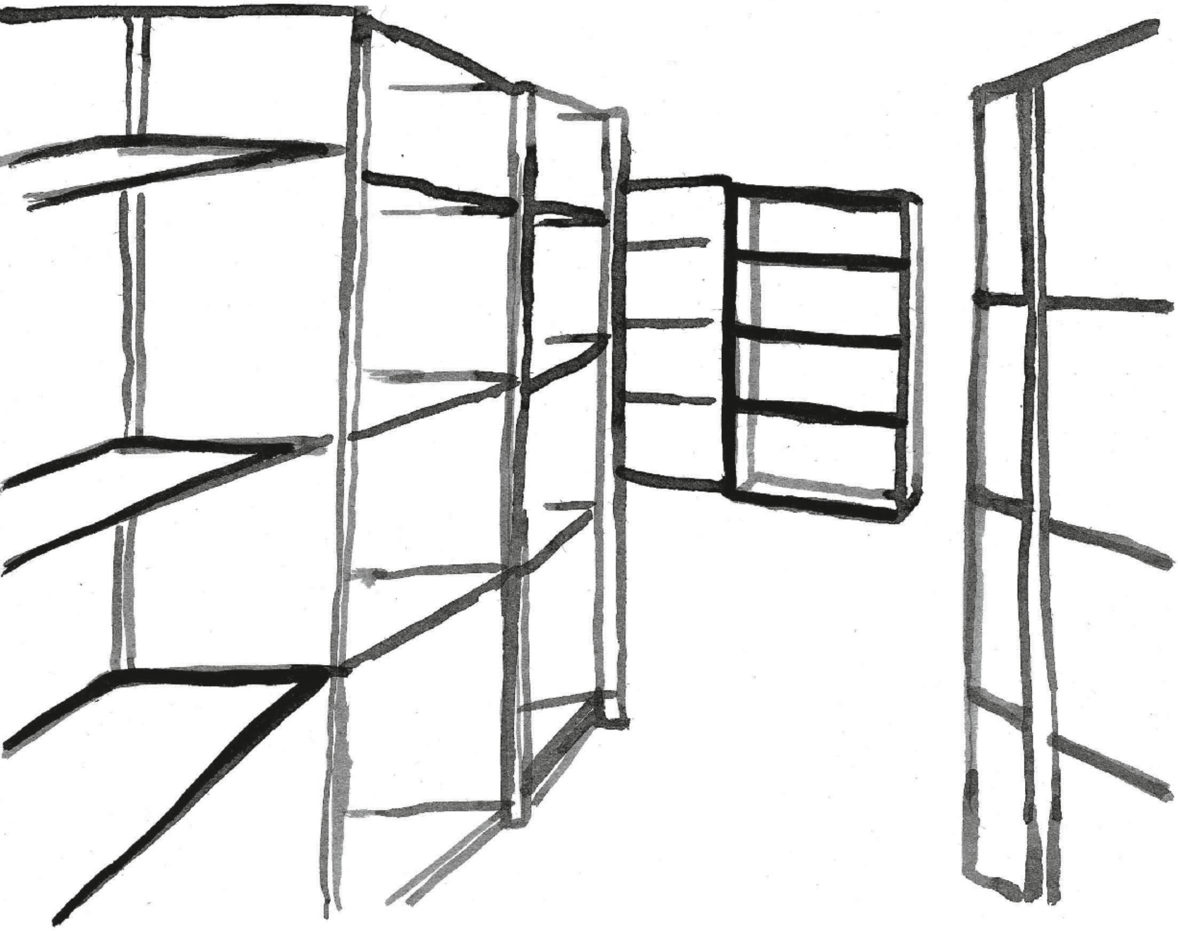
UNIVERSITEIT GENT



Tolerating Normative Conflicts in Deontic Logic

Mathieu Beirlaen

Proefschrift voorgedragen tot het bekomen van de graad van Doctor in de Wijsbegeerte
Promotoren: Prof. Dr. Joke Meheus en Dr. Christian Straßer

Submitted July 4th, 2012.

Supervisors:

Prof. Dr. Joke Meheus, Ghent University
Dr. Christian Straßer, Ghent University

Reading Committee:

Prof. Dr. Diderik Batens, Ghent University
Prof. Dr. Jan Broersen, Utrecht University
Prof. Dr. Allard Tamminga, University of Groningen / University of Oldenburg
Prof. Dr. Peter Verdée, Ghent University

UNIVERSITEIT GENT

Faculteit Letteren en Wijsbegeerte

Mathieu Beirlaen

Tolerating Normative Conflicts in Deontic Logic

Proefschrift voorgedragen tot het bekomen van de graad van
Doctor in de Wijsbegeerte

Promotoren: Prof. Dr. Joke Meheus en Dr. Christian Straßer

Acknowledgments

I would like to thank the following people:

- My supervisors Joke Meheus and Christian Straßer, and my former supervisor Diderik Batens. I thank Joke, Christian and Dirk for the time and effort they invested in my training as a logician, and for their ever critical attitude towards my thoughts and writings.
- My colleagues and collaborators at the Centre for Logic and Philosophy of Science, for the many discussions on logic, philosophy and anything-but-logic-and-philosophy. Special thanks go out to Frederik Van De Putte and Tjerk Gauderis. I am a proud co-author of Frederik, and a proud co-organizer of Tjerk.
- My friends and family, for providing the support and balance needed to endure life as a graduate student.
- Gitte Callaert, for helping out with the design of this book.

This thesis was typeset in LaTeX.
The artwork on the cover is by Dries Warlop:
`http://users.telenet.be/drieswarlop`.

Introduction

Subject and aims of the thesis

There are strong pragmatic grounds for organizing our norms in a coherent way, and, more generally, for reasoning in a way that is both inherently consistent and in accordance with our perception of the world. However, there is no denying that we sometimes face conflicts or requirements that we are unable to live up to. The god of the protestants, for instance, provided mankind with a law it was unable to keep, and then righteously damned all sinners for failing to keep it [104]. For a more 'earthly' feel, just replace this god with some legislator or commander-in-chief.

When confronted with a conflict of norms, we do not just give up on the constraints of consistency and coherence. We are not in general at loss as to how to act. In situations in which we have no uniquely action-guiding principle, we simply try to reason onwards and do what we can.

The general aim of this thesis is to show in a formally precise way how we can tolerate normative conflicts without these rendering our moral/ethical/legal theories useless. I will argue that we can behave in a perfectly rational way despite the presence of irresolvable conflicts, and in some cases even contradictions, in the norms that are supposed to guide our behavior. My toolbox in doing so consists of a set of logics. With Harry Gensler, I believe that logic helps us clarify, understand, and evaluate:

> Logic can help us understand our moral reasoning - how we go from premises to a conclusion. It can force us to clarify and spell out our presuppositions, to understand conflicting points of view, and to identify weak points in our reasoning. Logic is a useful discipline to sharpen our ethical thinking. [61, p. 38]

The branch of formal logic that studies our normative concepts is called *deontic logic*. Over the last decades several deontic logicians have tried, with varying degrees of success, to accommodate normative conflicts in their formal calculi. Their proposals vary in the formalisms used and in the rules for normative reasoning that are given up or restricted in order to accommodate conflicting normative directives.

I will present a new way of tackling the problem of tolerating normative conflicts in deontic logic. My approach is pluralist and contextual in spirit: I believe that different strategies and different degrees of conflict-tolerance are called for in different normative contexts. For instance, in the context of legal norms we need a logic that differs significantly from the one we need for accommodating conflicts in a moral context.

Despite its pluralist flavor, this approach will be spelled out and made precise within one and the same logical framework: the standard format for *adaptive logics*. The main technical merits of this thesis lie in its presentation and defense of various adaptive deontic logics that (i) are capable of tolerating normative conflicts, and (ii) are sufficiently powerful to account for our everyday normative inferencing. In their treatment of the trade-off between (i) and (ii), these adaptive systems outperform their competitors from the literature.

The main philosophical relevance of this thesis lies in its powerful formal clarification of the idea that we can reason logically and coherently despite the undeniable fact that every once in a while we are confronted with an unresolvable normative conflict. Moreover, given recent developments in the fields of artificial intelligence and legal science, the logics defined here may also serve a more practical purpose as specification devices for the development of 'ethical' computer programs and artificial legal reasoners.

Before I outline the general structure of the thesis, two more remarks are in order: First, all logics presented here are propositional monadic modal logics. Given the limited expressive resources of such logics, I am not doing justice to the complexity of the world when I say that the logics presented here can "account for our everyday normative reasoning". Surely, formalizing our *actual* normative reasoning would require additional resources in our formal language (at the very least this would require polyadic operators, predicates, etc.).

With Segerberg, I agree that "working with logical techniques pushes the requirement of rigour so high that pressures of complexity enforce a very narrow focus" [159, pp. 347-348]. Because of this narrow focus, additional expressive resources are not considered in this thesis. Hence the reader should be warned that when mentioning 'our everyday (normative) reasoning' I make abstraction of extra resources in the language and hence target a rather simplified account of our everyday reasoning.

Second, since many of the logics presented in this thesis result from collaborations with various colleagues, chapters 1-7 are written in the 'we' form for reasons of uniformity. Sections based on joint work are indicated at the beginning of each chapter.

Structure of the thesis

This thesis is structured as follows:

Chapter 1 introduces some qualifications and terminological distinctions that will be used throughout the thesis. It contains a first (informal) characterization of what a normative conflict is, as well as a motivation for devising conflict-tolerant deontic logics.

Chapter 2 is concerned with a presentation and discussion of the system best known as *Standard Deontic Logic*. This system is formally characterized and discussed at some length, with special attention to its treatment of normative conflicts and its well-known problems or 'paradoxes'.

In Chapter 3 we turn to the problem of accommodating normative conflicts in deontic logic, and discuss a number of strategies for preventing conflicts from rendering our premises trivial. From this discussion we distil a number of desiderata for adequate conflict-tolerant deontic logics, against which we will evaluate the logics presented later on.

In Chapter 4 we present the standard format for adaptive logics. The standard format provides a generic, unifying framework within which all adaptive logics presented here are defined. Logics characterized within this framework automatically inherit a dynamic proof theory, a characteristic semantics and a number of meta-theoretical properties. We illustrate each of these by means of a concrete example.

Chapters 5-7 contain the presentation and illustration of a number of adaptive conflict-tolerant deontic logics that are argued to meet the desiderata given in Chapter 3. Each of these chapters has a different focus.

In Chapter 5, we assess two logics that restrict the application of the rule that allows us to aggregate two or more obligations to a single one. The first system is inspired by Bernard Williams' characterization of the structure of moral conflict, and was already defined in Chapter 4. The second system has its roots in Sir David Ross' distinction between prima facie and all-things-considered obligations.

In Chapter 6, we turn to a different strategy for devising conflict-tolerant adaptive deontic logics. The logics defined in this chapter are built 'on top' of a logic that invalidates the *ex contradictione quodlibet* principle according to which a contradiction trivializes our premise set. Adaptive logics built 'on top' of such logics are usually called inconsistency-adaptive logics. We present two such systems. The first one allows for inconsistencies inside as well as outside the scope of its deontic operators. The second one allows for inconsistencies inside, but not outside the scope of its deontic operators. Moreover, the second system also invalidates the excluded middle principle inside the scope of its deontic operators.

Chapter 7 builds on the ideas presented in Chapter 6, and adds some expressive power to the picture. In it, we present an inconsistency-adaptive deontic logic capable of representing the actions of multiple agents. As such, the adaptive logic presented in this chapter is capable of modeling normative conflicts between different (groups of) agents.

We end this thesis with some concluding remarks on the merits of the systems presented earlier.

Contents

Chapter 1

Reasoning with normative conflicts

> The present world, and those worlds we should think we could bring about, are worlds of conflict
>
> Michael Stocker [168, p. 125]

✎ I am indebted to Joke Meheus and Christian Straßer for valuable comments on this chapter.

In this chapter, we introduce some terminology and some important distinctions that often surface in the literature on normative conflicts and/or deontic logic. Qualifications made here are used for the delineation of the topic and aim of this thesis. Moreover, the discussion in this chapter serves as a basis for the motivation of the logics defined in later chapters.

In Section 1.1 we informally characterize the concept 'norm', and briefly comment on the history and subject of deontic logic. In Section 1.2 we elaborate on some further distinctions often made when discussing the logical properties of norms.

In Section 1.3 we provide a first, intuitive, characterization of a normative conflict. We do so by means of a number of examples from the literature. Some further distinctions and qualifications often made in the literature are discussed in Section 1.4.

We turn to the debate on normative conflicts in a moral context in Section 1.5. Drawing on some of the nuances made earlier on, we defend the existence of so-called 'moral dilemmas'. In Section 1.6 we turn to normative conflicts in general. We defend both the existence of irresolvable normative conflicts and the need for accounting for them in our systems of deontic logic. Finally, we comment on the problems relating to the assignment of truth-values to norms (Section 1.7) and state some preliminaries on notation that will be used throughout this thesis (Section 1.8).

1.1 Norms and deontic logic

Norms can be conceived as directives that are issued by a norm-authority to direct the behavior of norm-subjects. As examples of norms, we can think of military commands, orders and permissions given by parents to children, traffic laws issued by a magistrate, etc. Norms can also be self-directed, issued and aimed at directing one's own behavior. More generally still, norms may arise from institutions, from traditions or religions.[1] For our present aims, we need not settle on a more precise definition of what a norm is. Instead, we will illustrate our claims by means of transparent examples and, where necessary, make clear which kind or type of norm we have in mind.

Norms appear to come in three main varieties: obligations, prohibitions, and permissions. Table 1.1 summarizes some roughly equivalent ways of expressing such norms (examples taken from [124, p. 3]).

Obligations	You ought to (should, must) attend the meeting. You have an obligation (duty) to attend. It is obligatory (required, mandatory, compulsory) that you attend. It ought to be (the case) that you attend. You are obligated (obliged, required) to attend.
Prohibitions	You are forbidden to attend. You are prohibited from attending. It is forbidden (prohibited) that you attend. You ought not to attend. You may not attend.
Permissions	You may attend. You are permitted (allowed, authorized, licensed, at liberty, free) to attend. It is permissible (acceptable, okay, legal) for you to attend. It is permitted (okay, acceptable) that you attend. You have permission to attend.

Table 1.1: Expressing obligations, prohibitions, and permissions.

Deontic logic is concerned with the logical properties displayed by these concepts and with the logical relations between them. More broadly, deontic logic can be seen as the logical study of the normative use of language [8].

The formal study of norms was stimulated in the previous century by Ernst Mally's 1926 monograph *Grundgesetze des Sollens: Elemente der Logik des Willens* [119] and by G.H. von Wright's 1951 article *Deontic Logic* [190]. Especially the latter was very influential, as it contains the building blocks of the system that is now known as *Standard Deontic Logic* (cfr. Chapter 2). It is hard to say

[1] More broadly still, norms are sometimes also taken to include rules, e.g. the rules of grammar, the rules of chess, and customs [192]. We stick to the more narrow conception here.

how far exactly the history of deontic logic dates back. For more details on the origin and emergence of this field, we refer to [106].

1.2 Some distinctions

1.2.1 Norms and agency

Consider the sentences "The window ought to be open" and "Someone ought to open the window". In the former sentence, the obligation holds of a state of affairs, whereas in the latter the obligation holds of an action. In the literature, this distinction is known as that between "*ought to be*" and "*ought to do*" (see e.g. [193]). Typically, "ought to do" statements are assumed to involve agency.

For reasons that we cannot yet spell out in detail, we find the "ought to be" and "ought to do" reading of deontic operators suboptimal. Instead, we opt to interpret obligations simply as statements of the form "it is obligatory that". We argue for this reading in Section 2.4.5, where we discuss nested occurrences of deontic operators.

The distinction between agentive and non-agentive norms remains in full force under this new reading. In a sentence of the form "It is obligatory that A" or "It is permitted that A", the term A may or may not be agentive. Compare "It is obligatory that the window remains open" to "It is obligatory that John sees to it that the window remains open".

For reasons of generality and convenience, we will forget about the distinction between agentive and non-agentive terms throughout most of this thesis, and assume that the formalisms that we present can be augmented with formal means for explicitly representing the notion of agency if necessary. In Chapter 7 we return to this problem, and illustrate how some of the systems presented earlier on can be enriched so as to explicitly represent agentive formulas.

1.2.2 Prescriptions and descriptions

In ordinary language, normative sentences exhibit a characteristic ambiguity. The very same words may be used to enunciate a norm (give a prescription) and to make a statement about norms (description) [192, pp. 104-106]. In deontic logic, it is important to carefully distinguish between this prescriptive and descriptive use of norms.

The distinction between prescriptions and descriptions is that between norms themselves and statements *about* norms. In what follows, we take the term *norm* to denote the former (prescriptive), and *norm-proposition* to denote the latter (descriptive) interpretation of normative statements.[2]

For now, it suffices to see that when norms are given or issued, we use prescriptions. When we report on or describe already existing norms, we use descriptions. We return to the distinction between norms and norm-propositions in Section 1.7.

[2]Von Wright [192] and Åqvist [8] cite Ingemar Hedenius as the first philosopher to note the distinction between norms and norm-propositions. According to Hedenius, norms are "genuine", and norm-propositions are "spurious" deontic sentences [85]. The distinction between norms and norm-propositions was later also drawn – among others – by Wedberg [199], Stenius [165], Alchourrón [1], and Hansson [78] (see also [8]).

1.2.3 Further distinctions

In what follows, we will sometimes distinguish between moral norms, legal norms, commands, etc. when motivating the context of applicability of some of the logics presented in this thesis. However, in general the level of analysis that we are concerned with is that of norms simpliciter, as specified in Section 1.1. Unless stated explicitly, we will make abstraction of more refined distinctions occurring 'inside' norms. We already mentioned that, throughout most of the thesis, we will not be concerned with the distinction between agentive and non-agentive norms. Similarly, we will not make a formal distinction between binding and non-binding norms, epistemic and non-epistemic norms (see e.g. [141]), etc. Interesting as these distinctions may be, they fall outside the scope of this thesis.

1.3 Normative conflicts

1.3.1 Intuitive characterization

Intuitively, a *normative conflict* occurs whenever we find ourselves in a situation in which our normative directives are inconsistent or not uniquely action-guiding in the sense that we are permitted or even obliged to do something that is forbidden. We say that a normative conflict is *escapable* if the conflict does not require us to violate any of our obligations. Otherwise the conflict is *inescapable*.

Example 1. Agamemnon is told by a seer that he must sacrifice his daughter to satisfy a goddess who is delaying his expedition against Troy. As a commander, Agamemnon ought to sacrifice his daughter in order to further the expedition. However, as a father, Agamemnon ought not to kill his daughter [203].

Example 2. According to his religious beliefs, Yilmaz is prohibited to drink alcohol. However, according to the laws of his country, he is permitted to drink alcohol.

In Example 1, Agamemnon faces an inescapable normative conflict since he cannot possibly satisfy both his obligations as a commander and his obligations as a father. In Example 2, Yilmaz faces an escapable normative conflict, since he can satisfy all of his obligations by not drinking alcohol.

Next to the distinction between escapable and inescapable conflicts, we can draw many more distinctions between 'types' of normative conflicts. We elaborate on some of these in Section 1.4. First, let us look at some more examples.

1.3.2 More examples

The examples presented in this section are chosen in function of the discussions that follow. A more comprehensive list of examples of normative conflicts (in no particular order) is contained in Appendix A.

1.3.2.1 Tragic fiction

In discussions on normative conflicts, authors often cite ready-made fictional examples from popular culture where tailor-made constructions involving conflicting moral obligations are presented in an often very dramatic setting. Nonetheless, the very possibility of such situations is a forceful argument in defense of the need for taking normative conflicts into account:

> Theories in which moral dilemmas are advanced in practical decision-making standardly draw on the literature of tragic fiction, and ingenious but farfetched imaginary obligation scenarios in which agents are caught between identically forceful contrary moral requirements. However remote and improbable, like the degenerate undecidable constructions in the metatheory of mathematical logic, the mere possibility of moral dilemmas challenges the consistency and completeness of systematic ethical judgment. [98, p. 43]

Example 3. In *Sophie's Choice*, a novel by William Styron, Sophie arrives with her two children at a Nazi concentration camp. A guard asks her to choose one child, and he tells her that the child she chooses will be killed, and the other child will live in the children's barracks. Sophie does not want to choose at all, but the guard tells her that, if she refuses to choose, both children will be killed [164].

Example 4. A person falls overboard from a ship in a wartime convoy; if the captain of the ship leaves his place in the convoy to pick him up, he puts the ship and all on board at risk from submarine attack; if he does not, the person will drown. In the film *The Cruel Sea*, a somewhat similar case occurs; the commander of a corvette is faced with a situation in which if he does not drop depth charges the enemy submarine will get away to sink more ships and kill more people; but if he does drop them he will kill the survivors in the water. In fact he drops them, and is depicted in the film as suffering anguish of mind [83, p. 29].

1.3.2.2 From the newspaper

The examples below illustrate that normative conflicts not only occur in tragic literature, but also in real life.

Example 5. SWIFT is a Belgium-based company with offices in the United States that operates a worldwide messaging system used to transmit, inter alia, bank transaction information. According to the U.S. Treasury, information derived from the use of SWIFT data has enhanced the United States' and third countries' ability to identify financiers of terrorism, to map terrorist networks and to disrupt the activities of terrorists and their supporters. However, in September 2006 the Belgian Data Protection Authority stated that SWIFT processing activities for the execution of interbank payments are in breach of Belgian data protection law. American diplomats and politicians claim that SWIFT ought to continue passing information to the U.S. Treasury, whereas according to Belgian law SWIFT ought not to pass this information, since this activity is in breach of Belgian data protection law.

Example 6. A team of Dutch scientists of the Erasmus Medical Center led by the virologist Ron Fouchier has created a highly contagious variant of the H5N1 ("bird flu") virus. The scientists have submitted their results for publication in *Science*, claiming that they have positively answered the question whether or not the H5N1 virus can possibly trigger a pandemic by mutating into a more transmissible variant.

On the one hand, many virologists support the publication of these results due to their potential benefits for public health. According to Fouchier, the U.S. National Institute of Health (NIH) has agreed to the publication of his team's results. On the other hand, representatives of the U.S. Government fear that the publication of the study will give terrorists new knowledge for constructing bio-weapons of mass destruction.

On December 20^{th} 2011, the U.S. National Science Advisory Board for Biosecurity (NSABB) ruled that all technical details must be left out for publication. The journals *Science* and *Nature* opposed this decision. Eventually, on March 30^{th} 2012, the NSABB revised its stance after a two-day meeting during which its members decided (after voting) that the full paper can be published after all.[3]

The following case describes a more 'tragic' real-life normative conflict.

Example 7. During the Battle of Britain, Churchill was faced with the following choice. Thanks to the British government's access to Germany's secret codes, he was informed in advance of many planned German air raids on populated areas. He could evacuate those areas, sparing many innocent lives, but doing so would, with a significant degree of probability, reveal to the Germans that their codes had been broken, seriously impairing the British war effort. He decided not to evacuate these areas [108, p. 214].

1.4 Important qualifications

1.4.1 Prima facie obligations and all-things-considered obligations

Inspired by Kant's distinction between *perfect* and *imperfect* obligations according to which only the imperfect ones admit of exceptions, moral philosophers sometimes distinguish between so-called *prima facie* duties on the one hand, and *actual, proper, all-things-considered duties* or *duties sans phrase* on the other. This terminology first arises in the context of the moral dilemmas debate (cfr. Section 1.5). The term 'prima facie duty' was coined by Sir David Ross in 1930 [153]. Against the utilitarians, Ross argued that (actual) duties are highly personal acts that arise in particular circumstances, and as such do not lend themselves to quantification according to some universal standard.[4] Whether an act is a duty 'sans phrase', 'actual' or 'proper' duty depends on *all* the morally significant kinds it is an instance of. In contrast, a 'prima facie duty' or 'conditional duty' refers to

[3]The controversy regarding this news item can be followed at
`http://www.nature.com/news/specials/mutantflu/index.html`.

[4]For details about Ross' intuitionist views on morality and his distinction between prima facie duties and actual duties, we refer to [153].

> the characteristic ... which an act has, in virtue of being of a certain kind (e.g. the keeping of a promise), of being an act which would be a duty proper if it were not at the same time of another kind which is morally significant. [153, p. 19]

The distinction between prima facie and actual duties was taken over by Hare, albeit that the latter applied it in a different sense, and used it to argue against Ross' intuitionism. Hare discerns an intuitive and a critical level of (moral) thinking. Prima facie principles are relatively simple principles used at the intuitive level. At the critical level, we select among prima facie principles and resolve conflicts between them [83].

Notwithstanding their diametrically opposed meta-ethical positions (Ross the intuitionist vs. Hare the utilitarian), Ross and Hare agree that, ultimately, all moral conflicts are resolvable. Their defense of this claim is similar to the extent in which they argue that, *prima facie*, conflicts arise between moral principles. After critical investigation, however, the conflicts disappear.

We pick up the discussion on the ultimate existence of moral dilemmas below in Section 1.5. In what follows, we use the terms 'prima facie' and 'all-things-considered' in a sense very similar to that of both Ross and Hare, except that we do not tie the concepts to any meta-ethical views. We take the distinction to apply to each two-level view on morality according to which (a) at the first 'prima facie' level, conflicts may arise between duties, and (b) in case no conflict arises, a prima facie obligation becomes actual.

Philosophers that deny the existence of moral dilemmas (such as Ross and Hare) use the distinction between prima facie and all-things-considered obligations to argue that, in case a conflict arises between two prima facie obligations, we can always, at the deeper, all-things-considered level, find certain features which distinguish both obligations, and make a choice in favor of one of the alternatives. We return to this claim in Section 1.5.

1.4.2 Overriding and overridden norms

If a moral philosopher claims that, all-things-considered, we can always make a choice in favor of one of two conflicting prima facie obligations, she claims that, ultimately, one of the two obligations will always *override* the other. Consider, for instance, the following variant of Plato's classic case (Republic 331c) of a person who ought to return a borrowed weapon (because he promised to do so), and who ought not to return it (because the lender has become insane).

Example 8. A friend leaves you with his gun saying he will be back for it in the evening, and you promise to return it when he calls. He arrives in a distraught condition, demands his gun, and announces he is going to shoot his wife because she has been unfaithful. You ought to return the gun, since you promised to do so – a case of obligation. And yet you ought not to do so, since to do so would be to be indirectly responsible for a murder [109, p. 148].

In this example, your obligation not to return the gun ultimately overrides your obligation to return it. Although *prima facie* both obligations are in conflict, everyone likely agrees that, all-things-considered, one obligation outweighs the other [70, 154].

The idea of norms overriding one another is broader than suggested by the distinction between prima facie and all-things-considered obligations. We could for instance use it to model a more fine-grained hierarchical normative structure, and introduce a partial or total order on various degrees of obligation. Conflicts between norms of different degrees can then be resolved in favor of the norms that are higher up in the hierarchy and that override the ones lower in the hierarchy. This approach is of great practical interest, but falls outside the scope of this thesis. For some examples of it, see e.g. [81, 182].

Another way of implementing the idea that one norm overrides another is to give precedence to norms that, given some normative context, are more *specific* than others. As an example, Horty cites the etiquette norms “Don't eat with your fingers” and “If your are served asparagus, eat it with your fingers”. When eating asparagus, the latter (more specific) norm overrides the former [91]. Logicians typically use *conditional* operators for modeling cases of specificity. This approach will not be pursued here, although we briefly return to it in Section 2.4.1. For some treatments of specificity cases in the literature that make use of conditional operators, see e.g. [42, 171].

1.4.3 Normative standards

In Examples 1 and 2 it seems that the normative conflict in question arises due to opposing directives originating from different normative standards. In Example 1, Agamemnon ought, *as a father*, not to sacrifice his daughter. *As a commander* however, he ought to further the expedition. Similarly, Yilmaz ought, *as a religious devotee*, not to drink alcohol. *As a citizen of his country* however, he is permitted to drink alcohol. Thus, one might object, Agamemnon and Yilmaz face no normative conflict *simpliciter*. Instead, the apparent conflict they face is relative to different normative codes.

In Section 5.2 we will present a logic capable of distinguishing between different normative standards or codes in view of which norms arise.

1.4.4 Further distinctions

Apart from their relative strength and the normative standards in view of which they arise, there are various other features on the basis of which we can distinguish between two or more norms. In legal practice, for instance, there are a number of ‘meta-norms’ for conflict-resolution according to which later laws may override earlier ones (lex posterior derogat priori), laws promulgated by higher or more competent authorities may override laws promulgated by lower authorities (lex superior derogat inferiori), and particular laws may override more general laws (lex specialis derogat generali) [3, 124].

Some authors also make the distinction between intrapersonal and interpersonal normative conflicts. *Intrapersonal conflicts* are conflicts in which the conflicting norms concern one and the same agent (or group of agents). *Interpersonal conflicts* are conflicts in which the conflicting norms hold for different (groups of) agents. Both Marcus and Sinnott-Armstrong refer to the following example as an interpersonal dilemma [121, 164].

Example 9. In Sophocles' *Antigone*, Creon declares the burial of Antigone's brother Polyneices illegal on the not unreasonable grounds that he was a traitor to the city and that his burial would mock the loyalists who defended the city, thereby causing civil disorder. At the same time, there is reason for Creon to respect the religious and familial obligation of Antigone to bury her brother [72, p. 4].

In the example, Creon's obligation to keep his word and preserve the peace conflicts with Antigone's obligation to bury her brother. In Chapter 7 we present a multi-agent logic that respects the distinction between inter- and intrapersonal normative conflicts.

There are various other properties of norms that we can use for distinguishing between them. For instance, next to the concrete agents for which a norm holds, we could also specify the *interest group* in view of which it holds [107]. Moreover, we could discern norms with different *probabilities* [40] or *degrees of utility* [93], etc.

1.5 Normative conflicts and moral dilemmas

As mentioned in Section 1.4.1, the distinction between prima facie and all-things-considered obligations was introduced by Ross in order to argue against the existence of moral dilemmas. A *moral dilemma* is any situation in which, at the same time,

(a) there is a moral requirement for an agent to adopt each of two alternatives,

(b) neither moral requirement is overridden in any morally relevant way,

(c) the agent cannot adopt both alternatives together, and

(d) the agent can adopt each alternative separately.

The characterization in terms of (a)-(d) is taken from [164]. Some authors add to it that the impossibility to adopt both alternatives together must be *circumstancial*: that the conflict arises *contingently* or that it is not *logically* impossible to realize both alternatives together [37, 203]. This additional demand is important for the discussion in Section 5.1, but we can ignore it for now. The characterization of moral dilemmas can be extended straightforwardly to situations in which more than two alternative moral requirements are in conflict (moral trilemmas, quadrilemmas, etc.).

Using the terminology from Section 1.4.1, opponents of the existence of moral dilemmas argue that in case a conflict arises between two (or more) prima facie obligations, we can always, at the all-things-considered level, find certain features which distinguish the obligations from each other, so that we can make a choice in favor of one of the alternatives. Using the terminology of Section 1.4.2, they argue that when faced with conflicting obligations, one of these ultimately overrides the other(s).

At least in theory, we can always construct counterexamples to the arguments of philosophers that deny the existence of moral dilemmas. Consider Example 3 above. In the novel, Sophie has to choose between the lives of her older and younger kid. The younger child is more dependent and thus less likely to survive in the children's barracks. This might constitute a (morally relevant?) reason for Sophie to let her oldest child live. However, the example can easily be modified so that there is no relevant difference between both alternatives. For the sake of argument, take Sophie's children to be identical twins. This assumption removes all morally relevant differences between Sophie's alternatives.

Cases like the modified Sophie case, in which no morally relevant differences can be found between two incompatible requirements, are sometimes called *symmetrical conflicts* [140, 164]. Note that in these cases none of the qualifications discussed in Section 1.4 is in force. These are normative conflicts of the same preference, arising from one and the same authority in view of one and the same normative standard, that hold in view of one and the same interest group in the same circumstances. In the words of Ruth Barcan Marcus:

> There is always the analogue of Buridan's ass. Under the single principle of promise keeping, I might make two promises in all good faith and reason that they will not conflict, but then they do, as a result of circumstances that were unpredictable and beyond my control. All other considerations may balance out. The lives of identical twins are in jeopardy, and, through force of circumstances, I am in a position to save only one. Make the situation as symmetrical as you please. A single-principled framework is not necessarily unlike the code with qualifications or priority rule, in that it would appear that, however strong our wills and complete our knowledge, we might be faced with a moral choice in which there are no moral grounds for favoring doing x over y [121, p. 125].

It seems, then, that despite the efforts of philosophers like Ross and Hare, moral dilemmas exist after all. If one moreover accepts that non-overridden moral requirements constitute all-things-considered obligations, then conflicting all-things-considered obligations exist as well. In any case, the possibility of symmetrical conflicts shows that not all conflicts between obligations are ultimately resolvable. This suffices for our discussion on moral dilemmas. Next, we will extend the discussion to irresolvable normative conflicts in general. For a good collection of texts on the topic of moral dilemmas, see [72].

A final remark is in order here. Even if one does not agree with the arguments presented above, the answer to the question concerning the existence of moral dilemmas bears no weight on the question concerning the existence of normative conflicts in general. Under our characterization of normative conflicts, even prima facie conflicting obligations constitute a normative conflict. Thus, whatever one thinks about moral dilemmas, the existence of normative conflicts in general is not at stake.

1.6 Why devise logics for tolerating normative conflicts?

We are concerned with the treatment of irresolvable normative conflicts in deontic logic. Here, we briefly motivate this aim and further delineate the scope of our investigation.

1.6.1 Irresolvable conflicts

By now, it is clear that not all normative conflicts are resolvable. First, normative conflicts can be irresolvable due to their symmetry. At least in theory, it is possible to construct a conflict between two (or more) normative requirements in such a way that none of the alternatives are normatively distinguishable from each other. We already encountered such a 'symmetric' example in a moral context in Section 1.5.

Second, normative conflicts can be irresolvable in practice due to the incomparability of the conflicting alternatives. Even if normatively relevant features *are* present by means of which we can at least try to investigate which of the alternatives outweigh the other(s), it is not always possible to do so. When, for instance, different normative standards are at work, it is not always possible to weigh each conflicting alternative against the other(s). Moreover, our normative theories may not be as well-developed as we need them to be in order to resolve complex situations of conflict. In international law for instance, "the avenues of norm conflict resolution are [...] at best rudimentary. It therefore knows conflicts that are both unavoidable and irresolvable" [136, p. 470].

In complex real-life settings it is not always clear how to proceed when a conflict arises. This is illustrated by the fact that different courts or governments can have diametrically opposed views on the same matter (e.g. the SWIFT case from Example 5). Institutions sometimes even change their minds when it comes to making a decision involving complex normative conflicts. Consider, for instance, the situation sketched in Example 6 in which the NSABB revised its stance on the publication of a controversial scientific result.

It seems, then, that if we want to investigate the logical relations between norms, we have to take into account the reality of irresolvable normative conflicts. This is the main task to be set out in the remainder of this thesis.

1.6.2 Modeling irresolvable conflicts

Resolving normative conflicts is not always possible. A fortiori, it is not possible to devise a logic that will – given some normative and factual information – always provide us with a consistent and uniquely action-guiding set of normative directives. All we can do is

(a) to devise logics that do not trivialize normative conflicts, and

(b) to devise logics that resolve some, but not all normative conflicts.

For target (b), we need to provide our obligations, prohibitions and permissions with some kind of weight and triggering condition such that, when faced with a conflict, we can give priority to norms that are more important or more specific given the context. This aim lies outside the scope of this thesis.

For target (a), we need to devise logics that, given their intended context of application, are *conflict-tolerant*: systems that consistently allow for the presence of irresolvable normative conflicts. This is exactly what we will do in the remainder of this thesis. As will become clear, the mere technical incorporation of normative conflicts in deontic logic poses some problems that are interesting and difficult in their own right. We say that a logic *tolerates* or *accommodates* normative conflicts in case such conflicts do not give rise to triviality or explosion in the logic. For now, this rather vague characterization suffices. In chapters to come, we further refine what it means for a logic to be conflict-tolerant.

Before we turn to deontic logic, we must still answer an important question. Why, if not for resolving them, need we at all develop logics for 'merely' tolerating normative conflicts?

Let us make a short detour concerning the general usefulness of logical methods. Sven Ove Hansson mentions the following advantages of formalization in philosophy:[5]

> When we formalize an informal discourse, we have to make up our minds on issues that are otherwise often neglected, such as the choice of basic concepts, the interdefinability of these concepts, and what principles of inference apply to them. Formalization also stimulates us to provide a reasonably complete account of the entities that we deal with. In particular, the rigorousness of a formal language makes it meaningful to search for a complete list of valid principles of inference. [86, pp. 99-100]

Hansson divides philosophical and interpretational discussions on formal models into three types:

(Type 1) New aspects on issues already discussed in informal philosophy.

(Type 2) New philosophical issues discovered in the formalism that have philosophical relevance apart from the formal models.

(Type 3) Issues peculiar to the chosen formalism that have no bearing on philosophical issues expressible without the formalism.

Suppose now that we formalize the sentence "A is obligatory" as $\mathsf{O}A$. We can then illustrate how the debate concerning normative conflicts in deontic logic gives rise to issues of all three types.

As an issue of type 1, consider the formalization of conflicting obligations. In [69] Goble distinguishes three such formalizations, each one more comprehensive than its predecessors. First, he considers formulas of the form $\mathsf{O}A \wedge \mathsf{O}\neg A$. Next, he considers a logically inconsistent state of affairs, A and B, both conjuncts of which are obligatory, i.e. $\mathsf{O}A$, $\mathsf{O}B$, yet $\vdash \neg(A \wedge B)$. Third, he considers situations

[5] For a more detailed account of Hansson's views on formalization in philosophy, see [80].

in which $\mathsf{O}A$ and $\mathsf{O}B$, yet $\neg \Diamond (A \wedge B)$, where $\Diamond$ abbreviates some sense of possibility (e.g. physical possibility). Assuming the principle (NM),

$$\vdash \neg \Diamond (A \wedge \neg B) \supset (\mathsf{O}A \supset \mathsf{O}B) \tag{NM}$$

Goble ultimately reduces all conflicting obligations to formulas of the form $\mathsf{O}A \wedge \mathsf{O}\neg A$. Suppose that $\mathsf{O}A, \mathsf{O}B$, and $\neg \Diamond (A \wedge B)$. Then, assuming the validity of all rules and axioms of (the propositional fragment of) Classical Logic (henceforth **CL**), we obtain $\mathsf{O}\neg B$ as well as $\mathsf{O}\neg A$ by the substitution-instances $\neg \Diamond (A \wedge B) \supset (\mathsf{O}A \supset \mathsf{O}\neg B)$ and $\neg \Diamond (B \wedge A) \supset (\mathsf{O}B \supset \mathsf{O}\neg A)$ of (NM).

Is (NM) not a tad too strong? In [129], it was argued that it is. We return to this point in Section 5.1. For now, it suffices to see that Goble's discussion illustrates that for issues of type 1, logic provides the clarification and precision to take philosophical discussions to a higher level. In trying to formalize conflicting obligations, we are faced with new questions concerning the validity of schemas like (NM).

But the use of formal methods in philosophy need not stop here. During the second half of the previous century, formal philosophers used arguments from deontic logic for arguing against the very possibility of conflicting (moral) obligations. Suppose, for instance, that we accept the unconditional validity of the aggregation rule (AND) and the 'Kantian' rule according to which 'ought' implies 'can' (OIC):

$$(\mathsf{O}A \wedge \mathsf{O}B) \supset \mathsf{O}(A \wedge B) \tag{AND}$$
$$\mathsf{O}A \supset \Diamond A \tag{OIC}$$

Suppose further that we accept all inferences of **CL**, and that we know that $\mathsf{O}A$, that $\mathsf{O}B$, and that $\neg \Diamond (A \wedge B)$. Then, using (AND), we obtain $\mathsf{O}(A \wedge B)$ from $\mathsf{O}A$ and $\mathsf{O}B$. However, using the contraposition rule from **CL**, we also obtain $\neg\mathsf{O}(A \wedge B)$ from $\neg \Diamond (A \wedge B)$ using (OIC). Thus, we end up with a contradiction. This shows that, on pain of inconsistency, one cannot accept all of **CL**, (AND), and (OIC) while agreeing that there are conflicting (moral) obligations which can be formalized as we did here.

This problem too will be treated in much more detail later on in this thesis. What matters for now is that we are here faced with a new and philosophically interesting problem – an issue of type 2 in Hansson's taxonomy – that arose through and has its origins in attempts to logically characterize the inferences underlying our everyday normative reasoning.

As a problem of type 3, Hansson cites the fact that in Standard Deontic Logic the formula $\mathsf{O}q$ is derivable from the formulas $\mathsf{O}p$ and $\mathsf{O}\neg p$. This is a logical artefact that, although of technical importance, has little to do with moral philosophy. In conclusion, Hansson states that "formal philosophy can only be successful if we have a strong emphasis on issues of types one and two" [86, p. 101].

Keeping in mind Hansson's desiderata for success, we will in the remainder stress the philosophical significance and interest of the logics presented in this thesis, and hope that the type 1 and type 2 examples above already illustrate to some extent the philosophical use of the task at hand. Some more examples of philosophical (type 1 and type 2) issues raised by the problem of normative conflicts in deontic logic are:

(i) What types of normative conflicts are particularly important under which circumstances? Are there contexts in which certain types of normative conflicts can be ignored?

(ii) To what extent should normative conflicts be isolated in deontic logics? Which rules of inference are applicable to conflicting norms?

(iii) Given the possibility of conflicting norms, which inferences should hold unrestrictedly in a conflict-tolerant deontic logic? Which inferences should be restricted? Which inferences should not be valid under any condition?

Apart from these philosophical issues, there is also an obvious practical interest in the development of logics capable of accommodating irresolvable normative conflicts. The development of systems capable of tolerating conflicting norms is considered an important challenge in the fields of deontic logic [77] and normative multi-agent systems [36], with very concrete applications in artificial intelligence [184, 133, 196, 202] and legal science (witness the fact that entire academic journals are devoted to the treatment of legal conflicts and to applications of AI and logic in law).[6]

1.7 Norms, truth-values and the possibility of deontic logic

Before we turn to deontic logic in all its formal details, we stop for one more moment to consider its very possibility. During the first half of the previous century, moral philosophers of the emotivist and prescriptivist persuasion have argued that there are no logical relations between imperatives,[7] while others argued that there are.[8] In the field of deontic logic, this debate is best known in the form of a puzzle made explicit by the Danish philosopher Jørgensen (the puzzle is also known as *Jørgensen's dilemma*) [99]:

(i) Logical operations only hold for sentences with truth values. Imperatives do not have truth values. Therefore there can be no logic of imperatives.

(ii) It seems evident that inferences can be formulated in which some or all of the premises are imperatives. From the sentences 'Love your neighbor as yourself' and 'Love yourself', the conclusion 'Love your neighbor' seems inescapable.

As formulated by Jørgensen, the puzzle only applies to imperatives, and not to norms in general. But if in (i) and (ii) we replace 'imperatives' by 'ought-statements', 'ought-sentences', or 'normative statements', then the argument still

[6] *The Journal of Conflict Resolution* (ISSN: 0022-0027) is an illustration of the former, *Artificial Intelligence and Law* (ISSN: 0924-8463) an illustration of the latter.

[7] E.g. Ayer in [10, pp. 107-109], Stevenson in [167, pp. 113-114].

[8] E.g. Hare in [82, Ch. 2].

bears some intuitive force. It seems, then, that Jørgensen's puzzle applies not just to imperatives, but to evaluative sentences in general.

Clearly, the thorn in the eye of the deontic logician is (i). If (i) is correct (and if Jørgensen's argument applies to evaluative sentences in general) then the very enterprise of deontic logic seems ill-conceived. So let us take a closer look at (i).

(a) Logical operations hold only for sentences with truth values.

(b) Imperatives do not have truth values.

If (a) and (b) are true, then (c) is unavoidable:

(c) There can be no logic of imperatives.

However, even if (c) is correct, we still need

(d) Jørgensen's argument applies to evaluative sentences in general.

in order to attain the conclusion

(e) There can be no logic of norms.

Thus, if (e) is to be avoided, then we must deny one of (a), (b), and (d). Let us examine each of these premises in turn, starting with (a). Supposing that logical deduction should not just be pursued as a purely formal game, but that instead logical operations are to some extent meaningful and subject to interpretation, we require some 'hereditary property' to be conferred upon the conclusion whenever this property is possessed by the premises. Usually, it is the truth value 'true' which constitutes this hereditary property [200]. This is why (a) seems intuitive.

We could, however, escape the conclusion (e) if we replace 'truth' with a different hereditary property in the specific case of norms. We could, for instance, replace it with the concept of validity (relative to some normative system) [200]. Or we could say that an imperative sentence is 'binding' if there is a reason for the agent to perform the required action [195].

Alternatively, one can take Jørgensen's puzzle seriously without acknowledging (e) by constructing a non-truth-functional semantics for norms. This is the main motivation behind e.g. the input/output logics devised by Makinson and van der Torre [77, 113, 115].

The price to pay for this solution is that we seem to 'broaden the conception of logic'. But this need not be too problematic. In the words of von Wright:

> Deontic logic gets part of its philosophic significance from the fact that norms and valuations, though removed from the realm of truth, yet are subject to logical law. This shows that logic, so to speak, has a wider reach than truth. [191, Introduction]

Instead of (a), we could also reject (b), the claim that imperatives do not have truth values. For instance, Kalinowksi argued that people normally treat moral and legal norms as true or false, and that such norms can be part of logical inference [100]. Hansen, however, claims that such considerations confuse the notion of truth with that of a legal or moral norm's validity, i.e. the 'external' recognition of a norm as valid in a certain society [76, pp. 5-6].

Another way out of the acceptance of (e) is to deny (d). This road was taken up by Hage, who argues that imperatives are not closely connected to norms, and that the relevance of Jørgensens puzzle for legal theory and deontic logic is very limited [74, Ch. 6].

As a final option (a last refuge for the deontic logician, if you wish), we mention that some philosophers have accepted (e), but not

(f) There can be no deontic logic.

One can acknowledge (e) while denying (f) by distinguishing between norms and norm-propositions (cfr. Section 1.2.2), and by arguing that deontic logic is the logic of norm-propositions. As such, (e) is acknowledged and Jørgensen's puzzle is taken seriously, while (f) is denied and deontic logic is saved once again.[9]

Throughout this dissertation, we take norms (including imperatives) to stand in relationships that parallel the logical relationships between propositions (Jørgensen too must acknowledge at least this, since it is implied by the second horn of his 'dilemma'). For instance, there exist negated orders, conjunctive commands, conditional requests, etc. Moreover, inferential-like relations hold amongst imperatives as well as between imperatives and (factual) statements.

Given these parallels, we will in the remainder use the same words ("logical","valid","invalid","inference", etc.) for studying normative inference as we do for studying the relations of implication between (non-normative) propositions. For 'purist' philosophers who wish to introduce a new set of terms for talking about normative inference, Castañeda provides the following suggestion:

> we may naturally use the old terms, which the purist philosopher applies to propositions, prefixed by the morpheme '*sh*–'. Thus, we would speak of imperative *sh*–reasonings, which divide into those which are *sh*–valid and those which are *sh*–invalid, the latter being those in which the *sh*–premises *sh*–imply the *sh*–conclusions, and so on. [44, p. 101]

Our approach is consistent with suggestions made for rejecting (a) or (b). Readers that acknowledge both (a) and (b) can still read on under the assumption that the logics defined in this dissertation are logics of norm-propositions instead of logics of norms. The latter claim holds especially for the systems defined in Section 6.2, which are presented explicitly as logics of norm-propositions.

Clearly, Jørgensens puzzle need not constitute the end of (semantical approaches to) deontic logic. What the puzzle illustrates instead is that we should perhaps not take norms to be 'true' in the same sense in which we take factual or analytical ('tautological') statements to be 'true'.

[9] A diverging position is taken up by Alchourrón and Bulygin, who take seriously the distinction between norms and norm-proposition while at the same time denying (e). Instead Alchourrón and Bulygin proceed to construct a logic of norm-propositions 'on top' of the logic of norms [1, 2, 4]. We discuss the approach of Alchourrón and Bulygin in Section 6.2.7.1.

1.8 Preliminaries

Let $\mathcal{W}^a$ be a denumerable stock $p, q, r, \ldots$ of atomic propositions. The set of literals is defined by $\mathcal{W}^l = \{A, \neg A \mid A \in \mathcal{W}^a\}$. The set $\mathcal{W}$ of well-formed formulas (wffs) of the propositional fragment of **CL** is defined recursively as follows:

$$\mathcal{W} := \qquad \mathcal{W}^a \mid \neg\langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \vee \langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \wedge \langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \supset \langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \equiv \langle\mathcal{W}\rangle \mid \bot$$

We use the notational convention that $\mathcal{W}^{\mathbf{L}}$ abbreviates the set of **L**-wffs. Where $\Gamma \subseteq \mathcal{W}^{\mathbf{L}}$ and $A \in \mathcal{W}^{\mathbf{L}}$, we write $\Gamma \vdash_{\mathbf{L}} A$ to denote that A is **L**-derivable from Γ, and $\vdash_{\mathbf{L}} A$ to denote that A is **L**-derivable from the empty premise set. M is an **L**-model of Γ iff $M \Vdash A$ for all $A \in \Gamma$. $\vDash_{\mathbf{L}} A$ iff all **L**-models verify A, and $\Gamma \vDash_{\mathbf{L}} A$ iff all **L**-models of Γ verify A.

Where unambiguous, we will sometimes use $\vdash$ and $\vDash$ without subscripts.

Chapter 2

Normative conflicts and Standard Deontic Logic

> Sometimes it seems as though the standard is only a target for deontic logicians to snipe at
>
> James Forrester [57, p. 1]

✎ I am indebted to Joke Meheus and Christian Straßer for valuable comments on this chapter.

In this chapter, we define and discuss the system best known as *Standard Deontic Logic* (**SDL**). After some preliminaries on deontic operators and on the syntax of **SDL** (Section 2.1), we provide an axiomatic and semantic characterization of **SDL** in Section 2.2.

In Section 2.3, we briefly discuss the formal treatment of normative conflicts by **SDL** as a foretaste of what is to come in the next chapter. We conclude with a section on some of the well-known problems and puzzles relating to **SDL** (Section 2.4). Readers familiar with these issues can safely skip Section 2.4. We conclude this chapter with a preliminary assessment of **SDL** in Section 2.5.

2.1 Preliminaries

2.1.1 Deontic operators

We denote obligations by means of the operator O, permissions by means of the operator P, and prohibitions by means of the operator F. A sentence like "It is obligatory that the street is clean" can be formalized as $\mathsf{O}C$, where C abbreviates "the street is clean". Similarly, $\mathsf{P}C$ denotes "It is permitted that the street is clean", and $\mathsf{F}C$ denotes "It is forbidden that the street is clean".

Many deontic logicians standardly assume that the following equivalences hold between obligations, permissions, and prohibitions:

$$\mathsf{P}A \equiv \neg\mathsf{O}\neg A \tag{2.1}$$

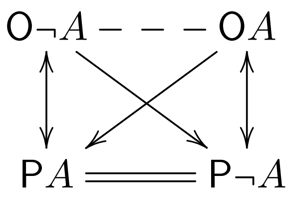


Figure 2.1: The deontic square of opposition.

$$\mathsf{P}A \equiv \neg \mathsf{F}A \tag{2.2}$$
$$\mathsf{F}A \equiv \neg \mathsf{P}A \tag{2.3}$$
$$\mathsf{F}A \equiv \mathsf{O}\neg A \tag{2.4}$$
$$\mathsf{O}A \equiv \neg \mathsf{P}\neg A \tag{2.5}$$
$$\mathsf{O}A \equiv \mathsf{F}\neg A \tag{2.6}$$

Assuming that the equivalences (2.1)-(2.6) hold, the relations between obligations and permissions are depicted graphically in the *deontic square of opposition* (Figure 2.1). In this figure, one-directional arrows represent implications, and two-directional arrows represent contradictories; the nodes connected by a dotted line are contraries, and those connected by a double line are subcontraries.[1]

As we shall see later on, not all of (2.1)-(2.6) are uncontested. The relation of the concept of permission to the concepts of obligation and prohibition is especially problematic. Equivalences (2.4) and (2.6) however are – to the best of our knowledge – uncontested. For these reasons, we shall in the remainder skip the F-operator and assume that it can be defined in terms of the O-operator in the following way: $\mathsf{F}A =_{\mathrm{df}} \mathsf{O}\neg A$. The P-operator remains an essential part of our main language schemas.

2.1.2 Language

The set $\mathcal{W}^{\mathsf{O}}$ of wffs of the fragment of **SDL** without iterated modalities is defined as:

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

We also define the sets $\mathcal{W}^{\mathsf{O}'}$ and $\mathcal{W}^{\mathsf{O}\setminus\mathsf{P}}$. The set $\mathcal{W}^{\mathsf{O}'}$ of wffs of full **SDL** (with iterated modalities) is defined as:

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

The set $\mathcal{W}^{\mathsf{O}\setminus\mathsf{P}}$ of **SDL**-formulas without a primitive P-operator is defined as:

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

[1] Two formulas are *contraries* if they cannot both be true; they are *subcontraries* if they cannot both be false.

In the remainder, we shall work mostly with the set $\mathcal{W}^{\mathsf{O}}$. We come back to the use of iterated deontic modalities in Section 2.4.5. For future reference, we also define the language $\mathcal{W}^{\mathsf{O}}_{\Box}$, obtained by adding to $\mathcal{W}^{\mathsf{O}}$ the alethic modality $\Box$:

$$\mathcal{W}^{\mathsf{O}}_{\Box} := \quad \mathcal{W}^{\mathsf{O}} \mid \Box\langle\mathcal{W}\rangle \mid \neg\langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle\vee\langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle\wedge\langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \supset \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \equiv \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle$$

2.2 SDL

As mentioned at the beginning of the previous chapter, von Wright's 1951 paper *Deontic Logic* [190] has stimulated a lot of later work on deontic logic. With one small modification (the addition of the axiom $\mathsf{O}(A \vee \neg A)$), the 'minimal' system of deontic logic proposed in this paper is now known as *Standard Deontic Logic* or **SDL**. **SDL** has an extremely elegant Kripke-style possible worlds semantics.[2] In this section, we present an axiomatic and semantic characterization of **SDL**, mention some of its meta-theoretical properties, and discuss some closely related systems of deontic logic that are sometimes called 'standard' as well.

2.2.1 Axiomatization

SDL is defined for the set of wffs $\mathcal{W}^{\mathsf{O}}$ by adding to **CL** the axiom schemas (K), (P) and (D), and the rule (NEC):

$$\mathsf{O}(A \supset B) \supset (\mathsf{O}A \supset \mathsf{O}B) \tag{K}$$

$$\mathsf{P}A \equiv \neg\mathsf{O}\neg A \tag{P}$$

$$\mathsf{O}A \supset \mathsf{P}A \tag{D}$$

$$\text{If } \vdash A \text{ then } \vdash \mathsf{O}A \tag{NEC}$$

Adding (K), (P), and (NEC) to **CL** gives us the basic normal modal logic **K** for the language schema $\mathcal{W}^{\mathsf{O}}$. **SDL** extends **K** by (D). For that reason, it is sometimes called **KD** or simply **D**.

(P) is often replaced by the definition $\mathsf{P}A =_{\mathrm{df}} \neg\mathsf{O}\neg A$. In this case the P-operator need not be a primitive symbol in the language. Since later on in this dissertation we will question some instances of (P) in specific deontic contexts, we stick to the axiomatic characterization of the P-operator here.

The following axiom schemas and rules are derivable in **SDL**, and are stated here for future reference:

$$(A \wedge \neg A) \supset B \tag{ECQ}$$

$$(\mathsf{O}A \wedge \mathsf{O}B) \supset \mathsf{O}(A \wedge B) \tag{AND}$$

$$\mathsf{O}A \supset \neg\mathsf{O}\neg A \tag{CONS}$$

$$\text{If } \vdash A \supset B \text{ then } \vdash \mathsf{O}A \supset \mathsf{O}B \tag{RM}$$

$$\text{If } \vdash A \equiv B \text{ then } \vdash \mathsf{O}A \equiv \mathsf{O}B \tag{RE}$$

Fact 1. (ECQ), (AND), (CONS), (RM), and (RE) are **SDL**-valid.

[2] In fact, developments in the newly established field of deontic logic played an important role in the invention of what we now call possible worlds semantics. See [205] for more details.

The proof of Fact 1 is safely left to the reader. Moreover, we leave it to the reader to check that **SDL** validates all logical relationships between the deontic operators that are displayed in the deontic square of opposition (Figure 2.1).

2.2.2 Semantics

We define **SDL** semantically by means of a Kripke-style possible worlds semantics with an actual or designated world. An **SDL**-model is a quadruple $\langle W, w_0, R, v\rangle$, where W is a set of worlds[3], $w_0 \in W$ is the actual world, R is a serial accessibility relation[4] on W and $v : \mathcal{W}^a \times W \to \{0,1\}$ is an assignment function. The valuation $v_M : \mathcal{W}^{\mathsf{O}} \times W \to \{0,1\}$, associated with the model M, is defined by:

(Ca) where $A \in \mathcal{W}^a, v_M(A,w) = 1$ iff $v(A,w) = 1$
(C¬) $v_M(\neg A, w) = 1$ iff $v_M(A,w) = 0$
(C∨) $v_M(A \vee B, w) = 1$ iff $v_M(A,w) = 1$ or $v_M(B,w) = 1$
(C∧) $v_M(A \wedge B, w) = 1$ iff $v_M(A,w) = v_M(B,w) = 1$
(C⊃) $v_M(A \supset B, w) = 1$ iff $v_M(A,w) = 0$ or $v_M(B,w) = 1$
(C≡) $v_M(A \equiv B, w) = 1$ iff $v_M(A,w) = v_M(B,w)$
(CO) $v_M(\mathsf{O}A, w) = 1$ iff $v_M(A,w') = 1$ for every w' such that Rww'
(CP) $v_M(\mathsf{P}A, w) = 1$ iff $v_M(A,w') = 1$ for some w' such that Rww'

An **SDL**-model $M = \langle W, w_0, R, v\rangle$ verifies A, $M \Vdash A$, iff $v_M(A, w_0) = 1$.

2.2.3 Meta-theory

Theorem 1. **SDL** *is reflexive, transitive and monotonic.*[5]

Theorem 2. **SDL** *is compact (if* $\Gamma \vdash_{\mathbf{SDL}} A$ *then* $\Gamma' \vdash_{\mathbf{SDL}} A$ *for some finite* $\Gamma' \subseteq \Gamma$*).*

Theorem 3. *If* $\Gamma \vdash_{\mathbf{SDL}} B$ *and* $A \in \Gamma$*, then* $\Gamma - \{A\} \vdash_{\mathbf{SDL}} A \supset B$ *(Generalized Deduction Theorem for* **SDL***).*

Theorem 4. *If* $\Gamma \vdash_{\mathbf{SDL}} A$*, then* $\Gamma \vDash_{\mathbf{SDL}} A$*. (Soundness of* **SDL***)*

Theorem 5. *If* $\Gamma \vDash_{\mathbf{SDL}} A$*, then* $\Gamma \vdash_{\mathbf{SDL}} A$*. (Strong Completeness of* **SDL***)*

Proofs for Theorems 1-5 can be found in any good introductory textbook on modal logic, e.g. [97].

---

[3]If one feels that the notion of a 'world' has too strong a metaphysical connotation, a more neutral word may be used to denote the elements of W, e.g. *points*.

[4]R is *serial* iff, for every $w \in W$ there is a $w' \in W$ such that Rww'.

[5]Where $\mathrm{Cn}_{\mathbf{L}}(\Gamma)$ denotes the consequence set of some premise set Γ for the logic **L**, a logic **L** is reflexive iff, for all premise sets Γ, $\Gamma \subseteq \mathrm{Cn}_{\mathbf{L}}(\Gamma)$; it is transitive iff, for all sets of wffs Γ and Γ', if $\Gamma' \subseteq \mathrm{Cn}_{\mathbf{L}}(\Gamma)$ then $\mathrm{Cn}_{\mathbf{L}}(\Gamma \cup \Gamma') \subseteq \mathrm{Cn}_{\mathbf{L}}(\Gamma)$; and it is monotonic iff, for all sets of wffs Γ and Γ', $\mathrm{Cn}_{\mathbf{L}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{L}}(\Gamma \cup \Gamma')$.

2.2.4 More 'standard' deontic logics

In [8], Åqvist widens the conception of what it is to be a 'standard' deontic logic. Roughly, he defines a *normal propositional monadic von Wright-type deontic logic* to be a logic that contains **CL**, (K), (P), and (NEC). A *strongly* normal propositional monadic von Wright-type deontic logic moreover contains (D). Åqvist's umbrella conception of 'standard' or 'normal' deontic logics (we use the terms interchangeably here) also includes strengthenings with iterated modalities that make use of axiom schemas like:

$$\mathsf{O}A \supset \mathsf{O}\mathsf{O}A \tag{2.7}$$

$$\mathsf{P}\mathsf{O}A \supset \mathsf{O}A \tag{2.8}$$

$$\mathsf{O}(\mathsf{O}A \supset A) \tag{2.9}$$

Furthermore, it allows him to regard conditional extensions of these systems as 'standard'. Thus, for Åqvist, **SDL** as defined here is just one of many 'standard' deontic logics.

Another interesting development worth mentioning in this context is Anderson's reduction of deontic logic to modal logic with only alethic modalities [5, 6]. This reduction is realized by adding to the basic normal modal logic **K** a constant proposition V representing a *violation* (penalty, sanction) relative to a normative system under investigation. Where $\Box$ is the necessity operator of **K**, the F-operator is then defined as $\mathsf{F}A =_{\mathrm{df}} \Box(A \supset \mathsf{V})$, i.e. A is forbidden if it entails a violation. The other operators are defined in terms of F: $\mathsf{O}A =_{\mathrm{df}} \mathsf{F}\neg A$, and $\mathsf{P}A =_{\mathrm{df}} \neg\mathsf{F}A$.

Around the same time and independently from Anderson, Kanger [102] proposed a roughly equivalent reduction by making use of a constant proposition Q for abbreviating that all normative demands are met (see [126] for a comparison with Anderson's reduction). The Anderson-Kangerian reduction of deontic to alethic modal logic is an extension of **SDL**. More precisely, **SDL** is the deontic fragment of the Anderson-Kanger systems (see [8, Section 14] for a rigorous stipulation of the deontic fragment of these systems, and for the proof that this fragment is equivalent to **SDL**).

We sympathize with Åqvist's, Anderson's and Kanger's characterization of 'standard' deontic logics and recognize their historic importance. Yet although many of the systems defined by these authors are very similar to **SDL**, and although many of the claims we make about **SDL** also hold for 'normal' deontic logics in these families, we will in the remainder keep on referring to the system **SDL** as defined in Sections 2.2.1 and 2.2.2 when we talk about Standard Deontic Logic. We briefly return to the Anderson-Kangerian reduction in Section 2.4.3.

2.3 SDL and normative conflicts

2.3.1 Formalizing normative conflicts

Many inescapable normative conflicts fit the general logical form $\mathsf{O}A \wedge \mathsf{O}\neg A$. Take, for instance, Example 1 from the previous chapter. Where d abbreviates

"Agamemnon sacrifices his daughter", it seems that, as a commander, Agamemnon faces the obligation $\mathsf{O}d$. As a father, however, Agamemnon ought not sacrifice his daughter, $\mathsf{O}\neg d$. Similarly, in Example 5, SWIFT ought not to pass the information ($\mathsf{O}\neg i$) according to Belgian data protection regulations, whereas SWIFT ought to pass the information ($\mathsf{O}i$) according to the U.S. Treasury. In general, we call conflicts of the type $\mathsf{O}A \wedge \mathsf{O}\neg A$ OO-*conflicts*.

Escapable conflicts do not fit the logical form $\mathsf{O}A \wedge \mathsf{O}\neg A$. In Example 2, Yilmaz ought not drink any alcohol ($\mathsf{O}\neg d$) according to his religious beliefs, whereas he is permitted to do so ($\mathsf{P}d$) according to the laws of his country. Thus, escapable normative conflicts generally appear to be of the form $\mathsf{O}A \wedge \mathsf{P}\neg A$ (or $\mathsf{O}\neg A \wedge \mathsf{P}A$). In general, we call conflicts of the type $\mathsf{O}A \wedge \mathsf{P}\neg A$ (or $\mathsf{O}\neg A \wedge \mathsf{P}A$) OP-*conflicts*.

In more complex situations, more than one proposition may be involved in a normative conflict. Consider the following example.

Example 10. Alice is throwing a party for her birthday. Since Bob and Charles are good friends of Alice, she ought to invite Bob ($\mathsf{O}b$) and to invite Charles ($\mathsf{O}c$) to her party. However, when Bob and Charles get together, they usually get drunk, and chances are that they will annoy the other guests. Hence Alice ought not invite both Bob and Charles to her party ($\mathsf{O}\neg(b \wedge c)$).

In view of (RE), $\mathsf{O}\neg(b \wedge c)$ is equivalent to $\mathsf{O}(\neg b \vee \neg c)$. The latter formula generates a conflict together with the obligations $\mathsf{O}b$ and $\mathsf{O}c$. Since two propositions (b and c) are involved, we say that we are dealing with a *binary* normative conflict.

Following [63], we say that, where $n > 1$, a conflict of the form $\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n \wedge \mathsf{O}(\neg A_1 \vee \ldots \vee \neg A_n)$ is an n-ary OO-conflict. Similarly, a conflict of the form $\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n \wedge \mathsf{P}(\neg A_1 \vee \ldots \vee \neg A_n)$ is an n-ary OP-conflict.

In Section 3.2.1, we elaborate further on how to formalize normative conflicts in deontic logic, and on the expressive resources needed for doing so. For now, we suppose for the sake of argument that – at least when restricted to the language schema $\mathcal{W}^{\mathsf{O}}$ – many examples of normative conflicts indeed bear the logical form $\mathsf{O}A \wedge \mathsf{O}\neg A$, $\mathsf{O}A \wedge \mathsf{P}\neg A$, or $\mathsf{O}\neg A \wedge \mathsf{P}A$.[6]

2.3.2 Explosion

We can formalize normative conflicts in **SDL** as OO-conflicts or OP-conflicts, but we cannot *consistently* do so. **SDL** trivializes OO- and OP-conflicts:

$$\mathsf{O}A \wedge \mathsf{O}\neg A \vdash_{\mathbf{SDL}} \perp \tag{OO-EX}$$

$$\mathsf{O}A \wedge \mathsf{P}\neg A \vdash_{\mathbf{SDL}} \perp \tag{OP-EX}$$

Suppose that we are facing an OO-conflict $\mathsf{O}A \wedge \mathsf{O}\neg A$. By (D), we can derive $\mathsf{P}A$ from $\mathsf{O}A$. By (P), we can derive $\neg\mathsf{O}\neg A$ from $\mathsf{P}A$. But then the contradiction $\mathsf{O}\neg A \wedge \neg\mathsf{O}\neg A$ is derivable, and, by (ECQ), it follows that $\perp$.

[6] This also holds for binary, ternary, etc. conflicts, since every n-ary OO-conflict $\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n \wedge \mathsf{O}(\neg A_1 \vee \ldots \vee \neg A_n)$ is **SDL**-equivalent to the formula $\mathsf{O}(A_1 \wedge \ldots \wedge A_n) \wedge \mathsf{O}\neg(A_1 \wedge \ldots \wedge A_n)$ in view of (AND) and (RE); similarly, every n-ary OP-conflict $\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n \wedge \mathsf{P}(\neg A_1 \vee \ldots \vee \neg A_n)$ is **SDL**-equivalent to the formula $\mathsf{O}(A_1 \wedge \ldots \wedge A_n) \wedge \mathsf{P}\neg(A_1 \wedge \ldots \wedge A_n)$.

Similarly, when faced with an OP-conflict $\mathsf{O}A \wedge \mathsf{P}\neg A$, we can derive $\neg\mathsf{O}\neg\neg A$ from $\mathsf{P}\neg A$ by (P). By (RE), it follows that $\neg\mathsf{O}A$. Again, we have derived a contradiction, and again $\perp$ follows by (ECQ).

Given the existence of normative conflicts, and assuming that they can be formalized as OO- and OP-conflicts, the explosion principles (OO-EX) and (OP-EX) show that **SDL** is incapable of modeling our actual normative reasoning.

For now, this is all we say about explosion. In the next chapter we will have a lot more to say about this phenomenon. There, we will tackle the remaining questions concerning the formalization of normative conflicts, and take a closer look at some of the inferences that lead to explosion in **SDL**. Moreover, we will define an additional set of more sophisticated explosion principles, all of which also arise in **SDL**.

2.4 More problems

Given the dominant role of **SDL** in this dissertation, it is important that, apart from its incapability to accommodate normative conflicts, we also mention its other problems and 'puzzles'. This section is largely meant to inform the reader about some well-known issues pertaining to **SDL**, and to facilitate the discussion in later chapters.

2.4.1 Chisholm's puzzle and contrary-to-duty obligations

Consider the following sentences, which appear both consistent and independent of one another [46]:

(i) It is obligatory that Jones goes to the assistance of his neighbours

(ii) It is obligatory that if Jones goes to the assistance of his neighbours, then he tells them he is coming

(iii) If Jones does not go to the assistance of his neighbours, then he ought not tell them he is coming

(iv) Jones does not go to the aid of his neighbours

The three most popular formalizations of (i)-(iv) are the following [126, 127]:

	Formalization 1	Formalization 2	Formalization 3
(i)	$\mathsf{O}g$	$\mathsf{O}g$	$\mathsf{O}g$
(ii)	$\mathsf{O}(g \supset t)$	$\mathsf{O}(g \supset t)$	$g \supset \mathsf{O}t$
(iii)	$\neg g \supset \mathsf{O}\neg t$	$\mathsf{O}(\neg g \supset \neg t)$	$\neg g \supset \mathsf{O}\neg t$
(iv)	$\neg g$	$\neg g$	$\neg g$

Formalization 1 is **SDL**-inconsistent. From (i) and (ii) we can derive $\mathsf{O}t$ by (K). From (iii) and (iv) we obtain $\mathsf{O}\neg t$. Together, $\mathsf{O}t$ and $\mathsf{O}\neg t$ imply $\perp$ by (OO-EX).

In formalization 2, the independence of the premises is lost, since (iii) is an **SDL**-consequence of (i). Similarly, in formalization 3, (ii) follows from (iv).

Thus, in none of the formalizations above, the premises are both consistent and independent of each other.[7] This is Chisholm's puzzle.

The sting in Chisholm's puzzle is caused by the *contrary-to-duty* obligation (iii). In order to solve the puzzle, many authors have relied on the stronger resources of dyadic deontic logic. Let $\mathsf{O}(A \mid B)$ express "it is obligatory that B under condition A". Then the following formalizations of (i)-(iv) seem to fare better:

	Formalization 4	Formalization 5
(i)	$\mathsf{O}g$	$\mathsf{O}g$
(ii)	$\mathsf{O}(g \supset t)$	$\mathsf{O}(g \mid t)$
(iii)	$\mathsf{O}(\neg g \mid \neg t)$	$\mathsf{O}(\neg g \mid \neg t)$
(iv)	$\neg g$	$\neg g$

Let $\mathcal{W}^{\mathsf{O}}_{C}$ be obtained by adding the conditional obligation operator to $\mathcal{W}^{\mathsf{O}}$ (and by closing it under the classical connectives), and let $\mathbf{SDL_C}$ be the logic obtained by the grammar $\mathcal{W}^{\mathsf{O}}_{C}$ and by the axioms and rules of **SDL**. Then formalizations 4 and 5 are better than formalizations 1-3 since they satisfy the demands of consistency and independence. However, the problem reappears when we try to *detach* conditional obligations. Consider the following detachment principles commonly occurring in the literature on dyadic deontic logic:

$$(A \wedge \mathsf{O}(A \mid B)) \supset \mathsf{O}B \tag{F-DET}$$

$$(\mathsf{O}A \wedge \mathsf{O}(A \mid B)) \supset \mathsf{O}B \tag{D-DET}$$

The *factual detachment* principle (F-DET) allows to derive an unconditional obligation from a conditional one whenever its condition holds. The *deontic detachment* principle (D-DET) allows to derive an unconditional obligation from a conditional one whenever the condition of the latter obligation is also obligatory.

If (F-DET) and (D-DET) were added to $\mathbf{SDL_C}$, then formalizations 4 and 5 become inconsistent again. In case of formalization 4, we could then derive $\mathsf{O}\neg t$ from (iii) and (iv) by (F-DET), while we could derive $\mathsf{O}t$ from (i) and (ii) by (K). By (OO-EX), it follows that $\perp$. In case of formalization 5, the argument is analogous, except that here we also need (D-DET) in order to derive $\mathsf{O}t$ from (i) and (ii).

The principles (F-DET) and (D-DET) are not easily given up. With Van Eck, we agree that it is hard to "take seriously a conditional obligation if it cannot, by way of detachment, lead to an unconditional obligation" [186, p. 263]. Hence, given the intuitive appeal of (F-DET) and (D-DET), it seems that Chisholm's puzzle reappears in the conditional setting as a dilemma of commitment and detachment. As this is clearly a problem for conditional logics, it need not concern us here. For a more detailed discussion of Chisholm's puzzle, and for some attempts at solving it, see [8, 173].

As a final remark on Chisholm's puzzle, note that, given the present level of analysis, the puzzle loses (part of) its sting in a logic capable of accommodating

[7] One may feel that an important alternative formalization is missing, i.e. $\{\mathsf{O}g, g \supset \mathsf{O}t, \mathsf{O}(\neg g \supset \neg t), \neg g\}$. However, it is readily seen that this formalization combines the disadvantages of formalizations 2 and 3.

OO-conflicts. For then the first formalization of premises (i)-(iv) is no longer inconsistent, and one might argue that Jones is simply facing two conflicting obligations here. On the one hand, he should tell his neighbours he is coming due to his obligation in (i). On the other hand, he should not tell them he is coming due to the fact that he does not go (iv). For a more detailed analysis of the situation at hand, we need additional expressive resources.

2.4.2 Ross' puzzle and free choice permissions

The following inference is **SDL**-valid:

$$\mathsf{O}A \vdash_{\mathbf{SDL}} \mathsf{O}(A \vee B) \tag{2.10}$$

According to (2.10), if a certain state of affairs A is obligatory, then $A \vee B$ too is obligatory. For instance, if I ought to mail a letter, then I also ought to mail the letter or burn it. Alf Ross felt that there is something paradoxical in this inference, since there is a way of fulfilling the second obligation without fulfilling the first one (namely to burn the letter without mailing it) [152]. Therefore Ross and others concluded that (2.10) should not be a theorem of deontic logic.

(2.10) is valid in view of the **CL**-theorem $A \supset (A \vee B)$, (NEC) and (K). To say that $\mathsf{O}A$ implies $\mathsf{O}(A \vee B)$, then, appears no more paradoxical than to say that A implies $A \vee B$. But is this sufficient to dismiss Ross's paradox?

According to some, it is. Castañeda, for instance, argues that Ross's puzzle arises from 'semantical atomism'. Against such atomists, Castañeda argues that no sentence is an island unto itself: "When one infers a conclusion one is considering one member of a related set – and one must remember the premises, or remember that the premises are still valid or true, or whatever property is supposed to be preserved in inference" [43, pp. 64]. According to Castañeda we must simply forget about Ross's puzzle.

Von Wright is less pleased with easy dismissals of Ross's puzzle, and relates it to the problem of *free choice permission*. Ross again has argued that there is a sense of permission (free choice permission) according to which from $\mathsf{P}(A \vee B)$ we are able to derive both $\mathsf{P}A$ and $\mathsf{P}B$. When the waiter at a restaurant tells you that "You may have steak or fish for lunch", then, normally, it follows that you may have steak for lunch. Similarly, it follows that you may have fish for lunch. However, both inferences are **SDL**-invalid:

$$\mathsf{P}(A \vee B) \nvdash_{\mathbf{SDL}} \mathsf{P}A \tag{2.11}$$

$$\mathsf{P}(A \vee B) \nvdash_{\mathbf{SDL}} \mathsf{P}B \tag{2.12}$$

According to von Wright, (2.10) strikes us absurd because "We incline to think of the obligatory as also permitted and we 'naturally' understand disjunctive permissions as free choice permissions. This being so, Ross's Paradox seems to allow the inference from the obligatoriness of a certain action to the permittedness of any other action" [193, p. 22]. In symbols, von Wright's inference is that from $\mathsf{O}A$ to $\mathsf{P}B$. From $\mathsf{O}A$ it follows that $\mathsf{O}(A \vee B)$ by (2.10). By (D), $\mathsf{P}(A \vee B)$. Now if $\mathsf{P}(A \vee B)$ were a free choice permission, it would follow that $\mathsf{P}B$, contrary to inference (2.12).

Von Wright solves the apparent paradox by adopting a pluralist position. He takes Ross's puzzle to arise from a confusion between different concepts of permission, namely the P-operator from **SDL** and a free choice permission operator P' for which (2.11) and (2.12) are valid:

> The moral to be drawn from these considerations is that there are *several concepts* of permission and obligation. The "paradoxes" arise through a confusion on the intuitive level between different concepts. When the concepts are clearly separated there are no "paradoxes". Their separation is achieved through the construction of a variety of deontic logics. [193, p. 33-34]

Føllesdal and Hilpinen disagree. In [52], they argue that (2.10), (2.11), and (2.12) may be explained by reference to general conventions regarding the use of language. One such convention is that it is generally assumed that people make as strong statements as they are in a position to make. Thus, if someone wants another person to mail a letter, it is very awkward for her to say that it is obligatory that the letter be mailed or burned, especially if the latter alternative is forbidden. Similarly, they argue that the logical force of the word "or" in "You may have steak or fish for lunch" is really the same as that of "and". Thus, the sentence should be formalized not as a disjunctive permission, $\mathsf{P}(A \vee B)$, but as a conjunction of two permissions, $\mathsf{P}A \wedge \mathsf{P}B$. For Føllesdal and Hilpinen, "There is no need to invent special notions of permission and obligation on the basis of this accidental interchangeability of the words 'or' and 'and' in ordinary language" [52, p. 23].

It seems, then, that Ross's puzzle can be resolved in various ways. With Åqvist, we agree that:

> Contrary to the view of its originator, the Alf Ross paradox does not seem to be a serious threat to the very possibility of constructing a viable deontic logic. But it usefully directs our attention to the ambiguity of normative phrases in natural language as a possible source of error and confusion – in viable deontic logics we should be able to express, to do justice to, and to pinpoint such ambiguities. For this reason I agree with von Wright in claiming that the puzzle deserves serious consideration. [8, p. 179]

2.4.3 The good Samaritan

As formulated by Arthur N. Prior, the paradox of the good Samaritan is the following:

> [H]elping someone who has been robbed with violence is an act that can only occur if the person has been so robbed ("x helps y who has been robbed" necessarily implies "y has been robbed"); but the robbery (being wrong) necessarily implies the sanction; therefore the succor (since it implies robbery) implies the sanction, too, and is also wrong. [148, p. 144]

Where h abbreviates "x helps y who has been robbed" and r abbreviates "y has been robbed", Prior formalizes the good Samaritan paradox as follows, using Anderson's reduction from Section 2.2.4:

(i) "x helps y who has been robbed" necessarily implies "y has been robbed" ($\Box(h \supset r)$)

(ii) It ought not be the case (is forbidden) that y has been robbed ($\Box(r \supset \mathsf{V})$)

But then, in view of the **K**-theorem

$$\Box(A \supset B) \supset (\Box(B \supset C) \supset \Box(A \supset C)) \tag{2.13}$$

(iii) It ought not be the case (is forbidden) that x helps y ($\Box(h \supset \mathsf{V})$)

Surely, (iii) is a weird conclusion to draw from (i) and (ii).

In the non-Andersonian language of **SDL**, the good Samaritan paradox is usually formalized as follows [7, 42, 126, 139]:

(i′) It is obligatory that x helps y who has been robbed ($\mathsf{O}(h \wedge r)$)

In view of (RM) and the **CL**-theorem $(h \wedge r) \supset r$, (i′) immediately gives us:

(ii′) It is obligatory that y has been robbed ($\mathsf{O}r$)

A third formalization of the good Samaritan paradox is found in e.g. [105]:

(i″) It is obligatory that x helps y who has been robbed ($\mathsf{O}h$)

(ii″) y has been robbed (r)

Since h entails r ($\vdash h \supset r$), we can derive $\mathsf{O}r$ from $\mathsf{O}h$ by (RM).

If the formalizations above are correct, then it is clear that the central principle underlying this 'paradox' is (RM) or – in case of the first formalization – its alethic counterpart (RM□):[8]

$$\text{If } \vdash A \supset B, \text{ then } \vdash \Box A \supset \Box B \tag{RM□}$$

However, there is something fishy about the above formalizations of the good Samaritan puzzle. In each of these formalizations, the paradoxical effect that "it is obligatory that y has been robbed" only occurs if the entire sentence "x helps y who has been robbed" occurs within the scope of a modal operator.

Suppose now that we formalize the situation as follows:

(i‴) If y has been robbed, then it is obligatory that x helps y ($r \supset \mathsf{O}h$)

(ii‴) y has been robbed (r)

[8]Theorem (2.13) arises after applying (RM□) and the alethic counterpart of (K) to the **CL**-theorem $(A \supset B) \supset ((B \supset C) \supset (A \supset C))$.

Then the paradox no longer arises. With Forrester, we agree that it feels more natural to read the clause "who has been robbed" as outside the scope of the O-operator [55].[9]

Are we then freed from all puzzles relating to the **SDL**-derivable rule (RM)? Not yet, since in the very same paper in which Forrester showed that the good Samaritan puzzle loses its bite, he goes on to develop an 'adverbial' variant in which the problems centering (RM) reappear: the gentle murderer puzzle.

2.4.4 The gentle murderer

The gentle murderer puzzle arises from the following premises:

(i) It is obligatory that Smith not murder Jones ($\mathsf{O}\neg m$)

(ii) If Smith murders Jones, it is obligatory that Smith murders Jones gently ($m \supset \mathsf{O}g$)

(iii) Smith murders Jones (m)

Clearly,

(iv) If Smith murders Jones gently, then Smith murders Jones ($\vdash g \supset m$)

From (ii) and (iii), we can derive $\mathsf{O}g$ by modus ponens (MP). From (iv), it follows that $\vdash_{\mathbf{SDL}} \mathsf{O}g \supset \mathsf{O}m$ by (RM). But then, by (MP) again, it follows that it is obligatory that Smith murders Jones ($\mathsf{O}m$). To make things worse, $\mathsf{O}m$ and (i) cause full explosion by (O-EX). This is Forrester's gentle murderer puzzle.

Castañeda famously called this puzzle "the deepest paradox of all" in deontic logic, but he nonetheless proposed a solution for it [45]. Other solutions were presented e.g. by Meyer [131] and by Sinnott-Armstrong [163]. The solutions of Castañeda and Sinnott-Armstrong were questioned by Goble in [64]. Like Forrester's, Goble's own solution to the gentle murderer involves rejecting (RM). We feel, however, that this is too drastic. We come back this point in Section 3.2.2.2.

Despite the trouble it causes for **SDL**, two remarks will suffice to show that the gentle murderer 'paradox' need not pose problems for us here. First, one might argue that Smith is simply facing two conflicting obligations here. On the one hand, Smith ought not to murder Jones, while on the other hand he ought to murder Jones (albeit gently). Thus, Forrester's paradox loses its sting in any logic capable of accommodating conflicting obligations.

Second, if one feels intuitively unsatisfied by the dissolution of the gentle murderer puzzle in conflict-tolerant deontic logics, there is still the possibility of enriching **SDL** with degrees of obligation. In the resulting enrichment, Smith's obligation not to murder should carry more weight than his obligation to murder gently. Moreover, the violation of the stronger obligation not to murder should not free Smith from his weaker obligation to murder gently.[10]

[9] Forrester's argument can be generalized to the epistemic variant of the good Samaritan puzzle formulated by Åqvist in [7]. See [55] for some discussion.

[10] Alternatively, one might argue that a dyadic operator is needed for formalizing Smith's 'conditional' obligation to murder gently, or that different 'kinds' of obligation are at stake here (see [173]).

2.4.5 Iterated deontic modalities

Iterated or *nested* deontic modalities are deontic operators that occur within the scope of another deontic operator. The set of formulas in which such iterations occur consists of those formulas that belong to $\mathcal{W}^{\mathsf{O}'}$, but not to $\mathcal{W}^{\mathsf{O}}$. The problem relating to iterated deontic operators is twofold. First, we need to know how such formulas are to be interpreted. Second, we need to find out if there are any theorems or 'truths' of deontic logic that pertain to the nesting of deontic operators.

To see why the first point is important, it suffices to try and pronounce (let alone interpret) a formula like $\mathsf{OPOFOP}A$. But we need not consider such complex iterations in order to realize that the interpretation of iterated deontic modalities is ambiguous. Consider, for instance, the seemingly sensible statement "Parking on highways ought to be forbidden". Taking h to denote "parking on highways", the statement could be formalized as:

$$\mathsf{OF}h \tag{2.14}$$

If we accept equivalences (2.1)-(2.6) from Section 2.1.1, then by some simple propositional manipulations (2.14) is equivalent to the formula

$$\mathsf{FP}h \tag{2.15}$$

In [120], Marcus argued that although (2.14) and (2.15) are equivalent assuming **CL** and (2.1)-(2.6), their meaning is quite different. Whereas (2.14) describes a desirable state of affairs ("It ought to be that ..."), (2.15) describes the very same state of affairs as if it already obtains ("It is forbidden that ..."). The conclusion seems to be that we cannot trust our pre-formal intuitions regarding iterated deontic modalities.

Marcus' confusion disappears if we only consider statements of the form "it is obligatory that A". Clearly, if a formula $\mathsf{O}A$ is read as "it ought to be that A", then $\mathsf{O}\neg A$ should not be equivalent to $\mathsf{F}A$ if the latter is read as "it is forbidden that A. For "it ought not be that A" is a very different statement from "it is forbidden that A". This is why we mentioned in Section 1.2.1 that we will not read a formula $\mathsf{O}A$ as "it ought to be that A", but rather as "it is obligatory that A".

