this set as follows:

- (i) W^c is the set of all maximal consistent and **DCLuN**-closed sets Δ ,
- (ii) $R^c = \{(\Delta, \Delta') \mid \{A \mid \mathbf{O}A \in \Delta\} \subseteq \Delta'\}$,
- (iii) for all $A \in \mathcal{S} \cup \{\sim A \mid A \in \mathcal{W}_\mathbf{O}^\sim\}$, for all $\Delta \in W^c$: $v_a(A, \Delta) = 1$ iff $A \in \Delta$.

To show that M_Γ^c is a **DCLuN-model**, we need to rely on excluded middle for \sim and the maximality of each $\Delta \in W^c$. For seriality, we rely on the (D)-axiom in the usual way. The proof of the truth lemma proceeds by a standard induction. So we can derive that all the members of Γ are satisfied at Γ in M_Γ^c .

Note that, since **CLuN** is a conservative extension of **CL**, **DCLuN** is also a conservative extension of **SDL**. However, if we consider the \neg -free fragment of

DCLuN, and treat \sim as the “proper” negation, then **DCLuN** is a proper fragment of **SDL**. When applying the logic **DCLuN** to concrete examples, we will use \sim to translate negations in natural language. Given this convention, the logic **DCLuN** is strongly conflict-tolerant.

$$OA \wedge O\sim A \not\vdash_{\mathbf{DCLuN}} OB$$

$$OA \wedge \sim OA \not\vdash_{\mathbf{DCLuN}} OB$$

In **DCLuN** we can define permission in various ways relative to our negation operators:

$$P_{\neg}^{\square} A =_{\text{df}} \neg O \neg A$$

$$P_{\sim}^{\square} A =_{\text{df}} \neg O \sim A$$

$$P_{\neg}^{\sim} A =_{\text{df}} \sim O \neg A$$

$$P_{\sim}^{\sim} A =_{\text{df}} \sim O \sim A$$

All of these permission operators tolerate conflicts between an obligation and a permission, as well as contradictory norms. Where $\dagger, \ddagger \in \{\sim, \neg\}$:

$$OA \wedge P_{\dagger}^{\ddagger} \sim A \not\vdash_{\mathbf{DCLuN}} OB \quad (43)$$

$$O\sim A \wedge P_{\dagger}^{\ddagger} A \not\vdash_{\mathbf{DCLuN}} OB \quad (44)$$

$$P_{\dagger}^{\ddagger} A \wedge \sim P_{\dagger}^{\ddagger} A \not\vdash_{\mathbf{DCLuN}} OB \quad (45)$$

In sum, **DCLuN** is very conflict-tolerant, especially compared to the logics discussed in previous sections. However, it is also rather weak. To be sure, the Jones argument, the Roberts arguments, and the (original and modified) Natascha argument are valid in **DCLuN** due to the validity of (Inh) and (Agg). Unfortunately, the Smith argument and the Thomas argument are not **DCLuN**-valid. More generally, all instances of the following inference schemas fail in **DCLuN**:

$$O(A \supset B) \not\vdash_{\mathbf{DCLuN}} O(\sim B \supset \sim A) \quad (46)$$

$$O(A \supset B), O\sim B \not\vdash_{\mathbf{DCLuN}} O\sim A \quad (47)$$

$$O(A \vee B), O\sim A \not\vdash_{\mathbf{DCLuN}} OB \quad (48)$$

The invalidity of (46)-(48) mirrors the invalidity of their non-deontic counterparts (40)-(42) in **CLuN**. So the main advantage of **DCLuN** goes hand in hand with its inability to validate seemingly intuitive inferences. This drawback is overcome by strengthening this system within the adaptive logics framework.

Going adaptive We strengthen **DCLuN** to the adaptive logic **DCLuN^x**, which is defined by the triple $\langle \mathbf{DCLuN}, \Omega^\sim, \mathbf{x} \rangle$, where

$$\Omega^\sim = \{A \wedge \sim A \mid A \in \mathcal{W}_\circ^\sim\} \cup \{P_\sqsupset(A \wedge \sim A) \mid A \in \mathcal{W}^\sim\}$$

Ω^\sim contains not only plain contradictions, but also formulas that express that in some deontically accessible world, a given contradiction is true. This allows us at once to validate the Smith argument and the Thomas argument. Here is a **DCLuN^x**-proof illustrating the validity of the Thomas argument:

1	$O(t \wedge (f \vee s))$	Prem	\emptyset
2	$O(\sim t \wedge \sim f)$	Prem	\emptyset
3	$O(f \vee s)$	1; RU	\emptyset
4	$O\sim f$	2; RU	\emptyset
5	O_s	3,4; RC	$\{P_\sqsupset(f \wedge \sim f)\}$

The inference made at line 5 holds in view of the **DCLuN**-valid inference

$$O(f \vee s), O\sim f \vdash O_s \vee P_\sqsupset(f \wedge \sim f) \quad (49)$$

Suppose that $O(f \vee s)$ and $O\sim f$. By (Agg), $O((f \vee s) \wedge \sim f)$. By normal modal logic properties, we can infer $O_s \vee \neg O\neg(f \wedge \sim f)$ so that we can derive O_s on the condition $P_\sqsupset(f \wedge \sim f)$.

Equations (50)-(55) illustrate that the **DCLuN**-invalid inferences (40)-(42) and (46)-(48) hold conditionally in **DCLuN^x**. The conditions on which these inferences can be made in a **DCLuN^x**-proof are indicated between square brackets.

$$p \supset q \vdash_{\mathbf{DCLuN}^\mathbf{x}} \sim q \supset \sim p \quad [q \wedge \sim q] \quad (50)$$

$$p \supset q, \sim q \vdash_{\mathbf{DCLuN}^\mathbf{x}} \sim p \quad [q \wedge \sim q] \quad (51)$$

$$p \vee q, \sim p \vdash_{\mathbf{DCLuN}^\mathbf{x}} q \quad [p \wedge \sim p] \quad (52)$$

$$O(p \supset q) \vdash_{\mathbf{DCLuN}^\mathbf{x}} O(\sim q \supset \sim p) \quad [P_\sqsupset(q \wedge \sim q)] \quad (53)$$

$$O(p \supset q), O\sim q \vdash_{\mathbf{DCLuN}^\mathbf{x}} O\sim p \quad [P_\sqsupset(q \wedge \sim q)] \quad (54)$$

$$O(p \vee q), O\sim p \vdash_{\mathbf{DCLuN}^\mathbf{x}} Oq \quad [P_\sqsupset(p \wedge \sim p)] \quad (55)$$

More generally, relative to premise sets from which no abnormalities are **DCLuN**-derivable \sim is as strong as \neg in **DCLuN^x**. That is, where $A \in \mathcal{W}_\circ^\sim$, let $\pi(A)$ be the result of replacing every occurrence of \sim in A with \neg . We lift this translation to sets of formulas in the usual way. We can now prove the following:

Theorem 7.1. *If Γ is normal, then $\Gamma \vdash_{\mathbf{DCLuN}^\mathbf{x}} A$ iff $\pi(\Gamma) \vdash_{\mathbf{SDL}} \pi(A)$.*

Proof. The upper limit logic of \mathbf{DCLuN}^\times is obtained by adding to \mathbf{DCLuN} all formulas $\neg A$ for which $A \in \Omega^\sim$. Call this logic \mathbf{UDCLuN} . By Theorem 3.14: If Γ is normal, then $\Gamma \vdash_{\mathbf{DCLuN}^\times} A$ iff $\Gamma \vdash_{\mathbf{UDCLuN}} A$. We show that $\Gamma \vdash_{\mathbf{UDCLuN}} A$ iff $\pi(\Gamma) \vdash_{\mathbf{SDL}} \pi(A)$.

(\Rightarrow) It is easily checked that, under the transformation given, all \mathbf{CLuN} -valid inferences are \mathbf{CL} -valid; (K), (D), and (N) are \mathbf{SDL} -valid; and all elements of $\pi(\{\neg A \mid A \in \Omega^\sim\})$ are \mathbf{SDL} -valid.

(\Leftarrow) Given the fact that \mathbf{UDCLuN} , like \mathbf{DCLuN} , extends \mathbf{SDL} , it suffices to show that \sim is as strong as \neg in \mathbf{UDCLuN} :

$$\vdash_{\mathbf{UDCLuN}} \sim A \supset \neg A \quad (56)$$

$$\vdash_{\mathbf{UDCLuN}} \mathbf{O}\sim A \supset \mathbf{O}\neg A \quad (57)$$

Ad. (56) Suppose $\sim A$. Then $\neg A \vee (A \wedge \sim A)$ since $\vdash_{\mathbf{CLuN}} \sim A \supset (\neg A \vee (A \wedge \sim A))$. We also know that $\vdash_{\mathbf{UDCLuN}} \neg(A \wedge \sim A)$, so by \mathbf{CL} -properties we obtain $\neg A$.

Ad. (57) By (N), $\vdash_{\mathbf{UDCLuN}} \mathbf{O}(\sim A \supset (\neg A \vee (A \wedge \sim A)))$. Suppose $\mathbf{O}\sim A$. By (K) and (MP), $\mathbf{O}(\neg A \vee (A \wedge \sim A))$. By \mathbf{SDL} -properties, $\mathbf{O}\neg A \vee \mathbf{P}\sqsupset(A \wedge \sim A)$. But then $\mathbf{O}\neg A$ follows in view of $\vdash_{\mathbf{UDCLuN}} \neg\mathbf{P}\sqsupset(A \wedge \sim A)$. \square

7.2 Semi-paraconsistent adaptive deontic logic

The logic \mathbf{DCLuN} and its adaptive extensions consistently accommodate all types of normative conflicts that we have encountered so far. But they also consistently accommodate plain contradictions between formulas not involving deontic operators, such as $p \wedge \sim p$. One could argue that this is overkill. Even if normative conflicts are part of life and should be accommodated in a deontic logic, there is no need to allow also for a non-deontic statement and its negation to be true at the same time.

In this section we mention two ways to adjust \mathbf{DCLuN} and its adaptive extensions so as to tolerate normative conflicts, without having to tolerate all outright contradictions of the form $A \wedge \sim A$. Casey McGinnis coined the term *semi-paraconsistent deontic logic* for paraconsistent deontic logics that meet this desideratum [64; 63].

Excluding non-deontic contradictions The logic \mathbf{DCLuN}_1 is obtained by closing \mathbf{DCLuN} under the axiom schema (Cons_1) .⁵³

$$\text{Where } A \in \mathcal{W}^\sim : \sim A \supset \neg A \quad (\text{Cons}_1)$$

⁵³Where $\vdash \subseteq \wp(\Phi) \times \Phi$ is a consequence relation and Δ is a set of axioms, we obtain \vdash_Δ , the closure of \vdash under Δ , as follows: $\Gamma \vdash_\Delta A$ iff $\Gamma \cup \Delta \vdash A$. This means that one cannot e.g. apply necessitation to members of Δ .

Where $A \in \mathcal{W}^\sim$, (Cons_1) takes care that $A \wedge \sim A$ is trivialized in \mathbf{DCLuN}_1 . This means that for non-deontic formulas, we obtain full \mathbf{CL} . Still, \mathbf{DCLuN}_1 , like \mathbf{DCLuN} , is highly conflict-tolerant. Where as before $\dagger, \ddagger \in \{\sim, \neg\}$:

$$OA \wedge O\sim A \not\vdash_{\mathbf{DCLuN}_1} OB \quad (58)$$

$$OA \wedge P_{\dagger}^{\ddagger} \sim A \not\vdash_{\mathbf{DCLuN}_1} OB \quad (59)$$

$$O\sim A \wedge P_{\dagger}^{\ddagger} A \not\vdash_{\mathbf{DCLuN}_1} OB \quad (60)$$

$$OA \wedge \sim OA \not\vdash_{\mathbf{DCLuN}_1} OB \quad (61)$$

$$P_{\dagger}^{\ddagger} A \wedge \sim P_{\dagger}^{\ddagger} A \not\vdash_{\mathbf{DCLuN}_1} OB \quad (62)$$

As desired, \mathbf{DCLuN}_1 consistently accommodates normative conflicts while trivializing contradictions between statements without occurrences of deontic operators.