The problem posed by Marcus was addressed in an agentive setting in [32, 197]. Although interesting because of the formalisms used there for representing agency in deontic logic, we do not believe that agentive modalities are required for addressing Marcus' original problem of parking on highways. We simply need to be aware of the ambiguity between statements of the form "it ought to be that A" and statements of the form "it is obligatory that A".

Depending on our interpretation of nested occurrences of deontic operators, we might wonder if there are any theorems of deontic logic that concern iterated modalities. Proposed candidates for such theorems include the schema (2.9) from Section 2.2.4, according to which "It is obligatory that what is obligatory is the case". Prior, for instance, takes this schema to be intuitively acceptable [149].

Other candidate schemas include (2.7) and (2.8) from Section 2.2.4, as well as for instance the schema

$$\mathsf{OO}A \supset \mathsf{O}A \tag{2.16}$$

The question whether or not to accept iterated modalities and (some) axiom schemas pertaining to them bears influence on the topic of this dissertation. For given a schema like (2.16) a formula $\mathsf{OO}A \wedge \mathsf{O}\neg A$ reduces to the OO-conflict $\mathsf{O}A \wedge \mathsf{O}\neg A$. However, such questions are secondary to the more general question which axioms pertaining to iterated deontic modalities should be valid in which contexts, if we should at all allow for such iterations. As the latter question remains largely unanswered, we will not say much about nested modalities in the remainder. In Chapter 7 however, we come back to nested modalities in the setting of multi-agent logics.

2.4.6 Permission

Remember from Section 2.4.2 that 'permission' is a very ambiguous term in our natural language. With Stenius, we agree that "it is much more difficult to get an intuitive grasp of "permission" than of "obligation", or, above all, "prohibition"" [166, pp. 66-67]. Apart from free choice permissions, we can distinguish at least two other senses of permission in our natural usage of the term.

The P-operator of **SDL** is typically conceived as a *weak* or *negative* permission operator. In this sense of the term, a permission to A merely denotes the absence of an obligation to the contrary. Thus, if P is an operator for weak permission, then $\mathsf{P}A$ is logically equivalent to $\neg\mathsf{O}\neg A$.

As opposed to weak or negative permissions, philosophers also speak of *strong* or *positive* permissions, i.e. permissions that are either explicitly stated as such, or permissions that are derivable from other explicitly stated permissions or obligations. In case of a strong permission, the usual interrelations between obligations and permissions (as displayed in the square of opposition at the beginning of this chapter) break down.

In Section 6.2 we have much more to say on the distinction between weak and strong permission, and define a logic that is capable of formalizing both concepts. Regarding other notions of permission (e.g. the concept of free choice permission from Section 2.4.2) we adopt a pluralist stance. Calculi of deontic logic can be enriched at will with various conceptions of permission, depending on the normative context for which they are devised. As long as we are aware of which intuitive concept we are dealing with, there is no problem.

2.5 A first assessment of SDL

Some of the 'puzzles' mentioned above constitute no real problem for **SDL** (e.g. Ross' puzzle and the Good Samaritan). For others, we seem to need additional expressive resources. Chisholm's puzzle seems to suggest that we need a conditional obligation operator, while the gentle murderer puzzle may be better addressed in a language capable of expressing various degrees or kinds of obligation. Moreover, problems relating to the concept of permission seem to require the definition of additional permission operators for strong permission, and perhaps for free choice permission. We also noted that two of the more serious 'paradoxes', namely Chisholm's puzzle and the gentle murderer, lose their bite (at least from a technical point of view) in systems capable of accommodating OO-conflicts.

From the discussion in this chapter, we can extract two main drawbacks of **SDL**. First and foremost, **SDL** is incapable of accommodating normative conflicts due to its validation of (OO-EX) and (OP-EX). Second, **SDL** lacks the expressive means to model some key features of our normative reasoning.

In the remainder of this thesis, we will tackle the first criticism. In chapters 5-7 we present a number of non-monotonic weakenings of **SDL** that overcome this problem. Nonetheless, **SDL** will have an important role to play in this thesis. Although its rules and axioms are not infallible, we argue in later chapters that **SDL** functions as a *standard of deduction* for deontic logic, the inferences of which have intuitive appeal and are valid in a *defeasible* manner.

The second drawback, concerning expressivity, can in principle be overcome by *enriching* the language of **SDL** with additional expressive resources, e.g. additional obligation and/or permission operators, so as to tackle some of the puzzles mentioned above. The resulting enrichments would still be 'standard' under Åqvist's umbrella conception from Section 2.2.4. As such, its lack of expressivity need not be fatal to **SDL**, although, admittedly, there is much work to be done in devising the appropriate enrichments.

SDL is an extremely simple and elegant tool for reasoning about norms. We think it remains intelligible as a 'standard' system of deontic logic, provided that we weaken it a bit by taking its inferences to hold in a defeasible manner, and provided that we allow for it to be enriched with additional expressive resources when required by the context of application at hand. For now, this suffices as a first assessment of the standard system of deontic logic.

Chapter 3

Avoiding explosion

> To see the harm of inconsistency, imagine that you wake up with contradictitis – a dreaded condition that makes you believe *not A* whenever you believe *A* [...] Contradictitis would be a living hell
>
> Harry Gensler [61, p. 30]

✎ Section 3.4 of this chapter is based on the paper *A Unifying Framework for Reasoning about Normative Conflicts* (in Michal Peliš and Vít Punčochář (eds.), *The Logica Yearbook 2011*) [27].

✎ I am indebted to Joke Meheus and Christian Straßer for valuable comments on this chapter.

In Chapter 1 we showed that there are irresolvable normative conflicts. In Chapter 2 it became clear that, when formalized as OO- or OP-conflicts, **SDL** cannot accommodate normative conflicts. In this chapter, we extend the discussion of Section 2.3 to deontic logic in general, and present some first proposed solutions to the problems posed by normative conflicts in deontic logic.

In Section 3.1 we define two explosion principles that are more refined than those presented in the previous chapter, and that will be very useful later on. In Section 3.2 we discuss the main strategies for avoiding (deontic) explosion. The list of approaches and systems evaluated in this chapter is not exhaustive. For instance, we limit the discussion in Section 3.2 to monotonic proposals. The discussion of technically more involving non-monotonic approaches is postponed until later chapters.

In Section 3.3, we distil from the preceding discussion a number of design requirements for devising conflict-tolerant deontic logics (CTDLs). Later proposals will then be tested against these desiderata.

Constructing a CTDL can be done in many ways. From this multitude of approaches we have no preference for one particular account. Instead, we embrace the pluralism and adopt a context-dependent approach. We conclude this chapter with some thoughts on this form of logical pluralism (Section 3.4).

3.1 Deontic explosion

In Section 2.3.2 we stated the explosion principles (OO-EX) and (OP-EX), and showed how they are **SDL**-valid in view of (D), (P), and (ECQ). A more refined, yet equally unintuitive explosion principle is that of *deontic explosion* [69]. Deontic explosion occurs when from a normative conflict it follows that everything is obligatory. The deontic explosion principle for OO-conflicts is given by (OO-DEX). For OP-conflicts, it is given by (OP-DEX):

$$\mathsf{O}A \wedge \mathsf{O}\neg A \vdash_{\mathbf{SDL}} \mathsf{O}B \tag{OO-DEX}$$

$$\mathsf{O}A \wedge \mathsf{P}\neg A \vdash_{\mathbf{SDL}} \mathsf{O}B \tag{OP-DEX}$$

Clearly, (OO-DEX) and (OP-DEX) are **SDL**-valid, since they are instances of (OO-EX) and (OP-EX) respectively. What's interesting about (OO-DEX) and (OP-DEX), however, is that they also arise in logics weaker than **SDL**. If, for instance, we were to remove (D) from **SDL**, (OO-DEX) would still be valid, whereas (OO-EX) would not be.[1]

(OO-DEX) holds in any logic that validates all of **CL**, (NEC), and (K). From the **CL**-theorem $A \supset (\neg A \supset B)$, it follows by (NEC) that $\vdash_{\mathbf{SDL}} \mathsf{O}(A \supset (\neg A \supset B))$. By (K), we get $\vdash_{\mathbf{SDL}} \mathsf{O}A \supset \mathsf{O}(\neg A \supset B)$ Suppose now that $\mathsf{O}A$. By (MP), $\mathsf{O}(\neg A \supset B)$. By (K), $\mathsf{O}\neg A \supset \mathsf{O}B$. Suppose that $\mathsf{O}\neg A$. Then, by (MP) again, $\mathsf{O}B$.

In Section 3.2, we present and discuss various ways of making sure that principles like (OO-DEX) and (OP-DEX) do not arise. Some clues as to how to proceed are clear already: from every set of principles that gives rise to (OO-(D)EX) or (OP-(D)EX), at least one principle must be restricted or given up in order to obtain a conflict-tolerant deontic logic.

3.2 Strategies for avoiding explosion

In the literature on normative conflicts, there is a strong tendency to focus on OO-conflicts. Presumably, this is due to the sense of moral urgency surrounding inescapable conflicts between two or more (moral) obligations. This tendency is reflected in the structure of this section, although we will occasionally extend the discussion to other types of conflicts.

There are two main strategies for averting the validity of (OO-(D)EX). The first is to enrich the formal language of **SDL** in order to distinguish between various features in view of which normative conflicts arise. The second strategy is to weaken **SDL** by rejecting or restricting some of its axiom schemas and/or inference rules.

[1] **SDL** without (D) would just be the deontic variant of the basic normal modal logic **K**. Thus, (OO-DEX) and (OP-DEX) are valid in any normal modal logic. Referring back to the discussion in Section 2.2.4, it is now clear that (OO-EX) holds only in *strongly* normal propositional monadic von Wright-type deontic logics, whereas (OO-DEX) holds in *all* normal propositional monadic von Wright-type deontic logics.

3.2.1 Formalizing normative conflicts in richer formal languages

3.2.1.1 Relativizing deontic operators

In [44], Hector-Neri Castañeda proposes to use indexed deontic operators in order to formalize normative conflicts. By relativizing the axioms of deontic logic to the agents in view of which norms hold, he avoids difficulties with conflicting norms – as long as those norms arise from different agents. Consider the following example:

Example 11. (i) Insofar as you promised to Jones, it is obligatory that you wait for Jones' friend ($\mathsf{O}f$). However, (ii) insofar as you promised to your wife, it is obligatory that you do not wait for Jones' friend ($\mathsf{O}\neg f$).

In formalizing (i) as $\mathsf{O}f$ and (ii) as $\mathsf{O}\neg f$, we end up with an OO-conflict, which causes explosion in **SDL**. Moreover, we lose the information that the obligation $\mathsf{O}f$ holds in view of your promise to Jones, whereas $\mathsf{O}\neg f$ holds in view of your promise to your wife. As a solution to these problems, Castañeda suggested that we relativize our obligations and their logical properties to the sources from which they arise. Following this suggestion, we can formalize (i) as $\mathsf{O}_j f$, and (ii) as $\mathsf{O}_w \neg f$, where the subscripts j and w denote 'Jones' and 'wife' respectively. Since in Castañeda's system we can no longer employ (AND) for aggregating obligations with different subscripts, explosion is avoided. Moreover, the subscripting gives us a more informative formalization.

Castañeda's proposal was taken up independently by Schotch and Jennings in [157]. The method of adding sub- and/or superscripts to deontic operators can also be used for indicating authorities, normative standards and/or interest groups in view of which (conflicting) norms hold. In a multi-agent setting it was followed by Kooi and Tamminga in [107]. The latter authors use indices for representing bearers and interest groups. In a legal setting, the idea was taken up by Herrestad and Krogh in [87], where deontic operators are relativized to their bearers and counterparties.

Another way of dealing with normative conflicts by enriching the expressive power of **SDL** is to introduce a preference ordering on our obligations and permissions, e.g. [81]. Doing this allows us to model situations in which more binding obligations or permissions override less binding ones. Yet another extension of **SDL** consists in making its deontic operators dyadic in order to properly express under which conditions our obligations and permissions hold true (cfr. Section 2.4.1).

The main tenet of all these proposals is the following: in making our formal language more expressive, we can distinguish in our formalization between different features of conflicting norms. In the words of Castañeda: "a conflict of duties is the truth of a conjunction of the form $\mathsf{O}_i A \wedge \mathsf{O}_j B$, where A and B are at least causally incompatible and $i \neq j$" [44, p. 264]. In the richer language, formalizations of this kind no longer cause (deontic) explosion.

These enrichments are very successful in increasing the expressive power of **SDL**. Furthermore, they effectively allow us to consistently model conflicts between norms of different hierarchies, norms issued by different authorities, norms arising from different normative standards, norms that hold in different circumstances, etc. However, two main problems remain.

The first problem is that in merely relativizing deontic operators we can no longer aggregate obligations with different indices. Suppose, for instance, that we use different indices for representing the sources in view of which your promises arise. Consider the following example.

Example 12. Inasmuch as you promised to Jones, it is obligatory that you invite him for your birthday party ($\mathsf{O}_j j$). Moreover, inasmuch as you promised to your wife, it is obligatory that you do not invite both Jones and Smith for your birthday party ($\mathsf{O}_w \neg(j \wedge s)$).

Provided that this is all the information available to you, it makes sense for you to conclude that you ought not invite Smith for your birthday party (for this is the only way for you to fulfill the promises you have made to Jones and to your wife). However, if, as in the example, subscripts are added to your obligations, and there is no way of aggregating obligations with different subscripts, the subscripted logic will not lead you to this conclusion.

The example illustrates that if we want to use indices for consistently allowing for normative conflicts, and if we want our logic to be capable of modeling actual normative reasoning, we will also need some sort of 'overarching' aggregation rule for aggregating obligations with different subscripts. Clearly, such a rule must somehow be restricted so as to avoid the aggregation of obligations that are, or that turn out to be, conflicting. This is a difficult, but not unsolvable problem which we will tackle in Section 5.2.

The second (and worse) problem is that the richer formal languages discussed above are still insufficiently rich for accommodating *all* normative conflicts. As explained in Section 1.6.1 normative conflicts may be irresolvable for reasons of symmetry. Conflicts may arise in which there are no relevant differences between each of the alternatives. In such cases, no subscripts/indices are available for distinguishing between different features of conflicting obligations. In the words of Forrester: "The promise of subscripting then proves to be an illusion. Subscripting does not prevent irresoluble conflicts of obligation from occurring, nor does it explain away all of the conflicts that seemingly do occur" [56, p. 41].

We conclude with a remark arising from practical, rather than philosophical considerations. There are contexts of application in which we simply lack the necessary formal means for distinguishing between various features of conflicting obligations. In a legal context, for instance, existing principles such as the aforementioned *lex specialis, lex posterior*, etc. may not be of any help. By the complexity of the world and by mere human mistakes, conflicts may arise between norms promulgated at the same time, by the same authority, etc. In the words of Alchourrón and Bulygin:

> Even one and the same authority may command that p and that *not* p at the same time, especially when a great number of norms are enacted on the same occasion. This happens when the legislature enacts a very extensive statute, e.g. a Civil Code, that usually contains four to six thousand dispositions. All of them are regarded as promulgated at the same time, by the same authority, so that there is no wonder

> that they sometimes contain a certain amount of explicit or implicit contradictions. [3, pp. 112-113]

3.2.1.2 Alethic modalities

Many normative conflicts arise from the *impossibility* to fulfill two or more obligations. This has led some philosophers to make use of both deontic and alethic modalities in formalizing normative conflicts. A conflict is then represented as a situation in which it is obligatory that A_1 and ... and obligatory that A_n, but in which it is impossible to realize all of $A_1, \ldots, A_n$. In symbols: $\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n \wedge \neg \Diamond (A_1 \wedge \ldots \wedge A_n)$ (where, following the usual convention, $\Diamond A$ abbreviates "it is possible that A").

Formalizations that make use of the alethic operator $\Diamond$ were presented in e.g. [88, 164, 203]. Bernard Williams in particular preferred to formalize normative conflicts by making use of the diamond, at least in a moral context. According to Williams, the basis of moral conflicts is contingent in the sense that it is the world, not logic, that makes it impossible for two (or more) conflicting obligations to be satisfied; we can consistently imagine a state of affairs in which they could all be satisfied, but the present factual situation makes it impossible to do so. Williams' concern lies *only* with conflicts that have a contingent basis, with conflict *via* the facts, and not with conflicts between logically incompatible obligations:

> I shall further omit any discussion of the possibility (if it exists) that a man should hold moral principles or general moral views which are intrinsically inconsistent with one another, in the sense that there could be no conceivable world in which anyone could act in accordance with both of them; as might be the case, for instance, with a man who thought that he ought not to go in for any blood-sport (as such) and that he ought to go in for foxhunting (as such). I doubt whether there are any interesting questions that are peculiar to this possibility. [203, p. 108]

Williams argues that, in case two moral obligations conflict, a situation should always be conceivable in which the very same obligations can be consistently satisfied. Moral conflicts between logically incompatible obligations, if they exist at all, are at best uninteresting.

Understood in this way, moral conflicts can take two basic forms: "One is that in which it seems that I ought to do each of two things, but I cannot do both. The other is that in which something which (it seems) I ought to do in respect of certain of its features also has other features in respect of which (it seems) I ought not do it"[203, p. 108].

Next, Williams argues that conflicts of the second form are reducible to conflicts of the first form. He illustrates this reduction by means of Example 1 from Section 1.3.1. In this example, the roots of the conflict are exposed by acknowledging that the conflict arises from the contingent incompatibility of Agamemnon's duties as a commander, respectively as a parent. Given this acknowledgment, Williams believes we can formalize Agamemnon's dilemma by making use of alethic modalities: "here again there is a double ought: the first, to further the expedition, the second, to refrain from the killing; and that as things are he [Agamemnon] cannot discharge both" [203, p. 119]. Seen in this way, the real roots of OO-conflicts are no longer concealed, and a more realistic picture is offered of how the situation is. As an upshot, moral conflicts need no longer wear the form of an inconsistency of the type "ought-ought not", and all moral conflicts are, ultimately, of the form $\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n \wedge \neg \Diamond (A_1 \wedge \ldots \wedge A_n)$.

Let us point to two extra considerations that arise in view of Williams' characterization of moral conflicts. First, it is important to realize that Williams' characterization is not generalizable to normative contexts in general. In legal contexts, for instance, it is perfectly possible that, in some specific situations, the law considers it mandatory for someone to act in two logically incompatible ways. Similarly, an authority may command someone to do the logically impossible. Williams does not argue against (the use of) constructing formal calculi that consistently allow for the presence of conflicts like these, which arise through human fault. Instead, Williams claims that there is nothing *morally* relevant to say in such cases.

Second, an extra difficulty arises from the use of the diamond operator: a new argument for explosion becomes available in view of the principle that 'ought' implies 'can' (OIC):

$$\mathsf{O}A \supset \Diamond A \tag{OIC}$$

Suppose that you are facing two obligations $\mathsf{O}A$ and $\mathsf{O}B$, and that it is impossible for you to fulfill both obligations ($\neg\Diamond(A\wedge B)$). By (AND), we can derive $\mathsf{O}(A\wedge B)$ from $\mathsf{O}A$ and $\mathsf{O}B$. By (OIC) and **CL**, however, we can derive $\neg\mathsf{O}(A \wedge B)$ from $\neg \Diamond (A \wedge B)$, which contradicts $\mathsf{O}(A \wedge B)$. Hence, if (OIC) is a valid principle of deontic reasoning, then moral conflicts as formalized by Williams cause explosion all over again:

$$\mathsf{O}A, \mathsf{O}B, \neg \Diamond (A \wedge B) \vdash \bot \tag{$\Diamond$-EX}$$

Altogether, Williams' formalization of normative conflicts is restricted in scope (i.e. confined to the moral context) and – in combination with (OIC) – explosive when combined with **SDL**. In Chapter 4, we use a logic that makes use of alethic modalities as an illustration of the standard format for adaptive logics. In Section 5.1 we come back to Williams' approach and discuss his solution for avoiding the validity of ($\Diamond$-EX).

3.2.2 Weakening SDL

Although they succeed in consistently formalizing many instances of normative conflicts, the approaches presented in Section 3.2.1 cannot accommodate each and every such instance. This is due to their verification of full **SDL**.

Instead of enriching the language of **SDL** in order to distinguish between different features of conflicting norms, we can also weaken **SDL** in order to increase the degree of conflict-tolerance of our logic. We already noted, for instance, how removing (D) from **SDL** results in a logic that invalidates (OO-EX). In this section, we evaluate some suggestions along this line. In the literature on **SDL**-weakened deontic logics we can distinguish between three dominant strategies for accommodating normative conflicts. The first strategy consists in restricting the aggregation schema (AND). The second strategy proceeds by restricting the inheritance principle (RM), and the third strategy is that of replacing **CL** with a weaker logic that invalidates (ECQ), and by building a deontic logic on top of this weaker (paraconsistent) alternative. After presenting these main strategies, we also consider some hybrid proposals made in the literature.

3.2.2.1 Weakening aggregation

Restricting or rejecting the aggregation schema (AND) is an intuitive solution for accommodating normative conflicts in deontic logic. Bernard Williams, for instance, argued that an agent facing two conflicting obligations thinks that she should fulfill each of these obligations, but not both of them [203, p. 120]. Lou Goble too stated that giving up (AND) is "perhaps the most natural suggestion for avoiding deontic explosion" [69, p. 466]. In several papers Goble advocated the use of one particular non-aggregative logic, namely the logic **P** [65, 66, 67]. **P** is a very well-behaved system and has a natural interpretation in a Kripke-like semantics.[2] Its semantics was constructed independently from Goble by Schotch & Jennings [157]. A system closely akin to **P** was also axiomatized by van Fraassen in [58].[3]

For the language $\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}$, **P** is axiomatized by adding to **CL** the rules (NEC), (RM), and (PN):

$$\text{If } \vdash A, \text{ then } \vdash \neg\mathsf{O}\neg A \tag{PN}$$

Where $\Gamma \subseteq \mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}$, we write $\Gamma \vdash_{\mathbf{P}} A$ to denote that A is **P**-derivable from Γ, and $\vdash_{\mathbf{P}} A$ to denote that A is **P**-derivable from the empty premise set.[4]

Semantically, a **P**-model M is a quadruple $\langle W, \mathcal{R}, v, w_0\rangle$ where W is a set of possible worlds, $\mathcal{R}$ is a non-empty set of serial accessibility relations R on W, $v : \mathcal{W}^a \times W \to \{0,1\}$ is an assignment function, and $w_0 \in W$ is the designated world. The valuation v_M defined by the model M is characterized by adding the clause (CO') to the **CL**-clauses (Ca)-(C$\equiv$) from Section 2.2.2:

(CO') $v_M(\mathsf{O}A, w) = 1$ iff, for some $R \in \mathcal{R}$, $v_M(A, w') = 1$ for all w' such that Rww'

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

A permission operator P for the logic **P** is defined as $\mathsf{P}A =_{\text{df}} \neg\mathsf{O}\neg A$. In [65], Goble proved soundness and (weak) completeness for **P**.

Next to its elegance and simplicity, a main advantage of **P** is that it invalidates (OO-DEX):

$$\mathsf{O}A \wedge \mathsf{O}\neg A \nvdash_{\mathbf{P}} \mathsf{O}B \tag{3.1}$$

A first disadvantage is that **P** is not fully conflict-tolerant. Suppose, for instance, that $\mathsf{O}A \wedge \mathsf{P}\neg A$. Due to the definition of its P-operator, $\mathsf{P}\neg A$ is **P**-equivalent to $\neg\mathsf{O}\neg\neg A$. Due to (RM) and the **CL**-theorem $A \supset \neg\neg A$, we can derive $\mathsf{O}\neg\neg A$ from

[2] Goble also proposed a preferential semantics for **P** in [65, 66].

[3] The main difference between van Fraassen's system and **P** is that the latter validates $\mathsf{O}\top$ for all tautologies $\top$, while the former does not (see also [65, footnote 3]).

[4] Goble is only interested in the theorems of his logic, not in a consequence relation. As we are mainly interested in the consequence relation, we define one for **P**. Semantically, we slightly modify Goble's semantics in such a way that we introduce a designated world in the models.

$\mathsf{O}A$, and we obtain a contradiction. Since **P** contains all of **CL**, it follows that $\bot$. Hence, ($\mathsf{O}\mathsf{P}$-EX) is valid in **P**.

A second disadvantage of **P** is that it is too weak to account for our everyday deontic reasoning. As an illustration, consider the following example by Horty [91]:

Example 13. An agent, Smith, is confronted with two obligations. First, it is obligatory that Smith fight in the army or perform alternative service to his country ($\mathsf{O}(f \vee s)$). Second, it is obligatory that Smith does not fight in the army ($\mathsf{O}\neg f$). The first obligation follows from Smith's duties as a citizen, whereas the second arises from his pacifist convictions. No conflict seems to be present between Smith's obligations: he can safely fulfill both simply by performing alternative service to his country. Hence, it is obligatory that Smith perform alternative service to his country ($\mathsf{O}s$).

The inference drawn in Example 13 is **SDL**-valid: $\mathsf{O}(f \vee s), \mathsf{O}\neg f \vdash_{\mathbf{SDL}} \mathsf{O}s$. However, **P** invalidates the inference: $\mathsf{O}(f \vee s), \mathsf{O}\neg f \nvdash_{\mathbf{P}} \mathsf{O}s$.

Making abstraction from the fact that **P** cannot accommodate $\mathsf{O}\mathsf{P}$-conflicts, its main problem appears to be that if (AND) is rejected in its entirety, we end up with a logic that is too weak. This led Horty to the observation that:

> Apparently, what is needed is some degree of agglomeration [aggregation], but not too much; and the problem of formulating a principle allowing for exactly the right amount of agglomeration [aggregation] raises delicate issues that have generally been ignored in the literature, which seems to contain only arguments favoring either wholesale acceptance or wholesale rejection. [95, p. 580]

In his [69], Goble is sceptical of the very idea of an aggregation principle that allows for "exactly the right amount of aggregation". He discusses various approaches (including an alternative presented by Horty), and comes to the conclusion that none of them lives up to this daunting task. Below, we briefly discuss two such proposals together with Goble's criticism.

First, suppose that we restrict (AND) by imposing a further consistency requirement which results in the 'consistent aggregation' rule (CAND):

$$\text{If } \nvdash A \supset \neg B, \text{ then } (\mathsf{O}A \wedge \mathsf{O}B) \supset \mathsf{O}(A \wedge B) \tag{CAND}$$

Although this suggestion appears natural, it is much too strong. In the presence of a normative conflict $\mathsf{O}A \wedge \mathsf{O}\neg A$ and some random formula B such that $\nvdash \neg B$, (CAND) allows one to derive $\mathsf{O}(A \wedge (\neg A \vee B))$, from which follows $\mathsf{O}B$. Hence (CAND) gives rise to the explosion principle

$$\text{If } \nvdash \neg B, \text{ then } \mathsf{O}A, \mathsf{O}\neg A \vdash \mathsf{O}B \tag{3.2}$$

Second, Goble discusses the weakened aggregation rule of permitted aggregation (PAND):

$$\mathsf{P}(A \wedge B) \supset ((\mathsf{O}A \wedge \mathsf{O}B) \supset \mathsf{O}(A \wedge B)) \tag{PAND}$$

Instead of allowing aggregation for obligations that are jointly compatible (as is the case for (CAND)), this alternative allows aggregation for obligations that are jointly permissible. Unfortunately (PAND) suffers from problems very similar to those of (CAND). Whereas in the case of (CAND) explosion follows from an OO-conflict $\mathsf{O}A \wedge \mathsf{O}\neg A$ in the presence of some contingent formula $\neg B$, in the case of (PAND) explosion arises when faced with an OO-conflict $\mathsf{O}A \wedge \mathsf{O}\neg A$ in the presence of some formula B such that B is permitted. To see why, note that $\mathsf{P}B \equiv \mathsf{P}((A \vee B) \wedge (\neg A \vee B))$, that from $\mathsf{O}A$ it follows that $\mathsf{O}(A \vee B)$ and that from $\mathsf{O}\neg A$ it follows that $\mathsf{O}(\neg A \vee B)$. Hence, by (PAND), we obtain $\mathsf{O}((A \vee B) \wedge (\neg A \vee B))$ (which is equivalent to $\mathsf{O}B$) from $\mathsf{O}A \wedge \mathsf{O}\neg A$. This yields

$$\mathsf{O}A, \mathsf{O}\neg A \vdash \mathsf{P}B \supset \mathsf{O}B \tag{3.3}$$

Both proposed weakenings ((CAND) and (PAND)) illustrate that avoiding unwanted consequences when weakening the aggregation rule of **SDL** is very difficult. Goble discusses two more classes of solutions to the aggregation-problem. The first is that of 'constrained consistent aggregation', as proposed by Horty [91, 92, 95] and van Fraassen [58].[5] As this approach is non-monotonic and technically more involving, we postpone a more detailed discussion of it until Section 5.3. The second solution is that of 'two-phase deontic logic', as proposed by van der Torre and Tan [185]. We discuss this proposal in Section 3.2.2.4.

In [63], Goble proposes a restricted aggregation schema that allows for the application of (AND) *unless* one of the formulas to be aggregated *or* a subformula of one of these formulas is 'tainted' by an OO-conflict. Let $\mathrm{U}A = \neg(\mathsf{O}A \wedge \mathsf{O}\neg A)$ abbreviate that A is unconflicted. Where $B_1, \ldots, B_n$ are all subformulas of A (including A itself), we write $\mho A$ to abbreviate the conjunction $\mathrm{U}B_1 \wedge \ldots \wedge \mathrm{U}B_n$. The *ultra-unconflicted aggregation* schema (UU-AND) is given by:

$$\mho(A \wedge B) \supset ((\mathsf{O}A \wedge \mathsf{O}B) \supset \mathsf{O}(A \wedge B)) \tag{UU-AND}$$

As opposed to (CAND) and (PAND), (UU-AND) seems to do the job. Nevertheless, (UU-AND) is not a very 'natural' aggregation rule. In complex settings, it requires a lot of calculations to know whether or not we can aggregate two obligations. We return to this point in Section 3.3.2.

3.2.2.2 Weakening inheritance

A second way of weakening **SDL** so as to make it more conflict-tolerant is to weaken the inheritance principle (RM). Like the first one, this approach is 'classical' in the sense that full **CL** remains valid in the resulting weakened logic.

A solution along this line was explicitly advocated by Goble in [68, 69], where he defined his family of **DPM**-systems. On the one hand, weakening (RM) seems intuitive due to the 'paradoxes' that hinge on this principle (cfr. Section 2.4). On the other hand, rejecting all instances of (RM) results in a very weak logic. The inference from $\mathsf{O}(A \wedge B)$ to $\mathsf{O}A$, for instance, fails when (RM) is given up.

In the **DPM**-systems, (RM) is replaced by a rule of *permitted inheritance* (RPM):

$$\text{If } \vdash A \supset B, \text{ then } \vdash \mathsf{P}A \supset (\mathsf{O}A \supset \mathsf{O}B) \tag{RPM}$$

[5] A conditional version of the proposals of Horty and van Fraassen was presented by Hansen in [75].

Moreover, (D) is invalid in all members of the **DPM**-family. In order to apply the weakened inheritance principle (RPM) to a **CL**-theorem $A \supset B$ and an obligation $\mathsf{O}A$, the logic requires that A is permitted.[6] For instance, in order to apply (RPM) to $\mathsf{O}(p \wedge q)$ in order to derive $\mathsf{O}q$, we also need $\mathsf{P}(p \wedge q)$. Since (P) is **DPM**-valid, this means that the obligation $\mathsf{O}(p \wedge q)$ cannot be involved in a normative conflict whenever (RPM) is applicable.

In cases in which the required permission statements are not derivable from the premises by means of **DPM**, we are faced with a dilemma. If we would add these permissions to the premise set, we run the risk of causing explosion. If we do not add them, then (RPM) is not applicable and we end up with a very weak consequence set. This is suboptimal for various reasons, which we will discuss in more detail in Section 3.3.3.

The problems concerning the applicability of (RPM) are resolved by the adaptive extensions of some systems in the **DPM**-family proposed in [175].

3.2.2.3 Going paraconsistent

The 'classical' approaches presented in Sections 3.2.2.1 and 3.2.2.2 may succeed in making deontic logic OO-conflict-tolerant, but OP-conflicts cannot be consistently allowed for by merely restricting or even rejecting (AND) and/or (RM). The reason is that, in view of the interdefinability principle (P) and **CL**, every OP-conflict $\mathsf{P}A \wedge \mathsf{O}\neg A$ is equivalent to a contradiction $\neg\mathsf{O}\neg A \wedge \mathsf{O}\neg A$. Thus, either the interdefinability of O and P must be given up in order to accommodate OP-conflicts, or we need a logic that invalidates certain axiom schemas and/or rules of **CL**.

Moreover, remember from Section 2.3.2 that, if full **CL** is kept valid, either (P) or (D) must be restricted or given up in order to avoid (OO-EX). In view of these considerations it seems reasonable to try and weaken **CL** in order to consistently allow for the presence of both OO- and OP-conflicts. The most obvious way to do so is to weaken the **CL**-negation to a paraconsistent negation connective. A logic is *paraconsistent* if it invalidates the schema (ECQ), i.e. if it consistently tolerates contradictions.

Several authors have presented paraconsistent deontic logics in order to account for normative conflicts. Da Costa & Carnielli [48], McGinnis [125, 124] and Priest [145] reject (ECQ) by weakening the negation of **CL**. Routley & Plumwood [154] reject (ECQ) by using a relevant implication instead of material implication.

An extra argument in favor of a paraconsistent approach is that paraconsistent deontic logics are capable of tolerating *contradictory permissions* and *contradictory obligations*: formulas of the forms $\mathsf{P}A \wedge \neg\mathsf{P}A$ or $\mathsf{O}A \wedge \neg\mathsf{O}A$ respectively. Contradictory obligations and permissions may look like exotic beasts, but in certain contexts of application (e.g. legal contexts, logics of command) it seems reasonable to take into account the possibility of contradictory norms. Consider the following example from [145, pp. 184-185].

[6] In some other systems in the **DPM**-family, this is also required for applying a weakened aggregation principle, cfr. Section 3.2.2.4.

Example 14. Suppose that, in some country, (i) women are not permitted to vote, and (ii) property holders are permitted to vote. Suppose further that (possibly due to a recent revision of the property law) women are allowed to hold property. Then (i) and (ii) cause an inconsistency in case there exists a female property holder, since the latter is both permitted and not permitted to vote ($\mathsf{P}v \wedge \neg\mathsf{P}v$).[7]

The drawback of the paraconsistent deontic logics presented in [48, 145, 154] is that they are rather weak. For instance, they cannot account for Example 13 from Section 3.2.2.1. The reason is that many paraconsistent logics invalidate principles like disjunctive syllogism (from $A \vee B$ and $\neg A$ to derive B), contraposition (from $\neg B$ and $A \supset B$ to derive $\neg A$), and even modus ponens.

In Chapters 6 and 7 we present some adaptive logics with a paraconsistent negation connective. These logics are both fully conflict-tolerant and capable of accounting for Example 13 and other 'toy' examples from the literature.

3.2.2.4 Mixed proposals

Of course, a CTDL need not merely weaken (AND), (RM), or **CL**. Some authors have suggested combined proposals. As mentioned, van der Torre and Tan opted for such an approach in [185]. In their 'two-phased' system of deontic logic, two O-operators O_1 and O_2 are distinguished. (CAND) applies to the first, whereas (RM) applies to the second. Both operators are related by the schema $\mathsf{O}_1 A \supset \mathsf{O}_2 A$.

For Example 13, the resulting two-phased system yields $\mathsf{O}_1(f \vee \neg s), \mathsf{O}_1\neg f \vdash \mathsf{O}_2 s$, whereas $\mathsf{O}_1(f \vee \neg s), \mathsf{O}_1\neg f \nvdash \mathsf{O}_1 s$ and $\mathsf{O}_2(f \vee \neg s), \mathsf{O}_2\neg f \nvdash \mathsf{O}_2 s$. With Goble, we agree that there is a certain ambiguity about this two-phased approach, and that it is not at all obvious

> that there is any such ambiguity of 'ought' as it occurs in the discourse that gives us the argument about Smiths service and inclines us to accept its being valid. Nor is it obvious that, as that argument is given, the premises should be taken in the first of the two senses rather than the second, even while the conclusion is taken in the second and not the first. [69, p. 471]

In the very same paper, Goble presents his own 'mixed' variant. It concerns the logic **DPM.2** of his **DPM** family (cfr. Section 3.2.2.2) in which both (RM) and (AND) are restricted. More precisely, (RM) is restricted to (RPM), while (AND) is restricted to (PAND).[8]

DPM.2 faces pretty much the same problem as its relatives in the **DPM**-family. In order to apply the weakened inheritance principle the user has to "manually" add permission statements. For **DPM.2**, we need to do this not only in order to apply (RPM), but also in order to apply (PAND). As promised, we will return to this problem in Section 3.3.3.

[7] In line with the discussion in Chapter 6, the P-operator should be interpreted prescriptively here.

[8] In [175] a variant of this logic is presented in which (PAND) is replaced by the schema (PAND'), $(\mathsf{O}A \wedge \mathsf{O}B \wedge \mathsf{P}A \wedge \mathsf{P}B) \supset \mathsf{O}(A \wedge B)$.

This concludes our overview of the different strategies for devising CTDLs. The discussion so far is restricted to monotonic solutions. In later chapters, we will assess the more involving (yet more promising) non-monotonic CTDLs. Why a non-monotonic approach is more promising than a monotonic one, will be clear by the end of this chapter.

3.3 Design requirements

Before we move on to present the standard format for adaptive logics and the CTDLs defined within this framework, we round up the above discussion by stating three desiderata for CTDLs.

3.3.1 Non-explosiveness

Clearly, any adequate CTDL should invalidate principles like (OO-EX) and (OO-DEX). Depending on the context, it might also be required to invalidate (OP-EX) and (OP-DEX). However, the cases of (CAND) and (PAND) show that we must also be on guard for 'weaker', more refined explosion principles. A logic that merely restricts (AND) to (CAND), validates the inference (3.2). A logic that merely restricts (AND) to (PAND), validates the inference (3.3). Both of these inferences pose serious problems for the systems in question.

In [174], some other more refined explosion principles were specified that can serve as benchmarks for measuring the conflict-tolerance of various deontic logics. Here are some examples:

$$\mathsf{O}A, \mathsf{O}\neg A \vdash \mathsf{O}B \vee \mathsf{O}\neg B \tag{3.4}$$

$$\mathsf{O}A, \mathsf{O}\neg A \vdash \mathsf{O}B \vee \mathsf{P}B \tag{3.5}$$

$$\mathsf{O}A, \mathsf{O}\neg A \vdash \mathsf{O}B \vee \neg\mathsf{O}\neg B \tag{3.6}$$

$$\mathsf{O}A, \mathsf{O}\neg A \vdash \mathsf{P}B \tag{3.7}$$

A further requirement is to demand not just that there is a non-trivial model that validates the conflicting norms, but to also impose certain normality conditions on this model. For instance, non-explosive models should also validate a non-conflicting obligation, e.g. $\mathsf{O}C$ and $\neg\mathsf{O}\neg C$, and/or a non-conflicting permission, e.g. $\mathsf{P}D$ and $\neg\mathsf{O}\neg D$, and/or there should be a proposition E such that neither E nor $\neg E$ is obligatory, i.e. $\neg\mathsf{O}E \wedge \neg\mathsf{O}\neg E$, and/or there is a proposition F such that both, F and $\neg F$, are allowed, i.e. $\mathsf{P}F \wedge \mathsf{P}\neg F$. These conditions obviously hold for the real world, so there should also be interpretations of deontic conflicts that satisfy these criteria. We denote such refinements by adding the additional requirements in set brackets after the basic principle, for instance, where $\gamma = \{\neg\mathsf{O}E, \neg\mathsf{O}\neg E\}$ and $\gamma' = \{\mathsf{O}C, \neg\mathsf{O}\neg C, \mathsf{P}D, \neg\mathsf{O}\neg D, \neg\mathsf{O}E, \neg\mathsf{O}\neg E, \mathsf{P}F, \mathsf{P}\neg F\}$,

$$\{\mathsf{O}A, \mathsf{O}\neg A\} \cup \gamma \vdash \mathsf{O}B \vee \mathsf{P}B \tag{3.8}$$

$$\{\mathsf{O}A, \mathsf{P}\neg A\} \cup \gamma' \vdash \mathsf{O}B \vee \mathsf{P}B \tag{3.9}$$

Indeed, any *truly* conflict-tolerant logic should be tolerant concerning any of these principles.

Another more fine-grained explosion principle involves the further requirement that non-explosive models invalidate the derivation of $\mathsf{O}B$ for any contingent formula C. For instance,

$$\text{If } \nvdash C, \text{ then } \{\mathsf{O}A, \mathsf{O}\neg A\} \cup \gamma \vdash \mathsf{O}B \vee \mathsf{P}B \tag{3.10}$$

There is no clear-cut test that a 'conflict-tolerant' deontic logic does not validate some very sophisticated, counter-intuitive explosion principle. The best we can do is to test any proposed candidate system not only for the validity of the principles (OO-EX), (OP-EX), (OO-DEX), and (OP-DEX), but also for the validity of more refined inferences like (3.4)-(3.10) above.

3.3.2 Non-monotonicity and inferential strength

A disadvantage of the systems presented in Section 3.2.2 is that, although capable of tolerating various types of normative conflicts, these systems are too weak to account for our natural deontic reasoning. All of the CTDLs discussed above are *monotonic* (remember that a logic $\mathbf{L}$ is monotonic iff, for all sets of $\mathbf{L}$-wffs Γ and Γ', $\mathrm{Cn}_{\mathbf{L}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{L}}(\Gamma \cup \Gamma')$).

Consider now the following inferences:

(a)	$\mathsf{O}(\neg p \vee q)$	(a')	$\mathsf{O}(\neg p \vee q)$
(b)	$\underline{\mathsf{O}p}$	(b')	$\mathsf{O}p$
(c)	$\mathsf{O}q$	(c')	$\underline{\mathsf{O}\neg p}$
		(d')	$\mathsf{O}q$

On the one hand, the inference from (a) and (b) to (c) seems reasonable as an instance of the *deontic disjunctive syllogism* (DDS) schema,

$$\mathsf{O}(\neg A \vee B), \mathsf{O}A \vdash \mathsf{O}B \tag{DDS}$$

In Example 13 for instance, we wanted to apply this inference in order to attain the intuitively correct conclusion. The inference from (a')-(c') to (d') on the other hand, seems dubious. In order to derive (d'), we need to rely essentially on premise (b'), which is directly involved in an OO-conflict. In view of the conflicting obligations at lines (b') and (c'), it seems better not to derive $\mathsf{O}q$ for at least two reasons.

First, we do not want to draw any conclusions from conflicting obligations. Instead, we want to *isolate* whichever premises behave abnormally, and use only the non-conflicting part of our premise set for inferring new conclusions.

The second reason is of a technical nature. Note that in any logic which validates the *deontic addition* (DA) schema,

$$\mathsf{O}A \vdash \mathsf{O}(A \vee B) \tag{DA}$$

premise (a') is derivable from (c'). If a logic validates both (DA) and (DDS), then (d') is derivable from (a') and (b'). But then (OO-DEX) is valid in this logic, since we have derived (d') from (b') and (c').

If normative conflicts are present, there is a trade-off between isolating conflicts in order to avoid explosion on the one hand, and allowing for those inferences that are intuitive in the absence of conflicts on the other hand. Some inferences sometimes ought to be blocked in order to avoid explosion, while at other times their application is harmless. The application of (DDS) to derive $\mathsf{O}q$, for instance, is harmless if all we know is that $\mathsf{O}p$ and $\mathsf{O}(\neg p \vee q)$. But it is problematic if we also know that $\mathsf{O}\neg p$.

If we want to infer (c) from (a) and (b) without being able to infer (d') from (a')-(c'), then we need a *non-monotonic* logic, i.e. a logic for which some conclusions derivable from a premise set may not be derivable anymore if further premises are added. Only non-monotonic logics can overcome the trade-off between the isolation of conflicts and the inferential power necessary to model our everyday normative reasoning.

In general, non-monotonic logics are better capable of dealing with conflicts due to their flexibility when new information is added to the premises. In Chapters 5-7 we present a number of non-monotonic CTDLs devised within the adaptive logics framework for defeasible reasoning, and compare these with other non-monotonic approaches from the literature.

3.3.3 User-friendliness

In Section 3.2.2.2 we already mentioned the suboptimality of the need to 'manually' add formulas to the premise set in order for further information to become available. Here, we further explain and motivate this claim, and illustrate it by means of the **DPM** logics from Section 3.2.2.2.

In all interesting cases, determining which statements can safely be added 'manually' to a set of premises (that is, in such a way that no explosion follows) requires *reasoning*. Suppose, for instance, that $\mathsf{O}(p \wedge q)$ and that $\mathsf{P}(p \wedge q)$ is not derivable from the premise set. Then $\mathsf{O}q$ is not **DPM**-derivable from the premises, unless we add $\mathsf{P}(p \wedge q)$ so that we can apply (RPM) to $\mathsf{O}(p \wedge q)$. But adding this statement is a non-trivial deed to say the least, since it might give rise to an inconsistency when conjoined with the other premises.

The application of (RPM) is especially problematic in cases where the premises are complex, plentiful, and/or tightly interwoven. In such complicated setups it might not be obvious at all that for instance $\mathsf{O}A \wedge \mathsf{O}\neg A$ is derivable. However, suppose that in this case the user naively added $\mathsf{P}A$ to the premises in order to apply (RPM) to $\mathsf{O}A$. Since in the **DPM** systems $\mathsf{P}A$ is equivalent to $\neg\mathsf{O}\neg A$ the user gives way to explosion by adding this permission to the premise set.

In short, the reasoning required for a sensible application of the inheritance principle falls partly outside the scope of the logic. Whenever the required permission statements are not derivable from the premises, additional consistency-checks are needed in order to be able to apply (RPM) without running the risk of causing explosion.

Ideally, we would like the logic to be more 'user-friendly' and do this reasoning in our place.

3.4 Logical pluralism

The aim of this thesis is not to present and defend one particular conflict-tolerant deontic logic that allows for the consistent possibility of *all* types of normative conflicts. We believe instead that the adequacy of a given CTDL is a context-dependent matter: both its rules of inference and its degree of conflict-tolerance depend on the concrete application of the logic. Let us illustrate this claim by means of three examples, each of which is situated in a different 'deontic' context.

(1) As a first example, consider a *moral* context. In discussions on moral dilemmas, philosophers have typically focussed on conflicting obligations. Moral dilemmas are conceived as situations in which an agent ought to adopt each of two or more alternatives which are equally compelling from a moral point of view, and in which the agent cannot do both (or all) of the actions [164].

In this context, there is nothing particularly 'dilemmatic' about an OP-conflict (here, the agent can still safely fulfill all of her moral requirements, i.e. all of her obligations). Thus, a CTDL for modeling moral dilemmas need not necessarily account for the possibility of OP-conflicts.

How should **SDL** be weakened in this context? One suggestion is to reject or restrict the aggregation schema (AND). In the moral context, (AND) was disputed (amongst others) by Bernard Williams, who argued that an agent facing conflicting obligations thinks that she should fulfill each of the obligations, but does not think that she should fulfill all [203].

(2) Next, consider the context of *normative systems*. When talking about norms belonging or not belonging to such a system, we use norm-propositions (cfr. Section 1.2.2). Thus, formulas of the form $\mathsf{O}A$ [$\mathsf{P}A$] are interpreted as "there exists a norm to the effect that A is mandatory [permitted]". In Section 6.2, we argue that normative systems often contain irresolvable conflicts between norms, and that these conflicts can be formalized as OO- or OP-conflicts (see also [1, 2]).

In this context, a formula $\neg\mathsf{O}p$ [$\neg\mathsf{P}p$] denotes the absence of a norm to the effect that p is mandatory [permitted].[9] Whereas a normative system may very well contain both a norm to the effect that p is mandatory as well as a norm to the effect that $\neg p$ is mandatory or permitted, it is less clear how such a system could both contain and not contain a norm to the effect that p is mandatory or permitted.[10] Thus it is reasonable to construct a logic of normative systems that takes into account the possibility of OO- and OP-conflicts, but not the possibility of contradictory norms.

Due to the possibility of OP-conflicts and the specific interpretation of the deontic operators in this context, a concrete CTDL for normative systems should invalidate the interdefinability schema (P) [192]. In Section 6.2 we further defend and illustrate the claims made here, and present a concrete CTDL that accommodates OO- and OP-conflicts and that invalidates (P).

(3) As a final illustration, consider the logic of commands. Since it is possible for a (confused) authority to assert that p is obligatory, and also that $\neg p$ is

[9] Formulas of the form $\mathsf{P}A$ are interpreted here as *strong* or *positive* permissions, in accordance with their interpretation in [1].

[10] Exceptions can be made, for instance, when one of two parties argues that system S *does* contain a norm to the effect that A is permitted, whereas the other argues that S *doesn't* contain such a norm. However, such a context is different from the one discussed here.

obligatory or permitted, OO- and OP-conflicts should be tolerated. Moreover, assuming the validity of (P), we should allow for the possibility of contradictory obligations and permissions in this setting, since a formula $\mathsf{O}A \wedge \mathsf{P}\neg A$ is equivalent to $\mathsf{O}A \wedge \neg\mathsf{O}A$ and $\neg\mathsf{P}\neg A \wedge \mathsf{P}\neg A$ in view of (P).[11] In Section 6.1, we present a CTDL that is *fully* conflict-tolerant in the sense that it tolerates OO-conflicts, OP-conflicts as well as contradictory obligations and permissions. This logic weakens **SDL** by turning its classical negation into a paraconsistent one, as discussed in Section 3.2.2.3.

One need not agree with all the details in illustrations (1)-(3) in order to be convinced by the main argument, namely that different normative contexts require different CTDLs. From the illustrations, it is also clear that the degree of conflict-tolerance of a given CTDL, i.e. the variety of types of normative conflicts that the CTDL should consistently allow for, is also context-dependent.

The remarks made here on logical pluralism partly answer the questions raised in point (i) of Section 1.6.2. Depending on the context, some types of normative conflicts need not be accommodated and may be ignored. When investigating the nature of moral dilemmas we may for instance leave OP-conflicts out of the picture. The types of inferences that are valid in the presence of normative conflicts are context-dependent as well. (P) for instance seems intuitive when dealing with commands, but contra-intuitive when dealing with norm-propositions. As promised, we will flesh out the details of some suggestions made here in chapters to come.

[11] Some authors have argued against the principle (P) given the possibility of normative conflicts or normative gaps (see the discussion in Section 6.2.2). However, their arguments presuppose a *descriptive* reading of the O- and P-operator, as opposed to the present prescriptive reading. We come back to this point in detail in Chapter 6.

Chapter 4

The standard format for adaptive logics

> A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend. The process of justification is the delicate one of making mutual adjustments between rules and accepted inferences
>
> Nelson Goodman [71, p. 64]

✎ I am indebted to Joke Meheus, Christian Straßer and Frederik Van De Putte for valuable comments on this chapter.

The standard format for adaptive logics (henceforth ALs) provides a unified characterization of ALs. All ALs defined later on in this thesis are defined within this format. An AL defined within the standard format is characterized as a triple consisting of a so-called lower limit logic, a set of abnormalities and an adaptive strategy. After introducing the standard format in Section 4.1, we discuss each of the elements in this triple characterization in turn. In Sections 4.2 and 4.3 we present some requirements on the lower limit logic and the set of abnormalities. In Sections 4.4 and 4.5 we discuss the two adaptive strategies currently defined within the standard format, the reliability strategy and the minimal abnormality strategy. A comparison between these strategies is provided in Section 4.6. Each element in the triple characterization of an AL is illustrated by means of a concrete example, the adaptive deontic logic $\mathbf{P}^{\mathbf{x}}_{\Diamond}$.

In Section 4.7 we state some meta-theoretical properties that come for free with ALs defined within the standard format. In Sections 4.8 and 4.9 we discuss two more features peculiar to ALs, their dynamic proofs and their so-called upper limit logic. We end this chapter with a brief mentioning of some adaptive strategies other than reliability and minimal abnormality (Section 4.10).

4.1 The standard format

ALs are tools for explicating and understanding defeasible human reasoning patterns. The first ALs arose from Diderik Batens' work on reasoning in the presence of inconsistent information [11, 12]. Later applications of ALs include inductive generalization [18], abductive reasoning [128, 59, 112], abstract argumentation [176], reasoning with vague premises [188], belief revision [183], and reasoning with prioritized beliefs [189] and prioritized norms [182]. Moreover, ALs have been used to characterize existing non-monotonic consequence relations [23, 25, 170, 189].

The standard format for ALs was introduced by Diderik Batens in [14, 16]. It provides a generic, unifying framework within which most existing ALs are defined. ALs characterized within this format automatically inherit a number of meta-theoretical properties that in earlier times had to be proven separately for each AL in question.

An AL in standard format consists of three elements:

1. A lower limit logic **LLL**,
2. A set of abnormalities Ω, and
3. An adaptive strategy (reliability or minimal abnormality).

Let $\mathbf{AL^r}$ denote the AL defined by $\langle \mathbf{LLL}, \Omega, \text{reliability}\rangle$, and $\mathbf{AL^m}$ the AL defined by $\langle \mathbf{LLL}, \Omega, \text{minimal abnormality}\rangle$. By $\mathbf{AL^x}$ we refer to either $\mathbf{AL^r}$ or $\mathbf{AL^m}$, i.e. $x \in \{r, m\}$. In Sections 4.2-4.5 we discuss each element in the definition of an AL in turn. We illustrate each element by means of a concrete AL, the deontic logic $\mathbf{P^x_\Diamond}$.

4.2 The lower limit logic

An adaptive logic $\mathbf{AL^x}$ is built 'on top' of a *lower limit logic* **LLL**. In order to be eligible as a lower limit logic (LLL) in the standard format for ALs, **LLL** must meet criteria (i)-(vi). Let $\Gamma, \Gamma' \subseteq \mathcal{W}^{\mathbf{LLL}}$ and $A \in \mathcal{W}^{\mathbf{LLL}}$:

(i) Reflexivity: $\Gamma \subseteq \mathrm{Cn}_{\mathbf{LLL}}(\Gamma)$.

(ii) Transitivity: if $\Gamma' \subseteq \mathrm{Cn}_{\mathbf{LLL}}(\Gamma)$ then $\mathrm{Cn}_{\mathbf{LLL}}(\Gamma \cup \Gamma') \subseteq \mathrm{Cn}_{\mathbf{LLL}}(\Gamma)$.

(iii) Monotonicity: $\mathrm{Cn}_{\mathbf{LLL}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{LLL}}(\Gamma \cup \Gamma')$.

(iv) Compactness: if $A \in \mathrm{Cn}_{\mathbf{LLL}}(\Gamma)$ then there is a finite $\Gamma' \subseteq \Gamma$ such that $A \in \mathrm{Cn}_{\mathbf{LLL}}(\Gamma')$.

(v) Supraclassicality: if $A \in \mathrm{Cn}_{\mathbf{CL}}(\Gamma)$ then $A \in \mathrm{Cn}_{\mathbf{LLL}}(\Gamma)$.

(vi) Soundness and completeness: $\Gamma \vdash_{\mathbf{LLL}} A$ iff $\Gamma \vDash_{\mathbf{LLL}} A$.

Any logic that meets these criteria is suitable as a LLL for an AL defined within the standard format.[1]

Its LLL is the monotonic base of an AL. Syntactically, an AL allows for the application of all LLL-valid inferences in an adaptive proof. As we explain below, ALs enhance the static proof theory of their LLL with a dynamic element, by means of which additional consequences are usually derivable. Semantically, ALs proceed by selecting a certain subset of the LLL-models of a given premise set. The exact set of **LLL**-models that is selected depends on the other two elements in the definition of $\mathbf{AL^x}$.

We promised to illustrate the workings of ALs by means of an example from deontic logic, the adaptive logic $\mathbf{P}^{\mathbf{x}}_{\Diamond}$. The LLL of $\mathbf{P}^{\mathbf{x}}_{\Diamond}$ is the logic $\mathbf{P}_{\Diamond}$. $\mathbf{P}_{\Diamond}$ strengthens Lou Goble's logic **P** from Section 3.2.2.1. It is defined for the language $\mathcal{W}^{\mathsf{O}}_{\Box}$ from Section 2.1.2. Where $\Diamond A =_{\mathrm{df}} \neg\Box\neg A$, $\mathbf{P}_{\Diamond}$ is obtained by adding to the axiomatization of **P** the schemas (CONS), (□K), (OIC) and (◇AND) as well as the rule (□NEC):

$$\Box(A \supset B) \supset (\Box A \supset \Box B) \tag{$\Box$K}$$

$$\text{If } \vdash A \text{ then } \vdash \Box A \tag{$\Box$NEC}$$

$$\mathsf{O}A \supset \Diamond A \tag{OIC}$$

$$(\mathsf{O}A \wedge \mathsf{O}B) \supset (\Diamond(A \wedge B) \supset \mathsf{O}(A \wedge B)) \tag{$\Diamond$AND}$$

In other words, $\mathbf{P}_{\Diamond}$ strengthens **P** with (CONS), with a **K**-operator for representing alethic modalities, and with the bridge principles (OIC) and (◇AND).[2]

Since $\mathbf{P}_{\Diamond}$ contains (CONS), it cannot accommodate OO-conflicts. OP-conflicts too cause explosion in $\mathbf{P}_{\Diamond}$. For suppose that $\mathsf{O}A$ and $\mathsf{P}\neg A$. Since (P) holds in **P**, $\mathsf{P}\neg A \equiv \neg\mathsf{O}\neg\neg A$. By (RE) and some simple propositional manipulations, $\neg\mathsf{O}\neg\neg A \equiv \neg\mathsf{O}A$. Thus, by modus ponens, we can derive $\neg\mathsf{O}A$ from $\mathsf{P}\neg A$, and we obtain a contradiction.

Remember from Section 3.2.1.2 that normative conflicts can also be formalized by making use of alethic modalities. Bernard Williams took conflicting moral obligations to be situations in which it is morally obliged to do two or more things ($\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n$), while at the same time it is (physically) impossible to fulfill the obligations ($\neg\Diamond(A_1 \wedge \ldots \wedge A_n)$). It is these types of conflicts which $\mathbf{P}_{\Diamond}$ is able to accommodate, and on which we will focus during this chapter while explaining how adaptive logics work. Note that (◇AND) blocks the aggregation of two obligations whenever it is not possible to fulfill both of them.

When more than two obligations need to be aggregated, we can use the following derived rule:

$$(\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n) \supset (\Diamond(A_1 \wedge \ldots \wedge A_n) \supset \mathsf{O}(A_1 \wedge \ldots \wedge A_n)) \tag{$\Diamond$AND'}$$

[1] For technical reasons, it is sometimes required that **LLL** is equipped with a distinct set of classical connectives. Since all classical connectives are present or definable in the logics presented in this thesis, this is a very straightforward operation. For reasons of didactics and convenience, we skip it here and refer the reader to [19, Sec. 4.3], [181, Sec. 2.7], or [172, Sec. 2.8] for more details.

[2] Alethic modalities are often characterized by modal logics like **T**, **S4** or **S5**. If the reader has a preference for a stronger modality, then the corresponding axiom schemas (T), (S4) and (S5) can be added to $\mathbf{P}_{\Diamond}$ at will. This is inessential for our present purposes, hence we stick to the weakest normal modal logic, **K**.

Fact 2. ($\Diamond$AND') is valid in $\mathbf{P}_\Diamond$.

Proof. Suppose that $\mathsf{O}A_1, \ldots, \mathsf{O}A_n$ and $\Diamond(A_1 \wedge \ldots \wedge A_n)$. By **K**-properties, (†) $\Diamond(\bigwedge \Theta)$ for all $\Theta \subseteq \{A_1, \ldots, A_n\}$. By $\mathsf{O}A_1, \mathsf{O}A_2$, (†) and ($\Diamond$AND), $\mathsf{O}(A_1 \wedge A_2)$. By $\mathsf{O}A_3$, (†) and ($\Diamond$AND), $\mathsf{O}(A_1 \wedge A_2 \wedge A_3)$, and so on until $\mathsf{O}(A_1 \wedge \ldots \wedge A_n)$. □

It is straightforward and left to the reader to show that $\mathbf{P}_\Diamond$ meets criteria (i)-(v). For (vi), we provide a semantical characterization of $\mathbf{P}_\Diamond$ and a proof outline of the following theorem in Appendix D:

Theorem 6. $\Gamma \vdash_{\mathbf{P}_\Diamond} A$ *iff* $\Gamma \vDash_{\mathbf{P}_\Diamond} A$

4.3 Abnormalities

Adaptive logics typically interpret a given premise set 'as normally as possible' with respect to some standard of normality. Intuitively, the set of abnormalities determines what it means to violate the standard of normality that an AL applies. For instance, if the aim of an AL is to interpret a set of premises 'as consistently as possible', then inconsistencies will typically give rise to an abnormality in the AL.

Formally, the *set of abnormalities* Ω of an adaptive logic $\mathbf{AL^x}$ is a set of **LLL**-wffs characterized by a (possibly restricted) logical form, or a union of such sets. To interpret a premise set 'as normally as possible', then, is to interpret this set in such a way that as few abnormalities as possible follow from it. Semantically, $\mathbf{AL^x}$ selects **LLL**-models of a given premise set that are 'as normal as possible' in terms of the abnormalities they verify. Syntactically, $\mathbf{AL^x}$ usually strengthens **LLL** by allowing for the application of a defeasible inference rule that considers abnormalities to be false 'whenever possible'. The phrases between inverted commas are disambiguated by the third element in the definition of $\mathbf{AL^x}$, the adaptive strategy.

Before we explain the workings of the different adaptive strategies, we return to our example, the logic $\mathbf{P}^{\mathbf{x}}_\Diamond$. Remember from Section 3.3.3 that we wish for implementations of our logics to be user-friendly in the sense that they should not leave any complicated reasoning processes to the user. However, in $\mathbf{P}_\Diamond$ the application of ($\Diamond$AND) is very demanding, since it requires a consistency-check on our premises: before we can aggregate two obligations, we need to ensure that they are not involved in a normative conflict. Ideally, this check should be done by the logic itself, and not by the user.

We will overcome this problem by letting the resulting adaptive extension $\mathbf{P}^{\mathbf{x}}_\Diamond$ of $\mathbf{P}_\Diamond$ do the required work in our place. $\mathbf{P}^{\mathbf{x}}_\Diamond$ allows us to aggregate obligations 'as much as possible' by allowing for its application *unless* applied to formulas which are (jointly) impossible. To this end, we could define the set Ω of $\mathbf{P}^{\mathbf{x}}_\Diamond$-abnormalities simply as the set of $\mathbf{P}_\Diamond$-wffs of the form $\neg \Diamond A$. It will prove more economical, however, to define Ω as follows:

$$\Omega = \left\{ \neg \Diamond \bigwedge \Delta \mid \Delta \subseteq \mathcal{W}^l, \text{ for all } A \in \mathcal{W}^a : \{A, \neg A\} \not\subseteq \Delta \right\}$$

The constraint that, for all $A \in \mathcal{W}^a$, $\{A, \neg A\} \nsubseteq \Delta$ is to ensure that Ω contains only $\mathbf{P}_{\Diamond}$-*contingent* formulas.[3] Thus, $\mathbf{P}_{\Diamond}$-valid formulas of the form $\neg \Diamond (A \wedge \neg A)$ do not give rise to abnormalities, nor do any of their consequences. For instance, neither the $\mathbf{P}_{\Diamond}$-wff $\neg \Diamond (p \wedge \neg p)$ nor its $\mathbf{P}_{\Diamond}$-consequence $\neg \Diamond (p \wedge \neg p \wedge q)$ belong to the set Ω.

Although Ω does not contain *all* $\mathbf{P}_{\Diamond}$-contingent wffs of the form $\neg \Diamond A$ (due to its restriction to conjunctions of literals within the scope of the $\Diamond$-operator), we can guarantee that each contingent wff of this logical form is $\mathbf{P}_{\Diamond}$-equivalent to an abnormality, or to a conjunction of abnormalities:

Fact 3. If $A \in \mathcal{W}_{\Box}^{\bigcirc}$ is of the logical form $\neg \Diamond B$ and B is **CL**-consistent, then A is $\mathbf{P}_{\Diamond}$-equivalent to a member of Ω, or to a conjunction of members of Ω.

Proof. Let $B_1 \vee \ldots \vee B_n$ be a disjunctive normal form of B such that, for all $i \in \{1, \ldots, n\}$, B_i is **CL**-consistent. Then, by **K**-properties, $\neg \Diamond B \dashv\vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (B_1 \vee \ldots \vee B_n) \dashv\vdash_{\mathbf{P}_{\Diamond}} (\neg \Diamond B_1 \wedge \ldots \wedge \neg \Diamond B_n)$. Since, for all $i \in \{1, \ldots, n\}$, B_i is a **CL**-consistent conjunction of members of $\mathcal{W}^l$, it follows that $\neg \Diamond B_i \in \Omega$. □

Fact 3 warrants our restriction of Ω to conjunctions of literals within the scope of a $\Diamond$-operator preceded by a negation. Although nothing prevents one from defining the set of abnormalities as the set of all formulas of the form $\neg \Diamond A$ (where $A \in \mathcal{W}$), the present definition is much more succinct.

4.4 The reliability strategy

Together with the set of abnormalities, the adaptive strategy stipulates what it means to interpret a premise set 'as normally as possible'. The two prominent adaptive strategies defined within the standard format are the reliability strategy and the minimal abnormality strategy. Reliability is slightly more 'cautious' than minimal abnormality. We discuss each strategy in turn, both from a proof theoretical and semantical point of view.

4.4.1 Proof theory

Before we get to the specifics of the reliability strategy, we need to introduce some general (strategy-independent) features of adaptive proofs. A line in an annotated adaptive proof consists of four elements: a line number i, a formula A, a justification (consisting of a series of line numbers and a derivation rule), and a condition Δ. The condition of a line is a (possibly empty) set of abnormalities. Intuitively, we interpret a line i at which formula A is derived on the condition Δ as "At line i of the proof, we have derived A on the assumption that all members of Δ are false".

The presence of a condition is part of what makes an adaptive proof *dynamic*. The dynamics of these proofs is controlled by attaching conditions to derived formulas and by introducing a marking definition. The rules determine which lines (consisting of the four aforementioned elements) may be added to a given

[3] A formula A is **L**-contingent iff neither A nor $\neg A$ is **L**-valid.

proof. The effect of the marking definition is that, at every stage[4] of the proof, certain lines may be marked whereas others are unmarked. Formulas occurring at lines that are marked at a certain stage s in the proof, are considered not derivable at that stage. The marking definition is different for each adaptive strategy.

Let us now introduce the rules of inference of an adaptive logic in standard format and the marking definition for the reliability strategy. The rules of inference reduce to three generic rules: a premise introduction rule PREM, an unconditional rule RU, and a conditional rule RC. Where $\Gamma \subseteq \mathcal{W}^{\mathbf{LLL}}$ is the set of premises, and where

$$A \quad \Delta$$

abbreviates that A occurs in the proof on the condition Δ, the inference rules PREM and RU are given by

PREM If $A \in \Gamma$: $\begin{array}{ll} \vdots & \vdots \\ \hline A & \emptyset \end{array}$

RU If $A_1, \ldots, A_n \vdash_{\mathbf{LLL}} B$: $\begin{array}{ll} A_1 & \Delta_1 \\ \vdots & \vdots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \end{array}$

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

Let a *Dab*-formula be a finite disjunction of members of Ω, and $Dab(\Theta) =_{\text{df}} \bigvee \Theta$, where Θ is a finite and non-empty set of abnormalities ($\Theta \subseteq \Omega$).[5] The conditional rule RC is defined as follows:

RC If $A_1, \ldots, A_n \vdash_{\mathbf{LLL}} B \vee Dab(\Theta)$: $\begin{array}{ll} A_1 & \Delta_1 \\ \vdots & \vdots \\ A_n & \Delta_n \\ \hline B & \Delta_1 \cup \ldots \cup \Delta_n \cup \Theta \end{array}$

In general, if $A_1, \ldots, A_n \vdash_{\mathbf{LLL}} B \vee Dab(\Theta)$ and $A_1, \ldots, A_n$ occur in a proof on the conditions $\Delta_1, \ldots, \Delta_n$ respectively, then, by the conditional rule RC, we can infer B on the condition $\Delta_1 \cup \ldots \cup \Delta_n \cup \Theta$. RC is the only rule that allows for the introduction of new conditions in an adaptive proof.

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

[5]If Θ is a singleton $\{A\}$ for some $A \in \Omega$, then $Dab(\Theta) = A$.

Since, for instance, $\vdash_{\mathbf{P}_\Diamond} \Diamond(p \wedge q) \vee \neg \Diamond (p \wedge q)$, and since $\neg \Diamond (p \wedge q) \in \Omega$, this means that in any $\mathbf{P}^{\mathbf{x}}_\Diamond$-proof we can derive $\Diamond(p \wedge q)$ via RC on the condition $\{\neg \Diamond (p \wedge q)\}$.

Suppose now that $\mathsf{O}p$ and $\mathsf{O}q$. Then, since $\mathsf{O}p, \mathsf{O}q \vdash_{\mathbf{P}_\Diamond} \Diamond(p \wedge q) \vee \neg \Diamond (p \wedge q)$, it follows by ($\Diamond$AND) and **CL** that $\mathsf{O}p, \mathsf{O}q \vdash_{\mathbf{P}_\Diamond} \mathsf{O}(p \wedge q) \vee \neg \Diamond (p \wedge q)$. Hence, in any $\mathbf{P}^{\mathbf{x}}_\Diamond$-proof in which we have derived both $\mathsf{O}p$ and $\mathsf{O}q$ on the conditions Δ and Θ respectively, we can apply RC to derive $\mathsf{O}(p \wedge q)$ on the condition $\Delta \cup \Theta \cup \{\neg \Diamond (p \wedge q)\}$.

Before we introduce the marking definition for the reliability strategy, we illustrate the ideas presented so far by means of a simple example. Let $\Gamma_1 = \{\mathsf{O}(p \wedge q), \mathsf{O}r, \mathsf{O}s, \neg \Diamond (p \wedge r) \vee \neg \Diamond (q \wedge s)\}$. We start a $\mathbf{P}^{\mathbf{r}}_\Diamond$-proof from Γ_1 by entering the premises:

1	$\mathsf{O}(p \wedge q)$	PREM	$\emptyset$
2	$\mathsf{O}r$	PREM	$\emptyset$
3	$\mathsf{O}s$	PREM	$\emptyset$
4	$\neg \Diamond (p \wedge r) \vee \neg \Diamond (q \wedge s)$	PREM	$\emptyset$

We can continue the proof as follows:

5	$\mathsf{O}p$	1; RU	$\emptyset$

Since $\mathsf{O}(p \wedge q) \vdash_{\mathbf{P}_\Diamond} \mathsf{O}p$, we can use RU to derive $\mathsf{O}p$ from the formula at line 1. Since the condition of this line is empty, the condition of line 5 is empty too. In an analogous fashion, we can apply RU to derive $\mathsf{O}q$:

6	$\mathsf{O}q$	1; RU	$\emptyset$

Suppose now that we want to aggregate $\mathsf{O}p$ and $\mathsf{O}s$ via the rule ($\Diamond$AND). For this, we need to know that p and s are jointly possible ($\Diamond(p \wedge s)$). Γ_1 does not provide us with that information, so we cannot derive the formula $\mathsf{O}(p \wedge s)$ by means of the LLL $\mathbf{P}_\Diamond$. However, we do know (by **CL**) that either $p \wedge s$ is possible, or that it is not:

7	$\Diamond(p \wedge s) \vee \neg \Diamond (p \wedge s)$	RU	$\emptyset$

Since $\neg \Diamond (p \wedge s) \in \Omega$, we can now move this formula to the condition set by means of RC:

8	$\Diamond(p \wedge s)$	7; RC	$\{\neg \Diamond (p \wedge s)\}$

Given this information, we can apply ($\Diamond$AND) to lines 3, 5 and 8, and derive $\mathsf{O}(p \wedge s)$ as desired:

9	$\mathsf{O}(p \wedge s)$	3,5,8; RU	$\{\neg \Diamond (p \wedge s)\}$

Note that the condition of line 8 gets carried over to line 9 in view of the definition of RU. Suppose now that we continue the proof as follows:

10	$\mathsf{O}(p \wedge r)$	2,5; RC	$\{\neg \Diamond (p \wedge r)\}\checkmark^4$

At line 10, we have derived $\mathsf{O}(p \wedge r)$ in a fashion analogous to the derivation of $\mathsf{O}(p \wedge s)$ at line 9 (we skipped the intermediate step of deriving $\Diamond(p \wedge r)$). In doing so we have assumed that $\neg \Diamond (p \wedge r)$ is false. However, we know from line 4 that either $\neg \Diamond (p \wedge r)$ holds, or $\neg \Diamond (q \wedge s)$ does.

In view of this information, our assumption at line 10 that $\neg\Diamond(p \wedge r)$ is false was too hasty. In the logic $\mathbf{P}^{\mathbf{r}}_{\Diamond}$, the presence of the condition of line 10 in the disjunction of abnormalities at line 4 causes the withdrawal of line 10 from the proof. This is taken care of by the marking definition, and is indicated by a checkmark sign ("✓") indexed by the number of the line in view of which line 10 is marked.

The marking definition for the reliability strategy proceeds in terms of the *minimal Dab-formulas* and the *unreliable formulas* derived at a stage of the proof:

Definition 1. $Dab(\Delta)$ is a *minimal Dab-formula* at stage s iff, at stage s, $Dab(\Delta)$ is derived on the condition $\varnothing$, and no $Dab(\Delta')$ with $\Delta' \subset \Delta$ is derived on the condition $\varnothing$.

Definition 2. Where $Dab(\Delta_1), Dab(\Delta_2), \ldots$ are the minimal Dab-formulas derived at stage s, $U_s(\Gamma) = \Delta_1 \cup \Delta_2 \cup \ldots$ is the set of formulas that are *unreliable* at stage s.

Marking for reliability is defined as follows:

Definition 3. Where Δ is the condition of line i, line i is marked at stage s iff $\Delta \cap U_s(\Gamma) \neq \varnothing$.

At stage 10 of our proof, a minimal Dab-formula was derived at line 4. By Definition 2, $U_{10}(\Gamma_1) = \{\neg\Diamond(p \wedge r), \neg\Diamond(q \wedge s)\}$. Since the element in the condition set of line 10 is a member of $U_{10}(\Gamma_1)$, the line is marked due to Definition 3. Note that lines 8 and 9 remain unmarked at this stage of the proof, since their condition does not overlap with the set of unreliable formulas at stage 10.

The marking definition for the reliability strategy is further illustrated in the following extension[6] of the proof (we repeat the proof from line 8 on):

8	$\Diamond(p \wedge s)$	RC	$\{\neg\Diamond(p \wedge s)\}$
9	$\mathsf{O}(p \wedge s)$	3,5,8; RU	$\{\neg\Diamond(p \wedge s)\}$
10	$\mathsf{O}(p \wedge r)$	2,5; RC	$\{\neg\Diamond(p \wedge r)\}\checkmark^{4}$
11	$\mathsf{O}(q \wedge s)$	3,6; RC	$\{\neg\Diamond(q \wedge s)\}\checkmark^{4}$
12	$\mathsf{O}(q \wedge r)$	2,6; RC	$\{\neg\Diamond(q \wedge r)\}$
13	$\mathsf{O}(r \wedge s)$	2,3; RC	$\{\neg\Diamond(r \wedge s)\}$

At stage 13, no new minimal Dab-formulas have been derived. Hence $U_{13}(\Gamma_1) = U_{10}(\Gamma_1)$.

We can continue the proof as follows:

14	$\mathsf{O}(p \wedge q \wedge r)$	1,2; RC	$\{\neg\Diamond(p \wedge q \wedge r)\}\checkmark^{20}$
15	$\mathsf{O}(p \wedge q \wedge s)$	1,3; RC	$\{\neg\Diamond(p \wedge q \wedge s)\}\checkmark^{21}$
16	$\mathsf{O}(p \wedge r \wedge s)$	2,3,5; RC	$\{\neg\Diamond(p \wedge r \wedge s)\}\checkmark^{19}$
17	$\mathsf{O}(q \wedge r \wedge s)$	2,3,6; RC	$\{\neg\Diamond(q \wedge r \wedge s)\}\checkmark^{22}$
18	$\mathsf{O}(p \wedge q \wedge r \wedge s)$	1,2,3; RC	$\{\neg\Diamond(p \wedge q \wedge r \wedge s)\}\checkmark^{25}$
19	$\neg\Diamond(p \wedge r \wedge s) \vee \neg\Diamond(q \wedge s)$	4; RU	$\varnothing$
20	$\neg\Diamond(p \wedge q \wedge r) \vee \neg\Diamond(q \wedge s)$	4; RU	$\varnothing$
21	$\neg\Diamond(p \wedge r) \vee \neg\Diamond(p \wedge q \wedge s)$	4; RU	$\varnothing$

[6] A stage s' of an adaptive proof is an *extension* of a stage s iff every line in s occurs in s'.

22	$\neg\Diamond(p\wedge r)\vee\neg\Diamond(q\wedge r\wedge s)$	4; RU	$\varnothing$
23	$\neg\Diamond(p\wedge q\wedge r)\vee\neg\Diamond(q\wedge r\wedge s)$	4; RU	$\varnothing$
24	$\neg\Diamond(p\wedge q\wedge r)\vee\neg\Diamond(p\wedge q\wedge s)$	4; RU	$\varnothing$
25	$\neg\Diamond(p\wedge q\wedge r\wedge s)$	4; RU	$\varnothing$

The formulas derived at lines 19-25 are **K**-consequences of the formula derived at line 4. At stage 25 of the proof, $U_{25}(\Gamma_1) = \{\neg\Diamond(p\wedge r), \neg\Diamond(q\wedge s), \neg\Diamond(p\wedge r\wedge s), \neg\Diamond(p\wedge q\wedge r), \neg\Diamond(p\wedge q\wedge s), \neg\Diamond(q\wedge r\wedge s), \neg\Diamond(p\wedge q\wedge r\wedge s)\}$. Hence, lines 10, 11 and 14-18 are marked in view of Definition 3.

Marking is a *dynamic* matter: marks may come and go in adaptive proofs. Suppose, for instance, that we add the premise $\neg\Diamond(p\wedge r)$ to Γ_1. Call the resulting premise set Γ_1'. Since lines 1-25 above form a valid $\mathbf{P}^{\mathbf{r}}_{\Diamond}$-proof from Γ_1', we can simply copy these lines to a proof for Γ_1', add the new premise at a new line and continue as follows (we repeat the proof from line 8 on):

8	$\Diamond(p\wedge s)$	RC	$\{\neg\Diamond(p\wedge s)\}$
9	$\mathsf{O}(p\wedge s)$	3,5,8; RU	$\{\neg\Diamond(p\wedge s)\}$
10	$\mathsf{O}(p\wedge r)$	2,5; RC	$\{\neg\Diamond(p\wedge r)\}\checkmark^{4}$
11	$\mathsf{O}(q\wedge s)$	3,6; RC	$\{\neg\Diamond(q\wedge s)\}$
12	$\mathsf{O}(q\wedge r)$	2,6; RC	$\{\neg\Diamond(q\wedge r)\}$
13	$\mathsf{O}(r\wedge s)$	2,3; RC	$\{\neg\Diamond(r\wedge s)\}$
14	$\mathsf{O}(p\wedge q\wedge r)$	1,2; RC	$\{\neg\Diamond(p\wedge q\wedge r)\}\checkmark^{28}$
15	$\mathsf{O}(p\wedge q\wedge s)$	1,3; RC	$\{\neg\Diamond(p\wedge q\wedge s)\}$
16	$\mathsf{O}(p\wedge r\wedge s)$	2,3,5; RC	$\{\neg\Diamond(p\wedge r\wedge s)\}\checkmark^{27}$
17	$\mathsf{O}(q\wedge r\wedge s)$	2,3,6; RC	$\{\neg\Diamond(q\wedge r\wedge s)\}$
18	$\mathsf{O}(p\wedge q\wedge r\wedge s)$	1,2,3; RC	$\{\neg\Diamond(p\wedge q\wedge r\wedge s)\}\checkmark^{25}$
19	$\neg\Diamond(p\wedge r\wedge s)\vee\neg\Diamond(q\wedge s)$	4; RU	$\varnothing$
20	$\neg\Diamond(p\wedge q\wedge r)\vee\neg\Diamond(q\wedge s)$	4; RU	$\varnothing$
21	$\neg\Diamond(p\wedge r)\vee\neg\Diamond(p\wedge q\wedge s)$	4; RU	$\varnothing$
22	$\neg\Diamond(p\wedge r)\vee\neg\Diamond(q\wedge r\wedge s)$	4; RU	$\varnothing$
23	$\neg\Diamond(p\wedge q\wedge r)\vee\neg\Diamond(q\wedge r\wedge s)$	4; RU	$\varnothing$
24	$\neg\Diamond(p\wedge q\wedge r)\vee\neg\Diamond(p\wedge q\wedge s)$	4; RU	$\varnothing$
25	$\neg\Diamond(p\wedge q\wedge r\wedge s)$	4; RU	$\varnothing$
26	$\neg\Diamond(p\wedge r)$	PREM	$\varnothing$
27	$\neg\Diamond(p\wedge r\wedge s)$	26; RU	$\varnothing$
28	$\neg\Diamond(p\wedge q\wedge r)$	26; RU	$\varnothing$

At stage 28 of the proof from Γ_1', the *Dab*-formulas at lines 4 and lines 19-24 are no longer minimal in view of Definition 1. At this stage, $U_{28}(\Gamma_1') = \{\neg\Diamond(p\wedge r), \neg\Diamond(p\wedge q\wedge r), \neg\Diamond(p\wedge r\wedge s), \neg\Diamond(p\wedge q\wedge r\wedge s)\}$. Accordingly, lines 11, 15 and 17 are unmarked at stage 28.

Due to the stage-dependency of the marking criterion in adaptive proofs, we can define a dynamic notion of derivation as follows:

Definition 4. A formula A has been derived at stage s of an adaptive proof iff, at that stage, A is the second element of some unmarked line i.

Since we want to define a syntactic consequence relation for ALs, we also need a static, stage-independent notion of derivability:

Definition 5. A is *finally derived* from Γ at line i of a proof at a finite stage s iff (i) A is the second element of line i, (ii) line i is not marked at stage s, and (iii) every extension of the proof in which line i is marked may be further extended in such a way that line i is unmarked.[7]

We can now define a syntactic consequence relation for an AL that makes use of the reliability strategy:

Definition 6. $\Gamma \vdash_{\mathbf{AL^r}} A$ (A is finally $\mathbf{AL^r}$-derivable from Γ) iff A is finally derived at a line of an $\mathbf{AL^r}$-proof from Γ.

Returning to our example, we can define a syntactic consequence relation for the logic $\mathbf{P}^{\mathbf{r}}_{\Diamond}$ by replacing "$\mathbf{AL^r}$" by "$\mathbf{P}^{\mathbf{r}}_{\Diamond}$" in Definition 6. Applying this definition to the example proof above, it can be shown that:

$$
\begin{array}{ll}
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r) & \Gamma_1' \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r) \\
\Gamma_1 \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge s) & \Gamma_1' \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge s) \\
\Gamma_1 \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r) & \Gamma_1' \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r) \\
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge s) & \Gamma_1' \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge s) \\
\Gamma_1 \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(r \wedge s) & \Gamma_1' \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(r \wedge s) \\
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r) & \Gamma_1' \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r) \\
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge s) & \Gamma_1' \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge s) \\
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r \wedge s) & \Gamma_1' \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r \wedge s) \\
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r \wedge s) & \Gamma_1' \vdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r \wedge s) \\
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r \wedge s) & \Gamma_1' \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r \wedge s)
\end{array}
$$

In the proofs from Γ_1 and Γ_1' we could have derived even more minimal *Dab*-formulas. Let us illustrate this point for the premise set Γ_1. The following inferences are $\mathbf{P}_{\Diamond}$-valid, since $\neg \Diamond A \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (A \wedge B_1 \wedge \ldots \wedge B_n)$ for all $A, B_1, \ldots, B_n \in \mathcal{W}^l$:

Table 4.1: More disjunctions of abnormalities for Γ_1.

$\neg \Diamond (p \wedge r) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge q \wedge r)$	$\neg \Diamond (q \wedge s) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge q \wedge s)$
$\neg \Diamond (p \wedge r) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge r \wedge s)$	$\neg \Diamond (q \wedge s) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (q \wedge r \wedge s)$
$\neg \Diamond (p \wedge r) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge \neg q \wedge r)$	$\neg \Diamond (q \wedge s) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (\neg p \wedge q \wedge s)$
$\neg \Diamond (p \wedge r) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge r \wedge \neg s)$	$\neg \Diamond (q \wedge s) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (q \wedge \neg r \wedge s)$
$\neg \Diamond (p \wedge r) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge \neg q \wedge r \wedge s)$	$\neg \Diamond (q \wedge s) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (\neg p \wedge q \wedge r \wedge s)$
$\neg \Diamond (p \wedge r) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge \neg q \wedge r \wedge \neg s)$	$\neg \Diamond (q \wedge s) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (\neg p \wedge q \wedge \neg r \wedge s)$
$\neg \Diamond (p \wedge r) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge q \wedge r \wedge \neg s)$	$\neg \Diamond (q \wedge s) \vdash_{\mathbf{P}_{\Diamond}} \neg \Diamond (p \wedge q \wedge \neg r \wedge s)$

Each of the formulas occurring in the left column of Table 4.1 can be placed in disjunction with a formula occurring in the right column in order to obtain a minimal *Dab*-formula derivable from Γ_1. For instance, all of the formulas $\neg \Diamond (p \wedge$

[7] Definition 5 has a game-theoretic flavor to it. In [17], this definition is interpreted as a two-player game in which the proponent has a winning strategy in case she has a reply to every counterargument by her opponent.

$r) \vee \neg\Diamond(\neg p \wedge q \wedge s)$, $\neg\Diamond(p \wedge \neg q \wedge r \wedge s) \vee \neg\Diamond(q \wedge s)$, and $\neg\Diamond(p \wedge \neg q \wedge r) \vee \neg\Diamond(\neg p \wedge q \wedge \neg r \wedge s)$ are $\mathbf{P}_\Diamond$-derivable from Γ_1, and each of them is minimal: none of their disjuncts is $\mathbf{P}_\Diamond$-derivable from Γ_1.