Semantically, the logic \mathbf{DCLuN}_1 is characterized by imposing the following additional condition on \mathbf{DCLuN} -models:

$$\text{For all } A \in \mathcal{W}^\sim : v(A, w_0) = 1 \text{ iff } v(\sim A, w_0) = 0 \quad (C_1^0)$$

Unlike \mathbf{DCLuN} , the logic \mathbf{DCLuN}_1 is not a normal modal logic, since it is not closed under the standard necessitation rule (N). That is, even though $\sim p \supset \neg p$ is a theorem of the logic, $O(\sim p \supset \neg p)$ is not. For similar reasons, the logic is not closed under Uniform Substitution. For instance, $\sim Op \supset \neg Op$ is not a theorem of \mathbf{DCLuN}_1 .

Adaptive logics based on \mathbf{DCLuN}_1 can be defined just as before. Mind however that abnormalities of the form $A \wedge \sim A$ for $A \in \mathcal{W}^\sim$ are vacuous in the resulting adaptive logics, since they are anyway trivialized by their lower limit logic, in view of (Cons_1) . These adaptive logics will perform just as well as \mathbf{DCLuN}^x , in that they validate all the inferences from our list of benchmark examples.

Excluding all contradictions at the actual world A second, stronger semi-paraconsistent deontic logic is obtained by closing \mathbf{DCLuN} under the unrestricted version of (Cons_1) :

$$\sim A \supset \neg A \quad (\text{Cons}_2)$$

Call the resulting logic \mathbf{DCLuN}_2 . Its semantics is obtained by imposing the following condition on \mathbf{DCLuN} -models:

$$v(\sim A, w_0) = 1 \text{ iff } v(A, w_0) = 0 \quad (C_2^0)$$

In the \mathbf{DCLuN}_2 -semantics, \sim and \neg are interchangeable at w_0 . At all other worlds, \neg remains strictly stronger than \sim . This means that contradictions outside the scope of O are trivialized, whereas contradictions within the scope of O are not.

The logic **DCLuN₂** is not as conflict-tolerant as **DCLuN₁**, since it trivializes conflicts of the form $OA \wedge \sim OA$ or $P_{\dagger}^{\ddagger}A \wedge \sim P_{\dagger}^{\ddagger}A$, where $\dagger, \ddagger \in \{\sim, \neg\}$. Since (Cons₂) and (C₂⁰) are no longer restricted to members of \mathcal{W}^{\sim} , the logic **DCLuN₂** satisfies the rule of uniform substitution, although necessitation (in its full generality) is still invalid.

Just as with **DCLuN** and **DCLuN₁**, we can use **DCLuN₂** as a lower limit logic of our adaptive logic. In this case, the set of abnormalities can be further simplified to the following:

$$\Omega_2^{\sim} = \{P_{\neg}^{\sim}(A \wedge \sim A) \mid A \in \mathcal{W}^{\sim}\}$$

7.3 Other paraconsistent negations

CLuN is the weakest logic which verifies the full positive fragment of **CL** as well as the principle of Excluded Middle (EM). Stronger paraconsistent logics can be obtained by adding to **CLuN** the double negation laws and/or de Morgan's laws for negation:

$$\begin{aligned} \sim\sim A &\equiv A && (A\sim\sim) \\ \sim(A \supset B) &\equiv (A \wedge \sim B) && (A\sim\supset) \\ \sim(A \wedge B) &\equiv (\sim A \vee \sim B) && (A\sim\wedge) \\ \sim(A \vee B) &\equiv (\sim A \wedge \sim B) && (A\sim\vee) \\ \sim(A \equiv B) &\equiv ((A \vee B) \wedge (\sim A \vee \sim B)) && (A\sim\equiv) \end{aligned}$$

Let **CLuNs** be obtained by adding all of these axioms to **CLuN**. Analogously to the construction of **DCLuN**, we can now construct the logic **DCLuNs** by enriching **CLuNs** with (K), (D), and (N).

One clear difference between **DCLuN**-based ALs and **DCLuNs**-based ALs is that the latter verify a number of additional inferences in a non-defeasible way. For instance, where $\Gamma = \{O(p \wedge q), O\sim(p \wedge q)\}$, one cannot **DCLuN^r**-derive $O(\sim p \vee \sim q)$ from Γ , since one cannot rely on the falsehood of the abnormality $P_{\neg}^{\sim}((p \wedge q) \wedge \sim(p \wedge q))$. In contrast, one can finally **DCLuNs^r**-derive $O(\sim p \vee \sim q)$ from the same premise set, simply in view of properties of **DCLuNs**.

We have to take care when constructing adaptive logics on the basis of **DCLuNs**. Suppose that we work with the set Ω^{\sim} of **DCLuN^x**-abnormalities.

1	Op	Prem	\emptyset
2	$O\sim p$	Prem	\emptyset
3	Oq	Prem	\emptyset
4	$O(\sim q \vee r)$	Prem	\emptyset
5	Or	3,4;RC	$\{P_{\square}^{\sim}(q \wedge \sim q)\} \checkmark^6$
6	$P_{\square}^{\sim}(q \wedge \sim q) \vee P_{\square}^{\sim}((p \wedge r) \wedge \sim(p \wedge r))$	1-4;RU	\emptyset

Line 5 is marked in view of the minimal **Dab**-formula derived at line 6. There is no extension of this proof in which to unmark line 5. The proof illustrates that Or is not finally derivable from the premises at lines 1-4. This is counter-intuitive.

If we are to build an adaptive logic on the basis of the lower limit logic **DCLuNs** and the set of abnormalities Ω^{\sim} , the resulting logic would exhibit flip-flop behavior (see Section 5 where we also encountered this problem). The solution is to restrict the set of abnormalities as follows:

$$\Omega_s^{\sim} = \{A \wedge \sim A \mid A \in \mathcal{S}\} \cup \{OA \wedge \sim OA \mid A \in \mathcal{W}^{\sim}\} \cup \{P_{\square}^{\sim}(A \wedge \sim A) \mid A \in \mathcal{S}\} \quad (63)$$

Given $(A\sim\sim)-(A\sim \equiv)$, inconsistencies between complex formulas in \mathcal{W} can be reduced to inconsistencies at the level of atoms in **DCLuNs**. In view of this, **DCLuNs**^x-abnormalities must be restricted accordingly, on pain of flip-flop behavior. That is, where $A \in \mathcal{W}$, $A \wedge \sim A$ and $P_{\square}^{\sim}(A \wedge \sim A)$ only counts as an abnormality when $A \in \mathcal{S}$.

The situation is different for formulas of the form $OA \wedge \sim OA$: within the scope of O , inconsistencies between complex formulas do *not* reduce to inconsistencies at the level of atoms. For instance, the inference from $O(p \wedge q) \wedge \sim O(p \wedge q)$ to $(Op \wedge \sim Op) \vee (Oq \wedge \sim Oq)$ is not **DCLuNs**-valid, since $\sim O(p \wedge q)$ does not **DCLuNs**-entail $\sim Op \vee \sim Oq$. More generally, where A is a complex formula, the formula $\sim OA$ cannot be further analysed in **DCLuNs**. So, as in **DCLuN**^x, all formulas of the form $OA \wedge \sim OA$ count as abnormalities in **DCLuNs**^x.

Let **DCLuNs**^x be the adaptive logic defined by the lower limit logic **DCLuNs**, the set of abnormalities Ω_s^{\sim} , and the strategy $\times \in \{r, m\}$. Then clearly the formula derived at line 6 of the proof above is no longer a minimal **Dab**-formula, and line 5 remains unmarked. We can still derive the **Dab**-formula $P_{\square}^{\sim}(q \wedge \sim q) \vee P_{\square}^{\sim}(p \wedge \sim p) \vee P_{\square}^{\sim}(r \wedge \sim r)$ from lines 1-4 via RU, in view of

$$P_{\square}^{\sim}((p \wedge r) \wedge \sim(p \wedge r)) \vdash_{\mathbf{DCLuNs}} P_{\square}^{\sim}(p \wedge \sim p) \vee P_{\square}^{\sim}(r \wedge \sim r) \quad (64)$$

However, this **Dab**-formula is not minimal, since its disjunct $P_{\square}^{\sim}(p \wedge \sim p)$ is a **DCLuNs**-consequence of the formulas Op and $O\sim p$ at lines 1 and 2. As a result, line 5 is finally derivable and Or is a **DCLuNs**^x-consequence of the premises.

Other than **CLuN** and **CLuNs**, there is a wide variety of paraconsistent logics that can serve as the core logic of an inconsistency-adaptive logic. We could, for instance, treat ‘ \sim ’ as a dummy operator for which not even (EM) holds by removing $(A\sim 1)$ in the axiomatization of **CLuN**. The resulting logic is called **CLoN** (for Classical Logic with both gluts and gaps for Negation). Extending **CLoN** with $(A\sim\sim)$ - $(A\sim\equiv)$ results in the logic **CLoNs**. These systems too can be extended deontically and adaptively. In addition, one can also consider semi-paraconsistent versions of **DCLuNs** and **DCLoNs**.

7.4 Further reading and open ends

For a general overview of paraconsistent logic, see e.g. [79; 80]. For an overview of (monotonic) paraconsistent deontic logic, we refer to [42, Sect. 6.1].

The first paper on inconsistency-adaptive logic – published in 1989, but written in 1981 – is [5], where the proof theory for the reliability strategy was first presented. The minimal abnormality strategy was first presented (semantically) in [4]. The (propositional) results of the two aforementioned papers were generalized to the predicative level in [7]. For an overview and more recent results within the inconsistency-adaptive program, see [12].

Inconsistency-adaptive deontic logics were presented in [15; 22], in [21], and in [43]. Most of these systems – in contrast to the ones presented in this section – allow for the following inference:⁵⁴

$$OA \wedge O\sim A \vdash \sim O\sim A \wedge O\sim A \tag{65}$$

That is, conflicts of the form $OA \wedge O\sim A$ entail plain contradictions. Goble is critical of such systems:

That seems an exceedingly strong commitment. It is easy to accept that there are normative conflicts, harder to suppose they all yield contradictions that are true. Even Priest, the hierarch of dialetheism, does not consider normative conflicts so paradoxical [43, Fn. 15].

The systems presented in this section circumvent Goble’s criticism by invalidating inferences like (65).

In [16] the semi-paraconsistent deontic logic **LNP** is presented and extended within the adaptive logics framework. **LNP** is a close cousin of **DCLoNs₂**, but has

⁵⁴(65) holds for the inconsistency-adaptive systems presented in [15; 22], and [21]. The closely related principle $OA \wedge O\sim A \vdash (O\sim A \wedge \sim O\sim A) \vee OB$ holds for those logics mentioned in [43] which satisfy the ‘deontic addition’ schema $OA \supset O(A \vee B)$.

a slightly different language in which the P-operator is primitive, and in which ‘ \neg ’ is allowed only outside the scope of deontic operators, while ‘ \sim ’ is allowed only inside the scope of deontic operators.

Once we are open to the possibility of changing the logic of the connectives, new questions arise. For instance, why should we always blame negation for the explosive behavior of a logic, and why not weaken the meaning of the other connectives? Why not e.g. give up addition for \vee (i.e., to derive $A \vee B$ from A or from B)? In [8], Batens shows that a whole range of interesting new logics come to the fore, once we generalize the idea of gluts and gaps to other connectives and logical operators. The application of all this to deontic reasoning is yet to be studied in detail, but it can draw on many existing results concerning corrective ALs.