Although each and every one of these disjunctions could be added to the proof from Γ_1, none of them would cause the unmarking of any of the formulas derived at stage 25 of this proof. The reason is that all of the disjunctions currently derived in the proof are already minimal in view of Γ_1. Hence, whatever happens at later stages in the proof, the members of these disjunctions will remain in the set of unreliable formulas at these later stages. Moreover, adding these disjunctions to the proof from Γ_1 would not cause the marking of any of the lines that are unmarked at stage 25 of the proof, since none of their disjuncts is a member of the condition of any of the currently unmarked lines.

We return in more detail to the dynamics of adaptive proofs in Section 4.8, but first we define and illustrate the semantics for $\mathbf{AL^r}$ as well as the proof theory and semantics for $\mathbf{AL^m}$.

4.4.2 Semantics

ALs employ a preferential semantics in the vein of Shoham [160] (see also [161, 162]). The idea is that an AL selects a 'preferred' subset of its LLL-models. In ALs, this preferred subset is the set of LLL-models that verify "as few abnormalities as possible" in view of the adaptive strategy. For the reliability strategy, this preferred subset is often a proper superset of the one selected by the minimal abnormality strategy.

We need to introduce some terminology first. Let a *Dab*-formula $Dab(\Delta)$ be a *Dab-consequence* of Γ if it is **LLL**-derivable from Γ; it is a *minimal Dab*-consequence of Γ if there is no $\Delta' \subset \Delta$ such that $Dab(\Delta')$ is a *Dab*-consequence of Γ. The set of formulas that are unreliable with respect to Γ, denoted by $U(\Gamma)$, is defined by:

Definition 7. Where $Dab(\Delta_1), Dab(\Delta_2), \ldots$ are the minimal *Dab*-consequences of Γ, $U(\Gamma) = \Delta_1 \cup \Delta_2 \cup \ldots$ is the set of formulas that are *unreliable* with respect to Γ.

Let us illustrate the workings of the $\mathbf{AL^r}$-semantics by means of the logic $\mathbf{P}^{\mathbf{r}}_\Diamond$ and the premise set Γ_1 from Section 4.4.1. Recall that $\Gamma_1 = \{\mathsf{O}(p \wedge q), \mathsf{O}r, \mathsf{O}s, \neg\Diamond(p \wedge r) \vee \neg\Diamond(q \wedge s)\}$.

The minimal *Dab*-consequences derivable from Γ_1 include the premise $\neg\Diamond(p \wedge r) \vee \neg\Diamond(q \wedge s)$ as well as its $\mathbf{P}_\Diamond$-consequences $\neg\Diamond(p \wedge r) \vee \neg\Diamond(\neg p \wedge q \wedge s)$, $\neg\Diamond(p \wedge \neg q \wedge r \wedge s) \vee \neg\Diamond(q \wedge s)$, $\neg\Diamond(p \wedge \neg q \wedge r) \vee \neg\Diamond(\neg p \wedge q \wedge \neg r \wedge s)$, etc. (cfr. supra). Thus, the set of minimal *Dab*-consequences of Γ_1 is the infinite set containing all $\mathbf{P}_\Diamond$-contingent disjunctions of the form $\neg\Diamond\bigwedge\Delta \vee \neg\Diamond\bigwedge\Delta'$, where $\Delta, \Delta' \subset \mathcal{W}^l$ and $p, r \in \Delta$ and $q, s \in \Delta'$. The set $U(\Gamma_1)$ of unreliable formulas of Γ_1 contains each disjunct occurring in a minimal *Dab*-consequence of Γ_1.

We define the *abnormal part* $Ab(M)$ of a model M as the set of abnormalities verified by M:

Definition 8. $Ab(M) = \{A \in \Omega \mid M \Vdash A\}$

We can now split the set of $\mathbf{P}_\Diamond$-models of Γ_1 into a partition of three clauses according to the abnormal parts of the models. Due to the minimal *Dab*-consequence $\neg\Diamond(p \wedge r) \vee \neg\Diamond(q \wedge s)$, all $\mathbf{P}_\Diamond$-models of Γ_1 verify $\neg\Diamond(p \wedge r)$ or $\neg\Diamond(q \wedge s)$ or both. $\mathbf{P}_\Diamond$-models of Γ_1 that verify $\neg\Diamond(p \wedge r)$, also verify $\neg\Diamond(p \wedge r \wedge A_1 \ldots \wedge A_n)$ for all $A_1, \ldots, A_n \in \mathcal{W}^l$. Analogously, $\mathbf{P}_\Diamond$-models of Γ_1 that verify $\neg\Diamond(q \wedge s)$ also verify $\neg\Diamond(q \wedge s \wedge B_1 \wedge \ldots \wedge B_m)$ for all $B_1, \ldots, B_m \in \mathcal{W}^l$. Let M_1 be a $\mathbf{P}_\Diamond$-model of Γ_1 that verifies $\neg\Diamond(p \wedge r)$ but not $\neg\Diamond(q \wedge s)$, M_2 a $\mathbf{P}_\Diamond$-model of Γ_1 that verifies $\neg\Diamond(q \wedge s)$ but not $\neg\Diamond(p \wedge r)$, and M_3 a $\mathbf{P}_\Diamond$-model of Γ_1 that verifies both $\neg\Diamond(p \wedge r)$ and $\neg\Diamond(q \wedge s)$. Suppose further that M_1 verifies no other abnormalities than $\neg\Diamond(p \wedge r)$ and its *Dab*-consequences, that M_2 verifies no other abnormalities than $\neg\Diamond(q \wedge s)$ and its *Dab*-consequences, and that M_3 verifies no other abnormalities than $\neg\Diamond(p \wedge r), \neg\Diamond(q \wedge s)$ and the *Dab*-consequences of $\neg\Diamond(p \wedge r) \wedge \neg\Diamond(q \wedge s)$. Then the abnormal parts of these models look as displayed Table 4.2.

$M_1 \Vdash \ldots$	$M_2 \Vdash \ldots$	$M_3 \Vdash \ldots$
$\neg\Diamond(p \wedge r)$	$\neg\Diamond(q \wedge s)$	$\neg\Diamond(p \wedge r), \neg\Diamond(q \wedge s)$
$\neg\Diamond(p \wedge q \wedge r)$	$\neg\Diamond(p \wedge q \wedge s)$	$\neg\Diamond(p \wedge q \wedge r), \neg\Diamond(p \wedge q \wedge s)$
$\neg\Diamond(p \wedge r \wedge s)$	$\neg\Diamond(q \wedge r \wedge s)$	$\neg\Diamond(p \wedge r \wedge s), \neg\Diamond(q \wedge r \wedge s)$
$\neg\Diamond(p \wedge \neg q \wedge r)$	$\neg\Diamond(\neg p \wedge q \wedge s)$	$\neg\Diamond(p \wedge \neg q \wedge r), \neg\Diamond(\neg p \wedge q \wedge s)$
$\neg\Diamond(p \wedge r \wedge \neg s)$	$\neg\Diamond(q \wedge \neg r \wedge s)$	$\neg\Diamond(p \wedge r \wedge \neg s), \neg\Diamond(q \wedge \neg r \wedge s)$
$\neg\Diamond(p \wedge q \wedge r \wedge s)$	$\neg\Diamond(p \wedge q \wedge r \wedge s)$	$\neg\Diamond(p \wedge q \wedge r \wedge s)$
$\neg\Diamond(p \wedge \neg q \wedge r \wedge s)$	$\neg\Diamond(\neg p \wedge q \wedge r \wedge s)$	$\neg\Diamond(p \wedge \neg q \wedge r \wedge s), \neg\Diamond(\neg p \wedge q \wedge r \wedge s)$
$\neg\Diamond(p \wedge \neg q \wedge r \wedge \neg s)$	$\neg\Diamond(\neg p \wedge q \wedge \neg r \wedge s)$	$\neg\Diamond(p \wedge \neg q \wedge r \wedge \neg s), \neg\Diamond(\neg p \wedge q \wedge \neg r \wedge s)$
$\neg\Diamond(p \wedge q \wedge r \wedge \neg s)$	$\neg\Diamond(p \wedge q \wedge \neg r \wedge s)$	$\neg\Diamond(p \wedge q \wedge r \wedge \neg s), \neg\Diamond(p \wedge q \wedge \neg r \wedge s)$
$\vdots$	$\vdots$	$\vdots$

Table 4.2: Abnormal parts of M_1, M_2, and M_3.

A *reliable* model of a given premise set is defined as follows:

Definition 9. A **LLL**-model M of Γ is *reliable* iff $Ab(M) \subseteq U(\Gamma)$.

Note that, for all $\mathbf{P}^{\mathbf{r}}_\Diamond$-abnormalities A displayed in the table above, $A \in U(\Gamma_1)$. Thus M_1-M_3 are reliable $\mathbf{P}_\Diamond$-models of Γ_1.

Suppose that some $\mathbf{P}_\Diamond$-model M_4 of Γ_1 verifies the abnormality $\neg\Diamond(p \wedge s)$. Then, since $\neg\Diamond(p \wedge s) \notin U(\Gamma_1)$, M_4 is not a reliable $\mathbf{P}_\Diamond$-model of Γ_1. Hence, for all reliable $\mathbf{P}_\Diamond$-models M of Γ_1, $M \nVdash \neg\Diamond(p \wedge s)$. Consequently, $M \Vdash \Diamond(p \wedge s)$. Since we also know that $M \Vdash \mathsf{O}p, \mathsf{O}s$, it follows that $M \Vdash \mathsf{O}(p \wedge s)$.

A semantic consequence relation of $\mathbf{AL}^{\mathbf{r}}$ is defined as follows:

Definition 10. $\Gamma \vDash_{\mathbf{AL}^{\mathbf{r}}} A$ iff A is verified by all reliable models of Γ.

Since $M \Vdash \mathsf{O}(p \wedge s)$ for all reliable models M of Γ_1, it follows that $\Gamma \vDash_{\mathbf{P}^{\mathbf{r}}_\Diamond} \mathsf{O}(p \wedge s)$. It is safely left to the reader to check that all of the following hold:

$\Gamma_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r)$	$\Gamma'_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r)$
$\Gamma_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge s)$	$\Gamma'_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge s)$
$\Gamma_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r)$	$\Gamma'_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r)$
$\Gamma_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge s)$	$\Gamma'_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge s)$
$\Gamma_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(r \wedge s)$	$\Gamma'_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(r \wedge s)$
$\Gamma_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r)$	$\Gamma'_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r)$
$\Gamma_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge s)$	$\Gamma'_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge s)$
$\Gamma_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r \wedge s)$	$\Gamma'_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge r \wedge s)$
$\Gamma_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r \wedge s)$	$\Gamma'_1 \vDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(q \wedge r \wedge s)$
$\Gamma_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r \wedge s)$	$\Gamma'_1 \nvDash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r \wedge s)$

4.5 The minimal abnormality strategy

The minimal abnormality strategy is a tad less 'cautious' than reliability. In this section, we define and illustrate the proof theory and semantics for the minimal abnormality strategy. In Section 4.6 we compare both strategies.

4.5.1 Proof theory

The proof theory for the minimal abnormality strategy differs from that for reliability only with respect to the marking definition. The marking definition for minimal abnormality uses the notion of minimal choice sets. A *choice set* of $\Sigma = \{\Delta_1, \Delta_2, \ldots\}$ is a set that contains one element out of each member of Σ. A *minimal choice set* of Σ is a choice set of Σ of which no proper subset is a choice set of Σ. Where $Dab(\Delta_1), Dab(\Delta_2), \ldots$ are the minimal *Dab*-formulas that are derived at stage s of a proof, $\Phi_s(\Gamma)$ is the set of minimal choice sets of $\{\Delta_1, \Delta_2, \ldots\}$.

Marking for minimal abnormality is defined as follows:

Definition 11. Where A is derived at line i of a proof from Γ on a condition Δ, line i is marked at stage s iff
(i) there is no $\Delta' \in \Phi_s(\Gamma)$ such that $\Delta' \cap \Delta = \varnothing$, or
(ii) for some $\Delta' \in \Phi_s(\Gamma)$, there is no line at which A is derived on a condition Θ for which $\Delta' \cap \Theta = \varnothing$.

Alternatively, this definition can be understood in the following 'dual' way: where A is derived on the condition Δ on line i, line i is *unmarked* at stage s iff (i) there is a $\Delta' \in \Phi_s(\Gamma)$ for which $\Delta' \cup \Delta = \varnothing$ and (ii) for every $\Delta' \in \Phi_s(\Gamma)$ there is a line at which A is derived on a condition Θ for which $\Delta' \cap \Theta = \varnothing$.

We illustrate the marking mechanism for minimal abnormality by means of our familiar example. Consider the following $\mathbf{P}^{\mathbf{m}}_{\Diamond}$-proof from Γ_1:

1	$\mathsf{O}(p \wedge q)$	PREM	$\varnothing$
2	$\mathsf{O}r$	PREM	$\varnothing$
3	$\mathsf{O}s$	PREM	$\varnothing$
4	$\neg\Diamond(p \wedge r) \vee \neg\Diamond(q \wedge s)$	PREM	$\varnothing$
5	$\mathsf{O}p$	1; RU	$\varnothing$

6	$\mathsf{O}q$	1; RU	$\varnothing$
7	$\mathsf{O}(p \wedge s)$	3,5; RC	$\{\neg \Diamond (p \wedge s)\}$
8	$\mathsf{O}(p \wedge r)$	2,5; RC	$\{\neg \Diamond (p \wedge r)\}\checkmark^{4}$
9	$\mathsf{O}(q \wedge s)$	3,6; RC	$\{\neg \Diamond (q \wedge s)\}\checkmark^{4}$
10	$\mathsf{O}(q \wedge r)$	2,6; RC	$\{\neg \Diamond (q \wedge r)\}$
11	$\mathsf{O}(r \wedge s)$	2,3; RC	$\{\neg \Diamond (r \wedge s)\}$
12	$\mathsf{O}(p \wedge q \wedge r)$	2,5,6; RC	$\{\neg \Diamond (p \wedge q \wedge r)\}\checkmark^{17}$
13	$\mathsf{O}(p \wedge q \wedge s)$	3,5,6; RC	$\{\neg \Diamond (p \wedge q \wedge s)\}\checkmark^{17}$
14	$\mathsf{O}(p \wedge q \wedge (r \vee s))$	2,5,6; RC	$\{\neg \Diamond (p \wedge q \wedge r)\}$
15	$\mathsf{O}(p \wedge q \wedge (r \vee s))$	3,5,6; RC	$\{\neg \Diamond (p \wedge q \wedge s)\}$
16	$\mathsf{O}(p \wedge q \wedge r \wedge s)$	2,3,5,6; RC	$\{\neg \Diamond (p \wedge q \wedge r \wedge s)\}\checkmark^{18}$
17	$\neg\Diamond(p\wedge q\wedge r)\vee\neg\Diamond(p\wedge q\wedge s)$	4; RU	$\varnothing$
18	$\neg \Diamond (p \wedge q \wedge r \wedge s)$	4; RU	$\varnothing$

At stage 18, three minimal *Dab*-formulas were derived in the proof (at lines 4, 17 and 18). These formulas give rise to the following 4 minimal choice sets:

$$\begin{aligned}
\varphi_1 &= \{\neg \Diamond (p \wedge r), \neg \Diamond (p \wedge q \wedge r), \neg \Diamond (p \wedge q \wedge r \wedge s)\}\\
\varphi_2 &= \{\neg \Diamond (p \wedge r), \neg \Diamond (p \wedge q \wedge s), \neg \Diamond (p \wedge q \wedge r \wedge s)\}\\
\varphi_3 &= \{\neg \Diamond (q \wedge s), \neg \Diamond (p \wedge q \wedge r), \neg \Diamond (p \wedge q \wedge r \wedge s)\}\\
\varphi_4 &= \{\neg \Diamond (q \wedge s), \neg \Diamond (p \wedge q \wedge s), \neg \Diamond (p \wedge q \wedge r \wedge s)\}
\end{aligned}$$

Thus, $\Phi_{18}(\Gamma_1) = \{\varphi_1, \varphi_2, \varphi_3, \varphi_4\}$. Lines 7, 10 and 11 remain unmarked because their condition does not overlap with any of $\varphi_1 - \varphi_4$. Line 16 is marked because of condition (i) in Definition 11: its condition overlaps with all minimal choice sets in $\Phi_{18}(\Gamma_1)$. Lines 8, 9, 12, and 13 are marked because of condition (ii) in Definition 11: the formulas derived at these lines are such that if their condition intersects with a minimal choice set of $\Phi_{13}(\Gamma_1)$, there is no line in the proof at which the formula was derived on a condition that does not intersect with this minimal choice set.

The situation is different for the formula $\mathsf{O}(p \wedge q \wedge (r \vee s))$ derived at lines 14 and 15. Take line 14. Although the condition of this line intersects with the minimal choice sets φ_1 and φ_3, we have also derived the formula $\mathsf{O}(p\wedge q\wedge(r\vee s))$ on a condition that does not intersect with any of these sets, namely the condition $\{\neg \Diamond (p \wedge q \wedge s)\}$ of line 15. Analogously, the condition of line 15 intersects with the minimal choice sets φ_2 and φ_4, yet this line remains unmarked because we have derived $\mathsf{O}(p \wedge q \wedge (r \vee s))$ on the condition $\{\neg \Diamond (p \wedge q \wedge r)\}$ at line 14, and this condition does not intersect with any of these minimal choice sets.

Note that if the above proof were a $\mathbf{P}^{\mathbf{r}}_{\Diamond}$-proof from Γ_1, then lines 14 and 15 would be marked in view of Definition 3.

As for the reliability strategy, we can use Definition 5 in order to establish final derivability in an $\mathbf{AL}^{\mathbf{m}}$-proof. Analogous to Definition 6, we define a syntactic consequence relation for $\mathbf{AL}^{\mathbf{m}}$ as follows:

Definition 12. $\Gamma \vdash_{\mathbf{AL}^{\mathbf{m}}} A$ (A is finally $\mathbf{AL}^{\mathbf{m}}$-derivable from Γ) iff A is finally derived at a line of an $\mathbf{AL}^{\mathbf{m}}$-proof from Γ.

Returning to our example, we can define a syntactic consequence relation for the logic $\mathbf{P}^{\mathbf{m}}_{\Diamond}$ by replacing "$\mathbf{AL}^{\mathbf{m}}$" by "$\mathbf{P}^{\mathbf{m}}_{\Diamond}$" in Definition 12.

Recall from Section 4.4.2 that each $\mathbf{P}_\Diamond$-contingent formula of the form $\neg\Diamond \bigwedge \Delta \vee \neg\Diamond \bigwedge \Delta'$, where $\Delta, \Delta' \subset \mathcal{W}^l$ and $p, r \in \Delta$ and $q, s \in \Delta'$, is a minimal *Dab*-consequence of Γ_1. Thus, there are infinitely many ways in which we could extend the proof from Γ_1 with new minimal *Dab*-formulas. Each time such a formula is added to the proof, the set of minimal choice sets is updated.

We could, for instance, add each disjunction $A \vee B$ to the proof, where $A \in \{\neg\Diamond(p\wedge r), \neg\Diamond(p\wedge q\wedge r), \neg\Diamond(p\wedge\neg q\wedge r), \neg\Diamond(p\wedge r\wedge s), \neg\Diamond(p\wedge r\wedge\neg s), \neg\Diamond(p\wedge\neg q\wedge r\wedge s), \neg\Diamond(p\wedge q\wedge r\wedge\neg s), \neg\Diamond(p\wedge\neg q\wedge r\wedge\neg s)\}$ and $B \in \{\neg\Diamond(q\wedge s), \neg\Diamond(p\wedge q\wedge s), \neg\Diamond(\neg p\wedge q\wedge s), \neg\Diamond(q\wedge r\wedge s), \neg\Diamond(q\wedge\neg r\wedge s), \neg\Diamond(\neg p\wedge q\wedge r\wedge s), \neg\Diamond(p\wedge q\wedge\neg r\wedge s), \neg\Diamond(\neg p\wedge q\wedge\neg r\wedge s)\}$. Since two of these disjunctions are already in the proof (lines 4 and 17), this would result in the addition of 62 new minimal *Dab*-formulas. At stage 18 + 62 of the proof, we would then obtain the following minimal choice sets of Γ_1:[8]

$$\varphi_1 = \{\neg\Diamond(p\wedge r), \neg\Diamond(p\wedge q\wedge r), \neg\Diamond(p\wedge\neg q\wedge r), \neg\Diamond(p\wedge r\wedge s), \neg\Diamond(p\wedge r\wedge\neg s), \neg\Diamond(p\wedge q\wedge r\wedge s), \neg\Diamond(p\wedge\neg q\wedge r\wedge s), \neg\Diamond(p\wedge q\wedge r\wedge\neg s), \neg\Diamond(p\wedge\neg q\wedge r\wedge\neg s)\}$$

$$\varphi_2 = \{\neg\Diamond(q\wedge s), \neg\Diamond(p\wedge q\wedge s), \neg\Diamond(\neg p\wedge q\wedge s), \neg\Diamond(q\wedge r\wedge s), \neg\Diamond(q\wedge\neg r\wedge s), \neg\Diamond(p\wedge q\wedge r\wedge s), \neg\Diamond(\neg p\wedge q\wedge r\wedge s), \neg\Diamond(p\wedge q\wedge\neg r\wedge s), \neg\Diamond(\neg p\wedge q\wedge\neg r\wedge s)\}$$

For the same reasons as before, lines 8 and 9 remain marked at stage 18 + 62 while lines 10 and 11 remain unmarked.

There are various ways in which we can even further extend the proof, but none of these would cause the unmarking of any of lines 8, 9, 12, 13, or 16: the formulas derived at lines 8 and 12 are not derivable on a condition that does not intersect with φ_1; the formulas derived at lines 9 and 13 are not derivable on a condition that does not intersect with φ_2; and the formula derived at line 16 is not derivable on a condition that does not intersect with φ_1 or φ_2. Moreover, any extension of the proof in which any of lines 7, 10, 11, 14 or 15 is marked can be further extended so that these lines are unmarked again. There are no minimal *Dab*-consequences of Γ_1 containing any of the conditions of lines 7, 10 or 11. Thus, lines 7, 10 or 11 can never be marked in an extension of the proof. Moreover, if the proof is extended in such a way that line 14 or line 15 is marked, we can further extend it in a way that the lines are unmarked again. In view of Definition 12:

$$\begin{array}{lll}
\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p\wedge r) & \Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(q\wedge s) & \Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p\wedge q\wedge s) \\
\Gamma_1 \vdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p\wedge s) & \Gamma_1 \vdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(r\wedge s) & \Gamma_1 \vdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p\wedge q\wedge (r\vee s)) \\
\Gamma_1 \vdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(q\wedge r) & \Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p\wedge q\wedge r) & \Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p\wedge q\wedge r\wedge s)
\end{array}$$

Note that $\Gamma_1 \vdash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p\wedge(q\vee r))$, whereas $\Gamma_1 \nvdash_{\mathbf{P}^{\mathbf{r}}_\Diamond} \mathsf{O}(p\wedge(q\vee r))$. We return in some detail to this difference between both strategies in Section 4.6.

4.5.2 Semantics

Semantically, the minimal abnormality strategy selects all **LLL**-models of a premise set Γ which have a *minimal* abnormal part (with respect to set-inclusion).

[8]Calculating these choice sets is straightforward, but rather tedious. We leave it to the skeptical reader to double-check our calculations.

Definition 13. A **LLL**-model M of Γ is *minimally abnormal* iff there is no **LLL**-model M' of Γ such that $Ab(M') \subset Ab(M)$.

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

Definition 14. $\Gamma \vDash_{\mathbf{AL^m}} A$ iff A is verified by all minimally abnormal models of Γ.

Reconsider Γ_1 and its models M_1, M_2, and M_3 from Section 4.4.2, the abnormal parts of which are displayed in Table 4.2. Clearly, $Ab(M_1) \subset Ab(M_3)$ and $Ab(M_2) \subset Ab(M_3)$. By Definition 13, M_3 cannot be a minimally abnormal $\mathbf{P}_\Diamond$-model of Γ_1. Moreover, $Ab(M_1) \not\subset Ab(M_2)$ and $Ab(M_2) \not\subset Ab(M_1)$.

Consider again a model M_4 that verifies the abnormality $\neg\Diamond(p \wedge s)$. Since M_4 must also verify either the abnormality $\neg\Diamond(p \wedge r)$ or the abnormality $\neg\Diamond(q \wedge s)$, and since neither M_1 nor M_2 verifies $\neg\Diamond(p \wedge s)$, we know that either $Ab(M_1) \subset Ab(M_4)$ or $Ab(M_2) \subset Ab(M_4)$. By Definition 13, M_4 cannot be a minimally abnormal $\mathbf{P}_\Diamond$-model of Γ_1.

The only minimally abnormal $\mathbf{P}_\Diamond$-models of Γ_1 are models that have the same abnormal part as M_1 or M_2. These models all verify either $\neg\Diamond(p \wedge q \wedge r)$ or $\neg\Diamond(p \wedge q \wedge s)$, *but not both.* Consequently, these models all verify either $\Diamond(p \wedge q \wedge r)$ or $\Diamond(p \wedge q \wedge s)$. Since $\Diamond(p \wedge q \wedge r) \vDash_{\mathbf{P}_\Diamond} \Diamond(p \wedge q \wedge (r \vee s))$ and $\Diamond(p \wedge q \wedge s) \vDash_{\mathbf{P}_\Diamond} \Diamond(p \wedge q \wedge (r \vee s))$, all minimally abnormal $\mathbf{P}_\Diamond$-models of Γ_1 verify $\Diamond(p \wedge q \wedge (r \vee s))$. Moreover, $\Gamma_1 \vDash_{\mathbf{P}_\Diamond} \mathsf{O}p \wedge \mathsf{O}q \wedge \mathsf{O}(r \vee s)$. By ($\Diamond$AND), all minimally abnormal $\mathbf{P}_\Diamond$-models of Γ_1 verify $\mathsf{O}(p \wedge q \wedge (r \vee s))$. By Definition 14, $\Gamma_1 \vDash_{\mathbf{P}^{\mathbf{m}}_\Diamond} \mathsf{O}(p \wedge q \wedge (r \vee s))$.

4.6 Comparing the strategies

As our example from the previous sections illustrates, there are premise sets from which some consequences are derivable by means of the minimal abnormality strategy, while they are not derivable by means of the reliability strategy. In fact we can prove a far stronger result. In [16] it is shown generically that $\mathbf{AL^x}$ is always at least as strong as **LLL**, and that $\mathbf{AL^m}$ is always at least as strong as $\mathbf{AL^r}$:[9]

Theorem 7. $\mathrm{Cn}_{\mathbf{LLL}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{AL^r}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{AL^m}}(\Gamma)$.

It seems, then, that if we literally want to implement a certain standard of normality 'as much as possible', we should always use the minimal abnormality strategy. Indeed, $\mathbf{AL^m}$ will often deliver more consequences than $\mathbf{AL^r}$. Should we then abandon reliability?

Let us briefly recapitulate the rationale underlying both strategies. Suppose that a formula A is derived in an adaptive proof on the condition Δ, and that later in the proof A is also derived on the condition Δ'. Suppose further that, at an even later stage, we derive the minimal *Dab*-formula $Dab(\Delta) \vee Dab(\Delta')$, that this is the only minimal *Dab*-formula derivable from the premise set, and that there are no other conditions on which we can derive A.

[9] **L** is *stronger* than $\mathbf{L}'$ ($\mathbf{L}'$ is *weaker* than **L**) iff for every $\Gamma \subseteq \mathcal{W}^{\mathbf{L}}$, $\mathrm{Cn}_{\mathbf{L}'}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{L}}(\Gamma)$.

In this case, A is not derivable if the proof is an $\mathbf{AL^r}$-proof, since each condition on which A is derived overlaps with the set of unreliable formulas. However, A *is* derivable if the proof is a $\mathbf{AL^m}$-proof, since for each of the two minimal choice sets $\{\Delta\}$ and $\{\Delta'\}$ there is a condition on which A is derived that does not intersect with the minimal choice set.

Conclusions such as A are instances of so-called *floating conclusions*. A conclusion is said to be floating if there is no single correct argument supporting it; instead, it is supported only by conflicting arguments not all of which can be jointly correct.[10]

Horty has argued that, since we do not know which of the conflicting arguments is correct, and since the inference from each argument to the conclusion is defeasible, it is in some cases recommended to opt for the skeptical approach of 'withholding judgment' instead of a less cautious approach by which floating conclusions are taken to be valid [94].

In Section 5.2.5 we discuss some examples in which it is not intuitively clear which strategy we should adopt. We refer to the literature on floating conclusions for more arguments pro and contra the use of a skeptical strategy like reliability.[11]

Since all ALs defined within the standard format can be equipped with either strategy, we need not decide the matter here. We only aim to show that one should not always blindly adopt the minimal abnormality strategy at the expense of the more skeptical reliability strategy.

Let us conclude by briefly mentioning two more differences between both strategies. First, the intuition behind the reliability strategy is easier to understand from a proof theoretical perspective. According to this strategy, a line in an adaptive proof is marked as soon as its condition intersects with Δ for some minimal *Dab*-formula $Dab(\Delta)$ derived in the proof. The intuition underlying minimal abnormality is easier to grasp from a semantical perspective. A **LLL**-model M of some premise set Γ is minimally abnormal if its abnormal part is minimal with respect to set-inclusion.

Second, the reliability strategy takes a lower place in the hierarchy of computational complexity. Whereas ALs making use of reliability are Σ^0_3-complex, ALs that use minimal abnormality can be up to Π^1_1-complex.[12] We refer to [21, 89, 187] for more details on the computational complexity of ALs.

4.7 Meta-theory of the standard format

The main advantage of formulating an AL within the standard format is that a number of meta-theoretical properties come for free for the resulting logic. We mention some of these properties below. In square brackets, we add the references to the literature where the theorems were formulated and proven.

Theorem 8. *[Soundness and completeness]* $\Gamma \vdash_{\mathbf{AL^x}} A$ *iff* $\Gamma \vDash_{\mathbf{AL^x}} A$. *[16, Cor. 2, Th. 9]*

[10] The term 'floating conclusion' was coined by Makinson and Schlechta in [114].

[11] See [62, 114] for arguments pro, and [94] for arguments contra the acceptance of floating conclusions.

[12] These are upper bounds. In concrete instances ALs are often less computationally complex.

Theorem 9. *[Reflexivity]* $\Gamma \subseteq \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)$. *[16, Th. 11.2]*

Theorem 10. *[Fixed point/idempotence]* $\mathrm{Cn}_{\mathbf{AL^x}}(\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)) = \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)$. *[16, Th. 11.6, Th. 11.7]*

In Sections 4.4.2 and 4.5.2 we explained how an AL selects a subset of its LLL-models. It was proven generically for ALs in the standard format that if Γ has **LLL**-models, then it has $\mathbf{AL^x}$-models. This property is called *reassurance*, since it 'reassures' that, unless Γ is **LLL**-trivial, $\mathbf{AL^x}$ will not trivialize Γ.

A slightly stronger result is that, if a **LLL**-model of Γ is not selected, then there is a **LLL**-model M' of Γ that *is* selected and for which $Ab(M') \subset Ab(M)$. This property, which entails the reassurance property, is called *strong reassurance*:

Theorem 11. *[Strong reassurance] If $M \in \mathcal{M}_{\mathbf{LLL}}(\Gamma) - \mathcal{M}_{\mathbf{AL^x}}(\Gamma)$, then there is a $M' \in \mathcal{M}_{\mathbf{AL^x}}(\Gamma)$ such that $Ab(M') \subset Ab(M)$. [16, Th. 4-5]*

The following two theorems further clarify the relation between ALs and their LLL:

Theorem 12. *[**LLL**-closure]* $\mathrm{Cn}_{\mathbf{LLL}}(\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)) = \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)$. *[16, Th. 11.8]*

Theorem 13. *[**LLL**-invariance]* $\mathrm{Cn}_{\mathbf{AL^x}}(\mathrm{Cn}_{\mathbf{LLL}}(\Gamma)) = \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)$. *[16, Th. 15.2]*

Theorem 14. *[Cumulative indifference] If $\Gamma' \subseteq \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)$, then $\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma) = \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma \cup \Gamma')$. [16, Th. 11.10]*

The cumulative indifference property warrants that whenever $\Gamma \vdash_{\mathbf{AL^x}} A$, the $\mathbf{AL^x}$-closure of $\Gamma \cup \{A\}$ is the same as the $\mathbf{AL^x}$-closure of Γ. Where $\Gamma' \subseteq \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)$, cumulative indifference is sometimes split up into the *cumulative monotonicity* property ($\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma \cup \Gamma')$) and the *cumulative transitivity* property ($\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma) \supseteq \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma \cup \Gamma')$).

In [24] the authors argue that ALs have certain advantages over other formal approaches to defeasible reasoning. The gist of their argument is that ALs are more transparent in their treatment of equivalent premise sets. Where **L** is a *Tarski-logic* if **L** is reflexive, monotonic and transitive, they specify three criteria of $\mathbf{AL^x}$-equivalence:

Theorem 15. $\mathrm{Cn}_{\mathbf{AL}}(\Gamma) = \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma')$ *if one of the following holds:*

- *(C1)* $\Gamma' \subseteq \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma)$ *and* $\Gamma \subseteq \mathrm{Cn}_{\mathbf{AL^x}}(\Gamma')$. *[24, Th. 6]*
- *(C2)* *Where **L** is a Tarksi-logic weaker than or identical to $\mathbf{AL^x}$:* $\mathrm{Cn}_{\mathbf{L}}(\Gamma) = \mathrm{Cn}_{\mathbf{L}}(\Gamma')$. *[24, Th. 7]*
- *(C3)* *Where **L** is a Tarski-logic and for every $\Theta \in \mathcal{W}^{\mathbf{L}}$,* $\mathrm{Cn}_{\mathbf{AL^x}}(\Theta) = \mathrm{Cn}_{\mathbf{L}}(\mathrm{Cn}_{\mathbf{AL^x}}(\Theta)) : \mathrm{Cn}_{\mathbf{L}}(\Gamma) = \mathrm{Cn}_{\mathbf{L}}(\Gamma')$. *[24, Th. 7]*

In view of the following theorem:

Theorem 16 (Maximality of **LLL**). *Every monotonic logic that is weaker than or identical to $\mathbf{AL^x}$ is weaker than or identical to **LLL**. [24, Th. 10]* criterion (C2) can be strengthened to (C2'):

(C2') Where **L** is a monotonic logic weaker than or identical to **LLL**: if Γ and Γ' are **L**-equivalent, then they are $\mathbf{AL^x}$-equivalent.

For a detailed discussion on the various criteria for equivalence, we refer to [24].

4.8 Internal and external dynamics of ALs

In adaptive proofs, inferences may be withdrawn for different reasons. First, a formula derived in an adaptive proof may be withdrawn in view of the availability of new information. Here, our reasoning process is non-monotonic: conclusions derivable from a premise set may not be derivable anymore if further premises are added. This non-monotonic aspect of our reasoning corresponds to what we call the *external dynamics* of defeasible reasoning.

Second, a formula derived in an adaptive proof may be withdrawn in view of an increased understanding of the premises. As we reason along, we may gain new insights in the premises even without the addition of genuinely new information in the form of new premises. This type of dynamics is called the *internal dynamics* of defeasible reasoning.

The external dynamics concerns the consequence relation of a logic, and the way it deals with the addition of new premises to those already present. The internal dynamics concerns the actual reasoning steps displayed by an agent, and how she stepwise obtains more insights in the premises.[13]

In order to model the internal dynamics, we use a proof theory. A unique feature of adaptive proofs is that they nicely explicate the internal dynamics of defeasible reasoning. We end this section with an illustration of this internal dynamics in a concrete proof. Let $\Gamma = \{\mathsf{O}p, \mathsf{O}r, \neg(\Diamond(p \wedge q) \wedge \Diamond(p \wedge r)), \mathsf{P}\neg(p \wedge r) \supset \neg\Diamond(p \wedge q)\}$. Consider the following $\mathbf{P}^{\mathbf{x}}_{\Diamond}$-proof from Γ:

1	$\mathsf{O}p$	PREM	$\varnothing$
2	$\mathsf{O}r$	PREM	$\varnothing$
3	$\neg(\Diamond(p \wedge q) \wedge \Diamond(p \wedge r))$	PREM	$\varnothing$
4	$\mathsf{P}\neg(p \wedge r) \supset \neg\Diamond(p \wedge q)$	PREM	$\varnothing$
5	$\mathsf{O}(p \wedge r)$	1,2; RC	$\{\neg\Diamond(p \wedge r)\}$

At stage 5 of the proof, line 5 is unmarked. However, at the next stage we have derived a minimal *Dab*-formula containing its condition, which causes the marking of line 5:

5	$\mathsf{O}(p \wedge r)$	1,2; RC	$\{\neg\Diamond(p \wedge r)\}\checkmark^{6}$
6	$\neg\Diamond(p \wedge q) \vee \neg\Diamond(p \wedge r)$	3; RU	$\varnothing$

Suppose now that we continue the proof as follows:

5	$\mathsf{O}(p \wedge r)$	1,2; RC	$\{\neg\Diamond(p \wedge r)\}$
6	$\neg\Diamond(p \wedge q) \vee \neg\Diamond(p \wedge r)$	3; RU	$\varnothing$
7	$\neg\Diamond(p \wedge r) \supset \neg\mathsf{O}(p \wedge r)$	RU	$\varnothing$
8	$\neg\Diamond(p \wedge r) \supset \neg\Diamond(p \wedge q)$	4,7; RU	$\varnothing$
9	$\neg\Diamond(p \wedge q)$	6,8; RU	$\varnothing$

[13] Pollock uses the terms 'synchronic defeasibility' and 'diachronic defeasibility' for referring to the external, respectively internal, dynamics of human defeasible reasoning [142].

The $\mathbf{P}_\Diamond$-valid formula derived at line 7 follows by contraposition from the instance $\mathsf{O}(p \wedge r) \supset \Diamond(p \wedge r)$ of the schema (OIC). The formula derived at line 8 follows from those derived at lines 4 and 7 by **CL** and the definition of the P-operator. Due to the derivation of a shorter *Dab*-formula at line 9, the *Dab*-formula derived at line 6 is no longer minimal. As a result, line 5 is unmarked again at stage 9.

Altogether, the example illustrates that, without the addition of new information to the premise set Γ, a formula can be derivable at some stage in the proof, not derivable at a later stage, and derivable again at an even later stage.

4.9 The upper limit logic

Above, we have talked extensively about 'standards of normality'. The standard format for ALs provides us with the formal machinery for making this standard technically precise. The standard of normality of an AL is called its *upper limit logic* (ULL).

The ULL of an AL is obtained by adding to the LLL one or more axiom schemas and/or rules that trivialize exactly those formulas that are members of Ω.

In case no *Dab*-formulas are derivable from a premise set by means of the lower limit logic, it is safe to consider all abnormalities as false. As a consequence, the adaptive logic will then yield the same consequence set as the ULL, i.e. the logic that interprets all abnormalities as false (or equivalently, the logic that unconditionally validates the inference rules whose application the adaptive logic only allows conditionally). In general, the upper limit logic **ULL** of $\mathbf{AL^x}$ is related to **LLL** as set out by the *Derivability Adjustment Theorem*:

Theorem 17. $\Gamma \vdash_{\mathbf{ULL}} A$ *iff (there is a* $\Delta \subseteq \Omega$ *for which* $\Gamma \vdash_{\mathbf{LLL}} A \vee Dab(\Delta)$ *or* $\Gamma \vdash_{\mathbf{LLL}} A$*).*

The set of *Dab*-consequences derivable from the premise set determines the extent to which the $\mathbf{AL^x}$-consequence set resembles the **ULL**-consequence set. This is why adaptive logicians say that $\mathbf{AL^x}$ *adapts* itself to a premise set. $\mathbf{AL^x}$ is always at least as strong as **LLL** and maximally as strong as **ULL**:

Theorem 18. $Cn_{\mathbf{LLL}}(\Gamma) \subseteq Cn_{\mathbf{AL^x}}(\Gamma) \subseteq Cn_{\mathbf{ULL}}(\Gamma)$.

Corollary 1. $Cn_{\mathbf{LLL}}(\Gamma) \subseteq Cn_{\mathbf{AL^r}}(\Gamma) \subseteq Cn_{\mathbf{AL^m}}(\Gamma) \subseteq Cn_{\mathbf{ULL}}(\Gamma)$.

Corollary 1 immediately follows from Theorems 7 and 18.

If Γ is *normal*, i.e. if $U(\Gamma) = \emptyset$, then we can even prove a stronger result:

Theorem 19. *If* Γ *is normal, then* $Cn_{\mathbf{AL^x}}(\Gamma) = Cn_{\mathbf{ULL}}(\Gamma)$.

For the proofs of Theorems 17-19, we refer to [16].

The upper limit logic $\mathbf{UP}_\Diamond$ of $\mathbf{P}_\Diamond$ is obtained by adding to the latter system the axiom (U$\Diamond$). Where $\Delta \subset \mathcal{W}^l$ and, for all $A \in \mathcal{W}^a$, $\{A, \neg A\} \nsubseteq \Delta$:

$$\Diamond \bigwedge \Delta \tag{U$\Diamond$}$$

It is easily checked that (U$\Diamond$) trivializes exactly those formulas that are in the set Ω of $\mathbf{P}^{\mathbf{x}}_{\Diamond}$-abnormalities. By Corollary 1,

Corollary 2. $Cn_{\mathbf{P}_{\Diamond}}(\Gamma) \subseteq Cn_{\mathbf{P}^{\mathbf{r}}_{\Diamond}}(\Gamma) \subseteq Cn_{\mathbf{P}^{\mathbf{m}}_{\Diamond}}(\Gamma) \subseteq Cn_{\mathbf{UP}_{\Diamond}}(\Gamma)$

Where $\Delta \subset \mathcal{W}^l$, the constraint that, for all $A \in \mathcal{W}^a$, $\{A, \neg A\} \nsubseteq \Delta$ warrants that $\mathbf{UP}_{\Diamond}$ does not trivialize every premise set (including the empty set). For suppose that we remove this constraint. Then, since it is valid in $\mathbf{P}_{\Diamond}$, the formula $\neg \Diamond (p \wedge \neg p)$ would be a member of Ω. Since $\mathbf{UP}_{\Diamond}$ falsifies all members of Ω, $\Diamond(p \wedge \neg p)$ would then be $\mathbf{UP}_{\Diamond}$-valid. Since $\mathbf{UP}_{\Diamond}$ extends $\mathbf{P}_{\Diamond}$, both $\Diamond(p \wedge \neg p)$ and its negation would be theorems of this logic. Hence by **CL** every member of $\mathcal{W}^{\mathsf{O}}_{\Box}$ would be $\mathbf{UP}_{\Diamond}$-valid.

Fact 4. $\mathsf{O}A, \mathsf{O}B \vdash_{\mathbf{UP}_{\Diamond}} \mathsf{O}(A \wedge B)$.

Proof. Suppose $\mathsf{O}A$ and $\mathsf{O}B$.
Case 1. $\{A, B\}$ is **CL**-consistent. Then $\nvdash_{\mathbf{UP}_{\Diamond}} \neg \Diamond (A \wedge B)$. Let $C_1 \vee \ldots \vee C_n$ be a disjunctive normal form of $A \wedge B$ of which each C_i is **CL**-consistent. By (U$\Diamond$), it follows that $\Diamond C_1$. By **K**-properties, it follows that $\Diamond(C_1 \vee \ldots \vee C_n)$. By **K** again, $\Diamond(A \wedge B)$. By ($\Diamond$AND), $\mathsf{O}(A \wedge B)$.
Case 2. $\{A, B\}$ is **CL**-inconsistent. Then (†) $\vdash_{\mathbf{CL}} \neg(A \wedge B)$. By $\mathsf{O}A$ and (OIC), $\Diamond A$. By (†) and **K**-properties, (‡) $\Diamond(A \wedge \neg(A \wedge B))$.
By (†) and (NEC), $\vdash_{\mathbf{UP}_{\Diamond}} \mathsf{O}\neg(A \wedge B)$. Hence, by $\mathsf{O}A$ and (‡), it follows by ($\Diamond$AND) that $\mathsf{O}(A \wedge \neg(A \wedge B))$. By (RM), it follows that (♯) $\mathsf{O}\neg B$.
By $\mathsf{O}B$ and (CONS), it follows that $\neg\mathsf{O}\neg B$, which contradicts (♯). Hence, by (ECQ), $\mathsf{O}(A \wedge B)$. □

We can now prove the yet stronger result that **SDL** is a fragment of $\mathbf{UP}_{\Diamond}$. Where $\Gamma \subseteq \mathcal{W}^{\mathsf{O}}$ and $A \in \mathcal{W}^{\mathsf{O}}$:

Theorem 20. *If* $\Gamma \vdash_{\mathbf{SDL}} A$, *then* $\Gamma \vdash_{\mathbf{UP}_{\Diamond}} A$.

Proof. We show that all axiom schemas and rules of **SDL** are valid in $\mathbf{UP}_{\Diamond}$. Since $\mathbf{UP}_{\Diamond}$ already contains **CL** and (NEC), we need only show that (K), (P) and (D) are $\mathbf{UP}_{\Diamond}$-valid. For (P), this is immediate in view of the definition of the P-operator. For (D), this is immediate in view of (CONS) and the definition of the P-operator.
Ad (K). Suppose $\mathsf{O}(A \supset B)$ and $\mathsf{O}A$. By Fact 4, $\mathsf{O}(A \wedge (A \supset B))$. By (RM), $\mathsf{O}B$. □

4.10 Some other strategies

Apart from the reliability and minimal abnormality strategies, some other adaptive strategies have been proposed in the literature on ALs. We briefly discuss three more proposals. Except for the simple strategy, these are not defined within the standard format.

1. The *simple strategy* is suitable whenever we know that $\Gamma \vdash_{\mathbf{LLL}} Dab(\Delta)$ iff $\Gamma \vdash_{\mathbf{LLL}} A$ for some $A \in \Delta$. In this case, the reliability and minimal

abnormality strategies come to the same. According to the simple strategy, a line l in an adaptive proof is marked at stage s iff an element of the condition of this line is derived at s on the empty condition.

Not surprisingly, the main interest of the simple strategy lies in its simplicity. In contexts in which the strategy is suitable, employing its straightforward marking criterion leads to a far less complex adaptive logic as compared to e.g. the reliability and minimal abnormality strategies. For some concrete ALs that make use of the simple strategy, see e.g. [59, 176].

2. We already mentioned that reliability is a more 'skeptical' strategy than minimal abnormality (cfr. Section 4.6). Scholars in artificial intelligence often use the term 'skeptical' in a different way in the context of non-monotonic reasoning. For any logic **L**, let a *maximally* **L***-consistent subset* be defined as follows:

   Definition 15. Where Γ and Δ are sets of **L**-formulas, Δ is a *maximally* **L***-consistent subset* of Γ iff (i) $\Delta \subseteq \Gamma$, (ii) Δ is **L**-consistent, and (iii) there is no **L**-consistent set Δ' such that $\Delta \subset \Delta'$ and $\Delta' \subseteq \Gamma$.

   Let A be a *skeptical* **L**-consequence of Γ iff A is an **L**-consequence of *each* maximally **L**-consistent subset of Γ. A is a *credulous* **L**-consequence of Γ iff A is an **L**-consequence of *some* maximally **L**-consistent subset of Γ.

   The *normal selections strategy* is more in the spirit of 'credulous' consequence relations. In adaptive proofs that make use of this strategy, a line l with condition Δ is marked at stage s iff $Dab(\Delta)$ has been derived at s on the empty condition. In [25], the normal selections strategy was used for characterizing a credulous consequence relation by means of an adaptive logic.

   In Section 5.3, we provide an example of a skeptical and a credulous consequence relation as defined in the AI-tradition, and compare its consequences to those derivable by means of adaptive logics making use of the reliability and minimal abnormality strategies.

3. So far, all adaptive strategies mentioned are qualitative. The *counting strategy* is an example of a quantitative strategy. The idea behind the counting strategy is that a **LLL**-model M is selected iff no other **LLL**-model M' verifies less abnormalities than M. As an application of this strategy, we can think of a number of witnesses in a trial. If equally trustworthy witnesses contradict each other, then according to the counting strategy a statement could be plausible if the witnesses affirming it outnumber those denying it. For an illustration of the counting strategy, see e.g. [147].

For a more detailed discussion on these strategies, and for some more examples of adaptive strategies not in the standard format, see [19, Section 6.1].

Chapter 5

Non-aggregative adaptive logics for normative conflicts

✎ Section 5.2 of this chapter is based on the paper *Non-Adjunctive Deontic Logics That Validate Aggregation as Much as Possible* (Journal of Applied Logic, conditionally accepted) [130], which is co-authored by Joke Meheus, Frederik Van De Putte and Christian Straßer.

✎ I am indebted to Joke Meheus and Christian Straßer for valuable comments on this chapter.

In developing formal systems capable of accommodating conflicting obligations, the historically dominant strategy is to restrict the aggregation principle (AND). In this chapter, we present some adaptive systems that allow for 'the right amount' of aggregation given their intended application contexts.

In Section 5.1, we assess the 'Williams-style' non-aggregative adaptive logic $\mathbf{P}^{\mathsf{x}}_{\Diamond}$ defined in the previous chapter. In doing so, we make use of the desiderata for CTDLs presented in Chapter 3.

We continue in Section 5.2 with the presentation and motivation of the logic $\mathbf{P2.2}^{\mathsf{x}}$ which allows for the aggregation of obligations arising from different normative standards. Alternatively $\mathbf{P2.2}^{\mathsf{x}}$ can be interpreted as a logic that allows us to derive all-things-considered obligations from a set of prima facie obligations. In presenting this logic, we also discuss some pitfalls that one should watch out for when constructing ALs. Moreover, we discuss at some length the treatment of incompatible obligations by $\mathbf{P2.2}^{\mathsf{x}}$.

We end this chapter with a comparison of the adaptive systems presented here with their main non-monotonic competitor (Section 5.3).

5.1 Tolerating moral dilemmas

In Section 3.2.1.2 we discussed Williams' claim that moral conflicts can be formalized as formulas of the form $\mathsf{O}A_1 \wedge \ldots \wedge \mathsf{O}A_n \wedge \neg \Diamond (A_1 \wedge \ldots \wedge A_n)$. We saw how defenders of this claim seem to be caught between a rock and a hard place, since they are required to either give up one of (AND) and (OIC), or accept the explosion principle ($\Diamond$-EX).[1]

Williams' own solution was to give up (AND). Although he does not claim to have a knock-down disproof of this principle, he talks of abandoning [203, p. 120], waiving [203, p. 122], and rejecting [203, p. 123] aggregation in order to obtain a more realistic picture of moral thought. Logicians too have proposed giving up the agglomeration principle in its entirety in order to make deontic logic conflict-tolerant, e.g. [58, 65, 157]. However, as we saw in Section 3.2.2.1 there are some problems with this approach.

Even if Williams is correct in claiming that "no agent, conscious of the situation of conflict, in fact thinks that he ought to do both of the things" [203, p. 120], we believe that the same agent will reason very differently in case the ought's in question do not conflict. This was already illustrated in Example 13. Recalling Horty's argument from Section 3.2.2.1, we seem to need an aggregation rule that allows for 'exactly the right amount of aggregation'. Recalling the design requirements from Section 3.3, the logic resulting after adding this rule should be non-explosive and non-monotonic, and should not require that the user needs to do some extra reasoning which is not supported by the formal logic in order to aggregate two obligations.

Taking Williams' formalization of moral conflicts as given, the logic $\mathbf{P}^{\mathbf{x}}_{\Diamond}$ defined in the previous chapter meets all requirements. First of all, where $\nvdash \neg(A \wedge B)$, ($\Diamond$-EX) is invalid in $\mathbf{P}^{\mathbf{x}}_{\Diamond}$:

$$\mathsf{O}A, \mathsf{O}B, \neg \Diamond (A \wedge B) \nvdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \bot \tag{5.1}$$

So are its weaker variants:

$$\mathsf{O}A, \mathsf{O}B, \neg \Diamond (A \wedge B) \nvdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{O}C \tag{5.2}$$

$$\mathsf{O}A, \mathsf{O}B, \neg \Diamond (A \wedge B) \nvdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{P}B \supset \mathsf{O}C \tag{5.3}$$

$$\text{If } \nvdash C, \text{ then } \mathsf{O}A, \mathsf{O}B, \neg \Diamond (A \wedge B) \nvdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{P}C \supset \mathsf{O}C \tag{5.4}$$

Similarly for the other variants of the explosion principles presented in Section 2.3.2.

Second, due to its non-monotonicity $\mathbf{P}^{\mathbf{x}}_{\Diamond}$ allows for the applicability of (AND) wherever it is intuitive to aggregate obligations. For instance, the following inferences are $\mathbf{P}^{\mathbf{x}}_{\Diamond}$-valid:

$$\mathsf{O}p, \mathsf{O}q \vdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{O}(p \wedge q) \tag{5.5}$$

$$\mathsf{O}p, \mathsf{O}q, \mathsf{O}r \vdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r) \tag{5.6}$$

$$\mathsf{O}p, \mathsf{O}q, \mathsf{O}r, \neg \Diamond (p \wedge q) \vdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{O}(p \wedge r) \tag{5.7}$$

$$\mathsf{O}(p \wedge q), \mathsf{O}r, \mathsf{O}s, \neg \Diamond (q \wedge r) \vdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{O}(p \wedge s) \tag{5.8}$$

The following inferences are $\mathbf{P}^{\mathbf{x}}_{\Diamond}$-invalid:

$$\mathsf{O}p, \mathsf{O}q, \neg \Diamond (p \wedge q) \nvdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{O}(p \wedge q) \tag{5.9}$$

[1] We refer to [26, 203] for further details on Williams' characterization of moral conflict and on his solution for accommodating moral conflicts in deontic logic.

$$\mathsf{O}p, \mathsf{O}q, \mathsf{O}r, \neg\Diamond(p \wedge q) \nvdash_{\mathbf{P}^{\mathbf{x}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge r) \tag{5.10}$$

In some cases, it depends on the adaptive strategy used whether a formula is $\mathbf{P}^{\mathbf{x}}_{\Diamond}$-derivable or not. As we saw in the previous chapter:

$$\mathsf{O}(p \wedge q), \mathsf{O}r, \mathsf{O}s, \neg\Diamond(p \wedge r) \vee \neg\Diamond(q \wedge s) \nvdash_{\mathbf{P}^{\mathbf{r}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge (r \vee s)) \tag{5.11}$$

$$\mathsf{O}(p \wedge q), \mathsf{O}r, \mathsf{O}s, \neg\Diamond(p \wedge r) \vee \neg\Diamond(q \wedge s) \vdash_{\mathbf{P}^{\mathbf{m}}_{\Diamond}} \mathsf{O}(p \wedge q \wedge (r \vee s)) \tag{5.12}$$

$\mathbf{P}^{\mathbf{x}}_{\Diamond}$ also treats Horty's Smith example (Example 13) the way it should:

1	$\mathsf{O}(f \vee s)$	PREM	$\varnothing$
2	$\mathsf{O}\neg f$	PREM	$\varnothing$
3	$\Diamond(\neg f \wedge s)$	RC	$\{\neg\Diamond(\neg f \wedge s)\}$
4	$\Diamond(\neg f \wedge (f \vee s))$	3; RU	$\{\neg\Diamond(\neg f \wedge s)\}$
5	$\mathsf{O}(\neg f \wedge (f \vee s))$	1,2,4; RU	$\{\neg\Diamond(\neg f \wedge s)\}$
6	$\mathsf{O}s$	5; RU	$\{\neg\Diamond(\neg f \wedge s)\}$

Finally, it is not required that the user add any extra information not contained in the premise set in order to aggregate two obligations in $\mathbf{P}^{\mathbf{x}}_{\Diamond}$. If no abnormality prevents one from doing so, two obligations can be aggregated without the 'manual' addition of extra premises.

Before we conclude this section, we briefly return to Williams' formalization of moral conflicts. One may argue against Williams that moral conflicts are better formalized by (or reduced to) conjunctions of the form $\mathsf{O}A \wedge \mathsf{O}\neg A$. This appears to be the approach of Goble in [69], where a conflict $\{\mathsf{O}A, \mathsf{O}B, \neg\Diamond(A \wedge B)\}$ reduces to $\{\mathsf{O}A, \mathsf{O}\neg A, \mathsf{O}B, \mathsf{O}\neg B\}$ in view of the principle (NM) introduced in Section 1.6.2.

Williams' formalization of conflicts within the richer language $\mathcal{W}^{\mathsf{O}}_{\Box}$ is more informative. A key feature of his characterization of moral conflicts is that such conflicts arise *via the facts*. In this sense, it is more natural to formalize moral conflicts by making use of an alethic possibility operator. Moreover, when using the formalization $\{\mathsf{O}A, \mathsf{O}\neg A, \mathsf{O}B, \mathsf{O}\neg B\}$ instead of $\{\mathsf{O}A, \mathsf{O}B, \neg\Diamond(A \wedge B)\}$, one loses the information that there is a *link* between the fulfillment of $\mathsf{O}A$ and the fulfillment of $\mathsf{O}B$.

When formalized without the use of alethic modalities, we obtain a more economical, simpler characterization of moral conflicts. When formalized with an alethic possibility operator, we obtain a richer, more expressive characterization. Which characterization is best depends on the context of application and need not be decided here. In the next section, we will present a non-aggregative deontic logic that allows for the formalization of normative conflicts as formulas of the form $\mathsf{O}A \wedge \mathsf{O}\neg A$.

5.2 Aggregating over different normative standards

5.2.1 Introduction

In this section, we present the non-adjunctive adaptive deontic logics $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$. Both $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$ are based on Goble's logic $\mathbf{SDL_aP_e}$ from [65], which we shall henceforth call **P2**. The system **P2** is a bimodal extension of the logic **P** from Section 3.2.2.1. The language of **P2** contains two distinct obligation operators: the operator O_e, which is the one from **P**, and the new operator O_a. The duals P_e and P_a are defined in the usual way, i.e. $\mathsf{P}_e A =_{\text{df}} \neg\mathsf{O}_e\neg A$ and $\mathsf{P}_a A =_{\text{df}} \neg\mathsf{O}_a\neg A$. Goble's motivation for this additional ought-operator is that $\mathsf{O}_e A$ expresses that, under *some* set of norms, A ought to be case, but cannot express that A holds 'universally', under *any* standard. The O_a-operator gives us exactly this. This results in a greater expressive power and also in different ways for formalizing conflicts (see Section 5.2.2).

The logic **P2** behaves exactly like **SDL** for the O_a-operator and like **P** for the O_e-operator. This seems to give the logic some advantages over **P**. Given the proper formalization, one can make sure that for all non-conflicting 'parts' of the premises, the same results are obtained as with **SDL**. For instance, in the Smith example, formalizing the premises as $\mathsf{O}_a(f \vee s)$ and $\mathsf{O}_a\neg f$ ensures that $\mathsf{O}_a s$ is derivable. This solution presupposes, however, that one knows in advance which premises can be safely formalized with the O_a-operator. As such, it presupposes that one knows in advance which 'parts' of the premises are problematic. This requires additional reasoning and is at odds with the user-friendliness requirement from Section 3.3.3.

Using the terminology from Section 1.4.1, we can also let the O_e-operator denote *prima facie* obligations and let the O_a-operator denote *actual* ("all-things-considered") obligations. The logics $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$ as well as their LLL **P2** accept the first main distinguishing feature of prima facie obligations: at this level, conflicts may arise between duties. Formulas of the form $\mathsf{O}_e A \wedge \mathsf{O}_e\neg A$ behave consistently in the systems **P2**, $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$.

The second distinguishing feature of prima facie obligations is that, in case no conflict arises, a prima facie obligation becomes actual. This feature is not caught by the logic **P2**, but is modeled in an intuitive way by $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$. The latter logics allow for the defeasible application of the rule "if $\mathsf{O}_e A$, then $\mathsf{O}_a A$".

The basic idea behind the two new logics is that O_e-obligations are interpreted "as much as possible" as O_a-obligations (that is, unless and until the premises explicitly prevent this). As is clear from the above, all classical operations can be applied to O_a-obligations (aggregation, disjunctive syllogism, ...). Which O_e-obligations are interpreted as O_a-obligations and which not is solely dependent on formal grounds. The logics adapt themselves to the set of premises and localize the conflicts. No interference of the user is required for this.

The systems $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$ are not the first adaptive logics that are based on **P**. In [129], the logic $\mathbf{P2.1^r}$ was presented. This system too was constructed to apply the schema $\mathsf{O}_e A \supset \mathsf{O}_a A$ 'as much as possible'. At first sight, $\mathbf{P2.1^r}$ is a very satisfactory system. It has all the nice properties of **P**, it leads to the same consequence set as **SDL** for conflict-free premise sets, and it allows one to deal with some well-known toy examples in their original formulation.

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

$$\mathsf{O}_e(p \vee q) \tag{5.13}$$

$$\mathsf{O}_e(r \vee s) \tag{5.14}$$
$$\neg\mathsf{O}_a((p \vee q) \wedge (r \vee s)) \tag{5.15}$$
$$\mathsf{O}_e t \tag{5.16}$$

There is clearly an incompatibility between (5.13) and (5.14) in view of (5.15). However, there is nothing wrong with (5.16). Hence one expects to be able to derive $\mathsf{O}_a t$ from this premise set, but $\mathbf{P2.1^r}$ does not allow for this inference. The logics $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$ solve this problem, while retaining all the nice properties of $\mathbf{P2.1^r}$.

In Section 5.2.2, we present Lou Goble's **P2**. Readers interested in the logic $\mathbf{P2.1^r}$ and the problems that it faces, can have a look at Section 5.2.3. In Section 5.2.4 we continue with the presentation of the systems $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$ which overcome the problems faced by $\mathbf{P2.1^r}$. We discuss the treatment of incompatible obligations in $\mathbf{P2.2^x}$ in Section 5.2.5, and state some further properties of $\mathbf{P2.2^x}$ in Section 5.2.6.

5.2.2 The logic P2

The set $\mathcal{W}^{\mathbf{P2}}$ of wffs of **P2** is defined as:

$$\mathcal{W}^{\mathbf{P2}} := \mathcal{W} \mid \mathsf{O}_e\langle\mathcal{W}\rangle \mid \mathsf{O}_a\langle\mathcal{W}\rangle \mid \neg\langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \vee \langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \wedge \langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \supset \langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \equiv \langle\mathcal{W}^{\mathbf{P2}}\rangle$$

The formal characterization of **P2** is exactly that of Goble's $\mathbf{SDL_aP_e}$ from [65], except for one minor aspect. Goble is only interested in the theorems of his logic, not in a consequence relation. As we are mainly interested in the consequence relation, and as we want to talk about the models of premise sets, we shall modify his characterization in such a way that we introduce an actual world in the models.

The idea behind **P2** is simple: in a Kripke-like semantics, aggregation is invalidated by considering a *set* of accessibility relations instead of only one. Intuitively, each accessibility relation can be thought of as corresponding to one of the normative systems an agent adheres to.

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

(Ca)	where $A \in \mathcal{W}^a$, $v_M(A,w) = v(A,w)$
(C$\neg$)	$v_M(\neg A, w) = 1$ iff $v_M(A,w) = 0$
(C$\vee$)	$v_M(A \vee B, w) = 1$ iff $v_M(A,w) = 1$ or $v_M(B,w) = 1$
(C$\wedge$)	$v_M(A \wedge B, w) = 1$ iff $v_M(A,w) = 1$ and $v_M(B,w) = 1$
(C$\supset$)	$v_M(A \supset B, w) = 1$ iff $v_M(A,w) = 0$ or $v_M(B,w) = 1$
(C$\equiv$)	$v_M(A \equiv B, w) = 1$ iff $v_M(A,w) = v_M(B,w)$
(CO_e)	$v_M(\mathsf{O}_e A, w) = 1$ iff, for some $R \in \mathcal{R}$, $v_M(A,w') = 1$ for all w' such that Rww'
(CO_a)	$v_M(\mathsf{O}_a A, w) = 1$ iff, for every $R \in \mathcal{R}$, $v_M(A,w') = 1$ for all w' such that Rww'

A **P2**-model M verifies A ($M \Vdash A$) iff $v_M(A, w_0) = 1$.

P2 is axiomatized by adding to **CL** the following axiom schemas and rules:

$$\mathsf{O}_a(A \supset B) \supset (\mathsf{O}_a A \supset \mathsf{O}_a B) \tag{K$_a$}$$
$$\mathsf{O}_a A \supset \neg\mathsf{O}_a \neg A \tag{D$_a$}$$
$$\text{If } \vdash A \text{ then } \vdash \mathsf{O}_a A \tag{N$_a$}$$
$$\text{If } \vdash A \supset B \text{ then } \vdash \mathsf{O}_e A \supset \mathsf{O}_e B \tag{RM$_e$}$$
$$\text{If } \vdash A \text{ then } \vdash \mathsf{O}_e A \tag{N$_e$}$$
$$\text{If } \vdash A \text{ then } \vdash \neg\mathsf{O}_e \neg A \tag{P$_e$}$$
$$\mathsf{O}_a(A \supset B) \supset (\mathsf{O}_e A \supset \mathsf{O}_e B) \tag{K$_{ae}$}$$

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

Where $\Gamma \subset \mathcal{W}^{\mathbf{P2}}$ is finite and $A \in \mathcal{W}^{\mathbf{P2}}$, we define $\Gamma \vdash_{\mathbf{P2}} A$ iff A is derivable from Γ by the axiom schemas and rules of **P2**, and $\vdash_{\mathbf{P2}} A$ iff A is derivable by the axiom schemas and rules of **P2** from the empty premise set.

In [65], Goble proved soundness and (weak) completeness for **P2**.

Theorem 21. *For any finite* $\Gamma \subset \mathcal{W}^{\mathbf{P2}}$, $\Gamma \vdash_{\mathbf{P2}} A$ *iff* $\Gamma \vDash_{\mathbf{P2}} A$.

The O_a-operator of **P2** is stronger than the O_e-operator:

$$\mathsf{O}_a A \vdash_{\mathbf{P2}} \mathsf{O}_e A \tag{5.17}$$
$$\mathsf{O}_e A \nvdash_{\mathbf{P2}} \mathsf{O}_a A \tag{5.18}$$
$$\mathsf{O}_e A \vdash_{\mathbf{P2}} \neg\mathsf{O}_a \neg A \tag{5.19}$$
$$\mathsf{O}_a A \vdash_{\mathbf{P2}} \neg\mathsf{O}_e \neg A \tag{5.20}$$

Proof. Ad (5.17). Suppose $\mathsf{O}_a A$. By the **CL**-theorem $A \supset ((((A \supset A) \supset A) \supset A) \supset A)$, (N$_a$), (K$_a$) and modus ponens, it follows that (†) $\mathsf{O}_a((((A \supset A) \supset A) \supset A) \supset A)$. Moreover, by the **CL**-theorem $((A \supset A) \supset A) \supset A$ and (N$_e$), we know that $\vdash_{\mathbf{P2}} \mathsf{O}_e(((A \supset A) \supset A) \supset A)$. By (†) and (K$_{ae}$), it follows that $\mathsf{O}_e A$.

The proofs for (5.18)-(5.20) are safely left to the reader. □

P2 tolerates OO-conflicts arising from different normative standards, but does not tolerate single-standard conflicts or conflicts between an O_e- and an O_a-obligation:

$$\mathsf{O}_e A, \mathsf{O}_e \neg A \nvdash_{\mathbf{P2}} \mathsf{O}_e B \tag{5.21}$$
$$\mathsf{O}_a A, \mathsf{O}_a \neg A \vdash_{\mathbf{P2}} \mathsf{O}_a B \tag{5.22}$$
$$\mathsf{O}_e A, \mathsf{O}_a \neg A \vdash_{\mathbf{P2}} \mathsf{O}_a B \tag{5.23}$$

Aggregation holds for O_a-obligations, but not for O_e-obligations. Moreover, from an O_a-obligation to do A and an O_e-obligation to do B, we can derive the O_e-obligation to do A and B, but not the O_a-obligation to do A and B:

$$\mathsf{O}_e A, \mathsf{O}_e B \nvdash_{\mathbf{P2}} \mathsf{O}_e(A \wedge B) \tag{5.24}$$
$$\mathsf{O}_a A, \mathsf{O}_a B \vdash_{\mathbf{P2}} \mathsf{O}_a(A \wedge B) \tag{5.25}$$
$$\mathsf{O}_a A, \mathsf{O}_e B \vdash_{\mathbf{P2}} \mathsf{O}_e(A \wedge B) \tag{5.26}$$

$$\mathsf{O}_a A, \mathsf{O}_e B \nvdash_{\mathbf{P2}} \mathsf{O}_a(A \wedge B) \tag{5.27}$$

All of these inferences are intuitive in view of the interpretation of the O_e- and O_a-operators. We will now extend **P2** so as to account for the intuition that O_e-obligations are interpreted "as much as possible" as O_a-obligations. A first attempt is made in Section 5.2.3. Although this attempt is ultimately unsuccessful, the discussion of it provides some insights concerning the problems that logicians are typically faced with when devising ALs. Readers not interested in these problems can safely skip this section and move immediately to Section 5.2.4, where a more successful solution is provided.

5.2.3 Excursion: the system $\mathbf{P2.1^r}$

In [129], we presented the adaptive logic $\mathbf{P2.1^r}$, which is defined by the LLL **P2**, the reliability strategy, and the set of abnormalities $\Omega = \Omega_1 \cup \Omega_2$, where

- $\Omega_1 = \{\mathsf{O}_e A \wedge \neg \mathsf{O}_a A \mid A \in \mathcal{W}^l\}$,
- $\Omega_2 = \{\mathsf{O}_e(A_1 \vee \ldots \vee A_n) \wedge \neg(\mathsf{O}_e A_1 \wedge \neg \mathsf{O}_a A_1) \wedge \ldots \wedge \neg(\mathsf{O}_e A_n \wedge \neg \mathsf{O}_a A_n) \wedge \neg \mathsf{O}_a(A_1 \vee \ldots \vee A_n) \mid A_1, \ldots, A_n \in \mathcal{W}^l, n \geq 2\}$.

Note that, by **CL**, for any $A \in \mathcal{W}^{\mathbf{P2}}$,

$$\mathsf{O}_e A \vdash_{\mathbf{P2}} \mathsf{O}_a A \vee (\mathsf{O}_e A \wedge \neg \mathsf{O}_a A) \tag{5.28}$$

Thus, if Ω were to contain all formulas of the form $\mathsf{O}_e A \wedge \neg \mathsf{O}_a A$, then in any $\mathbf{P2.1^r}$-proof we could derive $\mathsf{O}_a A$ from $\mathsf{O}_e A$ on the condition that $\mathsf{O}_e A \wedge \neg \mathsf{O}_a A$ is false.

However, as is clear from the definition of Ω, things are not that simple. First, the set Ω_1 contains a restriction to members of $\mathcal{W}^l$. The reason for this is easily demonstrated by means of an example. Consider the premise set $\Gamma_1 = \{\mathsf{O}_e p, \mathsf{O}_e \neg p, \mathsf{O}_e q\}$. In view of these premises, it seems natural to derive the all-things-considered obligation $\mathsf{O}_a q$, whereas we do not want to derive the all-things-considered obligations $\mathsf{O}_a p$ or $\mathsf{O}_a \neg p$ (since the latter are involved in a conflict). Let $\dagger(A)$ abbreviate $\mathsf{O}_e A \wedge \neg \mathsf{O}_a A$. Consider now the following $\mathbf{P2.1^r}$-proof from Γ_1:

1	$\mathsf{O}_e p$	PREM	$\emptyset$
2	$\mathsf{O}_e \neg p$	PREM	$\emptyset$
3	$\mathsf{O}_e q$	PREM	$\emptyset$
4	$\dagger(p)$	1,2; RU	$\emptyset$
5	$\dagger(\neg p)$	1,2; RU	$\emptyset$
6	$\mathsf{O}_a q$	3; RC	$\{\dagger(q)\}$
7	$\mathsf{O}_e(\neg p \vee \neg q)$	2; RU	$\emptyset$
8	$\dagger(q) \vee \dagger(\neg p \vee \neg q)$	1,3,7; RU	$\emptyset$

If Ω were to contain all formulas of the form $\dagger(A)$ where $A \in \mathcal{W}^{\mathbf{P2}}$, then line 6 in the proof would be marked at stage 8 in view of the minimal *Dab*-formula $\dagger(q) \vee \dagger(\neg p \vee \neg q)$. In fact, the logic resulting from defining Ω as the set of all formulas of the form $\dagger(A)$ would be a so-called flip-flop logic. An adaptive logic

$\mathbf{AL^x}$ (with lower limit logic **LLL** and upper limit logic **ULL**) is a *flip-flop logic* if, for all premise sets $\Gamma \subseteq \mathcal{W}^{\mathbf{LLL}}$, (i) if no *Dab*-formulas are **LLL**-derivable from Γ, then $\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma) = \mathrm{Cn}_{\mathbf{ULL}}(\Gamma)$, and (ii) if at least one *Dab*-formula is **LLL**-derivable from Γ, then $\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma) = \mathrm{Cn}_{\mathbf{LLL}}(\Gamma)$. Thus, if $\mathbf{AL^x}$ is a flip-flop logic, then either $\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma) = \mathrm{Cn}_{\mathbf{LLL}}(\Gamma)$ or $\mathrm{Cn}_{\mathbf{AL^x}}(\Gamma) = \mathrm{Cn}_{\mathbf{ULL}}(\Gamma)$.[2]

When trying to accommodate normative conflicts by means of adaptive logics, we clearly do not want these logics to be flip-flop logics. In case (ii), we usually want our logic to deliver a consequence set that is strictly stronger than that delivered by the LLL. Thus, we cannot simply define Ω as the set of all formulas of the form $\dagger(A)$. Let us now return to the actual definition of Ω for the logic $\mathbf{P2.1^r}$.

For conflicts relating to more complex obligations Ω_2 requires that none of the literals contained in such obligations is 'tainted' by a conflict. Note that the formula $\dagger(\neg p \vee \neg q)$ does not meet this requirement, since we already know that $\dagger(\neg p)$ is **P2**-derivable from Γ_1 (see line 5). Thus, $\dagger(q) \vee \dagger(\neg p \vee \neg q)$ does not give rise to a *Dab*-formula and line 6 remains unmarked. Moreover, there is no extension of the proof that would cause the marking of this line. Hence $\mathsf{O}_a q$ is finally $\mathbf{P2.1^r}$-derivable from Γ_1, as desired.

Unfortunately, the requirement that Ω_2-abnormalities be untainted by conflicts of the form $\dagger(A)$ for any of its subformulas $A \in \mathcal{W}^l$ is still insufficient. We illustrate this by means of the premise set $\Gamma_2 = \{\mathsf{O}_e(p \vee q), \mathsf{O}_e(r \vee s), \neg\mathsf{O}_a((p \vee q) \wedge (r \vee s)), \mathsf{O}_e t\}$:

1	$\mathsf{O}_e(p \vee q)$	PREM	$\emptyset$
2	$\mathsf{O}_e(r \vee s)$	PREM	$\emptyset$
3	$\neg\mathsf{O}_a((p \vee q) \wedge (r \vee s))$	PREM	$\emptyset$
4	$\mathsf{O}_e t$	PREM	$\emptyset$

The formulas at lines 1 and 2 are clearly incompatible in view of the formula at line 3. However, there is nothing wrong with the formula at line 4, so we would expect $\mathsf{O}_a t$ to be $\mathbf{P2.1^r}$-derivable from Γ_2.

Let $\ddagger(A_1 \vee \ldots \vee A_n)$ abbreviate $\mathsf{O}_e(A_1 \vee \ldots \vee A_n) \wedge \neg\dagger(A_1) \wedge \ldots \wedge \neg\dagger(A_n) \wedge \neg\mathsf{O}_a(A_1 \vee \ldots \vee A_n)$. Consider the following extension of the $\mathbf{P2.1^r}$-proof from Γ_2:

5	$\mathsf{O}_a t$	4; RC	$\{\dagger(t)\}\checkmark^6$
6	$\ddagger(p \vee q \vee \neg t) \vee \ddagger(r \vee s \vee \neg t) \vee \dagger(p) \vee \dagger(q) \vee \dagger(r) \vee \dagger(s) \vee \dagger(t) \vee \dagger(\neg t)$	1-4; RU	$\emptyset$

It is left to the reader to check that line 6 is indeed **P2**-derivable from Γ_2. Since the disjunction cannot be shortened in such a way that the result is still **P2**-derivable from the premises, the *Dab*-formula derived at line 6 is a minimal *Dab*-consequence of Γ_2. Hence, line 5 will remain marked in any extension of the proof, and $\Gamma_2 \nvdash_{\mathbf{P2.1^r}} \mathsf{O}_a t$.

Since there is no other condition on which the formula $\mathsf{O}_a t$ is derivable in a proof from Γ_2, switching strategies will not help us to derive this formula: $\Gamma_2 \nvdash_{\mathbf{P2.1^m}} \mathsf{O}_a t$.

[2]See [14, 22] for more information on flip-flop logics.

The reason why $\mathbf{P2.1^x}$ breaks down when faced with complex sets of obligations has to do with the definition of the set Ω_2. The logics $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$ overcome these problems by means of a more comprehensive definition of the set of abnormalities.

5.2.4 The logic $\mathbf{P2.2^x}$

Where $x \in \{r, m\}$, the logic $\mathbf{P2.2^x}$, like $\mathbf{P2.1^r}$, has Goble's $\mathbf{P2}$ as its LLL. However, it makes use of a different set Ω of abnormalities. Despite the relatively simple intuition behind it, the definition of this set needs some preparation. Where $\Theta \subset \mathcal{W}^l$ is finite and non-empty, and where

$$\sigma(\Theta) = \{\mathsf{O}_e(\bigvee \Theta') \wedge \neg\mathsf{O}_a(\bigvee \Theta') \mid \Theta' \subseteq \Theta \text{ and } \Theta' \neq \varnothing\}$$

the form of the abnormalities of $\mathbf{P2.2^x}$ is $\bigvee(\sigma(\Theta))$.

If $\Theta = \{p\}$, then $\bigvee(\sigma(\Theta))$ is simply the formula $\mathsf{O}_e p \wedge \neg\mathsf{O}_a p$. As a more complex example, consider the set $\Theta = \{p,\ q,\ \neg r\}$. In that case, $\bigvee(\sigma(\Theta))$ stands for the formula $(\mathsf{O}_e(p \vee q \vee \neg r) \wedge \neg\mathsf{O}_a(p \vee q \vee \neg r)) \vee (\mathsf{O}_e(p \vee q) \wedge \neg\mathsf{O}_a(p \vee q)) \vee (\mathsf{O}_e(p \vee \neg r) \wedge \neg\mathsf{O}_a(p \vee \neg r)) \vee (\mathsf{O}_e(q \vee \neg r) \wedge \neg\mathsf{O}_a(q \vee \neg r)) \vee (\mathsf{O}_e p \wedge \neg\mathsf{O}_a p) \vee (\mathsf{O}_e q \wedge \neg\mathsf{O}_a q) \vee (\mathsf{O}_e \neg r \wedge \neg\mathsf{O}_a \neg r)$. For reasons of transparency, we shall in the remainder use $\sharp(p \vee q \vee \neg r)$ instead of $\bigvee(\sigma(\{p, q, \neg r\}))$. More generally, we shall use $\sharp(A_1 \vee \ldots \vee A_n)$ (where $n \geq 1$) instead of $\bigvee(\sigma(\{A_1, \ldots, A_n\}))$.

The set Ω of $\mathbf{P2.2^x}$-abnormalities is defined as follows:

$$\Omega = \left\{\bigvee(\sigma(\Theta)) \mid \Theta \subset \mathcal{W}^l, \Theta \neq \varnothing, \Theta \text{ is finite}\right\}$$

The intuition behind the definition of Ω is that whenever a formula $\mathsf{O}_e A \in \mathcal{W}^{\mathbf{P2}}$ *or any of its subformulas* is involved in a conflict, then $\mathsf{O}_e A$ gives rise to a $\mathbf{P2.2^x}$-abnormality or to a disjunction of $\mathbf{P2.2^x}$-abnormalities.

As mentioned above, the idea behind $\mathbf{P2.2^x}$ is that O_e-obligations are interpreted "as much as possible" as O_a-obligations. In an adaptive proof, we can derive O_a-obligations from O_e-obligations via the conditional rule RC. Where $A \in \mathcal{W}^l$, we can derive $\mathsf{O}_a A \vee (\mathsf{O}_e A \wedge \neg\mathsf{O}_a A)$ from $\mathsf{O}_e A$ by $\mathbf{CL}$ (since $\mathsf{O}_a A \vee \neg\mathsf{O}_a A$ is $\mathbf{P2}$-valid). Since $\mathsf{O}_e A \wedge \neg\mathsf{O}_a A \in \Omega$, this means that we can derive $\mathsf{O}_a A$ from $\mathsf{O}_e A$ by RC on the condition $\dagger(A)$.

Where $A \notin \mathcal{W}^l$, the application of RC is slightly more involving. Suppose, for instance, that $\mathsf{O}_e(p \vee q)$. Then, as above, $\mathsf{O}_e(p \vee q) \vdash_{\mathbf{P2}} \mathsf{O}_a(p \vee q) \vee (\mathsf{O}_e(p \vee q) \wedge \neg\mathsf{O}_a(p \vee q)$. By $\mathbf{CL}$, $\mathsf{O}_e(p \vee q) \vdash_{\mathbf{P2}} \mathsf{O}_a(p \vee q) \vee ((\mathsf{O}_e(p \vee q) \wedge \neg\mathsf{O}_a(p \vee q) \vee (\mathsf{O}_e p \wedge \neg\mathsf{O}_a p) \vee (\mathsf{O}_e q \vee \neg\mathsf{O}_a q))$. In other words, $\mathsf{O}_e(p \vee q) \vdash_{\mathbf{P2}} \mathsf{O}_a(p \vee q) \vee \sharp(p \vee q)$. Hence $\mathsf{O}_a(p \vee q)$ is derivable from $\mathsf{O}_e(p \vee q)$ by means of RC on the condition $\{\sharp(p \vee q)\}$

Let us further illustrate the workings of $\mathbf{P2.2^x}$ by means of some examples.

Example 15. Suppose that Johnson faces the following three obligations:

O1 he ought to pay taxes, and fight in the army or perform alternative service to his country — $\mathsf{O}_e(t \wedge (f \vee s))$
O2 he ought not to pay taxes and not fight in the army — $\mathsf{O}_e(\neg t \wedge \neg f)$
O3 he ought to pay taxes or donate to charity — $\mathsf{O}_e(t \vee c)$

In view of (O1)-(O3), it seems intuitively clear that we want to derive $\mathsf{O}_e s$ and even $\mathsf{O}_a s$ from $\mathsf{O}_e(f \vee s)$ and $\mathsf{O}_e\neg f$, but that we do not want to derive $\mathsf{O}_a t$ or $\mathsf{O}_a\neg t$, since Johnson faces both the obligation to pay taxes and the obligation not to pay taxes. Moreover, it seems dubious to derive an all-things-considered obligation for Johnson to donate to charity, since in view of (O3) it seems that Johnson need only donate to charity if he does not pay taxes.

Let Γ_3 be the set of Johnson's obligations, i.e. $\Gamma_3 = \{\mathsf{O}_e(t \wedge (f \vee s)), \mathsf{O}_e(\neg t \wedge \neg f), \mathsf{O}_e(t \vee c)\}$. Consider the following $\mathbf{P2.2^r}$-proof from Γ_3:

1	$\mathsf{O}_e(t \wedge (f \vee s))$	PREM	$\emptyset$
2	$\mathsf{O}_e(\neg t \wedge \neg f)$	PREM	$\emptyset$
3	$\mathsf{O}_e(t \vee c)$	PREM	$\emptyset$
4	$\mathsf{O}_e(f \vee s)$	1; RU	$\emptyset$
5	$\mathsf{O}_e\neg f$	2; RU	$\emptyset$
6	$\mathsf{O}_e t$	1; RU	$\emptyset$
7	$\mathsf{O}_e\neg t$	2; RU	$\emptyset$
8	$\mathsf{O}_a\neg f$	5; RC	$\{\sharp(\neg f)\}$
9	$\mathsf{O}_a(f \vee s)$	4; RC	$\{\sharp(f \vee s)\}$
10	$\mathsf{O}_a s$	8,9; RU	$\{\sharp(\neg f), \sharp(f \vee s)\}$
11	$\mathsf{O}_a t$	6; RC	$\{\sharp(t)\}\checkmark^{16}$
12	$\mathsf{O}_a\neg t$	7; RC	$\{\sharp(\neg t)\}\checkmark^{15}$
13	$\mathsf{O}_a(t \vee c)$	3; RC	$\{\sharp(t \vee c)\}\checkmark^{17}$
14	$\mathsf{O}_a c$	12,13; RU	$\{\sharp(\neg t), \sharp(t \vee c)\}\checkmark^{15}$
15	$\sharp(\neg t)$	6,7; RU	$\emptyset$
16	$\sharp(t)$	6,7; RU	$\emptyset$
17	$\sharp(t \vee c)$	16; RU	$\emptyset$

As desired, $\Gamma_3 \vdash_{\mathbf{P2.2^r}} \mathsf{O}_a\neg f$ and $\Gamma_3 \vdash_{\mathbf{P2.2^r}} \mathsf{O}_a s$, yet $\Gamma_3 \nvdash_{\mathbf{P2.2^r}} \mathsf{O}_a t$, $\Gamma_3 \nvdash_{\mathbf{P2.2^r}} \mathsf{O}_a\neg t$ and $\Gamma_3 \nvdash_{\mathbf{P2.2^r}} \mathsf{O}_a c$.

As a further illustration, consider the following $\mathbf{P2.2^x}$-proof for the premise set Γ_2 from Section 5.2.3:

1	$\mathsf{O}_e(p \vee q)$	PREM	$\emptyset$
2	$\mathsf{O}_e(r \vee s)$	PREM	$\emptyset$
3	$\neg\mathsf{O}_a((p \vee q) \wedge (r \vee s))$	PREM	$\emptyset$
4	$\mathsf{O}_e t$	PREM	$\emptyset$
5	$\mathsf{O}_a(p \vee q)$	1; RC	$\{\sharp(p \vee q)\}\checkmark^{8}$
6	$\mathsf{O}_a(r \vee s)$	2; RC	$\{\sharp(r \vee s)\}\checkmark^{8}$
7	$\mathsf{O}_a t$	4; RC	$\{\sharp(t)\}$
8	$\sharp(p \vee q) \vee \sharp(r \vee s)$	1-3; RU	$\emptyset$
9	$\sharp(p \vee q \vee \neg t) \vee \sharp(r \vee s \vee \neg t)$	8; RU	$\emptyset$

The formula derived at line 8 abbreviates the long *Dab*-formula $(\mathsf{O}_e(p \vee q) \wedge \neg\mathsf{O}_a(p \vee q)) \vee (\mathsf{O}_e p \wedge \neg\mathsf{O}_a p) \vee (\mathsf{O}_e q \wedge \neg\mathsf{O}_a q) \vee (\mathsf{O}_e(r \vee s) \wedge \neg\mathsf{O}_a(r \vee s)) \vee (\mathsf{O}_e r \wedge \neg\mathsf{O}_a r) \vee (\mathsf{O}_e s \wedge \neg\mathsf{O}_a s)$. To see how this formula follows from the premises, note that $\neg\mathsf{O}_a((p \vee q) \wedge (r \vee s)) \vdash_{\mathbf{P2}} \neg\mathsf{O}_a(p \vee q) \vee \neg\mathsf{O}_a(r \vee s)$ (remember that the O_a-operator has all properties of the obligation operator of **SDL**). Thus, by **CL** and lines 1 and 2, $(\mathsf{O}_e(p \vee q) \wedge \neg\mathsf{O}_a(p \vee q)) \vee (\mathsf{O}_e(r \vee s) \wedge \neg\mathsf{O}_a(r \vee s))$. By **CL** again the longer disjunction $\sharp(p \vee q) \vee \sharp(r \vee s)$ follows immediately.

The formula derived at line 6 of the $\mathbf{P2.1^r}$-proof from Γ_2 in Section 5.2.3 is of course still **P2**-derivable from Γ_2, but it no longer constitutes a minimal *Dab*-formula in a $\mathbf{P2.2^x}$-proof from Γ_2 due to the entirely different definition of the set of abnormalities. Although the $\mathbf{P2.2^x}$-abnormality $\sharp(p \vee q \vee \neg t) \vee \sharp(r \vee s \vee \neg t) \vee \sharp(p) \vee \sharp(q) \vee \sharp(r) \vee \sharp(s) \vee \sharp(t) \vee \sharp(\neg t)$ is derivable from the premise set, it is no longer minimal in view of the formula derived at line 9. Hence $\Gamma_2 \vdash_{\mathbf{P2.2^x}} \mathsf{O}_a t$ as desired.

5.2.5 $\mathbf{P2.2^x}$ and incompatible obligations

In the language $\mathcal{W}^{\mathbf{P2}}$, incompatible obligations can be formalized in one of two ways, depending on the type of 'incompatibility' that is at stake. In what follows, we distinguish between two such types: incompatibility due to prohibition, and physical incompatibility.

5.2.5.1 Incompatibility due to prohibition

In this type of conflict, a number of propositions $A_1, \ldots, A_n$ is mandatory, whereas their conjunction is forbidden. A_1, ..., A_n can be jointly fulfilled, but there is an additional obligation not to fulfill all of them. Consider the following simple example of this type.

Example 16. Bob, at different moments in time, promised his two best friends, John and Peter, to invite them to his birthday party. However, he also promised his girlfriend not to invite them both. (John and Peter are known to quarrel over almost anything and Bob's girlfriend is afraid that this may put a damper on the party).

As there is no reason in this case to prefer one obligation over the other, we formalize all obligations involved as O_e-obligations:

(1) Bob has an obligation to invite John — $\mathsf{O}_e j$
(2) Bob has an obligation to invite Peter — $\mathsf{O}_e p$
(3) Bob has an obligation not to invite both Peter and John — $\mathsf{O}_e \neg(j \wedge p)$

Let $\Gamma_4 = \{\mathsf{O}_e j, \mathsf{O}_e p, \mathsf{O}_e \neg(j \wedge p)\}$, and consider the following $\mathbf{P2.2^x}$-proof from Γ_4:

1	$\mathsf{O}_e j$	PREM	$\varnothing$
2	$\mathsf{O}_e p$	PREM	$\varnothing$
3	$\mathsf{O}_e \neg(j \wedge p)$	PREM	$\varnothing$
4	$\mathsf{O}_a j$	1; RC	$\{\sharp(j)\}\checkmark^{10}$
5	$\mathsf{O}_a p$	2; RC	$\{\sharp(p)\}\checkmark^{10}$
6	$\mathsf{O}_a \neg(j \wedge p)$	3; RC	$\{\sharp(\neg j \vee \neg p)\}\checkmark^{16}$
7	$\mathsf{O}_a(j \vee p)$	4; RU	$\{\sharp(j)\}\checkmark^{10}$
8	$\mathsf{O}_a(j \vee p)$	5; RU	$\{\sharp(p)\}\checkmark^{10}$
9	$\mathsf{O}_a(j \vee p)$	1; RC	$\{\sharp(j \vee p)\}\checkmark^{13}$
10	$\sharp(j) \vee \sharp(p)$	1-3; RU	$\varnothing$
11	$\sharp(\neg j) \vee \sharp(p)$	1-3; RU	$\varnothing$
12	$\sharp(j) \vee \sharp(\neg p)$	1-3; RU	$\varnothing$
13	$\sharp(j \vee p)$	10; RU	$\varnothing$

14	$\sharp(\neg j \vee p)$	1-3; RU	$\varnothing$
15	$\sharp(j \vee \neg p)$	1-3; RU	$\varnothing$
16	$\sharp(\neg j \vee \neg p)$	1-3; RU	$\varnothing$

If the above proof is a $\mathbf{P2.2^r}$-proof, then $U_{16}(\Gamma_4) = \{\sharp(j), \sharp(p), \sharp(\neg j), \sharp(\neg p), \sharp(j \vee p), \sharp(\neg j \vee p), \sharp(j \vee \neg p), \sharp(\neg j \vee \neg p)\}$. Since none of these *Dab*-consequences of Γ_4 can be shortened in any way, the abnormalities in $U_{16}(\Gamma_4)$ are all members of $U(\Gamma_4)$. Thus, lines 4-9 remain marked in any extension of the proof, and neither $\mathsf{O}_a j$, $\mathsf{O}_a p$, nor $\mathsf{O}_a \neg(j \wedge p)$ is $\mathbf{P2.2^r}$-derivable from Γ_4.

If the above proof is a $\mathbf{P2.2^m}$-proof, then $\Phi_{16}(\Gamma_4) = \{\varphi_1, \varphi_2, \varphi_3\}$, where:

$$\varphi_1 = \{\sharp(j), \sharp(\neg j), \sharp(j \vee p), \sharp(\neg j \vee p), \sharp(j \vee \neg p), \sharp(\neg j \vee \neg p)\}$$
$$\varphi_2 = \{\sharp(p), \sharp(\neg p), \sharp(j \vee p), \sharp(\neg j \vee p), \sharp(j \vee \neg p), \sharp(\neg j \vee \neg p)\}$$
$$\varphi_3 = \{\sharp(j), \sharp(p), \sharp(j \vee p), \sharp(\neg j \vee p), \sharp(j \vee \neg p), \sharp(\neg j \vee \neg p)\}$$

By the marking definition for minimal abnormality, lines 4-9 are marked at stage 16 of the proof. In view of the final derivability criterion, $\mathsf{O}_a j$, $\mathsf{O}_a p$, nor $\mathsf{O}_a \neg(j \wedge p)$ is $\mathbf{P2.2^m}$-derivable from Γ_4. Thus, $\mathbf{P2.2^r}$ and $\mathbf{P2.2^m}$ deliver the same results for Γ_4.

5.2.5.2 Physical incompatibility

In this second type of conflict, the joint fulfillment of a certain series of obligations is not merely forbidden; it is simply impossible to fulfill them all. As an example, one may think of a typical Buridan's ass dilemma, e.g.

Example 17. Imagine a situation where two identical twins are drowning some distance apart from each other, and the situation is such that you can save either of them, but you cannot save both.

Example 18. Suppose that someone, Charlotte, ought to visit her daughter Abby at a certain time and in preparation for that, notify her she is coming. But it could also be that Charlotte ought also to visit her daughter Beth at that same time and notify her she is coming. However, since Abby and Beth live on opposite sides of the country, it is impossible for Charlotte to visit both daughters at that time ([69, p. 468], [95, p. 581]).

Given the language of **P2** – which lacks alethic operators for representing (im)possibility – there are different ways to formalize incompatible obligations of this second type. We shall concentrate on two of them. A formalization that immediately comes to mind is to express the impossibility to fulfill a certain number of obligations by the universal obligation not to fulfill them all. This would give us the following formalization in the drowning twin case:

(1) I have an obligation to save the first twin — $\mathsf{O}_e t_1$
(2) I have an obligation to save the second twin — $\mathsf{O}_e t_2$
(3) I have the *universal* obligation not to save both — $\mathsf{O}_a \neg(t_1 \wedge t_2)$

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

However, there are several objections possible. The first is that it is too strong since, given (1)–(3), one is able to infer that one has the obligation *not* to save the first twin and also the obligation *not* to save the second twin:

$$\mathsf{O}_e t_1, \mathsf{O}_e t_2, \mathsf{O}_a \neg(t_1 \wedge t_2) \vdash_{\mathbf{P2}} \mathsf{O}_e \neg t_1 \tag{5.29}$$

$$\mathsf{O}_e t_1, \mathsf{O}_e t_2, \mathsf{O}_a \neg(t_1 \wedge t_2) \vdash_{\mathbf{P2}} \mathsf{O}_e \neg t_2 \tag{5.30}$$

The second objection concerns the notion of a 'deontically perfect world'. The above formalization leads to a very strong restriction on what counts as a deontically perfect alternative for our world. One not only has to assume that a deontically perfect world has at least the same natural laws as our world (which is a reasonable requirement), but also that its past history is exactly as our world's history *up to the point* where at least one of the twins is actually drowning.

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

In view of this, we favor a weaker formalization of the twin example: we only require that it is not a universal obligation to save both. Thus, instead of (3), we obtain

(4) I do *not* have the universal obligation to save both — $\neg\mathsf{O}_a(t_1 \wedge t_2)$

This formalization has several advantages. One is that the link between the two incompatible obligations is preserved: there is no reduction to a series of direct conflicts (i.e. conflicts of the form $\mathsf{O}_e A \wedge \mathsf{O}_e \neg A$), since $\mathsf{O}_e \neg t_1$ and $\mathsf{O}_e \neg t_2$ are not **P2**-derivable from (1), (2) and (4). As we shall see below, this allows us to follow different 'strategies' when dealing with incompatible obligations of the second type. It also nicely agrees with a certain interpretation of the '*ought* implies *can*' principle. An obligation that is impossible to fulfill should not be a universal obligation, which is captured by (4).

Let now $\Gamma_5 = \{\mathsf{O}_e t_1, \mathsf{O}_e t_2, \neg\mathsf{O}_a(t_1 \wedge t_2)\}$, and consider the following $\mathbf{P2.2^m}$-proof from Γ_5:

1	$\mathsf{O}_e t_1$	PREM	$\varnothing$
2	$\mathsf{O}_e t_2$	PREM	$\varnothing$
3	$\neg\mathsf{O}_a(t_1 \wedge t_2)$	PREM	$\varnothing$
4	$\mathsf{O}_a t_1$	1; RC	$\{\sharp(t_1)\}\checkmark^6$
5	$\mathsf{O}_a t_2$	2; RC	$\{\sharp(t_2)\}\checkmark^6$
6	$\sharp(t_1) \vee \sharp(t_2)$	1-3; RU	$\varnothing$
7	$\mathsf{O}_a(t_1 \vee t_2)$	4; RU	$\{\sharp(t_1)\}$
8	$\mathsf{O}_a(t_1 \vee t_2)$	5; RU	$\{\sharp(t_2)\}$

The disjunctive universal obligation $\mathsf{O}_a(t_1 \vee t_2)$ is derivable on the condition $\{\sharp(t_1)\}$ and on the condition $\{\sharp(t_2)\}$. According to the reliability strategy, both of these conditions are considered unreliable. This causes the marking of lines 7 and 8 (in view of line 6) in a $\mathbf{P2.2^r}$-proof.

However, not so in a $\mathbf{P2.2^m}$-proof. According to the minimal abnormality strategy, we need not assume that *both* obligations in the minimal *Dab*-formula on line 6 behave abnormally. As only one of the disjuncts has to be true (in order for the premises to be true), we can assume either one of the two obligations to behave normally. This means that if, on the one hand, the formula $\sharp(t_1)$ in the condition of line 7 were true, then we can still assume the formula in the condition of line 8 to be false. This in turn means that we can still take $\mathsf{O}_a(t_1 \vee t_2)$ to be a $\mathbf{P2.2^m}$-consequence of our premises. If, on the other hand, the formula $\sharp(t_2)$ in the condition of line 8 were true, then we can still assume the formula in the condition of line 7 to be false. Again, we can take $\mathsf{O}_a(t_1 \vee t_2)$ to be a $\mathbf{P2.2^m}$-consequence of our premises. So, whichever disjunct of the *Dab*-formula on line 6 turns out to be true, we can still take $\mathsf{O}_a(t_1 \vee t_2)$ to be a $\mathbf{P2.2^m}$-consequence of our premises.

In our example, this outcome is a very desirable one: whichever one of the twins we eventually decide not to save, we will face an all-things-considered obligation to save the other one. Hence even though we cannot save both twins, we still face the obligation to save at least one of them.[3] However, in other situations this outcome may not be as desirable. Consider the following formalization of Example 18:

(1) Charlotte has an obligation to visit Abby — $\mathsf{O}_e a$
(2) Charlotte has an obligation to visit Beth — $\mathsf{O}_e b$
(3) Charlotte cannot visit both Abby and Beth — $\neg\mathsf{O}_a(a \wedge b)$

For a more tragic effect, we might add that the reason for Charlotte's visit is the wedding of Abby and Beth respectively. The dates of the weddings are fixed, and as things are Charlotte cannot attend both weddings. Analogously to the drowning twins example, we can derive the disjunctive obligation $\mathsf{O}_a(a \vee b)$ by means of the minimal abnormality strategy, but not by means of reliability.

In this case, do we really want to derive Charlotte's obligation to either visit Abby's wedding or visit Beth's wedding? There might be good reasons for Charlotte not to visit any of the weddings. For instance, she might want to treat her daughters equally and avoid arguments as to why she visited one wedding instead of the other. As pointed out in Section 4.6, we need not decide the matter here. Instead, we leave it to the intuitions of the reader to decide which adaptive strategy is best suited for modeling these examples.

In conclusion of this aside on the formalization of incompatible obligations, let us recapitulate our two main findings. First, different formalizations are preferable depending on whether the incompatibility arises due to a prohibition or due to the physical structure of the world. Second, depending on the formalization used different adaptive strategies may lead to different conclusions. This is not a drawback of the logic $\mathbf{P2.2^x}$, nor is it a cue for favoring one strategy over the other. Rather, it points to the different rationales that may underly our reasoning in specific situations.

[3] Several authors have argued that, in case of a conflict between two obligations $O_e A$ and $O_e B$, the obligation $\mathsf{O}_a(A \vee B)$ should be derivable. See, for instance, [38], [50], [95].

5.2.6 Further properties of $\mathbf{P2.2^x}$

Due to its definition within the standard format for adaptive logics, $\mathbf{P2.2^x}$ automatically inherits all properties discussed in Section 4.7.

Theorem 22. *Where* $\Gamma \subseteq \mathcal{W}^{\mathbf{P2}}$ *and* $A \in \mathcal{W}^{\mathbf{P2}}$*:*
(i) $\Gamma \vdash_{\mathbf{P2.2^x}} A$ *iff* $\Gamma \vDash_{\mathbf{P2.2^x}} A$ *(Soundness & completeness)*
(ii) $\Gamma \subseteq \mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma)$ *(Reflexivity)*
(iii) $\mathrm{Cn}_{\mathbf{P2.2^x}}(\mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma)) = \mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma)$ *(Fixed point/idempotence)*
(iv) *If* $M \in \mathcal{M}_{\mathbf{P2}}(\Gamma) - \mathcal{M}_{\mathbf{P2.2^x}}(\Gamma)$, *then there is a* $M' \in \mathcal{M}_{\mathbf{P2.2^x}}(\Gamma)$ *such that* $Ab(M') \subset Ab(M)$ *(Strong reassurance)*
(v) $\mathrm{Cn}_{\mathbf{P2}}(\mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma)) = \mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma)$ *(LLL-closure)*
(vi) $\mathrm{Cn}_{\mathbf{P2.2^x}}(\mathrm{Cn}_{\mathbf{P2}}(\Gamma)) = \mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma)$ *(LLL-invariance)*
(vii) *If* $\Gamma' \subseteq \mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma)$, *then* $\mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma) = \mathrm{Cn}_{\mathbf{P2.2^x}}(\Gamma \cup \Gamma')$ *(Cautious indifference/Cumulativity)*

The ULL of $\mathbf{P2.2^x}$ is the logic obtained by adding to **P2** the axiom schema (Uσ). Where $\Theta \subset \mathcal{W}^l, \Theta \neq \varnothing, \Theta$ is finite:

$$\neg \bigvee(\sigma(\Theta)) \tag{Uσ}$$

The logic resulting from adding (Uσ) to **P2** is the logic $\mathbf{SDL_{ae}}$, i.e. the logic in which both the O_e- and the O_a-operator behave exactly like the O-operator of **SDL**. Let $\pi(\Gamma)$ be obtained by replacing every occurrence of "O_e" and "O_a" in Γ with "O".

Theorem 23. *Where* $\Gamma \subseteq \mathcal{W}^{\mathbf{P2}}$ *and* $A \in \mathcal{W}^{\mathbf{P2}}$, $\Gamma \vdash_{\mathbf{SDL_{ae}}} A$ *iff* $\pi(\Gamma) \vdash_{\mathbf{SDL}} \pi(A)$.

Proof. Since the O_a-operator of **P2** is characterized exactly like the O-operator of **SDL**, the theorem follows immediately as soon as we can show that the O_e-operator of $\mathbf{SDL_{ae}}$ inherits all properties of the O_a-operator. Thus, we need to show that ($\natural$) holds in $\mathbf{SDL_{ae}}$:

$$\mathsf{O}_e A \equiv \mathsf{O}_a A \tag{$\natural$}$$

Left-Right. Suppose $\mathsf{O}_e A$. By (RM$_e$), $\mathsf{O}_e(A_1 \wedge \ldots \wedge A_n)$, where $A_1 \wedge \ldots \wedge A_n$ is a conjunctive normal form of A. By (RM$_e$) again, $\mathsf{O}_e A_i$ for each $i \in \{1, \ldots, n\}$. By (Uσ), it follows that $\neg(\mathsf{O}_e A_i \wedge \neg \mathsf{O}_a A_i)$ for each $i \in \{1, \ldots, n\}$. By **CL**, $\mathsf{O}_a A_i$ for each $i \in \{1, \ldots, n\}$. Since the O_a-operator is an **SDL**-operator, it follows that $\mathsf{O}_a A$.
Right-Left. Immediate in view of (5.17). □

5.3 A non-adaptive alternative: maximally consistent subsets

In Section 3.2.2.1 we already discussed some proposals made in the literature that reject or restrict (AND). However, the discussion was restricted to monotonic approaches. Here, we pick up the discussion and assess a non-monotonic strategy for restricting the aggregation rule of **SDL**.

The most important alternative to our adaptive approach from this chapter makes use of the notion of maximally consistent subsets of a set of obligations (cfr. Definition 15). The champion of this alternative approach is John Horty.[4] In his [90, 92], Horty defines two consequence relations $\vdash_{\mathbf{F}}$ and $\vdash_{\mathbf{S}}$ for modeling credulous deontic consequence (inspired by Van Fraassen's [58]), respectively skeptical deontic consequence. In [91, 95], he defines these relations for a 'two-faced' account where $\{\mathsf{O}_P A \mid A \in \mathcal{W}\}$ constitutes a set of prima facie obligations from which we try to derive a distinct set $\{\mathsf{O}_A A \mid A \in \mathcal{W}\}$ of all-things-considered obligations. Here, we follow the latter account.

Where Γ is a set of prima facie obligations and $\Gamma^{\mathsf{O}} = \{A \mid \mathsf{O}_P A \in \Gamma\}$, the credulous and skeptical deontic consequence relations are defined as follows:

Definition 16 (Credulous deontic consequence)**.** $\Gamma \vdash_{\mathbf{F}} \mathsf{O}_A A$ iff $A \in \mathrm{Cn}_{\mathbf{CL}}(\Delta)$ for some **CL**-maximally consistent subset Δ of Γ^{O}.

Definition 17 (Skeptical deontic consequence)**.** $\Gamma \vdash_{\mathbf{S}} \mathsf{O}_A A$ iff $A \in \mathrm{Cn}_{\mathbf{CL}}(\Delta)$ for each **CL**-maximally consistent subset Δ of Γ^{O}.

The following example illustrates the difference between the credulous and skeptical consequence relations:

$$\mathsf{O}_P(p \wedge \neg q), \mathsf{O}_P q \vdash_{\mathbf{F}} \mathsf{O}_A p \tag{5.31}$$

$$\mathsf{O}_P(p \wedge \neg q), \mathsf{O}_P q \nvdash_{\mathbf{S}} \mathsf{O}_A p \tag{5.32}$$

The set $\{p \wedge \neg q, q\}$ has two **CL**-maximally consistent subsets, $\{p \wedge \neg q\}$ and $\{q\}$. Since p is a **CL**-consequence of the first of these **CL**-maximally consistent subsets, it follows by Definition 16 that $\mathsf{O}_A p$ is a credulous consequence of the premise set. Since p is not a **CL**-consequence of the second **CL**-maximally consistent subset, it follows by Definition 17 that $\mathsf{O}_A p$ is not a skeptical consequence of the premise set.

An immediate difference between Horty's systems and the logics defined earlier on in this chapter is that Horty only takes into account a set of obligations, i.e. a set of formulas of the form $\mathsf{O}_P A$. His approach is restricted in the sense that premises cannot contain negated obligations (e.g. $\neg\mathsf{O}_P p$) or disjunctions of obligations (e.g. $\mathsf{O}_P p \vee \mathsf{O}_P q$). Moreover, permissions are not explicitly dealt with in Horty's framework.

Apart from this restriction, the main difference between the maximally consistent subset-approach and the adaptive systems defined in this chapter is that during the process of devising the **CL**-maximally consistent subsets of our premises, prima facie obligations are not further analyzed into shorter logical constituents. As a result, the maximally consistent subset-approach often delivers a rather weak consequence set. Consider, for instance, the premise set $\{\mathsf{O}_P(p \wedge q), \mathsf{O}_P(\neg p \wedge r)\}$, respectively the set $\{\mathsf{O}_e(p \wedge q), \mathsf{O}_e(\neg p \wedge r)\}$ of $\mathbf{P2.2^x}$-wffs.

$$\mathsf{O}_P(p \wedge q), \mathsf{O}_P(\neg p \wedge r) \nvdash_{\mathbf{F}} \mathsf{O}_A(q \wedge r) \tag{5.33}$$

$$\mathsf{O}_P(p \wedge q), \mathsf{O}_P(\neg p \wedge r) \nvdash_{\mathbf{S}} \mathsf{O}_A(q \wedge r) \tag{5.34}$$

$$\mathsf{O}_e(p \wedge q), \mathsf{O}_e(\neg p \wedge r) \vdash_{\mathbf{P2.2^x}} \mathsf{O}_a(q \wedge r) \tag{5.35}$$

[4]Horty's account is inspired by Reiter's default logic [150].

There is no **CL**-maximally consistent subset of $\{p \wedge q, \neg p \wedge r\}$ of which $q \wedge r$ is a **CL**-consequence. However, the following **P2.2$^{\mathbf{x}}$**-proof illustrates that $\mathsf{O}_a(q \wedge r)$ is a **P2.2$^{\mathbf{x}}$**-consequence of $\{\mathsf{O}_e(p \wedge q), \mathsf{O}_e(\neg p \wedge r)\}$.

1	$\mathsf{O}_e(p \wedge q)$	PREM	$\varnothing$
2	$\mathsf{O}_e(\neg p \wedge r)$	PREM	$\varnothing$
3	$\mathsf{O}_e p$	1; RU	$\varnothing$
4	$\mathsf{O}_e q$	1; RU	$\varnothing$
5	$\mathsf{O}_e \neg p$	2; RU	$\varnothing$
6	$\mathsf{O}_e r$	2; RU	$\varnothing$
7	$\mathsf{O}_a q$	4; RC	$\{\sharp(q)\}$
8	$\mathsf{O}_a r$	6; RC	$\{\sharp(r)\}$
9	$\mathsf{O}_a(q \wedge r)$	7,8; RU	$\{\sharp(q), \sharp(r)\}$
10	$\sharp(p)$	3,5; RU	$\varnothing$
11	$\sharp(\neg p)$	3,5; RU	$\varnothing$

Since no minimal *Dab*-consequence of $\{\mathsf{O}_e(p \wedge q), \mathsf{O}_e(\neg p \wedge r)\}$ contains $\sharp(q)$ or $\sharp(r)$ as one of its disjuncts, there is no extension of this proof in which line 9 is marked, and for which there is no further extension in which this line is unmarked again. Hence $\mathsf{O}_e(p \wedge q), \mathsf{O}_e(\neg p \wedge r) \vdash_{\mathbf{P2.2^x}} \mathsf{O}_a(q \wedge r)$.

As another illustration of the differences between the logics defined here and the maximally consistent subset-approach, consider Example 15. Translated to the grammar used by Horty, Johnson faces the set of obligations $\Gamma_{3'} = \{\mathsf{O}_P(t \wedge (f \vee s)), \mathsf{O}_P(\neg t \wedge \neg f), \mathsf{O}_P(t \vee c)\}$. $\Gamma^{\mathsf{O}}_{3'}$ gives rise to two **CL**-maximally consistent subsets: the sets $\{t \wedge (f \vee s), t \vee c\}$ and $\{\neg t \wedge \neg f, t \vee c\}$. By Definitions 16 and 17:

$$\Gamma_{3'} \nvdash_{\mathbf{F}} \mathsf{O}_A s \qquad \Gamma_{3'} \nvdash_{\mathbf{S}} \mathsf{O}_A s$$
$$\Gamma_{3'} \vdash_{\mathbf{F}} \mathsf{O}_A \neg f \qquad \Gamma_{3'} \nvdash_{\mathbf{S}} \mathsf{O}_A \neg f$$
$$\Gamma_{3'} \vdash_{\mathbf{F}} \mathsf{O}_A t \qquad \Gamma_{3'} \nvdash_{\mathbf{S}} \mathsf{O}_A t$$
$$\Gamma_{3'} \vdash_{\mathbf{F}} \mathsf{O}_A \neg t \qquad \Gamma_{3'} \nvdash_{\mathbf{S}} \mathsf{O}_A \neg t$$

Neither the skeptical nor the credulous consequence relations allow us to derive the intuitive $\mathsf{O}_A s$. Although the credulous consequence relation does deliver the intuitive $\mathsf{O}_A \neg f$, it also allows us to derive the all-things-considered obligations $\mathsf{O}_A t$ and $\mathsf{O}_A \neg t$.

The maximally consistent subset-approach was also used by Makinson and van der Torre in their input/output (I/O) framework. I/O-logics thus face the same problems as Horty. We discuss the I/O-logics in more detail in Section 6.2.7.2.

Altogether, the maximally consistent subset-approach is suboptimal in its treatment of various toy examples from the literature. Moreover, it is restricted to premise sets containing *only* prima facie obligations.

Chapter 6

Inconsistency-adaptive logics for normative conflicts

✎ Section 6.1 is based on the paper *An Inconsistency-Adaptive Deontic Logic for Normative Conflicts* (Journal of Philosophical Logic, in print) [31], which is co-authored by Christian Straßer and Joke Meheus.

✎ Section 6.2 is based on the paper *Two Adaptive Logics of Norm-Propositions* (Journal of Applied Logic, in print) [29], which is co-authored by Christian Straßer.

✎ I am indebted to Joke Meheus and Christian Straßer for valuable comments on this chapter.

In this chapter, we present two adaptive CTDLs that adopt the strategy of weakening **CL** to a paraconsistent logic (cfr. Section 3.2.2.3). ALs that are built on top of a paraconsistent logic are usually called *inconsistency-adaptive logics.*

In Section 6.1 we define the inconsistency-adaptive deontic logic $\mathbf{DP^x}$. This logic makes use of a paraconsistent negation connective instead of the classical one. As a result, it safely accommodates not only OO- and OP-conflicts, but also contradictory obligations and permissions. $\mathbf{DP^x}$ is especially suited for reasoning with conflicting commands or imperatives, or in other settings in which we may face inconsistent prescriptions.

In Section 6.2 we present the logic $\mathbf{LNP^x}$. $\mathbf{LNP^x}$ is a *semi-paraconsistent* and *semi-paracomplete* deontic logic [125]: outside the scope of its deontic operators, it makes use of classical negation; inside the scope of its deontic operators, it uses a negation connective that is not only paraconsistent but also *paracomplete* (i.e. it invalidates the excluded middle principle). As a result, $\mathbf{LNP^x}$ accommodates OO- and OP-conflicts as well as normative gaps, i.e. propositions that are neither positively permitted nor forbidden nor obligatory. This makes $\mathbf{LNP^x}$ very suitable for reasoning about norm-propositions.

To the best of our knowledge, the logics presented in this chapter are the first non-monotonic paraconsistent deontic logics. Hence the discussion of related approaches in this chapter is rather limited (although in Section 6.2.7 we compare the logic $\mathbf{LNP^x}$ to other logics of norm-propositions presented in the literature).

A discussion of some alternative ways of devising inconsistency-adaptive logics is postponed until Section 7.5.1 in the next chapter.

The systems **DP** and **LNP** are deontic extensions of the paraconsistent logic **CLuNs**$^{\bot}$ and the paraconsistent and paracomplete logic **CLoNs** respectively. The latter systems were devised by Diderik Batens for reasoning in the presence of possibly inconsistent information. In this chapter we will characterize **CLuNs**$^{\bot}$ and **CLoNs** only informally. However, a full formal characterization of these logics is contained in Appendix C.

6.1 Reasoning with contradictory obligations and permissions

In Section 3.2.2.3, we already provided some good reasons for weakening **CL** to a paraconsistent logic when devising CTDLs. In doing so, we can tolerate OO-conflicts, OP-conflicts as well as contradictory obligations and permissions. Moreover, we need not weaken any of the principles (D), (P), (AND) or (RM).

In Section 6.1.1 we present the paraconsistent deontic logic **DP**. **DP** is rather strong for a paraconsistent logic. It validates de Morgan's laws for negation, inheritance and necessitation for the deontic operators, and all of positive **CL** (i.e. **CL** without a negation connective). Nonetheless, it suffers from some weaknesses inherent to many paraconsistent logics (cfr. Section 3.2.2.3). For instance, the intuitive contraposition and disjunctive syllogism rules are invalidated both inside and outside the scope of its deontic operators.

The weaknesses that bother the logic **DP** are overcome by its inconsistency-adaptive extension $\mathbf{DP^x}$, which we present in Section 6.1.2. Like its LLL **DP**, $\mathbf{DP^x}$ safely accommodates OO-conflicts, OP-conflicts as well as contradictory obligations and permissions. Unlike **DP** however, $\mathbf{DP^x}$ allows for the conditional application of all **SDL**-valid inferences.

6.1.1 The logic DP

6.1.1.1 Semantics

DP is a proper extension of the (propositional fragment of the) non-modal paraconsistent logic **CLuNs**$^{\bot}$. The set $\mathcal{W}^{\sim}_{\bot}$ of wffs of **CLuNs**$^{\bot}$ is defined as:

$$\mathcal{W}^{\sim}_{\bot} := \mathcal{W}^a \mid {\sim}\langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \vee \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \wedge \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \supset \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \equiv \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \bot$$

We also define the set $\mathcal{W}^{\sim}_{l} =_{\mathrm{df}} \{A, {\sim}A \mid A \in \mathcal{W}^a\}$ of **CLuNs**$^{\bot}$-literals. **CLuNs**$^{\bot}$ makes use of a paraconsistent negation. In **CL**, both (D~1) and (D~2) are valid for all atomic propositions:

(D~1) If A is true, then ${\sim}A$ is false
(D~2) If A is false, then ${\sim}A$ is true

CLuNs$^{\bot}$ validates only (D~2), thereby allowing for both A and ${\sim}A$ to be true. The **CLuNs**$^{\bot}$-negation is fully characterized by (D~2) and de Morgan's laws. In fact, adding (D~2), the falsum constant and de Morgan's laws to the semantics of full positive **CL** (in the remainder, we abbreviate this fragment by

CL^pos^) is all that is needed in order to obtain the **CLuNs**$^\perp$-semantics. For a full formal characterization of the logic **CLuNs**$^\perp$, see Appendix C.

The set $\mathcal{W}^{\mathbf{DP}}$ of wffs of **DP** is defined as:

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

We also define the set $\mathcal{W}^{\not\sim}$ of wffs of **DP** that are not of the form ${\sim}A$. **DP** is a modal extension of **CLuNs**$^\perp$ that differs semantically from **SDL** only in the characterization of its negation.[1] A **DP**-model is a quadruple $\langle W, w_0, R, v\rangle$, where W is a set of worlds, $w_0 \in W$ is the actual world, R is a serial accessibility relation on W and $v : \mathcal{W}^{\sim}_{l} \times W \to \{0,1\}$ is an assignment function. The valuation $v_M : \mathcal{W}^{\mathbf{DP}} \times W \to \{0,1\}$, associated with the model M, is defined by:

(C$_a$)	where $A \in \mathcal{W}^a, v_M(A,w) = 1$ iff $v(A,w) = 1$
(C~1')	where $A \in \mathcal{W}^a, v_M({\sim}A,w) = 1$ iff $(v_M(A,w) = 0$ or $v({\sim}A,w) = 1)$
(C∨)	$v_M(A \vee B, w) = 1$ iff $(v_M(A,w) = 1$ or $v_M(B,w) = 1)$
(C∧)	$v_M(A \wedge B, w) = 1$ iff $v_M(A,w) = v_M(B,w) = 1$
(C⊃)	$v_M(A \supset B, w) = 1$ iff $(v_M(A,w) = 0$ or $v_M(B,w) = 1)$
(C≡)	$v_M(A \equiv B, w) = 1$ iff $v_M(A,w) = v_M(B,w)$
(CO)	$v_M(\mathsf{O}A, w) = 1$ iff $v_M(A,w') = 1$ for every w' such that Rww'
(C~~)	$v_M({\sim}{\sim}A, w) = v_M(A,w)$
(C~⊃)	$v_M({\sim}(A \supset B), w) = v_M(A \wedge {\sim}B, w)$
(C~∧)	$v_M({\sim}(A \wedge B), w) = v_M({\sim}A \vee {\sim}B, w)$
(C~∨)	$v_M({\sim}(A \vee B), w) = v_M({\sim}A \wedge {\sim}B, w)$
(C~≡)	$v_M({\sim}(A \equiv B)) = v_M((A \vee B) \wedge ({\sim}A \vee {\sim}B))$
(C~O)	where $A \in \mathcal{W}^{\sim}_{l} \cup \mathcal{W}^{\not\sim}$, $v_M({\sim}\mathsf{O}A, w) = 1$ iff there is a w' such that Rww' and $v_M({\sim}A, w') = 1$
(C~~′)	$v_M({\sim}\mathsf{O}{\sim}{\sim}A, w) = v_M({\sim}\mathsf{O}A, w)$
(C~⊃′)	$v_M({\sim}\mathsf{O}{\sim}(A \supset B), w) = v_M({\sim}\mathsf{O}(A \wedge {\sim}B), w)$
(C~∧′)	$v_M({\sim}\mathsf{O}{\sim}(A \wedge B), w) = v_M({\sim}\mathsf{O}({\sim}A \vee {\sim}B), w)$
(C~∨′)	$v_M({\sim}\mathsf{O}{\sim}(A \vee B), w) = v_M({\sim}\mathsf{O}({\sim}A \wedge {\sim}B), w)$
(C~≡′)	$v_M({\sim}\mathsf{O}{\sim}(A \equiv B)) = v_M({\sim}\mathsf{O}((A \vee B) \wedge ({\sim}A \vee {\sim}B)))$
(C⊥)	$v_M(\perp, w) = 0$

The permission operator P is defined by $\mathsf{P}A =_{\text{df}} {\sim}\mathsf{O}{\sim}A$. Clauses (C$_a$) and (C∨)-(CO) are as in the usual Kripke semantics for **SDL**. Where $A \in \mathcal{W}^a$, (C~1') makes it possible for both A and ${\sim}A$ to be true at a world. (C~~)-(C~≡) guarantee that de Morgan's laws are valid outside the scope of a deontic operator. (C~O)-(C~≡′) guarantee that de Morgan's laws are valid inside the scope of a deontic operator, and that O and P are interdefinable as in **SDL**, e.g. ${\sim}\mathsf{O}A \equiv \mathsf{P}{\sim}A$.

A **DP**-model $M = \langle W, w_0, R, v\rangle$ verifies A, $M \Vdash A$, iff $v_M(A, w_0) = 1$.

6.1.1.2 Syntactic characterization of DP

Syntactically, **CLuNs**$^\perp$ is obtained by adding to **CL^pos^** the following axiom schemas:

(A~1) $(A \supset {\sim}A) \supset {\sim}A$

[1] For some other modal extensions of **CLuNs**, see [111].

(A$\sim\sim$)	$\sim\sim A \equiv A$
(A$\sim\supset$)	$\sim(A \supset B) \equiv (A \wedge \sim B)$
(A$\sim\wedge$)	$\sim(A \wedge B) \equiv (\sim A \vee \sim B)$
(A$\sim\vee$)	$\sim(A \vee B) \equiv (\sim A \wedge \sim B)$
(A$\sim\equiv$)	$\sim(A \equiv B) \equiv ((A \vee B) \wedge (\sim A \vee \sim B))$
(A$\bot$1)	$\bot \supset A$

DP is fully axiomatized by adding to **CLuNs**$^{\bot}$ the principles (K), (NEC), and the following axiom schemas and rules, all of which are also valid in **SDL**:

(A$\sim\sim'$)	$\sim\mathsf{O}\sim\sim A \equiv \sim\mathsf{O}A$
(A$\sim\supset''$)	$\sim\mathsf{O}(A \wedge \sim B) \supset \sim\mathsf{O}\sim(A \supset B)$
(A$\sim\equiv''$)	$\sim\mathsf{O}((A \vee B) \wedge (\sim A \vee \sim B)) \supset \sim\mathsf{O}\sim(A \equiv B)$
(A$\bot$2)	$\sim\mathsf{O}\sim\bot \supset A$
(CONS$^{\sim}$)	$\mathsf{O}A \supset \sim\mathsf{O}\sim A$
(KP)	$\mathsf{O}(A \supset B) \supset (\sim\mathsf{O}\sim A \supset \sim\mathsf{O}\sim B)$
(OD)	$\mathsf{O}(A \vee B) \supset (\mathsf{O}A \vee \sim\mathsf{O}\sim B)$
(PD)	$\sim\mathsf{O}\sim(A \vee B) \supset (\sim\mathsf{O}\sim A \vee \sim\mathsf{O}\sim B)$

(A$\sim\sim$), (A$\sim\supset''$) and (A$\sim\equiv''$) are necessary in order to ensure that de Morgan's laws hold inside and outside the scope of a deontic operator. Similarly, (A$\bot$2) ensures that (A$\bot$1) holds inside the scope of a permission. (CONS$^{\sim}$) is **DP**-equivalent to the principle (D), $\mathsf{O}A \supset \mathsf{P}A$. (KP), (OD), and (PD) further characterize permissions in **DP**.

6.1.1.3 Meta-theory of DP

Theorem 24. **DP** *is reflexive, transitive and monotonic.*

Theorem 25. **DP** *is compact (if* $\Gamma \vdash_{\mathbf{DP}} A$ *then* $\Gamma' \vdash_{\mathbf{DP}} A$ *for some finite* $\Gamma' \subseteq \Gamma$*).*

Theorem 26. *If* $\Gamma \vdash_{\mathbf{DP}} B$ *and* $A \in \Gamma$*, then* $\Gamma - \{A\} \vdash_{\mathbf{DP}} A \supset B$ *(Generalized Deduction Theorem for* **DP***).*

For the proofs of the reflexivity, transitivity, monotonicity, compactness of **CLuNs** and the validity of the Generalized Deduction Theorem for **CLuNs**, see [20]. Since **DP** adds only some standard axioms and rules to **CLuNs**, the proofs of Theorems 24-26 are straightforward.

Theorem 27. *If* $\Gamma \vdash_{\mathbf{DP}} A$*, then* $\Gamma \vDash_{\mathbf{DP}} A$*. (Soundness of* **DP***)*

Theorem 28. *If* $\Gamma \vDash_{\mathbf{DP}} A$*, then* $\Gamma \vdash_{\mathbf{DP}} A$*. (Strong Completeness of* **DP***)*

Proofs for Theorem 27 and Theorem 28 are contained in Appendix E.

6.1.1.4 Further properties and discussion

It is easy to see (and proven in Lemma 1 in Appendix E) that all instances of the following axiom schemas are valid in **DP**:

(A$\sim\supset'$) $\quad \sim\mathsf{O}\sim(A \supset B) \equiv \sim\mathsf{O}(A \wedge \sim B)$

(A$\sim\wedge'$) $\sim\mathsf{O}\sim(A \wedge B) \equiv \sim\mathsf{O}(\sim A \vee \sim B)$
(A$\sim\vee'$) $\sim\mathsf{O}\sim(A \vee B) \equiv \sim\mathsf{O}(\sim A \wedge \sim B)$
(A$\sim\equiv'$) $\sim\mathsf{O}\sim(A \equiv B) \equiv \sim\mathsf{O}((A \vee B) \wedge (\sim A \vee \sim B))$

All of the following inferences are **DP**-valid:

$$(\mathsf{O}A \wedge \mathsf{O}B) \vdash_{\mathbf{DP}} \mathsf{O}(A \wedge B) \tag{6.1}$$
$$(\mathsf{O}A \wedge \mathsf{P}B) \vdash_{\mathbf{DP}} \mathsf{P}(A \wedge B) \tag{6.2}$$
$$\vdash_{\mathbf{DP}} \mathsf{P}(A \supset A) \tag{6.3}$$
$$\mathsf{P}A \vdash_{\mathbf{DP}} \sim\mathsf{O}\sim A \tag{6.4}$$
$$\sim\mathsf{P}A \vdash_{\mathbf{DP}} \mathsf{O}\sim A \tag{6.5}$$
$$\mathsf{O}A \vdash_{\mathbf{DP}} \sim\mathsf{P}\sim A \tag{6.6}$$
$$\sim\mathsf{O}A \vdash_{\mathbf{DP}} \mathsf{P}\sim A \tag{6.7}$$

(6.1) and (6.2) are shown to hold in Fact 5 in Appendix E. It is safely left to the reader to check that (6.3)-(6.7) are **DP**-valid.

DP is fully conflict-tolerant: (OO-DEX) and (OP-DEX) are invalidated, and contradictory obligations and permissions do not lead to explosion either:

$$\mathsf{O}A \wedge \mathsf{O}\sim A \nvdash_{\mathbf{DP}} \mathsf{O}B \tag{6.8}$$
$$\mathsf{O}A \wedge \mathsf{P}\sim A \nvdash_{\mathbf{DP}} \mathsf{O}B \tag{6.9}$$
$$\mathsf{O}A \wedge \sim\mathsf{O}A \nvdash_{\mathbf{DP}} \mathsf{O}B \tag{6.10}$$
$$\mathsf{P}A \wedge \sim\mathsf{P}A \nvdash_{\mathbf{DP}} \mathsf{O}B \tag{6.11}$$

Moreover, **DP** verifies all of (D), (P), (AND), (K), and (NEC). However, as compared to **SDL**, **DP** is still rather weak. Although modus ponens holds, intuitive inferences such as disjunctive syllogism and contraposition are invalid in **DP**:

$$A, \sim A \vee B \nvdash_{\mathbf{DP}} B \tag{6.12}$$
$$A \supset B, \sim B \nvdash_{\mathbf{DP}} \sim A \tag{6.13}$$

Like the paraconsistent systems discussed in Section 3.2.2.3, **DP** also invalidates the deontic disjunctive syllogism principle (from $\mathsf{O}(A \vee B)$ and $\mathsf{O}\sim A$ to derive $\mathsf{O}B$). Consequently, **DP** cannot account for Horty's Smith example (i.e. $\mathsf{O}(f \vee s), \mathsf{O}\sim f \nvdash_{\mathbf{DP}} \mathsf{O}s$).

These weaknesses make it hard for **DP** to model our everyday normative reasoning. We will now address and solve this problem by extending **DP** within the adaptive logics framework, and by demonstrating how the resulting extensions $\mathbf{DP^r}$ and $\mathbf{DP^m}$ resolve the inferential weaknesses that bother the basic logic **DP**.

6.1.2 The logics $\mathbf{DP^r}$ and $\mathbf{DP^m}$

6.1.2.1 Definition and illustration

Where $\mathcal{W}^{\sim}_{\mathsf{O}} = \{\mathsf{O}A \mid A \in \mathcal{W}^{\sim}_{l}\}$, $\mathbf{DP^x}$ is defined as a triple:

(1) Lower limit logic: **DP**.[2]
(2) Set of abnormalities: $\Omega = \{A \wedge {\sim}A \mid A \in \mathcal{W}^a \cup \mathcal{W}^{\sim}_{\mathsf{O}}\}$.
(3) Adaptive strategy: $x \in \{r, m\}$.

Since the lower limit logic of $\mathbf{DP^x}$ is **DP**, we know that $Cn_{\mathbf{DP}}(\Gamma) \subseteq Cn_{\mathbf{DP^x}}(\Gamma)$ for any premise set Γ.

Any inconsistency in the language is **DP**-equivalent to a member of Ω, or to a disjunction of members of Ω. For instance, $(\mathsf{P}q \wedge {\sim}\mathsf{P}q) \equiv ({\sim}\mathsf{O}{\sim}q \wedge \mathsf{O}{\sim}q)$, and $((p \wedge q) \wedge {\sim}(p \wedge q)) \equiv ((p \wedge {\sim}p) \vee (q \wedge {\sim}q))$. Ω is constructed in such a way that every normative conflict gives rise to an abnormality in $\mathbf{DP^x}$ in view of the LLL. For instance, from an OO-conflict $\mathsf{O}q \wedge \mathsf{O}{\sim}q$, the abnormality $\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q$ is **DP**-derivable. Similarly, $\mathsf{O}p \wedge \mathsf{P}{\sim}p \vdash_{\mathbf{DP}} \mathsf{O}p \wedge {\sim}\mathsf{O}p$. Moreover, complex normative conflicts are always reducible to a disjunction of abnormalities, e.g. $\mathsf{O}(p \vee q) \wedge \mathsf{O}({\sim}p \wedge {\sim}q) \vdash_{\mathbf{DP}} (\mathsf{O}{\sim}p \wedge {\sim}\mathsf{O}{\sim}p) \vee (\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q)$, $\mathsf{O}(r \wedge s) \wedge {\sim}\mathsf{O}(r \wedge s) \vdash_{\mathbf{DP}} (\mathsf{O}r \wedge {\sim}\mathsf{O}r) \vee (\mathsf{O}s \wedge {\sim}\mathsf{O}s)$.

$\mathbf{DP^x}$ is an inconsistency-adaptive logic, i.e. a logic that interprets (possibly) inconsistent sets of premises 'as consistently as possible'. $\mathbf{DP^x}$ is the first inconsistency-adaptive logic that aims to explicate normative reasoning. Other inconsistency-adaptive logics have been presented, for instance, in [19, 22, 111].

We now illustrate the workings of the logic $\mathbf{DP^x}$ by means of an example.

Consider the set of formulas $\Gamma = \{\mathsf{O}(p \vee q), \mathsf{O}({\sim}r \vee {\sim}s), \mathsf{O}{\sim}q, \mathsf{O}r, \mathsf{O}t \supset {\sim}\mathsf{O}p, \mathsf{P}(t \wedge s)\}$. We start a $\mathbf{DP^x}$-proof from Γ by introducing the premises:

1	$\mathsf{O}(p \vee q)$	PREM	$\varnothing$
2	$\mathsf{O}({\sim}r \vee {\sim}s)$	PREM	$\varnothing$
3	$\mathsf{O}{\sim}q$	PREM	$\varnothing$
4	$\mathsf{O}r$	PREM	$\varnothing$
5	$\mathsf{O}t \supset {\sim}\mathsf{O}p$	PREM	$\varnothing$
6	$\mathsf{P}(t \wedge s)$	PREM	$\varnothing$

From the formulas on lines 1 and 3 in the proof, the formula $\mathsf{O}p \vee (\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q)$ is **DP**-derivable.[3] Since $\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q$ is a *Dab*-formula, we can introduce the following line using the conditional rule RC:

7	$\mathsf{O}p$	1,3; RC	$\{\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q\}$

At stage 7 of the proof, we have derived the obligation $\mathsf{O}p$ on the assumption that the abnormality $\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q$ is false. In a similar fashion, we can apply RC to lines 2 and 4 as follows:

8	$\mathsf{O}{\sim}s$	2,4; RC	$\{\mathsf{O}r \wedge {\sim}\mathsf{O}r\}$

---

[2]In Chapter 4, footnote 1 we stated that the lower limit logic of an adaptive logic in standard format should contain **CL**. As defined here, the **DP**-connectives $\vee, \wedge, \supset$, and $\equiv$ already behave classically. Moreover, **DP** features a classical negation in the sense that the classical negation $\neg A$ of a formula A is definable as $\neg A =_{\text{df}} A \supset \bot$. The implicit definability of all **CL**-connectives in **DP** is sufficient for $\mathbf{DP^r}$ to be in the standard format.

We formalize negations in a $\mathbf{DP^r}$-premise set by means of ${\sim}$ and not by means of $\neg$. Otherwise normative conflicts would be rendered trivial all over again.

[3]From $\mathsf{O}(p \vee q)$ it follows by (OD) that $\mathsf{O}p \vee {\sim}\mathsf{O}{\sim}q$. From $\mathsf{O}p \vee {\sim}\mathsf{O}{\sim}q$ and $\mathsf{O}{\sim}q$ it follows by $\mathbf{CL^{pos}}$ that $\mathsf{O}p \vee (\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q)$.

Since (AND) is valid in **DP**, we can apply it unconditionally by means of RU:

9	$\mathsf{O}(p \wedge {\sim}s)$	7,8; RU	$\{\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q, \mathsf{O}r \wedge {\sim}\mathsf{O}r\}$

Consider now the following extension of the proof (we repeat the proof from line 7 on):

7	$\mathsf{O}p$	1,3; RC	$\{\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q\}$
8	$\mathsf{O}{\sim}s$	2,4; RC	$\{\mathsf{O}r \wedge {\sim}\mathsf{O}r\}\checkmark^{11}$
9	$\mathsf{O}(p \wedge {\sim}s)$	7,8; RU	$\{\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q, \mathsf{O}r \wedge {\sim}\mathsf{O}r\}\checkmark^{11}$
10	$\mathsf{P}s$	6; RU	$\varnothing$
11	$(\mathsf{O}r \wedge {\sim}\mathsf{O}r) \vee (\mathsf{O}{\sim}s \wedge {\sim}\mathsf{O}{\sim}s)$	2,4,10; RU	$\varnothing$

In order to infer the formulas $\mathsf{O}{\sim}s$ and $\mathsf{O}(p \wedge {\sim}s)$ at lines 8 and 9, we have relied on the consistent behavior of $\mathsf{O}r$, viz. on the falsity of $\mathsf{O}r \wedge {\sim}\mathsf{O}r$. However, at line 11 it has become clear that either $\mathsf{O}r$ behaves inconsistently, or $\mathsf{O}{\sim}s$ does.[4] In view of this new information, it is appropriate to withdraw our conclusions drawn at lines 8 and 9.

We can now further extend the proof as follows:

12	$\mathsf{O}p \supset {\sim}\mathsf{O}t$	5; RC	$\{\mathsf{O}p \wedge {\sim}\mathsf{O}p\}$
13	${\sim}\mathsf{O}t$	7,12; RU	$\{\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q, \mathsf{O}p \wedge {\sim}\mathsf{O}p\}$
14	$\mathsf{P}{\sim}t$	13; RU	$\{\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q, \mathsf{O}p \wedge {\sim}\mathsf{O}p\}$

Line 12 illustrates the conditional applicability of the contraposition rule in a $\mathbf{DP^x}$-proof.

There are various ways in which we can further extend the proof from Γ, but in no such extension will any of lines 8 or 9 ever be unmarked. The reason for this is that $\Gamma \nvdash_{\mathbf{DP}} \mathsf{O}{\sim}s \wedge {\sim}\mathsf{O}{\sim}s$. If the formula $\mathsf{O}{\sim}s \wedge {\sim}\mathsf{O}{\sim}s$ were **DP**-derivable from Γ, lines 8 or 9 could become unmarked in an extension of the proof.

All of the formulas derived at unmarked lines at this stage of the proof are finally derivable from Γ. The formulas $\mathsf{O}{\sim}q \wedge {\sim}\mathsf{O}{\sim}q$ and $\mathsf{O}p \wedge {\sim}\mathsf{O}p$, which were used as conditions in the illustration, are not members of any minimal disjunction of abnormalities derivable from Γ. Consequently, $\Gamma \vdash_{\mathbf{DP^x}} \mathsf{O}p$, $\Gamma \vdash_{\mathbf{DP^x}} {\sim}\mathsf{O}t$, and $\Gamma \vdash_{\mathbf{DP^x}} \mathsf{P}{\sim}t$, whereas $\Gamma \nvdash_{\mathbf{DP^x}} \mathsf{O}{\sim}s$, and $\Gamma \nvdash_{\mathbf{DP^x}} \mathsf{O}(p \wedge {\sim}s)$. It is safely left to the reader to check that for the derivability of these consequences it does not make a difference whether we use the reliability or minimal abnormality strategy.

In Section 4.8 we mentioned that marking is a dynamic matter. Lines that are marked at a stage of a proof, may be unmarked again at a later stage. Suppose, for instance, that the premise set Γ were extended with a new premise $\mathsf{O}{\sim}s$. Call this extended premise set Γ'. Then $\Gamma' \vdash_{\mathbf{DP}} \mathsf{O}{\sim}s \wedge {\sim}\mathsf{O}{\sim}s$. If we would extend

[4] From the $\mathbf{CLuNs}^{\perp}$-theorem $(t \wedge s) \supset s$ it follows by (NEC) and (KP) that $\vdash_{\mathbf{DP}} \mathsf{P}(t \wedge s) \supset \mathsf{P}s$. Hence the formula $\mathsf{P}s$ at line 10 follows from the formula $\mathsf{P}(t \wedge s)$ at line 6 by a simple application of modus ponens.

From the formula $\mathsf{O}({\sim}r \vee {\sim}s)$ at line 2 it follows by (OD) and $\mathbf{CL^{pos}}$ that $\mathsf{O}{\sim}s \vee \mathsf{P}{\sim}r$. Since $\mathsf{P}{\sim}r$ is equivalent to ${\sim}\mathsf{O}r$ (by (P) and (A${\sim}{\sim}'$)), and since we know that $\mathsf{O}r$ (line 4) and ${\sim}\mathsf{O}{\sim}s$ (by (P) and line 10), it follows by $\mathbf{CL^{pos}}$ that $(\mathsf{O}r \wedge {\sim}\mathsf{O}r) \vee (\mathsf{O}{\sim}s \wedge {\sim}\mathsf{O}{\sim}s)$.

the proof above with the formula $\mathsf{O}{\sim}s \wedge {\sim}\mathsf{O}{\sim}s$, the formula on line 11 would no longer be a minimal *Dab*-formula. Consequently, lines 8 and 9 would no longer be marked.

Above it was illustrated how $\mathbf{DP^x}$ interprets a given premise set 'as normally as possible'. Whenever a line is inferred in a $\mathbf{DP^x}$-proof, whether or not this inference is considered to be reliable depends on whether or not its condition 'behaves normally'. According to the reliability strategy, a condition behaves normally as long as it does not contain a member of the set of unreliable formulas of the premise set. As we have seen, this requirement is loosened a bit for the minimal abnormality strategy. The behavior of a condition of a line in a $\mathbf{DP^x}$-proof is independent of the rule that was applied at this line. This explains why in a $\mathbf{DP^x}$-proof some applications of a rule are marked whereas other applications of the same rule remain unmarked throughout the proof and any of its extensions.

The illustration above already shows how rules like deontic disjunctive syllogism (line 7) and contraposition (line 12) are conditionally applicable in $\mathbf{DP^x}$. But we can prove a far stronger result. Next, we show that all inferences valid in **SDL** are either unconditionally or conditionally applicable in $\mathbf{DP^x}$. Consequently, for premise sets from which no abnormalities are derivable, the logic $\mathbf{DP^x}$ is just as strong as **SDL**.

6.1.2.2 $\mathbf{DP^x}$ and SDL

In interpreting a set of premises 'as normally as possible', we implicitly make use of a certain standard of normality. In this section we make clear that for $\mathbf{DP^x}$ this standard of normality is **SDL**.

In Section 4.9 we stated that the upper limit logic **ULL** of an adaptive logic $\mathbf{AL^x}$ is obtained by adding to its lower limit logic one or more axiom schemas that trivialize all $\mathbf{AL^x}$-abnormalities.

The upper limit logic **UDP** of $\mathbf{DP^x}$ is obtained by adding to **DP** the axiom schema (UDP), which trivializes all abnormalities in Ω. Where $A \in \mathcal{W}^a \cup \mathcal{W}^{\sim}_{\mathsf{O}}$, $B \in \mathcal{W}^{\mathbf{DP}}$:

$$(A \wedge {\sim}A) \supset B \tag{UDP}$$

UDP trivializes contradictions, thus promoting "~" to a fully classical negation connective. In fact, **UDP** is just **SDL** in disguise. Where $\Gamma \subseteq \mathcal{W}^{\mathbf{DP}}$, define $\Gamma^{\neg}$ by replacing every $A \in \Gamma$ by $\pi(A)$, where $\pi(A)$ is the result of replacing every occurrence of "~" in A by "¬". Then:

Theorem 29. $\Gamma \vdash_{\mathbf{UDP}} A$ *iff* $\Gamma^{\neg} \vdash_{\mathbf{SDL}} \pi(A)$.

A proof outline of Theorem 29 is contained in Appendix E.3.

$\mathbf{DP^x}$ interprets a given premise set in terms of **SDL** 'whenever possible'. The negation connective ~ of $\mathbf{DP^x}$ is hence strengthened to a fully classical negation connective 'as much as possible'. Moreover, all **SDL**-rules can be applied either conditionally or unconditionally in a $\mathbf{DP^x}$-proof. Those applications of **SDL**-rules that are considered safe according to the adaptive logic constitute the final $\mathbf{DP^x}$-consequences of the premise set.

Corollary 3. *Where* Γ *is normal and* $A, B \in \mathcal{W}^{\mathbf{DP}}$, $C, D \in \mathcal{W}^{\sim}_{\perp}$*:*

(i) *If* $\Gamma \vdash_{\mathbf{DP^x}} A \vee B$ *and* $\Gamma \vdash_{\mathbf{DP^x}} {\sim}A$*, then* $\Gamma \vdash_{\mathbf{DP^x}} B$
(ii) *If* $\Gamma \vdash_{\mathbf{DP^x}} \mathsf{O}(C \vee D)$ *and* $\Gamma \vdash_{\mathbf{DP^x}} \mathsf{O}{\sim}C$*, then* $\Gamma \vdash_{\mathbf{DP^x}} \mathsf{O}D$
(iii) *If* $\Gamma \vdash_{\mathbf{DP^x}} A \supset B$*, then* $\Gamma \vdash_{\mathbf{DP^x}} {\sim}B \supset {\sim}A$

Corollary 3 follows immediately in view of Theorem 29. Remember that a premise set Γ is normal iff no *Dab*-formulas are derivable from it, or, equivalently, if $U(\Gamma) = \emptyset$. (i) and (iii) illustrate that, for normal premise sets, $\mathbf{DP^x}$ validates all instances of disjunctive syllogism and contraposition. (ii) illustrates that, for normal premise sets, $\mathbf{DP^x}$ validates all instances of deontic disjunctive syllogism. Note that Horty's Smith example is an instance of (ii).

Not only normal premise sets, but also *non-normal* premise sets usually have more $\mathbf{DP^x}$-consequences than $\mathbf{DP}$-consequences (cfr. the example proof in the previous section). Note that by Theorem 17 and Theorem 29:

Corollary 4. $\Gamma^{\neg} \vdash_{\mathbf{SDL}} \pi(A)$ *iff there is a* $\Delta \subseteq \Omega$ *for which* $\Gamma \vdash_{\mathbf{DP}} A \vee Dab(\Delta)$.

Hence whenever a formula $\pi(A)$ is an $\mathbf{SDL}$-consequence of some premise set $\Gamma^{\neg}$, we can construct a $\mathbf{DP^x}$-proof from Γ such that, at some line i of this proof, A is the second element and Δ the fourth.

6.2 Reasoning about norms

In this section, we present the logic of norm-propositions $\mathbf{LNP^x}$. In sections 6.2.1 and 6.2.2, we introduce in an informal way some of the key concepts that feature in this normative context, and that are studied in more detail later on. In Section 6.2.3 we introduce the logic $\mathbf{LNP}$, a semi-paraconsistent and semi-paracomplete deontic logic that serves as the LLL of $\mathbf{LNP^x}$. The latter system is defined in Section 6.2.4.

We further illustrate the workings of $\mathbf{LNP^x}$ in Section 6.2.5, and discuss its meta-theory and its relation to $\mathbf{SDL}$ in Section 6.2.6. In Section 6.2.7 we compare $\mathbf{LNP}$ to some other logics of norm-propositions presented in the literature.

6.2.1 Normative conflicts and normative gaps

Ideally, sets of norms issued by agents, authorities, legislators, etc. are both consistent and complete. In our everyday practice, however, such sets often contain normative conflicts and normative gaps.

In legal contexts, existence of normative conflicts is nicely motivated by a passage written by Alchourrón and Bulygin which we already cited in Section 3.2.1.1:

> Even one and the same authority may command that p and that *not* p at the same time, especially when a great number of norms are enacted on the same occasion. This happens when the legislature enacts a very extensive statute, e.g. a Civil Code, that usually contains four to six thousand dispositions. All of them are regarded as promulgated at the same time, by the same authority, so that there is no wonder that they sometimes contain a certain amount of explicit or implicit contradictions. [3, pp. 112-113]

Normative conflicts also arise where both an obligation to do something and a (positive) permission not to do it are promulgated [1, 3, 35, 194].

The adaptive logics to be presented in this section deal in an adequate way with both normative conflicts and normative gaps. We say that a set of norms contains a *normative gap* with respect to a formula A if A is neither positively permitted nor forbidden nor obliged. For a defense of the existence of normative gaps, see e.g. [2, Chapters 7,8], [41].

Note that the formulation refers to *positive permissions* (also, *strong permissions*), i.e. permissions that are either explicitly stated as such, or permissions that are derivable from other explicitly stated permissions or obligations. This is to be distinguished from so-called *weak* or *negative permissions*: A is weakly permitted in case A is not forbidden. Would we replace "positive permission" by "weak permission" in the definition of normative gaps then the concept would be vacuous since each A is either forbidden or not forbidden (and hence, weakly permitted).

The practical use of the distinction between positive and negative permission can be illustrated by means of the legal principle *nullum crimen sine lege.* According to this principle anything which is not forbidden is permitted.[5] Alternatively, the principle states that a negative permission to do A implies a positive permission to do A. Typically, the nullum crimen principle is understood as a rule of closure permitting all the actions not prohibited by penal law [2, pp. 142-143]. We return to this principle in Section 6.2.3.1.

We will in the remainder tacitly assume that in case A is obligatory then A is positively permitted. In this case, there is a normative gap with respect to A iff A is neither positively permitted nor forbidden.

Another way to think about normative gaps is in terms of normative determination: A is *normatively determined* if and only if A is either positively permitted or forbidden, which is to say that there is no normative gap with respect to A.[6] We say that a set of norms is *normatively complete* if all of its norms are normatively determined, i.e. if there are no gaps with respect to any of its norms. From the existence of incomplete legal systems, Bulygin concludes that legal gaps are perfectly possible:

> It is not true that all legal systems are necessarily complete. The problem of completeness is an empirical, contingent, question, whose truth depends on the contents of the system. So legal gaps due to the silence of the law [...] are perfectly possible. [41, p. 28]

6.2.2 Norm-propositions and their formal representation

As pointed out in Section 1.2.2, it is important to distinguish between norms and norm-propositions in deontic logic. As a norm, a formula of the form $\ulcorner \mathsf{O}A \urcorner$ means something like "you ought to do $\ulcorner A \urcorner$", or "it is obligatory that $\ulcorner A \urcorner$", and a formula of the form $\ulcorner \mathsf{P}A \urcorner$ means something like "you may do $\ulcorner A \urcorner$", or "it

[5] Legal philosophers also refer to this principle as the *sealing legal principle.* We thank an anonymous referee for pointing this out.

[6] The notion of normative determination is adopted from [193].

is permitted that $\ulcorner A \urcorner$".[7] As a norm-proposition, a formula of the form $\ulcorner \mathsf{O}A \urcorner$ [$\ulcorner \mathsf{P}A \urcorner$] means something like "there is a norm to the effect that $\ulcorner A \urcorner$ is obligatory [permitted]". Thus, in our descriptive reading a formula $\ulcorner \mathsf{P}A \urcorner$ always denotes a strong permission.

According to Alchourrón and Bulygin [1, 2, 3], any perceived harmony between norms and norm-propositions in deontic logic is merely apparent. Instead of using the same calculus of deontic logic for reasoning with both norms and norm-propositions, we need two separate logics: a logic of norms and a logic of norm-propositions. In this section we are concerned with the characterization of a logic of norm-propositions.

OO- and OP-conflicts between norm-propositions are expressed as before by formulas such as $\ulcorner \mathsf{O}A \wedge \mathsf{O}\ \textit{not}\ A \urcorner$ in case two obligations conflict, and $\ulcorner \mathsf{O}A \wedge \mathsf{P}\ \textit{not}\ A \urcorner$ in case an obligation conflicts with a permission.

Normative gaps occur if neither $\ulcorner \mathsf{P}A \urcorner$ nor $\ulcorner \mathsf{O}\ \textit{not}\ A \urcorner$ is the case. A full formal characterization of normative gaps is presented after the definition of our formal language. As pointed out above, the permission in question is a strong permission. Weak permissions may be defined as the modal duals to O by $\ulcorner \textit{not}\ \mathsf{O}\ \textit{not}\ A \urcorner$. The latter expresses that "there is no norm to the effect that $\ulcorner \textit{not}\ A \urcorner$ is obligatory" and hence it expresses the descriptive meaning of a weak permission. However, we need an independent permission operator P in order to express strong permissions. From $\ulcorner \mathsf{P}A \urcorner$ we cannot infer $\ulcorner \textit{not}\ \mathsf{O}\ \textit{not}\ A \urcorner$ due to the possible existence of an OP-conflict. Similarly we cannot, vice versa, infer $\ulcorner \mathsf{P}A \urcorner$ from $\ulcorner \textit{not}\ \mathsf{O}\ \textit{not}\ A \urcorner$ since, despite the absence of a norm that expresses that $\ulcorner \textit{not}\ A \urcorner$ is obliged, $\ulcorner A \urcorner$ may not be positively permitted.[8]

In the remainder we show how each of the concepts presented in this introductory section is formalized and treated by the logics defined later on. In Section 6.2.3 we define the Logic of Norm-Propositions **LNP**. This logic is sufficiently expressive to formalize both normative conflicts and normative gaps without having to resort to the meta-language. Inside the scope of its deontic operators, **LNP** makes use of a paraconsistent and paracomplete negation connective for dealing with normative conflicts and normative gaps.

As a result of the weakness of this negation connective, **LNP** is not powerful enough for capturing many intuitive normative inferences. We deal with this problem in Section 6.2.4, where we strengthen **LNP** within the adaptive logics framework. This results in two adaptive logics which interpret a given premise set 'as consistently and as completely as possible'.

Next, we further illustrate the workings of these logics (Section 6.2.5) and provide some meta-theoretical properties (Section 6.2.6). In Section 6.2.7, we compare the logics defined here to other approaches taken up in the literature on norm-propositions.

---

[7] Until our formal language is defined, we use brackets "$\ulcorner$" and "$\urcorner$" for denoting formulas in order to avoid possible confusions.

[8] See [2, 192] for further arguments against the equivalence of $\ulcorner \mathsf{P}A \urcorner$ and $\ulcorner \textit{not}\ \mathsf{O}\ \textit{not}\ A \urcorner$ in a descriptive setting.

6.2.3 The logic LNP

6.2.3.1 Syntax

In the setting of norm-propositions, negation behaves differently depending on whether it occurs inside or outside the scope of an operator O or P. Outside the scope of a deontic operator, negation behaves classically. A formula $\ulcorner$*not* $\mathsf{O}p\urcorner$ is read as "it is *not* the case that there is a norm to the effect that p is obligatory". Under this reading, $\ulcorner$*not* $\mathsf{O}p\urcorner$ is incompatible with $\ulcorner\mathsf{O}p\urcorner$: $\ulcorner\mathsf{O}p\urcorner$ and $\ulcorner$*not* $\mathsf{O}p\urcorner$ cannot both be the case. Moreover, one of $\ulcorner\mathsf{O}p\urcorner$ or $\ulcorner$*not* $\mathsf{O}p\urcorner$ must hold: either there is a norm to the effect that p is obligatory, or there is not.

Things change when we turn to negations inside the scope of O or P. Here, both $\ulcorner\mathsf{O}p\urcorner$ and $\ulcorner\mathsf{O}$ *not* $p\urcorner$ are verified by the same set of norm-propositions if this set contains an OO-conflict with respect to p. Moreover, neither $\ulcorner\mathsf{P}p\urcorner$ nor $\ulcorner\mathsf{O}$ *not* $p\urcorner$ are verified by a given set of norm-propositions that contains a normative gap with respect to p. Given the standard characterizations of O and P, this means that – inside the scope of O or P– both the consistency and the completeness constraint for negation fail in some instances: $\ulcorner\mathsf{P}(p\wedge$ *not* $p)\urcorner$ is true in case of a normative conflict, and $\ulcorner\mathsf{O}(p\vee$ *not* $p)\urcorner$ is false in case of a normative gap.

The logic **LNP** is defined in such a way that it respects this distinction: outside the scope of a deontic operator, only the classical negation connective "$\neg$" occurs. Inside the scope of a deontic operator, **LNP** makes use of the connective "$\sim$", which is a paraconsistent and paracomplete "negation" connective, i.e. it invalidates both $\ulcorner(A\wedge{\sim}A)\supset B\urcorner$ (*ex contradictione quodlibet*) and $\ulcorner A\vee{\sim}A\urcorner$ (*excluded middle*).[9]

Let $\mathcal{W}^{\sim}$ be the $\langle\sim,\vee,\wedge,\supset,\equiv\rangle$-closure of $\mathcal{W}^a$, and:

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

We do not allow for nested occurrences of the modal operators in our language. The set $\mathcal{W}^{\mathbf{LNP}}$ of well-formed formulas of **LNP** is defined as the $\langle\neg,\vee,\wedge,\supset,\equiv\rangle$-closure of $\mathcal{W}\cup\mathcal{W}_{\mathsf{O}}^{\sim}$.
Since the denotation of formulas is no longer ambiguous now that our language $\mathcal{W}^{\mathbf{LNP}}$ is defined, we skip the $\ulcorner\urcorner$-marks in the remainder.

Both normative conflicts and normative gaps are expressible in the object language $\mathcal{W}^{\mathbf{LNP}}$. A normative conflict occurs relating to a formula $A\in\mathcal{W}^{\sim}$ whenever we can derive one of $\mathsf{O}A\wedge\mathsf{O}{\sim}A$ or $\mathsf{O}A\wedge\mathsf{P}{\sim}A$. A normative gap occurs relating to A whenever we can derive $\neg\mathsf{P}A\wedge\neg\mathsf{O}{\sim}A$, i.e. whenever there is no norm to the effect that A is permitted or forbidden.

The P-operator functions as an operator for positive permission. A proposition A is said to be negatively permitted if there is no obligation to the contrary, i.e. if $\neg\mathsf{O}{\sim}A$. The nullum crimen principle can be formalized as an axiom schema:

$$\neg\mathsf{O}{\sim}A\supset\mathsf{P}A \tag{NC}$$

[9] "$\sim$" as defined below is actually a "dummy" connective rather than a negation connective: it has no properties at all, except that it validates de Morgan's laws. However, below we show that "$\sim$" functions as a negation connective in the adaptive strengthenings of the logic **LNP**.

Clearly, (NC) a priori excludes the possibility of normative gaps. That is why it is not validated by any gap-tolerant logic of norm-propositions, including **LNP**.

6.2.3.2 Semantics

LNP is characterizable within a Kripke-style semantics with a set of worlds or points W and a designated or 'actual' world $w_0 \in W$. In w_0, negation is defined classically by means of the connective "$\neg$". In the other worlds, negation is defined by the paraconsistent and paracomplete connective "$\sim$".[10]

An **LNP**-model is a tuple $\langle W, w_0, R, v_0, v \rangle$, where $R = \{w_0\} \times (W \setminus \{w_0\})$ is a non-empty accessibility relation, and $v_0 : \mathcal{W}^a \times \{w_0\} \to \{0,1\}$ and $v : \mathcal{W}_l^\sim \times (W \setminus \{w_0\}) \to \{0,1\}$ are assignment functions. v_0 assigns truth-values to atomic propositions at the actual world w_0. Since all logical connectives (including negation) behave classically at this world, truth values for complex formulas can be defined in terms of a valuation function in the usual way. The situation is slightly different for other worlds. In the latter, the $\sim$-connective does not behave classically and truth values are assigned to all $\sim$-literals, i.e. all atomic propositions p and their $\sim$-negation $\sim p$.

Let $w \in W, w' \in W \setminus \{w_0\}$. Then the valuation $v_M : (\mathcal{W}^{\mathbf{LNP}} \times \{w_0\}) \cup (\mathcal{W}^\sim \times W \setminus \{w_0\}) \to \{0,1\}$, associated with the model M, is defined by:

(C_0)	where $A \in \mathcal{W}^a, v_M(A, w_0) = 1$ iff $v_0(A, w_0) = 1$
(C_l)	where $A \in \mathcal{W}_l^\sim, v_M(A, w') = 1$ iff $v(A, w') = 1$
(C$\neg$)	$v_M(\neg A, w_0) = 1$ iff $v_M(A, w_0) = 0$
(C$\sim\sim$)	$v_M(\sim\sim A, w') = 1$ iff $v_M(A, w') = 1$
(C$\sim\supset$)	$v_M(\sim(A \supset B), w') = 1$ iff $v_M(A \wedge \sim B, w') = 1$
(C$\sim\wedge$)	$v_M(\sim(A \wedge B), w') = 1$ iff $v_M(\sim A \vee \sim B, w') = 1$
(C$\sim\vee$)	$v_M(\sim(A \vee B), w') = 1$ iff $v_M(\sim A \wedge \sim B, w') = 1$
(C$\sim\equiv$)	$v_M(\sim(A \equiv B), w') = 1$ iff $v_M((A \vee B) \wedge (\sim A \vee \sim B), w') = 1$
(C$\supset$)	$v_M(A \supset B, w) = 1$ iff $v_M(A, w) = 0$ or $v_M(B, w) = 1$
(C$\wedge$)	$v_M(A \wedge B, w) = 1$ iff $v_M(A, w) = v_M(B, w) = 1$
(C$\vee$)	$v_M(A \vee B, w) = 1$ iff $v_M(A, w) = 1$ or $v_M(B, w) = 1$
(C$\equiv$)	$v_M(A \equiv B, w) = 1$ iff $v_M(A, w) = v_M(B, w)$
(CO)	$v_M(\mathsf{O}A, w_0) = 1$ iff $v_M(A, w') = 1$ for every w' such that Rw_0w'
(CP)	$v_M(\mathsf{P}A, w_0) = 1$ iff $v_M(A, w') = 1$ for some w' such that Rw_0w'

(C_0) and (C_l) simply take over the values of the assignment functions v_0 and v respectively. (C$\neg$) determines truth values for the classical negation connective "$\neg$" in w_0. (C$\sim\sim$)-(C$\sim\equiv$) guarantee that de Morgan's laws hold for "$\sim$" in accessible worlds. Where $A \in \mathcal{W}^a$, the interpretation of $\sim A$ is provided directly by the assignment function v. Where A is a complex formula, its negation $\sim A$ can be reduced to simpler constituents in view of (C$\sim\sim$)-(C$\sim\equiv$). (C$\supset$)-(C$\equiv$) determine truth values for the other classical connectives $\supset, \wedge, \vee$, and $\equiv$ in all worlds, and (CO) and (CP) define the deontic operators O and P in the usual way.

A semantic consequence relation for **LNP** is defined in terms of truth preservation at the actual world: an **LNP**-model M verifies A ($M \Vdash A$) iff $v_M(A, w_0) =$

[10]The semantic clauses for accessible worlds are inspired by those for (the propositional fragment of) Batens' paraconsistent and paracomplete logic **CLoNs**, a variation on the paraconsistent logic **CLuNs** as found in e.g. [20]. **CLoNs** is defined in Appendix C.

1. In Section 6.2.3.4, we discuss the workings of **LNP** in more detail and provide some illustrations. But first we define its syntactic consequence relation.

6.2.3.3 Axiomatization and meta-theory

Inside the scope of O and P, we want to allow for the consistent possibility of contradictions and gaps. In order to do so, we make use of the propositional fragment of the logic **CLoNs** (cfr. footnote 10). **CLoNs** is defined by adding de Morgan's laws for "~" to $\mathbf{CL^{pos}}$:[11]

$$\sim\sim A \equiv A \tag{A$\sim\sim$}$$
$$\sim(A \supset B) \equiv (A \wedge \sim B) \tag{A$\sim\supset$}$$
$$\sim(A \wedge B) \equiv (\sim A \vee \sim B) \tag{A$\sim\wedge$}$$
$$\sim(A \vee B) \equiv (\sim A \wedge \sim B) \tag{A$\sim\vee$}$$
$$\sim(A \equiv B) \equiv ((A \vee B) \wedge (\sim A \vee \sim B)) \tag{A$\sim\equiv$}$$

Except for de Morgan's laws, "~" has no properties at all. The logic **LNP** is fully axiomatized by **CL** (with the classical negation connective "¬") plus:

$$\mathsf{O}(A \supset B) \supset (\mathsf{O}A \supset \mathsf{O}B) \tag{K}$$
$$\mathsf{O}A \supset \mathsf{P}A \tag{D}$$
$$\text{If } \vdash_{\mathbf{CLoNs}} A \text{ then } \vdash \mathsf{O}A \tag{NEC$^\sim$}$$
$$\mathsf{O}(A \supset B) \supset (\mathsf{P}A \supset \mathsf{P}B) \tag{KP}$$
$$\mathsf{O}(A \vee B) \supset (\mathsf{O}A \vee \mathsf{P}B) \tag{OD}$$
$$\mathsf{P}(A \vee B) \supset (\mathsf{P}A \vee \mathsf{P}B) \tag{PD}$$

LNP resembles **SDL** in the sense that it contains (K), (D), and a necessitation rule. However, it is non-standard in the sense that its necessitation rule (NEC$^\sim$) is defined in terms of theoremhood in **CLoNs** instead of theoremhood in **CL**. Moreover, in **LNP** the permission operator P is not definable in terms of the obligation operator O. Instead, the P-operator is characterized by the axiom schemas (KP), (OD), and (PD), all of which also hold in **SDL**.
The axiom schemas (O-AND) and (P-AND) are derivable in **LNP** (their derivability is shown in Fact 7 in Section F.1 of the Appendix):

$$\mathsf{O}A, \mathsf{O}B \vdash_{\mathbf{LNP}} \mathsf{O}(A \wedge B) \tag{O-AND}$$
$$\mathsf{O}A, \mathsf{P}B \vdash_{\mathbf{LNP}} \mathsf{P}(A \wedge B) \tag{P-AND}$$

Theorem 30. *If* $\Gamma \vdash_{\mathbf{LNP}} A$*, then* $\Gamma \vDash_{\mathbf{LNP}} A$*. (Soundness of* **LNP***)*

Theorem 31. *If* $\Gamma \vDash_{\mathbf{LNP}} A$*, then* $\Gamma \vdash_{\mathbf{LNP}} A$*. (Strong Completeness of* **LNP***)*

Proofs for Theorem 30 and Theorem 31 are contained in Section F.2 of the Appendix.

[11] Remember that axiomatizations of $\mathbf{CL^{pos}}$ and **CLoNs** are contained in Appendix C.

6.2.3.4 Discussion

LNP allows for the consistent possibility of normative conflicts and normative gaps, and invalidates deontic explosion:

$$\mathsf{O}p \wedge \mathsf{O}{\sim}p \nvdash_{\mathbf{LNP}} \mathsf{O}q \tag{6.14}$$

$$\mathsf{O}p \wedge \mathsf{P}{\sim}p \nvdash_{\mathbf{LNP}} \mathsf{O}q \tag{6.15}$$

$$\neg\mathsf{P}p \wedge \neg\mathsf{O}{\sim}p \nvdash_{\mathbf{LNP}} \mathsf{O}q \tag{6.16}$$

In accordance with the discussion in Section 6.2.2, the following interdependencies between the O- and P-operators are invalid in **LNP**:

$$\mathsf{P}p \nvdash_{\mathbf{LNP}} \neg\mathsf{O}{\sim}p \tag{6.17}$$

$$\neg\mathsf{P}p \nvdash_{\mathbf{LNP}} \mathsf{O}{\sim}p \tag{6.18}$$

$$\mathsf{O}p \nvdash_{\mathbf{LNP}} \neg\mathsf{P}{\sim}p \tag{6.19}$$

$$\neg\mathsf{O}p \nvdash_{\mathbf{LNP}} \mathsf{P}{\sim}p \tag{6.20}$$

(6.17)-(6.20) correspond to the characterization of the P-operator as an operator for positive permission. (6.17) fails in the presence of an OP-conflict $\mathsf{P}p \wedge \mathsf{O}{\sim}p$. (6.18) fails in the presence of a gap $\neg\mathsf{P}p \wedge \neg\mathsf{O}{\sim}p$. (6.19) fails in the presence of a conflict $\mathsf{O}p \wedge \mathsf{P}{\sim}p$, and (6.20) fails in the presence of a gap $\neg\mathsf{P}{\sim}p \wedge \neg\mathsf{O}p$.

The conflict- and gap-tolerance of **LNP**, as well as the non-interdefinability of its O- and P-operators, all depend crucially on the paraconsistency and paracompleteness of the "~"-connective. However, the very weak characterization of "~"also causes the **LNP**-invalidity of the following inferences:

$$\mathsf{O}(p \vee q), \mathsf{O}{\sim}q \nvdash_{\mathbf{LNP}} \mathsf{O}p \tag{6.21}$$

$$\mathsf{O}(p \vee q), \mathsf{O}({\sim}p \vee q) \nvdash_{\mathbf{LNP}} \mathsf{O}q \tag{6.22}$$

$$\mathsf{O}(p \supset q), \mathsf{O}{\sim}q \nvdash_{\mathbf{LNP}} \mathsf{O}{\sim}p \tag{6.23}$$

Indeed, except for de Morgan's laws **LNP** invalidates all classically valid inferences that somehow depend on the properties of the ~-connective, e.g. the disjunctive syllogism or contraposition rules. (6.21) is invalid because the possibility of an OO-conflict $\mathsf{O}q \wedge \mathsf{O}{\sim}q$ cannot be excluded. In that case, $\mathsf{O}p$ need not follow from the premises $\mathsf{O}(p \vee q)$ and $\mathsf{O}{\sim}q$. Likewise, (6.22) is invalid since $\mathsf{O}q$ need not follow from $\mathsf{O}(p \vee q)$ and $\mathsf{O}({\sim}p \vee q)$ in the presence of an OO-conflict $\mathsf{O}p \wedge \mathsf{O}{\sim}p$.

(6.23) fails (i) in case of a normative conflict relating to q *or* (ii) in case of a normative gap relating to p. Suppose that $\mathsf{O}(p \supset q)$ and $\mathsf{O}{\sim}q$ are true at the actual world. Then $p \supset q$ and ${\sim}q$ are true at all accessible worlds. In case (i), both q and ${\sim}q$ are true in at least one accessible world. In this world, $p \supset q$ is true in view of (C$\supset$), and ${\sim}p$ need not be true. In case ${\sim}p$ is false at an accessible world, we have a model in which $\mathsf{O}{\sim}p$ is false at the actual world. In case (ii), both p and ${\sim}p$ are false in at least one accessible world. Again we have a model in which $\mathsf{O}{\sim}p$ is false at the actual world.

For similar reasons all of the following 'variants' of (6.21)-(6.23) are invalid in **LNP**:

$$\mathsf{O}(p \vee q), \mathsf{P}{\sim}q \nvdash_{\mathbf{LNP}} \mathsf{P}p \tag{6.24}$$

$$\mathsf{P}(p \vee q), \mathsf{O}{\sim}q \nvdash_{\mathbf{LNP}} \mathsf{P}p \tag{6.25}$$
$$\mathsf{O}(p \vee q), \mathsf{P}({\sim}p \vee q) \nvdash_{\mathbf{LNP}} \mathsf{P}q \tag{6.26}$$
$$\mathsf{P}(p \vee q), \mathsf{O}({\sim}p \vee q) \nvdash_{\mathbf{LNP}} \mathsf{P}q \tag{6.27}$$
$$\mathsf{O}(p \supset q), \mathsf{P}{\sim}q \nvdash_{\mathbf{LNP}} \mathsf{P}{\sim}p \tag{6.28}$$
$$\mathsf{P}(p \supset q), \mathsf{O}{\sim}q \nvdash_{\mathbf{LNP}} \mathsf{P}{\sim}p \tag{6.29}$$
$$\mathsf{O}(p \supset q) \nvdash_{\mathbf{LNP}} \mathsf{O}({\sim}q \supset {\sim}p) \tag{6.30}$$
$$\mathsf{P}(p \supset q) \nvdash_{\mathbf{LNP}} \mathsf{P}({\sim}q \supset {\sim}p) \tag{6.31}$$

In spite of the rationale behind their invalidity (i.e. the possibility of normative conflicts/gaps), all of (6.21)-(6.31) have some intuitive appeal. In real life, we tend to *assume* that norms behave consistently and that propositions are normatively regulated. Normative conflicts and normative gaps are *anomalies*. We *rely* on inferences like (6.21)-(6.31) in our everyday reasoning processes, albeit in a *defeasible* way.