In [17], a very rich paraconsistent deontic logic is presented, one that allows the user to express not only obligations that concern states of affairs, but also obligations that concern agency. The language of these systems contains modal operators \Box_J for “the group of agents J brings it about that”, inspired by existing work on logics of agency [86; 23; 31]. This in turn allows one to distinguish between various different types of *inter-personal* and *intra-personal* deontic conflicts:⁵⁵

$$O\Box_i A \wedge O\Box_j \sim A \tag{66}$$

$$O\Box_i A \wedge P\Box_j \sim A \tag{67}$$

$$O\Box_i A \wedge O\Box_i \sim A \tag{68}$$

$$O\Box_i A \wedge P\Box_i \sim A \tag{69}$$

$$O\Box_i A \wedge O\sim\Box_i A \tag{70}$$

$$O\Box_i A \wedge P\sim\Box_i A \tag{71}$$

$$O\Box_i A \wedge \sim O\Box_i A \tag{72}$$

$$P\Box_i A \wedge \sim P\Box_i A \tag{73}$$

One further advantage of such richer formal languages in the context of adaptive reasoning is that they allow us to prioritize the minimization of certain types of conflicts over that of others. For instance, we may consider conflicts of type (68) worse than those of type (66) and (67), since the former clearly violate the principle that if an agent ought to bring about A , then that agent is also able to see to A

⁵⁵An inter-personal conflict is one that holds between the obligations of different agents, whereas an intra-personal conflict obtains between the obligations of a single agent. One famous example of an inter-personal normative conflict can be found in Sophocles’ *Antigone*, where due to the city’s laws, Creon is obliged to prevent the burial of Antigone’s brother Polyneices, but Antigone faces a religious and familial obligation to bury Polyneices [62; 44].

– assuming agents cannot bring about contradictions. Such a prioritized reasoning can be modeled in terms of a lexicographic AL (cf. Section 3.4).

8 Conflict-tolerant adaptive logics: round-up

In this section, we give an overview of the main features of the logics discussed so far. We start by giving an overview of the performance of revisionist ALs with respect to the criteria introduced in Section 4.2. In Section 8.2 we return to the logics from Section 3. We show how these can be evaluated using similar criteria, and how they can be enriched in various ways.

8.1 Revisionist deontic adaptive logics: overview

The behavior of the revisionist adaptive logics with respect to the criteria from Section 4.2 is summarized in Tables 1 and 2. Principles (arguments) that are valid in a given logic receive a 3, invalid principles (arguments) receive a 7.⁵⁶ Where the premises of an argument are trivialized by a given logic, we write a \perp in Table 2.

	DEX	DEX-O \perp	DEX-P \perp	DEX-OP \neg	DEX-O \neg P
LUM.a ^x	\times	\times	\checkmark	\checkmark	\times
LUM.b ^x	\times	\checkmark	\checkmark	\checkmark	\times
LUM.c ^x	\times	\checkmark	\checkmark	\checkmark	\times
P ^x	\times	\checkmark	\checkmark	\checkmark	\times
DCLuN ^x	\times	\times	\times	\times	\times
DCLuN ₁ ^x	\times	\times	\times	\times	\times
DCLuN ₂ ^x	\times	\times	\times	\times	\checkmark

Table 1: Behavior of deontic ALs with respect to various explosion principles.

It should be noted here once more (in line with our remarks in Section 4.2) that whether a given AL validates some form of deontic explosion or a specific inference should not be seen as conclusive evidence in favour of or against such a logic. The above tables are mostly for purposes of comparison and classification, and do not serve as strict criteria of the relative success or failure of the respective systems or their purposes. For example, with a view to supporting ought-implies-can, a system might be *designed* to consider $O(A \wedge \neg A)$ inconsistent even while $OA \wedge O\neg A$ is

⁵⁶As noted before, for the logics from Section 7 we assume that the principles (arguments) in question are formalized using the paraconsistent negation sign \sim .

	S	J	R1	R2	T	N1	N2
LUM.a^r	✓	✓	✗	✗	✗	⊥	✗
LUM.a^m	✓	✓	✗	✗	✗	⊥	✓
LUM.b^r	✓	✓	✗	✗	✗	⊥	✗
LUM.b^m	✓	✓	✗	✗	✗	⊥	✓
LUM.c^r	✓	✓	✗	✗	✗	✗	✗
LUM.c^m	✓	✓	✗	✗	✗	✓	✓
P^x	✓	✓	✓	✓	✓	✓	✓
DCLuN^x	✓	✓	✓	✓	✓	✓	✓
DCLuN₁^x	✓	✓	✓	✓	✓	✓	✓
DCLuN₂^x	✓	✓	✓	✓	✓	⊥	✓
DCLuNs^x	✓	✓	✓	✓	✓	✓	✓

Table 2: Behavior of deontic ALs with respect to the Smith (S), Jones (J), Roberts (R1 and R2), Thomas (T), and Natascha (N1 and N2) arguments from Section 4.)

consistent. In that case, that the system validates (DEX- $O\perp$) may be taken as a virtue rather than a vice. Likewise, the validation of (DEX- $OP\neg$) would be embraced by one with a classical point of view (and given the standard interdefinability of O and P).

Let us close this overview with a technical point. All ALs discussed in Sections 5–7 have a monotonic, conflict-tolerant deontic logic as their lower limit logic. The latter logics are mutually incomparable, in the sense that none is stronger than any other.⁵⁷ For instance, the logic **LUM.a** from Section 5 invalidates (Inh) but validates (Agg); conversely, the logic **P** that is discussed in Section 6 invalidates (Agg) but validates (Inh). It can easily be shown that any two ALs that are based on such incomparable lower limit logics, are themselves equally incomparable. This is an immediate corollary of the following:⁵⁸

Theorem 8.1. *Let \mathbf{AL}_1 and \mathbf{AL}_2 be two ALs in standard format, defined by the triples $\langle \mathbf{LLL}_1, \Omega_1, \times_1 \rangle$, resp. $\langle \mathbf{LLL}_2, \Omega_2, \times_2 \rangle$, over a given formal language. If $\vdash_{\mathbf{AL}_1} \subseteq \vdash_{\mathbf{AL}_2}$, then $\vdash_{\mathbf{LLL}_1} \subseteq \vdash_{\mathbf{LLL}_2}$.*

Proof. By contraposition: suppose that $\vdash_{\mathbf{LLL}_1} \not\subseteq \vdash_{\mathbf{LLL}_2}$. Let Γ, A be such that

⁵⁷A small warning is in place here. The paraconsistent deontic logics of the **DCLuN**-family, presented in Section 7, work with a richer language that contains both a paraconsistent and a classical negation. The claim we make here concerns the fragment of those logics *without* the classical negation.

⁵⁸Theorem 8.1 generalizes one direction of Theorem 3.3 in [104].

(i) $\Gamma \vdash_{\mathbf{LLL}_1} A$ but (ii) $\Gamma \not\vdash_{\mathbf{LLL}_2} A$. By (i) and the monotonicity of \mathbf{LLL}_1 , (iii) $\Gamma \cup \{\neg A\} \vdash_{\mathbf{LLL}_1} A$. by (ii), $\Gamma \cup \{\neg A\}$ is \mathbf{LLL}_2 -consistent, and hence by \mathbf{CL} -properties, (iv) $\Gamma \cup \{\neg A\} \not\vdash_{\mathbf{LLL}_2} A$. By (iii) and Theorem 3.15, $\Gamma \cup \{\neg A\} \vdash_{\mathbf{AL}_1} A$. By (iv) and Theorem 3.11, $\Gamma \cup \{\neg A\} \not\vdash_{\mathbf{AL}_2} A$. Hence, $\vdash_{\mathbf{AL}_1} \not\subseteq \vdash_{\mathbf{AL}_2}$. \square

As a result, the ALs discussed in Sections 5-7 are incomparable, i.e. an AL belonging to one of these three types cannot in general be stronger or weaker than an AL belonging to another of the three types.

8.2 *Prima facie* obligations revisited

Explosion principles To apply the criteria from Section 4.2 to the logics from Section 3.1, we need some more preparation. We take it that the premises of the explosion principles, resp. arguments under consideration are all concerned with *prima facie* obligations, whereas their conclusion concerns actual obligations. Under this translation, \mathbf{SDL}_p^r and \mathbf{SDL}_p^m invalidate the analogues of (DEX) and (DEX- \perp):

$$\text{OP}A \wedge \text{OP}\neg A \vdash \text{OB} \quad (74)$$

$$\text{OP}(A \wedge \neg A) \vdash \text{OB} \quad (75)$$

The other explosion principles cannot as easily be translated to these systems, because in Section 3.1 we did not define a corresponding *prima facie* permission operator for the logics \mathbf{SDL}_p^x .

Suppose that we add a second dummy operator PP to the language of \mathbf{SDL}_p . For the adaptive extension of the resulting logic, we re-define the set of abnormalities Ω_p by including both formulas of the form $\text{OP}A \wedge \neg \text{OA}$ and formulas of the form $\text{PP}A \wedge \neg \text{PA}$. In the resulting logic, the following analogues of the explosion principles (DEX- \perp) and (DEX- $\text{OP}\neg$) are invalid:

$$\text{PP}(A \wedge \neg A) \vdash \text{OB} \quad (76)$$

$$\text{OP}A \wedge \text{PP}\neg A \vdash B \quad (77)$$

$$\text{OP}A \wedge \neg \text{PP}A \vdash B \quad (78)$$

Note that conflicts of the form $\text{OP}A \wedge \text{PP}\neg A$ give rise to disjunctions of abnormalities in this logic:

$$\text{OP}A \wedge \text{PP}\neg A \vdash (\text{OP}A \wedge \neg \text{OA}) \vee (\text{PP}\neg A \wedge \neg \text{P}\neg A) \quad (79)$$

In case there is a conflict between a *prima facie* obligation and a *prima facie* permission, the adaptive logic will not prioritize one over the other. This is in line

with [?], where it is argued that permission should not take priority over obligations or conversely. Should one nevertheless want a logic that does treat one type of conflict as “worse” than the other, then one can turn to the format of lexicographic ALs as sketched in Section 3.4.

Benchmark examples First, in both \mathbf{SDL}_p^r and \mathbf{SDL}_p^m , the Smith and Jones arguments are \mathbf{SDL}_p^x -valid, while Roberts and Thomas are not.

$$\mathbf{O}^p(f \vee s), \mathbf{O}^p\neg f \vdash_{\mathbf{SDL}_p^x} \mathbf{O}s \quad (\text{Smith})$$

$$\mathbf{O}^p(j \wedge s) \vdash_{\mathbf{SDL}_p^x} \mathbf{O}j \quad (\text{Jones})$$

$$\mathbf{O}^p(t \wedge r), \mathbf{O}^p(\neg t \wedge v) \not\vdash_{\mathbf{SDL}_p^x} \mathbf{O}r \wedge \mathbf{O}v \quad (\text{Roberts 1})$$

$$\mathbf{O}^p(t \wedge r), \mathbf{O}^p(\neg t \wedge v) \not\vdash_{\mathbf{SDL}_p^x} \mathbf{O}(r \wedge v) \quad (\text{Roberts 2})$$

$$\mathbf{O}^p(t \wedge (f \vee s)), \mathbf{O}^p(\neg t \wedge \neg f) \not\vdash_{\mathbf{SDL}_p^x} \mathbf{O}s \quad (\text{Thomas})$$

In order to infer the conclusions of the Roberts and Thomas arguments, we would need to detach the obligations $\mathbf{O}(t \wedge r)$ and $\mathbf{O}(t \wedge (f \vee s))$ respectively. But we cannot do that in view of the following minimal **Dab**-consequences of the respective premise sets:

$$(\mathbf{O}^p(t \wedge r) \wedge \neg \mathbf{O}(t \wedge r)) \vee (\mathbf{O}^p(\neg t \wedge v) \wedge \neg \mathbf{O}(\neg t \wedge v)) \quad (80)$$

$$(\mathbf{O}^p(t \wedge (f \vee s)) \wedge \neg \mathbf{O}(t \wedge (f \vee s))) \vee (\mathbf{O}^p(\neg t \wedge \neg f) \wedge \neg \mathbf{O}(\neg t \wedge \neg f)) \quad (81)$$

One way of accounting for the Roberts and Thomas arguments is to strengthen \mathbf{SDL}_p^x by closing the operator \mathbf{O}^p under a number of further rules. For instance, we could add a principle permitting the inference from $\mathbf{O}^p(A \wedge B)$ to \mathbf{O}^pA , such as (Inh). That would enable us to infer \mathbf{O}^pr given $\mathbf{O}^p(t \wedge r)$, and $\mathbf{O}r$ given \mathbf{O}^pr (on the condition $\mathbf{O}^pr \wedge \neg \mathbf{O}r$). Clearly, however, not anything goes when closing \mathbf{O}^p under additional rules. For one thing, we do not want to end up with full **SDL** or even **K** for *prima facie* obligations, as this would completely annihilate our initial objective. But also if we characterize \mathbf{O}^p in terms of weaker logics like the ones presented in Sections 5-7, we should be careful. After all, the richer one’s lower limit logic, the more likely one is to end up with flip-flop problems that will require further tinkering with the set of abnormalities, much as we had to do in previous sections.