It seems then, that **LNP** is too weak to account for our normative reasoning. Inferences like (6.21)-(6.31) should only be blocked once we can reasonably assume that one of the norm-propositions needed in the inference behaves abnormally, i.e. that there might be a conflict or gap relating to this norm-proposition. Note that this reasoning process is non-monotonic: new premises may provide the information that there is a conflict or gap relating to some norm-proposition that was previously deemed to behave normally. Consider, for instance, the inference from $\mathsf{O}(p \vee q)$ and $\mathsf{O}{\sim}p$ to $\mathsf{O}q$. This inference is intuitive assuming that there is no normative conflict relating to p. If, however, we obtain the new information that there *is* a normative conflict relating to p, then the inference should be blocked, since we do not want to rely on conflicted norm-propositions in deriving new information.

In the next section, we strengthen **LNP** in a non-monotonic fashion in order to overcome the problems mentioned here, and to make formally precise the idea of 'assuming' norm-propositions to behave 'normally'.

6.2.4 The logics LNP$^{\mathbf{r}}$ and LPN$^{\mathbf{m}}$

For any **LNP**-model M and $A \in \mathcal{W}^a$, the classical negation connective "$\neg$" satisfies the following semantic conditions at the actual world:

(†) If $v_M(A, w_0) = 1$, then $v_M(\neg A, w_0) = 0$,
(‡) If $v_M(A, w_0) = 0$, then $v_M(\neg A, w_0) = 1$.

(†) guarantees the consistency of A: A and $\neg A$ cannot both be true at w_0. (‡) imposes a completeness condition on A: at least one of A and $\neg A$ is true at w_0.

As is clear from the **LNP**-semantics, (†) and (‡) fail for "$\sim$" at accessible worlds. Instead of (†) and (‡), only the weaker conditions (†') and (‡') hold for "$\sim$" at a world $w \in W \setminus \{w_0\}$:

(†') If $v_M(A, w) = 1$, then either $v_M({\sim}A, w) = 0$ or $v_M(A \wedge {\sim}A, w) = 1$,
(‡') If $v_M(A, w) = 0$, then either $v_M({\sim}A, w) = 1$ or $v_M(A \vee {\sim}A, w) = 0$.

In view of the semantic clauses for **LNP** it is easily checked that whenever a normative conflict occurs relating to a proposition p, the formula $p \wedge {\sim}p$ is true at some accessible world. In case of an OP-conflict $\mathsf{O}p \wedge \mathsf{P}{\sim}p$ or $\mathsf{O}{\sim}p \wedge \mathsf{P}p$, this follows in view of (CO), (CP), and (C$\wedge$). In case of an OO-conflict $\mathsf{O}p \wedge \mathsf{O}{\sim}p$, it follows in view of (CO), (C$\wedge$) and the non-emptiness of the accessibility relation.

In a similar fashion, we can check that whenever a normative gap occurs relating to p, the formula $p \vee {\sim}p$ is false at some accessible world. Suppose, for instance, that $\neg\mathsf{O}p \wedge \neg\mathsf{P}{\sim}p$ is true at w_0. Then by (C$\neg$), both $\mathsf{O}p$ and $\mathsf{P}{\sim}p$ are false at w_0. By (CO), there is a world w such that Rw_0w and $v_M(p, w) = 0$. By (CP), ${\sim}p$ too is false at this world: $v_M({\sim}p, w) = 0$. By (C$\vee$), $v_M(p \vee {\sim}p, w) = 0$.

Normative conflicts create truth-value gluts, whereas normative gaps create truth-value gaps at accessible worlds.[12] Suppose now that we label such gluts and gaps as *abnormal*, and that we try to interpret our worlds *as normally as possible*. Then, in view of (†') and (‡'), normal behavior corresponds to the satisfaction of the consistency and completeness demands (†) and (‡) for "~" at accessible worlds.

The adaptive logic $\mathbf{LNP^x}$ exploits the above idea in making the assumption that norm-propositions behave 'normally' unless and until we find out that they are involved in some normative conflict or gap. $\mathbf{LNP^x}$ is defined as a triple:

(1) Lower limit logic: **LNP**.
(2) Set of abnormalities: $\Omega = \Omega_1 \cup \Omega_2$, where $\Omega_1 = \{\mathsf{P}(A \wedge {\sim}A) \mid A \in \mathcal{W}^a\}$ and $\Omega_2 = \{\neg\mathsf{O}(A \vee {\sim}A) \mid A \in \mathcal{W}^a\}$.
(3) Adaptive strategy: $x \in \{r, m\}$.

Ω_1 is the set of atomic gluts true at some accessible world. Note that, in view of the validity of de Morgan's laws for "~", more complex gluts can be reduced to (disjunctions of) atomic gluts by the LLL, e.g. if $v_M((p \vee q) \wedge {\sim}(p \vee q), w) = 1$, then $v_M((p \wedge {\sim}p) \vee (q \wedge {\sim}q), w) = 1$. Consequently, whenever some **LNP**-model verifies an OO- or OP-conflict, it also validates an abnormality in the set Ω_1.

In view of the **LNP**-semantics, $p \vee {\sim}p$ is false at some accessible world whenever $\neg\mathsf{O}(p \vee {\sim}p)$ is true at the actual world. Thus Ω_2 is the set of atomic gaps true at some accessible world. Again, complex instances of gaps are **LNP**-reducible to a (disjunction of) atomic gap(s), e.g. if $v_M((p \vee q) \vee {\sim}(p \vee q), w) = 0$, then $v_M((p \vee {\sim}p) \wedge (q \vee {\sim}q), w) = 0$. Hence whenever some **LNP**-model verifies a normative gap, it also validates an abnormality in the set Ω_2.

For any atomic proposition p, the Ω_2-abnormality $\neg\mathsf{O}(p \vee {\sim}p)$ expresses that there is an accessible world in which *neither* p nor ${\sim}p$ is verified, whereas the Ω_1-abnormality $\mathsf{P}(p \wedge {\sim}p)$ expresses that there is an accessible world in which *both* p and ${\sim}p$ are verified. Thus, in **LNP** both gluts and gaps in accessible worlds constitute abnormalities. In view of the discussion at the beginning of this section, this means that both normative conflicts and normative gaps constitute abnormalities in **LNP**.

[12] A *truth-value glut* for p occurs when both p and ${\sim}p$ are assigned the value 1; a *truth-value gap* for p occurs when both p and ${\sim}p$ are assigned the value 0.

6.2.5 Some illustrations

6.2.5.1 Semantics

Example 19. Let $\Gamma_1 = \{\mathsf{O}p, \mathsf{O}({\sim}p \vee q)\}$. Then, for all **LNP**-models M of Γ_1, $M, w_0 \vDash \mathsf{O}p$ and $M, w_0 \vDash \mathsf{O}({\sim}p \vee q)$. By (C$\mathsf{O}$), $M, w \vDash p$ and $M, w \vDash {\sim}p \vee q$ for all worlds w such that Rw_0w. The possible truth values for $p, {\sim}p, q$, and ${\sim}q$ at accessible worlds in M are depicted in Table 6.1a. Let $R(w_0)$ abbreviate the set of worlds $w \in W \setminus \{w_0\}$ such that Rw_0w. Then each $w \in R(w_0)$ is of one of types (1)-(6).

Table 6.1: Accessible worlds for Γ_1 and Γ_3. Grey cells indicate propositions that behave abnormally in w.

(a) Accessible worlds for Γ_1

w	p	${\sim}p$	q	${\sim}q$
(1)	1	0	1	0
(2)	1	0	1	1
(3)	1	1	0	0
(4)	1	1	0	1
(5)	1	1	1	0
(6)	1	1	1	1

(b) Accessible worlds for Γ_3

w	p	${\sim}p$	q	${\sim}q$
(1)	0	0	0	1
(2)	0	0	1	1
(3)	0	1	0	1
(4)	0	1	1	1
(5)	1	0	1	1
(6)	1	1	1	1

If at least one $w \in R(w_0)$ is of one of types (3)-(6), then, by (C$\wedge$) and (CP), $M, w_0 \vDash \mathsf{P}(p \wedge {\sim}p)$, and $\mathsf{P}(p \wedge {\sim}p) \in Ab(M)$. Similarly, if at least one $w \in R(w_0)$ is of type (2) or type (6), then $\mathsf{P}(q \wedge {\sim}q) \in Ab(M)$. Moreover, if some $w \in R(w_0)$ is of type (3), then, by (C$\vee$), (CO) and (C$\neg$), $M, w_0 \vDash \neg\mathsf{O}(q \vee {\sim}q)$, and $\neg\mathsf{O}(q \vee {\sim}q) \in Ab(M)$.

If, however, all worlds $w \in R(w_0)$ are of type (1), then M verifies no abnormalities relating to p or q. In view of Definition 13, only models for which all worlds $w \in R(w_0)$ are of type (1) qualify as minimally abnormal **LNP**-models of Γ_1. Note that, for all type (1)-worlds $w \in R(w_0)$, $M, w \vDash q$. By (CO), $M, w_0 \vDash \mathsf{O}q$. By Definition 14, $\Gamma_1 \vDash_{\mathbf{LNP^m}} \mathsf{O}q$.

Since Γ_1 has **LNP**-models M of which all accessible worlds $w \in R(w_0)$ are such that, for all $A \in \mathcal{W}^a$, $M, w \nvDash A \wedge {\sim}A$ and $M, w \vDash A \vee {\sim}A$, we can conclude that Γ_1 has **LNP**-models M such that $Ab(M) = \varnothing$. It follows that Γ_1 has no minimal *Dab*-consequences. In view of Definition 7, $U(\Gamma_1) = \varnothing$. By Definition 9, $Ab(M) = \varnothing$ for all reliable **LNP**-models M of Γ_1. Again, only models for which all worlds $w \in R(w_0)$ are of type (1) qualify as reliable **LNP**-models of Γ_1. By Definition 10, $\Gamma_1 \vDash_{\mathbf{LNP^r}} \mathsf{O}q$.

Example 20. Let $\Gamma_2 = \{\mathsf{O}p, \mathsf{O}({\sim}p \vee q), \mathsf{O}{\sim}p\}$. It is easily checked that $\Gamma_2 \vDash_{\mathbf{LNP}} \mathsf{P}(p \wedge {\sim}p)$. Consequently, all **LNP**-models verify this abnormality, including the minimally abnormal and reliable ones. Hence all accessible worlds in all **LNP**-models of Γ_2 are of one of types (3)-(6) in Table 6.1a. Since $\mathsf{P}(p \wedge {\sim}p)$ is the only *Dab*-consequence of Γ_2, the selected $\mathbf{LNP^x}$-models for both strategies are those which verify exactly this abnormality, i.e. models of which all accessible worlds are of type (4) or (5). In all of these models, $p, {\sim}p \vee q$, and ${\sim}p$ are true at all accessible worlds. Since q need not be true at some of these worlds, Γ_2 has **LNP^x**-models in which $\mathsf{O}q$ is false. Hence $\Gamma_2 \nvDash_{\mathbf{LNP^x}} \mathsf{O}q$.

Note that Examples 19 and 20 illustrate the non-monotonicity of **LNP^x**: adding the premise $\mathsf{O}{\sim}p$ to Γ_1 blocks the derivation of $\mathsf{O}q$.

Example 21. Let $\Gamma_3 = \{\mathsf{O}(p \supset q), \mathsf{O}{\sim}q\}$, and let M be an **LNP**-model of Γ_3. The possible truth values for $p, {\sim}p, q$, and ${\sim}q$ at accessible worlds in M are depicted in Table 6.1b.

If at least one $w \in R(w_0)$ is of one of types (1) or (2), then $\neg\mathsf{O}(p \vee {\sim}p) \in Ab(M)$. If at least one $w \in R(w_0)$ is of one of types (2), (4), (5) or (6), then $\mathsf{P}(q \wedge {\sim}q) \in Ab(M)$. Only if all $w \in R(w_0)$ are of type (3) it is possible that $Ab(M) = \emptyset$. In view of Definition 13, only models of which all $w \in R(w_0)$ are of type (3) qualify as minimally abnormal models. But then $M, w_0 \vDash \mathsf{O}{\sim}p$, and, by Definition 14, $\Gamma_3 \vDash_{\mathbf{LNP^m}} \mathsf{O}{\sim}p$. It is safely left to the reader to check that, in view of Definitions 9 and 10, $\Gamma_3 \vDash_{\mathbf{LNP^r}} \mathsf{O}{\sim}p$.

Example 22. Let $\Gamma_4 = \{\mathsf{O}(p \wedge q), \mathsf{O}({\sim}(p \vee q) \vee r), \mathsf{P}({\sim}p \vee {\sim}q)\}$, and let M be an **LNP**-model of Γ_4. By (CO) we know that, for all $w \in R(w_0)$ in M, $M, w \vDash p \wedge q$ and $M, w \vDash {\sim}(p \vee q) \vee r$. Hence every $w \in R(w_0)$ is of one of types (1)-(10) depicted in Table 6.2.

w	p	${\sim}p$	q	${\sim}q$	r	${\sim}r$
(1)	1	0	1	0	1	0
(2)	1	0	1	0	1	1
(3)	1	0	1	1	1	0
(4)	1	0	1	1	1	1
(5)	1	1	1	0	1	0
(6)	1	1	1	0	1	1
(7)	1	1	1	1	0	0
(8)	1	1	1	1	0	1
(9)	1	1	1	1	1	0
(10)	1	1	1	1	1	1

Table 6.2: Accessible worlds for Γ_4.

By (CP), we also know that there is at least one world w such that $w \in R(w_0)$ and $M, w \vDash {\sim}p \vee {\sim}q$. Thus, w cannot be of type (1) or type (2). If w is of type (3), then $\mathsf{P}(q \wedge {\sim}q) \in Ab(M)$. If w is of type (5), then $\mathsf{P}(p \wedge {\sim}p) \in Ab(M)$. It is easily checked that if w is of type (4), (6), (7), (8), (9), or (10), then M validates more than one abnormality, i.e. either $\{\mathsf{P}(p \wedge {\sim}p)\} \subset Ab(M)$ or $\{\mathsf{P}(q \wedge {\sim}q)\} \subset Ab(M)$.

In general, it follows by Definition 13 that M only qualifies as a minimally abnormal **LNP**-model of Γ_4 if either w is of type (3) and all $w' \in R(w_0) \setminus \{w\}$ are of type (1) or type (3), or w is of type (5) and all $w' \in R(w_0) \setminus \{w\}$ are of type (1) or type (5). Hence if M is minimally abnormal, then all accessible worlds in M are of type (1), type (3), or type (5). But then, by (CO), $M, w_0 \vDash \mathsf{O}r$ and, by Definition 14, $\Gamma_4 \vDash_{\mathbf{LNP^m}} \mathsf{O}r$.

Since at least one accessible world w in M is of types (3)-(10), it follows by (C$\wedge$), (CP), and (C$\vee$) that $M, w_0 \vDash \mathsf{P}(p \wedge {\sim}p) \vee \mathsf{P}(q \wedge {\sim}q)$ (for any model M of

Γ_4). On the other hand, there exist models M of Γ_4 such that $M, w_0 \nvDash \mathsf{P}(p \wedge {\sim}p)$, and there exist models M of Γ_4 such that $M, w_0 \nvDash \mathsf{P}(q \wedge {\sim}q)$. Thus, it follows that $\mathsf{P}(p \wedge {\sim}p) \vee \mathsf{P}(q \wedge {\sim}q)$ is a minimal *Dab*-consequence of Γ_4. By Definition 7, $\mathsf{P}(p \wedge {\sim}p), \mathsf{P}(q \wedge {\sim}q) \in U(\Gamma_4)$.

Suppose now that all $w \in R(w_0)$ are of type (8), and that, for all $A \in \mathcal{W}^a \smallsetminus \{p, q, r\}$, $M, w \nvDash A \wedge {\sim}A$ and $M, w \vDash A \vee {\sim}A$. Then it is easily verified that the only abnormalities verified by M are $\mathsf{P}(p \wedge {\sim}p)$ and $\mathsf{P}(q \wedge {\sim}q)$. Thus, $Ab(M) \subseteq U(\Gamma_4)$. By Definition 9, M is reliable. However, $M, w_0 \nvDash \mathsf{O}r$. Thus, by Definition 10, $\Gamma_4 \nvDash_{\mathbf{LNP^r}} \mathsf{O}r$.

Example 22 illustrates that there are premise sets $\Gamma \subseteq \mathcal{W}^{\mathbf{LNP}}$ and formulas $A \in \mathcal{W}^{\mathbf{LNP}}$ such that $\Gamma \nvDash_{\mathbf{LNP^r}} A$ and $\Gamma \vDash_{\mathbf{LNP^m}} A$.

6.2.5.2 Proof theory

By now, readers are familiar with the workings of the adaptive proof theory. For this reason, the adaptive proofs in this section are provided without any extra information concerning the derivations made at each stage. However, the following table containing some **LNP**-valid inferences of the form $\Gamma \vdash_{\mathbf{LNP}} A \vee Dab(\Delta)$ is helpful in figuring out which moves can be made in a proof by means of the conditional rule RC:

$$\mathsf{O}p \vdash_{\mathbf{LNP}} \neg\mathsf{P}{\sim}p \vee \mathsf{P}(p \wedge {\sim}p) \tag{6.32}$$
$$\mathsf{P}p \vdash_{\mathbf{LNP}} \neg\mathsf{O}{\sim}p \vee \mathsf{P}(p \wedge {\sim}p) \tag{6.33}$$
$$\neg\mathsf{O}p \vdash_{\mathbf{LNP}} \mathsf{P}{\sim}p \vee \neg\mathsf{O}(p \vee {\sim}p) \tag{6.34}$$
$$\neg\mathsf{P}p \vdash_{\mathbf{LNP}} \mathsf{O}{\sim}p \vee \neg\mathsf{O}(p \vee {\sim}p) \tag{6.35}$$
$$\mathsf{O}(p \vee q), \mathsf{O}{\sim}q \vdash_{\mathbf{LNP}} \mathsf{O}p \vee \mathsf{P}(q \wedge {\sim}q) \tag{6.36}$$
$$\mathsf{O}(p \vee q), \mathsf{P}{\sim}q \vdash_{\mathbf{LNP}} \mathsf{P}p \vee \mathsf{P}(q \wedge {\sim}q) \tag{6.37}$$
$$\mathsf{P}(p \vee q), \mathsf{O}{\sim}q \vdash_{\mathbf{LNP}} \mathsf{P}p \vee \mathsf{P}(q \wedge {\sim}q) \tag{6.38}$$
$$\mathsf{O}(p \supset q), \mathsf{O}{\sim}q \vdash_{\mathbf{LNP}} \mathsf{O}{\sim}p \vee \neg\mathsf{O}(p \vee {\sim}p) \vee \mathsf{P}(q \wedge {\sim}q) \tag{6.39}$$
$$\mathsf{O}(p \supset q), \mathsf{P}{\sim}q \vdash_{\mathbf{LNP}} \mathsf{P}{\sim}p \vee \neg\mathsf{O}(p \vee {\sim}p) \vee \mathsf{P}(q \wedge {\sim}q) \tag{6.40}$$
$$\mathsf{P}(p \supset q), \mathsf{O}{\sim}q \vdash_{\mathbf{LNP}} \mathsf{P}{\sim}p \vee \neg\mathsf{O}(p \vee {\sim}p) \vee \mathsf{P}(q \wedge {\sim}q) \tag{6.41}$$
$$\mathsf{O}(p \vee q), \mathsf{O}({\sim}p \vee q) \vdash_{\mathbf{LNP}} \mathsf{O}q \vee \mathsf{P}(p \wedge {\sim}p) \tag{6.42}$$
$$\mathsf{O}(p \vee q), \mathsf{P}({\sim}p \vee q) \vdash_{\mathbf{LNP}} \mathsf{P}q \vee \mathsf{P}(p \wedge {\sim}p) \tag{6.43}$$
$$\mathsf{P}(p \vee q), \mathsf{O}({\sim}p \vee q) \vdash_{\mathbf{LNP}} \mathsf{P}q \vee \mathsf{P}(p \wedge {\sim}p) \tag{6.44}$$

As a first illustration, consider the following $\mathbf{LNP^x}$-proof from the premise set $\Gamma_5 = \{\mathsf{O}{\sim}p, \mathsf{P}({\sim}q \wedge ({\sim}r \vee s)), \mathsf{O}(r \vee s), \mathsf{O}({\sim}q \supset p)\}$:

2cm	1	$\mathsf{O}{\sim}p$	PREM	$\emptyset$
	2	$\mathsf{P}({\sim}q \wedge ({\sim}r \vee s))$	PREM	$\emptyset$
	3	$\mathsf{O}(r \vee s)$	PREM	$\emptyset$
	4	$\mathsf{O}({\sim}q \supset p)$	PREM	$\emptyset$
	5	$\mathsf{P}{\sim}q$	2; RU	$\emptyset$
	6	$\mathsf{P}({\sim}r \vee s)$	2; RU	$\emptyset$
	7	$\neg\mathsf{O}q$	5; RC	$\{\mathsf{P}(q \wedge {\sim}q)\}$

8	$\mathsf{P}s$	3,6; RC	$\{\mathsf{P}(r \wedge {\sim}r)\}$
9	$\mathsf{O}q$	1,4; RC	$\{\neg\mathsf{O}(q \vee {\sim}q), \mathsf{P}(p \wedge {\sim}p)\}\checkmark^{11}$
10	$\mathsf{P}p$	4,5; RU	$\emptyset$
11	$\mathsf{P}(p \wedge {\sim}p)$	1,10; RU	$\emptyset$

Since no other minimal *Dab*-formulas are **LNP**-derivable from Γ_5, it follows that $\Gamma_5 \vdash_{\mathbf{LNP^x}} \neg\mathsf{O}q$, $\Gamma_5 \vdash_{\mathbf{LNP^x}} \mathsf{P}s$ and $\Gamma_5 \nvdash_{\mathbf{LNP^x}} \mathsf{O}q$.

The following proof illustrates that $\neg\mathsf{O}({\sim}p \vee {\sim}q), \mathsf{O}({\sim}q \vee r), \neg\mathsf{P}p \vdash_{\mathbf{LNP^r}} \mathsf{P}r$:

1	$\neg\mathsf{O}({\sim}p \vee {\sim}q)$	PREM	$\emptyset$
2	$\mathsf{O}({\sim}q \vee r)$	PREM	$\emptyset$
3	$\neg\mathsf{P}p$	PREM	$\emptyset$
4	$\neg\mathsf{O}{\sim}p$	1; RU	$\emptyset$
5	$\neg\mathsf{O}{\sim}q$	1; RU	$\emptyset$
6	$\mathsf{P}p$	4; RC	$\{\neg\mathsf{O}(p \vee {\sim}p)\}\checkmark^{8}$
7	$\mathsf{P}q$	5; RC	$\{\neg\mathsf{O}(q \vee {\sim}q)\}$
8	$\neg\mathsf{O}(p \vee {\sim}p)$	3,4; RU	$\emptyset$
9	$\mathsf{P}r$	2,7; RC	$\{\neg\mathsf{O}(q \vee {\sim}q), \mathsf{P}(q \wedge {\sim}q)\}$

At this point, it is useful to come back to the nullum crimen principle (NC) as defined in Section 6.2.3.1. As is illustrated in the derivation of lines 6 and 7 in the proof above, $\mathbf{LNP^x}$ allows for the conditional application of (NC). In general, if A is not prohibited, then we can derive $\mathsf{P}A$ *on the condition that there is no normative gap relating to* A.

Finally, here is a $\mathbf{LNP^m}$-proof for the premise set Γ_4 from Section 6.2.5.1:

1	$\mathsf{O}(p \wedge q)$	PREM	$\emptyset$
2	$\mathsf{O}({\sim}(p \vee q) \vee r)$	PREM	$\emptyset$
3	$\mathsf{P}({\sim}p \vee {\sim}q)$	PREM	$\emptyset$
4	$\mathsf{O}p$	1; RU	$\emptyset$
5	$\mathsf{O}(({\sim}(p \vee q) \vee r) \wedge p)$	2,4; RU	$\emptyset$
6	$\mathsf{O}(r \vee (p \wedge {\sim}p))$	5; RU	$\emptyset$
7	$\mathsf{O}r \vee \mathsf{P}(p \wedge {\sim}p)$	6; RU	$\emptyset$
8	$\mathsf{O}r$	7; RC	$\{\mathsf{P}(p \wedge {\sim}p)\}$
9	$\mathsf{O}r$	1,2; RC	$\{\mathsf{P}(q \wedge {\sim}q)\}$
10	$\mathsf{P}(p \wedge {\sim}p) \vee \mathsf{P}(q \wedge {\sim}q)$	1,3; RU	$\emptyset$

In the $\mathbf{LNP^m}$-proof from Γ_4, the set $\Phi_{10}(\Gamma_4)$ of minimal choice sets of Γ_4 at stage 10 consists of the sets $\{\mathsf{P}(p \wedge {\sim}p)\}$ and $\{\mathsf{P}(q \wedge {\sim}q)\}$. In view of the marking definition for the minimal abnormality strategy, lines 8 and 9 remain unmarked. Since there is no way to extend the proof in such a way that these lines become marked, it follows that $\Gamma_4 \vdash_{\mathbf{LNP^m}} \mathsf{O}r$. Note that, if the above proof were a $\mathbf{LNP^r}$-proof from Γ_4, lines 8 and 9 would be marked due to the minimal *Dab*-formula derived at line 10.

6.2.6 Meta-theoretical properties of $\mathbf{LNP^x}$

Due to Theorem 8 and its definition within the standard format for ALs, $\mathbf{LNP^x}$ is sound and complete with respect to its semantics:

Corollary 5. *$\Gamma \vdash_{\mathbf{LNP^x}} A$ iff $\Gamma \vDash_{\mathbf{LNP^x}} A$.*

The upper limit **ULNP** of $\mathbf{LNP^x}$ is obtained by adding to **LNP** the axiom schemas (ULNP1) and (ULNP2), which trivialize all members of Ω_1 and Ω_2 respectively. Where $A \in \mathcal{W}^a$ and $B \in \mathcal{W}^{\mathbf{LNP}}$:

$$\mathsf{P}(A \wedge {\sim}A) \supset B \tag{ULNP1}$$

$$\neg\mathsf{O}(A \vee {\sim}A) \supset B \tag{ULNP2}$$

ULNP is related to **LNP** as set out by Theorem 17:

Corollary 6. *$\Gamma \vdash_{\mathbf{ULNP}} A$ iff (there is a $\Delta \subseteq \Omega$ for which $\Gamma \vdash_{\mathbf{LNP}} A \vee Dab(\Delta)$ or $\Gamma \vdash_{\mathbf{LNP}} A$).*

The set of *Dab*-consequences derivable from the premise set determines the amount to which the $\mathbf{LNP^x}$-consequence set will resemble the **ULNP**-consequence set. This is why adaptive logicians say that $\mathbf{LNP^x}$ *adapts* itself to a premise set. By Theorem 18, $\mathbf{LNP^x}$ will always be at least as strong as **LNP** and maximally as strong as **ULNP**:

Corollary 7. *$Cn_{\mathbf{LNP}}(\Gamma) \subseteq Cn_{\mathbf{LNP^x}}(\Gamma) \subseteq Cn_{\mathbf{ULNP}}(\Gamma)$.*

In view of Theorem 7, it follows immediately that:

Corollary 8. *$Cn_{\mathbf{LNP}}(\Gamma) \subseteq Cn_{\mathbf{LNP^r}}(\Gamma) \subseteq Cn_{\mathbf{LNP^m}}(\Gamma) \subseteq Cn_{\mathbf{ULNP}}(\Gamma)$.*

If Γ is *normal*, i.e. if Γ has no *Dab*-consequences, then, by Theorem 19:

Corollary 9. *If Γ is normal, then $Cn_{\mathbf{LNP^x}}(\Gamma) = Cn_{\mathbf{ULNP}}(\Gamma)$.*

The reader may have noticed that **ULNP** trivializes both gluts and gaps at accessible worlds, thus promoting "$\sim$" to a fully classical negation connective. It should come as no surprise then, that **ULNP** is just **SDL** in disguise. Where $\Gamma \subseteq \mathcal{W}^{\mathbf{LNP}}$, define $\Gamma^{\neg}$ by replacing every $A \in \Gamma$ by $\pi(A)$, where $\pi(A)$ is the result of replacing every occurrence of "$\sim$" in A by "$\neg$". Then:

Theorem 32. *$\Gamma \vdash_{\mathbf{ULNP}} A$ iff $\Gamma^{\neg} \vdash_{\mathbf{SDL}} \pi(A)$.*

A proof outline for Theorem 32 is contained in Section F.3 of the Appendix.

6.2.7 Related work

6.2.7.1 Alchourrón and Bulygin

In [1, 2, 3, 4], Alchourrón and Bulygin present a logic of norm-propositions that is built 'on top' of a logic of norms.[13] A norm-proposition "there exists a norm to the effect that A is permitted" is formalized as $\mathsf{NP}A$, where the operator N behaves like a quantifier over the norm $\mathsf{P}A$. The latter formula (without N) is read simply as "A is permitted". Only obligations and permissions can occur

[13] Alchourrón and Bulygin's logic of norm-propositions is inspired by Rescher's assertion logic from [151].

inside the scope of the N-operator; formulas of the form $\mathsf{N}A$ where A is not of the form $\mathsf{O}B$ or $\mathsf{P}B$ are not well-formed.

Alchourrón and Bulygin's logic of norms is just **SDL**. Their logic of norm-propositions **NL** extends **SDL** by adding to it the axiom schema (NK) and the rule (NRM):

$$\mathsf{N}(A \supset B) \supset (\mathsf{N}A \supset \mathsf{N}B) \tag{NK}$$

$$\text{If } \vdash A \supset B \text{ then } \vdash \mathsf{N}A \supset \mathsf{N}B \tag{NRM}$$

In **NL**, OO-conflicts are formulas of the form $\mathsf{NO}A \wedge \mathsf{NO}\neg A$. Similarly, OP-conflicts are formulas of the form $\mathsf{NO}A \wedge \mathsf{NP}\neg A$. As opposed to normative conflicts, normative gaps cannot be expressed in the object language of **NL**. Instead, Alchourrón and Bulygin define a normative gap as a situation in which, for some **CL**-formula A, we cannot derive $\mathsf{NP}A$ nor $\mathsf{NO}\neg A$, i.e. $\nvdash_{\mathbf{NL}} \mathsf{NP}A \vee \mathsf{NO}\neg A$. Normative conflicts and gaps do not cause full explosion in **NL**. Where A and B are well-formed **NL**-formulas:[14]

$$\mathsf{NO}A \wedge \mathsf{NO}\neg A \nvdash_{\mathbf{NL}} B \tag{6.45}$$

$$\mathsf{NO}A \wedge \mathsf{NP}\neg A \nvdash_{\mathbf{NL}} B \tag{6.46}$$

$$\nvdash_{\mathbf{NL}} \mathsf{NP}A \vee \mathsf{NO}\neg A \tag{6.47}$$

However, the following variants of deontic explosion are valid in **NL**:

$$\mathsf{NO}A \wedge \mathsf{NO}\neg A \vdash_{\mathbf{NL}} \mathsf{NO}B \tag{6.48}$$

$$\mathsf{NO}A \wedge \mathsf{NP}\neg A \vdash_{\mathbf{NL}} \mathsf{NO}B \tag{6.49}$$

With Alchourrón, Bulygin, and von Wright, we agree that "experience seems to testify that mutually contradictory norms may co-exist within one and the same legal order – and also that there are a good many "gaps" in any such order" [194, p. 32]. But if conflicting normative propositions indeed often coexist within a normative order, then deontic explosion should be avoided by any logic of normative propositions. No judge will agree that a normative order containing one or more conflicts contains norms to the effect that anything whatsoever is obligatory. Hence (6.48) and (6.49) cause serious problems for **NL**.

(6.48) and (6.49) follow by applications of (NRM) and (NK) to the **SDL**-theorems $\vdash \mathsf{O}A \supset (\mathsf{O}{\sim}A \supset \mathsf{O}B)$ and $\vdash \mathsf{O}A \supset (\mathsf{P}{\sim}A \supset \mathsf{O}B)$ respectively. This led von Wright to questioning the presupposition of **SDL** by **NL** [194, footnote 2].

As opposed to **NL**, $\mathbf{LNP^x}$ is not built 'on top' of the **CL**-based logic **SDL**. Although **LNP** contains full **CL**, its 'deontic' formulas make use of the much weaker logic **CLoNs** inside the scope of the O- and P-operator. This way, $\mathbf{LNP^x}$ avoids deontic explosion.

Interestingly, Alchourrón and Bulygin point out that under the assumptions of consistency and completeness, the logic of norm-propositions is 'isomorphic' to

[14] Alchourrón and Bulygin allow for iterated/nested deontic and normative operators. Nothing in principle prevents the occurrence of such nestings in $\mathbf{LNP^x}$. This requires some modifications of the language $\mathcal{W}^{\mathbf{LNP}}$ and of the sets Ω_1 and Ω_2 such that e.g. $\mathsf{PP}(p \wedge {\sim}p)$ is also considered an abnormality.

SDL: if we dismiss the possibility of normative conflicts and normative gaps, the differences between both logics disappear [1, 4]. In Section F.3 of the Appendix we prove this isomorphism for $\mathbf{LNP^x}$ by showing that for normal (consistent and complete) premise sets, $\mathbf{LNP^x}$ is just as strong as **SDL**.

6.2.7.2 Input/output logic

In Section 5.3 we already mentioned that input/output-logics (I/O logics) employ a technique similar to that used by Horty for obtaining a set of output obligations from a set of input obligations. As such, they suffer from the same problems concerning their treatment of normative conflicts (see Section 5.3). Since I/O logics were originally motivated as logics of norm-propositions, this is the right place for discussing their further properties in more detail.

In I/O logic, norms are represented as ordered pairs of formulas (a, x), where each coordinate of a pair is a **CL**-formula.[15] The body of such a pair constitutes an input consisting of some condition or factual situation. The head constitutes an output representing what the norm tells us to be desirable, obligatory or permitted in that situation. A normative order or system is a set G of input/output pairs. G is seen as a 'transformation device' in which **CL** functions as its 'secretarial assistant' [118, p. 2].

In [115], Makinson and van der Torre define various operations of the form $out(G, A)$ for making up the output of G given a set A containing factual information (input). In [116], the authors add constraints to these systems for dealing with contrary-to-duty scenarios and conflicting norms. In [117], the framework is extended for dealing with permissions. Constrained I/O logics make use of maximally consistent subsets. In doing so, they avoid explosion when dealing with conflicting conditional obligations, even if e.g. the norms (a, x) and $(a, \neg x)$ tell us that both x and $\neg x$ are obligatory under the same circumstances.

The treatment of obligation-permission conflicts by constrained I/O logics is less straightforward. In [169], Stolpe noted that the constrained systems deontically explode when facing a conflict between an obligation (a, x) and a positive permission $(a, \neg x)$.[16] Stolpe's solution to this problem is to treat positive permissions as *derogations*: "a positive permission suspends, annuls or obstructs a covering prohibition, thereby generating a corresponding set of liberties" [169, p. 99].

Stolpe's solution creates an asymmetry between obligations and permissions. In obligation-obligation conflicts, both norms may still be of equal importance. In obligation-permission conflicts however, the permission always overrides the obligations it is in conflict with. Although certainly of interest in legal contexts, where the concept of derogation is a very important one, we doubt that *all* obligation-permission conflicts can be dealt with in this way.

[15] The framework of I/O logic was initially developed for dealing with conditional norms. We do not discuss its merits as a conditional logic here. Instead, we focus on issues related to conflict- and gap-tolerance. For a discussion of the representation of conditional norms in I/O logic, see [201].

[16] Translated to the I/O setting, deontic explosion ensues from a given input if – under certain circumstances invoked by the input – everything becomes obligatory in the output.

In the literature on I/O logic, normative gaps are left unmentioned. However, it seems possible to model gaps in this framework. For instance, we could say that there is a normative gap relating to proposition x in circumstances a if neither the obligations to do x or $\neg x$, nor the positive permissions to do x or $\neg x$ are in the output of a given set of norms. One drawback seems to be that, whichever I/O operation we pick, both the obligation to do $x \vee \neg x$ and the positive permission to do $x \vee \neg x$ will always be in the output set. This is due to the closure of the output set under **CL**. Furthermore, as with Alchourrón and Bulygin's approach, normative gaps cannot be modeled at the object level in I/O logic.

Another difference between I/O logic and $\mathbf{LNP^x}$ is that for I/O operations the input is restricted to simple norm-bases, i.e. sets of input-output pairs. More complex formulas such as disjunctions between norms or negated norms cannot be fed into the system. $\mathbf{LNP^x}$ is more flexible in this sense, since it can easily deal with premise sets containing formulas such as $\neg \mathsf{O}p$, $\mathsf{O}q \vee \mathsf{P}r$, etc.

Chapter 7

Multi-agent adaptive logics for normative conflicts

✎ The content of this chapter is based on the paper *Nonmonotonic Reasoning with Normative Conflicts in Multi-Agent Deontic Logic* [28], which is coauthored by Christian Straßer.

✎ I am indebted to Joke Meheus and Christian Straßer for valuable comments on this chapter.

By now, the reader is familiar with the inconsistency-adaptive approach from the previous chapter. In this chapter, we use the inconsistency-adaptive approach for modeling interactions in a multi-agent normative setting. We start off with the presentation of a simple and elegant multi-agent logic of action, the logic **ML** (Section 7.1). Next, we extend **ML** to the deontic multi-agent logic of action **MDL**. The latter adds deontic operators to the language of **ML** and allows us to model multi-agent normative reasoning.

The logics **ML** and **MDL** are conflict-intolerant. They trivialize all premise sets containing normative conflicts, be they conflicts between (groups of) agents or conflicting directives faced by one and the same agent or group. We deal with this problem by (i) weakening **MDL** to the paraconsistent deontic multi-agent logic **PMDL**, and (ii) strengthening **PMDL** within the adaptive logics framework.

In realizing (i), we 'reconstruct' our logic on top of the paraconsistent logic **LP**, resulting in the logic **PMDL** (Section 7.3).[1] Like the monotonic paraconsistent (deontic) logics encountered before, **PMDL** is too weak to account for many intuitive and classically valid inference patterns. Hence, we use it as the LLL of a stronger adaptive extension: the logic $\mathbf{PMDL^x}$ defined and illustrated in Section 7.4.

The research presented in this chapter builds on earlier work on agentive adaptive logics from [30]. There too, we presented an inconsistency-adaptive multi-agent deontic logic. However, the system $\mathbf{PMDL^x}$ defined here improves

[1]**LP** abbreviates 'Logic of Paradox'. It was devised by Priest [143]. See [145] for more information.

on this earlier work in various ways. In Section 7.5, we compare the present approach to this earlier alternative and make some general remarks on paraconsistent and inconsistency-adaptive (deontic) logics. Moreover, we discuss its relation to some of the main paradigms in the logical study of agency.

7.1 ML, a multi-agent logic of action

7.1.1 Definition

7.1.1.1 Language and conventions

We use a finite non-empty set $I = \{i_1, \ldots, i_n\}$ of agents. Since we will in the remainder often refer to groups of agents J in I, i.e. non-empty subsets of I, the following notation is useful for this: $J \subseteq_\varnothing I$ iff $J \neq \varnothing$ and $J \subseteq I$. We also introduce the notation $J \subset_\varnothing I$ for denoting *proper* non-empty subsets J of I: $J \subset_\varnothing I$ iff $J \neq \varnothing$ and $J \subset I$. Where $J \subseteq_\varnothing I$, the set $\mathcal{W}^{\mathbf{ML}}$ of wffs of **ML** is defined recursively as follows:

$$\mathcal{W}^{\mathbf{ML}} := \langle \mathcal{W}^a \rangle \mid \neg\langle \mathcal{W}^{\mathbf{ML}} \rangle \mid \langle \mathcal{W}^{\mathbf{ML}} \rangle \vee \langle \mathcal{W}^{\mathbf{ML}} \rangle \mid \langle \mathcal{W}^{\mathbf{ML}} \rangle \wedge \langle \mathcal{W}^{\mathbf{ML}} \rangle \mid \Box_J \langle \mathcal{W}^{\mathbf{ML}} \rangle \mid \Diamond_J \langle \mathcal{W}^{\mathbf{ML}} \rangle$$

Note that we do not define the $\Diamond_J$-operators in terms of their dual $\Box_J$-operators. Instead, the diamond operators are primitive in our language. The reason for this will become clear in Section 7.3. Where $A, B \in \mathcal{W}^{\mathbf{ML}}$, we define the implication by $A \supset B =_{\mathrm{df}} \neg A \vee B$ and the equivalence relation by $A \equiv B =_{\mathrm{df}} (A \supset B) \wedge (B \supset A)$. A formula $\Box_J A$ is interpreted as "group or agent J brings about A by a joint effort". A formula $\Diamond_J A$ is interpreted (rather weakly) as "A is compatible with the (joint) actions of group or agent J" (cfr. Section 7.1.2). Where $i \in I$, we abbreviate $\Box_{\{i\}} A$ by $\Box_i A$.

Unless stated differently, we presuppose throughout this section that $A, B \in \mathcal{W}^{\mathbf{ML}}$, $\Gamma \subseteq \mathcal{W}^{\mathbf{ML}}$, and $J, K \subseteq_\varnothing I$.

7.1.1.2 Axiomatization

ML is axiomatized by adding the following axiom schemas and rules to **CL**:

$$\Box_J(A \supset B) \supset (\Box_J A \supset \Box_J B) \qquad (\mathrm{AK}\Box_J)$$

$$\Box_J A \supset \Box_J \Box_J A \qquad (\mathrm{A4}\Box_J)$$

$$\Box_J A \supset A \qquad (\mathrm{AT}\Box_J)$$

$$\Diamond_J A \equiv \neg \Box_J \neg A \qquad (\mathrm{ADf}\Diamond_J)$$

$$\text{If } \vdash A, \text{ then } \vdash \Box_J A \qquad (\mathrm{NEC}\Box_J)$$

The modal operators of **ML** are **S4**-operators. In agreement with the characterization of the $\Diamond_J$-operators as separate modal operators not defined in terms of their duals (cfr. supra), we also need (ADf$\Diamond_J$) in order to obtain the usual properties for the diamond operators.

7.1.1.3 Semantics

An **ML**-model is a tuple $\langle W, \langle R_J \rangle_{J \subseteq_\varnothing I}, v, w_0 \rangle$, where W is a set of points referred to as 'worlds', each $R_J \subseteq W \times W$ is a transitive and reflexive accessibility relation[2], $v : \mathcal{W}^a \to \wp(W)$ is an assignment function, and $w_0 \in W$ is the 'home' or 'actual' world.

Truth at a world w is defined as follows:

(Ca)	where $A \in \mathcal{W}^a$, $M, w \vDash A$ iff $w \in v(A)$
(C$\wedge$)	$M, w \vDash A \wedge B$ iff $M, w \vDash A$ and $M, w \vDash B$
(C$\vee$)	$M, w \vDash A \vee B$ iff $M, w \vDash A$ or $M, w \vDash B$
(C$\neg$)	$M, w \vDash \neg A$ iff $M, w \nvDash A$
(C$\Box_J$)	$M, w \vDash \Box_J A$ iff for all w' such that $R_J ww'$, $M, w' \vDash A$
(C$\Diamond_J$)	$M, w \vDash \Diamond_J A$ iff there is a w' such that $R_J ww'$ and $M, w' \vDash A$

An **ML**-model M verifies A ($M \vDash_{\mathbf{ML}} A$) iff $M, w_0 \vDash A$.

7.1.2 Further discussion

As mentioned above, we read $\Box_J A$ as "group or agent J brings about A by a joint effort". $\Diamond_J A$ is read as "A is compatible with J's actions", instead of the stronger "J has the ability to bring about A".[3] The reason for this weaker reading has to do with the following inferences:

$$\Diamond_J (A \vee B) \vdash_{\mathbf{ML}} \Diamond_J A \vee \Diamond_J B \tag{7.1}$$

$$A \vdash_{\mathbf{ML}} \Diamond_J A \tag{7.2}$$

Kenny noted in [103] that (7.1) and (7.2) are too strong for formalizing the 'can' of ability. (7.1) is violated by anyone who has the ability to pick a card from a pack of cards without having the ability to pick a red card or the ability to pick a black one. (7.2) is violated by any hopeless darts player who – by accident – hits the bull's eye but lacks the ability to repeat his deed [103, 159]. For this reason, we prefer our weaker reading of the $\Diamond_J$-operators.

As the modal operators of **ML** are **S4**-modalities, we can aggregate over actions:

$$\Box_J A \wedge \Box_J B \vdash_{\mathbf{ML}} \Box_J (A \wedge B) \tag{7.3}$$

The opposite direction of (7.3) also holds:

$$\Box_J (A \wedge B) \vdash_{\mathbf{ML}} \Box_J A \wedge \Box_J B \tag{7.4}$$

ML invalidates the stronger axiom schemas (A5$\Box_J$) and (AB$\Box_J$):

$$\Diamond_J A \supset \Box_J \Diamond_J A \tag{A5$\Box_J$}$$

$$A \supset \Box_J \Diamond_J A \tag{AB$\Box_J$}$$

[2] R is transitive iff, for each w, w', and w'', whenever Rww' and $Rw'w''$, also Rww''; R is reflexive iff, for each w, Rww.

[3] Note that due to (NEC$\Box_J$) we have $\Box_J A$ for all **ML**-theorems A, for every $J \subseteq_\varnothing I$. This adds a non-deliberative flavor to our $\Box_J$ operator similar to the non-deliberative character of the Chellas-stit (see e.g., [93]), cfr. infra. In view of this a more refined reading of $\Box_J A$ is "group or agent J brings about A or A is logically necessary".

A's being compatible with J's actions need not imply that J brings it about that A is compatible with his/her/its actions. Moreover, A's being the case need not imply that – for all agents and groups J – J takes care (or brings it about) that A is compatible with J's actions.

As indicated in Section 7.1.1.1, group actions are *joint* actions in **ML**. A formula $\Box_J A$ is true only if *all* members of J bring about A *together*. Where $J \subseteq_\varnothing K$:

$$\Box_J A \nvdash_{\mathbf{ML}} \Box_K A \tag{7.5}$$

$$\Diamond_J A \nvdash_{\mathbf{ML}} \Diamond_K A \tag{7.6}$$

$$\Box_K A \nvdash_{\mathbf{ML}} \Box_J A \tag{7.7}$$

$$\Diamond_K A \nvdash_{\mathbf{ML}} \Diamond_J A \tag{7.8}$$

Thus, **ML**'s agency operators do not allow for the inclusion of 'free riders' in their actions: for each action $\Box_J A$, each member of the group J is essential to J's bringing about A.[4] In the **ML**-semantics, individual agents and groups of agents each have their own (possibly disjoint) accessibility relations. From an individual's acting so-and-so, we do not gain any information about the group actions this individual takes part in.

The only constraints present on the actions of individuals and groups in **ML** is that they need to be compatible with the actions of other agents and groups, and with the facts. For all J, K:

$$A \vdash_{\mathbf{ML}} \Diamond_J A \tag{7.9}$$

$$A \vdash_{\mathbf{ML}} \neg \Box_J \neg A \tag{7.10}$$

$$\Box_J A \vdash_{\mathbf{ML}} \Diamond_K A \tag{7.11}$$

$$\Box_J A \vdash_{\mathbf{ML}} \neg \Box_K \neg A \tag{7.12}$$

Following [93], we define an agent or group's refraining from A as $\Box_J \neg \Box_J A$. Refrainment is stronger than simple non-action:

$$\Box_J \neg \Box_J A \vdash_{\mathbf{ML}} \neg \Box_J A \tag{7.13}$$

(7.13) follows immediately by (AT$\Box_J$). Its converse, however, does not hold in **ML**:

$$\neg \Box_J A \nvdash_{\mathbf{ML}} \Box_J \neg \Box_J A \tag{7.14}$$

This is as it should be: in not bringing about a state of affairs, we need not 'actively' do so. Von Wright notes, for instance, that this is especially true in situations in which acting so-and-so is beyond our capacity. For example, while it may be true that an agent does not alter the course of a tornado, it seems incorrect to say that she refrains from doing so [192].

The $\Box_J$-operator is not a 'deliberative' action operator in the sense of [96], since for instance the following not so intuitive formulas are **ML**-theorems:

$$\vdash_{\mathbf{ML}} \Box_J (A \vee \neg A) \tag{7.15}$$

[4]The concept of free riders is borrowed from [33].

$$\vdash_{\mathbf{ML}} \Box_J(\Box_J A \vee \neg \Box_J A) \tag{7.16}$$

If we were to add to **ML** a normal modal operator "$\Box$" for representing (physical) necessity and call the resulting logic **ML′**, then, in line with the literature on deliberative agency, a *deliberative* agency-operator $\triangle_J$ can be defined in **ML′** by $\triangle_J A =_{\mathrm{df}} (\Box_J A \wedge \neg \Box A)$. The analogues to (7.15) and (7.16) are invalid for this new operator:

$$\nvdash_{\mathbf{ML'}} \triangle_J(A \vee \neg A) \tag{7.17}$$
$$\nvdash_{\mathbf{ML'}} \triangle_J(\triangle_J A \vee \neg \triangle_J A) \tag{7.18}$$

For convenience, we will in the remainder continue to use the $\Box_J$-operators instead of the more involving $\triangle_J$-operators.

7.2 Adding deontic modalities: the logic MDL

7.2.1 Definition

The language $\mathcal{W}^{\mathbf{MDL}}$ of **MDL** is obtained by adding the deontic operators O and P to the language of **ML**:

$$\mathcal{W}^{\mathbf{MDL}} := \langle \mathcal{W}^{\mathbf{ML}} \rangle \mid \neg\langle \mathcal{W}^{\mathbf{MDL}} \rangle \mid \langle \mathcal{W}^{\mathbf{MDL}} \rangle \vee \langle \mathcal{W}^{\mathbf{MDL}} \rangle \mid \langle \mathcal{W}^{\mathbf{MDL}} \rangle \wedge \langle \mathcal{W}^{\mathbf{MDL}} \rangle \mid \Box_J\langle \mathcal{W}^{\mathbf{MDL}} \rangle \mid \Diamond_J\langle \mathcal{W}^{\mathbf{MDL}} \rangle \mid \mathsf{O}\langle \mathcal{W}^{\mathbf{MDL}} \rangle \mid \mathsf{P}\langle \mathcal{W}^{\mathbf{MDL}} \rangle$$

As for **ML**, we do not define the P-operator as the dual of the O-operator, but add it separately to the language of **MDL**.

Where $A \in \mathcal{W}^{\mathbf{MDL}}$, a formula $\mathsf{O}A$ is read as "it is obligatory that A". $\mathsf{P}A$ is read as "it is permitted that A".

Unless stated differently, we presuppose throughout this section that $A, B \in \mathcal{W}^{\mathbf{MDL}}$, $\Gamma \subseteq \mathcal{W}^{\mathbf{MDL}}$, and $J, K \subseteq_{\varnothing} I$.

MDL is axiomatized by adding to **ML** the axiom schemas (K), (D), (P), and the rule (NEC). In other words, **MDL** is obtained by adding to **ML** all axioms and rules of **SDL** from Section 2.2.1.

A semantical characterization of **MDL** is obtained just as easily. An **MDL**-model is a tuple $\langle W, \langle R_J \rangle_{J \subseteq_{\varnothing} I}, R_{\mathsf{O}}, v, w_0 \rangle$, where W, $\langle R_J \rangle_{J \subseteq_{\varnothing} I}$, v and w_0 are as before, and where $R_{\mathsf{O}} \subseteq W \times W$ is a serial accessibility relation. Truth at a world w is defined by adding to clauses (Ca)-(C$\Diamond_J$) from Section 7.1.1.3 the clauses (CO) and (CP):

(CO) $M, w \models \mathsf{O}A$ iff for all w' such that $R_{\mathsf{O}}ww'$, $M, w' \models A$
(CP) $M, w \models \mathsf{P}A$ iff for some w' such that $R_{\mathsf{O}}ww'$, $M, w' \models A$

As before, an **MDL**-model M verifies A ($M \models_{\mathbf{MDL}} A$) iff $M, w_0 \models A$.

7.2.2 Discussion

As the O-operator is a normal modal operator, we can aggregate over obligations:

$$\mathsf{O}A \wedge \mathsf{O}B \vdash_{\mathbf{MDL}} \mathsf{O}(A \wedge B) \tag{7.19}$$
$$\mathsf{O}\Box_J A \wedge \mathsf{O}\Box_K B \vdash_{\mathbf{MDL}} \mathsf{O}(\Box_J A \wedge \Box_K B) \tag{7.20}$$
$$\mathsf{O}\Box_J A \wedge \mathsf{O}\Box_J B \vdash_{\mathbf{MDL}} \mathsf{O}\Box_J (A \wedge B) \tag{7.21}$$

The deontic analogues of (7.5)-(7.8) remain invalid in **MDL**:

$$\mathsf{O} \,\Box_J\, A \nvdash_{\mathbf{ML}} \mathsf{O} \,\Box_K\, A \tag{7.22}$$

$$\mathsf{O} \,\Diamond_J\, A \nvdash_{\mathbf{ML}} \mathsf{O} \,\Diamond_K\, A \tag{7.23}$$

$$\mathsf{O} \,\Box_K\, A \nvdash_{\mathbf{ML}} \mathsf{O} \,\Box_J\, A \tag{7.24}$$

$$\mathsf{O} \,\Diamond_K\, A \nvdash_{\mathbf{ML}} \mathsf{O} \,\Diamond_J\, A \tag{7.25}$$

And similarly for permissions. Thus, obligations and permissions are not closed under weakening or strengthening via the addition or subtraction of agents to the group. Collective obligations of the kind interpreted by **MDL** are called *strict collective obligations* by Dignum & Royakkers [49]. A strict collective obligation to bring about A is satisfied only if *all* agents in the collective bring about A *together*.

Next to strict collective obligations, Dignum & Royakkers also define *weak collective obligations*. A weak collective obligation to bring about A is satisfied as soon as any subset of the collective brings about A. Given the language $\mathcal{W}^{\mathbf{MDL}}$, we can define an operator O^{w} for expressing weak collective obligations as follows:

$\mathsf{O}^{\mathsf{w}} \,\Box_J\, A =_{\mathrm{df}} \mathsf{O}(\bigvee_{K \subseteq_{\varnothing} J} \Box_K A)$

The weak collective obligation operator O^{w} captures the intended meaning that if it is obligatory for a group of agents to bring about a certain state of affairs, then this state of affairs ought to be brought about by some subset of the group.

Where $J \subset_{\varnothing} K$, the following weakening and strengthening properties hold for the O^{w}-operator in **MDL**:

$$\mathsf{O}^{\mathsf{w}} \,\Box_J\, A \vdash_{\mathbf{MDL}} \mathsf{O}^{\mathsf{w}} \,\Box_K\, A \tag{7.26}$$

$$\mathsf{O}^{\mathsf{w}} \,\Box_K\, A \nvdash_{\mathbf{MDL}} \mathsf{O}^{\mathsf{w}} \,\Box_J\, A \tag{7.27}$$

Another form of interaction between agents occurs when actions get nested or iterated. In line with the (literal) reading of $\Box_J A$ and $\mathsf{O}A$, we read a formula $\Box_J \mathsf{O} \,\Box_K\, A$ as "J brings it about that it is obligatory that K brings it about that A". Alternatively, we can interpret this formula as "J issues the obligation for K to bring about A".

$$\mathsf{O} \,\Box_J\, \mathsf{O} \,\Box_K\, A \nvdash_{\mathbf{MDL}} \mathsf{O} \,\Box_J\, A \tag{7.28}$$

(7.28) expresses that if it is obligatory for J to issue the obligation for K to bring about A, then it need not be obligatory for J to realize A. This is as it should be, since J might realize his/her/their duty and issue the obligation to K, without K realizing his/her/their duty to actually bring about A. Hence it is not up to J to bring about A. Thus, we cannot derive $\mathsf{O} \,\Box_J\, A$ from $\mathsf{O} \,\Box_J\, \mathsf{O} \,\Box_K\, A$.

So far, the treatment of actions, obligations, and action-obligation compounds by **MDL** seems fine. Things change, however, when we turn to more 'messy' settings in which the requirements on agents can conflict.

7.3 Dealing with normative conflicts

7.3.1 MDL and normative conflicts

In Example 9 from Section 1.4.4, Creon declares the burial of Antigone's brother Polyneices illegal on the grounds that he was a traitor to the city, and that his burial would mock the loyalists who defended the city. At the same time however, Antigone faces a religious and familial obligation to bury her brother. The conflicting obligations of Antigone and Creon to bury and not bury Polyneices can be formalized as $\mathsf{O}\,\Box_a\, B$ and $\mathsf{O}\,\Box_c\,\neg B$ respectively (where 'a' abbreviates 'Antigone', 'c' abbreviates Creon, and 'B' abbreviates the statement "Polyneices is buried"). Conflicts between obligations for different agents or groups to bring about some state of affairs are called *interpersonal conflicts* in [121, 164].

Interpersonal obligation-obligation conflicts or OO-conflicts of the kind displayed above cannot be consistently formalized in **MDL**, due to the validity of (7.29). Where $J \neq K$:

$$\mathsf{O}\,\Box_J\, A \wedge \mathsf{O}\,\Box_K\,\neg A \vdash_{\mathbf{MDL}} B \tag{7.29}$$

Similarly for interpersonal obligation-permission conflicts or OP-conflicts:

$$\mathsf{O}\,\Box_J\, A \wedge \mathsf{P}\,\Box_K\,\neg A \vdash_{\mathbf{MDL}} B \tag{7.30}$$

As has been argued extensively by moral philosophers and deontic logicians, single agents as well as groups can face (*intra-personal*) OO- or OP-conflicts (see e.g. [58, 70, 109, 203]). An adult muslim living in Western Europe might for instance be permitted to drink alcohol (by law) and forbidden to drink alcohol (by his or her religion) (cfr. Example 2). However, such situations too cause explosion when formalized in **MDL**, due to the validity of:

$$\mathsf{O}\,\Box_J\, A \wedge \mathsf{O}\,\Box_J\,\neg A \vdash_{\mathbf{MDL}} B \tag{7.31}$$

$$\mathsf{O}\,\Box_J\, A \wedge \mathsf{P}\,\Box_J\,\neg A \vdash_{\mathbf{MDL}} B \tag{7.32}$$

The same story applies to the slightly weaker inferences (7.33) and (7.34), and to 'nested' OO- or OP-conflicts:

$$\mathsf{O}\,\Box_J\, A \wedge \mathsf{O}\neg\,\Box_J\, A \vdash_{\mathbf{MDL}} B \tag{7.33}$$

$$\mathsf{O}\,\Box_J\, A \wedge \mathsf{P}\neg\,\Box_J\, A \vdash_{\mathbf{MDL}} B \tag{7.34}$$

$$\mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\, A \wedge \mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,\neg A \vdash_{\mathbf{MDL}} B \tag{7.35}$$

$$\mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\, A \wedge \mathsf{O}\,\Box_J\,\mathsf{P}\,\Box_K\,\neg A \vdash_{\mathbf{MDL}} B \tag{7.36}$$

In general, the following explosion principles are **MDL**-valid:

$$\mathsf{O}A \wedge \mathsf{O}\neg A \vdash_{\mathbf{MDL}} B \tag{7.37}$$

$$\mathsf{O}A \wedge \mathsf{P}\neg A \vdash_{\mathbf{MDL}} B \tag{7.38}$$

$$\text{If } \vdash_{\mathbf{MDL}} \neg(A_1 \wedge \ldots \wedge A_n),\ \text{then } \mathsf{O}\,\Box_{J_1}\, A_1 \wedge \ldots \wedge \mathsf{O}\,\Box_{J_n}\, A_n \vdash_{\mathbf{MDL}} B \tag{7.39}$$

$$\text{If } \vdash_{\mathbf{MDL}} \neg(A_1 \wedge \ldots \wedge A_n),\ \text{then } \mathsf{O}\,\Box_{J_1}\, A_1 \wedge \ldots \wedge \mathsf{P}\,\Box_{J_n}\, A_n \vdash_{\mathbf{MDL}} B \tag{7.40}$$

Unfortunately, real life is abundant with (inter- and intra-personal) OO- and OP-conflicts between (groups of) agents [107]. Hence we should be able to accommodate such conflicts within our logic. In Section 7.3.2, we weaken **MDL** to a logic that invalidates the explosion principles (7.37)-(7.40).

7.3.2 Enters paraconsistency: the logic PMDL

The solution adopted here for the problem of accommodating normative conflicts, is again to weaken the negation connective of **MDL** to a *paraconsistent* negation connective (cfr. Chapter 6). Below we introduce a logic that weakens $\neg$ to a paraconsistent negation connective $\sim$, namely the logic **PMDL**. **PMDL** is built on top of the propositional fragment of the paraconsistent logic **LP**.

The set $\mathcal{W}^{\mathbf{PMDL}}$ of wffs of **PMDL** is defined by replacing the connective $\neg$ of $\mathcal{W}^{\mathbf{MDL}}$ with the connective $\sim$. Where:

$$\mathcal{W}^{\mathbf{ML}}_{\sim} := \langle\mathcal{W}^a\rangle \mid \sim\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle\vee\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle\wedge\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \Box_J\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \Diamond_J\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle$$

The set $\mathcal{W}^{\mathbf{PMDL}}$ is defined by:

$$\mathcal{W}^{\mathbf{PMDL}} := \langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \sim\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \vee \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \wedge \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \Box_J\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \Diamond_J\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \mathsf{O}\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \mathsf{P}\langle\mathcal{W}^{\mathbf{PMDL}}\rangle$$

For reasons of transparency, we first characterize **PMDL** semantically. The **PMDL**-semantics differs from that of **MDL** in that (i) we broaden the domain of the assignment function v so that it includes the set of all literals $\mathcal{W}^{\sim}_l$, i.e. we define $v : \mathcal{W}^{\sim}_l \to \wp(W)$, (ii) we replace clause (C$\neg$) by (C$\sim$) and add de Morgan's laws to the semantics (clauses (C$\sim\sim$)-(C-$\sim\vee$)), and (iii) we add clauses (C$\sim\Box_J$), (C$\sim\Diamond_J$), (C$\sim$O), and (C$\sim$P) which give us the usual interrelations between dual operators. Thus, we keep clauses (Ca), (C$\wedge$), (C$\vee$), (C$\Box_J$), (C$\Diamond_J$), (CO), and (CP) and add the following:

(C$\sim$)	Where $A \in \mathcal{W}^a$, $M, w \models {\sim}A$ iff $M, w \not\models A$ or $w \in v({\sim}A)$
(C$\sim\sim$)	$M, w \models {\sim}{\sim}A$ iff $M, w \models A$
(C$\sim\wedge$)	$M, w \models {\sim}(A \wedge B)$ iff $M, w \models {\sim}A \vee {\sim}B$
(C$\sim\vee$)	$M, w \models {\sim}(A \vee B)$ iff $M, w \models {\sim}A \wedge {\sim}B$
(C$\sim\Box_J$)	$M, w \models {\sim}\Box_J A$ iff $M, w \models \Diamond_J{\sim}A$
(C$\sim\Diamond_J$)	$M, w \models {\sim}\Diamond_J A$ iff $M, w \models \Box_J{\sim}A$
(C$\sim$O)	$M, w \models {\sim}\mathsf{O}A$ iff $M, w \models \mathsf{P}{\sim}A$
(C$\sim$P)	$M, w \models {\sim}\mathsf{P}A$ iff $M, w \models \mathsf{O}{\sim}A$

As before, a **PMDL**-model M verifies A ($M \models_{\mathbf{PMDL}} A$) iff $M, w_0 \models A$.

The addition of clauses (C$\sim\Box_J$), (C$\sim\Diamond_J$), (C$\sim$O), and (C$\sim$P) is necessary in order to guarantee the interdefinability of the modal operators. If, for instance, the P-operator were simply defined as the dual of the O-operator (i.e. $\mathsf{P}A =_{\mathrm{df}} {\sim}\mathsf{O}{\sim}A$), then, due to the paraconsistency of "$\sim$" we would no longer be able to derive $\mathsf{P}{\sim}A$ from ${\sim}\mathsf{O}A$. Similarly for the $\Box_J$- and $\Diamond_J$-operators. This is why all modalities are primitive in our language, and why extra semantic clauses are added in order to guarantee their usual interrelations.

Syntactically, the negation connective of **LP** is defined by de Morgan's laws (including double negation) and excluded middle (EM):

$$A \vee {\sim}A \tag{EM}$$

Since **LP** no longer validates (ECQ), it can consistently allow for contradictions $A \wedge {\sim}A$. A consequence of this weakening is that **LP** invalidates modus ponens, due to its definition of the implication connective in terms of the disjunction and negation connectives. A full syntactical characterization of **LP** is contained in Section G.1 of the Appendix.

Where $\boxdot \in \{\mathsf{O}\} \cup \{\Box_J \mid J \subseteq_\varnothing I\}$ and $\diamonddot \in \{\mathsf{P}\} \cup \{\Diamond_J \mid J \subseteq_\varnothing I\}$, the logic **PMDL** is defined by adding to **LP** the rules (4$\Box_J$)–(T$\Diamond_J$) for every $J \subseteq_\varnothing I$, (D$\mathsf{O}$), and (AND$\boxdot$)–(INH$\diamonddot$):

$\Box_J A \vdash \Box_J \Box_J A$	(4$\Box_J$)	$\diamonddot(A \vee B) \vdash \diamonddot A \vee \diamonddot B$	(OR$\diamonddot$)
$\Diamond_J \Diamond_J A \vdash \Diamond_J A$	(4$\Diamond_J$)	$\boxdot(A \vee B) \vdash \boxdot A \vee \diamonddot B$	(OR$\boxdot$)
$\Box_J A \vdash A$	(T$\Box_J$)	$\sim \boxdot A \vdash \diamonddot{\sim} A$	(R$\sim\boxdot$)
$A \vdash \Diamond_J A$	(T$\Diamond_J$)	$\diamonddot{\sim} A \vdash \sim \boxdot A$	(R$\diamonddot\sim$)
$\mathsf{O}A \vdash \mathsf{P}A$	(DO)	$\boxdot{\sim} A \vdash \sim \diamonddot A$	(R$\boxdot\sim$)
$\boxdot A, \boxdot B \vdash \boxdot(A \wedge B)$	(AND$\boxdot$)	$\sim \diamonddot A \vdash \boxdot{\sim} A$	(R$\sim\diamonddot$)
$\boxdot A, \diamonddot B \vdash \diamonddot(A \wedge B)$	(AND$'\boxdot$)	If $A \vdash B$, then $\boxdot A \vdash \boxdot B$	(INH$\boxdot$)
		If $A \vdash B$, then $\diamonddot A \vdash \diamonddot B$	(INH$\diamonddot$)

In the case of (INH$\boxdot$) and (INH$\diamonddot$) we also allow for the case that A is the empty string, in which case we stipulate that also $\boxdot A$ resp. $\diamonddot A$ is the empty string.

Note that all of the rules of **PMDL** are **MDL**-valid (after replacing occurrences of $\sim$ with $\neg$). As we will illustrate below, **PMDL** is strictly weaker than **MDL**. The reason why the properties of **PMDL** are introduced as *rules* – and not as axiom schemas – is that the implication connective of **PMDL** is not detachable: modus ponens is invalid in **PMDL** due to its failure in **LP**. For instance, if instead of (T$\Box_J$) only its weaker variant $\Box_J A \supset A$ were to hold, then A would not be **PMDL**-derivable from $\Box_J A$ and $\Box_J A \supset A$.

Theorem 33. $\Gamma \vdash_{\mathbf{PMDL}} A$ *iff* $\Gamma \vDash_{\mathbf{PMDL}} A$.

A proof of Theorem 33 is contained in Section G.2 of the Appendix.

In accordance with the goal set out for this logic, **PMDL** tolerates all types of normative conflicts mentioned in Section 7.3.1; in other words, **PMDL** invalidates the explosion principles (7.37)-(7.40):

$$\mathsf{O}A \wedge \mathsf{O}{\sim}A \nvdash_{\mathbf{PMDL}} B \quad (7.41)$$

$$\mathsf{O}A \wedge \mathsf{P}{\sim}A \nvdash_{\mathbf{PMDL}} B \quad (7.42)$$

$$\text{If } \vdash_{\mathbf{MDL}} {\sim}(A_1 \wedge \ldots \wedge A_n), \text{ then } \mathsf{O}\,\Box_{J_1} A_1 \wedge \ldots \wedge \mathsf{O}\,\Box_{J_n} A_n \nvdash_{\mathbf{PMDL}} B \quad (7.43)$$

$$\text{If } \vdash_{\mathbf{MDL}} {\sim}(A_1 \wedge \ldots \wedge A_n), \text{ then } \mathsf{O}\,\Box_{J_1} A_1 \wedge \ldots \wedge \mathsf{P}\,\Box_{J_n} A_n \nvdash_{\mathbf{PMDL}} B \quad (7.44)$$

7.3.3 A price to pay?

Although **PMDL** provides a consistent treatment of normative conflicts, this treatment comes at a high price. Not only does **PMDL** invalidate inferences (7.37)-(7.40) (as was desired); alas it also invalidates many other – less unwanted – **MDL**-valid inferences:

$$\mathsf{O}\,\Box_J A, \mathsf{O}\,\Box_J ({\sim}A \vee B) \nvdash_{\mathbf{PMDL}} \mathsf{O}\,\Box_J B \quad (7.45)$$

$$\mathsf{P} \Box_J A \nvdash_{\mathbf{PMDL}} {\sim}\mathsf{O}{\sim} \Box_J A \tag{7.46}$$
$$\mathsf{O} \Box_J A \nvdash_{\mathbf{PMDL}} {\sim}\mathsf{P}{\sim} \Box_J A \tag{7.47}$$
$$\mathsf{O}{\sim}A, \mathsf{O}(A \vee B) \nvdash_{\mathbf{PMDL}} \mathsf{O}B \tag{7.48}$$
$$\mathsf{P}A \nvdash_{\mathbf{PMDL}} {\sim}\mathsf{O}{\sim}A \tag{7.49}$$
$$\mathsf{O}A \nvdash_{\mathbf{PMDL}} {\sim}\mathsf{P}{\sim}A \tag{7.50}$$
$$\Box_J A, \Box_J({\sim}A \vee B) \nvdash_{\mathbf{PMDL}} \Box_J B \tag{7.51}$$
$$\Diamond_J A \nvdash_{\mathbf{PMDL}} {\sim} \Box_J {\sim}A \tag{7.52}$$
$$\Box_J A \nvdash_{\mathbf{PMDL}} {\sim} \Diamond_J {\sim}A \tag{7.53}$$

In general, the disjunctive syllogism and modus ponens rules fail in **PMDL**:

$$A, {\sim}A \vee B \nvdash_{\mathbf{PMDL}} B \tag{7.54}$$
$$A, A \supset B \nvdash_{\mathbf{PMDL}} B \tag{7.55}$$

This is a very high price to pay for the conflict-tolerance of **PMDL**. **PMDL** is way too poor to account for our everyday normative and non-normative, agentive and non-agentive reasoning.

Thus **PMDL** suffers from a trade-off: its paraconsistent negation connective allows for the accommodation of normative conflicts, but it drastically weakens the logic. In Section 7.4 we propose to overcome this trade-off by non-monotonically strengthening **PMDL** within the standard format for adaptive logics. The resulting adaptive logics $\mathbf{PMDL^r}$ and $\mathbf{PMDL^m}$ interpret a given premise set 'as consistently as possible'. On the one hand, these logics allow us to defeasibly apply all **MDL**-valid inference steps on the condition that the formulas to which we apply them behave consistently. On the other hand, $\mathbf{PMDL^r}$ and $\mathbf{PMDL^m}$ remain fully conflict-tolerant.

7.4 Two inconsistency-adaptive multi-agent deontic logics

7.4.1 Intuition and definition

Let us take a look at the reasons why some intuitive applications of certain inference rules fail in **PMDL**. First, reconsider (7.45)-(7.47). Although these inferences are **PMDL**-invalid, the following hold in **PMDL**:

$$\mathsf{O} \Box_J ({\sim}A \vee B), \mathsf{O} \Box_J A \vdash_{\mathbf{PMDL}} \mathsf{O} \Box_J B \vee \mathsf{P} \Diamond_J (A \wedge {\sim}A) \tag{7.56}$$
$$\mathsf{P} \Box_J A \vdash_{\mathbf{PMDL}} {\sim}\mathsf{O}{\sim} \Box_J A \vee \mathsf{P} \Diamond_J (A \wedge {\sim}A) \tag{7.57}$$
$$\mathsf{O} \Box_J A \vdash_{\mathbf{PMDL}} {\sim}\mathsf{P}{\sim} \Box_J A \vee \mathsf{P} \Diamond_J (A \wedge {\sim}A) \tag{7.58}$$

Analogously, while (7.48)-(7.53) are **PMDL**-invalid, the following hold:

$$\mathsf{O}{\sim}A, \mathsf{O}(A \vee B) \vdash_{\mathbf{PMDL}} \mathsf{O}B \vee \mathsf{P}(A \wedge {\sim}A) \tag{7.59}$$
$$\mathsf{P}A \vdash_{\mathbf{PMDL}} {\sim}\mathsf{O}{\sim}A \vee \mathsf{P}(A \wedge {\sim}A) \tag{7.60}$$
$$\mathsf{O}A \vdash_{\mathbf{PMDL}} {\sim}\mathsf{P}{\sim}A \vee \mathsf{P}(A \wedge {\sim}A) \tag{7.61}$$
$$\Box_J A, \Box_J({\sim}A \vee B) \vdash_{\mathbf{PMDL}} \Box_J B \vee \Diamond_J(A \wedge {\sim}A) \tag{7.62}$$

$$\Diamond_J A \vdash_{\mathbf{PMDL}} {\sim}\Box_J {\sim}A \vee \Diamond_J (A \wedge {\sim}A) \tag{7.63}$$
$$\Box_J A \vdash_{\mathbf{PMDL}} {\sim}\Diamond_J {\sim}A \vee \Diamond_J (A \wedge {\sim}A) \tag{7.64}$$

Moreover, **PMDL** allows for the following 'weak' variants of modus ponens and disjunctive syllogism:

$$A, {\sim}A \vee B \vdash_{\mathbf{PMDL}} B \vee (A \wedge {\sim}A) \tag{7.65}$$
$$A, A \supset B \vdash_{\mathbf{PMDL}} B \vee (A \wedge {\sim}A) \tag{7.66}$$

Whereas (7.45)-(7.55) all fail for **PMDL**, their weaker versions (7.56)-(7.66) are **PMDL**-valid. In all of these 'weakened' cases, the discussed inferences hold in **PMDL** in disjunction with a formula that expresses some counterintuitive consequence. For (7.56)-(7.58), this is the formula $\mathsf{P}\Diamond_J (A \wedge {\sim}A)$, expressing that it is permitted that the inconsistency $A \wedge {\sim}A$ is compatible with J's actions. For (7.59)-(7.61), it is the formula $\mathsf{P}(A \wedge {\sim}A)$, expressing that the inconsistency $A \wedge {\sim}A$ is permitted. For (7.62)-(7.64), the counterintuitive alternative is the formula $\Diamond_J (A \wedge {\sim}A)$, expressing that $A \wedge {\sim}A$ is compatible with J's actions. For (7.65) and (7.66), it is the plain contradiction $A \wedge {\sim}A$. What all these counterintuitive disjuncts have in common, is that, semantically, they express that a contradiction is verified at some accessible world in every **PMDL**-model of the premises.

By now, the reader is sufficiently familiar with the adaptive logics framework to see that inferences (7.45)-(7.55) can be defeasibly applied by an adaptive logic. Such a logic should (i) use **PMDL** as its LLL, and (ii) ensure that each right-hand disjunct of the formulas derived in inferences (7.56)-(7.66) gives rise to an abnormality. This is taken care of by the logic **PMDL$^{\mathbf{x}}$**, which is defined as follows (where $i \in \{1, \ldots, n\}$):

(1) Lower limit logic: **PMDL**.
(2) Set of abnormalities: $\Omega = \{\diamond_1 \ldots \diamond_n (A \wedge {\sim}A) \mid A \in \mathcal{W}^a, \diamond_i \in \{\mathsf{P}\} \cup \{\Diamond_J \mid J \subseteq_\varnothing I\}\}$.
(3) Adaptive strategy: $x \in \{r, m\}$.

Intuitively, Ω is the set each member of which verifies an inconsistency in some accessible world in the **PMDL**-semantics.

Since our aim is to interpret a given set of premises as consistently as possible the set Ω is defined in such a way that each normative conflict gives rise to a (disjunction of) abnormalities in **PMDL**. This is illustrated in the following list. Let $A \in \mathcal{W}^a$:

$$\mathsf{O}\Box_J A \wedge \mathsf{O}\Box_K {\sim}A \vdash_{\mathbf{PMDL}} \mathsf{P}(A \wedge {\sim}A) \tag{7.67}$$
$$\mathsf{O}\Box_J A \wedge \mathsf{P}\Box_K {\sim}A \vdash_{\mathbf{PMDL}} \mathsf{P}(A \wedge {\sim}A) \tag{7.68}$$
$$\mathsf{O}\Box_J A \wedge \mathsf{O}\Box_J {\sim}A \vdash_{\mathbf{PMDL}} \mathsf{P}\Diamond_J (A \wedge {\sim}A) \tag{7.69}$$
$$\mathsf{O}\Box_J A \wedge \mathsf{P}\Box_J {\sim}A \vdash_{\mathbf{PMDL}} \mathsf{P}\Diamond_J (A \wedge {\sim}A) \tag{7.70}$$
$$\mathsf{O}\Box_J A \wedge \mathsf{O}{\sim}\Box_J A \vdash_{\mathbf{PMDL}} \mathsf{P}\Diamond_J (A \wedge {\sim}A) \tag{7.71}$$
$$\mathsf{O}\Box_J A \wedge \mathsf{P}{\sim}\Box_J A \vdash_{\mathbf{PMDL}} \mathsf{P}\Diamond_J (A \wedge {\sim}A) \tag{7.72}$$
$$\mathsf{O}A \wedge \mathsf{O}{\sim}A \vdash_{\mathbf{PMDL}} \mathsf{P}(A \wedge {\sim}A) \tag{7.73}$$
$$\mathsf{O}A \wedge \mathsf{P}{\sim}A \vdash_{\mathbf{PMDL}} \mathsf{P}(A \wedge {\sim}A) \tag{7.74}$$

Where $A \notin \mathcal{W}^a$, it is easy to see that due to the validity of de Morgan's laws the inferences in this table can be generalized to conflicts between more complex formulas. These will typically give rise to disjunctions of abnormalities. Let for instance $A = A_1 \vee A_2$. Then, for example:

$$\mathsf{O} \Box_J A, \mathsf{O} \Box_K {\sim} A \vdash_{\mathbf{PMDL}} \mathsf{P}(A_1 \wedge {\sim} A_1) \vee \mathsf{P}(A_2 \wedge {\sim} A_2) \tag{7.75}$$

If $A_1, A_2 \in \mathcal{W}^a$, then $\mathsf{P}(A_1 \wedge {\sim} A_1), \mathsf{P}(A_2 \wedge {\sim} A_2) \in \Omega$. Otherwise, $\mathsf{P}(A_1 \wedge {\sim} A_1) \vee \mathsf{P}(A_2 \wedge {\sim} A_2)$ can be further analyzed into a (longer) disjunction of abnormalities.

7.4.2 Illustrations

A first example illustrates that the disjunctive syllogism rule is applicable in $\mathbf{PMDL^x}$ inside the scope of its modal operators.

Example 23. Let $\Gamma_1 = \{\mathsf{O} \Box_J {\sim} p, \mathsf{O} \Box_J (p \vee q)\}$. The following $\mathbf{PMDL^x}$-proof from Γ_1 illustrates that $\Gamma_1 \vdash_{\mathbf{PMDL^x}} \mathsf{O} \Box_J q$:

1	$\mathsf{O} \Box_J {\sim} p$	PREM	$\varnothing$
2	$\mathsf{O} \Box_J (p \vee q)$	PREM	$\varnothing$
3	$\mathsf{O} \Box_J q$	1,2;RC	$\{\mathsf{P} \Diamond_J (p \wedge {\sim} p)\}$

The application of RC at line 3 follows in view of (7.56) above. The latter in turn follows from the **LP**-valid inference ${\sim} p, p \vee q \vdash q \vee (p \wedge {\sim} p)$ by applications of (INH$\Box_J$), (INHO), (OR$\Diamond_J$) and (ORP).

Example 24. Let $\Gamma_2 = \{\mathsf{O} \Box_J {\sim} p, \mathsf{O} \Box_J (p \vee q), \mathsf{P} \Box_J p\}$. The following $\mathbf{PMDL^x}$-proof from Γ_2 illustrates that $\Gamma_2 \nvdash_{\mathbf{PMDL^x}} \mathsf{O} \Box_J q$:

1	$\mathsf{O} \Box_J {\sim} p$	PREM	$\varnothing$
2	$\mathsf{O} \Box_J (p \vee q)$	PREM	$\varnothing$
3	$\mathsf{P} \Box_J p$	PREM	$\varnothing$
4	$\mathsf{O} \Box_J q$	1,2;RC	$\{\mathsf{P} \Diamond_J (p \wedge {\sim} p)\} \checkmark^5$
5	$\mathsf{P} \Diamond_J (p \wedge {\sim} p)$	1,3;RU	$\varnothing$

Example 24 highlights the non-monotonicity of $\mathbf{PMDL^x}$. By adding a new premise to the set Γ_1 from Example 23, the formula $\mathsf{O} \Box_J q$ is no longer $\mathbf{PMDL^x}$-derivable from the new premise set.

Example 25. Let $\Gamma_3 = \{\mathsf{O}({\sim} \Box_K p \supset \Box_J p), \mathsf{O} \Box_K {\sim} \Box_K p\}$. The following $\mathbf{PMDL^x}$-proof from Γ_3 illustrates that $\Gamma_3 \vdash_{\mathbf{PMDL^x}} \mathsf{O} \Box_J p$:

1	$\mathsf{O}({\sim} \Box_K p \supset \Box_J p)$	PREM	$\varnothing$
2	$\mathsf{O} \Box_K {\sim} \Box_K p$	PREM	$\varnothing$
3	$\mathsf{O} {\sim} \Box_K p$	2; RU	$\varnothing$
4	$\mathsf{O} \Box_J p$	1,3; RC	$\{\mathsf{P} \Diamond_K (p \wedge {\sim} p)\}$

Since $\Box_K {\sim} \Box_K p \vdash_{\mathbf{PMDL}} {\sim} \Box_K p$ (by (T$\Box_K$)), it follows by (INHO) that $\mathsf{O} \Box_K {\sim} \Box_K p \vdash_{\mathbf{PMDL}} \mathsf{O} {\sim} \Box_K p$. This justifies the application of RU at line 3. Modus ponens fails in **PMDL**, but by the weaker inference ${\sim} \Box_K p \supset \Box_J p, {\sim} \Box_K p \vdash_{\mathbf{PMDL}} \Box_J p \vee \Diamond_K (p \wedge {\sim} p)$ it follows by (INHO) and (ORP) that $\mathsf{O}({\sim} \Box_K p \supset \Box_J p), \mathsf{O} {\sim} \Box_K p \vdash_{\mathbf{PMDL}} \mathsf{O} \Box_J p \vee \mathsf{P} \Diamond_K (p \wedge {\sim} p)$. This motivates the application of RC at line 4.

Example 26. Let $\Gamma_4 = \{\mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,p, \mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,({\sim}p \vee q), \mathsf{O}\,\Box_J\,r, \mathsf{P}\,\Box_K\,{\sim}r\}$. The following $\mathbf{PMDL^x}$-proof from Γ_4 illustrates that $\Gamma_4 \vdash_{\mathbf{PMDL^x}} \mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,q$:

1	$\mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,p$	PREM	$\emptyset$
2	$\mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,({\sim}p \vee q)$	PREM	$\emptyset$
3	$\mathsf{O}\,\Box_J\,r$	PREM	$\emptyset$
4	$\mathsf{P}\,\Box_K\,{\sim}r$	PREM	$\emptyset$
5	$\mathsf{P}(r \wedge {\sim}r)$	3,4; RU	$\emptyset$
6	$\mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,q$	1,2; RC	$\{\mathsf{P}\,\Diamond_J\,\mathsf{P}\,\Diamond_K\,(p \wedge {\sim}p)\}$

Example 26 illustrates that even for non-normal premise sets $\mathbf{PMDL^x}$ often delivers a stronger consequence set than its LLL. Although an abnormality is **PMDL**-derivable from the premises, the application of RC at line 6 remains unmarked, and $\Gamma_4 \vdash_{\mathbf{PMDL^x}} \mathsf{O}\,\Box_J\,\mathsf{O}\,\Box_K\,q$.

Examples 27 and 28 below show that $\mathbf{PMDL^m}$ is slightly stronger than $\mathbf{PMDL^r}$. $\Gamma_5 \vdash_{\mathbf{PMDL^m}} \mathsf{O}\,\Box_J\,r$, whereas $\Gamma_5 \nvdash_{\mathbf{PMDL^r}} \mathsf{O}\,\Box_J\,r$.

Example 27. Let $\Gamma_5 = \{\mathsf{O}\,\Box_J\,(p \vee r), \mathsf{O}\,\Box_J\,(q \vee r), \mathsf{O}\,\Box_J\,({\sim}p \wedge {\sim}q), \mathsf{P}\,\Box_J\,(p \vee q)\}$. The following $\mathbf{PMDL^r}$-proof from Γ_5 illustrates that $\Gamma_5 \nvdash_{\mathbf{PMDL^r}} \mathsf{O}\,\Box_J\,r$:

1	$\mathsf{O}\,\Box_J\,(p \vee r)$	PREM	$\emptyset$
2	$\mathsf{O}\,\Box_J\,(q \vee r)$	PREM	$\emptyset$
3	$\mathsf{O}\,\Box_J\,({\sim}p \wedge {\sim}q)$	PREM	$\emptyset$
4	$\mathsf{P}\,\Box_J\,(p \vee q)$	PREM	$\emptyset$
5	$\mathsf{O}\,\Box_J\,r$	1,3;RC	$\{\mathsf{P}\,\Diamond_J\,(p \wedge {\sim}p)\}\checkmark^6$
6	$\mathsf{P}\,\Diamond_J\,(p \wedge {\sim}p) \vee \mathsf{P}\,\Diamond_J\,(q \wedge {\sim}q)$	3,4;RU	$\emptyset$

At stage 6 of the $\mathbf{PMDL^r}$-proof from Γ_5, $U_6(\Gamma_5) = \{\mathsf{P}\,\Diamond_J\,(p \wedge {\sim}p), \mathsf{P}\,\Diamond_J\,(q \wedge {\sim}q)\}$. Hence line 5 is marked in view of Definition 3. Since no other minimal *Dab*-formulas are **PMDL**-derivable from Γ_5, the proof cannot be extended in such a way that line 5 is unmarked. Hence $\mathsf{O}\,\Box_J\,r$ is not a final $\mathbf{PMDL^r}$-consequence of Γ_5.

Example 28. The following $\mathbf{PMDL^m}$-proof from Γ_5 illustrates that $\Gamma_5 \vdash_{\mathbf{PMDL^m}} \mathsf{O}\,\Box_J\,r$:

1	$\mathsf{O}\,\Box_J\,(p \vee r)$	PREM	$\emptyset$
2	$\mathsf{O}\,\Box_J\,(q \vee r)$	PREM	$\emptyset$
3	$\mathsf{O}\,\Box_J\,({\sim}p \wedge {\sim}q)$	PREM	$\emptyset$
4	$\mathsf{P}\,\Box_J\,(p \vee q)$	PREM	$\emptyset$
5	$\mathsf{O}\,\Box_J\,r$	1,3;RC	$\{\mathsf{P}\,\Diamond_J\,(p \wedge {\sim}p)\}$
6	$\mathsf{P}\,\Diamond_J\,(p \wedge {\sim}p) \vee \mathsf{P}\,\Diamond_J\,(q \wedge {\sim}q)$	3,4;RU	$\emptyset$
7	$\mathsf{O}\,\Box_J\,r$	2,3;RC	$\{\mathsf{P}\,\Diamond_J\,(q \wedge {\sim}q)\}$

At stage 7 of the $\mathbf{PMDL^m}$-proof from Γ_5, $\Phi_7(\Gamma_5) = \{\{\mathsf{P}\,\Diamond_J\,(p \wedge {\sim}p)\}, \{\mathsf{P}\,\Diamond_J\,(q \wedge {\sim}q)\}\}$. By Definition 11, lines 5 and 7 remain unmarked at this stage. Since the formula derived at line 7 is the only minimal *Dab*-consequence of Γ_5, $\mathsf{O}\,\Box_J\,r$ is finally $\mathbf{PMDL^m}$-derivable from Γ_5.

7.4.3 Meta-theoretical properties

Due to Theorem 8 and its definition within the standard format for ALs, $\mathbf{PMDL^x}$ is sound and complete with respect to its semantics:

Corollary 10. $\Gamma \vdash_{\mathbf{PMDL^x}} A$ *iff* $\Gamma \vDash_{\mathbf{PMDL^x}} A$.

The upper limit logic **UPMDL** of **PMDL$^\mathbf{x}$** is obtained by adding to **PMDL** the rule (UPMDL), which trivializes all abnormalities in Ω. Where $A \in \mathcal{W}^a, B \in \mathcal{W}^{\mathbf{PMDL}}$ and, for all $i \in \{1, \ldots, n\}$, $\diamond_i \in \{\mathsf{P}\} \cup \{\Diamond_J \mid J \subseteq_\varnothing I\}$:

$$\diamond_1 \ldots \diamond_n (A \wedge {\sim}A) \vdash B \qquad \text{(UPMDL)}$$

Due to the definition of **PMDL$^\mathbf{x}$** within the standard format for adaptive logics it follows by Theorems 7 and 18 that:

Corollary 11. $Cn_{\mathbf{PMDL}}(\Gamma) \subseteq Cn_{\mathbf{PMDL^r}}(\Gamma) \subseteq Cn_{\mathbf{PMDL^m}}(\Gamma) \subseteq Cn_{\mathbf{UPMDL}}(\Gamma)$.

If Γ is *normal*, i.e. if Γ has no *Dab*-consequences, then, by Theorem 19:

Corollary 12. *If* Γ *is normal, then* $Cn_{\mathbf{PMDL^x}}(\Gamma) = Cn_{\mathbf{UPMDL}}(\Gamma)$.

Like the upper limit **UPD** of the logic **DP$^\mathbf{x}$** from Section 6.1, **UPMDL** trivializes contradictions, thus promoting "~" to a fully classical negation connective. We can even show that **UPMDL** is just **MDL** in disguise. Where $\Gamma \subseteq \mathcal{W}^{\mathbf{PMDL}}$, define $\Gamma^\neg$ by replacing every $A \in \Gamma$ by $\pi(A)$, where $\pi(A)$ is the result of replacing every occurrence of "~" in A by "¬". Then:

Theorem 34. $\Gamma \vdash_{\mathbf{UPMDL}} A$ *iff* $\Gamma^\neg \vdash_{\mathbf{MDL}} \pi(A)$.

A proof outline of Theorem 34 is contained in Section G.3 of the Appendix.

7.5 Related work

7.5.1 Paraconsistent logic

Apart from its capability to represent actions, a major difference between the logic **PMDL$^\mathbf{x}$** presented in this chapter and the logic **DP$^\mathbf{x}$** presented in Chapter 6 is that the former is built 'on top' of the paraconsistent logic **LP** whereas the latter is built 'on top' of the paraconsistent logic **CLuNs$^\perp$**.

The main difference between **LP** and **CLuNs$^\perp$** is that the former does not feature a detachable implication whereas the latter does (since it is an extension of **CL$^\mathbf{pos}$**). As a result, modus ponens holds unconditionally in **DP$^\mathbf{x}$**, whereas it holds only conditionally in **PMDL$^\mathbf{x}$**.

The main advantage of having a non-detachable implication is that it provides a better isolation of normative conflicts. For instance, from $\mathsf{O}A, \mathsf{O}{\sim}A$, and $\mathsf{O}(A \supset B)$ we cannot derive $\mathsf{O}B$ by means of **PMDL$^\mathbf{x}$**, but we can derive $\mathsf{O}B$ by means of **DP$^\mathbf{x}$**. The main disadvantages are (i) that we need to use the more involving conditional rule for applying modus ponens in unproblematic cases, and (ii) that we lose expressive power by not having an implication connective not definable in terms of the other connectives in the language.

We remain indifferent as to which approach is best followed in which context. It suffices to recognize that adaptive logics can be constructed on top of both paraconsistent logics with a detachable implications and paraconsistent logics with a non-detachable implication, as we have illustrated in this chapter and the preceding one.

A related matter concerns the (unconditional) validity of de Morgan's laws in paraconsistent logics underlying the definition of a system of deontic logic. All of the paraconsistent logics defined in Chapters 6 and 7 validate these laws for their paraconsistent negation connectives. However, there are paraconsistent logics which do not feature them, e.g. the logic **CLuN** defined in Appendix C.

Here too, it is easily checked that adaptive logics can be constructed on top of both paraconsistent logics with and without de Morgan's laws for their respective negation connectives (an example is the logic **ACLuN1** from e.g. [15]). And here too, we remain open to the possibility of fruitfully implementing logics without de Morgan's laws for certain normative contexts of application. A possible motivation for such logics could be that conflicts are better isolated in case we abstain from applying de Morgan's laws to conflicting information.

7.5.2 Multi-agent adaptive deontic logic

As mentioned in the introduction, this paper builds on earlier work on agentive adaptive deontic logics. More specifically, it continues the task set out in [30] of constructing a multi-agent adaptive deontic logic capable of tolerating normative conflicts. The system $\mathbf{PMDL^x}$ differs from the semantics defined in [30] in various ways.

First, $\mathbf{PMDL^x}$ is built 'on top' of the paraconsistent logic **LP**, whereas the logic $\mathbf{MDP^m}$ from [30] is built 'on top' of the paraconsistent logic $\mathbf{CLuNs}^{\perp}$.

Second, the language of $\mathbf{PMDL^x}$ has no restrictions whatsoever on nested modal operators. This flexibility makes it easier to extend the language in various ways by adding extra modalities for representing e.g. knowledge, beliefs, commitments of agents and groups.

Third, as opposed to $\mathbf{MDP^m}$, $\mathbf{PMDL^x}$ does not allow for distribution over disjunctive actions:

$$\Box_J(A \vee B) \nvdash_{\mathbf{PMDL^x}} \Box_J A \vee \Box_J B \tag{7.76}$$

Suppose, for instance, that an agent flips a coin. In doing so she guarantees that either heads or teals will be the outcome, but she cannot determine the exact outcome of the flip. Hence she does not bring it about that heads is the outcome or bring it about that tails is the outcome.

Fourth, $\mathbf{PMDL^x}$ is equipped with an adaptive proof theory, whereas $\mathbf{MDP^m}$ was only characterized semantically. Moreover, unlike $\mathbf{MDP^m}$, $\mathbf{PMDL^x}$ has a regular Kripke-semantics. Altogether, this makes $\mathbf{PMDL^x}$ the first Kripke-style agentive adaptive (deontic) logic.

7.5.3 Logics of action and *stit*-logic

The logics presented in this paper are not defined within one of the two 'main' paradigms for representing actions in (deontic) logic, i.e. *stit*-logic [33, 34, 93, 107] and dynamic logic [39, 132].[5] Nonetheless, our $\Box_J$ operators resemble in some

[5]We paradigmatically consider the logic **MDL** as representative for our notion of agency defended in this section. All arguments below are equally valid for the logics **PMDL** and $\mathbf{PMDL^x}$.

respects the Chellas *stit* or *cstit* operators used in *stit* logic. In our framework, a formula $\Box_J A$ is interpreted as "J brings about A". In *stit*-logic, a formula $[J\ stit : A]$ is interpreted as "J sees to it that A". On both accounts, A is a state of affairs, and not an action nominal as is the case in e.g. dynamic logic. Moreover, the notions of refrainment and deliberative agency as defined in Section 7.1.2 are analogous to those of *stit* logic.

A first major difference between the logics defined here and *stit*-logics is that the *stit*-framework is temporal/prospective, while we work in an atemporal setting. It is a question for future research to extend the framework given here with the ability to reason about future (and maybe past) states.

A second difference between both approaches is that the $\Box_J$ operators defined here are **S4**-modalities, while *cstit* operators – their analogues in *stit* logic – are **S5**-modalities. Thus, in **MDL** the (5_J) schema is invalid:

$$\Diamond_J A \supset \Box_J \Diamond_J A \tag{5_J}$$

Note that if (5_J) were valid, then the 'Brouwerian' schema (B_J) too would hold for our agentive operators:

$$A \supset \Box_J \Diamond_J A \tag{B_J}$$

Intuitively, (B_J) requires that if A is the case, then all agents guarantee that A is compatible with their actions. This is a very strong requirement. If A is indeed the case, then *normally* we try to act on this fact as much as possible. But there are exceptions. We might, for instance, not know that A is the case, we might not be aware of it etc. In such cases, A need not be compatible with our actions. Therefore we opted to leave (B_J) (and, consequently, (5_J)) out of our axiomatization.[6]

A third difference worth pointing out is that our systems differ from *stit* logics in their treatment of collective actions and obligations. In *stit* logic, operators for agency are closed under 'weakening' by the addition of further agents: If $J \subset K \subseteq_\varnothing I$, then if J sees to it that A, then K sees to it that A. As illustrated in (7.5) and (7.6), this kind of weakening is invalid in the logics defined here. Consequently, a statement like (7.77) is **ML**- and **MDL**-consistent, while its *stit* analogue would cause explosion:[7]

$$\Box_i A \wedge \Box_j A \wedge \neg \Box_{\{i,j\}} A \tag{7.77}$$

Let us further illustrate this property by generalizing it to the deontic setting. Suppose that two agents i and j are divorced and that they work for the same company. Then we can imagine that, when faced with a certain task A, it makes sense for the boss k to issue the following obligations:

$$\Box_k(\mathsf{O}(\Box_i A \vee \Box_j A) \wedge \neg \mathsf{P} \Box_{\{i,j\}} A) \tag{7.78}$$

[6] A very welcome consequence of not having (5_J) is that – as opposed to refrainment for the *cstit*-operator – refrainment for the $\Box_J$-operator does not collapse into simple non-action: $\Box_J \neg \Box_J A \vdash_{\mathbf{MDL}} \neg \Box_J A$, but $\neg \Box_J A \nvdash_{\mathbf{MDL}} \Box_J \neg \Box_J A$.

[7] A notable exception here is the *sstit* or 'strictly sees to it that' operator for joint agency as defined in [33].

Thus, one of i and j should bring about A, but they should not do it together (because since the divorce they are no longer on speaking terms).

Altogether, these differences motivate our approach as a pursuit-worthy alternative for existing logics of action.

Chapter 8

Concluding remarks

Forget your perfect offering
There is a crack in everything
That's how the light gets in

Leonard Cohen

A world without conflicts is a highly idealized world. Unlike such a perfect world, real life is messy and full of conflicts. In modeling the structure of human reasoning, we need formalisms that can account for life's sometimes disorderly nature. Such formalisms should try to capture an agent's reasoning processes not only in the absence, but also in the presence of conflicting information.

The adaptive deontic logics presented in this thesis aim to show that we can have our cake and eat it too. In the presence of normative conflicts, these logics allow us to distinguish between sensible and insensible applications of inference rules that are unrestrictedly valid in a conflict-free setting. On the one hand, these systems are sufficiently conflict-tolerant given their intended context of application. On the other hand, they account for all inferences that **SDL** would account for as long as the premises to which the inference is applied are untainted by some normative conflict.

In this concluding chapter, I recapitulate the main merits of the logics presented in the preceding chapters (Section 8.1), and glimpse beyond with some suggestions for future work (Section 8.2).

8.1 Merits of this thesis

In chapters 4-7, five adaptive systems for dealing with normative conflicts in deontic logic were presented and discussed. All of these meet the design requirements proposed in Section 3.3.

First, the lower limit logics of these adaptive systems are sufficiently conflict-tolerant given their intended context of application:

- $\mathbf{P}_\Diamond$ accommodates conflicting moral requirements as characterized by Bernard Williams,

- **P2** tolerates $\mathsf{O}\mathsf{O}$-conflicts arising from possibly different normative standards; alternatively, it tolerates conflicting prima facie obligations,

- **DP** tolerates $\mathsf{O}\mathsf{O}$-conflicts, $\mathsf{O}\mathsf{P}$-conflicts as well as contradictory prescriptions, be they obligations or permissions,

- **LNP** tolerates $\mathsf{O}\mathsf{O}$-conflicts and $\mathsf{O}\mathsf{P}$-conflicts between norm-propositions,

- and **PMDL** tolerates intra-personal and inter-personal $\mathsf{O}\mathsf{O}$-conflicts, $\mathsf{O}\mathsf{P}$-conflicts and contradictory obligations and permissions in a multi-agent setting.

None of these logics validate any of the explosion principles stated in Section 3.3.1 for any of the types of conflicts which they are tolerant of.

Second, all of the systems $\mathbf{P}^{\mathbf{x}}_{\Diamond}$, $\mathbf{P2.2}^{\mathbf{x}}$, $\mathbf{DP}^{\mathbf{x}}$, $\mathbf{LNP}^{\mathbf{x}}$ and $\mathbf{PMDL}^{\mathbf{x}}$ are non-monotonic and allow for the conditional application of any **SDL**-valid inference. For instance, none of these logics unconditionally allows for the application of the deontic disjunctive syllogism (DDS) schema in an adaptive proof. However, each of them allows for its conditional application, thus validating all unproblematic instances of this inference rule.

Third, the reasoning processes underlying the application of the adaptive logics presented here are fully explicable in terms of the logics themselves. No tailoring the premises is required for applying the conditional rule in an adaptive proof. For instance, no intervention from outside is required for aggregating two or more obligations in the logics $\mathbf{P}^{\mathbf{x}}_{\Diamond}$ and $\mathbf{P2.2}^{\mathbf{x}}$. The localization of conflicts and the check for the applicability of certain inferences is fully taken care of by the logics.

As mentioned in Section 1.6.2, the philosophical relevance of this thesis lies in its raising and addressing some new questions concerning the structure of normative conflicts. I brought up the following:

(i) What types of normative conflicts are particularly important under which circumstances? Are there contexts in which certain types of normative conflicts can be ignored?

(ii) To what extent should normative conflicts be isolated in deontic logics? Which rules of inference are applicable to conflicting norms?

(iii) Given the possibility of conflicting norms, which inferences should hold unrestrictedly in a conflict-tolerant deontic logic? Which inferences should be restricted? Which inferences should not be valid under any condition?

The very raising of these questions has philosophical importance. Although I did not aim to answer each and every one of them, I have at some places in this thesis provided partial and tentative answers.

In reply to (i) it is clear from Section 3.4 that in different contexts we want to focus our attention on different types of conflicts. For instance, when dealing with moral norms, philosophers have traditionally focused on conflicting obligations. As discussed in Chapter 5, these can be formalized directly as $\mathsf{O}\mathsf{O}$-conflicts or indirectly as formulas of the form $\mathsf{O}A, \mathsf{O}B, \neg \Diamond (A \wedge B)$. Since permissions do not bear any sense of 'moral urgency', it seems safe to ignore these in the formal study of e.g. moral dilemmas.

The situation is different in a legal context, where OP-conflicts should also be accounted for (cfr. Section 6.2), and in the context of conflicting commands or imperatives, where, apart from OO-conflicts, we may also want to accommodate OP-conflicts as well as contradictory obligations and permissions (cfr. Section 6.1).

The questions in (ii) remain largely unanswered. In Section 7.5.1 I addressed the differences between the logics $\mathbf{CLuNs}^{\perp}$ and $\mathbf{LP}$ and their consequences for the inconsistency-adaptive logics $\mathbf{DP^x}$ and $\mathbf{PMDL^x}$ respectively. For instance, $\mathsf{O}A, \mathsf{O}{\sim}A, \mathsf{O}(A \supset B) \vdash \mathsf{O}B$ is valid in $\mathbf{DP^x}$, but not in $\mathbf{PMDL^x}$.

Similarly, we may wonder whether an inference like $\mathsf{O}A, \mathsf{O}\neg A \vdash \mathsf{O}(A \vee B)$ should or should not be invalidated by a logic that accommodates OO-conflicts (remember from Section 2.4.2 that even the inference $\mathsf{O}A \vdash \mathsf{O}(A \vee B)$ is contested). All adaptive logics presented here validate this inference, but there are conflict-tolerant adaptive deontic logics which invalidate the inference, e.g. the **ADPM**-systems from [175].

My answer to the questions posed in (iii) is again relative to the normative context at hand. In the context of moral obligations, Williams argued against the aggregation principle (AND). As discussed in Section 3.2.2.1, this answer is a bit too harsh. (AND) need not be rejected. It suffices to restrict its application to those instances in which none of the obligations to be conjoined is involved in a conflict.

In the context of commands or imperatives, it is less clear why (AND) should be restricted for obligations arising from one and the same source, even if the obligations are conflicting. Here, it is in my opinion justified to defend the unrestricted applicability of (AND), at least in the a-temporal setting assumed in this thesis. Finally, in answer to the last question it is quite clear that, given its intended context of application, a deontic logic should be sufficiently conflict-tolerant, i.e. it should invalidate all explosion principles relating to those types of conflicts that it aims to accommodate.

8.2 Further work

As is clear from the (partial) answers to the questions raised in (i)-(iii) in the previous section, a lot of work remains to be done. Here, I add two more roads for further research that seem particularly promising.

- In the introduction to this thesis I mentioned that my focus is very narrow. If we want to do justice to the complex structure of the world, we will need to extend the results from the previous section to languages with more expressive power. A first attempt in this direction was undertaken in Chapter 7, where we added indices for representing (groups of) agents and modalities for representing the actions of agents.

  As an illustration of the type of further strengthenings that can be realized within the adaptive logics framework, consider the extension of the logic $\mathbf{P2.2^x}$ to a logic for prioritized normative reasoning from [182], where obligations and permissions are equipped with an index expressing their priority, and norms with higher priority overrule norms of lower priorities.

Another extension that seems particularly interesting in relation to the topic of normative conflicts, is to use dyadic operators for representing conditional obligations. In doing so, we might be able to model cases in which conflicts are resolved by giving priority to the obligations that are most 'specific' given the situation at hand. A promising approach here might be to combine the logics defined here with the adaptive approach for modeling the detachment of conditional obligations from [173].

- Deontic logic is not the only branch of modal logic in which the accommodation of conflicts bears practical and philosophical interest. As for obligations, it is possible for our desires, beliefs and intentions to be in conflict. Moreover, a desire may be incompatible with an agent's intention, an obligation may be incompatible with an agent's belief, etc.

  The modal operators of the logics $\mathbf{P}_{\Diamond}$, **P2**, **DP**, **LNP** and **PMDL** presented here can all be adjusted by adding axiom schemas (and corresponding constraints on their accessibility relations) for properties such as reflexivity, transitivity etc. In this way, these operators can be reinterpreted as representing our beliefs, desires, intentions and so on. Using the adaptive logics framework, we can then use the resulting systems as the lower limit of new conflict-tolerant adaptive logics. As such, insights from this thesis can be applied for the accommodation of other types of conflicts in modal logic.

Appendix A

A list of normative conflicts

Here is a list of examples of normative conflicts (tragic ones, real-life ones, (hypothetical) toy examples etc.) in no particular order. Some of these examples are only 'prima facie' conflicting, others are more severe.

1. Suppose that someone, Jones, ought to visit his daughter Abby at a certain time and in preparation for that, notify her he is coming. But it could also be that Jones ought also to visit his daughter Beth at that same time and notify her he is coming. However, since Abby and Beth live on opposite sides of the country, it is impossible for Jones to visit both daughters at that time. Thus he faces a deontic dilemma ([69, p. 468], [95, p. 581]).

2. A friend leaves you with his gun saying he will be back for it in the evening, and you promise to return it when he calls. He arrives in a distraught condition, demands his gun, and announces he is going to shoot his wife because she has been unfaithful. You ought to return the gun, since you promised to do so – a case of obligation. And yet you ought not to do so, since to do so would be to be indirectly responsible for a murder, and your moral principles are such that you regard this as wrong [109, p. 148].[1]

3. Morty promises to meet a friend at the station by 3 pm. On his way there, he sees a seriously injured child in an alley; and helping the child will make Morty late. Morty ought to help children in need, but he also ought to keep his promises. So it seems that Morty ought to help the child *and* be at the station by 3 pm, even if he cannot do both [140, p. 489].[2]

4. A French student during WWII ought to join the Resistance and fight against the Nazis to liberate his country, but he also ought to remain at home to care for his mother, and he cannot do both [70, p. 455].[3]

[1] This is a variant of Plato's classic case (Republic 331c) of a person who ought to return a borrowed weapon (because he promised to do so), and who ought not to return it (because the lender has become insane). The example also appears in [70, 121, 154].

[2] This example varies on an example of Ross that illustrates the possibility of conflicting *prima facie* obligations [153, pp. 17-18].

[3] This example also appears in [122, 95, 109, 154, 164]. It is often referred to as 'Sartre's student' after the original formulation of the example by Sartre in [155].

5. If the law requires Eleanor to report Franks marijuana use to the police, then presumably Eleanor ought to report that, and if common decency requires that she not report Franks marijuana use, then she ought not to. For such a situation Eleanor seems to have a genuine conflict [70, p. 456].

6. A body of law, or the rules of a university, say, might easily demand one thing of a person, e.g., that Daniel park his car overnight in Lot A, and also demand the opposite, e.g., that Daniel not park overnight in Lot A [70, p. 456].

7. Jephthah had vowed to God, permissibly according to Mosaic law as he understood it, that if he should be granted victory over the Ammonites he would, on his return, offer as a burnt sacrifice the first living creature that should leave his doors to greet him. On his return after winning the victory, his daughter was the first living creature to leave his doors to greet him. By Mosaic law as Jephthah understood it, he was morally bound, on one hand, not to break his vow, and on the other not to commit murder – that is, not to kill the innocent: in other words, to kill his daughter and not to kill her ([51, p. 13], [138, pp. 47-48]).

8. In 1842, a ship struck an iceberg and more than 30 survivors were crowded into a lifeboat intended to hold 7. As a storm threatened, it became obvious that the lifeboat would have to be lightened if anyone were to survive. The captain reasoned that the right thing to do in this situation was to force some individuals to go over the side and drown. Such an action, he reasoned, was not unjust to those thrown overboard, for they would have drowned anyway. If he did nothing, however, he would be responsible for the deaths of those whom he could have saved. Some people opposed the captain's decision. They claimed that if nothing were done and everyone died as a result, no one would be responsible for these deaths. On the other hand, if the captain attempted to save some, he could do so only by killing others and their deaths would be his responsibility; this would be worse than doing nothing and letting all die. The captain rejected this reasoning. Since the only possibility for rescue required great efforts of rowing, the captain decided that the weakest would have to be sacrificed. In this situation it would be absurd, he thought, to decide by drawing lots who should be thrown overboard. As it turned out, after days of hard rowing, the survivors were rescued and the captain was tried for his action ([73, pp. 7-8]; shorter version in [60]).

9. You are an inmate in a concentration camp. A sadistic guard is about to hang your son who tried to escape and wants you to pull the chair from underneath him. He says that if you don't he will not only kill your son but some other innocent inmate as well. You don't have any doubt that he means what he says [73, p. 8].

10. A fat man leading a group of people out of a cave on a coast is stuck in the mouth of that cave. In a very short time high tide will be upon them, and unless he is promptly unstuck, they will all be drowned except the fat man,

whose head is out of the cave. But, fortunately, or unfortunately, someone has with him a stick of dynamite. There seems no way to get the fat man loose from the opening without using that dynamite which will inevitably kill him; but if they do not use it everyone will drown [73, p. 8].

11. In Victor Hugo's Les Miserables, the hero, Jean Valjean, is an ex-convict, living illegally under an assumed name and wanted for a robbery he committed many years ago. [Actually, no – he is only wanted for breaking parole.] Although he will be returned to the galleys – probably [in fact, actually] for life – if he is caught, he is a good man who does not deserve to be punished. He has established himself in a town, becoming mayor and a public benefactor. One day, Jean learns that another man, a vagabond, has been arrested for a minor crime and identified as Jean Valjean. Jean is first tempted to remain quiet, reasoning to himself that since he had nothing to do with the false identification of this hapless vagabond, he has no obligation to save him. Perhaps this man's false identification, Jean reflects, is "an act of Providence meant to save me." Upon reflection, however, Jean judges such reasoning "monstrous and hypocritical." He now feels certain that it is his duty to reveal his identity, regardless of the disastrous personal consequences. His resolve is disturbed, however, as he reflects on the irreparable harm his return to the galleys will mean to so many people who depend upon him for their livelihood – especially troubling in the case of a helpless woman and her small child to whom he feels a special obligation. He now reproaches himself for being too selfish, for thinking only of his own conscience and not of others. The right thing to do, he now claims to himself, is to remain quiet, to continue making money and using it to help others. The vagabond, he comforts himself, is not a worthy person, anyway. Still unconvinced and tormented by the need to decide, Jean goes to the trial and confesses [73, pp. 8-9].

12. Roger Smith, a quite competent swimmer, is out for a leisurely stroll. During the course of his walk he passes by a deserted pier from which a teenage boy who apparently cannot swim has fallen into the water. The boy is screaming for help. Smith recognizes that there is absolutely no danger to himself if he jumps in to save the boy; he could easily succeed if he tried. Nevertheless, he chooses to ignore the boy's cries. The water is cold and he is afraid of catching a cold – he doesn't want to get his good clothes wet either. "Why should I inconvenience myself for this kid," Smith says to himself, and passes on [73, p. 9].

13. Suppose some terrorists poison the water supply of a large city. Liz is an official who can prevent anyone from being killed, but only by torturing the child of a terrorist in order to get the terrorist to tell her what and where the poison is [164, p. 44].[4]

14. You are a psychiatrist and your patient has just confided to you that he intends to kill a woman. You're inclined to dismiss the threat as idle, but

[4] A variant of this conflict is given in [73].

you aren't sure. Should you report the threat to the police and the woman or should you remain silent as the principle of confidentiality between psychiatrist and patient demands [73, pp. 10-11]?[5]

15. Physicians and families who believe that human life should not be deliberately shortened and that unpreventable pain should not be tolerated face a conflict in deciding whether to withdraw life support from a dying patient [123].

16. In *Sophie's Choice*, a novel by William Styron, Sophie arrives with her two children at a Nazi concentration camp. A guard asks her to choose one child, and he tells her that the child she chooses will be killed, and the other child will live in the children's barracks. Sophie does not want to choose at all, but the guard tells her that, if she refuses to choose, both children will be killed [164, p. 54].[6]

17. A friend of yours has confided to you that he has committed a particular crime and you have promised never to tell. Discovering that an innocent man has been accused of your friend's crime, you plead with the latter to give himself up to the authorities. He refuses and reminds you of your promise. What should you do [73, p. 12]?

18. In Sophocles' *Antigone*, Creon declares the burial of Antigone's brother Polyneices illegal on the not unreasonable grounds that he was a traitor to the city and that his burial would mock the loyalists who defended the city, thereby causing civil disorder. At the same time, there is reason for Creon to respect the religious and familial obligation of Antigone to bury her brother [72, p. 4].[7]

19. In Shakespeare's *Julius Caesar*, Brutus defends the slaying of Caesar, his friend but ambitious leader, as follows: "not that I loved Caesar less, but that I loved Rome more"[72, p. 4].

20. In Shaw's *Major Barbara*, the main character, Barbara, has to choose between discontinuing her efforts on behalf of the bodily and spiritual salvation of the poor and accepting donations that have their origin in profits of a liquor and a munitions manufacturer [72, p. 4].

21. In 'work and role dilemma's', one's employment calls for activities which are morally repugnant (such as deception, or participation in the production of

[5] A legal variant of this conflict is given in [123]: "The criminal defense attorney is said to have an obligation to hold in confidence the disclosures made by a client and to be required to conduct herself with candor before the court (where the latter requires that the attorney inform the court when her client commits perjury)."

[6] For alternative formulations of this conflict, see e.g. [108, 123].

[7] This example is discussed in detail in [44, pp. 26-31], where Castañeda takes it to be a case of conflicting normative standards (legal vs. religious obligation), and where the religious obligation ultimately overrides the legal obligation; the example also appears in [108, 121, 164]. Both Marcus and Sinnott-Armstrong refer to this example as an 'interpersonal' dilemma.

nuclear or chemical weapons). These situations may arise either contractually, e.g. one signed up for the position, or from duty, e.g. to provide for one's family [154, p. 661].[8]

22. Suppose that it is forbidden to kill one's parents and forbidden to allow them to die. A dilemma would arise in a situation in which unless one kills one's mother, she will kill one's father. In such a situation it would be forbidden to kill one's mother, but also forbidden to do anything else (since that would allow one's father to die) [180, p. 114].

23. In a co-operative industrial association, is it just or not that talent or skill should give a title to superior remuneration? On the negative side it is argued, that whoever does the best he can, deserves equally well, and ought not in justice to be put in a position of inferiority for no fault of his own; that superior abilities have already advantages more than enough, in the admiration they excite, the personal influence they command, and the internal sources of satisfaction attending them, without adding to these a superior share of the world's goods; and that society is bound in justice rather to make compensation to the less favoured, for this unmerited inequality of advantages, than to aggravate it. On the contrary side it is contended, that society receives more from the more efficient labourer; that his services being more useful, society owes him a larger return for them; that a greater share of the joint result is actually his work, and not to allow his claim to it is a kind of robbery; that if he is only to receive as much as others, he can only be justly required to produce as much, and to give a smaller amount of time and exertion, proportioned to his superior efficiency. Who shall decide between these appeals to conflicting principles of justice [137, pp. 253-254]?

24. In Shakespeare's *Measure for Measure*, Angelo, the deputy of the duke of Vienna, condemns to death one of his subjects, Claudio, for the crime of lechery. Isabella, Claudio's sister, goes to plead for her brother's life. She is a devout worshipper and a nun. Angelo tells her that he will free her brother only on the condition that she will sleep with him. As a sister and one devoted to her family, Isabella believes that she must do what is in her power to save her brother's life. As a nun, however, she is morally committed to preserving her virginity. Whatever she does, she believes that she will be doing something wrong [72, p. 159] (reprinted version of the original [122]).[9]

25. In his *Two Cheers for Democracy*, E.M. Forster wrote "if I had to choose between betraying my country and betraying my friend, hope I should have

[8] Also relevant here is Lemmon's statement that "Duty conflicts with principle every time that we are called on in our jobs to do things which we find morally repugnant" [109, p. 150].

[9] Sinnott-Armstrong adds an extra dimension to the example by discussing it as a case of ignorance: "The Duke has returned to Vienna, and at one point he stands right next to Isabella. If she tells him her story, he will let her brother go. But she does not recognize the Duke, since he is disguised. Thus, she has the opportunity and physical ability to save her brother without breaking her vow, but she lacks the necessary factual knowledge" [164, p. 27].

the courage to betray my country". Marcus considers the following remark of (the fictional) A.B. Worster: "if I had to choose between betraying my country and betraying my friend, hope I should have the courage to betray my friend"[121].

26. Agamemnon is told by a seer that he must sacrifice his daughter to satisfy a goddess who is delaying at Aulis his expedition against Troy. As a commander, Agamemnon ought to sacrifice his daughter in order to further the expedition. However, as a father, Agamemnon ought not to kill his daughter [203].[10]

27. According to his religious beliefs, Yilmaz is prohibited to drink alcohol. However, according to the laws of his country, he is permitted to drink alcohol [129].

28. SWIFT is a Belgium-based company with offices in the United States that operates a worldwide messaging system used to transmit, inter alia, bank transaction information. According to the U.S. Treasury, information derived from the use of SWIFT data has enhanced the United States and third countries ability to identify financiers of terrorism, to map terrorist networks and to disrupt the activities of terrorists and their supporters. However, in September 2006 the Belgian Data Protection Authority stated that SWIFT processing activities for the execution of interbank payments are in breach of Belgian data protection law. American diplomats and politicians claim that SWIFT ought to continue passing information to the U.S. Treasury, whereas according to Belgian law SWIFT ought not to pass this information, since this activity is in breach of Belgian data protection law [174].

29. Alice is throwing a party for her birthday. Since Bob and Charles are good friends of Alice, it ought to be that Alice invite Bob and that Alice invite Charles to her party. However, when Bob and Charles get together, they usually get drunk, and chances are that they will annoy the other guests. Hence Alice ought not invite both Bob and Charles to her party.

30. Having a thousand dollars in my office safe and five hundred in my pocket, and owing Smith and Jones five hundred dollars each, I promise each that I will repay my debt at my office tomorrow, only to find next day that I cannot do so, because overnight my safe has been emptied by a burglar [50, p. 302].

31. A person falls overboard from a ship in a wartime convoy; if the master of the ship leaves his place in the convoy to pick him up, he puts the ship and all on board at risk from submarine attack; if he does not, the person will drown. In the film *The Cruel Sea*, a somewhat similar case occurs; the commander of a corvette is faced with a situation in which if he does not drop depth charges the enemy submarine will get away to sink more ships and kill more people; but if he does drop them he will kill the survivors in

[10] This example also appears in [47, 140, 164].

the water. In fact he drops them, and is depicted in the film as suffering anguish of mind [83, p. 29].

32. During the Battle of Britain, Churchill was faced with the following choice. Thanks to the British government's access to Germany's secret codes, he was informed in advance of many planned German air raids on populated areas. He could evacuate those areas, sparing many innocent lives, but doing so would, with a significant degree of probability, reveal to the Germans that their codes had been broken, seriously impairing the British war effort. He decided not to evacuate these areas [108, p. 214].

33. Elmer had murdered his grandfather (for which crime he was convicted). The grandfather's will bequeathed a substantial sum to Elmer. The grandfather's will violated none of the explicit provision of estate law in New York, and no statute explicitly justified withholding the inheritance from Elmer. Nonetheless, the court ruled the bequest invalid, appealing to the real but unstated intentions of the lawmakers [108, p. 218].[11]

34. A train is moving at a speed of 150 miles per hour. All of a sudden the conductor notices a light on the panel indicating complete brake failure. Straight ahead of him on the track are five hikers, walking with their backs turned, apparently unaware of the train. The conductor notices that the track is about to fork, and another hiker is on the side track. The conductor must make a decision: He can let the train continue on its current course, thereby killing the five hikers, or he can redirect the train onto the side track and thereby kill one hiker but save five. Is it morally permissible for the conductor to take the side track [84, p. 32]?[12]

35. A surgeon walks into the hospital as a nurse rushes forward with the following case. "Doctor! An ambulance just pulled in with five people in critical condition. Two have a damaged kidney, one a crushed heart, one a collapsed lung, and one a completely ruptured liver. We don't have time to search for possible organ donors, but a healthy young man just walked in to donate blood and is sitting in the lobby. We can save all five patients if we take the needed organs from this young man. Of course he won't survive, but we will save all five patients" [84, p. 32].

36. Suppose I have simultaneously arranged to have a private dinner this evening with each of two identical and identically situated twins, both of whom would now be equally disappointed by my cancelation; the situation can be made arbitrarily symmetrical. The resulting prima facie ought's – to have dinner with one twin, and to have dinner with the other – issue from the same source of value, and can meaningfully be compared in importance.

[11] Koons and Seung took this real-life example from Ronald Dworkin's *Law's Empire*. It concerns the 1889 case Riggs v. Palmer, 115 N.Y. 506, 22 N.E. 188.

[12] This is an instance of a famous class of problems called *trolley problems*. Trolley problems, first introduced by Philippa Foot in [53], are thought experiments that present moral dilemmas in which the permissibility to harm one or more persons for the purpose of saving others is questioned. For more discussion on trolley cases, see e.g. [53, 54, 101, 177, 178, 179]. See [135] for some variants of this example.

But in light of the symmetry, what reason could there be for preferring one over the other [95, p. 564]?

37. I contract with party X to be present at a certain spot at a certain time. Separately, I contract with party Y not to be present at that spot at that time. Both contracts are validated in the usual way, by witnessing, etc. I may do this with or without ill intention. It may be my intention to deceive one of the parties. On the other hand, I may just be absentminded. In such circumstances I am legally obliged both to be and not to be at this spot at this time. (And if it be suggested that this is not a case of inconsistent obligations simpliciter, since I am obliged to X to be at that spot and obliged to Y not to be, just take X and Y to be the same person.)

    How can one be sure that I am committed to inconsistent obligations in the situation described? The answer is simple. If, after the event, I am sued by the party of whichever contract I do not comply with, the court will hold me in breach of obligation and award damages appropriately [145, p. 182].

38. Suppose that someone contracts to bring about a more complex inconsistency, say, the squaring of the circle. Suppose that they contracted to do this before it was known to be impossible, and that they failed to fulfill the contract. Would a court hold them in default? The answer is 'yes'. Suppose it were proved to be impossible after signing the contract but before the court hearing? The answer is still 'yes' [145, p. 183].

39. Suppose that there is a pair of statutes, one of which requires a car owner to change registration plates on January 1st, and the other of which forbids working on a Sunday. About every seven years the average law-abiding citizen is embarrassed [145, p. 184].

40. Suppose that there is a certain country which has a constitutional parliamentary system of government. And suppose that its constitution contains the following clauses. In a parliamentary election:

    (1) no person of the female sex shall have the right to vote;

    (2) all property holders shall have the right to vote.

    We may also suppose that it is part of common law that women may not legally possess property. As enlightenment creeps over the country, this part of common law is revised to allow women to hold property ... Inevitably, sooner or later, a woman, whom we will call 'Jan', turns up at a polling booth for a parliamentary election claiming the right to vote on the ground that she is a property holder. A test case ensues. Patently, the law is inconsistent. Jan, it would seem, both does and does not have the right to vote in this election [145, pp. 184-185].

41. Let us suppose that the priority law of a certain state is as follows. At an unmarked junction at which two vehicles arrive simultaneously, (1) any female driver shall have priority over any male driver; and (2) any older person shall have priority over any younger person.

If now an occasion arises when Mr X, of age 40, meets Ms Y, of age 30, at a junction, then Ms Y has priority by (1), whereas Mr X has priority by (2). So X and Y both have and do not have priority [145, p. 185].

42. Suppose you are a doctor faced with a mentally competent patient who has refused a treatment you think represents her best hope of survival. Should you try again to persuade her (a possible violation of respect for the patient's autonomy) or should you accept her decision (a possible violation of your duty to provide the most beneficent care) [196, p. 27]?

43. Suppose the company you work for licenses some new, expensive computer software, say Adobe's Photoshop. After becoming comfortable with the new software package at work, you feel the urge to copy it onto your home computer. An internal dialog commences, but not necessarily as wholly verbal and grammatical as what follows. "Let's bring Photoshop home and load the program on my Mac." "You shouldn't do that. That would be illegal and stealing." "But I'd use it for work-related projects that benefit my company, which owns the software." "Yes, but you'd also use it for personal projects with no relation to the company." "True, but most of the work would be company related." And so on and on [196, p. 181].

44. An ancient paradox is about the famous Greek law teacher Protagoras and goes like this: Protagoras and Euathlus agree that the former is to instruct the latter in rhetoric and is to receive a certain fee which is to be paid *if and only if* Euathlus wins his first court-case (in some versions: *as soon as* he has won his first case). Well, Euathlus completed his course but did not take any law cases. Some time elapsed and Protagoras sued his student for the sum. The following arguments were presented to the judge in court.

    Protagoras: If I win this case, then Euathlus has to pay me by virtue of your verdict. On the other hand, if he wins the case, then he will won [*sic*] his first case, hence he has to pay me, this time by virtue of our agreement. In either case, he has to pay me. Therefore, he is obliged to pay me my fee.

    Euathlus: If I win this case, then, by your verdict, I don't have to pay. If, however, Protagoras wins the case, then I will not yet have won my first case, so, by our agreement, I don't have to pay. Hence I am not obliged to pay the fee [8, p. 147].

45. People have described the situation in Vietnam as follows (not that it really *is* this way): If the Americans withdraw from Vietnam, a large number of people will be killed. If the Americans stay in (do not withdraw from) Vietnam, a large number of people will be killed [154, p. 659].[13]

46. If Z does not go to war, he fails to help his friends and fellow countrymen when they are in desperate need. If Z goes to war, he will be involved in killing people he has nothing morally or otherwise against [154, p. 659].

[13] Routley and Plumwood also present the following variant of this dilemma: "If the company commander invades the hamlet, a large number of his troops will be killed. If he does not invade the hamlet the prisoners held therein will be killed" [154, p. 659].

47. Suppose that we have a duly qualified principle to the effect that things which lead to the feeding of starving people should happen. Suppose also that we have limited resources, and have decided to distribute these as follows: If the coin comes down heads (p), then group A of starving people will be fed. But if the coin does not come down heads ($\neg p$), then group B of starving people will be fed. Plainly in an ideal situation both p and $\neg p$ will occur, that is the coin will come down heads and it will not come down heads, so that both groups of people will be fed. And if the moral principle cited is correctly applied then a case of $\mathsf{O}p$ and $\mathsf{O}\neg p$ occurs [154, p. 660].

48. It ought to happen that the forest is chopped down, for the benefit of the presently starving (or etc.), and it ought to happen that the forest is preserved by positive action for the benefit of coming generations (or etc.) [154, p. 660].

49. Consider the dilemma of a feminist environmentalist as regards Aboriginal women. On the one hand, suppression of women in Aboriginal society should be opposed; on the other hand major western interference in Aboriginal affairs should be avoided; but changing the position of women would constitute major interference [154, p. 660].

50. The University opposes the government. The Vice-Chancellor, or rather his advisor on logic, argues both the following: on the one hand opposing the government is wrong, because the University's funds will be restricted and the students and learning in general will suffer. On the other hand, not opposing the government is wrong because it will strengthen the present iniquitous status quo [154, pp. 660-661].

51. As the play opens, Philoctetes has suffered for years with a disfiguring disease; he had wandered into a forbidden garden, through no fault of his own, and had been punished by the gods. Banished to a remote island, he has nothing left but his bow. But the gods reveal to Odysseus that only that bow can win the Trojan War. So, Odysseus orders Neoptolemus to trick Philoctetes out of his bow. Neoptolemus obeys. Overcome with regret, however, he decides to return the bow. Neoptolemus tricks Philoctetes for serious reasons: to obey Odysseus, his commander, and to win the war. But those reasons, he concludes, cannot justify the cruelty to the anguished Philoctetes [9, p. 19].

52. A coach may consider himself consistent when telling each of several athletes that they ought to win a contest [79, p. 344].

53. A soldier may claim that the commanders-in-chief of two armies at war both ought to bring about the victory of their respective side [79, p. 344].

54. This problem arises when someone in possession of real estate – which is owned not by him but by someone else – transfers it (by way of sale or gift) to a third person. Then comes the question whether (and if so, in what circumstances) the owner of real estate may recover its possession from the third holder. Or to put the question in other terms: in what circumstances

has the third holder the obligation to restore it to its owner and in what circumstances (if any) may he keep it, i.e. be allowed to refused to restore it [2, p. 9-10]?

55. Suppose that the UK becomes the belligerent occupant of a territory that has Sharia as part of its domestic law, e.g. Iran. The Penal Law of Iran prescribes stoning as a punishment for adultery. On the one hand, International Human Rights Law commands the UK to take all possible measures to prevent the stoning of adulterers in the territory that it has occupied. On the other hand, because it considers occupation to be a temporary situation that requires deference to the displaced sovereign, International Humanitarian Law prohibits the UK from changing the laws of the occupied country, particularly its penal laws [136, p. 480].

56. In 1986 German national Jens Söring committed a double murder in Virginia, USA, after which he fled to the UK, where he was ultimately arrested. The European Convention on the Protection of Human Rights and Fundamental Freedoms interpreted Article 3 of the European Court of Human Rights as setting out a non-refoulement obligation, prohibiting the UK from transferring Söring to the US if a real risk of that person being subjected to inhuman or degrading treatment in the US was established. On the other end was a valid extradition treaty between the UK and the US, which obliged the UK to extradite Söring, and which specified no exception to that obligation [136, pp. 470-471].

57. The Third and Fourth amended versions of the Declaration maintained that it is unethical to assign patients to receive a placebo when effective treatment exists: "In any medical study, every patient – including those of a control group – should be assured of the best proven diagnostic and therapeutic method". This is in clear opposition to current practice of the US Food and Drug Administration (FDA). Despite the mandates of the Declaration of Helsinki and concern from ethicists and scientists, the FDA continues to demand and defend placebo-controlled evidence of efficacy and safety for the development of new pharmaceuticals, even if effective therapy exists [134, pp. 188-189].

58. In Aeschylus's *Choephoroe*, Orestes's mother, Clytaemestra, has killed his father, Agamemnon. She and her lover, Aegisthus, rule in Argos. Orestes returns secretly from exile, gains entrance to the palace, and has no difficulty – physical or moral – in killing Aegisthus. But then Clytaemestra confronts Orestes. He clearly can kill her, but should he do so? Orestes ought to kill his mother, because he owes his father the deed of vengeance. But he ought not to kill his mother, for killing a parent is a terrible crime. As Aeschylus makes clear in the *Eumenides*, Orestes faces punishment from the avenging Furies, whatever he chooses to do [57, pp. 116-117].

59. In the country of Freedonia, there are only two possible forms of government: rule by the people's choice from the candidates put forward by Party A and Party B; or military dictatorship. An impartial observer asserts that

democratic rule is clearly better than dictatorship. For this reason, says the observer, it ought to be that either the candidate of Party A rules the country or that the candidate of Party B rules the country. However, thinks the observer, both candidates are equally rotten. Surely it ought to be that the candidate of Party A does not rule the country. But that hardly means that it ought to be that the candidate of Party B rules the country. And yet, given the terrors of military dictatorship, it ought to be that one or the other candidate rules [57, p. 48].

60. Suppose that both my brother and my sister have a disease that in a few cases may lead to kidney failure and to the need for a transplantation. Suppose further that I have solemnly promised each of them that one of my kidneys will be available for transplantation if that should be medically called for. Let D_iA [D_iB] denote my action of making one of my kidneys available for my sister [brother]. If both my brother and my sister turn out to need a transplantation, then both $\mathsf{O}D_iA$ and $\mathsf{O}D_iB$ apply as prima facie duties, but I cannot reasonably be said to have a prima facie duty $\mathsf{O}(D_iA \wedge D_iB)$ [79, pp. 344-345].[14]

61. Consider a Buridan's ass-type moral dilemma in which not both of two identical twins (of identical moral status) can be saved from being crushed to death by a heavy rock. The twins are pinned down in such a way that only one can be pulled free at a time. If nothing is done the rock will soon kill both, but if either twin is removed, the shifting increased weight will immediately kill the other [98, p. 44].

62. Imagine the following situation: you are a heart surgeon treating newborn Siamese twins who are grown together at the chest in such a way that they share the same heart. Apart from sharing this vital organ, they have complete sets of separate organs. The heart is too weak to support both little bodies but perfectly strong enough to support one of them. So, if they are not separated within the next 24 hours, they will both die. There is no way of deciding which of the babies to give the organ to - each of them has exactly the same fair chance of survival with the organ and will probably be able to live a long and happy life, and, of course, each of them will certainly die without a heart. From this, the following question arises: During the operation that separates the two, which one will you give the heart to? And which child will you leave to die[37, p. 78]?[15]

63. Imagine a situation where two identical twins are drowning some distance apart from each other, and the situation is such that you can save either of them, but you cannot save both [130, 129, 123].

64. Consider, for example, the controversies surrounding non-spontaneous abortion. Philosophers are often criticized for inventing bizarre examples and counterexamples to make a philosophical point. But no contrived example

[14] Hansen uses this example to argue against the application of (AND) to prima facie duties.
[15] This example also appears in [198, p. 241].

can equal the complexity and the puzzles generated by the actual circumstances of foetal conception, parturation, and ultimate birth of a human being. We have an organism, internal to and parasitic upon a human being, hidden from view but relentlessly developing into a human being, which at some stage of development can live, with nurture, outside of its host. There are arguments that recognize competing claims: the right to life of the foetus (at some stage) versus the right of someone to determine what happens to his body. Arguments that justify choosing the mother over the foetus (or vice-versa) where their survival is in competition. Arguments in which foetuses that are defective are balanced against the welfare of others. Arguments in which the claims to survival of others will be said to override survival of the foetus under conditions of great scarcity. There are even arguments that deny prima facie conflicts altogether on some metaphysical grounds, such as that the foetus is not a human being or a person until quickening, or until it has recognizable human features, or until its life can be sustained external to its host, or until birth, or until after birth when it has interacted with other persons. Various combinations of such arguments are proposed in which the resolution of a dilemma is seen as more uncertain, the more proximate the foetus is to whatever is defined as being human or being a person. What all the arguments seem to share is the assumption that there is, despite uncertainty, a resolution without residue; that there is a correct set of metaphysical claims, principles, and priority rankings of principles which will justify the choice. Then, given the belief that one choice is justified, assignment of guilt relative to the overridden alternative is seen as inappropriate, and feelings of guilt or pangs of conscience are viewed as, at best, sentimental. But as one tries to unravel the tangle of arguments, it is clear that to insist that there is in every case a solution without residue is false to the moral facts [121, pp.131-132].[16]

65. A team of Dutch scientists of the Erasmus Medical Center led by the virologist Ron Fouchier has created a highly contagious variant of the H5N1 ("bird flu") virus. The scientists have submitted their results for publication in *Science*, claiming that they have positively answered the question whether or not the H5N1 virus can possibly trigger a pandemic by mutating into a more transmissible variant.

    On the one hand, many virologists support the publication of these results due to their potential benefits for public health. According to Fouchier, the U.S. National Institute of Health (NIH) has agreed to the publication of his team's results. On the other hand, representatives of the U.S. Government fear that the publication of the study will give terrorists new knowledge for constructing bio-weapons of mass destruction.

    On December 20^{th} 2011, the U.S. National Science Advisory Board for Biosecurity ruled that all technical details must be left out for publication. The journals *Science* and *Nature* opposed this decision. After months of debate about whether the benefits of publishing the research outweigh the

[16] This example by Marcus is integrally cited in [48, pp. 294-295].

risks, the paper of Fouchier's group was published in *Science* on June 21st 2012.[17]

[17] For a more detailed oversight regarding this controversy, see: `http://www.nature.com/news/specials/mutantflu/index.html`.

Appendix B

Overview of formal languages

The following table provides an oversight of the grammars defined in this thesis and shows which logics use which grammar.

$\mathcal{W}^a$	$\{p, q, r, \ldots\}$	
$\mathcal{W}^l$	$\{A, \neg A \mid A \in \mathcal{W}^a\}$	
$\mathcal{W}^{\sim}_l$	$\{A, {\sim}A \mid A \in \mathcal{W}^a\}$	
$\mathcal{W}$	$\mathcal{W}^a \mid \neg\langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \vee \langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \wedge \langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \supset \langle\mathcal{W}\rangle \mid \langle\mathcal{W}\rangle \equiv \langle\mathcal{W}\rangle \mid \bot$	**CL**
$\mathcal{W}^{\text{pos}}$	$\mathcal{W}^a \mid \langle\mathcal{W}^{\text{pos}}\rangle \vee \langle\mathcal{W}^{\text{pos}}\rangle \mid \langle\mathcal{W}^{\text{pos}}\rangle \wedge \langle\mathcal{W}^{\text{pos}}\rangle \mid \langle\mathcal{W}^{\text{pos}}\rangle \supset \langle\mathcal{W}^{\text{pos}}\rangle \mid \langle\mathcal{W}^{\text{pos}}\rangle \equiv \langle\mathcal{W}^{\text{pos}}\rangle$	$\mathbf{CL^{pos}}$
$\mathcal{W}^{\sim}$	$\mathcal{W}^a \mid {\sim}\langle\mathcal{W}^{\sim}\rangle \mid \langle\mathcal{W}^{\sim}\rangle \vee \langle\mathcal{W}^{\sim}\rangle \mid \langle\mathcal{W}^{\sim}\rangle \wedge \langle\mathcal{W}^{\sim}\rangle \mid \langle\mathcal{W}^{\sim}\rangle \supset \langle\mathcal{W}^{\sim}\rangle \mid \langle\mathcal{W}^{\sim}\rangle \equiv \langle\mathcal{W}^{\sim}\rangle$	**CLuN**, **CLuNs**, **CLaN**, **CLaNs**, **CLoN**, **CLoNs**
$\mathcal{W}^{\sim}_{\bot}$	$\mathcal{W}^a \mid {\sim}\langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \vee \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \wedge \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \supset \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \langle\mathcal{W}^{\sim}_{\bot}\rangle \equiv \langle\mathcal{W}^{\sim}_{\bot}\rangle \mid \bot$	$\mathbf{CLuN}^{\bot}$, $\mathbf{CLuNs}^{\bot}$, $\mathbf{CLaN}^{\bot}$, $\mathbf{CLaNs}^{\bot}$, $\mathbf{CLoN}^{\bot}$, $\mathbf{CLoNs}^{\bot}$
$\mathcal{W}^{\mathsf{O}}$	$\mathcal{W} \mid \mathsf{O}\langle\mathcal{W}\rangle \mid \mathsf{P}\langle\mathcal{W}\rangle \mid \neg\langle\mathcal{W}^{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}\rangle \vee \langle\mathcal{W}^{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}\rangle \wedge \langle\mathcal{W}^{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}\rangle \supset \langle\mathcal{W}^{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}\rangle \equiv \langle\mathcal{W}^{\mathsf{O}}\rangle$	**SDL**
$\mathcal{W}^{\mathsf{O}'}$	$\mathcal{W} \mid \mathsf{O}\langle\mathcal{W}^{\mathsf{O}'}\rangle \mid \mathsf{P}\langle\mathcal{W}^{\mathsf{O}'}\rangle \mid \neg\langle\mathcal{W}^{\mathsf{O}'}\rangle \mid \langle\mathcal{W}^{\mathsf{O}'}\rangle \vee \langle\mathcal{W}^{\mathsf{O}'}\rangle \mid \langle\mathcal{W}^{\mathsf{O}'}\rangle \wedge \langle\mathcal{W}^{\mathsf{O}'}\rangle \mid \langle\mathcal{W}^{\mathsf{O}'}\rangle \supset \langle\mathcal{W}^{\mathsf{O}'}\rangle \mid \langle\mathcal{W}^{\mathsf{O}'}\rangle \equiv \langle\mathcal{W}^{\mathsf{O}'}\rangle$	
$\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}$	$\mathcal{W} \mid \mathsf{O}\langle\mathcal{W}\rangle \mid \neg\langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \vee \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \wedge \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \supset \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \mid \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle \equiv \langle\mathcal{W}^{\mathsf{O}\smallsetminus\mathsf{P}}\rangle$	
$\mathcal{W}^{\mathsf{O}}_{\Box}$	$\mathcal{W}^{\mathsf{O}} \mid \Box\langle\mathcal{W}\rangle \mid \neg\langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \vee \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \wedge \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \supset \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \mid \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle \equiv \langle\mathcal{W}^{\mathsf{O}}_{\Box}\rangle$	$\mathbf{P}_{\Diamond}$, $\mathbf{P}^{\mathbf{x}}_{\Diamond}$

$\mathcal{W}^{\mathbf{P2}}$	$\mathcal{W} \mid \mathsf{O}_e\langle\mathcal{W}\rangle \mid \mathsf{O}_a\langle\mathcal{W}\rangle \mid \neg\langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \vee \langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \wedge \langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \supset \langle\mathcal{W}^{\mathbf{P2}}\rangle \mid \langle\mathcal{W}^{\mathbf{P2}}\rangle \equiv \langle\mathcal{W}^{\mathbf{P2}}\rangle$	**P2**, **P2.1$^{\mathbf{x}}$**, **P2.2$^{\mathbf{x}}$**, **SDL$_{\mathbf{a}}$P$_{\mathbf{e}}$**, **SDL$_{\mathbf{ae}}$**
$\mathcal{W}^{\mathbf{DP}}$	$\mathcal{W}^{\sim}_{\perp} \mid \mathsf{O}\langle\mathcal{W}^{\sim}_{\perp}\rangle \mid {\sim}\langle\mathcal{W}^{\mathbf{DP}}\rangle \mid \langle\mathcal{W}^{\mathbf{DP}}\rangle \vee \langle\mathcal{W}^{\mathbf{DP}}\rangle \mid \langle\mathcal{W}^{\mathbf{DP}}\rangle \wedge \langle\mathcal{W}^{\mathbf{DP}}\rangle \mid \langle\mathcal{W}^{\mathbf{DP}}\rangle \supset \langle\mathcal{W}^{\mathbf{DP}}\rangle \mid \langle\mathcal{W}^{\mathbf{DP}}\rangle \equiv \langle\mathcal{W}^{\mathbf{DP}}\rangle$	**DP**,**DP$^{\mathbf{x}}$**
$\mathcal{W}^{\not\sim}$	$\{A \mid A \in \mathcal{W}^{\mathbf{DP}}$ and A is not of the form ${\sim}B$, where $B \in \mathcal{W}^{\mathbf{DP}}\}$	
$\mathcal{W}^{\sim}_{\mathsf{O}}$	$\{\mathsf{O}A \mid A \in \mathcal{W}^{\sim}_{\perp}\}$	
$\mathcal{W}^{\neg}_{\mathsf{O}}$	$\mathsf{O}\langle\mathcal{W}^{\sim}\rangle \mid \mathsf{P}\langle\mathcal{W}^{\sim}\rangle \mid \neg\langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \vee \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \wedge \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \supset \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \mid \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle \equiv \langle\mathcal{W}^{\neg}_{\mathsf{O}}\rangle$	
$\mathcal{W}^{\mathbf{LNP}}$	$\mathcal{W} \mid \mathcal{W}^{\neg}_{\mathsf{O}} \mid \neg\langle\mathcal{W}^{\mathbf{LNP}}\rangle \mid \langle\mathcal{W}^{\mathbf{LNP}}\rangle \vee \langle\mathcal{W}^{\mathbf{LNP}}\rangle \mid \langle\mathcal{W}^{\mathbf{LNP}}\rangle \wedge \langle\mathcal{W}^{\mathbf{LNP}}\rangle \mid \langle\mathcal{W}^{\mathbf{LNP}}\rangle \supset \langle\mathcal{W}^{\mathbf{LNP}}\rangle \mid \langle\mathcal{W}^{\mathbf{LNP}}\rangle \equiv \langle\mathcal{W}^{\mathbf{LNP}}\rangle$	**LNP**,**LNP$^{\mathbf{x}}$**
$\mathcal{W}^{\mathbf{ML}}$	$\langle\mathcal{W}^{a}\rangle \mid \neg\langle\mathcal{W}^{\mathbf{ML}}\rangle \mid \langle\mathcal{W}^{\mathbf{ML}}\rangle \vee \langle\mathcal{W}^{\mathbf{ML}}\rangle \mid \langle\mathcal{W}^{\mathbf{ML}}\rangle \wedge \langle\mathcal{W}^{\mathbf{ML}}\rangle \mid \Box_J\langle\mathcal{W}^{\mathbf{ML}}\rangle \mid \Diamond_J\langle\mathcal{W}^{\mathbf{ML}}\rangle$	**ML**
$\mathcal{W}^{\mathbf{MDL}}$	$\langle\mathcal{W}^{\mathbf{ML}}\rangle \mid \neg\langle\mathcal{W}^{\mathbf{MDL}}\rangle \mid \langle\mathcal{W}^{\mathbf{MDL}}\rangle \vee \langle\mathcal{W}^{\mathbf{MDL}}\rangle \mid \langle\mathcal{W}^{\mathbf{MDL}}\rangle \wedge \langle\mathcal{W}^{\mathbf{MDL}}\rangle \mid \Box_J\langle\mathcal{W}^{\mathbf{MDL}}\rangle \mid \Diamond_J\langle\mathcal{W}^{\mathbf{MDL}}\rangle \mid \mathsf{O}\langle\mathcal{W}^{\mathbf{MDL}}\rangle \mid \mathsf{P}\langle\mathcal{W}^{\mathbf{MDL}}\rangle$	**MDL**
$\mathcal{W}^{\mathbf{ML}}_{\sim}$	$\langle\mathcal{W}^{a}\rangle \mid {\sim}\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \vee \langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \wedge \langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \Box_J\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid \Diamond_J\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle$	
$\mathcal{W}^{\mathbf{PMDL}}$	$\langle\mathcal{W}^{\mathbf{ML}}_{\sim}\rangle \mid {\sim}\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \vee \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \wedge \langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \Box_J\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \Diamond_J\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \mathsf{O}\langle\mathcal{W}^{\mathbf{PMDL}}\rangle \mid \mathsf{P}\langle\mathcal{W}^{\mathbf{PMDL}}\rangle$	**PMDL**, **PMDL$^{\mathbf{x}}$**

Appendix C

$\mathbf{CL^{pos}}$, CLuN(s), CLaN(s), and CLoN(s)

CLoN, **CLuN**, and **CLaN** are extensions of $\mathbf{CL^{pos}}$, the positive (negation-free) fragment of **CL**. These logics allow for negation-gaps and/or negation-gluts. **CLoN** abbreviates "classical logic with both gluts and gaps for negation", **CLuN** abbreviates "classical logic with gluts for negation", and **CLaN** abbreviates "classical logic with gaps for negation". For some presentations of these systems in the literature, see e.g. [13, 19, 20].

Below we provide an axiomatic and semantic characterization of the propositional fragment of these logics and of their extensions **CLoNs**, **CLuNs**, and **CLaNs** (which validate de Morgan's laws for negation). We also define the logics obtained by adding the falsum constant $\bot$ to these systems.

C.1 Axiomatizations

Let:

$\mathcal{W}^{\text{pos}} := \mathcal{W}^a \mid \langle\mathcal{W}^{\text{pos}}\rangle \vee \langle\mathcal{W}^{\text{pos}}\rangle \mid \langle\mathcal{W}^{\text{pos}}\rangle \wedge \langle\mathcal{W}^{\text{pos}}\rangle \mid \langle\mathcal{W}^{\text{pos}}\rangle \supset \langle\mathcal{W}^{\text{pos}}\rangle \mid \langle\mathcal{W}^{\text{pos}}\rangle \equiv \langle\mathcal{W}^{\text{pos}}\rangle$

$\mathbf{CL}^{\text{pos}}$, the positive fragment of **CL**, is defined by the language schema $\mathcal{W}^{\text{pos}}$, the rule modus ponens (MP) $(A, A \supset B/B)$ and the following axiom schemas:

(A⊃1)	$A \supset (B \supset A)$
(A⊃2)	$(A \supset (B \supset C)) \supset ((A \supset B) \supset (A \supset C))$
(A⊃3)	$((A \supset B) \supset A) \supset A$
(A∧1)	$(A \wedge B) \supset A$
(A∧2)	$(A \wedge B) \supset B$
(A∧3)	$A \supset (B \supset (A \wedge B))$
(A∨1)	$A \supset (A \vee B)$
(A∨2)	$B \supset (A \vee B)$
(A∨3)	$(A \supset C) \supset ((B \supset C) \supset ((A \vee B) \supset C))$
(A≡1)	$(A \equiv B) \supset (A \supset B)$
(A≡2)	$(A \equiv B) \supset (B \supset A)$
(A≡3)	$(A \supset B) \supset ((B \supset A) \supset (A \equiv B))$

Adding the classical negation "$\neg$" to $\mathbf{CL}^{\text{pos}}$ is sufficient in order to obtain full **CL**. This can be done by replacing the set $\mathcal{W}^{\text{pos}}$ of $\mathbf{CL}^{\text{pos}}$-wffs by the set $\mathcal{W}$ of **CL**-wffs, and by adding the axiom schemas ($A\neg 1$) and ($A\neg 2$):

(A¬1) $(A \supset \neg A) \supset \neg A$
(A¬2) $A \supset (\neg A \supset B)$

For the logics **CLuN(s)**, **CLaN(s)**, and **CLoN(s)** we make use of the negation connective "$\sim$". **CLoN** is defined by simply adding this connective to the language schema of $\mathbf{CL}^{\text{pos}}$, i.e. by replacing the set $\mathcal{W}^{\text{pos}}$ of $\mathbf{CL}^{\text{pos}}$-wffs with the set $\mathcal{W}^{\sim}$:

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

CLuN is defined by adding to **CLoN** the axiom schema (A~1):

(A~1) $(A \supset \sim A) \supset \sim A$

CLaN is defined by adding to **CLoN** the axiom schema (A~2):

(A~2) $A \supset (\sim A \supset B)$

The logics **CLoNs**, **CLuNs**, and **CLaNs** are defined by adding the axiom schemas (A~~)-(A~ ≡) to the logics **CLoN**, **CLuN**, and **CLaN** respectively:

(A~~) $\sim\sim A \equiv A$
(A~ ⊃) $\sim(A \supset B) \equiv (A \wedge \sim B)$
(A~∧) $\sim(A \wedge B) \equiv (\sim A \vee \sim B)$
(A~∨) $\sim(A \vee B) \equiv (\sim A \wedge \sim B)$
(A~ ≡) $\sim(A \equiv B) \equiv ((A \vee B) \wedge (\sim A \vee \sim B))$

Finally, where:

$$\mathcal{W}^{\sim}_{\perp} := \mathcal{W}^{a} \mid \sim\langle\mathcal{W}^{\sim}_{\perp}\rangle \mid \langle\mathcal{W}^{\sim}_{\perp}\rangle \vee \langle\mathcal{W}^{\sim}_{\perp}\rangle \mid \langle\mathcal{W}^{\sim}_{\perp}\rangle \wedge \langle\mathcal{W}^{\sim}_{\perp}\rangle \mid \langle\mathcal{W}^{\sim}_{\perp}\rangle \supset \langle\mathcal{W}^{\sim}_{\perp}\rangle \mid \langle\mathcal{W}^{\sim}_{\perp}\rangle \equiv \langle\mathcal{W}^{\sim}_{\perp}\rangle \mid \perp$$

The logics $\mathbf{CLoN(s)}^{\perp}$, $\mathbf{CLuN(s)}^{\perp}$, and $\mathbf{CLaN(s)}^{\perp}$ are defined by
(i) replacing in the logics **CLoN(s)**, **CLuN(s)**, and **CLaN(s)** respectively the language schema $\mathcal{W}^{\sim}$ with the language schema $\mathcal{W}^{\sim}_{\perp}$, and
(ii) adding the schema (A⊥1):

(A⊥1) $\perp \supset A$

Due to the addition of the falsum constant to $\mathbf{CLoN(s)}^{\perp}$, $\mathbf{CLuN(s)}^{\perp}$ and $\mathbf{CLaN(s)}^{\perp}$, the classical negation connective is definable in these logics by $\neg A =_{\text{df}} A \supset \perp$. Thus, all classical connectives become definable in these logics. This is important in view of the remark made in footnote 1 in Section 4.2.

C.2 Semantics

A $\mathbf{CL}^{\text{pos}}$-model M is a tuple $\langle\mathcal{W}^{a}, v\rangle$, where $v : \mathcal{W}^{a} \to \{0,1\}$ is an assignment function. The valuation function $v_M : \mathcal{W}^{\text{pos}} \to \{0,1\}$ associated with M is defined by:

(Ca) where $A \in \mathcal{W}^{a}, v_M(A) = 1$ iff $v(A) = 1$
(C∨) $v_M(A \vee B) = 1$ iff $v_M(A) = 1$ or $v_M(B) = 1$
(C∧) $v_M(A \wedge B) = 1$ iff $v_M(A) = 1$ and $v_M(B) = 1$
(C⊃) $v_M(A \supset B) = 1$ iff $v_M(A) = 0$ or $v_M(B) = 1$
(C≡) $v_M(A \equiv B) = 1$ iff $v_M(A) = v_M(B)$

A $\mathbf{CL}^{\mathrm{pos}}$-model M verifies A, $M \Vdash A$, iff $v_M(A) = 1$.
The logic **CLoN** is obtained by letting the assignment function map both atoms and negated formulas into the set $\{0, 1\}$. Let $\mathcal{F}^{\sim} = \{{\sim}A \mid A \in \mathcal{W}\}$. A **CLoN**-model M is a tuple $\langle \mathcal{W}^a \cup \mathcal{F}^{\sim}, v\rangle$, where $v : \mathcal{W}^a \cup \mathcal{F}^{\sim} \to \{0, 1\}$ is an assignment function. The valuation function $v_M : \mathcal{W}^{\sim} \to \{0, 1\}$ associated with M is defined by adding the clause (C~0) to the clauses (Ca), (C$\vee$), (C$\wedge$), (C$\supset$), and (C$\equiv$):

(C~0) $\quad v_M({\sim}A) = 1$ iff $v({\sim}A) = 1$

A **CLoN**-model M verifies A, $M \Vdash A$, iff $v_M(A) = 1$.
A **CLuN**-model M too is a tuple $\langle \mathcal{W}^a \cup \mathcal{F}^{\sim}, v\rangle$, where $v : \mathcal{W}^a \cup \mathcal{F}^{\sim} \to \{0, 1\}$ is an assignment function. The valuation function $v_M : \mathcal{W}^{\sim} \to \{0, 1\}$ associated with M is defined by adding the clause (C~1) to the clauses (Ca), (C$\vee$), (C$\wedge$), (C$\supset$), and (C$\equiv$):

(C~1) $\quad v_M({\sim}A) = 1$ iff ($v_M(A) = 0$ or $v({\sim}A) = 1$)

A **CLuN**-model M verifies A, $M \Vdash A$, iff $v_M(A) = 1$.
A **CLaN**-model M again is a tuple $\langle \mathcal{W}^a \cup \mathcal{F}^{\sim}, v\rangle$, where $v : \mathcal{W}^a \cup \mathcal{F}^{\sim} \to \{0, 1\}$ is an assignment function. The valuation function $v_M : \mathcal{W}^{\sim} \to \{0, 1\}$ associated with M is defined by adding the clause (C~2) to the clauses (Ca), (C$\vee$), (C$\wedge$), (C$\supset$), and (C$\equiv$):

(C~2) $\quad v_M({\sim}A) = 1$ iff ($v_M(A) = 0$ and $v({\sim}A) = 1$)

A **CLaN**-model M verifies A, $M \Vdash A$, iff $v_M(A) = 1$.
The semantics for **CLoNs**, **CLuNs**, and **CLaNs** is obtained by:
(i) letting the assignment function v assign truth values to literals, i.e. $v : \mathcal{W}^{\sim}_l \to \{0, 1\}$, (ii) replacing in the respective logics the clauses (C~0), (C~1), and (C~2) with the clauses (C~0'), (C~1'), and (C~2'):

(C~0') Where $A \in \mathcal{W}^a$, $v_M({\sim}A) = 1$ iff $v({\sim}A) = 1$
(C~1') Where $A \in \mathcal{W}^a$, $v_M({\sim}A) = 1$ iff ($v_M(A) = 0$ or $v({\sim}A) = 1$)
(C~2') Where $A \in \mathcal{W}^a$, $v_M({\sim}A) = 1$ iff ($v_M(A) = 0$ and $v({\sim}A) = 1$)

and (iii) adding the clauses (C~~)-(C~ $\equiv$) to the clauses for **CLoN**, **CLuN** and **CLaN** respectively:

(C~~) $\quad v_M({\sim}{\sim}A) = v_M(A)$
(C~$\vee$) $\quad v_M({\sim}(A \vee B) = v_M({\sim}A \wedge {\sim}B)$
(C~$\wedge$) $\quad v_M({\sim}(A \wedge B) = v_M({\sim}A \vee {\sim}B)$
(C~ $\supset$) $\quad v_M({\sim}(A \supset B) = v_M(A \wedge {\sim}B)$
(C~ $\equiv$) $\quad v_M({\sim}(A \equiv B) = v_M((A \vee B) \wedge ({\sim}A \vee {\sim}B))$

Finally, the semantics for $\mathbf{CLoN(s)}^{\perp}$, $\mathbf{CLuN(s)}^{\perp}$, and $\mathbf{CLaN(s)}^{\perp}$ is obtained by adding the clause (C$\perp$) to the clauses for **CLoN(s)**, **CLuN(s)**, and **CLaN(s)** respectively:

(C$\perp$) $\quad v_M(\perp) = 0$

Appendix D

(Meta-)properties of the logic $\mathbf{P}_\Diamond$

In this Appendix, we provide a semantic characterization of the logic $\mathbf{P}_\Diamond$ in terms of neighborhoods. We rely on results from the literature for the proof of Theorem 6 (cfr. infra).

A $\mathbf{P}_\Diamond$-*frame* is a tuple $\langle W, R, \mathcal{N}\rangle$ where W is a set of points (worlds), $R \subseteq W \times W$ is an accessibility relation and $\mathcal{N} : W \to \wp(\wp(W))$ is a neighborhood function that satisfies the following conditions for each $w \in W$. Let $Rw = \{w' \mid Rww'\}$. $\mathbf{P}_\Diamond$-frames satisfy the following frame conditions:

(F-NEC) $W \in \mathcal{N}(w)$

(F-RM) If $X \in \mathcal{N}(w)$ and $X \subseteq Y$ then $Y \in \mathcal{N}(w)$

(F-PN) $\varnothing \notin \mathcal{N}(w)$

(F-OIC) If $Rw \subseteq X$ then $W \setminus X \notin \mathcal{N}(w)$

(F-$\Diamond$AND) If $X \in \mathcal{N}(w)$, $Y \in \mathcal{N}(w)$ and $Rw \cap (X \cap Y) \neq \varnothing$ then $X \cap Y \in \mathcal{N}(w)$

A $\mathbf{P}_\Diamond$-*model* is a tuple $\langle W, R, \mathcal{N}, v, w_0\rangle$ where $\langle W, R, \mathcal{N}\rangle$ is a $\mathbf{P}_\Diamond$-frame, $w_0 \in W$ is the actual world, and $v : \mathcal{W}^a \to \wp(W)$ is an assignment function. Truth at a world is defined in the following way:

(Ca)	Where $A \in \mathcal{W}^a$, $M, w \Vdash A$ iff $A \in v(w)$
(C$\neg$)	$M, w \Vdash \neg A$ iff $M, w \nVdash A$
(C$\vee$)	$M, w \Vdash A \vee B$ iff ($M, w \Vdash A$ or $M, w \Vdash B$)
(C$\wedge$)	$M, w \Vdash A \wedge B$ iff ($M, w \Vdash A$ and $M, w \Vdash B$)
(C$\supset$)	$M, w \Vdash A \supset B$ iff ($M, w \nVdash A$ or $M, w \Vdash B$)
(C$\equiv$)	$M, w \Vdash A \equiv B$ iff ($M, w \Vdash A$ iff $M, w \Vdash B$)
(CO)	$M, w \Vdash \mathsf{O}A$ iff $\lvert A\rvert_M \in \mathcal{N}(w)$ where $\lvert A\rvert_M =_{\text{df}} \{w \in W \mid M, w \models A\}$
(C$\Box$)	$M, w \Vdash \Box A$ iff for all $w' \in Rw$, $M, w' \models A$

Soundness and completeness is proven generically by Pattinson and Schröder for all rank-1 modal logics with respect to their canonical neighborhood semantics in [158]. Since their result applies to our semantics, Theorem 6 follows immediately.

Appendix E

(Meta-)properties of the logic DP

In this Appendix we prove some further properties of the logic **DP** (Section E.1), provide the soundness and completeness proof for this logic (Section E.2) and outline the proof of Theorem 29 (Section E.3). In some of the proofs contained below, we make extensive use of the axioms and rules of the logic $\mathbf{CLuNs}^{\perp}$ as defined in Appendix C.

E.1 Some facts about DP

Fact 5. The following are **DP**-valid:

(i) $\vdash_{\mathbf{DP}} (\mathsf{O}A \wedge \mathsf{O}B) \supset \mathsf{O}(A \wedge B)$ ((AND) is **DP**-derivable)
(ii) $\vdash_{\mathbf{DP}} (\mathsf{O}A \wedge \mathsf{P}B) \supset \mathsf{P}(A \wedge B)$
(iii) If $\vdash_{\mathbf{DP}} A' \supset A$ then $A \supset B \vdash_{\mathbf{DP}} A' \supset B$
(iv) If $\vdash_{\mathbf{DP}} B \supset B'$ then $A \supset B \vdash_{\mathbf{DP}} A \supset B'$
(v) $\vdash_{\mathbf{DP}} A \equiv ((A \supset \bot) \supset \bot)$
(vi) $\vdash_{\mathbf{DP}} (A \supset (A \supset \bot)) \supset (A \supset \bot)$
(vii) $\vdash_{\mathbf{DP}} A \vee (A \supset \bot)$
(viii) If $\vdash_{\mathbf{DP}} A \supset B$ then $\vdash_{\mathbf{DP}} \mathsf{O}A \supset \mathsf{O}B$
(ix) $\vdash_{\mathbf{DP}} (A \supset \bot) \supset {\sim}A$
(x) $\vdash_{\mathbf{DP}} A \vee {\sim}A$
(xi) $\vdash_{\mathbf{DP}} {\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B)) \equiv ((A \wedge B) \vee ({\sim}A \wedge {\sim}B))$
(xii) $A \vee B, A \supset C, B \supset C \vdash_{\mathbf{DP}} C$
(xiii) If $\vdash_{\mathbf{DP}} A \equiv B$ then $\vdash_{\mathbf{DP}} \mathsf{P}A \equiv \mathsf{P}B$
(xiv) If $\vdash_{\mathbf{DP}} A \supset B$ then $\vdash_{\mathbf{DP}} \mathsf{P}A \supset \mathsf{P}B$
(xv) $\vdash_{\mathbf{DP}} (A \supset B) \supset {\sim}(A \wedge {\sim}B)$
(xvi) $\vdash_{\mathbf{DP}} (A \wedge B) \supset (A \equiv B)$
(xvii) $A \vee B, A \supset C, B \supset D \vdash_{\mathbf{DP}} C \vee D$

Proof. By (A∧3), $\vdash_{\mathbf{DP}} A \supset (B \supset (A \wedge B))$. By (NEC), it follows that $\vdash_{\mathbf{DP}} \mathsf{O}(A \supset (B \supset (A \wedge B)))$. By (K), (†) $\vdash_{\mathbf{DP}} \mathsf{O}A \supset \mathsf{O}(B \supset (A \wedge B))$.

Ad (i): Suppose $\mathsf{O}A$ and $\mathsf{O}B$. By (†) and (MP), $\mathsf{O}(B \supset (A \wedge B))$. By (K), $\mathsf{O}B \supset \mathsf{O}(A \wedge B)$. By (MP), $\mathsf{O}(A \wedge B)$. The rest follows by Theorem 26.

Ad (ii): Suppose $\mathsf{O}A$ and $\mathsf{P}B$. By (†) and (MP), $\mathsf{O}(B \supset (A \wedge B))$. By (KP), $\mathsf{P}B \supset \mathsf{P}(A \wedge B)$. By (MP) we get $\mathsf{P}(A \wedge B)$. The rest follows by Theorem 26.

Ad (iii): Suppose $\vdash_{\mathbf{DP}} A' \supset A$. By (A⊃2), $\vdash_{\mathbf{DP}} (A' \supset (A \supset B)) \supset ((A' \supset A) \supset (A' \supset B))$. By (A⊃1) and (MP), $A \supset B \vdash_{\mathbf{DP}} A' \supset (A \supset B)$. The rest follows by multiple applications of (MP).

Ad (iv): The proof is similar and left to the reader.

Ad (v): *Left-to-right:* By (MP), $A, A \supset \bot \vdash_{\mathbf{DP}} \bot$. The rest follows by Theorem 26.

Right-to-left: By (iv) and (A⊥1), $(A \supset \bot) \supset \bot \vdash_{\mathbf{DP}} (A \supset \bot) \supset A$. By (A⊃3) and (MP), $(A \supset \bot) \supset \bot \vdash_{\mathbf{DP}} A$.

Ad (vi): By (MP), $A, A \supset (A \supset \bot) \vdash_{\mathbf{DP}} A \supset \bot$. By (MP), $A, A \supset (A \supset \bot) \vdash_{\mathbf{DP}} \bot$. By Theorem 26, $A \supset (A \supset \bot) \vdash_{\mathbf{DP}} A \supset \bot$, $\vdash_{\mathbf{DP}} (A \supset (A \supset \bot)) \supset (A \supset \bot)$.

Ad (vii): By (A∨1), $A \supset (A \vee (A \supset \bot))$. By (iii), $(A \vee (A \supset \bot)) \supset \bot \vdash_{\mathbf{DP}} A \supset \bot$. By Theorem 26, $\vdash_{\mathbf{DP}} ((A \vee (A \supset \bot)) \supset \bot) \supset (A \supset \bot)$. By (A∨2), $(A \supset \bot) \supset (A \vee (A \supset \bot))$. Hence, by (iv), $\vdash_{\mathbf{DP}} ((A \vee (A \supset \bot)) \supset \bot) \supset (A \vee (A \supset \bot))$. By (v), $\vdash_{\mathbf{DP}} (A \vee (A \supset \bot)) \equiv (((A \vee (A \supset \bot)) \supset \bot) \supset \bot)$. Thus, by (iv), $\vdash_{\mathbf{DP}} ((A \vee (A \supset \bot)) \supset \bot) \supset (((A \vee (A \supset \bot)) \supset \bot) \supset \bot)$. By (vi) and (MP), $\vdash_{\mathbf{DP}} ((A \vee (A \supset \bot)) \supset \bot) \supset \bot$. By (v), (A≡2), and (MP), $\vdash_{\mathbf{DP}} A \vee (A \supset \bot)$.

Ad (viii): Let $\vdash_{\mathbf{DP}} A \supset B$. By (NEC), $\vdash_{\mathbf{DP}} \mathsf{O}(A \supset B)$. By (K) and (MP), $\vdash_{\mathbf{DP}} \mathsf{O}A \supset \mathsf{O}B$.

Ad (ix): Suppose (†) $A \supset \bot$. Suppose ($\star$) A then by (†) and (MP), $\bot$. By (A⊥1), $\bot \supset {\sim}A$. By (MP), ${\sim}A$. By Theorem 26 and supposition ($\star$), $A \supset {\sim}A$. By (A~1) and (MP), ${\sim}A$. By Theorem 26 and supposition (†), $(A \supset \bot) \supset {\sim}A$.

Ad (x): This follows by simple propositional manipulations via (ix) and (vii).

Ad (xi): By (A~∧), $\vdash_{\mathbf{DP}} {\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B)) \equiv ({\sim}(A \vee B) \vee {\sim}({\sim}A \vee {\sim}B))$. By (A~∨), $\vdash_{\mathbf{DP}} {\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B)) \equiv (({\sim}A \wedge {\sim}B) \vee ({\sim}{\sim}A \wedge {\sim}{\sim}B))$. By (A~~) and some simple propositional manipulations, $\vdash_{\mathbf{DP}} {\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B)) \equiv ((A \wedge B) \vee ({\sim}A \wedge {\sim}B))$.

Ad (xii): Suppose $A \vee B$, $A \supset C$ and $B \supset C$. By the latter two, (A∨3) and simple propositional manipulations, $(A \vee B) \supset C$. By (MP), C.

Ad (xiii): Let $\vdash_{\mathbf{DP}} A \equiv B$. By (A≡1) and (A≡2) we get $\vdash_{\mathbf{DP}} A \supset B$ and $\vdash_{\mathbf{DP}} B \supset A$. Hence, by (NEC), $\vdash_{\mathbf{DP}} \mathsf{O}(A \supset B)$ and $\vdash_{\mathbf{DP}} \mathsf{O}(B \supset A)$. By (KP) and (MP), $\vdash_{\mathbf{DP}} \mathsf{P}A \supset \mathsf{P}B$ and $\vdash_{\mathbf{DP}} \mathsf{P}B \supset \mathsf{P}A$. By (A≡3) and (MP), $\vdash_{\mathbf{DP}} \mathsf{P}A \equiv \mathsf{P}B$.

Ad (xiv): Similar to the previous proof.

Ad (xv): Suppose (1) $A \supset B$. By (x), (2) $A \vee {\sim}A$. Suppose (3) A. By (1) and (MP), B. By (A~~), ${\sim}{\sim}B$. By (A∨2), ${\sim}{\sim}B \supset ({\sim}A \vee {\sim}{\sim}B)$ and whence by (MP), ${\sim}A \vee {\sim}{\sim}B$. By Theorem 26 and supposition (3), (4) $A \supset ({\sim}A \vee {\sim}{\sim}B)$. By (A∨1), (5) ${\sim}A \supset ({\sim}A \vee {\sim}{\sim}B)$. By (2), (4), (5) and (xii), ${\sim}A \vee {\sim}{\sim}B$. By (A~∧), ${\sim}(A \wedge {\sim}B)$. By Theorem 26 and supposition (1), $(A \supset B) \supset {\sim}(A \wedge {\sim}B)$.