For the Natascha argument, one can translate the impossibility of $s \wedge m$ using the operator \mathbf{O} for actual obligations. The underlying idea is that constraints concerning what is practically (im)possible only have a bearing on actual obligations, not on the *prima facie* obligations. This can again be done in two different ways, giving rise to two different premise sets. For both, the validity of the argument will depend on the adaptive strategy:

$$\begin{array}{ll}
 \mathcal{O}^P s, \mathcal{O}^P m, \mathcal{O}^P(s \supset t), \mathcal{O}^P(m \supset t), \neg \mathcal{O}(s \wedge m) \not\vdash_{\mathbf{SDL}_F} \mathcal{O}t & \text{(Natascha 1)} \\
 \mathcal{O}^P s, \mathcal{O}^P m, \mathcal{O}^P(s \supset t), \mathcal{O}^P(m \supset t), \neg \mathcal{O}(s \wedge m) \vdash_{\mathbf{SDL}_M} \mathcal{O}t & \text{(Natascha 1)} \\
 \mathcal{O}^P s, \mathcal{O}^P m, \mathcal{O}^P(s \supset t), \mathcal{O}^P(m \supset t), \mathcal{O}\neg(s \wedge m) \not\vdash_{\mathbf{SDL}_F} \mathcal{O}t & \text{(Natascha 2)} \\
 \mathcal{O}^P s, \mathcal{O}^P m, \mathcal{O}^P(s \supset t), \mathcal{O}^P(m \supset t), \mathcal{O}\neg(s \wedge m) \vdash_{\mathbf{SDL}_M} \mathcal{O}t & \text{(Natascha 2)}
 \end{array}$$

In Sections 4–8 we defined and discussed a large variety of conflict-tolerant deontic logics that can be developed within the AL framework. More variation is possible, as there are other ways still to define conflict-tolerant deontic logics – by moving to a hyperintensional framework, for instance – and strengthen them adaptively. Moreover, existing systems can be altered by making them more expressive, e.g. by considering the interplay between deontic modalities and alethic, doxastic, or epistemic modalities. All this goes to show that adaptive logics provide a versatile and modular framework for conflict-tolerant normative reasoning, and that their applications to this problem are far from exhausted.

9 Conditional obligations and adaptive detachment

SDL is inadequate not just for accommodating normative conflicts in deontic logic, but also for representing deontic conditionals, as we will explain below.⁵⁹ Within the vast literature on such conditionals, one can distinguish three general approaches. The first is to represent them by means of a *dyadic* obligation operator $\mathcal{O}(\cdot \mid \cdot)$, and to read a formula $\mathcal{O}(B \mid A)$ as ‘If A , then B is obligatory’. A second approach is to treat the problems surrounding deontic conditionals as symptomatic of the bigger challenge of how to formalize conditional statements in general. The third approach is more abstract: it treats deontic conditionals as pairs connecting a given “input” with an “output”, and defines specific proof theories and an operational semantics (based on the principle of detachment and **CL**) for such connections.

We will discuss these three different approaches in Sections 9.1-9.3 respectively, showing how the framework of ALs can be useful in each of them. Our discussion will be mainly tentative; we provide pointers to more technical results and fully worked-out proposals in the literature at the end of each subsection.

⁵⁹We will only sketch the latter inadequacy here. It is discussed at length in Section 8.5. and in the Appendix of [48, Chapter 1]. For other overviews of this problem, see for instance [3; 28].

9.1 Adaptive dyadic deontic logics

Helping one’s neighbours Let us illustrate the distinctive problems surrounding deontic conditionals by means of a so-called Chisholm scenario – after [30]. This scenario can be represented as follows in the dyadic setting:

- (i) It is obligatory that Jones goes to the aid of his neighbours ($\mathbf{O}g$).
- (ii) It is obligatory that if Jones goes to the aid of his neighbours, then he tells them he is coming ($\mathbf{O}(t \mid g)$).
- (iii) If Jones does not go to the aid of his neighbours, then he ought not to tell them he is coming ($\mathbf{O}(\neg t \mid \neg g)$).
- (iv) Jones does not go to the aid of his neighbours ($\neg g$).

Recall now the principles of *factual detachment* (FD) and *deontic detachment* (DD) from Section 1:

$$A, \mathbf{O}(B \mid A) \vdash \mathbf{O}B \quad (\text{FD})$$

$$\mathbf{O}A, \mathbf{O}(B \mid A) \vdash \mathbf{O}B \quad (\text{DD})$$

Given premises (iii) and (iv), we can use (FD) to infer an obligation $\mathbf{O}\neg t$ for Jones not to tell his neighbours he is coming. However, given premises (i) and (ii), we can also use (DD) to infer an obligation $\mathbf{O}t$ for Jones to tell his neighbours he is coming.

But now we face a dilemma. Jones cannot both tell and not tell his neighbours he is coming. So, each of (DD) and (FD) has some intuitive appeal, but together they lead to a deontic conflict, and hence explosion if the logic of \mathbf{O} is **SDL**. This is the dilemma of deontic and factual detachment, also known in the literature as “the dilemma of detachment and commitment” [3; 106]. In fact, one should rather speak here of a *trilemma*, since one may deny that **SDL** is an appropriate logic for obligations, and insist that both (FD) and (DD) should be unconditionally valid. This means one needs a conflict-tolerant deontic logic for \mathbf{O} , much as those discussed in preceding sections. Here, we will first focus on the other two horns of the trilemma and exclude conflicts at the level of \mathbf{O} .

Since each of (DD) and (FD) seems reasonable in isolation, Hilpinen and McNamara argue that we cannot just pick one of them at the expense of the other, and that we need to move to a more nuanced position beyond this choice [48, p. 119]. One solution is to make the detachment – via (DD) or (FD) – of unconditional obligations subject to further conditions, such as joint consistency. The AL framework allows us to make this idea exact, and to study its pros and cons.

A simple solution Let \mathbf{SDL}_d be the logic obtained by replacing the unary *prima facie* operator $\mathbf{O}^p(\cdot)$ of \mathbf{SDL}_p with the conditional operator $\mathbf{O}(\cdot \mid \cdot)$. As we did with the \mathbf{O}^p -operator of \mathbf{SDL}_p , we treat the new conditional operator like a dummy operator in \mathbf{SDL}_d .

Some authors treat unconditional obligations $\mathbf{O}A$ on the same foot as conditional obligations of the type $\mathbf{O}(A \mid \top)$. Note that in \mathbf{SDL}_d these are not equivalent. For instance, the conjunction $\mathbf{O}(A \mid \top) \wedge \mathbf{O}(\neg A \mid \top)$ is \mathbf{SDL}_d -consistent, while the conjunction $\mathbf{O}A \wedge \mathbf{O}\neg A$ is not. In line with the interpretation in Section 3, $\mathbf{O}(A \mid \top)$ expresses something like “ A is an unconditional *prima facie* obligation”, whereas the intended reading of $\mathbf{O}A$ is that “ A is an actual obligation”.

In order to detach unconditional obligations from conditional obligations, we strengthen \mathbf{SDL}_d adaptively to the logics \mathbf{SDL}_d^x , which are defined by the triple $\langle \mathbf{SDL}_d, \Omega_d, x \rangle$, with $x \in \{r, m\}$ and $\Omega_d = \Omega_{fd} \cup \Omega_{dd}$:

$$\begin{aligned}\Omega_{fd} &= \{\mathbf{O}(B \mid A) \wedge A \wedge \neg \mathbf{O}B \mid A, B \in \mathcal{W}\} \\ \Omega_{dd} &= \{\mathbf{O}(B \mid A) \wedge \mathbf{O}A \wedge \neg \mathbf{O}B \mid A, B \in \mathcal{W}\}\end{aligned}$$

In view of the \mathbf{SDL}_d -valid inferences (82) and (83), the adaptive logics \mathbf{SDL}_d^x allow for the conditional application of (FD) and (DD):

$$A, \mathbf{O}(B \mid A) \vdash \mathbf{O}B \vee (\mathbf{O}(B \mid A) \wedge A \wedge \neg \mathbf{O}B) \quad (82)$$

$$\mathbf{O}A, \mathbf{O}(B \mid A) \vdash \mathbf{O}B \vee (\mathbf{O}(B \mid A) \wedge \mathbf{O}A \wedge \neg \mathbf{O}B) \quad (83)$$

We illustrate the resulting logic by applying it to the Chisholm scenario in (i)-(iv):

1	$\mathbf{O}g$	Prem	\emptyset
2	$\mathbf{O}(t \mid g)$	Prem	\emptyset
3	$\mathbf{O}(\neg t \mid \neg g)$	Prem	\emptyset
4	$\neg g$	Prem	\emptyset
5	$\mathbf{O}t$	1,2; RC	$\{\mathbf{O}(t \mid g) \wedge \mathbf{O}g \wedge \neg \mathbf{O}t\} \checkmark^7$
6	$\mathbf{O}\neg t$	3,4; RC	$\{\mathbf{O}(\neg t \mid \neg g) \wedge \neg g \wedge \neg \mathbf{O}\neg t\} \checkmark^7$
7	$(\mathbf{O}(t \mid g) \wedge \mathbf{O}g \wedge \neg \mathbf{O}t) \vee$ $(\mathbf{O}(\neg t \mid \neg g) \wedge \neg g \wedge \neg \mathbf{O}\neg t)$	1-4; RU	\emptyset

Lines 4 and 5 remain marked in any extension of this proof, so that neither $\mathbf{O}t$ nor $\mathbf{O}\neg t$ is an \mathbf{SDL}_d^x -consequence of the premises at lines 1-4. Thus, in cases of conflict, the applications of (FD) and (DD) that lead to the conflict are rejected.

Some have taken a bolder stance here by arguing that when factual and deontic detachment lead to a conflict, (FD) overrules (DD) or vice versa. We will not go into this discussion here – see [48, p. 112-124] for an overview of the various positions.

However, let us briefly indicate how this idea of overruling can be modeled with the AL framework.

Recall the lexicographic ALs that were introduced in Section 3.4. Consider the lexicographic ALs defined in terms of the lower limit logic \mathbf{SDL}_d and the sequence $\langle \Omega_{fd}, \Omega_{dd} \rangle$. The idea is that we treat abnormalities with respect to factual detachment as “worst”, and hence give priority to (FD) over (DD). For instance, in the Chisholm case, the abnormality $O(\neg t \mid \neg g) \wedge \neg g \wedge \neg O\neg t$ will be avoided, and hence the abnormality $O(t \mid g) \wedge Og \wedge \neg Ot$ will be assumed to hold. Thus, in such logics, one can conclude that Jones ought not to tell his neighbours he is coming. Other applications of (DD) that do not result in conflicting obligations will remain valid in such logics. Finally, if two different applications of (FD) conflict, they will both be blocked in the adaptive logics.

A (prioritized) combination of various sorts of adaptive reasoning may also be useful for those who insist on the intuitiveness of (FD) and (DD), and use these to cast doubt on the validity of full \mathbf{SDL} for O (cf. our discussion of the trilemma of detachment and commitment, *supra*). Here, one may combine insights and techniques from Sections 5–7 with those from the present section, treating each of (FD), (DD), and (some or all) rules and axioms of \mathbf{SDL} as defeasible. This way one cannot only accommodate deontic conflicts that arise from an applications of either (FD) or (DD) or both – by invalidating those applications – but also conflicting obligations that happen to be simply there, “unconditionally”. In such a setting, one may e.g. prioritize the standard behavior of O over the applicability of (FD) and (DD), thus capturing the intuition that even if they are sometimes to be accepted, deontic conflicts should be avoided whenever possible.

Open problems and further reading The first monotonic dyadic deontic logics were introduced in Bengt Hansson’s seminal paper [46; 75]. Hansson-style dyadic deontic logics typically invalidate (FD), while some of them validate (DD).