Ad (xvi): Suppose $A \wedge B$. By (A∧1) and (MP), A. By (A⊃1) and (MP), $B \supset A$. In an analogous way we get $A \supset B$. By (A≡3) and some simple propositional manipulations, $A \equiv B$.

Ad (xvii): Suppose $A \vee B$, $A \supset C$ and $B \supset D$. By (A∨1), $C \supset (C \vee D)$. By (A∨2), $D \supset (C \vee D)$. By (iv), $A \supset (C \vee D)$ and $B \supset (C \vee D)$. The rest follows by (xii). □

Lemma 1. *The following are* **DP***-valid:*

(i) $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \vee B) \equiv {\sim}\mathsf{O}({\sim}A \wedge {\sim}B)$.
(ii) $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \wedge B) \equiv {\sim}\mathsf{O}({\sim}A \vee {\sim}B)$.
(iii) $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \supset B) \supset {\sim}\mathsf{O}(A \wedge {\sim}B)$
(iv) $\vdash_{\mathbf{DP}} (A \equiv B) \supset {\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B))$.
(v) $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \equiv B) \supset {\sim}\mathsf{O}((A \vee B) \wedge ({\sim}A \vee {\sim}B))$.

Proof. Ad (i): By (A${\sim}\wedge$), ($\dagger_1$) $\vdash_{\mathbf{DP}} ({\sim}{\sim}A \vee {\sim}{\sim}B) \equiv {\sim}({\sim}A \wedge {\sim}B)$. By means of (A${\sim}{\sim}$) it is easy to see that ($\dagger_2$) $\vdash_{\mathbf{DP}} (A \vee B) \equiv ({\sim}{\sim}A \vee {\sim}{\sim}B)$. By ($\dagger_1$), ($\dagger_2$) and some simple manipulations, $\vdash_{\mathbf{DP}} (A \vee B) \equiv {\sim}({\sim}A \wedge {\sim}B)$. Hence, by Fact 5, $\vdash_{\mathbf{DP}} \mathsf{P}(A \vee B) \equiv \mathsf{P}{\sim}({\sim}A \wedge {\sim}B)$. By the definition of P, $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \vee B) \equiv {\sim}\mathsf{O}{\sim}{\sim}({\sim}A \wedge {\sim}B)$. By (A${\sim}{\sim}'$) and some simple manipulations, $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \vee B) \equiv {\sim}\mathsf{O}({\sim}A \wedge {\sim}B)$.

Ad (ii): Analogous to the previous proof.

Ad (iii): By Fact 5.xiv and 5.xv, $\vdash_{\mathbf{DP}} \mathsf{P}(A \supset B) \supset \mathsf{P}{\sim}(A \wedge {\sim}B)$. By the definition of P, $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \supset B) \supset {\sim}\mathsf{O}{\sim}{\sim}(A \wedge {\sim}B)$. By (A${\sim}{\sim}'$) and some simple modifications, $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \supset B) \supset {\sim}\mathsf{O}(A \wedge {\sim}B)$.

Ad (iv): Suppose (1) $A \equiv B$. By Fact 5.vii, (2) $A \vee (A \supset \bot)$. Suppose (3) $A \supset \bot$. By (A$\bot$1), $\bot \supset A$ and whence by (A$\equiv$3), $A \equiv \bot$. By the latter and (1) and simple propositional manipulations, $B \equiv \bot$ and whence by (A$\equiv$1), $B \supset \bot$. Hence by Fact 5.ix, (4) ${\sim}B$. Also by Fact 5.ix and (3), (5) ${\sim}A$. By Theorem 26, supposition (3), (4), (5) and some simple propositional manipulations we have (6) $(A \supset \bot) \supset ({\sim}A \wedge {\sim}B)$.

Now suppose (7) A. By (1), (A$\equiv$1) and (MP), (8) B. By Theorem 26, supposition (7), (8) and some simple propositional manipulations, (9) $A \supset (A \wedge B)$.

Now by (2), (6), (9) and Fact 5.xvii, $(A \wedge B) \vee ({\sim}A \wedge {\sim}B)$. By Fact 5.xi, ${\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B))$.

Ad (v): By Fact 5.xiv and (iv), $\vdash_{\mathbf{DP}} \mathsf{P}(A \equiv B) \supset \mathsf{P}{\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B))$. By the definition of P, $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \equiv B) \supset {\sim}\mathsf{O}{\sim}{\sim}((A \vee B) \wedge ({\sim}A \vee {\sim}B))$. By (A${\sim}{\sim}'$) and some simple manipulations, $\vdash_{\mathbf{DP}} {\sim}\mathsf{O}{\sim}(A \equiv B) \supset {\sim}\mathsf{O}((A \vee B) \wedge ({\sim}A \vee {\sim}B))$. □

E.2 Proofs of Theorems 27–28

In order to simplify the notation in the following meta-proofs we define $Rw = \{w' \mid Rww'\}$.

Lemma 2. *Where $M = \langle W, w_0, R, v\rangle$ is a* **DP***-model, we have: for all $w \in W$, if $v_M(A, w) = 0$ then $v_M({\sim}A, w) = 1$.*

Proof. We show this by an induction over the length of A. Let $A \in \mathcal{W}^a$. Suppose $v_M(A, w) = 0$. By (C${\sim}$1'), $v_M({\sim}A, w) = 1$.

For the induction step let first $A = B \wedge C$. Suppose $v_M(B \wedge C, w) = 0$. By (C$\wedge$), $v_M(B, w) = 0$ or $v_M(C, w) = 0$. By the induction hypothesis, $v_M({\sim}B, w) = 1$ or $v_M({\sim}C, w) = 1$. By (C$\vee$), $v_M({\sim}B \vee {\sim}C, w) = 1$. By (C${\sim}\wedge$), $v_M({\sim}(B \wedge C), w) = 1$. The cases $A = B \vee C$, $A = B \supset C$, $A = B \equiv C$ and $A = {\sim}B$ are similar and left to the reader. Let $A = \mathsf{O}B$. Suppose $v_M({\sim}\mathsf{O}B, w) = 0$. By (C${\sim}\mathsf{O}$) there is no $w' \in Rw$ such that $v_M({\sim}B, w') = 1$. Hence, for all $w' \in Rw$, $v_M({\sim}B, w') = 0$. By the induction hypothesis, for all $w' \in Rw$, $v_M({\sim}{\sim}B, w') = 1$. By (C${\sim}{\sim}$), for all $w' \in Rw$, $v_M(B, w') = 1$. By (CO), $v_M(\mathsf{O}B, w) = 1$. □

Lemma 3. *Where $M = \langle W, w, R, v\rangle$ is a* **DP***-model:* $v_M(\sim\mathsf{O}A, w) = 1$ *iff there is a* $w' \in Rw$ *such that* $v_M(\sim A, w') = 1$.

Proof. Case 1. If $A \in \mathcal{W}_l^{\sim} \cup \mathcal{W}^{\not\dagger}$, then this is immediate in view of (C~O).
Case 2. $A = \sim D$, where $D = (B\pi C)$ and $\pi \in \{\wedge, \vee, \supset, \equiv\}$.
Case 2.1. $\pi = \wedge$. Then $v_M(\sim\mathsf{O}\sim(B \wedge C), w) = 1$ iff (by (C~∧′)) $v_M(\sim\mathsf{O}(\sim B \vee \sim C), w) = 1$ iff (by (C~O)) there is a $w' \in Rw$ such that $v_M(\sim(\sim B \vee \sim C), w') = 1$ iff (by (C~∧)) $v_M(\sim\sim B \wedge \sim\sim C, w') = 1$ iff (by (C∧)) $v_M(\sim\sim B, w') = v_M(\sim\sim C, w') = 1$ iff (by (C~~)) $v_M(B, w') = v_M(C, w') = 1$ iff (by (C∧)) $v_M(B \wedge C, w') = 1$ iff (by (C~~)) $v_M(\sim\sim(B \wedge C), w') = 1$ iff $v_M(\sim A, w') = 1$.

The cases for $\pi \in \{\vee, \supset, \equiv\}$ are analogous and are safely left to the reader.
Case 3. $A = \sim\sim D$. Let D' be the result of removing all pairs '~~' by which D is prefixed. Then either $D' \in \mathcal{W}_l^{\sim} \cup \mathcal{W}^{\not\dagger}$ or $D' = \sim(B\pi C)$. $v_M(\sim\mathsf{O}A, w) = 1$ iff (by (multiple applications of) (C~~′)) $v_M(\sim\mathsf{O}D', w) = 1$ iff (by Case 1 or 2) there is a $w' \in Rw$ such that $v_M(\sim D', w') = 1$ iff (by (multiple applications of) (C~~)) $v_M(\sim A, w') = 1$. □

Proof of Theorem 27. Let in the following $M = \langle W, w_0, R, v\rangle$ be a **DP**-model.

It is easy to see that all instances of (A⊃1), (A⊃2), (A⊃3) hold in M due to (C⊃). For instance, $M \Vdash A \supset (B \supset A)$ iff ($M \nVdash A$ or $M \Vdash B \supset A$) iff ($M \nVdash A$ or ($M \nVdash B$ or $M \Vdash A$)) iff ($M \nVdash A$ or $M \Vdash A$). Thus, $M \Vdash A \supset (B \supset A)$.

Similarly it can be shown that all instances of (A∧1), (A∧2) and (A∧3) hold in M due to (C⊃) and (C∧); all instances of (A∨1), (A∨2), and (A∨3) hold in M due to (C∨) and (C⊃); all instances of (A≡1), (A≡2), and (A≡3) hold in M due to (C≡) and (C⊃); where $\pi \in \{\sim, \supset, \wedge, \vee, \equiv\}$, all instances of (A~$\pi$) and of (A~$\pi'$) hold in M due to (C~π) and (C~π'); (A⊥1) holds in M due to (C⊥) and (C⊃).

Ad (A⊥2): Note that by (C⊥) there is no $w \in W$ for which $v_M(\bot, w) = 1$. Thus, by (C~~), there is no $w \in W$ such that $v_M(\sim\sim\bot, w) = 1$. Hence, by (C~O), $v_M(\sim\mathsf{O}\sim\bot, w) = 0$ for all $w \in W$. Hence, $v_M(\mathsf{P}\bot, w) = 0$ for all $w \in W$. Hence by (C⊃), $v_M(\mathsf{P}\bot \supset A, w_0) = 1$, and whence, $M \Vdash \mathsf{P}\bot \supset A$.

Ad (A~1): We have $M \Vdash (A \supset \sim A) \supset \sim A$ iff ($M \nVdash A \supset \sim A$ or $M \Vdash \sim A$) iff (not ($M \nVdash A$ or $M \Vdash \sim A$) or $M \Vdash \sim A$) iff (($M \Vdash A$ and $M \nVdash \sim A$) or $M \Vdash \sim A$) iff ($M \Vdash A$ or $M \Vdash \sim A$). The latter holds due to Lemma 2.

Ad (D): Suppose $M \Vdash \mathsf{O}A$. Then for all $w \in Rw_0$, $v_M(A, w) = 1$. By (C~~) and the seriality of R, there is a $w \in Rw_0$ for which $v_M(\sim\sim A, w) = 1$. By Lemma 3, $M \Vdash \sim\mathsf{O}\sim A$ and whence $M \Vdash \mathsf{P}A$. Altogether by (C⊃), $M \Vdash \mathsf{O}A \supset \mathsf{P}A$.

Ad (K): Suppose $M \Vdash \mathsf{O}(A \supset B)$. By (CO) and (C⊃), for all $w \in Rw_0$, $v_M(A, w) = 0$ or $v_M(B, w) = 1$. Suppose $M \Vdash \mathsf{O}A$, then for all $w \in Rw_0$, $v_M(A, w) = 1$. Hence, for all $w \in Rw_0$, $v_M(B, w) = 1$. Thus by (CO), $M \Vdash \mathsf{O}B$. Hence, by (C⊃), $M \Vdash \mathsf{O}A \supset \mathsf{O}B$. Altogether by (C⊃), $M \Vdash \mathsf{O}(A \supset B) \supset (\mathsf{O}A \supset \mathsf{O}B)$.

Ad (KP): Suppose $M \Vdash \mathsf{O}(A \supset B)$. By (CO) and (C⊃), (†) for all $w \in Rw_0$, $v_M(A, w) = 0$ or $v_M(B, w) = 1$. Suppose $M \Vdash \mathsf{P}A$ and whence $M \Vdash \sim\mathsf{O}\sim A$. Hence, by Lemma 3 and (C~~), (‡) there is a $w \in Rw_0$ for which $v_M(A, w) = 1$. Hence, by (†), (‡) and (C~~), there is a $w \in Rw_0$ such that $v_M(\sim\sim B, w) = 1$. Thus, by Lemma 3, $M \Vdash \sim\mathsf{O}\sim B$ and whence $M \Vdash \mathsf{P}B$. Hence, by (C⊃), $M \Vdash \mathsf{P}A \supset \mathsf{P}B$. Altogether by (C⊃), $M \Vdash \mathsf{O}(A \supset B) \supset (\mathsf{P}A \supset \mathsf{P}B)$.

Ad (OD): Suppose $M \Vdash \mathsf{O}(A \vee B)$. Hence, by (C$\mathsf{O}$) and (C$\vee$), (†) for all $w \in Rw_0$, $v_M(A, w) = 1$ or $v_M(B, w) = 1$. Suppose $M \nVdash \mathsf{P}B$ and whence $M \nVdash {\sim}\mathsf{O}{\sim}B$. By Lemma 3 and (C$\sim\sim$), (‡) for all $w \in Rw_0$, $v_M(B, w) = 0$. By (†) and (‡), for all $w \in Rw_0$, $v_M(A, w) = 1$. Thus $M \Vdash \mathsf{O}A$. Hence, by (C$\vee$), $M \Vdash \mathsf{O}A \vee \mathsf{P}B$. Altogether, by (C$\supset$), $M \Vdash \mathsf{O}(A \vee B) \supset (\mathsf{O}A \vee \mathsf{P}B)$.

Ad (PD): This is similar to the previous case and is left to the reader.

Ad (NEC): Let $\vDash_{\mathbf{DP}} A$. Suppose $v_M(A, w) = 0$ for some $w \in W$. However, then $\langle W, w, R, v\rangle \nVdash A$,— a contradiction. Hence, $v_M(A, w) = 1$ for all $w \in W$. Thus, by (CO), $M \Vdash \mathsf{O}A$.

We now know that all the axioms of **DP** are semantically valid. That $\Gamma \vdash_{\mathbf{DP}} A$ implies $\Gamma \vDash_{\mathbf{DP}} A$ can now be shown via the usual induction on the length of the proof of A. This is safely left to the reader. □

Let for the remainder W^c be the **DP**-deductively closed and maximally **DP**-non-trivial subsets of $\mathcal{W}^{\mathbf{DP}}$. For the completeness proof of **DP**, we make use of the following lemmas.[1]

Lemma 4. *If $\Delta \in W^c$, then (i) Δ is prime, i.e. $A \vee B \in \Delta$ iff $A \in \Delta$ or $B \in \Delta$; (ii) if $A \notin \Delta$ then ${\sim}A \in \Delta$; (iii) $A \notin \Delta$ iff $A \supset \bot \in \Delta$; (iv) $A, B \in \Delta$ iff $A \wedge B \in \Delta$; (v) $A \equiv B \in \Delta$ iff $A, B \in \Delta$ or $A, B \notin \Delta$.*

Proof. Ad (i): Suppose that, for a $\Delta \in W^c$, $A \vee B \in \Delta$ and $A \notin \Delta$ and $B \notin \Delta$. Then, since Δ is maximally **DP**-non-trivial, $\Delta \cup \{A\}$ is trivial and $\Delta \cup \{B\}$ is trivial. Then, for any C, $\Delta \cup \{A\} \vdash_{\mathbf{DP}} C$ and $\Delta \cup \{B\} \vdash_{\mathbf{DP}} C$. Then, by Theorem 26, $\Delta \vdash_{\mathbf{DP}} A \supset C$ and $\Delta \vdash_{\mathbf{DP}} B \supset C$. But then, by (MP) and (A$\vee$3), $\Delta \vdash_{\mathbf{DP}} (A \vee B) \supset C$. Since $A \vee B \in \Delta$, by (MP) $\Delta \vdash_{\mathbf{DP}} C$, and since Δ is **DP**-deductively closed, $C \in \Delta$. This contradicts the supposition. Hence if $A \vee B \in \Delta$, then $A \in \Delta$ or $B \in \Delta$. The other direction is shown in a similar way. This is left to the reader.

Ad (ii): Suppose $A \notin \Delta$. Since $\Delta \in W^c$, $\Delta \cup \{A\}$ is **DP**-trivial. Hence, $\Delta \cup \{A\} \vdash_{\mathbf{DP}} {\sim}A$. By Theorem 26, $\Delta \vdash_{\mathbf{DP}} A \supset {\sim}A$. By (A$\sim$1) and (MP), $\Delta \vdash_{\mathbf{DP}} {\sim}A$. Since Δ is **DP**-deductively closed, ${\sim}A \in \Delta$.

Ad (iii): Suppose $A \notin \Delta$. Assume $A \supset \bot \notin \Delta$. By (ii), ${\sim}(A \supset \bot) \in \Delta$. By (A$\sim\supset$) and the deductive closure of Δ, $A \wedge {\sim}\bot \in \Delta$ and whence by (A$\wedge$1) and (MP) $A \in \Delta$,—a contradiction. Suppose now that $A \supset \bot \in \Delta$. Assume $A \in \Delta$ then by (MP), $\bot \in \Delta$,—a contradiction to the fact that Δ is consistent.

Ad (iv): This follows by means of (A$\wedge$3), (A$\wedge$1) and (A$\wedge$2).

Ad (v): Suppose $A \equiv B \in \Delta$. Assume that neither $A, B \in \Delta$ nor $A, B \notin \Delta$. Without loss of generality let $A \notin \Delta$. Hence by the assumption $B \in \Delta$. Since $A \equiv B \in \Delta$ and by (A$\equiv$2) and (MP), $B \supset A$. By (MP), $A \in \Delta$,—a contradiction. Suppose now that $A, B \in \Delta$. By Fact 5.xvi and (MP), $A \equiv B \in \Delta$. Suppose now that $A, B \notin \Delta$. Assume $A \equiv B \notin \Delta$. By (ii), ${\sim}(A \equiv B) \in \Delta$. By (A$\sim\equiv$) and (MP) also $(A \vee B) \wedge ({\sim}A \vee {\sim}B) \in \Delta$. By (iv), $A \vee B \in \Delta$. By (i), $A \in \Delta$ or $B \in \Delta$,—a contradiction. □

[1] The proof of Lemma 6 is an adaptation of the proof of Lemma 1.7.1 from [19]. The proofs of Lemma 7 and Lemma 9 rely on insights from [110, pp. 205-207] and [144, p. 341].

Where $\Gamma \in W^c$ and $A \in \mathcal{W}^{\sim}_{\perp}$, we will use the following abbreviations: $\Gamma^A_O = \{B \mid \mathsf{O}B \in \Gamma\} \cup \{A\}$ and $\Gamma_P = \{B \mid \mathsf{P}B \notin \Gamma\}$.

Lemma 5. *Let* $\Gamma \in W^c$. *(i) If* $\mathsf{P}A \in \Gamma$ *then* $\mathrm{Cn}_{\mathbf{DP}}(\Gamma^A_O) \cap \Gamma_P = \varnothing$*; (ii) If* $\mathsf{P}A \in \Gamma$ *then* Γ^A_O *is* **DP***-non-trivial; (iii) If* $\mathsf{O}A \notin \Gamma$*, then* $\mathsf{P}(A \supset \perp) \in \Gamma$*; (iv) If* $\mathsf{O}A \notin \Gamma$*, then* $\Gamma^{A\supset\perp}_O$ *is* **DP***-non-trivial.*

Proof. Ad (i): We show by reductio that $\Gamma^A_O \nvdash_{\mathbf{DP}} C$ for all $C \in \Gamma_P$. Suppose thus that $\Gamma^A_O \vdash_{\mathbf{DP}} C$ for some $C \in \Gamma_P$. Then $\Gamma' \vdash_{\mathbf{DP}} C$ for some finite $\Gamma' \subseteq \Gamma^A_O$ (given the compactness of **DP**). Then $\Gamma' \cup \{A\} \vdash_{\mathbf{DP}} C$ by the monotonicity of **DP**. Then $\vdash_{\mathbf{DP}} (\bigwedge \Gamma' \wedge A) \supset C$ by Theorem 26. Then $\vdash_{\mathbf{DP}} \mathsf{O}((\bigwedge\Gamma' \wedge A) \supset C)$ by (NEC). Then $\vdash_{\mathbf{DP}} \mathsf{P}(\bigwedge\Gamma' \wedge A) \supset \mathsf{P}C$ by (KP) and (MP). By the supposition, $\{\mathsf{O}B \mid B \in \Gamma'\} \subseteq \Gamma$ and $\mathsf{P}A \in \Gamma$. Given the deductive closure of Γ and $\vdash_{\mathbf{DP}} (\mathsf{O}(\bigwedge\Gamma') \wedge \mathsf{P}A) \supset \mathsf{P}(\bigwedge\Gamma' \wedge A)$ (which follows from Fact 5 (ii)), it follows that $\mathsf{P}(\bigwedge\Gamma' \wedge A) \in \Gamma$. Hence $\mathsf{P}C \in \Gamma$, since Γ is deductively closed and $\vdash_{\mathbf{DP}} \mathsf{P}(\bigwedge\Gamma' \wedge A) \supset \mathsf{P}C$. But $\mathsf{P}C \notin \Gamma$ in view of the construction of Γ_P. This is a contradiction. So $\Gamma^A_O \nvdash_{\mathbf{DP}} C$ for all formulas $C \in \Gamma_P$. (ii) follows immediately due to (i) and the fact that $\perp \in \Gamma_P$.

Ad (iii): Assume $\mathsf{O}A \notin \Gamma$ and $\mathsf{P}(A \supset \perp) \notin \Gamma$. Then, by Lemma 4 (ii), ${\sim}\mathsf{P}(A \supset \perp) \in \Gamma$. Hence by the definition of P, (A~~) and the deductive closure of Γ, $\mathsf{O}{\sim}(A \supset \perp) \in \Gamma$. By (A~⊃) and (A∧1), $\vdash_{\mathbf{DP}} {\sim}(A \supset \perp) \supset A$. By Fact 5 (viii), $\vdash_{\mathbf{DP}} \mathsf{O}{\sim}(A \supset \perp) \supset \mathsf{O}A$. By (MP), $\mathsf{O}A \in \Gamma$,—a contradiction. Hence $\mathsf{P}(A \supset \perp) \in \Gamma$. (iv) follows by (ii) and (iii). □

Lemma 6. *Let* $\Gamma \in W^c$ *and* $\Gamma \vdash_{\mathbf{DP}} \mathsf{P}A$*. There is a* $\Delta \subseteq \mathcal{W}^{\mathbf{DP}}$ *for which (i)* $\Gamma^A_O \subseteq \Delta$*, (ii)* $\Gamma_P \cap \Delta = \varnothing$*, and (iii)* $\Delta \in W^c$*.*

Proof. Where $\langle B_1, B_2, \ldots\rangle$ is a list of the members of $\mathcal{W}^{\mathbf{DP}}$, define

$$\Delta_0 = \mathrm{Cn}_{\mathbf{DP}}(\Gamma^A_O)$$
$$\Delta_{i+1} = \begin{cases} \mathrm{Cn}_{\mathbf{DP}}(\Delta_i \cup \{B_{i+1}\}) & \text{if, for all } B \in \Gamma_P, \\ & B \notin \mathrm{Cn}_{\mathbf{DP}}(\Delta_i \cup \{B_{i+1}\}) \\ \Delta_i & \text{otherwise} \end{cases}$$
$$\Delta = \Delta_0 \cup \Delta_1 \cup \ldots$$

Ad (i): this holds by the construction.

Ad(ii): By Lemma 5 (i) $\Delta_0 \cap \Gamma_P = \varnothing$. The rest follows by the construction of the Δ_i's.

Ad (iii): We first show that if $B_i \notin \Delta$ then $B_i \supset \perp \in \Delta$. Suppose $B_i \notin \Delta$. Hence, by the construction and the monotonicity of **DP**, (†) $\Delta \cup \{B_i\} \vdash_{\mathbf{DP}} B$ for some $B \in \Gamma_P$. By Theorem 26, $\Delta \vdash_{\mathbf{DP}} B_i \supset B$. Assume that $B_i \supset \perp \notin \Delta$. Hence, by the construction and the monotonicity of **DP**, $\Delta \cup \{B_i \supset \perp\} \vdash_{\mathbf{DP}} C$ for some $C \in \Gamma_P$. By Theorem 26, (‡) $\Delta \vdash_{\mathbf{DP}} (B_i \supset \perp) \supset C$. By simple propositional manipulations, (†) and (‡), $\Delta \vdash_{\mathbf{DP}} (B_i \vee (B_i \supset \perp)) \supset (B \vee C)$. By Fact 5 (vii), $\vdash_{\mathbf{DP}} B_i \vee (B_i \supset \perp)$. By (MP), $\Delta \vdash_{\mathbf{DP}} B \vee C$. By (ii), $B \vee C \notin \Gamma_P$. Hence, $\Gamma \vdash_{\mathbf{DP}} \mathsf{P}(B \vee C)$. By (PD), $\Gamma \vdash_{\mathbf{DP}} \mathsf{P}B \vee \mathsf{P}C$. By Lemma 4 (i), $\mathsf{P}B \in \Gamma$ or $\mathsf{P}C \in \Gamma$,—a contradiction. Hence, $B_i \supset \perp \in \Delta$.

Note that by Lemma 5 (ii), Δ_0 is **DP**-non-trivial. Hence, Δ is **DP**-non-trivial by the construction. Suppose $B \notin \Delta$. By Lemma 4 (iii), $B \supset \perp \in \Delta$. Thus, by (MP) and (A⊥1), $\Delta \cup \{B\}$ is **DP**-trivial. Hence, Δ is maximally **DP**-non-trivial. □

Definition 18. The binary relation $\mathsf{R} \subseteq (W^c \times W^c)$ is defined as follows: $\mathsf{R}\Gamma\Delta$ iff, the following two conditions are met

(a) if $\mathsf{O}A \in \Gamma$ then $A \in \Delta$, and

(b) if $A \in \Delta$ then $\mathsf{P}A \in \Gamma$.

Lemma 7. *Let $\Gamma \in W^c$. $\mathsf{P}A \in \Gamma$ iff there is a $\Delta \in W^c$ for which $\mathsf{R}\Gamma\Delta$ and $A \in \Delta$.*

Proof. Left-right: Suppose $\mathsf{P}A \in \Gamma$. Then, by Lemma 6, there is a $\Delta \subseteq \mathcal{W}^{\mathbf{DP}}$ such that (i) $\Gamma^A_O \subseteq \Delta$, (ii) for all $C \in \Gamma_P$, $C \notin \Delta$, (iii) $\Delta \in W^c$. We now show that $\mathsf{R}\Gamma\Delta$. Ad (a): if, for some D, $\mathsf{O}D \in \Gamma$ then $D \in \Gamma^A_O$, hence $D \in \Delta$ by (i). Ad (b): suppose $\mathsf{P}E \notin \Gamma$ for some E. Then $E \in \Gamma_P$, hence $E \notin \Delta$ by (ii).

Right-left: Follows directly by Definition 19. □

Lemma 8. *Let $\Gamma \in W^c$. There is a $\Delta \in W^c$ such that $\mathsf{R}\Gamma\Delta$ (i.e. R is serial).*

Proof. By Fact 5 (vii), $\vdash_{\mathbf{DP}} A \vee (A \supset \bot)$ and hence by (NEC), (D) and (MP), $\vdash_{\mathbf{DP}} \mathsf{P}(A \vee (A \supset \bot))$. Thus, $\mathsf{P}(A \vee (A \supset \bot)) \in \Gamma$. By Lemma 7, there is a $\Delta \in W^c$ such that $\mathsf{R}\Gamma\Delta$ and $A \vee (A \supset \bot) \in \Delta$. □

Lemma 9. *Where $\Gamma \in W^c$, $\mathsf{O}A \in \Gamma$ iff, for all $\Delta \in W^c$ such that $\mathsf{R}\Gamma\Delta$, $A \in \Delta$.*

Proof. Left-right: This is an immediate consequence of Definition 19.
Right-left: For some $\Gamma \in W^c$ and some $A \in \mathcal{W}^{\sim}_{\bot}$, suppose that (†) $\mathsf{O}A \notin \Gamma$ and $\mathsf{R}\Gamma\Delta$. By Lemma 5 (iii), $\mathsf{P}(A \supset \bot) \in \Gamma$. By Lemma 7 there is a Δ such that $\mathsf{R}\Gamma\Delta$ and $A \supset \bot \in \Delta$. Since Δ is **DP**-non-trivial, $A \notin \Delta$. □

Lemma 10. *If $\Delta \in W^c$, then there is a **DP**-model M such that $M \Vdash A$ for all $A \in \Delta$ and $M \nVdash A$ for all $A \in \mathcal{W}^{\mathbf{DP}} - \Delta$.*

Proof. Let $\Delta \in W^c$. We construct a **DP**-model $M = \langle W^c, \Delta, \mathsf{R}, v\rangle$ where v is defined as follows:

(i) For all $A \in \mathcal{W}^a$ and all $w \in W^c$, $v(A, w) = 1$ iff $A \in w$

(ii) For all $A \in \mathcal{W}^a$ and all $w \in W^c$, $v({\sim}A, w) = 1$ iff $A, {\sim}A \in w$

Note that due to the seriality of R (Lemma 8) M is indeed a **DP**-model. We now show that:

(*) for all $A \in \mathcal{W}^{\mathbf{DP}}$ and for all $w \in W^c$, $v_M(A, w) = 1$ iff $A \in w$

The proof proceeds by an induction on the complexity of A. Let $w \in W^c$. If $A \in \mathcal{W}^a$, then $v(A, w) = 1$ iff $A \in w$ (by (i)). Then, by (C_a), $v_M(A, w) = 1$ iff $A \in w$. Hence, (*) is valid for all $A \in \mathcal{W}^a$ and all $w \in W^c$.

We proceed with the induction step. Depending on the logical form of A, we distinguish 6 cases (for the 5 connectives $\sim, \vee, \wedge, \supset, \equiv$ and for O) and show for each of them that $v_M(A, w) = 1$ iff $A \in w$.

Case 1: A is of the form ${\sim}B$.

Case 1.1: Suppose $B \in \mathcal{W}^a$. $v_M(\sim B, w) = 1$ iff [by (C~1')] ($v_M(B, w) = 0$ or $v(\sim B, w) = 1$) iff [by the induction hypothesis and (ii)] ($B \notin w$ or $B, \sim B \in w$) iff [by Lemma 4(ii)] $\sim B \in w$.

Case 1.2: Suppose $B \in \mathcal{W}^{\mathbf{DP}} \setminus \mathcal{W}^a$.

Case 1.2.1: If $B = \mathsf{O}C$, then $v_M(\sim\mathsf{O}C, w) = 1$ iff (by Lemma 3) there is a $w' \in \mathsf{R}w$ for which $v_M(\sim C, w') = 1$ iff [by the induction hypothesis] there is a $w' \in \mathsf{R}w$ for which $\sim C \in w'$ iff [by Definition 19] $\mathsf{P}\sim C \in w$ iff [by the definition of P] $\sim\mathsf{O}\sim\sim C \in w$ iff [by (A~~')] $\sim\mathsf{O}C \in w$.

Case 1.2.2: Otherwise B is of one of the following forms: $C \wedge D$, $C \vee D$, $C \supset D$, $C \equiv D$, or $\sim C$. Let $B = C \wedge D$. $v_M(\sim(C \wedge D), w) = 1$ iff [by (C~∧)] $v_M(\sim C \vee \sim D, w) = 1$ iff [by (C∨)] ($v_M(\sim C, w) = 1$ or $v_M(\sim D, w) = 1$) iff [by the induction hypothesis] ($\sim C \in w$ or $\sim D \in w$) iff [by Lemma 4(i)] $\sim C \vee \sim D \in w$ iff [by (A~∧) and since w is deductively closed] $\sim(C \wedge D) \in w$.

To give another example let $B = C \equiv D$. $v_M(\sim(C \equiv D), w) = 1$ iff [by (C~≡)] $v_M((C \vee D) \wedge (\sim C \vee \sim D), w) = 1$ iff [by (C∧)] $v_M(C \vee D, w) = v_M(\sim C \vee \sim D, w) = 1$ iff [by (C∨)] (($v_M(C, w) = 1$ or $v_M(D, w) = 1$) and ($v_M(\sim C, w) = 1$ or $v_M(\sim D, w) = 1$)) iff [by the induction hypothesis] (($C \in w$ or $D \in w$) and ($\sim C \in w$ or $\sim D \in w$)) iff [Lemma 4.i] ($C \vee D \in w$ and $\sim C \vee \sim D \in w$) iff [by Lemma 4.iv] $(C \vee D) \wedge (\sim C \vee \sim D) \in w$ iff [by (A~≡)] $\sim(C \equiv D) \in w$.

The other cases are analogous and are left to the reader.

Case 2. A is of the form $B \vee C$. $v_M(B \vee C, w) = 1$ iff [by (C∨)] ($v_M(B, w) = 1$ or $v_M(C, w) = 1$) iff [by the induction hypothesis] ($B \in w$ or $C \in w$) iff [by Lemma 4(i)] $B \vee C \in w$.

Case 3. A is of the form $B \equiv C$. $v_M(B \equiv C, w) = 1$ iff [by (C≡)] $v_M(B, w) = v_M(C, w)$ iff ((i) $v_M(B, w) = v_M(C, w) = 1$) or ((ii) $v_M(B, w) = v_M(C, w) = 0$) iff [by the induction hypotheses] (((i) $B, C \in w$) or ((ii) $B, C \notin w$)) iff [by Lemma 4.v] $B \equiv C \in w$.

The proof for the other classical connectives (cases 4-5) is similar and left to the reader. We proceed with the modal operator O.

Case 6. $\mathsf{O}A \in w$ iff [by Lemma 9] $A \in w'$ for all $w' \in \mathsf{R}w$ iff [by the induction hypothesis] $v_M(A, w') = 1$ for all $w' \in \mathsf{R}w$ iff [by (CO)] $v_M(\mathsf{O}A, w) = 1$.

The rest follows since our actual world is Δ and due to (*). □

Lemma 11. *Let $\Gamma \subseteq \mathcal{W}^{\mathbf{DP}}$ and $\Gamma \nvdash_{\mathbf{DP}} A$. There is a $\Delta \subseteq \mathcal{W}^{\mathbf{DP}}$ such that (i) $\Gamma \subseteq \Delta$, (ii) $A \notin \Delta$, and (iii) $\Delta \in W^c$.*

Proof. Where $\langle B_1, B_2, \ldots \rangle$ is a list of the members of $\mathcal{W}^{\mathbf{DP}}$, define

$$\Delta_0 = \mathrm{Cn}_{\mathbf{DP}}(\Gamma)$$

$$\Delta_{i+1} = \begin{cases} \mathrm{Cn}_{\mathbf{DP}}(\Delta_i \cup \{B_{i+1}\}) & \text{if } A \notin \mathrm{Cn}_{\mathbf{DP}}(\Delta_i \cup \{B_{i+1}\}) \\ \Delta_i & \text{else} \end{cases}$$

$$\Delta = \Delta_0 \cup \Delta_1 \cup \ldots$$

Ad (i): This holds by the definition of Δ_0.

Ad (ii): This holds by the construction and since $A \notin \mathrm{Cn}_{\mathbf{DP}}(\Gamma)$.

Ad (iii): Suppose $B \notin \Delta$. Assume that $B \supset \bot \notin \Delta$. By the construction of Δ and the monotonicity of **DP**, $\Delta \cup \{B \supset \bot\} \vdash_{\mathbf{DP}} A$ and whence by Theorem 26, $\Delta \vdash_{\mathbf{DP}} (B \supset \bot) \supset A$. Analogously, $\Delta \vdash_{\mathbf{DP}} B \supset A$. By (A∨3), $\Delta \vdash_{\mathbf{DP}} (B \vee (B \supset$

$\bot)) \supset A$. Since by Fact 5 (vii), $\vdash_{\mathbf{DP}} B \vee (B \supset \bot)$, we have by (MP), $\Delta \vdash_{\mathbf{DP}} A$,—a contradiction to (ii). Hence, $B \supset \bot \in \Delta$. Thus, $\Delta \cup \{B\} \vdash_{\mathbf{DP}} \bot$ by (MP). □

Proof of Theorem 28. Suppose $\Gamma \nvdash_{\mathbf{DP}} A$. Then, by Lemma 11, there is a $\Delta \supseteq \Gamma$ such that $A \notin \Delta$ and $\Delta \in W^c$. Then, by Lemma 10, there is a **DP**-model M such that $M \Vdash B$ for all $B \in \Gamma$ and $M \nVdash A$. Hence $\Gamma \nvDash_{\mathbf{DP}} A$. □

E.3 Proof outline of Theorem 29

Lemma 12. $\vdash_{\mathbf{UDP}} (A \wedge {\sim}A) \supset B$ *for all* $A, B \in \mathcal{W}^{\mathbf{DP}}$

Proof outline: This is shown by an induction over the complexity of A. For $A \in \mathcal{W}^a \cup \mathcal{W}^{\sim}_{\mathsf{O}}$ this holds due to (UDP). For the induction step we paradigmatically consider the case $A = C \wedge D$. By the induction hypothesis, $\vdash_{\mathbf{UDP}} (C \wedge {\sim}C) \supset B$ and $\vdash_{\mathbf{UDP}} (D \wedge {\sim}D) \supset B$. By (A$\vee$3), (†) $\vdash_{\mathbf{UDP}} ((C \wedge {\sim}C) \vee (D \wedge {\sim}D)) \supset B$. By some simple propositional manipulations it is easy to see that $\vdash_{\mathbf{UDP}} ((C \wedge D) \wedge {\sim}(C \wedge D)) \supset ((C \wedge {\sim}C) \vee (D \wedge {\sim}D))$. By the latter, (†) and (MP), $\vdash_{\mathbf{UDP}} ((C \wedge D) \wedge {\sim}(C \wedge D)) \supset B$. The other cases are similar and left to the reader. □

Fact 6. (i) $A \supset (B \supset C) \vdash_{\mathbf{DP}} (A \wedge B) \supset C$ and (ii) $(A \wedge B) \supset C \vdash_{\mathbf{DP}} A \supset (B \supset C)$.

The proof of Fact 6 is straightforward in view of the definition of $\mathbf{CLuNs}^{\bot}$, and is safely left to the reader.

Let $\mathbf{SDL}^{\sim}$ be **SDL** with the negation symbol ~ (similarly for $\mathbf{CL}^{\sim}$).

Proof outline of Theorem 29: Left-right: By its definition, **UDP** contains $\mathbf{CL}^{\mathbf{pos}}$ and (A~1). Moreover, By Lemma 12 and Fact 6 (ii), $\vdash_{\mathbf{UDP}} A \supset ({\sim}A \supset B)$ for all $A, B \in \mathcal{W}^{\mathbf{DP}}$. Hence **UDP** contains **CL**. By definition it also verifies (K), (D) and (NEC).

Right-left: Since $\mathbf{SDL}^{\sim}$ is a strengthening of $\mathbf{CL}^{\sim}$, and $\mathbf{CL}^{\sim}$ is a strengthening of $\mathbf{CLuNs}^{\bot}$, $\mathbf{SDL}^{\sim}$ also strengthens $\mathbf{CLuNs}^{\bot}$. Obviously, $\mathbf{SDL}^{\sim}$ also verifies (D), (K) and (NEC). Moreover, it is easily seen that $\mathbf{SDL}^{\sim}$ verifies (KP), (OD) and (PD). Hence $\mathbf{SDL}^{\sim}$ verifies all the axioms and rules of **DP**. By Fact 6 (i) and (A~2), $\mathbf{SDL}^{\sim}$ also verifies (UDP). Hence, it verifies all the axioms and rules of **UDP**. □

Appendix F

(Meta-)properties of the logic LNP

In this Appendix we prove some further properties of the logic **LNP** (Section F.1), provide the soundness and completeness proof for this logic (Section F.2) and provide the proof of Theorem 32 (Section F.3). In some of the proofs contained below, we make extensive use of the axioms and rules of the logic **CLoNs** as defined in Appendix C.

F.1 Some facts about LNP and CLoNs

The following theorems will come in handy for the proofs of Theorems 30 and 31. Let in the remainder $\mathbf{L} \in \{\mathbf{CLoNs}, \mathbf{LNP}\}$:

Theorem 35. $\mathbf{L}$ *is reflexive, transitive and monotonic.*

Theorem 36. $\mathbf{L}$ *is compact (if* $\Gamma \vdash_{\mathbf{L}} A$ *then* $\Gamma' \vdash_{\mathbf{L}} A$ *for some finite* $\Gamma' \subseteq \Gamma$*).*

Theorem 37. *If* $\Gamma \vdash_{\mathbf{L}} B$ *and* $A \in \Gamma$*, then* $\Gamma - \{A\} \vdash_{\mathbf{L}} A \supset B$ *(Generalized Deduction Theorem for* $\mathbf{L}$*).*

The proofs of Theorems 35 – 37 are straightforward and safely left to the reader.

Fact 7. (i) $\mathsf{O}A, \mathsf{O}B \vdash_{\mathbf{LNP}} \mathsf{O}(A \wedge B)$
(ii) $\mathsf{O}A, \mathsf{P}B \vdash_{\mathbf{LNP}} \mathsf{P}(A \wedge B)$
(iii) $\vdash_{\mathbf{LNP}} (\mathsf{O}A \wedge \mathsf{O}B) \supset \mathsf{O}(A \wedge B)$
(iv) $\vdash_{\mathbf{LNP}} (\mathsf{O}A \wedge \mathsf{P}B) \supset \mathsf{P}(A \wedge B)$
(v) $\vdash_{\mathbf{LNP}} \mathsf{P}(A \supset A)$
(vi) If $\vdash_{\mathbf{CLoNs}} A' \supset A$ then $A \supset B \vdash_{\mathbf{CLoNs}} A' \supset B$.
(vii) If $\vdash_{\mathbf{CLoNs}} B \supset B'$ then $A \supset B \vdash_{\mathbf{CLoNs}} A \supset B'$.
(viii) $\vdash_{\mathbf{CLoNs}} (A \vee (A \supset B)) \equiv (((A \vee (A \supset B)) \supset B) \supset B)$
(ix) $\vdash_{\mathbf{CLoNs}} (A \supset (A \supset B)) \supset (A \supset B)$
(x) $\vdash_{\mathbf{CLoNs}} A \vee (A \supset B)$.

Proof. Ad(i). Suppose $\mathsf{O}A$ and $\mathsf{O}B$. By (A∧3), $\vdash_{\mathbf{CLoNs}} A \supset (B \supset (A \wedge B))$. By (NEC˜), it follows that $\vdash_{\mathbf{LNP}} \mathsf{O}(A \supset (B \supset (A \wedge B)))$. By (K), $\vdash_{\mathbf{LNP}} \mathsf{O}A \supset$

$\mathsf{O}(B \supset (A \wedge B))$. By (MP), $\mathsf{O}(B \supset (A \wedge B))$. By (K), $\mathsf{O}B \supset \mathsf{O}(A \wedge B)$. By (MP), $\mathsf{O}(A \wedge B)$.

Ad(ii). Suppose $\mathsf{O}A$ and $\mathsf{P}B$. By (A∧3), $\vdash_{\mathbf{CLoNs}} A \supset (B \supset (A \wedge B))$. By (NEC˜), $\vdash_{\mathbf{LNP}} \mathsf{O}(A \supset (B \supset (A \wedge B)))$. By (K), $\vdash_{\mathbf{LNP}} \mathsf{O}A \supset \mathsf{O}(B \supset (A \wedge B))$. By (MP), $\mathsf{O}(B \supset (A \wedge B))$. By (KP), $\mathsf{P}B \supset \mathsf{P}(A \wedge B)$. By (MP), $\mathsf{P}(A \wedge B)$.

Ad(iii)-(iv). Immediate in view of (i),(ii), and Theorem 37.

Ad(v). Since $A \supset A$ is a theorem of the positive fragment of **CL**, it is also a **CLoNs**-theorem. By (NEC˜), $\vdash_{\mathbf{LNP}} \mathsf{O}(A \supset A)$. By (D), $\vdash_{\mathbf{LNP}} \mathsf{P}(A \supset A)$.

Ad (vi): Suppose $\vdash_{\mathbf{CLoNs}} A' \supset A$. By (A⊃2), $\vdash_{\mathbf{CLoNs}} (A' \supset (A \supset B)) \supset ((A' \supset A) \supset (A' \supset B))$. By (A⊃1) and (MP), $A \supset B \vdash_{\mathbf{CLoNs}} A' \supset (A \supset B)$. The rest follows by multiple applications of (MP).

Ad (vii): The proof is similar and left to the reader.

Ad (viii): *Left-to-right*: By (MP), $(A \vee (A \supset B)) \supset B, A \vee (A \supset B) \vdash_{\mathbf{CLoNs}} B$. The rest follows by Theorem 37. *Right-to-left*: By (A⊃1), (†) $\vdash_{\mathbf{CLoNs}} B \supset (A \supset B)$. By (A∨2), (‡) $\vdash_{\mathbf{CLoNs}} (A \supset B) \supset (A \vee (A \supset B))$. Altogether, by (†), (‡), (vii) and (MP), $\vdash_{\mathbf{CLoNs}} B \supset (A \vee (A \supset B))$. Hence, by (vii), $\vdash_{\mathbf{CLoNs}} ((A \vee (A \supset B) \supset B) \supset B) \supset ((A \vee (A \supset B) \supset B) \supset (A \vee (A \supset B)))$. By (A⊃3), $\vdash_{\mathbf{CLoNs}} ((A \vee (A \supset B) \supset B) \supset (A \vee (A \supset B))) \supset (A \vee (A \supset B))$. Hence, again by (vii), $((A \vee (A \supset B) \supset B) \supset B) \supset (A \vee (A \supset B))$.

Ad (ix): By (MP), $A, A \supset (A \supset B) \vdash_{\mathbf{CLoNs}} A \supset B$. By (MP), $A, A \supset (A \supset B) \vdash_{\mathbf{CLoNs}} B$. By Theorem 37, $A \supset (A \supset B) \vdash_{\mathbf{CLoNs}} A \supset B$, $\vdash_{\mathbf{CLoNs}} (A \supset (A \supset B)) \supset (A \supset B)$.

Ad (x): By (A∨1), $\vdash_{\mathbf{CLoNs}} A \supset (A \vee (A \supset B))$. By (vi), $\vdash_{\mathbf{CLoNs}} (A \vee (A \supset B)) \supset B \vdash_{\mathbf{CLoNs}} A \supset B$. By Theorem 37, $\vdash_{\mathbf{CLoNs}} ((A \vee (A \supset B)) \supset B) \supset (A \supset B)$. By (A∨2), $\vdash_{\mathbf{CLoNs}} (A \supset B) \supset (A \vee (A \supset B))$. Hence, by (vii), $\vdash_{\mathbf{CLoNs}} ((A \vee (A \supset B)) \supset B) \supset (A \vee (A \supset B))$. By (viii), $\vdash_{\mathbf{CLoNs}} (A \vee (A \supset B)) \equiv (((A \vee (A \supset B)) \supset B) \supset B)$. Thus, by (vii), $\vdash_{\mathbf{CLoNs}} ((A \vee (A \supset B)) \supset B) \supset (((A \vee (A \supset B)) \supset B) \supset B)$. By (ix) and (MP), $\vdash_{\mathbf{CLoNs}} ((A \vee (A \supset B)) \supset B) \supset B$. By (viii), (A≡2), and (MP), $\vdash_{\mathbf{CLoNs}} A \vee (A \supset B)$.

□

F.2 Proofs of Theorems 30 and 31

In order to simplify the notation in the following meta-proofs we define $R(w) = \{w' \mid Rww'\}$.

Proof of Theorem 30. Let in the following $M = \langle W, w_0, R, v_0, v \rangle$ be an **LNP**-model.

It is easy to check that all **CL**-axiom schemas hold at w_0 in M due to (C_0), (C¬), and (C⊃)-(C≡). Similarly, (†) where $w \in W \smallsetminus \{w_0\}$, all **CLoNs**-axiom schemas hold at w in M due to (C_l) and (C∼∼)-(C≡).

Ad (NEC˜). Let $\vDash_{\mathbf{CLoNs}} A$. By (C$\mathsf{O}$), (†) and the definition of R, $v_M(\mathsf{O}A, w_0) = 1$.

Ad (K). Suppose $M \Vdash \mathsf{O}(A \supset B)$. By (C$\mathsf{O}$) and (C⊃), for all $w \in R(w_0)$, $v_M(A, w) = 0$ or $v_M(B, w) = 1$. Suppose $M \Vdash \mathsf{O}A$, then for all $w \in R(w_0)$, $v_M(A, w) = 1$. Hence, for all $w \in R(w_0)$, $v_M(B, w) = 1$. Thus by (CO), $M \Vdash \mathsf{O}B$. Hence, by (C⊃), $M \Vdash \mathsf{O}A \supset \mathsf{O}B$. Altogether, by (C⊃), $M \Vdash \mathsf{O}(A \supset B) \supset (\mathsf{O}A \supset \mathsf{O}B)$.

Ad (D). Suppose $M \Vdash \mathsf{O}A$. Hence for all $w \in R(w_0)$, $v_M(A, w) = 1$ (by (CO)). By the non-emptiness of R, there is a $w \in R(w_0)$ for which $v_M(A, w) = 1$. By (CP), $M \Vdash \mathsf{P}A$. By (C$\supset$), $M \Vdash \mathsf{O}A \supset \mathsf{P}A$.

Ad (KP). Suppose $M \Vdash \mathsf{O}(A \supset B)$. By (C$\mathsf{O}$) and (C$\supset$), (‡) for all $w \in R(w_0)$, $v_M(A, w) = 0$ or $v_M(B, w) = 1$. Suppose $M \Vdash \mathsf{P}A$. Then, by (CP) there is a $w \in R(w_0)$ for which $v_M(A, w) = 1$. Hence, by (‡), there is a $w \in R(w_0)$ such that $v_M(B, w) = 1$. Thus, by (CP) $M \Vdash \mathsf{P}B$ and, by (C$\supset$), $M \Vdash \mathsf{P}A \supset \mathsf{P}B$. Altogether, by (C$\supset$), $M \Vdash \mathsf{O}(A \supset B) \supset (\mathsf{P}A \supset \mathsf{P}B)$.

Ad (OD). Suppose $M \Vdash \mathsf{O}(A \vee B)$. By (C$\mathsf{O}$) and (C$\vee$), ($\star$) for all $w \in R(w_0)$, $v_M(A, w) = 1$ or $v_M(B, w) = 1$. Suppose $M \nVdash \mathsf{P}B$. By (CP): for all $w \in R(w_0)$, $v_M(B, w) = 0$. By ($\star$), for all $w \in R(w_0)$, $v_M(A, w) = 1$. Thus, by (CO), $M \Vdash \mathsf{O}A$. Hence, by (C$\vee$), $M \Vdash \mathsf{O}A \vee \mathsf{P}B$. Altogether, by (C$\supset$), $M \Vdash \mathsf{O}(A \vee B) \supset (\mathsf{O}A \vee \mathsf{P}B)$.

Ad (PD). This is similar to the previous case and is left to the reader.

We now know that all axiom schemas and rules of **LNP** are semantically valid. That $\Gamma \vdash_{\mathbf{LNP}} A$ implies $\Gamma \vDash_{\mathbf{LNP}} A$ can now be shown via the usual induction on the length of the proof of A. This is safely left to the reader. □

Let in the remainder W_c be the **LNP**-deductively closed and maximally **LNP**-non-trivial subsets of $\mathcal{W}^{\mathbf{LNP}}$. Moreover, let $W_c^{\sim}$ be the **CLoNs**-deductively closed subsets Γ of $\mathcal{W}^{\sim}$ where Γ is prime, i.e. for each $A \vee B \in \Gamma$ either $A \in \Gamma$ or $B \in \Gamma$.

For the completeness proof of **LNP**, we make use of the following lemmas.[1]

Lemma 13. *If $\Delta \in W_c$, then Δ is prime.*

Proof. Suppose that, for a $\Delta \in W_c$, $A \vee B \in \Delta$ and $A \notin \Delta$ and $B \notin \Delta$. Then, since Δ is maximally **LNP**-non-trivial, $\Delta \cup \{A\}$ is trivial and $\Delta \cup \{B\}$ is trivial. Then, for any $C \in \mathcal{W}^{\mathbf{LNP}}$, $\Delta \cup \{A\} \vdash_{\mathbf{LNP}} C$ and $\Delta \cup \{B\} \vdash_{\mathbf{LNP}} C$. Then, by Theorem 37, $\Delta \vdash_{\mathbf{LNP}} A \supset C$ and $\Delta \vdash_{\mathbf{LNP}} B \supset C$. But then, by (MP) and (A$\vee$3), $\Delta \vdash_{\mathbf{LNP}} (A \vee B) \supset C$. Since $A \vee B \in \Delta$, by (MP) $\Delta \vdash_{\mathbf{LNP}} C$, and since Δ is **LNP**-deductively closed, $C \in \Delta$. This contradicts the supposition. Hence if $A \vee B \in \Delta$, then $A \in \Delta$ or $B \in \Delta$. □

Where $\Gamma \in W_c$ and $A \in \mathcal{W}^{\sim}$, we will use the following abbreviations: $\Gamma_O = \{B \mid \mathsf{O}B \in \Gamma\}$, $\Gamma_O^A = \Gamma_O \cup \{A\}$, $\Gamma_P = \{B \mid \mathsf{P}B \notin \Gamma\}$, ${}^{\vee}\Gamma_P = \{\bigvee \Theta \mid \Theta \subseteq \Gamma_P, \Theta \text{ is finite}\}$ and ${}^{\vee}\Gamma_P^B = \{\bigvee \Theta \mid \Theta \subseteq \Gamma_P \cup \{B\}, \Theta \text{ is finite}\}$.

Lemma 14. *Let $\Gamma \in W_c$. (i) If $C \in \mathrm{Cn}_{\mathbf{CLoNs}}(\Gamma_O)$ then $\mathsf{O}C \in \Gamma$. (ii) Where $\mathsf{P}A \in \Gamma$, if $C \in \mathrm{Cn}_{\mathbf{CLoNs}}(\Gamma_O^A)$ then $\mathsf{P}C \in \Gamma$.*

Proof. Ad (i): Suppose that $\Gamma_O \vdash_{\mathbf{CLoNs}} C$. Then $\Gamma' \vdash_{\mathbf{CLoNs}} C$ for some finite $\Gamma' \subseteq \Gamma_O$ (given the compactness of **CLoNs**). Hence, $\vdash_{\mathbf{CLoNs}} \left(\bigwedge \Gamma'\right) \supset C$ by Theorem 37. Thus, $\vdash_{\mathbf{LNP}} \mathsf{O}\left(\left(\bigwedge \Gamma'\right) \supset C\right)$ by (NEC$^{\sim}$). By (K), $\vdash_{\mathbf{CLoNs}} \mathsf{O} \bigwedge \Gamma' \supset \mathsf{O}C$. By the deductive closure of Γ, the fact that $\Gamma' \subseteq \Gamma$ and Fact 7 (i), $\mathsf{O} \bigwedge \Gamma' \in \Gamma$. By (MP), $\mathsf{O}C \in \Gamma$.

Ad (ii): Suppose that $\Gamma_O^A \vdash_{\mathbf{CLoNs}} C$. Then $\Gamma' \vdash_{\mathbf{CLoNs}} C$ for some finite $\Gamma' \subseteq \Gamma_O^A$ (given the compactness of **CLoNs**). Then $\Gamma' \cup \{A\} \vdash_{\mathbf{CLoNs}} C$ by the

[1] The proof of Lemma 16 is inspired by the proof of Lemma 1.7.1 from [19].

monotonicity of **CLoNs**. Then $\vdash_{\mathbf{CLoNs}} \big(\bigwedge \Gamma' \wedge A\big) \supset C$ by Theorem 37. Then $\vdash_{\mathbf{LNP}} \mathsf{O}\big(\big(\bigwedge \Gamma' \wedge A\big) \supset C\big)$ by (NEC$^{\sim}$). Then $\vdash_{\mathbf{LNP}} \mathsf{P}\big(\bigwedge \Gamma' \wedge A\big) \supset \mathsf{P}C$ by (KP) and (MP). By the supposition, $\{\mathsf{O}B \mid B \in \Gamma'\} \subseteq \Gamma$ and $\mathsf{P}A \in \Gamma$. Given the deductive closure of Γ and $\vdash_{\mathbf{LNP}} (\mathsf{O}(\bigwedge \Gamma') \wedge \mathsf{P}A) \supset \mathsf{P}(\bigwedge \Gamma' \wedge A)$ (which follows from Fact 7 (ii)), it follows that $\mathsf{P}\big(\bigwedge \Gamma' \wedge A\big) \in \Gamma$. Hence $\mathsf{P}C \in \Gamma$, since Γ is deductively closed and $\vdash_{\mathbf{LNP}} \mathsf{P}\big(\bigwedge \Gamma' \wedge A\big) \supset \mathsf{P}C$. □

Lemma 15. *Let* $\Gamma \in W_c$. *(i) Where* $\mathsf{P}A \in \Gamma$, ${}^{\vee}\Gamma_P \cap \mathrm{Cn}_{\mathbf{CLoNs}}(\Gamma_O^A) = \varnothing$. *(ii) Where* $B \notin \Gamma_O$, ${}^{\vee}\Gamma_P^B \cap \mathrm{Cn}_{\mathbf{CLoNs}}(\Gamma_O) = \varnothing$.

Proof. Ad (i): Let $C = \bigvee \Theta$ where $\Theta = \{C_1, \ldots, C_n\} \subseteq \Gamma_P$. Suppose $C \in \mathrm{Cn}_{\mathbf{CLoNs}}(\Gamma_O^A)$ then by Lemma 14 (ii), $\mathsf{P}\bigvee \Theta \in \Gamma$. Hence, by (PD), $\bigvee_{i=1}^n \mathsf{P}C_i \in \Gamma$. Since Γ is prime, there is an $i \in \{1, \ldots, n\}$ for which $\mathsf{P}C_i \in \Gamma$ and hence $C_i \notin \Gamma_P$,—a contradiction.

Ad (ii): Let $C = \bigvee \Theta$ where $\Theta = \{C_1, \ldots, C_n\} \subseteq (\Gamma_P \cup \{B\})$. Suppose $C \in \mathrm{Cn}_{\mathbf{CLoNs}}(\Gamma_O)$. By Lemma 14 (i), $\mathsf{O}\bigvee \Theta \in \Gamma$. Assume that, where $i \in \{1, \ldots, n\}$, all $C_i \in \Gamma_P$. By (D), $\mathsf{P}\bigvee \Theta \in \Gamma$. By (PD), $\bigvee_{i=1}^n \mathsf{P}C_i \in \Gamma$. Hence, since Γ is prime, there is an $i \in \{1, \ldots, n\}$ such that $\mathsf{P}C_i \in \Gamma$ and hence $C_i \notin \Gamma_P$,—a contradiction. Hence there is a non-empty $J \subseteq \{1, \ldots, n\}$ such that for each $j \in J$, $C_j = B$. Hence, by (OD), $\mathsf{O}B \vee \mathsf{P}\bigvee_{\{1,\ldots,n\} \setminus J} C_i \in \Gamma$. Thus, by (PD), $\mathsf{O}B \vee \bigvee_{\{1,\ldots,n\} \setminus J} \mathsf{P}C_i \in \Gamma$. Since $B \notin \Gamma_O$ and since Γ is prime, there is an $i \in \{1, \ldots, n\} \setminus J$ such that $\mathsf{P}C_i \in \Gamma$ and hence $C_i \notin \Gamma_P$,—a contradiction. □

Lemma 16. *Let* $\Gamma \in W_c$.

1. *Where* $\mathsf{P}A \in \Gamma$, *there is a* $\Delta \subseteq \mathcal{W}^{\sim}$ *for which (i)* $\Gamma_O^A \subseteq \Delta$, *(ii)* ${}^{\vee}\Gamma_P \cap \Delta = \varnothing$, *and (iii)* $\Delta \in W_c^{\sim}$.
2. *Where* $B \notin \Gamma_O$, *there is a* $\Delta \subseteq \mathcal{W}^{\sim}$ *for which (i)* $\Gamma_O \subseteq \Delta$, *(ii)* ${}^{\vee}\Gamma_P^B \cap \Delta = \varnothing$, *and (iii)* $\Delta \in W_c^{\sim}$.

Proof. Let $\langle \mathbf{\Gamma_O}, \mathbf{\Gamma_P} \rangle \in \{\langle \Gamma_O^A, {}^{\vee}\Gamma_P \rangle, \langle \Gamma_O, {}^{\vee}\Gamma_P^B \rangle\}$. Where $\langle B_1, B_2, \ldots \rangle$ is a list of the members of $\mathcal{W}^{\sim}$, define $\Delta_0 = \mathrm{Cn}_{\mathbf{CLoNs}}(\mathbf{\Gamma_O})$ and $\Delta = \Delta_0 \cup \Delta_1 \cup \ldots$ where

$$\Delta_{i+1} = \begin{cases} \mathrm{Cn}_{\mathbf{CLoNs}}(\Delta_i \cup \{B_{i+1}\}) & \text{if } \mathbf{\Gamma_P} \cap \mathrm{Cn}_{\mathbf{CLoNs}}(\Delta_i \cup \{B_{i+1}\}) = \varnothing \\ \Delta_i & \text{otherwise} \end{cases}$$

Ad (i): this holds by the construction and the reflexivity of **CLoNs**.

Ad (ii): By Lemma 15 $\Delta_0 \cap \mathbf{\Gamma_P} = \varnothing$. The rest follows by the construction.

Ad (iii): We first show that Δ is **CLoNs**-deductively closed. Suppose there is a $B_i \notin \Delta$ such that $\Delta \vdash_{\mathbf{CLoNs}} B_i$. Then, by the construction of Δ, there is a $D \in \mathbf{\Gamma_P}$ such that $\Delta \cup \{B_i\} \vdash_{\mathbf{CLoNs}} D$ and hence by Theorem 37, $\Delta \vdash_{\mathbf{CLoNs}} B_i \supset D$. However, by (MP) also $\Delta \vdash_{\mathbf{CLoNs}} D$. By the compactness of **CLoNs** there is a Δ_i for which $\Delta_i \vdash_{\mathbf{CLoNs}} D$. By the construction $\Delta_i = \mathrm{Cn}_{\mathbf{CLoNs}}(\Delta_i)$ and whence $D \in \Delta_i$. Hence, $D \in \Delta$,—a contradiction with (ii).

We now show that Δ is prime. Suppose $A_1 \vee A_2 \in \Delta$. Assume $A_1, A_2 \notin \Delta$. Hence, by the construction of Δ, $\Delta \cup \{A_1\} \vdash_{\mathbf{CLoNs}} D_1$ and $\Delta \cup \{A_2\} \vdash_{\mathbf{CLoNs}} D_2$ for some $D_1, D_2 \in \mathbf{\Gamma_P}$. By Theorem 37, $\Delta \vdash_{\mathbf{CLoNs}} A_1 \supset D_1$ and $\Delta \vdash_{\mathbf{CLoNs}} A_2 \supset D_2$. By some simple propositional manipulations, $\Delta \vdash_{\mathbf{CLoNs}} (A_1 \vee A_2) \supset$

$(D_1 \vee D_2)$. By (MP), $\Delta \vdash_{\mathbf{CLoNs}} D_1 \vee D_2$ and hence $D_1 \vee D_2 \in \Delta$. However, by the definition of $\mathbf{\Gamma_P}$, $D_1 \vee D_2 \in \mathbf{\Gamma_P}$,—a contradiction with (ii). □

Definition 19. The binary relation $\mathsf{R} \subseteq \left(W_c \times W_c^{\sim}\right)$ is defined as follows: $\mathsf{R}\Gamma\Delta$ iff the following two conditions are met

(a) if $\mathsf{O}A \in \Gamma$ then $A \in \Delta$, and

(b) if $A \in \Delta$ then $\mathsf{P}A \in \Gamma$.

In view of the definition of R, the following holds:

Lemma 17. *Where $\Gamma \in W_c$, $\mathsf{P}A \in \Gamma$ iff there is a $\Delta \in W_c^{\sim}$ such that $\mathsf{R}\Gamma\Delta$ and $A \in \Delta$.*

Proof. Left-right: Suppose $\mathsf{P}A \in \Gamma$. Then, by Lemma 16.1, there is a $\Delta \subseteq \mathcal{W}^{\sim}$ such that (i) $\Gamma_O^A \subseteq \Delta$, (ii) for all $C \in \Gamma_P$, $C \notin \Delta$, (iii) $\Delta \in W_c^{\sim}$. We now show that $\mathsf{R}\Gamma\Delta$. Ad (a): if, for some D, $\mathsf{O}D \in \Gamma$ then $D \in \Gamma_O^A$, hence $D \in \Delta$ by (i). Ad (b): suppose $\mathsf{P}E \notin \Gamma$ for some $E \in \mathcal{W}^{\sim}$. Then $E \in \Gamma_P$, hence $E \notin \Delta$ by (ii).

Right-left: Follows directly by Definition 19. □

Lemma 18. *For every $\Gamma \in W_c$, there is a $\Delta \in W_c^{\sim}$ such that $\mathsf{R}\Gamma\Delta$ (i.e. R is non-empty).*

Proof. By Fact 7 (v), $\vdash_{\mathbf{LNP}} \mathsf{P}(A \supset A)$. Hence, $\mathsf{P}(A \supset A) \in \Gamma$ for every $\Gamma \in W_c$. But then, by Lemma 17, there is a $\Delta \in W_c^{\sim}$ such that $\mathsf{R}\Gamma\Delta$ and $A \supset A \in \Delta$. Hence R is non-empty as required. □

Lemma 19. *Where $\Gamma \in W_c$, $\mathsf{O}A \in \Gamma$ iff, for all $\Delta \in W_c^{\sim}$ such that $\mathsf{R}\Gamma\Delta$, $A \in \Delta$.*

Proof. Left-right: This is an immediate consequence of Definition 19.

Right-left: Suppose $\mathsf{O}A \notin \Gamma$. Hence, $A \notin \Gamma_O$. By Lemma 16.2, there is a $\Delta \subseteq \mathcal{W}^{\sim}$ for which (i) $\Gamma_O \subseteq \Delta$, (ii) $(\Gamma_P \cup \{A\}) \cap \Delta = \varnothing$, and (iii) $\Delta \in W_c^{\sim}$. We now show that $\mathsf{R}\Gamma\Delta$. Ad (a): if, for some D, $\mathsf{O}D \in \Gamma$ then $D \in \Gamma_O$, hence $D \in \Delta$ by (i). Ad (b): suppose $\mathsf{P}E \notin \Gamma$ for some $E \in \mathcal{W}^{\sim}$. Then $E \in \Gamma_P$, hence $E \notin \Delta$ by (ii). □

Lemma 20. *Where $\Delta \in W_c$, there is an **LNP**-model M such that $M \Vdash A$ for all $A \in \Delta$ and $M \nVdash A$ for all $A \in \mathcal{W}^{\mathbf{LNP}} \setminus \Delta$.*

Proof. Let $\Delta \in W_c$. We construct an **LNP**-model $M = \langle \{\Delta\} \cup W_c^{\sim}, w_0, \mathsf{R}, v_0, v\rangle$ such that:

(i) $w_0 = \Delta$

(ii) For all $A \in \mathcal{W}^a$, $v_0(A, w_0) = 1$ iff $A \in w_0$

(iii) For all $A \in \mathcal{W}_l^{\sim}$ and all $w \in W_c^{\sim}, v(A, w) = 1$ iff $A \in w$

By Lemma 18, R is non-empty. We now show that:

(*) (a) for all $A \in \mathcal{W}^{\mathbf{LNP}}$, $v_M(A, w_0) = 1$ iff $A \in w_0$,
(b) for all $A \in \mathcal{W}^{\sim}$ and all $w \in W_c^{\sim}$, $v_M(A, w) = 1$ iff $A \in w$.

The proof proceeds as usual by an induction on the complexity of A. Let $w \in \{w_0\} \cup W_c^{\sim}$, and $A \in \mathcal{W}^a$. If $w = w_0$, then, by (ii), $v_0(A, w_0) = 1$ iff $A \in w_0$. By

(C_0), it follows that $v_M(A,w) = 1$ iff $A \in w$. If $w \neq w_0$, then, by (iii), $v(A,w) = 1$ iff $A \in w$. By (C_l), it follows that $v_M(A,w) = 1$ iff $A \in w$. Hence, for all $w \in \{w_0\} \cup W_c^{\sim}$, $v_M(A,w) = 1$ iff $A \in w$ and (*) is valid for all $A \in \mathcal{W}^a$.

Depending on the logical form of A, we distinguish 8 cases (6 for the connectives $\sim, \neg, \vee, \wedge, \supset, \equiv$, and 2 for the modal operators O and P) and show for each of them that $v_M(A,w) = 1$ iff $A \in w$.

Case 1. Let $w \in W_c^{\sim}$. We show that $v_M({\sim}A,w) = 1$ iff ${\sim}A \in w$. Either ${\sim}A \in \mathcal{W}_l^{\sim}$, or A has one of the forms ${\sim}B, B \vee C, B \wedge C, B \supset C$, or $B \equiv C$ (note that, since $w \neq w_0$, A cannot have the form $\mathsf{O}B$ or $\mathsf{P}B$).

If ${\sim}A \in \mathcal{W}_l^{\sim}$, then, by ($C_l$), $v_M({\sim}A,w) = 1$ iff $v({\sim}A,w) = 1$. By (iii), it follows that $v_M({\sim}A,w) = 1$ iff ${\sim}A \in w$.

If A has the form ${\sim}B$, then, by (C$\sim\sim$), $v_M({\sim\sim}B,w) = 1$ iff $v_M(B,w) = 1$. By the induction hypothesis, $v_M({\sim\sim}B,w) = 1$ iff $B \in w$. By (A$\sim\sim$), $v_M({\sim}A,w) = 1$ iff ${\sim}A \in w$.

If A has the form $B \vee C$, then, by (C$\sim\vee$), $v_M({\sim}(B \vee C),w) = 1$ iff $v_M({\sim}B \wedge {\sim}C,w) = 1$ iff [by (C$\wedge$)] $v_M({\sim}B,w) = 1$ and $v_M({\sim}C,w) = 1$ iff [by the induction hypothesis] ${\sim}B \in w$ and ${\sim}C \in w$ iff [by (A$\wedge$1), (A$\wedge$2), and (A$\wedge$3)] ${\sim}B \wedge {\sim}C \in w$ iff [by (A$\sim\vee$)] ${\sim}A \in w$.

The cases where A is of one of the forms $B \wedge C, B \supset C$, or $B \equiv C$ are similar and left to the reader.

Case 2. Let $w = w_0$. Suppose $v_M(\neg A,w) = 1$. By (C$\neg$), $v_M(A,w) = 0$. By the induction hypothesis, $A \notin w$. Then, since w is maximally **LNP**-non-trivial, $w \cup \{A\}$ is **LNP**-trivial and $w \cup \{A\} \vdash_{\mathbf{LNP}} \neg A$. By Theorem 37, it follows that $w \vdash_{\mathbf{LNP}} A \supset \neg A$. Then, since w is **LNP**-deductively closed, $A \supset \neg A \in w$ and, by (A$\neg$1) and (MP), $\neg A \in w$.

Suppose $\neg A \in w$. We show via reductio that $A \notin w$. Suppose thus that $A \in w$. Then, by (A$\neg$2), (MP), and since w is **LNP**-deductively closed, $B \in w$ for any $B \in \mathcal{W}^{\mathbf{LNP}}$. This contradicts the non-triviality of w, hence $A \notin w$. But then, by the induction hypothesis $v_M(A,w) = 0$ and, by (C$\neg$), $v_M(\neg A,w) = 1$.

Case 3. Let $w \in \{w_0\} \cup W_c^{\sim}$. Suppose $v_M(A \vee B,w) = 1$. Then, by (C$\vee$), $v_M(A,w) = 1$ or $v_M(B,w) = 1$. By the induction hypothesis, $A \in w$ or $B \in w$. Hence, by (A$\vee$1), (A$\vee$2), (MP), and the fact that w is **LNP**-(in case $w = w_0$)/**CLoNs**-(in case $w \in W_c^{\sim}$)-deductively closed, $A \vee B \in w$.

Suppose $A \vee B \in w$. If $w \neq w_0$, then, by the definition of $W_c^{\sim}$, $A \in w$ or $B \in w$. If $w = w_0$, then, by Lemma 13, $A \in w$ or $B \in w$. By the induction hypothesis, $v_M(A,w) = 1$ or $v_M(B,w) = 1$. Hence, by (C$\vee$), $v_M(A \vee B,w) = 1$.

Case 4. Let $w \in \{w_0\} \cup W_c^{\sim}$. Suppose $v_M(A \supset B,w) = 1$. Then by (C$\supset$), $v_M(A,w) = 0$ or $v_M(B,w) = 1$. By the induction hypothesis, $A \notin w$ or $B \in w$. Let now $w \in W_c^{\sim}$. If $A \notin w$, then, since $\vdash_{\mathbf{CLoNs}} A \vee (A \supset B)$ by Fact 7 (x) and since w is prime, also $A \supset B \in w$. If $B \in w$, then since by (A$\supset$1) $\vdash_{\mathbf{CLoNs}} B \supset (A \supset B)$ and by (MP), also $A \supset B \in w$. The same argument applies to $w = w_0$ since also $\vdash_{\mathbf{LNP}} A \vee (A \supset B)$, and (A$\supset$1) and (MP) are also valid in **LNP**.

Suppose $A \supset B \in w$. By (MP), if $A \in w$ then $B \in w$. By the induction hypothesis, if $v_M(A,w) = 1$ then $v_M(B,w) = 1$. Hence, by (C$\supset$), $v_M(A \supset B,w) = 1$.

The proof for the other classical connectives (cases 4-6) is similar and left to the reader. We proceed with the cases for O and P.

Case 7. Let $w = w_0$. By Lemma 19, $\mathsf{O}A \in w_0$ iff $A \in w$ for all w such that $\mathsf{R}w_0w$. Hence, by the induction hypothesis, $\mathsf{O}A \in w_0$ iff $v_M(A, w) = 1$ for all w such that $\mathsf{R}w_0w$. But then, by (CO), $\mathsf{O}A \in w_0$ iff $v_M(\mathsf{O}A, w_0) = 1$.

Case 8. Let $w = w_0$. By Lemma 17, $\mathsf{P}A \in w_0$ iff $A \in w$ for some w such that $\mathsf{R}w_0w$. Hence, by the induction hypothesis, $\mathsf{P}A \in w_0$ iff $v_M(A, w) = 1$ for some w such that $\mathsf{R}w_0w$. But then, by (CP), $\mathsf{P}A \in w_0$ iff $v_M(\mathsf{P}A, w_0) = 1$.

The rest follows by (i) and (*). □

Lemma 21. *Let* $\Gamma \subseteq \mathcal{W}^{\mathbf{LNP}}$ *and* $\Gamma \nvdash_{\mathbf{LNP}} A$. *There is a* $\Delta \subseteq \mathcal{W}^{\mathbf{LNP}}$ *such that (i)* $\Gamma \subseteq \Delta$, *(ii)* $A \notin \Delta$, *and (iii)* $\Delta \in W_c$.

Proof. Where $\langle B_1, B_2, \ldots \rangle$ is a list of the members of $\mathcal{W}^{\mathbf{LNP}}$, define $\Delta_0 = \mathrm{Cn}_{\mathbf{LNP}}(\Gamma)$ and $\Delta = \Delta_0 \cup \Delta_1 \cup \ldots$ where

$$\Delta_{i+1} = \begin{cases} \mathrm{Cn}_{\mathbf{LNP}}(\Delta_i \cup \{B_{i+1}\}) & \text{if } A \notin \mathrm{Cn}_{\mathbf{LNP}}(\Delta_i \cup \{B_{i+1}\}) \\ \Delta_i & \text{else} \end{cases}$$

Ad (i): This holds by the construction of Δ and the reflexivity of **LNP**.

Ad (ii): This holds by the construction and since $A \notin \mathrm{Cn}_{\mathbf{LNP}}(\Gamma)$.

Ad (iii): Assume that $B \notin \Delta$ and $\Delta \vdash_{\mathbf{LNP}} B$. Hence, by the construction of Δ, $\Delta \cup \{B\} \vdash_{\mathbf{LNP}} A$ and whence by Theorem 37, $\Delta \vdash_{\mathbf{LNP}} B \supset A$. But then by (MP), $\Delta \vdash_{\mathbf{LNP}} A$. Thus, by the compactness of **LNP** and since each $\Delta_i = \mathrm{Cn}_{\mathbf{LNP}}(\Delta_i)$, there is a Δ_i such that $A \in \Delta_i$,—a contradiction to (ii).

Suppose $B \notin \Delta$. Assume that $\neg B \notin \Delta$. By the construction of Δ and the monotonicity of **LNP**, $\Delta \cup \{\neg B\} \vdash_{\mathbf{LNP}} A$ and whence by Theorem 37, $\Delta \vdash_{\mathbf{LNP}} \neg B \supset A$. Analogously, $\Delta \vdash_{\mathbf{LNP}} B \supset A$. By (A∨3), $\Delta \vdash_{\mathbf{LNP}} (B \vee \neg B) \supset A$. Since $\vdash_{\mathbf{CL}} B \vee \neg B$, also $\Delta \vdash_{\mathbf{LNP}} B \vee \neg B$. By (MP), $\Delta \vdash_{\mathbf{LNP}} A$,—a contradiction to (ii). Hence, $\neg B \in \Delta$. Thus, $\Delta \cup \{B\}$ is **CL**-trivial and hence also **LNP**-trivial. □

Proof of Theorem 31. Suppose $\Gamma \nvdash_{\mathbf{LNP}} A$. Then, by Lemma 21, there is a $\Delta \supseteq \Gamma$ such that $A \notin \Delta$ and $\Delta \in W_c$. Then, by Lemma 20, there is an **LNP**-model M such that $M \Vdash B$ for all $B \in \Gamma$ and $M \nVdash A$. Hence $\Gamma \nvDash_{\mathbf{LNP}} A$. □

F.3 Proof outline of Theorem 32

We first show that (ULNP1) and (ULNP2) can be generalized to their derived schemas (ULNP1') and (ULNP2'), which hold without the restriction that A is an atomic proposition:

Lemma 22. *The schemas (ULNP1') and (ULNP2') are* **ULNP***-valid:*

$$\mathsf{P}(A \wedge {\sim}A) \supset B \tag{ULNP1'}$$

$$\neg\mathsf{O}(A \vee {\sim}A) \supset B \tag{ULNP2'}$$

Proof outline: This is shown by an induction over the complexity of A. Where $A \in \mathcal{W}^a$, this holds due to (ULNP1) and (ULNP2). For the induction step, the proof proceeds analogously to the proof of Lemma 12 and is safely left to the reader. □

Lemma 23. *If* $\Gamma \vdash_{\mathbf{ULNP}} A$ *then* $\Gamma^{\neg} \vdash_{\mathbf{SDL}} \pi(A)$.

Proof. It easily checked that, under the transformation given in Section 6.2.6, all of (K), (D), (KP), (OD), (PD), (ULNP1), and (ULNP2) are **SDL**-valid. Moreover, since **CLoNs** is a proper fragment of **CL**, (NEC˜) too is valid in **SDL** (assuming again the transformation from Section 6.2.6). □

Lemma 24. *If* $\Gamma^{\neg} \vdash_{\mathbf{SDL}} \pi(A)$ *then* $\Gamma \vdash_{\mathbf{ULNP}} A$.

Proof. By the definition, **ULNP** verifies (K) and (D). It remains to show that **ULNP** verifies (i) all instances of $\mathsf{P}A \equiv \neg\mathsf{O}{\sim}A$ and (ii) the rule "If $\vdash_{\mathbf{CL}^{\sim}} A$ then $\vdash_{\mathbf{ULNP}} \mathsf{O}A$", where $\mathbf{CL}^{\sim}$ is classical propositional logic with the negation symbol ~ behaving classically.

Ad (i). *Left-Right.* By (A∧3), $\mathsf{P}A \supset (\mathsf{O}{\sim}A \supset (\mathsf{P}A \wedge \mathsf{O}{\sim}A))$. By Fact 7 (iv), $(\mathsf{P}A \wedge \mathsf{O}{\sim}A) \supset \mathsf{P}(A \wedge {\sim}A)$. Thus, by some propositional manipulations in **CL**, $\mathsf{P}A \supset (\mathsf{O}{\sim}A \supset \mathsf{P}(A \wedge {\sim}A))$, which is **CL**-equivalent to (†) $\mathsf{P}A \supset (\neg\mathsf{O}{\sim}A \vee \mathsf{P}(A \wedge {\sim}A))$. Suppose now that $\mathsf{P}A$. By (†), $\neg\mathsf{O}{\sim}A \vee \mathsf{P}(A \wedge {\sim}A)$. Moreover, by (ULNP1'), $\mathsf{P}(A \wedge {\sim}A) \supset \neg\mathsf{O}{\sim}A$. Thus, by (MP) and some simple **CL**-manipulations, we obtain $\neg\mathsf{O}{\sim}A$.
Right-left. By **CL**, $\mathsf{O}(A \vee {\sim}A) \vee \neg\mathsf{O}(A \vee {\sim}A)$. By (OD), $\mathsf{O}(A \vee {\sim}A) \supset (\mathsf{O}{\sim}A \vee \mathsf{P}A)$. Thus, by some propositional manipulations in **CL**, $(\mathsf{O}{\sim}A \vee \mathsf{P}A) \vee \neg\mathsf{O}(A \vee {\sim}A)$. The latter formula is **CL**-equivalent to $\neg\mathsf{O}{\sim}A \supset (\mathsf{P}A \vee \neg\mathsf{O}(A \vee {\sim}A))$. Suppose now that $\neg\mathsf{O}{\sim}A$. By (MP), $\mathsf{P}A \vee \neg\mathsf{O}(A \vee {\sim}A)$. By (ULNP2'), $\neg\mathsf{O}(A \vee {\sim}A) \supset \mathsf{P}A$. Thus, by (MP) and some simple **CL**-manipulations, $\mathsf{P}A$.

Ad (ii). Note that $A \in \mathcal{W}^{\sim}$ iff $\pi(A) \in \mathcal{W}$. Thus, where

(A~1): $(A \supset {\sim}A) \supset {\sim}A$,

(A~2): $A \supset ({\sim}A \supset B)$,

it follows by the definitions of **CLoNs** and **CL** that $\vdash_{\mathbf{CLoNs}\cup\{(A\sim1),(A\sim2)\}} A$ iff $\vdash_{\mathbf{CL}} \pi(A)$. We show that (i) if $\vdash_{\mathbf{CLoNs}} A$, then $\vdash_{\mathbf{ULNP}} A$, (ii) $\vdash_{\mathbf{ULNP}} \mathsf{O}((A \supset {\sim}A) \supset {\sim}A)$, and (iii) $\vdash_{\mathbf{ULNP}} \mathsf{O}(A \supset ({\sim}A \supset B))$.
(i) In case A is a **CLoNs**-theorem, $\mathsf{O}A$ follows immediately in view of (NEC˜).
(ii) $(A \vee {\sim}A) \supset ((A \supset {\sim}A) \supset {\sim}A)$ is an instance of the theorem $(A \vee B) \supset ((A \supset B) \supset B)$ of positive **CL**, thus it is a **CLoNs**-theorem. By (NEC˜), $\vdash_{\mathbf{ULNP}} \mathsf{O}((A \vee {\sim}A) \supset ((A \supset {\sim}A) \supset {\sim}A))$. By (K), $\vdash_{\mathbf{ULNP}} \mathsf{O}(A \vee {\sim}A) \supset \mathsf{O}((A \supset {\sim}A) \supset {\sim}A)$. By **CL**, (†) $\vdash_{\mathbf{ULNP}} \mathsf{O}((A \supset {\sim}A) \supset {\sim}A) \vee \neg\mathsf{O}(A \vee {\sim}A)$. We know by (ULNP2') that $\neg\mathsf{O}(A \vee {\sim}A) \supset \mathsf{O}((A \supset {\sim}A) \supset {\sim}A)$. Hence, by (†) and **CL**, $\vdash_{\mathbf{ULNP}} \mathsf{O}((A \supset {\sim}A) \supset {\sim}A)$.
(iii) $(A \supset ({\sim}A \supset B)) \vee (A \wedge {\sim}A)$ is an instance of the theorem $(A \supset (B \supset C)) \vee (A \wedge B)$ of positive **CL**, thus it is a **CLoNs**-theorem. By (NEC˜), $\vdash_{\mathbf{ULNP}} \mathsf{O}((A \supset ({\sim}A \supset B)) \vee (A \wedge {\sim}A))$. By (OD), (‡) $\vdash_{\mathbf{ULNP}} \mathsf{O}(A \supset ({\sim}A \supset B)) \vee \mathsf{P}(A \wedge {\sim}A)$. We know by (ULNP1') that $\mathsf{P}(A \wedge {\sim}A) \supset \mathsf{O}(A \supset ({\sim}A \supset B))$. Hence, by (‡) and **CL**, $\vdash_{\mathbf{ULNP}} \mathsf{O}(A \supset ({\sim}A \supset B))$. □

Theorem 32 follows immediately by Lemmas 23 and 24.

Appendix G

(Meta-)properties of the logic PMDL

In this Appendix we provide a syntactic characterization of the logic **DP** (Section G.1), prove soundness and completeness for the logic **PMDL** (Section G.2) and outline the proof of Theorem 34 (Section G.3).