More recently, van Benthem, Grossi and Liu have investigated the relation between modal logics of preferences, priority structures, and dyadic deontic logic more generally [97]. In this account, the factual information in the antecedent of (FD) is formalized as a dynamic epistemic event, rather than as a “mere” factual (propositional) statement. This way, the non-monotonicity of reasoning with dyadic obligations is formalized at the object-level, rather than as a property of the consequence relation.

Our focus in this section was on the defeasible application of the detachment principles (FD) and (DD), in a language with both a dyadic operator $O(\cdot \mid \cdot)$ for conditional obligations and an independent, monadic operator O that satisfies full \mathbf{SDL} . We did not discuss other logical properties of $O(\cdot \mid \cdot)$, and instead treated

it as a dummy operator much like we treated the O^P -operator from Section 3. But we may of course wonder whether there are no logical properties which the dyadic operator ought to satisfy unrestrictedly. Possible candidates include, for instance, the dyadic versions of the aggregation and inheritance principles:

$$\begin{aligned} (O(B \mid A) \wedge O(C \mid A)) \supset (O(B \wedge C \mid A)) & \quad (\text{DAgg}) \\ \text{From } O(B \mid A) \text{ and } \vdash B \supset C, \text{ to infer } O(C \mid A) & \quad (\text{DInh}) \end{aligned}$$

However, one has to be careful again, since enriching one’s lower limit logic may easily give rise to flip-flop-problems, analogous to the monadic deontic logics presented in previous sections. The solutions that were discussed in those sections may in turn be transferred to the dyadic setting.

Different preferences regarding the characterization of $O(\cdot \mid \cdot)$ have given rise to a wide variety of dyadic systems, including a range of conflict-tolerant dyadic systems which could in turn be extended adaptively so as to gain further inferential power. For instance, in [90] and [87, Ch. 11], Christian Straßer studied conditional versions of some of the **LUM**-systems from Section 5, and presented a number of adaptive extensions of these logics. In [91] and [87, Chapters 11–12], Straßer presents a general method for turning dyadic deontic logics into ALs which allow for the conditional application of (FD), paying special attention to Chisholm-scenarios.

Finally, it should also be noted that, even if we leave (FD) and (DD) aside, all the observations and techniques from Sections 5–7 could be applied just as well to the case of dyadic deontic logics as developed, building on Goble’s work in [39; 40]. Here again, we may use adaptive logics to steer a middle course between all-too-weak conflict tolerant dyadic systems and deontic explosion.

9.2 Adaptive reasoning with conditionals

Adaptive detachment, generalized Instead of using a binary operator for conditional obligation, one may also introduce a new conditional \Rightarrow , so that the logic of deontic conditionals derives from the logic for this new conditional and the logic for the monadic operator O of one’s choice. In this section we focus on this second approach.

Suppose we formalize “If A , then B is obligatory” as $A \Rightarrow OB$.⁶⁰ Then at the very least we want to be able to factually detach OB given A and $A \Rightarrow OB$, *absent further information*.⁶¹ But we may not want unrestricted detachment (or full modus

⁶⁰One may also represent the conditional obligation “If A , then it is obligatory that B ” by $O(A \Rightarrow B)$ or $OA \Rightarrow OB$. We will have little to say about the first of these two alternatives; we briefly return to the second at the end of this section.

⁶¹We consider deontic detachment at the end of this section.

ponens) for the conditional \Rightarrow . For instance, given the premises $p, q, p \Rightarrow Or$, and $q \Rightarrow O\neg r$, we may not want to be able to detach both Or and $O\neg r$, unless perhaps we move to a non-standard characterization of O . So if we stick to a standard characterization of O as an **SDL**-operator, we will want to allow for some, but not all instances of modus ponens for \Rightarrow .

In other words, we only want to apply detachment in a defeasible way. This can be done as follows in terms of ALs. We first enrich the language of **SDL** with a default conditional, where nested occurrences of \Rightarrow are disallowed:

$$\begin{aligned} \mathcal{W}^{\Rightarrow} := & \mathcal{W}^d \mid \langle \mathcal{W}^d \rangle \Rightarrow \langle \mathcal{W}^d \rangle \mid \neg \langle \mathcal{W}^{\Rightarrow} \rangle \mid \langle \mathcal{W}^{\Rightarrow} \rangle \vee \langle \mathcal{W}^{\Rightarrow} \rangle \mid \langle \mathcal{W}^{\Rightarrow} \rangle \wedge \\ & \langle \mathcal{W}^{\Rightarrow} \rangle \mid \\ & \langle \mathcal{W}^{\Rightarrow} \rangle \supset \langle \mathcal{W}^{\Rightarrow} \rangle \mid \langle \mathcal{W}^{\Rightarrow} \rangle \equiv \langle \mathcal{W}^{\Rightarrow} \rangle \end{aligned}$$

Next, let $\mathbf{SDL}_{\Rightarrow}$ be just **SDL**, but defined over this richer language. Hence, \Rightarrow has no properties in $\mathbf{SDL}_{\Rightarrow}$. We then define our ALs on the basis of $\mathbf{SDL}_{\Rightarrow}$, by the set of abnormalities

$$\Omega_{\Rightarrow} =_{\text{df}} \{(A \Rightarrow B) \wedge A \wedge \neg B \mid A, B \in \mathcal{W}^d\}$$

So whenever the conditional $A \Rightarrow B$ is true and A is true, then we assume that also B is true. Note that A and B can be arbitrary members of \mathcal{W}^d , hence also A can be a deontic statement such as Op – we return to this point below.

Let us call the resulting adaptive logics $\mathbf{SDL}_{\Rightarrow}^x$. As the following proof illustrates, conditional obligations are detachable in $\mathbf{SDL}_{\Rightarrow}^x$ as long as no conflicts are generated. (For the sake of readability, we abbreviate $(A \Rightarrow B) \wedge A \wedge \neg B$ as $A \not\Rightarrow B$.)

1	$p \wedge q$	Prem	\emptyset
2	$p \Rightarrow Or$	Prem	\emptyset
3	$q \Rightarrow O\neg r$	Prem	\emptyset
4	$(p \wedge q) \Rightarrow Os$	Prem	\emptyset
5	Or	1,2;RC	$\{p \not\Rightarrow Or\} \checkmark^8$
6	$O\neg r$	1,3;RC	$\{q \not\Rightarrow O\neg r\} \checkmark^8$
7	Os	1,4;RC	$\{(p \wedge q) \not\Rightarrow Os\}$
8	$(p \not\Rightarrow Or) \vee (q \not\Rightarrow O\neg r)$	1-3;RU	\emptyset

The conditional \Rightarrow of $\mathbf{SDL}_{\Rightarrow}$ is of course very weak – we can only make use of it by going adaptive. We can however strengthen the lower limit logic by adding

further rules. Here are some candidates:

$$\begin{aligned}
 & \text{If } A \Rightarrow C \text{ and } B \Rightarrow C, \text{ then } (A \vee B) \Rightarrow C && \text{(Or)} \\
 & \text{If } A \Rightarrow B \text{ and } B \Rightarrow C, \text{ then } A \Rightarrow C && \text{(Tra)} \\
 & \text{If } A \Rightarrow B \text{ and } (A \wedge B) \Rightarrow C, \text{ then } A \Rightarrow C && \text{(CTra)} \\
 & \text{If } A \vdash B \text{ and } B \Rightarrow C, \text{ then } A \Rightarrow C && \text{(SA)}
 \end{aligned}$$

Each of these rules can be added to our logic if desired. However, one should be careful here, as adding more properties to one’s lower limit logic often generates flip-flop problems, as explained in the previous sections of this paper.

Unlike the dyadic deontic operator of \mathbf{SDL}_d from Section 9.1, the conditional \Rightarrow of $\mathbf{SDL}_{\Rightarrow}^x$ is completely independent of the way we formalize obligations. We can read a statement $A \Rightarrow B$ as ‘If A , then normally B ’ as we would do for defeasible conditionals in general. In $\mathbf{SDL}_{\Rightarrow}^x$ we detach obligations via defeasible modus ponens, just like we defeasibly detach conclusions in default logic or in your preferred calculus of non-monotonic logic. So this approach is very unifying, treating deontic reasoning as just one specific type of defeasible reasoning in general.

However, the approach has the disadvantage that it cannot as easily accommodate deontic detachment (DD) (cf. Section 9.1). Consider the following three inferences:

$$p, p \Rightarrow Oq \vdash Oq \tag{84}$$

$$Op, Op \Rightarrow Oq \vdash Oq \tag{85}$$

$$Op, p \Rightarrow Oq \vdash Oq \tag{86}$$

(84) and (85) are derivable $\mathbf{SDL}_{\Rightarrow}^x$ -rules: we can apply these rules conditionally in $\mathbf{SDL}_{\Rightarrow}^x$. However, (86) is not a derivable rule in $\mathbf{SDL}_{\Rightarrow}^x$. Some have argued that this is how it should be (see e.g. the discussion and references in [25]). Still, (86) has some intuitive force.

One way to defend $\mathbf{SDL}_{\Rightarrow}^x$ is by arguing that, whenever we think deontic detachment should be allowed, the appropriate translation of the conditional is as in (85). More generally, such conditionals are of the form: if A is obligatory, then also B is obligatory ($OA \Rightarrow OB$). However, that would mean that in many cases we need a kind of “double translation” of deontic conditionals – as $(A \Rightarrow OB) \wedge (OA \Rightarrow OB)$ – which seems highly artificial. Moreover, it would go against the spirit of the adaptive logic approach, where the idea is that the logic should determine which applications of deontic detachment are rational. So altogether, it seems that the second approach is less suited to accommodate (DD).

Further reading The literature on the formalization of defeasible conditionals is vast. For some good entry points, see e.g. [55; 61]. In this section we only presented a basic mechanism for the defeasible detachment of obligations via a new conditional. For more information on the types of rules that can be studied via this mechanism, we refer to [87, Chapter 6].

9.3 Adaptive Characterizations of input/output logic

Input/output logic The third approach to deontic conditionals that we will discuss here goes under the name input/output logic (henceforth I/O logic). Technically speaking, I/O logics (without constraints, cf. *infra*) are operations that map every pair $\langle \mathcal{A}, \mathcal{G} \rangle$ to an “output” $\mathcal{O} \subseteq \mathcal{W}$, where (i) $\mathcal{G} \subseteq \mathcal{W} \times \mathcal{W}$ is a set of “input/output pairs” (A, B) ; (ii) $\mathcal{A} \subseteq \mathcal{W}$ is the “input”. For instance, given the input $\mathcal{A} = \{p, q\}$ and the set of conditionals $\mathcal{G} = \{(p, r), (q, s)\}$, the output \mathcal{O} will consist of r , s , and everything that follows from their conjunction.

In a deontic setting, \mathcal{A} usually represents factual information, \mathcal{G} is a set of conditional obligations, and the output consists of what is obligatory, given the facts at hand and given the conditional obligations that make up our normative system. The idea of factual detachment thus lies at the very core of I/O-logics.

Different I/O-logics are obtained by varying on the rules under which \mathcal{G} is closed, before one applies factual detachment. These rules are themselves highly similar to the ones used to characterize default conditionals (cf. Section 9.2). For example, by assuming that \mathcal{G} is closed under the rule (OR)

$$\text{If } (A, C) \text{ and } (B, C), \text{ then } (A \vee B, C) \quad (\text{OR})$$

we can obtain r in the output of $\mathcal{A} = \{p \vee q\}$ and $\mathcal{G} = \{(p, r), (q, r)\}$. Similarly, if \mathcal{G} is closed under the rule (Tra), one can validate deontic detachment (DD):

$$\text{If } (A, B) \text{ and } (B, C), \text{ then } (A, C) \quad (\text{Tra})$$

So for instance, given closure under (Tra), we can obtain q in the output of $\mathcal{A} = \emptyset$ and $\mathcal{G} = \{(\top, p), (p, q)\}$.

Both (FD) and (DD) are accommodated within the I/O-systems presented [58]. However, this framework cannot handle conflicts that arise from the application of (FD) or (DD) or both: e.g. $\mathcal{A} = \{p, q\}$ and $\mathcal{G} = \{(p, r), (q, \neg r)\}$ will generate a trivial output.

To deal with such cases, Makinson and van der Torre introduced a set \mathcal{C} of “constraints” in their [59]. Depending on the application context \mathcal{C} may represent

physical constraints, human rights, practical considerations, etc. \mathcal{C} can restrict the output in two ways, each corresponding to a different style of reasoning. We can require consistency of $\mathcal{O} \cup \mathcal{C}$, or we can impose the weaker requirement that for each $A \in \mathcal{O}$, $\{A\} \cup \mathcal{C}$ is consistent. In the border case where $\mathcal{C} = \emptyset$, this simply means that we require the \mathcal{O} to be consistent, or that each $A \in \mathcal{O}$ is consistent. The first approach is called *meet* constrained output; the second is the *join* constrained output.

The adaptive characterization In [88], I/O-logics are characterized in terms of deductive systems within a rich modal language. We explain how this works for constrained I/O-logics (the case for unconstrained I/O-logics is simpler). The language uses unary modal operators *in*, *out*, *con* to represent input, output, and constraints respectively. Input/output pairs (A, B) are represented by means of *in*, *out* and a conditional \rightarrow , as follows:

$$\text{in}A \rightarrow \text{out}B$$

The principle of detachment and the rules for input/output-pairs are then translated into the object level. This gives us rules and axioms such as the following:

$$\begin{aligned} \text{If } \text{in}A \text{ and } \text{in}A \rightarrow \text{out}B, \text{ then } \text{out}B & \quad (\text{DET}') \\ ((\text{in}A \rightarrow \text{out}C) \wedge (\text{in}B \rightarrow \text{out}C)) \supset (\text{in}(A \vee B) \rightarrow \text{out}C) & \quad (\text{OR}') \\ ((\text{in}A \rightarrow \text{out}B) \wedge (\text{in}B \rightarrow \text{out}C)) \supset (\text{in}A \rightarrow \text{out}C) & \quad (\text{Tra}') \end{aligned}$$

The fact that the output should be consistent with the set of constraints is captured by

$$\text{con}A \supset \neg \text{out} \neg A \quad (\text{ROC})$$

Finally, to mimic the selection of maximal consistent sets of conditionals, a dummy operator \bullet is introduced and used in much the same way as we did in Section 3. That is, conditionals $(A, B) \in \mathcal{G}$ are translated into formulas of the form $\bullet(\text{in}A \rightarrow \text{out}B)$. The adaptive logics then allow one to “activate” such conditionals by removing the dummy, whence one can apply rules like (DET’), (OR’), or (Tra’) to them.

Suppose, for instance, that we are given the following set of inputs, I/O-pairs, and constraints: $\mathcal{A} = \{p, q\}$, $\mathcal{G} = \{(p, r), (q, s), (p, t)\}$, $\mathcal{C} = \{\neg r \vee \neg s\}$. In the language from [88], this gives us the following premise set:

$$\Gamma = \{\text{inp}, \text{inq}, \bullet(\text{inp} \rightarrow \text{outr}), \bullet(\text{inq} \rightarrow \text{outs}), \bullet(\text{inp} \rightarrow \text{outt}), \text{con}(\neg r \vee \neg s)\}$$

In an adaptive proof from Γ , we can finally derive *outt*. Depending on the strategy, we can also finally derive $\text{out}(r \vee s)$ or even *outr* and *outs*.

Let us illustrate this with an object-level proof. To enhance readability, we use $\star(A, B)$ to abbreviate $\bullet(\text{in}A \rightarrow \text{out}B) \wedge \neg(\text{in}A \rightarrow \text{out}B)$. Moreover, we use superscripts r, m to indicate the strategy under which certain lines are (not) marked:⁶²

1	<i>inp</i>	Prem	\emptyset
2	<i>inq</i>	Prem	\emptyset
3	$\bullet(\text{inp} \rightarrow \text{outr})$	Prem	\emptyset
4	$\bullet(\text{inq} \rightarrow \text{outs})$	Prem	\emptyset
5	$\bullet(\text{inp} \rightarrow \text{outt})$	Prem	\emptyset
6	$\text{con}(\neg r \vee \neg s)$	Prem	\emptyset
7	$\text{inp} \rightarrow \text{outr}$	3; RC	$\{\star(p, r)\} \checkmark^{r,m}$
8	$\text{inq} \rightarrow \text{outs}$	4; RC	$\{\star(q, s)\} \checkmark^{r,m}$
9	$\text{inp} \rightarrow \text{outt}$	5; RC	$\{\star(p, t)\}$
10	<i>outr</i>	1,7; RU	$\{\star(p, r)\} \checkmark^{r,m}$
11	<i>outs</i>	2,8; RU	$\{\star(q, s)\} \checkmark^{r,m}$
12	<i>outt</i>	1,9; RU	$\{\star(p, t)\}$
13	$\text{outr} \vee \text{outs}$	10; RU	$\{\star(p, r)\} \checkmark^r$
14	$\text{outr} \vee \text{outs}$	11; RU	$\{\star(q, s)\} \checkmark^r$
15	$\star(p, r) \vee \star(q, s)$	1-4,6; RU	\emptyset

Under the modal translation, the minimal abnormality strategy corresponds to the operation of meet constrained output; normal selections (cf. Section 3.4 corresponds to the join constrained output. The reliability strategy has no counterpart in the original framework of [59]; however, as shown in [88], one can also define a procedural semantics for the corresponding operation, much in the spirit of Makinson and van der Torre's original setting.

Further reading I/O-logic was introduced by Makinson and van der Torre [58; 59] as a formal tool for modeling non-monotonic reasoning with conditionals. We refer to [76] for an introduction to this approach and its applications to deontic reasoning.

The framework presented here is not only sufficient to characterize many well-known I/O logics, but it allows one to go beyond the expressive means of I/O logics so as to express useful notions in deontic logic such as violations and sanctions. We refer to [88] for the many details, and for an elaborate presentation and discussion of these advantages.

⁶²The formulas at lines 10-12 are derivable in view of (DET'). The formula at line 15 is derivable in view of (DET'), modal properties of the **KD**-operator *out*, and the axiom schema (ROC).

10 Deontic compatibility

10.1 Adaptive logics for deontic compatibility

We saw how ALs are useful for reasoning in the presence of normative conflicts, and for detaching conditional obligations. A different context of application for ALs that was mentioned in Section 1 concerns the implementation of the *nullum crimen sine lege* principle (henceforth NCSL). This principle expresses that no crimes occur where there is no law: that which is not forbidden, is permitted. Typically, NCSL is understood as a rule of closure permitting all the actions not prohibited by penal law [1, pp. 142–143]. It is a fundamental principle of law, the roots of which go back at least as far as the French Revolution. In the twentieth century it was incorporated in various human rights instruments as a non-derogable right [70].

Logicians and computer scientists are very familiar with the concept of “negation by default”, according to which a piece of information represented by some variable is taken to be absent unless and until we include it in our database. For instance, where a variable x abbreviates that there is a train leaving for Ghent at 14:14, we may conclude that $\neg x$ unless x is mentioned on the timetable at the train station. Similarly, we can think of NCSL as “permission by default”. Formally, this can be expressed as follows, where we take our premise set Γ to represent a given normative system or law, and where \vdash is an ordinary (Tarskian) deontic logic:

$$\Gamma \vdash PA \text{ iff } \Gamma \not\vdash \neg PA$$

Assume that we want to implement this equivalence against the background of full **SDL**. Then, on pain of inconsistency, the equivalence can at best hold *defeasibly*. Suppose, for instance, that we are given a premise set Γ such that $\Gamma \vdash \neg Pp \vee \neg Pq$, while $\Gamma \not\vdash \neg Pp$ and $\Gamma \not\vdash \neg Pq$. Then we cannot preserve consistency *and* apply NCSL to derive Pp as well as Pq . What we want, then, is a logic that preserves consistency and applies NCSL *as much as possible*.

This motivates an adaptive logic of deontic compatibility which implements NCSL by taking **SDL** as its lower limit logic, and Ω_P as its set of abnormalities:

$$\Omega_P = \{\neg PA \mid A \in \mathcal{W}\}$$

We call the resulting logic **SDL_{nc}^x** with **nc** for *nullum crimen* and $x \in \{r, m\}$. In view of the **SDL**-validity of $PA \vee \neg PA$, **SDL_{nc}^x** allows for the inference of jointly compatible permissions relative to a given premise set. The following object level proof further illustrates the ways this logic works.

1	$O(\neg p \vee \neg q)$	Prem	\emptyset
2	$O(\neg s \wedge t)$	Prem	\emptyset
3	$Pt \supset (Pu \supset O\neg v)$	Prem	\emptyset
4	Pp	RC	$\{\neg Pp\}$
5	$P\neg p$	RC	$\{\neg P\neg p\}$
6	Pq	RC	$\{\neg Pq\}$
7	$P\neg q$	RC	$\{\neg P\neg q\}$
8	Pr	RC	$\{\neg Pr\}$
9	$P\neg r$	RC	$\{\neg P\neg r\}$
10	Ps	RC	$\{\neg Ps\}\checkmark^{18}$
11	$P\neg s$	2; RU	\emptyset
12	Pt	2; RU	\emptyset
13	$P\neg t$	RC	$\{\neg P\neg t\}\checkmark^{19}$
14	Pu	RC	$\{\neg Pu\}\checkmark^{20}$
15	$P\neg u$	RC	$\{\neg P\neg u\}$
16	Pv	RC	$\{\neg Pv\}\checkmark^{20}$
17	$P\neg v$	RC	$\{\neg P\neg v\}$
18	$\neg Ps$	2; RU	\emptyset
19	$\neg P\neg t$	2;RU	\emptyset
20	$\neg Pu \vee \neg Pv$	2;3;RU	\emptyset

One nice feature of this logic is its simplicity, when restricted to premise sets of the form $\{OA \mid A \in \Delta\}$ for $\Delta \subseteq \mathcal{W}$. Indeed, for such cases, the strategies *reliability* and *minimal abnormality* will coincide, since every minimal **Dab**-consequence of such premise sets contains only one disjunct $A \in \Omega_{\mathcal{P}}$. This is itself an immediate corollary of the following:

Proposition 10.1. *If $\Gamma = \{OA \mid A \in \Delta\}$ for $\Delta \subseteq \mathcal{W}$, then $\Gamma \vdash_{\mathbf{SDL}} (\neg PA_1 \vee \dots \vee \neg PA_n)$ iff there is an $i \in \{1, \dots, n\}$ such that $\Gamma \vdash_{\mathbf{SDL}} \neg PA_i$.*

In more complex cases such as our example proof above, the two strategies may well differ. In either case, the resulting consequence set will be closed under **SDL** and consistent.

One may wonder whether the idea of deontic compatibility should necessarily be phrased in terms of the underlying logic **SDL** – after all, legal conflicts are a fact of life, and as soon as such conflicts are modeled in **SDL**, everything becomes obligatory and permissible. This motivates a logic that defeasibly applies NCSL and that accommodates conflicts much as the logics presented in Sections 5-7.

Let us illustrate this by means of the paraconsistent deontic logics from Section 7. One option is to just take a monotonic paraconsistent deontic logic – say **DCLuN**, to keep things relatively simple – and to use as a set of abnormalities

$$\Omega = \{OA \mid A \in \mathcal{W}^\sim\}$$

However, the resulting logic will be too strong, in the sense that it will allow one to derive permissions that should intuitively not be derivable, even if we take NCSL seriously. With such a logic, one can e.g. derive $P_\sqsupset \sim p$ from $\Gamma = \{Op\}$. The underlying reason is that in these logics, Op does not entail $O\neg\sim p$ (just like the truth of p does not entail the falsehood of $\sim p$ in their paraconsistent propositional base), and hence one can consistently assume that $O\neg\sim p$ is false even when Op is true. But the mere fact that we want to allow for the logical possibility of conflicts, should not entail that everything is permissible.

A more plausible combination of conflict-tolerance and *nullum crimen* can be obtained if we combine the *adaptive* logics **DCLuN^x** from Section 7 with NCSL, using the format of lexicographic ALs that was introduced in Section 3.4. This means that the logic first minimizes inconsistencies (which implies i.a. that we derive further obligations), and only after that do we maximize permissions. In this way we can e.g. explain why in view of $\Gamma' = \{Op, O(\sim p \vee q), Or, O\sim r\}$ we can derive Oq , Op and $\neg P_\sqsupset \sim p$, $\neg P_\sqsupset \sim q$, but also $P_\sqsupset s$, $P_\sqsupset \sim s$, and $P_\sqsupset r$, $P_\sqsupset \sim r$.

Analogously, one may enrich the logics from Sections 5 and 6 with a default version of NCSL. For similar reasons as in the paraconsistent case, it seems best to first apply the adaptive mechanisms from those sections, and only after that to apply NCSL. For instance, in the case of non-aggregative deontic logics, we would not want to infer $P\neg(p \wedge q)$ from $\Gamma = \{Op, Oq\}$. Likewise, in the context of the **LUM**-logics, we would not want to infer $P\neg p$ from $\Gamma' = \{O(p \wedge q)\}$. The full development of such rich ALs for deontic compatibility is still very much open; it should by now be clear that a broad range of options are to be considered, and that the devil may well be in the many details.

10.2 Further reading

Adaptive logics for classical compatibility were among the first ampliative adaptive logics to be published – see [13]. Although these logics were not formulated in the standard format, one can do this by means of the triple

$$\langle \mathbf{S5}, \{\neg \diamond A \mid A \text{ is a non-modal formula} \}, x \in \{r, m\} \rangle$$

The relation between classical compatibility and the logics in question is then expressed in terms of a modal translation: A is compatible with Γ iff $\{\Box A \mid A \in \Gamma\} \vdash_{\mathbf{AL}} \diamond A$.

In [65], the basic idea behind these logics is used in order to develop a formal account of paraconsistent compatibility, i.e., what it means that a given formula is compatible with a certain (possibly inconsistent) scientific theory. As Meheus argues there, one also first needs to minimize inconsistencies before checking compatibility with the resulting maximally consistent interpretation of the theory.

11 Summary and outlook

This paper started with two simple adaptive logics that can handle deontic conflicts. We then discussed in some detail more sophisticated conflict-tolerant ALs, as well as ALs for reasoning with conditional obligations and the problems of detachment that are associated with these. Finally, we broadened the picture by presenting ALs for the inherently defeasible *nullum crimen sine lege* principle. This should convince the reader of the generality and the flexibility of the adaptive logic framework.

It is important to realize, however, that this does not exhaust the possibilities of adaptive logics for the domain of normative reasoning. This requires more explanation.

All logics presented in this paper share important constraints. One of them is that we only considered the two main deontic modalities, “it is obligatory that” and “it is permitted that”, and we moreover restricted our formal languages to non-nested occurrences of those modalities. Another one is that we took it for granted that we can start from premise sets that merely consist of very specific and very concrete normative statements, like “Nathan ought to take Lisa to that particular movie on Saturday afternoon”.

Because of these constraints, the logics allow us to explicate only a very small part of the normative reasoning one finds in actual cases. Already the everyday examples from Nathan’s life (that are recognizable to many of us) suffice to illustrate this. In Nathan’s first predicament (the prelude), his normative reasoning does not start from the statements that he ought to take Lisa to the movie in the afternoon, that he ought to look after Ben in the afternoon and that he ought to take Lisa for a veggie burger in the evening. These statements are themselves *derived* from other statements, in this case concrete promises by Nathan and the general rule “One ought to keep one’s promises”. Also in Nathan’s second predicament (Section 3.1), the specific normative statements are not given at the outset, but are the result of reasoning. In this case, not only general rules play a role (like “One ought to return favors”), but also commands uttered by an authority (i.e. Nathan’s father). None of the logics presented here allows us to explicate the reasoning from general rules to their instances or from commands (uttered by one person) to obligations

(for another person) – to mention only two possible origins of specific normative statements.

There is more. Some readers may have noticed that, while presenting our conflict-tolerant logics, we used the term “*prima facie* obligations”, but never used the term “all-things-considered obligations” which is, at least since Ross’ [83], associated with it. Instead we consistently used the term “actual obligations”. The reason is that none of our logics enables us to explicate the reasoning from *prima facie* obligations to all-things-considered obligations, where the latter is taken to mean something like “obligations that are, after careful deliberation, considered to be binding”. Our logics only give us those binding obligations for which relatively little deliberation is needed. For instance, “if a *prima facie* obligation is unconflicted, it should be binding” or “if two *prima facie* obligations are unconflicted, also their conjunction should be binding”, etc.

In order to explicate the reasoning that goes on in resolving a predicament and finding out what one’s all-things-considered obligations are (or should be), we need much more than just deontic operators. For instance, whatever Nathan’s solution for his first predicament may be, it will involve certain beliefs (for instance, what Nathan believes will happen if he does not keep the promise he made to his mother). None of our logics can handle interactions between deontic modalities on the one hand, and doxastic or epistemic modalities on the other.⁶³

Does this mean we have gone all this way for nothing? Certainly not. We are convinced that the logics presented here are good candidates to explicate *part of* the reasoning that goes on in specific deontic contexts. They moreover provide a first stepping stone to more complex, richer accounts of deontic reasoning. So there is still hope for Nathan, or at least for us to fully understand how he should reason.

References

- [1] Carlos E. Alchourrón and Eugenio Bulygin. *Normative Systems*. Springer-Verlag, Wien/New York, 1971.
- [2] Patrick Allo. Adaptive logic as a modal logic. *Studia Logica*, 101(5):933–958, 2013.
- [3] Lennart Åqvist. Deontic logic. In Dov Gabbay and Franz Guenther, editors, *Handbook of Philosophical Logic (2nd edition)*, volume 8, pages 147–264. Kluwer Academic Publishers, 2002.
- [4] D. Batens. Dialectical dynamics within formal logics. *Logique et Analyse*, 114:161–173, 1986.

⁶³See e.g. [73] for a study of the interaction between epistemic and deontic modalities.

- [5] D. Batens. Dynamic dialectical logics. In G. Priest, R. Routley, and J. Norman, editors, *Paraconsistent Logic. Essays on the Inconsistent*, pages 187–217. Philosophia Verlag, München, 1989.
- [6] Diderik Batens. Inconsistencies and beyond. A logical-philosophical discussion. *Revue Internationale de Philosophie*, 200:259–273, 1997.
- [7] Diderik Batens. Inconsistency-adaptive logics. In Ewa Orłowska, editor, *Logic at Work. Essays dedicated to the memory of Helena Rasiowa*, pages 445–472. Physica Verlag (Springer), Heidelberg, New York, 1999.
- [8] Diderik Batens. Zero logic adding up to classical logic. *Logical Studies*, 2:15, 1999.
- [9] Diderik Batens. A general characterization of adaptive logics. *Logique et Analyse*, 173–175:45–68, 2001. Appeared 2003.
- [10] Diderik Batens. A universal logic approach to adaptive logics. *Logica Universalis*, 1:221–242, 2007.
- [11] Diderik Batens. Logics for qualitative inductive generalization. *Studia Logica*, 97:61–80, 2011.
- [12] Diderik Batens. Tutorial on inconsistency-adaptive logics. In Jean-Yves Béziau, Mihir Chakraborty, and Soma Dutta, editors, *Springer Proceedings in Mathematics & Statistics*, volume 152, pages 3–38. Springer, 2015.
- [13] Diderik Batens and Joke Meheus. The adaptive logic of compatibility. *Studia Logica*, 66:327–348, 2000.
- [14] Diderik Batens, Christian Straßer, and Peter Verdée. On the transparency of defeasible logics: Equivalent premise sets, equivalence of their extensions, and maximality of the lower limit. *Logique et Analyse*, 207:281–304, 2009.
- [15] M. Beirlaen. *Tolerating Normative Conflicts in Deontic Logic*. Dissertation, Ghent University, 2012. Available online at <http://www.clps.ugent.be/research/doctoral-dissertations>.
- [16] M. Beirlaen and C. Straßer. Two adaptive logics of norm-propositions. *Journal of Applied Logic*, 11(2):147–168, 2013.
- [17] M. Beirlaen and C. Straßer. Nonmonotonic reasoning with normative conflicts in multi-agent deontic logic. *Journal of Logic and Computation*, 24:1179–1207, 2014.
- [18] M. Beirlaen and C. Straßer. A structured argumentation framework for detaching conditional obligations. In O. Roy, A. Tamminga, and M. Willer, editors, *Proceedings of the 13th International Conference on Deontic Logic and Normative Systems (ΔEON 2016, Bayreuth, Germany)*, pages 32–48. College Publications, 2016.
- [19] Mathieu Beirlaen and Atocha Aliseda. A conditional logic for abduction. *Synthese*, 191(15):3733–3758, 2014.
- [20] Mathieu Beirlaen, Bert Leuridan, and Frederik Van De Putte. A logic for the discovery of deterministic causal regularities. *Synthese*, 195:367–399, 2018.
- [21] Mathieu Beirlaen and Christian Straßer. A paraconsistent multi-agent framework for dealing with normative conflicts. In Joao Leite, Paolo Torroni, Thomas Agotnes, Guido Boella, and Leon van der Torre, editors, *Computational Logic in Multi-Agent*

- Systems*, volume 6814 of *Lecture Notes in Computer Science*, pages 312–329. Springer, Berlin/Heidelberg, 2011.
- [22] Mathieu Beirlaen, Christian Straßer, and Joke Meheus. An inconsistency-adaptive deontic logic for normative conflicts. *Journal of Philosophical Logic*, 42(2):285–315, 2013.
- [23] N. Belnap and M. Perloff. In the realm of agents. *Annals of Mathematics and Artificial Intelligence*, 9:25–48, 1993.
- [24] Patrick Blackburn, Maarten De Rijke, and Yde Venema. *Modal Logic*. Cambridge Tracts in Theoretical Computer Science, 2001.
- [25] Daniel Bonevac. Defaulting on reasons. *Noûs*, 2016.
- [26] D. Brink. Moral conflict and its structure. *The Philosophical Review*, 103:215–247, 1994.
- [27] Fabrizio Cariani. “Ought” and resolution semantics. *Noûs*, 47(3):534–558, 2013.
- [28] J. Carmo and A. Jones. Deontic logic and contrary-to-duties. In Dov Gabbay and Franz Guenther, editors, *Handbook of Philosophical Logic (2nd edition)*, volume 8, pages 265–343. Kluwer Academic Publishers, 2002.
- [29] Brian Chellas. *Modal Logic: an Introduction*. Cambridge: Cambridge university press, 1980.
- [30] R. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 27:33–36, 1963.
- [31] Dag Elgesem. The modal logic of agency. *Nordic Journal of Philosophical Logic*, 2(2):1–46, 1997.
- [32] Kit Fine. Angelic content. *Journal of Philosophical Logic*, 45(2):199–226, April 2016.
- [33] Dov M. Gabbay. Bipolar argumentation frames and contrary to duty obligations, preliminary report. In M. Fisher, L. van der Torre, M. Dastani, and G. Governatori, editors, *Computational Logic in Multi-Agent Systems*, pages 1–24. Springer, 2012.
- [34] L. Goble. A logic of “good”, “should”, and “would”: Part I. *Journal of Philosophical Logic*, 19:169–199, 1990.
- [35] L. Goble. A proposal for dealing with deontic dilemmas. In A. Lomuscio and D. Nute, editors, *7th International Workshop on Deontic Logic in Computer Science*, volume 3065 of *Lecture Notes in Computer Science*, pages 74–113. Springer, 2004.
- [36] L. Goble. Normative conflicts and the logic of ought. *Noûs*, 43:450–489, 2009.
- [37] Lou Goble. A logic of “good”, “should”, and “would”: Part II. *Journal of Philosophical Logic*, 19:253–76, 1990.
- [38] Lou Goble. Multiplex semantics for deontic logic. *Nordic Journal of Philosophical Logic*, 5:113–134, 2000.
- [39] Lou Goble. Preference semantics for deontic logic. Part I: Simple models. *Logique et Analyse*, 183–184:383–418, 2003.
- [40] Lou Goble. Preference semantics for deontic logic. Part II: Multiplex models. *Logique et Analyse*, 185–188:335–363, 2004.

- [41] Lou Goble. A logic for deontic dilemmas. *Journal of Applied Logic*, 3:461–483, 2005.
- [42] Lou Goble. Prima facie norms, normative conflicts, and dilemmas. In Dov Gabbay, Leon van der Torre, John Horty, and Xavier Parent, editors, *Handbook of Deontic Logic and Normative Systems*, volume 1, chapter 4, pages 241–351. College Publications, 2013.
- [43] Lou Goble. Deontic logic (adapted) for normative conflicts. *Logic Journal of the IGPL*, 22(2):206–235, 2014.
- [44] C.W. Gowans, editor. *Moral Dilemmas*. Oxford University Press, 1987.
- [45] J. Hansen. Imperative logic and its problems. In Dov Gabbay, Leon van der Torre, John Horty, and Xavier Parent, editors, *Handbook of Deontic Logic and Normative Systems*, volume 1, chapter 2, pages 137–192. College Publications, 2013.
- [46] Bengt Hansson. An analysis of some deontic logics. *Nous*, 3:373–398, 1969.
- [47] Jesse Heyninck and Christian Straßer. Relations between assumption-based approaches in nonmonotonic logic and formal argumentation. In Gabriele Kern-Isberner and Renata Wassermann, editors, *16th International Workshop on Non-Monotonic Reasoning, Cape Town, South Africa*, pages 65–76, 2016.
- [48] Risto Hilpinen and Paul McNamara. Deontic logic: a historical survey and introduction. In Dov Gabbay, Leon van der Torre, John Horty, and Xavier Parent, editors, *Handbook of Deontic Logic and Normative Systems*, volume 1, chapter 1, pages 3–136. College Publications, 2013.
- [49] J. Horty. Moral dilemmas and nonmonotonic logic. *Journal of Philosophical Logic*, 23(1):35–66, 1994.
- [50] J. Horty. Nonmonotonic foundations for deontic logic. In Donald Nute, editor, *Defeasible Deontic Logic: Essays in Nonmonotonic Normative Reasoning*, pages 17–44. Kluwer Academic Publishers, 1997.
- [51] J. Horty. Skepticism and floating conclusions. *Artificial Intelligence*, 135:55–72, 2002.
- [52] J. Horty. Reasoning with moral conflicts. *Nous*, 37:557–605, 2003.
- [53] J. Horty. *Reasons as Defaults*. Oxford University Press, 2012.
- [54] Frank Jackson. On the semantics and logic of obligation. *Mind*, 94:177–195, 1985.
- [55] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.
- [56] Bert Leuridan. Causal discovery and the problem of ignorance. An adaptive logic approach. *Journal of Applied Logic*, 7(2):188–205, 2009.
- [57] D. Makinson and K. Schlechta. Floating conclusions and zombie paths: two deep difficulties in the “directly skeptical” approach to defeasible inheritance nets. *Artificial Intelligence*, 48:199–209, 1991.
- [58] D. Makinson and L. van der Torre. Input/output logics. *Journal of Philosophical Logic*, 29:383–408, 2000.
- [59] D. Makinson and L. van der Torre. Constraints for input/output logics. *Journal of Philosophical Logic*, 30:155–185, 2001.

- [60] David Makinson. General patterns in nonmonotonic reasoning. In *Handbook of Logic in Artificial Intelligence and Logic Programming, vol. III*. Clarendon Press, 1994.
- [61] David Makinson. *Bridges from Classical to Nonmonotonic Logic*, volume 5 of *Texts in Computing*. King’s College Publications, London, 2005.
- [62] Ruth Barcan Marcus. Moral dilemmas and consistency. *Journal of Philosophy*, 77:121–136, 1980. Reprinted in [44].
- [63] C. McGinnis. *Paraconsistency and Deontic Logic: Formal Systems for Reasoning with Normative Conflicts*. Dissertation, University of Minnesota, 2007.
- [64] C. McGinnis. Semi-paraconsistent deontic logic. In Jean-Yves Béziau, Walter Carnielli, and Dov Gabbay, editors, *Handbook of Paraconsistency*, pages 81–99. College Publications, London, 2007.
- [65] Joke Meheus. Paraconsistent compatibility. *Logique et Analyse*, 183–184:251–287, 2003.
- [66] Joke Meheus, Mathieu Beirlaen, and Frederik Van De Putte. Avoiding deontic explosion by contextually restricting aggregation. In Guido Governatori and Giovanni Sartor, editors, *Deontic Logic in Computer Science*, volume 6181 of *Lecture Notes in Computer Science*, pages 148–165. Springer Berlin Heidelberg, 2010.
- [67] Joke Meheus, Mathieu Beirlaen, Frederik Van De Putte, and Christian Straßer. Non-adjunctive deontic logics that validate aggregation as much as possible. *Journal of Applied Logic (conditionally accepted)*, 2012. Preprint available at <http://www.clps.ugent.be/research/publications>.
- [68] Joke Meheus, Christian Straßer, and Peter Verdée. Which style of reasoning to choose in the face of conflicting information? *Journal of Logic and Computation*, 26(1):361–380, 2016.
- [69] Joke Meheus, Liza Verhoeven, Maarten Van Dyck, and Dagmar Provijn. Ampliative adaptive logics and the foundation of logic-based approaches to abduction. In L. Magnani, N.J. Nersessian, and Claudio Pizzi, editors, *Logical and Computational Aspects of Model-Based Reasoning*, pages 39–71. Kluwer Academic, Dordrecht, 2002.
- [70] Ali Mokhtar. Nullum crimen, nulla poena sine lege: Aspects and prospects. *Statute Law Review*, 26(1):41–55, 2005.
- [71] D. Nute. Norms, priorities, and defeasibility. In Paul McNamara and Henri Prakken, editors, *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science.*, pages 201–218. IOS Press, 1999.
- [72] Sergei Odintsov and Stanislav Speranski. Computability issues for adaptive logics in expanded standard format. *Studia Logica*, 101(6):1237–1262, 2013.
- [73] Eric Pacuit, Rohit Parikh, and Eva Cogan. The logic of knowledge based obligation. *Synthese*, 149(2):311–341, 2006.
- [74] X. Parent and L. van der Torre. “Sing and dance!” Input/output logics without weakening. In F. Cariani, D. Grossi, J. Meheus, and X. Parent, editors, *DEON (12th International Conference on Deontic Logic in Computer Science)*, volume 8554 of *Lecture Notes in Artificial Intelligence*, pages 149–165. Springer, 2014.

- [75] Xavier Parent. A complete axiom set for Hansson’s deontic logic DSDL2. *Logic Journal of the IGPL*, 18(3):422–429, 2010.
- [76] Xavier Parent and Leendert van der Torre. Input/output logic. In Dov Gabbay, Jeff Horty, Xavier Parent, Ron van der Meyden, and Leendert van der Torre, editors, *Handbook of Deontic Logic and Normative Systems*, volume 1, chapter 8, pages 499–544. College Publications, 2013.
- [77] H. Prakken and G. Sartor. Law and logic: A review from an argumentation perspective. *Artificial Intelligence*, 227:214–245, 2015.
- [78] Henri Prakken. Intuitions and the modelling of defeasible reasoning: some case studies. In *Proceedings of the Ninth International Workshop on Nonmonotonic Reasoning*, pages 91–99, Toulouse, 2002.
- [79] G. Priest. Paraconsistent logic. In D. Gabbay and F. Guentchner, editors, *Handbook of Philosophical Logic (2nd edition)*, volume 8, pages 287–393. Kluwer Academic Publishers, 2002.
- [80] G. Priest, K. Tanaka, and Z. Weber. Paraconsistent logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2015 edition, 2015. <http://plato.stanford.edu/archives/spr2015/entries/logic-paraconsistent/>.
- [81] Graham Priest. *In Contradiction. A Study of the Transconsistent*. Nijhoff, Dordrecht, 1987.
- [82] N. Rescher and R. Manor. On inferences from inconsistent premises. *Theory and Decision*, 1:179–217, 1970.
- [83] W. David Ross. *The Right and the Good*. Clarendon Press, 1930.
- [84] 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.
- [85] K. Segerberg. An essay in classical modal logic, 1971.
- [86] K. Segerberg. Getting started: beginnings in the logic of action. *Studia Logica*, 51:347–378, 1992.
- [87] C. Straßer. *Adaptive Logic and Defeasible Reasoning. Applications in Argumentation, Normative Reasoning and Default Reasoning*. Springer, 2014.
- [88] C. Straßer, M. Beirlaen, and F. Van De Putte. Dynamic proof theories for input/output logic. *Studia Logica*, 104:869–916, 2016.
- [89] C. Straßer, A. Knoks, and Joke Meheus. Deontic reasoning on the basis of consistency considerations. Under review, 2017.
- [90] Christian Straßer. An adaptive logic framework for conditional obligations and deontic dilemmas. *Logic and Logical Philosophy*, 19(1-2):95–128, 2010.
- [91] Christian Straßer. A deontic logic framework allowing for factual detachment. *Journal of Applied Logic*, 9:61–80, 2011.
- [92] Christian Straßer and Ofer Arieli. Normative reasoning by sequent-based argumentation. *Journal of Logic and Computation*, 2015 (online first).
- [93] Christian Straßer and Mathieu Beirlaen. Towards more conflict-tolerant deon-

- tic logics by relaxing the interdefinability between obligations and permissions. Preprint available at <http://www.clps.ugent.be/research/publications>, unpublished manuscript.
- [94] Christian Straßer, Joke Meheus, and Mathieu Beirlaen. Tolerating deontic conflicts by adaptively restricting inheritance. *Logique et Analyse*, 219:477–506, 2012.
 - [95] Christian Straßer and Dunja Šešelja. Towards the Proof-theoretic Unification of Dung’s Argumentation Framework: an Adaptive Logic Approach. *Journal of Logic and Computation*, 21:133–156, 2010.
 - [96] Johan van Benthem. What one may come to know. *Analysis*, 64(282):95–105, 2004.
 - [97] Johan van Benthem, Davide Grossi, and Fenrong Liu. Priority structures in deontic logic. *Theoria*, 80(2):116–152, 2014.
 - [98] Frederik Van De Putte. *Generic Formats for Prioritized Adaptive Logics. With Applications in Deontic Logic, Abduction and Belief Revision*. Dissertation, Ghent University, 2012. Available at <http://www.clps.ugent.be/research/doctoral-dissertations>.
 - [99] Frederik Van De Putte. Default assumptions and selection functions: A generic framework for non-monotonic logics. In Felix Castro, Alexander Gelbukh, and Miguel Gonzalez, editors, *Advances in Artificial Intelligence and Its Applications*, volume 8265 of *Lecture Notes in Computer Science*, pages 54–67. Springer, 2013.
 - [100] Frederik Van De Putte. Coarse Deontic Logic. In Allard Tamminga and Malte Willer, editors, *Deontic Logic and Normative Systems: 13th International Conference, DEON 2016, Bayreuth, Germany*, pages 256–271. College Publications, July 2016.
 - [101] Frederik Van De Putte. Coarse Deontic Logic (Extended Version). Under review, 2016.
 - [102] Frederik Van De Putte and Christian Straßer. Extending the standard format of adaptive logics to the prioritized case. *Logique et Analyse*, 220:601–641, 2012.
 - [103] Frederik Van De Putte and Christian Straßer. A logic for prioritized normative reasoning. *Journal of Logic and Computation*, 23(3):563–583, 2013.
 - [104] Frederik Van De Putte and Christian Straßer. Adaptive logics: a parametric approach. *Logic Journal of IGPL*, 22(6):905–932, 2014.
 - [105] L. van der Torre and S. Villata. An ASPIC-based legal argumentation framework for deontic reasoning. In *Computational Models of Argument (Proceedings of COMMA 14)*, pages 421–432. IOS Press, 2014.
 - [106] J.A. van Eck. A system of temporally relative modal and deontic predicate logic and its philosophical applications. *Logique et Analyse*, 99:249–290, 1982.
 - [107] Bas C. van Fraassen. Values and the heart’s command. *Journal of Philosophy*, 70(1):5–19, 1973.
 - [108] Peter Verdée. Non-monotonic set theory as a pragmatic foundation of mathematics. *Foundations of Science*, 18(4):655–680, Nov 2013.
 - [109] Georg Henrik von Wright. Deontic logic. *Mind*, 60:1–15, 1951.
 - [110] P. Vranas. I ought, therefore I can. *Philosophical Studies*, 136:167–216, 2007.

- [111] Bernard Williams and W. F. Atkinson. Symposium: Ethical consistency. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 39:103–138, 1965.