G.1 The rules of LP

LP is axiomatized as follows:

$$
\begin{array}{rlrl}
A, B \vdash A \wedge B & \text{(AND)} & {\sim}A \vee {\sim}B \vdash {\sim}(A \wedge B) & \text{(DM2)} \\
A \wedge B \vdash A & \text{(AN1)} & {\sim}A \wedge {\sim}B \vdash {\sim}(A \vee B) & \text{(DM3)} \\
A \wedge B \vdash B & \text{(AN2)} & {\sim}(A \vee B) \vdash {\sim}A \wedge {\sim}B & \text{(DM4)} \\
A \vdash A \vee B & \text{(OR1)} & A \vdash {\sim}{\sim}A & \text{(DN1)} \\
B \vdash A \vee B & \text{(OR2)} & {\sim}{\sim}A \vdash A & \text{(DN2)} \\
{\sim}(A \wedge B) \vdash {\sim}A \vee {\sim}B & \text{(DM1)} & \vdash A \vee {\sim}A & \text{(EM)}
\end{array}
$$

and

$$\text{If } A, B \vdash D \text{ and } A, C \vdash D, \text{ then } A, B \vee C \vdash D. \qquad \text{(RBC)}$$

For a semantical characterization of **LP**, see e.g. [145, 146].

In footnote 1 in Section 4.2 we required – for technical reasons – that every LLL of an AL in standard format contains all classical connectives. Strictly speaking, **PMDL** does not feature the classical negation and implication connectives due to its definition ‘on top’ of **LP**. However, these connectives can easily be ‘superimposed’ on **LP**, as is done in [19, Ch. 7]. For more information on superimposing the classical connectives in order to obtain an AL in standard format, see e.g. [19, Sec. 4.3], [172, Sec. 2.8] or [181, Sec. 2.7].

G.2 Soundness and completeness of PMDL

Fact 8. (i) If $\Delta \cup \{A\} \vdash_{\mathbf{PMDL}} C$ and $\Delta \cup \{B\} \vdash_{\mathbf{PMDL}} C$ then $\Delta \cup \{A \vee B\} \vdash_{\mathbf{PMDL}} C$.

(ii) The following axiom is **PMDL**-derivable for each $J \subseteq_{\varnothing} I$:

$$\Box_J A \vdash \Diamond_J A \tag{D$\Box_J$}$$

It is easy to see that (i) follows by means of (RBC) and (ii) by means of (T$\Box_J$) and (T$\Diamond_J$).

As **ML** and **MDL** are fairly standard normal modal logics, we do not prove soundness and completeness theorems for these logics. Instead, we prove soundness and completeness for the more complex system **PMDL**.

Where $R \subseteq W \times W$ we use in the remainder the notation $Rw = \{w' \in W \mid Rww'\}$.

Lemma 25. *Where $M = \langle W, R_{\mathsf{O}}, \langle R_J \rangle_{J \subseteq_{\varnothing} I}, v, w_0 \rangle$ is a* **PMDL***-model, we have: for all $w \in W$, if $M, w \nvDash A$ then $M, w \vDash {\sim} A$.*

Proof. We show this by an induction over the length of A. Let $A \in \mathcal{W}^a$. Suppose $M, w \nvDash A$. By (C$\sim$), $M, w \vDash {\sim} A$.

For the induction step let first $A = B \wedge C$. Suppose $M, w \nvDash B \wedge C$. By (C$\wedge$), $M, w \nvDash B$ or $M, w \nvDash C$. By the induction hypothesis, $M, w \vDash {\sim} B$ or $M, w \vDash {\sim} C$. By (C$\vee$), $M, w \vDash {\sim} B \vee {\sim} C$. By (C$\sim\wedge$), $M, w \vDash {\sim}(B \wedge C)$. The cases $A = B \vee C$, and $A = {\sim} B$ are similar and left to the reader.

Let $A = \mathsf{O}B$. Suppose $M, w \nvDash {\sim}\mathsf{O}B$. By (C$\sim\mathsf{O}$) $M, w \nvDash \mathsf{P}{\sim}B$. By (C$\mathsf{P}$) there is no $w' \in R_{\mathsf{O}}w$ for which $M, w' \vDash {\sim}B$. By the induction hypothesis, for all $w' \in R_{\mathsf{O}}w$, $M, w' \vDash {\sim\sim}B$ and hence by (C$\sim\sim$), $M, w' \vDash B$. Thus, by (CO), $M, w \vDash \mathsf{O}B$.

The case $A = \Box_J B$ is analogous and left to the reader.

Let $A = \mathsf{P}B$. Suppose $M, w \nvDash \mathsf{P}B$. Hence, by (CP) there is no $w' \in R_{\mathsf{O}}w$ for which $M, w' \vDash B$. By the induction hypothesis, for all $w' \in R_{\mathsf{O}}w$, $M, w' \vDash {\sim}B$. Hence, by (CO), $M, w \vDash \mathsf{O}{\sim}B$. By (C$\sim\mathsf{P}$), $M, w \vDash {\sim}\mathsf{P}B$.

The case $A = \Diamond_J B$ is analogous and left to the reader. □

Theorem 38 (Soundness of **PMDL**). *If $\Gamma \vdash_{\mathbf{PMDL}} A$ then $\Gamma \Vdash_{\mathbf{PMDL}} A$.*

Proof. Let in the following $M = \langle W, R_{\mathsf{O}}, \langle R_J \rangle_{J \subseteq_{\varnothing} I}, v, w_0 \rangle$ be a **PMDL**-model, $w \in W$, and $J \subseteq_{\varnothing} I$.

Ad (AND): Suppose $M, w \vDash A, B$, then by (C$\wedge$), $M, w \vDash A \wedge B$.

Ad (AN1): Suppose $M, w \vDash A \wedge B$, then by (C$\wedge$), $M, w \vDash A$. The proof for (AN2) is analogous.

Ad (OR1): Suppose $M, w \vDash A$, then by (C$\vee$), $M, w \vDash A \vee B$. The proof for (OR2) is analogous.

Ad (DM1): Suppose $M, w \vDash {\sim}(A \wedge B)$, then by (C$\sim\wedge$), $M, w \vDash {\sim}A \vee {\sim}B$. The proof for (DM2), (DM3) and (DM4) is analogous.

Ad (DN1): Suppose $M, w \vDash A$, then by (C$\sim\sim$), $M, w \vDash {\sim\sim}A$. The proof for (DN2) is analogous.

Ad (EM): This holds by (C$\vee$) and Lemma 25 for all $w \in W$.

Ad (4$\Box_J$): Suppose $M, w \vDash \Box_J A$. Hence by (C$\Box_J$), for all $w' \in R_J w$, $M, w' \vDash A$. Let for some $w' \in R_J w$, $R_J w' w''$. Then by the transitivity of R_J also $R_J w w''$ and hence $M, w'' \vDash A$. Hence, $M, w \vDash \Box_J \Box_J A$.

Ad ($4\Diamond_J$): Suppose $M, w \vDash \Diamond_J \Diamond_J A$. Hence by (C$\Diamond_J$), there is a $w' \in R_J w$ and a $w'' \in R_J w'$ such that $M, w'' \vDash A$. By the transitivity of R_J also $R_J w w''$ and hence $M, w \vDash \Diamond_J A$.

Ad (AND$\Box_J$): Suppose $M, w \vDash \Box_J A$ and $M, w \vDash \Box_J B$. Hence, by (C$\Box_J$) and (C$\wedge$), for all $w' \in R_J w$, $M, w' \vDash A \wedge B$. Hence, again by (C$\Box_J$), $M, w \vDash \Box_J (A \wedge B)$.

Ad (ANDO): The proof is analogous to the one for (AND$\Box_J$).

Ad (AND$'\Box_J$): Suppose $M, w \vDash \Box_J A, \Diamond_J B$. Hence, by (C$\Box_J$) for all $w' \in R_J w$, $M, w' \vDash A$. Moreover, by (C$\Diamond_J$) there is a $w'' \in R_J w$ for which $M, w'' \vDash B$. Thus by (C$\wedge$), $M, w'' \vDash A \wedge B$. By (C$\Diamond_J$), $M, w \vDash \Diamond_J (A \wedge B)$.

Ad (AND$'\mathsf{O}$): The proof is analogous to the one for (AND$'\Box_J$).

Ad (OR$\Diamond_J$): Suppose $M, w \vDash \Diamond_J (A \vee B)$. Hence there is a $w' \in R_J w$ for which $M, w' \vDash A \vee B$. By (C$\vee$), $M, w' \vDash A$ or $M, w' \vDash B$. Hence, by (C$\Diamond_J$) and (C$\vee$), $M, w \vDash \Diamond_J A \vee \Diamond_J B$.

Ad (ORP): The proof is analogous to the one for (OR$\Diamond_J$).

Ad (OR$\Box_J$): Suppose $M, w \vDash \Box_J (A \vee B)$. Hence, for all $w' \in R_J w$, $M, w' \vDash A \vee B$. Suppose there is no $w' \in R_J w$ for which $M, w' \vDash B$. Then by (C$\vee$) for all $w' \in R_J w$, $M, w' \vDash A$ and hence $M, w \vDash \Box_J A$. Suppose there is a $w' \in R_J w$ for which $M, w' \vDash B$ then by (C$\Diamond_J$), $M, w \vDash \Diamond_J B$. Altogether, by (C$\vee$), $M, w \vDash \Box_J A \vee \Diamond_J B$.

Ad (ORO): The proof is analogous to the one for (OR$\Box_J$).

Ad (DO): Suppose $M, w \vDash \mathsf{O}A$. Hence for all $w' \in R_{\mathsf{O}} w$, $M, w' \vDash A$. By the seriality of R_{O} there is such a $w' \in R_{\mathsf{O}} w$ and hence by (CP), $M, w \vDash \mathsf{P}A$.

Ad (T$\Box_J$): Suppose $M, w \vDash \Box_J A$. By (C$\Box_J$) for all $w' \in R_J w$, $M, w \vDash A$. Since R_J is reflexive, $R_J w w$ and hence $M, w \vDash A$.

Ad (T$\Diamond_J$): Suppose $M, w \vDash A$. By the reflexivity of R_J, also $R_J w w$. Hence, by (C$\Diamond_J$), $M, w \vDash \Diamond_J A$.

Ad (R${\sim}\mathsf{O}$) and (R$\mathsf{P}{\sim}$): Note that by (C${\sim}\mathsf{O}$), $M, w \vDash {\sim}\mathsf{O}A$ iff $M, w \vDash \mathsf{P}{\sim}A$.

Ad (R${\sim}\Box$) and (R$\Diamond{\sim}$): The proof is analogous and left to the reader.

Ad (R${\sim}\mathsf{P}$) and (R$\mathsf{O}{\sim}$): Note that by (C${\sim}\mathsf{P}$) $M, w \vDash \mathsf{O}{\sim}A$ iff $M, w \vDash {\sim}\mathsf{P}A$.

Ad (R$\Box{\sim}$) and (R${\sim}\Diamond$): The proof is analogous and left to the reader.

We now show by means of an induction on the number of inference steps needed to derive some formula A that each $w \in W$ is **PMDL**-deductively closed.

For proofs of length 1 this has been demonstrated already above.

For the induction step suppose $\Gamma \vdash_{\mathbf{PMDL}} A$ and A is derived in $n+1$ steps. In case A is derived by (AND), (AN1), (AN2), (OR1), (OR2), (DM1), (DM2), (DM3), (DM4), (DN1), (DN2), (EM), ($4\Box_J$), ($4\Diamond_J$), (AND$\Box_J$), (AND$'\Box_J$), (OR$\Box_J$), (OR$\Diamond_J$), (D$\Box_J$), (T$\Box_J$), (T$\Diamond_J$), (ANDO), (AND$'\mathsf{O}$), (ORP), (ORO), (R${\sim}\mathsf{O}$), (R$\mathsf{P}{\sim}$), (R$\mathsf{O}{\sim}$), (R${\sim}\mathsf{P}$), (R${\sim}\Box$), (R$\Diamond{\sim}$), (R$\Box{\sim}$), or (R${\sim}\Diamond$), we have already shown above that all $w \in W$ are closed under these rules.

Suppose A is derived by means of (RBC) from D and $B \vee C$ and the fact that $\{D, B\} \vdash_{\mathbf{PMDL}} A$ and $\{D, C\} \vdash_{\mathbf{PMDL}} A$. By the induction hypothesis $M, w \vDash B, D$ implies $M, w \vDash A$ and $M, w \vDash C, D$ implies $M, w \vDash A$. Suppose $M, w \vDash D, B \vee C$. By (C$\vee$), $M, w \vDash D, B$ or $M, w \vDash D, C$ and hence $M, w \vDash A$.

Suppose $A = \Box_J A'$ is derived by means of (INH$\Box_J$) from $\Box_J B$ and the fact that $B \vdash_{\mathbf{PMDL}} A$. Suppose $M, w \vDash \Box_J B$. Hence, by (C$\Box_J$), for all $w' \in R_J w$, $M, w' \vDash B$. By the induction hypothesis if $M, w' \vDash B$ then $M, w' \vDash A'$. Hence, $M, w' \vDash A'$. Hence, by (C$\Box_J$), $M, w \vDash \Box_J A'$.

Suppose $A = \Diamond_J A'$ is derived by means of (INH$\Diamond_J$) from $\Diamond_J B$ and the fact that $B \vdash_{\mathbf{PMDL}} A'$. Suppose $M, w \vDash \Diamond_J B$. Hence, by (C$\Diamond_J$), there is a $w' \in R_J w$ for which $M, w' \vDash B$. By the induction hypothesis, if $M, w' \vDash B$ then $M, w' \vDash A'$. Hence, $M, w' \vDash A'$. Hence, by (C$\Diamond_J$), $M, w \vDash \Diamond_J A'$.

The arguments for the rules (INHO) and (INHP) are analogous and left to the reader. □

Definition 20. A set Γ of formulas is *prime* iff for all $A \vee B \in \Gamma$, $\Gamma \cap \{A, B\} \neq \varnothing$.

Where $\mathbf{L}$ is a logic, Γ is $\mathbf{L}$-*deductively closed* iff, if $\Gamma \vdash_{\mathbf{L}} A$ then $A \in \Gamma$.

Definition 21. Let $\Psi_{\mathbf{PMDL}}$ be the set of all prime and **PMDL**-deductively closed subsets of $\mathcal{W}^{\mathbf{MDL}}$.

Definition 22. We define $\mathbf{R}_J \subseteq \Psi_{\mathbf{PMDL}} \times \Psi_{\mathbf{PMDL}}$ as follows: $\mathbf{R}_J \Gamma\Delta$ iff (a) whenever $\Box_J A \in \Gamma$ then $A \in \Delta$, and (b) whenever $A \in \Delta$ then $\Diamond_J A \in \Gamma$.

Definition 23. We define $\mathbf{R}_{\mathsf{O}} \subseteq \Psi_{\mathbf{PMDL}} \times \Psi_{\mathbf{PMDL}}$ as follows: $\mathbf{R}_{\mathsf{O}} \Gamma\Delta$ iff (a) whenever $\mathsf{O}A \in \Gamma$ then $A \in \Delta$, and (b) whenever $A \in \Delta$ then $\mathsf{P}A \in \Gamma$.

Lemma 26. *For all* $\Gamma \subseteq \mathcal{W}^{\mathbf{MDL}}, B \in \mathcal{W}^{\mathbf{MDL}}$ *we have:*

(i) If $\Gamma \vdash_{\mathbf{PMDL}} B$*, then* $\{\mathsf{O}A \mid A \in \Gamma\} \vdash_{\mathbf{PMDL}} \mathsf{O}B$.

(ii) Where Γ *is finite, if* $\Gamma \vdash_{\mathbf{PMDL}} B$*, then* $\mathsf{P} \bigwedge \Gamma \vdash_{\mathbf{PMDL}} \mathsf{P}B$.

(iii) If $\Gamma \vdash_{\mathbf{PMDL}} B$*, then* $\{\Box_J A \mid A \in \Gamma\} \vdash_{\mathbf{PMDL}} \Box_J B$.

(iv) Where Γ *is finite, if* $\Gamma \vdash_{\mathbf{PMDL}} B$*, then* $\Diamond_J \bigwedge \Gamma \vdash_{\mathbf{PMDL}} \Diamond_J B$.

Proof. Ad (i): We prove the statement by means of an induction on the number of inference steps n needed to derive B from Γ in **PMDL**.

"$n = 1$": In case B is derived by a rule $R \notin \{$(AND), (ANDO), (AND$'\mathsf{O}$), (AND$\Box_J$), (AND$'\Box_J$) $\mid J \subseteq_{\varnothing} I\}$ from some $A \in \Gamma$, then $A \vdash_{\mathbf{PMDL}} B$ and hence by (INHO) also $\mathsf{O}A \vdash_{\mathbf{PMDL}} \mathsf{O}B$.

Suppose $R =$ (AND) and B is derived from $A_1, A_2 \in \Gamma$. Note that $\mathsf{O}A_1, \mathsf{O}A_2 \vdash_{\mathbf{PMDL}} \mathsf{O}(A_1 \wedge A_2)$ by (ANDO).

Suppose $R =$ (ANDO) and $B = \mathsf{O}(A_1 \wedge A_2)$ is derived from $\mathsf{O}A_1, \mathsf{O}A_2 \in \Gamma$. Then by (AND$\mathsf{O}$), $\mathsf{OO}A_1, \mathsf{OO}A_2 \vdash_{\mathbf{PDML}} \mathsf{O}(\mathsf{O}A_1 \wedge \mathsf{O}A_2)$. By (INH$\mathsf{O}$), $\mathsf{O}(\mathsf{O}A_1 \wedge \mathsf{O}A_2) \vdash_{\mathbf{PMDL}} \mathsf{OO}(A_1 \wedge A_2)$. Altogether, $\mathsf{OO}A_1, \mathsf{OO}A_2 \vdash_{\mathbf{PMDL}} \mathsf{OO}(A_1 \wedge A_2)$.

Suppose $R =$ (AND$'\mathsf{O}$) and $B = \mathsf{P}(A_1 \wedge A_2)$ is derived from $\mathsf{O}A_1, \mathsf{P}A_2 \in \Gamma$. By (AND$\mathsf{O}$), $\mathsf{OO}A_1, \mathsf{OP}A_2 \vdash_{\mathbf{PMDL}} \mathsf{O}(\mathsf{O}A_1 \wedge \mathsf{P}A_2)$. By (INH$\mathsf{O}$), $\mathsf{O}(\mathsf{O}A_1 \wedge \mathsf{P}A_2) \vdash_{\mathbf{PMDL}} \mathsf{OP}(A_1 \wedge A_2)$. Hence, altogether, $\mathsf{OO}A_1, \mathsf{OP}A_2 \vdash_{\mathbf{PMDL}} \mathsf{OP}(A_1 \wedge A_2)$.

The arguments for $R \in \{$(AND$\Box_J$), (AND$'\Box_J$)$\mid J \subseteq_{\varnothing} I\}$ are analogous and left to the reader.

"$n \Rightarrow n+1$": Suppose B is derived from Γ in $n+1$ inference steps from $A_1, \ldots, A_m$ by means of rule R. By the induction hypothesis $\{\mathsf{O}A \mid A \in \Gamma\} \vdash_{\mathbf{PMDL}} \mathsf{O}A_i$ for all $i \leq m$. If $R \notin \{$(AND), (ANDO), (AND$'\mathsf{O}$), (AND$\Box_J$), (AND$'\Box_J$) $\mid J \subseteq_{\varnothing} I\}$, then $m = 1$ and $A_1 \vdash_{\mathbf{PMDL}} B$ and hence by (INHO) also $\mathsf{O}A_1 \vdash_{\mathbf{PMDL}} \mathsf{O}B$. In case $R \in \{$(AND), (ANDO), (AND$'\mathsf{O}$), (AND$\Box_J$), (AND$'\Box_J$) $\mid J \subseteq_{\varnothing} I\}$ the argument is analogous to the one given above and is left to the reader.

Ad (iii): This case is analogous to case (i) and left to the reader.

Ad (ii): We again proceed by means of an induction similar as in (i). Where $\Gamma \vdash^i A$ means that A is derived from Γ in **PDML** in i steps, we show by an induction that for any $A_1, \ldots, A_m$ for which $\Gamma \vdash^i A_1, \ldots, A_m$ we have $\mathsf{P} \bigwedge \Gamma \vdash_{\mathbf{PDML}} \mathsf{P}(A_1 \wedge \ldots \wedge A_m)$. We show that this holds for any $i \in \mathbb{N}$ and statement (ii) follows immediately.

"$i = 1$": In case B is derived by a rule $R \notin \{$(AND), (ANDO), (AND$'\mathsf{O}$), (AND$\Box_J$), (AND$'\Box_J$) $\mid J \subseteq_\varnothing I\}$ from some $A \in \Gamma$, then $A \vdash_{\mathbf{PMDL}} B$ and hence by (INHP) also $\mathsf{P}A \vdash_{\mathbf{PMDL}} \mathsf{P}B$. By (INHP) also $\mathsf{P} \bigwedge \Gamma \vdash_{\mathbf{PMDL}} \mathsf{P}B$.

Suppose $R =$ (AND) and B is derived from $A_1, A_2 \in \Gamma$. Obviously, $\mathsf{P}(A_1 \wedge A_2) \vdash_{\mathbf{PMDL}} \mathsf{P}(A_1 \wedge A_2)$. Thus, by (INHP), $\mathsf{P} \bigwedge \Gamma \vdash_{\mathbf{PMDL}} \mathsf{P}B$.

Suppose $R =$ (ANDO) and $B = \mathsf{O}(A_1 \wedge A_2)$ is derived from $\mathsf{O}A_1$ and $\mathsf{O}A_2$. By (INHP), $\mathsf{P}(\mathsf{O}A_1 \wedge \mathsf{O}A_2) \vdash_{\mathbf{PMDL}} \mathsf{PO}(A_1 \wedge A_2)$. By (INHP), $\mathsf{P} \bigwedge \Gamma \vdash_{\mathbf{PMDL}} \mathsf{P}B$.

Suppose $R =$ (AND$'\mathsf{O}$) and $B = \mathsf{P}(A_1 \wedge A_2)$ is derived from $\mathsf{O}A_1$ and $\mathsf{P}A_2$. By (INHP), $\mathsf{P}(\mathsf{O}A_1 \wedge \mathsf{P}A_2) \vdash_{\mathbf{PMDL}} \mathsf{P}B$ and hence again by (INHP), $\mathsf{P} \bigwedge \Gamma \vdash_{\mathbf{PMDL}} \mathsf{P}B$.

The arguments for $R \in \{$(AND$\Box_J$), (AND$'\Box_J$)$\mid J \subseteq_\varnothing I\}$ are analogous and left to the reader.

"$i \Rightarrow i+1$": Suppose B is derived from A in $n+1$ inference steps from $A_1, \ldots, A_m$ by means of rule R. By the induction hypothesis $\mathsf{P} \bigwedge \Gamma \vdash_{\mathbf{PMDL}} \mathsf{P}(A_1 \wedge \ldots \wedge A_m)$. If $R \notin \{$(AND), (ANDO), (AND$'\mathsf{O}$), (AND$\Box_J$), (AND$'\Box_J$) $\mid J \subseteq_\varnothing I\}$ then $m = 1$ and $A_1 \vdash_{\mathbf{PMDL}} B$ and hence by (INHP) also $\mathsf{P}A_1 \vdash_{\mathbf{PMDL}} \mathsf{P}B$. In case $R \in \{$(AND), (ANDO), (AND$'\mathsf{O}$), (AND$\Box_J$), (AND$'\Box_J$) $\mid J \subseteq_\varnothing I\}$ the argument is analogous to the one given above and is left to the reader.

Ad (iv): This case is analogous to case (ii) and left to the reader. □

Definition 24. Where $\Gamma \in \Psi_{\mathbf{PMDL}}$ and $A \in \mathcal{W}^{\mathbf{MDL}}$, let

$$
\begin{array}{ll}
\Gamma_{\mathsf{O}} = \{B \mid \mathsf{O}B \in \Gamma\}, & \Gamma_{\Box}^{J} = \{B \mid \Box_J B \in \Gamma\}, \\
\Gamma_{\mathsf{O}}^{A} = \Gamma_{\mathsf{O}} \cup \{A\}, & \Gamma_{\Box}^{J,A} = \Gamma_{\Box}^{J} \cup \{A\}, \\
\Gamma_{\mathsf{P}} = \{B \mid \mathsf{P}B \notin \Gamma\}, & \Gamma_{\Diamond}^{J} = \{B \mid \Diamond_J B \notin \Gamma\}, \\
{}^{\vee}\Gamma_{\mathsf{P}} = \{\bigvee_I B_i \mid B_i \in \Gamma_{\mathsf{P}}\}, & {}^{\vee}\Gamma_{\Diamond}^{J} = \{\bigvee_I B_i \mid B_i \in \Gamma_{\Diamond}^{J}\}, \\
{}^{\vee}\Gamma_{\mathsf{P}}^{A} = \{\bigvee_I B_i \mid B_i \in \Gamma_{\mathsf{P}} \cup \{A\}\}, & {}^{\vee}\Gamma_{\Diamond}^{J,A} = \{\bigvee_I B_i \mid B_i \in \Gamma_{\Diamond}^{J} \cup \{A\}\}.
\end{array}
$$

Lemma 27. *Let $\Gamma \in \Psi_{\mathbf{PMDL}}$. (i) If $\Gamma_{\mathsf{O}} \vdash_{\mathbf{PMDL}} C$, then $\mathsf{O}C \in \Gamma$. (ii) Where $\mathsf{P}A \in \Gamma$, if $\Gamma_{\mathsf{O}}^{A} \vdash_{\mathbf{PMDL}} C$ then $\mathsf{P}C \in \Gamma$.*

Proof. Ad (i): Suppose that $\Gamma_{\mathsf{O}} \vdash_{\mathbf{PMDL}} C$. By Lemma 26i, $\Gamma \vdash_{\mathbf{PMDL}} \mathsf{O}C$. Since Γ is **PMDL**-deductively closed, $\mathsf{O}C \in \Gamma$.

Ad (ii): Suppose $\Gamma_{\mathsf{O}}^{A} \vdash_{\mathbf{PMDL}} C$. Then there is a finite $\Theta \subseteq \Gamma_{\mathsf{O}}$ for which (†) $\Theta \cup \{A\} \vdash_{\mathbf{PMDL}} C$. Since $\Theta \subseteq \Gamma_{\mathsf{O}}$, $\mathsf{O} \bigwedge \Theta \in \Gamma$ by (ANDO) and the deductive closure of Γ. Since also $\mathsf{P}A \in \Gamma$, also $\mathsf{P}(\bigwedge \Theta \wedge A) \in \Gamma$ by (AND$'\mathsf{O}$) and the deductive closure of Γ. Hence, by Lemma 26ii, (†), and the deductive closure of Γ, $\mathsf{P}C \in \Gamma$. □

Lemma 28. *Let $\Gamma \in \Psi_{\mathbf{PMDL}}$. (i) If $\Gamma_{\Box} \vdash_{\mathbf{PMDL}} C$, then $\Box_J C \in \Gamma$. (ii) Where $\Diamond_J A \in \Gamma$, if $\Gamma_{\Box}^{J,A} \vdash_{\mathbf{PMDL}} C$ then $\Diamond_J C \in \Gamma$.*

Proof. Analogous to Lemma 27. We use (AND$\Box_J$), (AND$'\Box_J$), Lemma 26iii, and Lemma 26iv instead of (ANDO), (AND$'\mathsf{O}$), Lemma 26i, and Lemma 26ii respectively. $\Box$

Lemma 29. *Let* $\Gamma \in \Psi_{\mathbf{PMDL}}$. *(i) Where* $\mathsf{P}A \in \Gamma$, ${}^{\vee}\Gamma_{\mathsf{P}} \cap \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma^{A}_{\mathsf{O}}) = \varnothing$. *(ii) Where* $B \notin \Gamma_{\mathsf{O}}$, ${}^{\vee}\Gamma^{B}_{\mathsf{P}} \cap \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma_{\mathsf{O}}) = \varnothing$.

Proof. Ad (i): Let $C = \bigvee_I C_i$ where $C_i \in \Gamma_{\mathsf{P}}$ for all $i \in I$. Assume $C \in \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma^{A}_{\mathsf{O}})$ then by Lemma 27ii, $\mathsf{P}\bigvee_I C_i \in \Gamma$. Hence, by (ORP) and the deductive closure of Γ, also $\bigvee_I \mathsf{P}C_i \in \Gamma$. Since Γ is prime, there is an $i \in I$ such that $\mathsf{P}C_i \in \Gamma$ and hence $C_i \notin \Gamma_{\mathsf{P}}$,—a contradiction.

Ad (ii): Let $C = \bigvee_I C_i$ where $C_i \in \Gamma_{\mathsf{P}} \cup \{B\}$ for all $i \in I$. Assume $C \in \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma_{\mathsf{O}})$. By Lemma 27i and the deductive closure of Γ, ($\star$) $\mathsf{O}\bigvee_I C_i \in \Gamma$. Assume that all $C_i \in \Gamma_{\mathsf{P}}$. By (D$\mathsf{O}$), $\mathsf{P}\bigvee_I C_i \in \Gamma$. By (ORP), $\bigvee_I \mathsf{P}C_i \in \Gamma$. Since Γ is prime there is a $i \in I$ for which $\mathsf{P}C_i \in \Gamma$ and hence $C_i \notin \Gamma_{\mathsf{P}}$,—a contradiction. Thus, there is a non-empty $J \subseteq I$ such that for each $j \in J$, $C_j = B$, and for each $j \in I \smallsetminus J, C_j \neq B$. By ($\star$), (OR$\mathsf{O}$) and the deductive closure of Γ, $\mathsf{O}B \vee \mathsf{P}\bigvee_{I\smallsetminus J} C_i \in \Gamma$. Hence, by (ORP), $\mathsf{O}B \vee \bigvee_{I\smallsetminus J} \mathsf{P}C_i \in \Gamma$. Since $B \notin \Gamma_{\mathsf{O}}$ and since Γ is prime, there is an $i \in I \smallsetminus J$ such that $\mathsf{P}C_i \in \Gamma$ and hence $C_i \notin \Gamma_{\mathsf{P}}$,—a contradiction. $\Box$

Lemma 30. *Let* $\Gamma \in \Psi_{\mathbf{PMDL}}$. *(i) Where* $\Diamond_J A \in \Gamma$, ${}^{\vee}\Gamma_{\Diamond} \cap \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma^{J,A}_{\Box}) = \varnothing$. *(ii) Where* $B \notin \Gamma^{J}_{\Box}$, ${}^{\vee}\Gamma^{J,B}_{\Diamond} \cap \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma^{J}_{\Box}) = \varnothing$.

Proof. The proof is analogous to the proof of Lemma 29. We use Lemma 28, (OR$\Diamond_J$), (D$\Box_J$), and (OR$\Box_J$) instead of Lemma 27, (ORP), (DO), and (ORO) respectively. $\Box$

Lemma 31. *Let* $\Gamma \in \Psi_{\mathbf{PMDL}}$.

(i) Where $\mathsf{P}A \in \Gamma$, *there is a* $\Delta \subseteq \mathcal{W}^{\mathbf{MDL}}$ *for which (1)* $\Gamma^{A}_{\mathsf{O}} \subseteq \Delta$, *(2)* ${}^{\vee}\Gamma_{\mathsf{P}} \cap \Delta = \varnothing$, *and (3)* $\Delta \in \Psi_{\mathbf{PMDL}}$.

(ii) Where $B \notin \Gamma_{\mathsf{O}}$, *there is a* $\Delta \subseteq \mathcal{W}^{\mathbf{MDL}}$ *for which (1)* $\Gamma_{\mathsf{O}} \subseteq \Delta$, *(2)* ${}^{\vee}\Gamma^{B}_{\mathsf{P}} \cap \Delta = \varnothing$, *and (3)* $\Delta \in \Psi_{\mathbf{PMDL}}$.

Proof. Let $\langle \boldsymbol{\Gamma_{\mathsf{O}}}, \boldsymbol{\Gamma_{\mathsf{P}}} \rangle \in \{\langle \Gamma^{A}_{\mathsf{O}}, {}^{\vee}\Gamma_{\mathsf{P}}\rangle, \langle \Gamma_{\mathsf{O}}, {}^{\vee}\Gamma^{B}_{\mathsf{P}}\rangle\}$. Where $\langle B_1, B_2, \ldots \rangle$ is a list of all the members of $\mathcal{W}^{\mathbf{MDL}}$, define $\Delta_0 = \mathrm{Cn}_{\mathbf{PMDL}}(\boldsymbol{\Gamma_{\mathsf{O}}})$ and $\Delta = \Delta_0 \cup \Delta_1 \cup \ldots$ where

$$\Delta_{i+1} = \begin{cases} \mathrm{Cn}_{\mathbf{PMDL}}(\Delta_i \cup \{B_{i+1}\}) & \text{if } \boldsymbol{\Gamma_{\mathsf{P}}} \cap \mathrm{Cn}_{\mathbf{PMDL}}(\Delta_i \cup \{B_{i+1}\}) = \varnothing \\ \Delta_i & \text{else} \end{cases}$$

Ad (1): This holds by the definition of Δ_0 and since $\Delta_0 \subseteq \Delta$.

Ad (2): By Lemma 29, $\Delta_0 \cap \boldsymbol{\Gamma_{\mathsf{P}}} = \varnothing$. The rest follows by the construction of Δ.

Ad (3): We first show that Δ is **PMDL**-deductively closed. Assume there is a $B_i \notin \Delta$ for which (†) $\Delta \vdash_{\mathbf{PMDL}} B_i$. Then, by the construction, there is a $D \in \boldsymbol{\Gamma_{\mathsf{P}}}$ for which $\Delta \cup \{B_i\} \vdash_{\mathbf{PMDL}} D$. Hence, by (†), $\Delta \vdash_{\mathbf{PMDL}} D$. Hence, there is a $j \in \mathbb{N}$ such that $\Delta_j \vdash_{\mathbf{PMDL}} D$. By the construction $\Delta_j = \mathrm{Cn}_{\mathbf{PMDL}}(\Delta_j)$ and thus, $D \in \Delta_j$. Hence, $D \in \Delta$,—a contradiction with (2).

We now show that Δ is prime. Suppose $A_1 \vee A_2 \in \Delta$. Assume that $A_1, A_2 \notin \Delta$. By the construction, $\Delta \cup \{A_1\} \vdash_{\mathbf{PMDL}} D_1$ and $\Delta \cup \{A_2\} \vdash_{\mathbf{PMDL}} D_2$ for some $D_1, D_2 \in \mathbf{\Gamma_P}$. Hence, by (OR1), $\Delta \cup \{A_1\} \vdash_{\mathbf{PMDL}} D_1 \vee D_2$ and by (OR2) $\Delta \cup \{A_2\} \vdash_{\mathbf{PMDL}} D_1 \vee D_2$. Hence, by Fact 8, $\Delta \cup \{A_1 \vee A_2\} \vdash_{\mathbf{PMDL}} D_1 \vee D_2$ and since $A_1 \vee A_2 \in \Delta$, $\Delta \vdash_{\mathbf{PMDL}} D_1 \vee D_2$ and hence $D_1 \vee D_2 \in \Delta$ by the deductive closure of Δ. However, $D_1 \vee D_2 \in \mathbf{\Gamma_P}$,—a contradiction with (2). □

Lemma 32. *Let* $\Gamma \in \Psi_{\mathbf{PMDL}}$.

(i) Where $\Diamond_J A \in \Gamma$, *there is a* $\Delta \subseteq \mathcal{W}^{\mathbf{MDL}}$ *for which (1)* $\Gamma_{\Box}^{J,A} \subseteq \Delta$, *(2)* ${}^{\vee}\Gamma_{\Diamond}^{J} \cap \Delta = \varnothing$, *and (3)* $\Delta \in \Psi_{\mathbf{PMDL}}$.

(ii) Where $B \notin \Gamma_{\Box}^{J}$, *there is a* $\Delta \subseteq \mathcal{W}^{\mathbf{MDL}}$ *for which (1)* $\Gamma_{\Box}^{J} \subseteq \Delta$, *(2)* ${}^{\vee}\Gamma_{\Diamond}^{J,B} \cap \Delta = \varnothing$, *and (3)* $\Delta \in \Psi_{\mathbf{PMDL}}$.

Proof. The proof is analogous to the proof of Lemma 31. Instead of making use of Lemma 29 we now use Lemma 30. □

Lemma 33. *Where* $\Gamma \in \Psi_{\mathbf{PMDL}}$, $\mathsf{P}A \in \Gamma$ *iff there is a* $\Delta \in \Psi_{\mathbf{PMDL}}$ *such that* $\mathbf{R}_{\mathsf{O}}\Gamma\Delta$ *and* $A \in \Delta$.

Proof. Left-Right: Suppose $\mathsf{P}A \in \Gamma$. By Lemma 31i there is a $\Delta \subseteq \mathcal{W}^{\mathbf{MDL}}$ for which (1) $\Gamma_{\mathsf{O}}^{A} \subseteq \Delta$, (2) for all $C \in \Gamma_{\mathsf{P}}$, $C \notin \Delta$, and (3) $\Delta \in \Psi_{\mathbf{PMDL}}$. We now show that $\mathbf{R}_{\mathsf{O}}\Gamma\Delta$. Ad (a): if, for some D, $\mathsf{O}D \in \Gamma$, then $D \in \Gamma_{\mathsf{O}}^{A}$ and hence $D \in \Delta$ by (1). Ad (b): suppose $\mathsf{P}E \notin \Gamma$ for some $E \in \mathcal{W}^{\mathbf{MDL}}$. Then $E \in \Gamma_{\mathsf{P}}$ and thus $E \notin \Delta$ by (2).

Right-Left: follows directly by the definition of $\mathbf{R}_{\mathsf{O}}$. □

Lemma 34. *Where* $\Gamma \in \Psi_{\mathbf{PMDL}}$, $\Diamond_J A \in \Gamma$ *iff there is a* $\Delta \in \Psi_{\mathbf{PMDL}}$ *such that* $\mathbf{R}_J\Gamma\Delta$ *and* $A \in \Delta$.

Proof. The proof is analogous to the proof of Lemma 33, except that we use Lemma 32i instead of Lemma 31i. □

Lemma 35. *For every* $\Gamma \in \Psi_{\mathbf{PMDL}}$ *there is a* $\Delta \in \Psi_{\mathbf{PMDL}}$ *such that* $\mathbf{R}_{\mathsf{O}}\Gamma\Delta$. *(*$\mathbf{R}_{\mathsf{O}}$ *is serial.)*

Proof. By (EM) and (INHP), $\vdash_{\mathbf{PDML}} \mathsf{P}(A \vee {\sim}A)$. Hence, $\mathsf{P}(A \vee {\sim}A) \in \Gamma$ by the deductive closure of Γ. By Lemma 33, there is a $\Delta \in \Psi_{\mathbf{PMDL}}$ such that $\mathbf{R}_{\mathsf{O}}\Gamma\Delta$ and $A \vee {\sim}A \in \Delta$. □

Lemma 36. *Where* $\Gamma \in \Psi_{\mathbf{PMDL}}$, $\mathsf{O}A \in \Gamma$ *iff, for all* $\Delta \in \Psi_{\mathbf{PMDL}}$ *such that* $\mathbf{R}_{\mathsf{O}}\Gamma\Delta$, $A \in \Delta$.

Proof. Left-Right: This is an immediate consequence of the definition of $\mathbf{R}_{\mathsf{O}}$.

Right-Left: Suppose $\mathsf{O}A \notin \Gamma$. Hence $A \notin \Gamma_{\mathsf{O}}$. By Lemma 31ii, there is a $\Delta \subseteq \mathcal{W}^{\mathbf{MDL}}$ for which (1) $\Gamma_{\mathsf{O}} \subseteq \Delta$, (2) $(\Gamma_{\mathsf{P}} \cup \{A\}) \cap \Delta = \varnothing$, and (3) $\Delta \in \Psi_{\mathbf{PMDL}}$. We now show that $\mathbf{R}_{\mathsf{O}}\Gamma\Delta$. Ad (a): if, for some $D \in \mathcal{W}^{\mathbf{MDL}}$, $\mathsf{O}D \in \Gamma$, then $D \in \Gamma_{\mathsf{O}}$ and thus $D \in \Delta$ by (1). Ad (b): Suppose $\mathsf{P}E \notin \Gamma$ and hence $E \in \Gamma_{\mathsf{P}}$. Thus, $E \notin \Delta$ by (2). □

Lemma 37. *Where* $\Gamma \in \Psi_{\mathbf{PMDL}}$, $\Box_J A \in \Gamma$ *iff, for all* $\Delta \in \Psi_{\mathbf{PMDL}}$ *such that* $\mathbf{R}_J\Gamma\Delta$, $A \in \Delta$.

Proof. The proof is analogous to the proof of Lemma 36, except that instead of Lemma 31ii we make use of Lemma 32ii. □

Lemma 38. *Where* $\Gamma \in \Psi_{\mathbf{PMDL}}$, $\mathbf{R}_J\Gamma\Gamma$. *($\mathbf{R}_J$ is reflexive.)*

Proof. Assume there is a $\Gamma \in \Psi_{\mathbf{PMDL}}$ for which $\mathbf{R}_J\Gamma\Gamma$ is not the case. Then, either (1) there is a $\Box_J A \in \Gamma$ such that $A \notin \Gamma$, or (2) there is a $A \in \Gamma$ such that $\Diamond_J A \notin \Gamma$. Ad (1): Since Γ is **PMDL**-deductively closed, and by (T$\Box_J$), $\Box_J A \vdash A$, also $A \in \Gamma$. Ad (2): Since Γ is **PMDL**-deductively closed, and by (T$\Diamond_J$), $A \vdash \Diamond_J A$, also $\Diamond_J A \in \Gamma$. Since neither (1) nor (2) we reached a contradiction. □

Lemma 39. *If* $\mathbf{R}_J\Gamma\Delta$ *and* $\mathbf{R}_J\Delta\Delta'$ *then* $\mathbf{R}_J\Gamma\Delta'$. *($\mathbf{R}_J$ is transitive.)*

Proof. Suppose $\mathbf{R}_J\Gamma\Delta$ and $\mathbf{R}_J\Delta\Delta'$. Assume not $\mathbf{R}_J\Gamma\Delta'$. Thus, either (1) there is a $\Box_J A \in \Gamma$ for which $A \notin \Delta'$, or (2) there is a $A \in \Delta'$ for which $\Diamond_J A \notin \Gamma$. Ad (1): Suppose $\Box_J A \in \Gamma$. By (4$\Box_J$) and the **PMDL**-deductive closure of Γ, also $\Box_J \Box_J A \in \Gamma$. Hence, by (a) in the definition of $\mathbf{R}_J$, $\Box_J A \in \Delta$. Hence, again by (a) in the definition of $\mathbf{R}_J$ and since $\mathbf{R}_J\Delta\Delta'$, $A \in \Delta'$. Ad (2): Suppose $A \in \Delta'$. Hence $\Diamond_J A \in \Delta$ by (b) in the definition of $\mathbf{R}_J$ and since $\mathbf{R}_J\Delta\Delta'$. Hence, $\Diamond_J \Diamond_J A \in \Gamma$ by (b) in the definition of $\mathbf{R}_J$ and since $\mathbf{R}_J\Gamma\Delta$. Since (4$\Diamond_J$) is valid in Γ, also $\Diamond_J A \in \Gamma$.

Since neither (1) nor (2) is the case we reached a contradiction. □

Lemma 40. *Where* $\Delta \in \Psi_{\mathbf{PMDL}}$, *there is a* **PMDL**-*model* M *such that* $M \models A$ *for all* $A \in \Delta$ *and* $M \not\models A$ *for all* $A \in \mathcal{W}^{\mathbf{MDL}} \setminus \Delta$.

Proof. Let $\Delta \in \Psi_{\mathbf{PMDL}}$. We construct a **PMDL**-model

$$M = \langle \Psi_{\mathbf{PMDL}}, \mathbf{R}_{\mathsf{O}}, \langle \mathbf{R}_J \rangle_{J \subseteq_{\varnothing} I}, v, \Delta \rangle$$

such that (†) for all $A \in \mathcal{W}_l^{\sim}$, $w \in v(A)$ iff $A \in w$. By Lemmas 35, 38, 39, $\mathbf{R}_{\mathsf{O}}$ and $\mathbf{R}_J$ (for all $J \subseteq_{\varnothing} I$) have the needed properties for M to be a **PMDL**-model.

We now show by an induction that for all $w \in \Psi_{\mathbf{PMDL}}$ and for all $A \in \mathcal{W}^{\mathbf{MDL}}$, $M, w \models A$ iff $A \in w$. The induction is in terms of the length of the formulas $A \in \mathcal{W}^{\mathbf{MDL}}$ in question.

Let $A \in \mathcal{W}^a$. By (†), $A \in w$ iff $w \in v(A)$ iff [by (Ca)] $M, w \models A$.

For the induction step let first A be of the form $\sim B$.

Let first $B \in \mathcal{W}^a$. By (†), $\sim B \in w$ iff $w \in v(\sim B)$. By (C$\sim$), if $\sim B \in w$ and hence $w \in v(\sim B)$, then $M, w \models \sim B$. Suppose now that $M, w \models \sim B$. By (C$\sim$) either $w \in v(\sim B)$ and hence $\sim B \in w$, or $M, w \not\models B$. In the second case, by the induction hypothesis, $B \notin w$. Since w is prime and since $B \vee \sim B \in w$ (since w is **PMDL**-deductively closed), $\sim B \in w$.

Now let $B = \sim B'$. $M, w \models \sim\sim B'$ iff [by (C$\sim\sim$)] $M, w \models B'$ iff [by the induction hypothesis] $B' \in w$ iff [since w is **DPML**-deductively closed, (DN1) and (DN2)] $\sim\sim B' \in w$.

The cases $B \in \{B_1 \wedge B_2, B_1 \vee B_2\}$ are similar and left to the reader.

Let now $B = \mathsf{O}B'$. $M, w \vDash {\sim}\mathsf{O}B'$ iff [by (C~O)] $M, w \vDash \mathsf{P}{\sim}B'$ iff [by (CP)] there is a $w' \in \mathbf{R}_{\mathsf{O}}w$ for which $M, w' \vDash {\sim}B'$ iff [by the induction hypothesis] ${\sim}B' \in w'$ iff [by the definition of $\mathbf{R}_{\mathsf{O}}$] $\mathsf{P}{\sim}B' \in w$ iff [by (R~O) and (RP~)] ${\sim}\mathsf{O}B' \in w$.

The case $B = \Box_J B'$ is analogous.

Let now $B = \mathsf{P}B'$. $M, w \vDash {\sim}\mathsf{P}B'$ iff [by (C~P)] $M, w \vDash \mathsf{O}{\sim}B'$ iff [by (CO)] for all $w' \in \mathbf{R}_{\mathsf{O}}w$, $M, w' \vDash {\sim}B'$ iff [by the induction hypothesis] for all $w' \in \mathbf{R}_{\mathsf{O}}w$, ${\sim}B' \in w'$ iff [by Lemma 36] $\mathsf{O}{\sim}B' \in w$ iff [by (R~P) and (RO~)] ${\sim}\mathsf{P}B' \in w$.

The case $B = \Diamond_J B'$ is analogous (except that we use Lemma 37 instead of Lemma 36).

Let now $A = B \wedge C$. $M, w \vDash B \wedge C$ iff [by (C$\wedge$)] $M, w \vDash B, C$ iff [by the induction hypothesis] $B, C \in w$ iff [by (AND), (AN1), (AN2) and the fact that w is **PMDL**-deductively closed] $B \wedge C \in w$.

The case $A = B \vee C$ is similar and left to the reader.

Let $A = \mathsf{O}B$. $M, w \vDash \mathsf{O}B$ iff [by (CO)] for all $w' \in \mathbf{R}_{\mathsf{O}}w$, $M, w' \vDash B$ iff [by the induction hypothesis] for all $w' \in \mathbf{R}_{\mathsf{O}}w$, $B \in w'$ iff [by Lemma 36] $\mathsf{O}B \in w$.

The case $A = \Box_J B$ is analogous and left to the reader (just we use Lemma 37 instead of Lemma 36).

Let $A = \mathsf{P}B$. $M, w \vDash \mathsf{P}B$ iff [by (CP)] there is a $w' \in \mathbf{R}_{\mathsf{O}}w$ for which $M, w' \vDash B$ iff [by the induction hypothesis] $B \in w'$ iff [by Lemma 33] $\mathsf{P}B \in w$.

The case $A = \Diamond_J B$ is analogous (just we use Lemma 34 instead of Lemma 33). □

Lemma 41. *Let* $\Gamma \subseteq \mathcal{W}^{\mathbf{MDL}}$ *and* $\Gamma \nvdash_{\mathbf{PMDL}} A$. *There is a* $\Delta \subseteq \mathcal{W}^{\mathbf{MDL}}$ *such that (i)* $\Gamma \subseteq \Delta$, *(ii)* $A \notin \Delta$, *and (iii)* $\Delta \in \Psi_{\mathbf{PMDL}}$.

Proof. Where $\langle B_1, B_2, \ldots \rangle$ is a list of the members of $\mathcal{W}^{\mathbf{MDL}}$, define $\Delta_0 = \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma)$ and $\Delta = \Delta_0 \cup \Delta_1 \cup \ldots$, where

$$\Delta_{i+1} = \begin{cases} \mathrm{Cn}_{\mathbf{PMDL}}(\Delta_i \cup \{B_{i+1}\}) & \text{if } A \notin \mathrm{Cn}_{\mathbf{PMDL}}(\Delta_i \cup \{B_{i+1}\}) \\ \Delta_i & \text{else} \end{cases}$$

Ad (i): This holds by the definition of Δ_0 and since $\Delta_0 \subseteq \Delta$.
Ad (ii): This holds since $A \notin \mathrm{Cn}_{\mathbf{PMDL}}(\Gamma)$ and by the construction of Δ.
Ad (iii): Assume that some $B_i \notin \Delta$ but $\Delta \vdash_{\mathbf{PMDL}} B_i$. Hence, by the construction of Δ, $\Delta_{i-1} \cup \{B_i\} \vdash_{\mathbf{PMDL}} A$ and hence $\Delta \cup \{B_i\} \vdash_{\mathbf{PMDL}} A$. Since also $\Delta \vdash_{\mathbf{PMDL}} B_i$, $\Delta \vdash_{\mathbf{PDML}} A$,—a contradiction with (ii). Hence Δ is **PMDL**-deductively closed.

Suppose $B \vee C \in \Delta$. Assume $B, C \notin \Delta$. Hence, $\Delta \cup \{B\} \vdash_{\mathbf{PMDL}} A$ and $\Delta \cup \{C\} \vdash_{\mathbf{PMDL}} A$. Hence, by Fact 8, $\Delta \cup \{B \vee C\} \vdash_{\mathbf{PMDL}} A$ and since $B \vee C \in \Delta$ also $\Delta \vdash_{\mathbf{PMDL}} A$,—a contradiction with (ii). Hence, Δ is prime. □

Theorem 39 (Strong Completeness of **PMDL**.). *If* $\Gamma \Vdash_{\mathbf{PMDL}} A$ *then* $\Gamma \vdash_{\mathbf{PMDL}} A$.

Proof. Suppose $\Gamma \nvdash_{\mathbf{PMDL}} A$. By Lemma 41 there is a $\Delta \supseteq \Gamma$ such that $A \notin \Delta$ and $\Delta \in \Psi_{\mathbf{PMDL}}$. By Lemma 40, there is a **PMDL**-model M for which $M \vDash B$ for all $B \in \Delta$ and $M \nvDash A$. □

G.3 Proof outline of Theorem 34

The following fact holds since **MDL** strengthens classical propositional logic.

Fact 9. $\Gamma \vdash_{\mathbf{MDL}} \neg A \vee B$ iff $\Gamma \cup \{A\} \vdash_{\mathbf{MDL}} B$.

Lemma 42. *$A \wedge {\sim}A \vdash_{\mathbf{UPMDL}} B$ for all $A, B \in \mathcal{W}^{\mathbf{PMDL}}$*

Proof outline: This is shown by an induction over the complexity of A. Where $A \in \mathcal{W}^a, \diamonddot_i \in \{\mathsf{P}\} \cup \{\Diamond_J \mid J \subseteq_\emptyset I\}$ this holds due to (UPMDL). For the induction step we paradigmatically consider three cases.

(i) Let $A = C \wedge D$. By the induction hypothesis, $C \wedge {\sim}C \vdash_{\mathbf{UPMDL}} B$ and $D \wedge {\sim}D \vdash_{\mathbf{UPMDL}} B$. By (RBC), $(C \wedge {\sim}C) \vee (D \wedge {\sim}D) \vdash_{\mathbf{UPMDL}} B$. By some simple **LP**-manipulations it is easy to see that $(C \wedge D) \wedge {\sim}(C \wedge D) \vdash_{\mathbf{UPMDL}} (C \wedge {\sim}C) \vee (D \wedge {\sim}D)$. Altogether $(C \wedge D) \wedge {\sim}(C \wedge D) \vdash_{\mathbf{UPMDL}} B$.

(ii) Let $A = \mathsf{O}C$. Suppose $\mathsf{O}C \wedge {\sim}\mathsf{O}C$. By (R${\sim}\boxdot$), $\mathsf{O}C \wedge \mathsf{P}{\sim}C$. By (AND'$\boxdot$), $\mathsf{P}(C \wedge {\sim}C)$. By (UPMDL) and the induction hypothesis, B.

(iii) Where $J \subseteq_\emptyset I$, let $A = \Diamond_J C$. Suppose $\Diamond_J C \wedge {\sim} \Diamond_J C$. By (R${\sim}\diamonddot$), $\Diamond_J C \wedge \Box_J {\sim}C$. By (AND'$\boxdot$), $\Diamond_J(C \wedge {\sim}C)$. By (UPMDL) and the induction hypothesis, B.

The other cases are similar and left to the reader. □

Lemma 43. *The following is valid in* **UPMDL***:*

(i) $A, A \supset B \vdash_{\mathbf{UPMDL}} B$

(ii) If $A \vdash_{\mathbf{UPMDL}} B$ *then* $\vdash_{\mathbf{UPMDL}} A \supset B$.

(iii) If $\vdash_{\mathbf{UPMDL}} A \supset B$ *then* $A \vdash_{\mathbf{UPMDL}} B$.

(iv) If $A \vdash_{\mathbf{UPMDL}} B$ *then* ${\sim}B \vdash_{\mathbf{UPMDL}} {\sim}A$.

Proof. Ad (i): Suppose A and ${\sim}A \vee B$. (1) Suppose ${\sim}A$. By A and ${\sim}A$ we get B by Lemma 42. (2) Suppose now B, then by (AND) and (AN1), B. By (1), (2), (RBC), and the supposition, B. *Ad (ii):* Suppose $A \vdash_{\mathbf{UPMDL}} B$. (1) Hence, by (OR2) and the supposition, $A \vdash_{\mathbf{UPMDL}} {\sim}A \vee B$. (2) By (OR1) ${\sim}A \vdash_{\mathbf{UPMDL}} {\sim}A \vee B$. (3) By (EM), $A \vee {\sim}A$. By (1), (2), (3) and (RBC), $\vdash_{\mathbf{UPMDL}} {\sim}A \vee B$. *Ad (iii):* Suppose $\vdash_{\mathbf{UPMDL}} {\sim}A \vee B$. Suppose A. By (i), B. Hence $A \vdash_{\mathbf{UPMDL}} B$. *Ad (iv):* Suppose $A \vdash_{\mathbf{UPMDL}} B$. By (ii), $\vdash_{\mathbf{UPMDL}} {\sim}A \vee B$. By (RBC) and (DN1), $\vdash_{\mathbf{UPMDL}} {\sim}A \vee {\sim}{\sim}B$. By (OR1), (OR2), and (RBC), $\vdash_{\mathbf{UPMDL}} {\sim}{\sim}B \vee {\sim}A$. By (iii), ${\sim}B \vdash_{\mathbf{UPMDL}} {\sim}A$. □

Let $\mathbf{MDL}^{\sim}$ be **MDL** with the negation symbol ${\sim}$ (similarly for $\mathbf{CL}^{\sim}$).

Proof outline of Theorem 32. We first show that all the $\mathbf{MDL}^{\sim}$ axioms are valid in **UPMDL**.

By Lemma 43.i and the fact that all classical theorems are theorems of **LP** (see e.g., [144]), **UPMDL** strengthens $\mathbf{CL}^{\sim}$. Let in the following $\boxdot \in \{\mathsf{O}, \Box_J \mid J \subseteq_\emptyset I\}$ and $\diamonddot \in \{\mathsf{P}, \Diamond_J \mid J \subseteq_\emptyset I\}$.

Ad (AK$\boxdot$): By simple propositional manipulations (henceforth, SPM), $\boxdot({\sim}A \vee B) \vdash_{\mathbf{UPMDL}} \boxdot(B \vee {\sim}A)$. By (OR$\boxdot$), $\boxdot({\sim}A \vee B) \vdash_{\mathbf{UPMDL}} \boxdot B \vee \diamonddot{\sim}A$. By (R$\diamonddot{\sim}$) and some SPM, $\boxdot({\sim}A \vee B) \vdash_{\mathbf{UPMDL}} {\sim} \boxdot A \vee \boxdot B$. By Lemma 43.ii, $\vdash_{\mathbf{UPMDL}} \boxdot({\sim}A \vee B) \supset ({\sim} \boxdot A \vee \boxdot B)$. *Ad (A4$\Box_J$):* This follows by Lemma 43.ii and (4$\Box_J$). *Ad (AT$\Box_J$):* This follows by Lemma 43.ii and (AT$\Box_J$). *Ad (ADf$\diamonddot$):* By (R$\boxdot{\sim}$)

and Lemma 43.iv, $\sim\sim \diamonddot A \vdash_{\mathbf{UPMDL}} \sim \boxdot \sim A$. By (DN1), $\diamonddot A \vdash_{\mathbf{UPMDL}} \sim\sim \diamonddot A$. Hence, $\diamonddot A \vdash_{\mathbf{UPMDL}} \sim \boxdot \sim A$. By Lemma 43.ii, $\vdash_{\mathbf{UPMDL}} \diamonddot A \supset \sim \boxdot \sim A$. In a similar way we get $\vdash_{\mathbf{UPMDL}} \sim \boxdot \sim A \supset \diamonddot A$. By (AND), $\vdash_{\mathbf{UPMDL}} \diamonddot A \equiv \sim \boxdot \sim A$. *Ad (NEC$\boxdot$):* This follows by (INH$\boxdot$). *Ad (AD$\mathsf{O}$):* This follows by (D$\mathsf{O}$) and Lemma 43.ii.

We now show that all the **UPMDL** axioms are valid in $\mathbf{MDL}^{\sim}$.

All the rules and axioms of **LP** hold trivially in $\mathbf{MDL}^{\sim}$ due to the fact that $\mathbf{MDL}^{\sim}$ strengthens $\mathbf{CL}^{\sim}$.

Ad (4$\Box_J$), (4$\Diamond_J$), (T$\Box_J$), (T$\Diamond_J$), (DO): This follows by Fact 9 and (A4$\Box_J$), (A4$\Diamond_J$), (AT$\Box_J$), and (ADO). *Ad (INH$\boxdot$):* This follows by (NEC$\boxdot$), (AK$\boxdot$) and SPM. *Ad (INH$\diamonddot$):* This follows by (INH$\boxdot$), (ADf$\diamonddot$) and SPM. *Ad (AND$\boxdot$):* This follows by (NEC$\boxdot$), (AK$\boxdot$) and by Fact 9. *Ad (AND$'\boxdot$):* By (ADfP), (AND$\boxdot$), by Fact 9 and SPM, $\boxdot A, \sim \diamonddot (A \wedge B) \vdash_{\mathbf{MDL}^{\sim}} \boxdot(A \wedge \sim B)$. By (INH$\boxdot$), $\boxdot(A \wedge \sim B) \vdash_{\mathbf{MDL}^{\sim}} \boxdot \sim B$. By (ADf$\diamonddot$) and SPM, $\boxdot(A \wedge \sim B) \vdash_{\mathbf{MDL}^{\sim}} \sim \diamonddot B$. Altogether, $\boxdot A, \sim \diamonddot (A \wedge B) \vdash_{\mathbf{MDL}^{\sim}} \sim \diamonddot B$. By SPM, $\boxdot A, \diamonddot B \vdash_{\mathbf{MDL}^{\sim}} \diamonddot(A \wedge B)$. *Ad (R$\sim\boxdot$), (R$\diamonddot\sim$), (R$\boxdot\sim$), (R$\sim\diamonddot$):* This follows by (ADf$\diamonddot$) and SPM. *Ad (OR$\diamonddot$):* By (AND$\boxdot$), $\boxdot\sim A \wedge \boxdot \sim B \vdash_{\mathbf{MDL}^{\sim}} \boxdot(\sim A \wedge \sim B)$. By contraposition, (ADf$\diamonddot$), and SPM, $\diamonddot(A \vee B) \vdash_{\mathbf{MDL}^{\sim}} \diamonddot A \vee \diamonddot B$. *Ad (OR$\boxdot$):* By SPM, $\boxdot(A \vee B) \vdash_{\mathbf{MDL}^{\sim}} \boxdot(\sim B \supset A)$. By (AK$\boxdot$) and by Fact 9, $\boxdot(\sim B \supset A) \vdash_{\mathbf{MDL}^{\sim}} \sim \boxdot \sim B \vee \boxdot A$. By SPM and (ADf$\diamonddot$), $\sim \boxdot \sim B \vee \boxdot A \vdash_{\mathbf{MDL}^{\sim}} \boxdot A \vee \diamonddot B$. Altogether, $\boxdot(A \vee B) \vdash_{\mathbf{MDL}^{\sim}} \boxdot A \vee \diamonddot B$. $\Box$

Bibliography

[1] Carlos E. Alchourrón. Logic of norms and logic of normative propositions. *Logique & Analyse*, 47:242–268, 1969.

[2] Carlos E. Alchourrón and Eugenio Bulygin. *Normative Systems*. Springer-Verlag, Wien/New York, 1971.

[3] Carlos E. Alchourrón and Eugenio Bulygin. The expressive conception of norms. In Risto Hilpinen, editor, *New Studies in Deontic Logic*, pages 95–124. D. Reidel Publishing Company, Dordrecht, 1981.

[4] Carlos E. Alchourrón and Eugenio Bulygin. Von Wright on deontic logic and the philosophy of law. In P.A. Schilpp and L.E. Hahn, editors, *The Philosophy of Georg Henrik Von Wright*, pages 665–693. Open Court Publishing Company, 1989.

[5] A.R. Anderson. The reduction of deontic logic to alethic modal logic. *Mind*, 67:100–103, 1958.

[6] A.R. Anderson. The formal analysis of normative systems. In Nicolas Rescher, editor, *The Logic of Decision and Action*, pages 147–213. University of Pittsburgh Press, 1967.

[7] Lennart Åqvist. Good samaritans, contrary-to-duty imperatives, and epistemic obligations. *Noûs*, 1:361–379, 1967.

[8] Lennart Åqvist. Deontic logic. In Dov Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic (2nd edition)*, volume 8, pages 147–264. Kluwer Academic Publishers, 2002.

[9] Nicolas Asher and Daniel Bonevac. Prima facie obligation. *Studia Logica*, 57(1):19–45, 1996.

[10] Alfred Jules Ayer. *Language, Truth and Logic*. Victor Gollancz, London, 1947.

[11] Diderik Batens. Paraconsistent extensional propositional logics. *Logique et Analyse*, 90–91:195–234, 1980.

[12] Diderik Batens. Dynamic dialectical logics. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic. Essays on the Inconsistent*, pages 187–217. Philosophia Verlag, München, 1989.

[13] Diderik Batens. A survey of inconsistency-adaptive logics. In Diderik Batens, Graham Priest, and Jean-Paul van Bendegem, editors, *Frontiers of Paraconsistent Logic*, pages 49–73. Baldock: Research Studies Press, Kings College Publication, 2000.

[14] Diderik Batens. A general characterization of adaptive logics. *Logique & Analyse*, 173–175:45–68, 2001.

[15] Diderik Batens. A procedural criterion for final derivability in inconsistency-adaptive logics. *Journal of Applied Logic*, 3:221–250, 2005.

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

[17] Diderik Batens. Towards a dialogic interpretation of dynamic proofs. In Cédric Dégremont, Laurent Keiff, and Helge Rückert, editors, *Dialogues, Logics and Other Strange Things. Essays in Honour of Shahid Rahman*, pages 27–51. College Publications, London, 2009.

[18] Diderik Batens. Logics for qualitative inductive generalization. *Studia Logica*, 97:61–80, 2011.

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

[20] Diderik Batens and Kristof De Clercq. A rich paraconsistent extension of full positive logic. *Logique et Analyse*, 185-188:227–257, 2004.

[21] Diderik Batens, Kristof De Clercq, Peter Verdée, and Joke Meheus. Yes fellows, most human reasoning is complex. *Synthese*, 166(1):113–131, 2009.

[22] Diderik Batens and Joke Meheus. Recent results by the inconsistency-adaptive labourers. In Jean-Yves Béziau, Walter Carnielli, and Dov Gabbay, editors, *Handbook of Paraconsistency*, pages 81–99. College Publications, London, 2007.

[23] Diderik Batens, Joke Meheus, and Dagmar Provijn. An adaptive characterization of signed systems for paraconsistent reasoning. Under review. Preprint available at `http://logica.ugent.be/centrum/writings/`.

[24] 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 & Analyse*, 207:281–304, 2009.

[25] Diderik Batens and Timothy Vermeir. Direct dynamic proofs for the Rescher-Manor consequence relations: the flat case. *Journal of Applied Non-Classical Logics*, 12:63–84, 2002.

[26] Mathieu Beirlaen. Ethical consistency and the logic of ought. *Theoretical and Applied Ethics*, 1(2):45–51, 2011.

[27] Mathieu Beirlaen. A unifying framework for reasoning about normative conflicts. In Michal Peliš and Vít Punčochář, editors, *The Logica Yearbook 2011*, pages 1–14. College Publications, London, 2012.

[28] Mathieu Beirlaen and Christian Straßer. Nonmonotonic reasoning with normative conflicts in multi-agent deontic logic. Under review. Preprint available at `http://logica.ugent.be/centrum/writings/`.

[29] Mathieu Beirlaen and Christian Straßer. Two adaptive logics of norm-propositions. *Journal of Applied Logic*. In print. Preprint available at `http://logica.ugent.be/centrum/writings/`.

[30] Mathieu Beirlaen and Christian Straßer. A paraconsistent multi-agent framework for dealing with normative conflicts. In João Leite, Paolo Torroni, Thomas Ågotnes, Guido Boella, and Leon van der Torre, editors, *Computational Logic in Multi-Agent Systems. Proceedings of the* 12^{th} *International Workshop CLIMA XII*, volume 6814 of *Lecture Notes in Artificial*

Intelligence, pages 312–329. Springer-Verlag, 2011.

[31] Mathieu Beirlaen, Christian Straßer, and Joke Meheus. An inconsistency-adaptive deontic logic for normative conflicts. *Journal of Philosophical Logic*. In print. DOI: 10.1007/s10992-011-9221-3.

[32] Nuel Belnap and Paul Bartha. Marcus and the problem of nested deontic modalities. In W. Sinnott-Armstrong, D. Raffman, and N. Asher, editors, *Modality, Morality, and Belief. Essays in Honour of Ruth Barcan Marcus*, pages 174–197. Cambridge University Press, Cambridge, 1995.

[33] Nuel Belnap and Michael Perloff. In the realm of agents. *Annals of Mathematics and Artificial Intelligence*, 9:25–48, 1993.

[34] Nuel Belnap, Michael Perloff, and Ming Xu. *Facing the Future. Agents and Choices in Our Indeterminist World.* Oxford University Press, 2001.

[35] Guido Boella and Leon van der Torre. Permissions and obligations in hierarchical normative systems. In *Proceedings of the 9th International Conference on Artificial Intelligence and Law*, ICAIL '03, pages 109–118, New York, NY, USA, 2003. ACM.

[36] Guido Boella, Leon van der Torre, and Harko Verhagen. Introduction to the special issue on normative multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 17:1–10, August 2008.

[37] Helen Bohse. A paraconsistent solution to the problem of moral dilemmas. *South African Journal of Philosophy*, 24(2):77–86, 2005.

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

[39] Jan Broersen. Action negation and alternative reductions for dynamic deontic logics. *Journal of Applied Logic*, 2:153–168, 2004.

[40] Jan Broersen. Probabilistic stit logic. In Weiru Liu, editor, *Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, volume 6717 of *Lecture Notes in Computer Science*, pages 521–531. Springer Berlin / Heidelberg, 2011.

[41] Eugenio Bulygin. On legal gaps. In P.Comanducci and R.Guastini, editors, *Ricerche Di Giurisprudenza Analitica*, pages 21–28. Giappichelli, Turin, 2003.

[42] José Carmo and Andrew Jones. Deontic logic and contrary-to-duties. In Dov Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic (2nd edition)*, volume 8, pages 265–343. Kluwer Academic Publishers, 2002.

[43] Hector-Neri Castañeda. The paradoxes of deontic logic: the simplest solution to all of them in one fell swoop. In Risto Hilpinen, editor, *New Studies in Deontic Logic*, pages 37–85. D. Reidel Publishing Company, Dordrecht, 1981.

[44] Hector-Neri Castañeda. *Thinking and Doing.* Reidel Publishing Company, Dordrecht, 1982.

[45] Hector-Neri Castañeda. Aspectual actions and Davidson's theory of events. In E. LePore and B.P. McLaughlin, editors, *Aspects and Events*, pages 294–310. Blackwell, Oxford, 1985.

[46] Roderick M. Chisholm. Contrary-to-duty imperatives and deontic logic.

Analysis, 27:33–36, 1963.

[47] Earl Conee. Against moral dilemmas. *The Philosophical Review*, 91:87–97, 1982. Reprinted in [72].

[48] Newton Da Costa and Walter Carnielli. On paraconsistent deontic logic. *Philosophia*, 16:293–305, 1986.

[49] Frank Dignum and Lambèr Royakkers. Collective commitment and obligation. In C. Ciampi and E. Marinai, editors, *Proceedings of 5th Int. conference on Law in the Information Society, Firenze, Italy*, pages 1008–1021, 1998.

[50] Alan Donagan. Consistency in rationalist moral systems. *The Journal of Philosophy*, 81:291–309, 1984. Reprinted in [72].

[51] Alan Donagan. Moral dilemmas, genuine and spurious: a comparative anatomy. *Ethics*, 104:7–21, 1993.

[52] Dagfinn Føllesdal and Risto Hilpinen. Deontic logic: an introduction. In Risto Hilpinen, editor, *Deontic Logic: Introductory and Systematic Readings*, pages 1–35. D. Reidel Publishing Company, Dordrecht, 1971.

[53] Philippa Foot. The problem of abortion and the doctrine of double effect. *Oxford Review*, 5:5–15, 1967. Reprinted in [54].

[54] Philippa Foot. *Virtues and Vices*. Basil Blackwell, Oxford, 1978.

[55] James W. Forrester. Gentle murder, or the adverbial samaritan. *The Journal of Philosophy*, 81:193–197, 1984.

[56] James William Forrester. Conflicts of obligation. *American Philosophical Quarterly*, 32(1):31–44, 1995.

[57] James William Forrester. *Being Good & Being Logical*. M.E. Sharpe, Armonk/London, 1996.

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

[59] Tjerk Gauderis. Modelling abduction in science by means of a modal adaptive logic. *Foundations of Science*. In print. DOI: 10.1007/s10699-012-9293-8.

[60] Peter Geach. Whatever happened to deontic logic? In Peter Geach, editor, *Logic and Ethics*, pages 33–48. Kluwer Academic Publishers, 1991.

[61] Harry J. Gensler. *Formal Ethics*. Routledge, London/New York, 1996.

[62] Matt Ginsberg. *Essentials of Artificial Intelligence*. Morgan Kaufmann Publishers Inc., San Francisco, 1994.

[63] Lou Goble. Notes on adaptive deontic logics for normative conflicts. *Unpublished paper*.

[64] Lou Goble. Murder most gentle: the paradox deepens. *Philosophical Studies*, 64:217–227, 1991.

[65] Lou Goble. Multiplex semantics for deontic logic. *Nordic Journal of Philosophical Logic*, 5(2):113–134, 2000.

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

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

[68] Lou Goble. A proposal for dealing with deontic dilemmas. In Alessio

Lomuscio and Donald Nute, editors, *DEON*, volume 3065 of *Lecture Notes in Computer Science*, pages 74–113. Springer, 2004.

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

[70] Lou Goble. Normative conflicts and the logic of ought. *Noûs*, 43:450–489, 2009.

[71] Nelson Goodman. *Fact, Fiction, and Forecast.* London, 1955.

[72] Christopher W. Gowans, editor. *Moral Dilemmas.* Oxford University Press, 1987.

[73] Victor Grassian. *Moral Reasoning. Ethical Theory and Some Contemporary Moral Problems.* Prentice Hall, 1981.

[74] Jaap Hage. *Studies in Legal Logic.* Springer, Berlin, 2005.

[75] Jörg Hansen. Conflicting imperatives and dyadic deontic logic. *Journal of Applied Logic*, 3:484–511, 2005.

[76] Jörg Hansen. *Imperatives and Deontic Logic. On the Semantic Foundations of Deontic Logic.* Dissertation, University of Leipzig, 2008.

[77] Jörg Hansen, Gabriella Pigozzi, and Leon van der Torre. Ten philosophical problems in deontic logic. In Guido Boella, Leon van der Torre, and Harko Verhagen, editors, *Normative Multi-agent Systems*, number 07122 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

[78] Bengt Hansson. An analysis of some deontic logics. *Noûs*, 3:373–398, 1969.

[79] Sven Ove Hansson. Deontic logic without misleading alethic analogies – part 1. *Logique & Analyse*, 123-124:337–353, 1988.

[80] Sven Ove Hansson. Formalization in philosophy. *Bulletin of Symbolic Logic*, 6:162–175, 2000.

[81] Sven Ove Hansson. *The Structure of Values and Norms.* Cambridge University Press, 2001.

[82] R.M. Hare. *The Language of Morals.* Clarendon Press, Oxford, 1961. First published 1952.

[83] R.M. Hare. *Moral Thinking.* Clarendon Press, Oxford, 1981. The relevant chapter on moral conflicts was reprinted in [72].

[84] Marc D. Hauser. *Moral Minds : How Nature Designed Our Universal Sense of Right and Wrong.* Ecco, New York, 2006.

[85] Ingemar Hedenius. *Om Rätt och Moral.* Tiden, Stockholm, 1941.

[86] Vincent F. Hendricks and John Symons, editors. *Formal Philosophy. Aims, Scope, Direction. 5 Questions for Formal Philosophers.* Automatic Press, 2005.

[87] Henning Herrestad and Christen Krogh. Obligations directed from bearers to counterparts. In *Proceedings of the 5th international conference on Artificial intelligence and law*, ICAIL '95, pages 210–218, New York, NY, USA, 1995. ACM.

[88] John Holbo. Moral dilemmas and the logic of obligation. *American Philosophical Quarterly*, 39:259–274, 2002.

[89] Leon Horsten and Philip Welch. The undecidability of propositional adap-

tive logic. *Synthese*, 169:217–218, 2009.

[90] John F. Horty. Deontic logic as founded on nonmonotonic logic. *Annals of Mathematics and Artificial Intelligence*, 9:69–91, 1993.

[91] John F. Horty. Moral dilemmas and nonmonotonic logic. *Journal of Philosophical Logic*, 23(1):35–66, 1994.

[92] John F. 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.

[93] John F. Horty. *Agency and Deontic Logic.* Oxford University Press, 2001.

[94] John F. Horty. Skepticism and floating conclusions. *Artificial Intelligence*, 135:55–72, 2002.

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

[96] John F. Horty and Nuel Belnap. The deliberative stit: a study of action, omission, ability, and obligation. *Journal of Philosophical Logic*, 24:583–644, 1995.

[97] G.E. Hughes and M.J. Cresswell. *A New Introduction to Modal Logic.* Routledge, London/New York, 1996.

[98] Dale Jacquette. Moral dilemmas, disjunctive obligations, and Kant's principle that 'ought' implies 'can'. *Synthese*, 88(1):43–55, 1991.

[99] Jørgen Jørgensen. Imperatives and logic. *Erkenntnis*, 7:288–296, 1937-8.

[100] Georges Kalinowski. *La Logique des Normes.* Presses Universitaires de France, Paris, 1972.

[101] Francis Myrna Kamm. Harming some to save others. *Philosophical Studies*, 57:227–260, 1989.

[102] Stig Kanger. New foundations for ethical theory. In Risto Hilpinen, editor, *Deontic Logic: Introductory and Systematic Readings*, pages 36–58. D. Reidel Publishing Company, Dordrecht, 1971. First published as a privately distributed pamphlet in 1957.

[103] Anthony Kenny. *Will, Freedom and Power.* Basil Blackwell, Oxford, 1975.

[104] Anthony Kenny. *The God of the Philosophers.* Clarendon Press, Oxford, 1979.

[105] Steven Orla Kimbrough. A note on the good samaritan paradox and the disquotation theory of propositional content. In John F. Horty and Andrew J.I. Jones, editors, *Proceedings of ΔEON'02, Sixth International Workshop on Deontic Logic in Computer Science*, pages 139–148, 2002.

[106] Simo Knuuttila. The emergence of deontic logic in the fourteenth century. In Risto Hilpinen, editor, *New Studies in Deontic Logic*, pages 225–248. D. Reidel Publishing Company, Dordrecht, 1981.

[107] Barteld Kooi and Allard Tamminga. Moral conflicts between groups of agents. *Journal of Philosophical Logic*, 37:1–21, 2008.

[108] Robert Koons and Thomas Seung. Defeasible reasoning and moral dilemmas. In Donald Nute, editor, *Defeasible Deontic Logic: Essays in Nonmonotonic Normative Reasoning*, pages 205–222. Kluwer Academic Publishers, 1997.

[109] E.J. Lemmon. Moral dilemmas. *The Philosophical Review*, 70:139–158, 1962.

[110] Hans Lycke. Fitch-style natural deduction for modal paralogics. *Logique et Analyse*, 207:193–218, 2009.

[111] Hans Lycke. Inconsistency-adaptive modal logics. On how to cope with modal inconsistency. *Logic and Logical Philosophy*, 19:1–31, 2010.

[112] Hans Lycke. A formal explication of the search for explanations: the adaptive logics approach to abductive reasoning. *Logic Journal of IGPL*, 20(2):497–516, 2012.

[113] David Makinson. On a fundamental problem of deontic logic. In Paul McNamara and Henri Prakken, editors, *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science*, pages 29–53. IOS Press, 1999.

[114] David Makinson and Karl Schlechta. Floating conclusions and zombie paths: two deep difficulties in the “directly skeptical” approach to defeasible inheritance nets. *Artificial Intelligence*, 48:199–209, 1991.

[115] David Makinson and Leon van der Torre. Input/output logics. *Journal of Philosophical Logic*, 29:383–408, 2000.

[116] David Makinson and Leon van der Torre. Constraints for input/output logics. *Journal of Philosophical Logic*, 30:155–185, 2001.

[117] David Makinson and Leon van der Torre. Permission from an input/output perspective. *Journal of Philosophical Logic*, 32(4):391–416, 2003.

[118] David Makinson and Leon van der Torre. What is input/output logic? input/output logic, constraints, permissions. In Guido Boella, Leon van der Torre, and Harko Verhagen, editors, *Normative Multi-agent Systems*, number 07122 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

[119] Ernst Mally. *Grundgesetze des Sollens: Elemente der Logik des Willens.* Leuschner and Lubensky, Graz, 1926.

[120] Ruth Barcan Marcus. Iterated deontic modalities. *Mind (New Series)*, 75:580–582, 1966.

[121] Ruth Barcan Marcus. Moral dilemmas and consistency. *Journal of Philosophy*, 77:121–136, 1980. Reprinted in [72].

[122] Terrance McConnell. Moral dilemmas and consistency in ethics. *Canadian Journal of Philosophy*, 8:269–287, 1978. Reprinted in [72].

[123] Terrance McConnell. Moral dilemmas. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Summer 2010 edition, 2010.

[124] Casey McGinnis. *Paraconsistency and Deontic Logic: Formal Systems for Reasoning with Normative Conflicts.* Dissertation, University of Minnesota, 2007.

[125] Casey 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.

[126] Paul McNamara. Deontic logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* 2010.

[127] Paul McNamara and Henri Prakken. Introduction. In Paul McNamara and Henri Prakken, editors, *Norms, Logics and Information Systems. New*

Studies on Deontic Logic and Computer Science, pages 1–14. IOS Press, 1999.

[128] Joke Meheus and Diderik Batens. A formal logic for abductive reasoning. *Logic Journal of The IGPL*, 14:221–236, 2006.

[129] Joke Meheus, Mathieu Beirlaen, and Frederik Van De Putte. Avoiding deontic explosion by contextually restricting aggregation. In Guido Governatori and Giovanni Sartor, editors, *DEON (10th International Conference on Deontic Logic in Computer Science)*, volume 6181 of *Lecture Notes in Artificial Intelligence*, pages 148–165. Springer, 2010.

[130] Joke Meheus, Mathieu Beirlaen, Frederik Van De Putte, and Christian Straßer. Non-adjunctive deontic logics that validate aggregation as much as possible. Under review. Preprint available at `http://logica.ugent.be/centrum/writings/`.

[131] J.-J. C. Meyer. A simple solution to the "deepest" paradox in deontic logic. *Logique & Analyse*, 117–118:81–90, 1987.

[132] J.-J. C. Meyer. A different approach to deontic logic: deontic logic viewed as a variant of dynamic logic. *Notre Dame Journal of Formal Logic*, 29:109–136, 1988.

[133] J.-J. C. Meyer and R.J. Wieringa. Applications of deontic logic in computer science: A concise overview. In J.-J. C. Meyer and Roel J. Wieringa, editors, *Deontic Logic in Computer Science: Normative System Specification*, pages 17–40. John Wiley & Sons, 1993.

[134] Karin B. Michels and Kenneth J. Rothman. Update on unethical use of placebos in randomised trials. *Bioethics*, 17:188–204, 2003.

[135] John Mikhail. Universal moral grammar: theory, evidence and the future. *Trends in Cognitive Science*, 11:143–152, 2007.

[136] Marko Milanović. A norm conflict perspective on the relationship between international humanitarian law and human rights law. *Journal of Conflict & Security Law*, 14:459–483, 2010.

[137] John Stuart Mill. Utilitarianism. In J.M. Robson, editor, *Collected Works of John Stuart Mill, Vol. X: Essays on Ethics, Religion and Society*, pages 203–259. University of Toronto Press, 1969.

[138] Michael Morreau. Prima facie and seeming duties. *Studia Logica*, 57(1):47–71, 1996.

[139] P.H. Nowell-Smith and E.J. Lemmon. Escapism: the logical basis of ethics. *Mind (New Series)*, 69:289–300, 1960.

[140] Paul Pietroski. Prima facie obligations, ceteris paribus laws in moral theory. *Ethics*, 103:489–515, 1993.

[141] John Pollock. Epistemic norms. *Synthese*, 71(1):61–95, 1987.

[142] John. Pollock. Defeasible reasoning. In Jonathan E. Adler and Lance J. Rips, editors, *Reasoning. Studies of Human Inference and Its Foundations*, pages 451–470. Cambridge University Press, 2008.

[143] Graham Priest. The logic of paradox. *Journal of Philosophical Logic*, 8:219–241, 1979.

[144] Graham Priest. Paraconsistent logic. In Dov Gabbay and Franz Guenthner, editors, *Handbook of Philosophical Logic (2nd edition)*, volume 8, pages

287–393. Kluwer Academic Publishers, 2002.

[145] Graham Priest. *In Contradiction: A Study of the Transconsistent (2nd Edition)*. Oxford University Press, 2006.

[146] Graham Priest. *An Introduction To Non-Classical Logic (2nd Edition)*. Cambridge University Press, 2008.

[147] Giuseppe Primiero and Joke Meheus. Majority merging by adaptive counting. *Synthese*, 165(2):203–223, 2008.

[148] Arthur N. Prior. Escapism: the logical basis of ethics. In A.I. Melden, editor, *Essays in Moral Philosophy*, pages 135–146. University of Washington Press, Seattle, 1958.

[149] Arthur N. Prior. *Formal Logic*. Oxford University Press, Oxford, 1962.

[150] Raymond Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[151] Nicholas Rescher. *Topics in Philosophical Logic*. Reidel, Dordrecht, 1968.

[152] Alf Ross. Imperatives and logic. *Philosophy of Science*, 11:30–46, 1944.

[153] David W. Ross. *The Right and the Good*. Oxford University Press, 1963.

[154] Richard Routley and Val Plumwood. Moral dilemmas and the logic of deontic notions. In Graham Priest, Richard Routley, and Jean Norman, editors, *Paraconsistent Logic. Essays on the Inconsistent*, pages 653–702. Philosophia Verlag, München, 1989.

[155] Jean-Paul Sartre. *L'Existentialisme est un Humanisme*. Editions Nagel, Paris, 1946.

[156] Geoffrey Sayre-McCord. *Essays on Moral Realism*. Cornell University Press, New York, 1988.

[157] Peter Schotch and Raymond Jennings. Non-Kripkean deontic logic. In Risto Hilpinen, editor, *New Studies In Deontic Logic*, pages 149–162. D. Reidel Publishing Company, Dordrecht, 1981.

[158] Lutz Schröder and Dirk Pattinson. Rank-1 modal logics are coalgebraic. *Journal of Logic and Computation*, 5(20):1113–1147, December 2010.

[159] Krister Segerberg. Getting started: beginnings in the logic of action. *Studia Logica*, 51:347–378, 1992.

[160] Yoav Shoham. Nonmonotonic logics: Meaning and utility. In *Proceedings of IJCAI'87*, pages 388–393, 1987.

[161] Yoav Shoham. A semantical approach to nonmonotonic logics. In Matthew L. Ginsberg, editor, *Readings in Nonmonotonic Reasoning*, pages 227–250. Morgan Kaufmann Publishers, 1987.

[162] Yoav Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*. M.I.T. Press, Cambridge,Mass., 1988.

[163] Walter Sinnott-Armstrong. A solution to Forrester's paradox of gentle murder. *The Journal of Philosophy*, 82:162–168, 1982.

[164] Walter Sinnott-Armstrong. *Moral Dilemmas*. Basil Blackwell, Oxford/New York, 1988.

[165] Erik Stenius. The principles of a logic of normative systems. *Acta Philosophica Fennica*, 16:247–260, 1963.

[166] Erik Stenius. Ross' paradox and well-formed codices. *Theoria*, 48:49–77, 1982.

[167] Charles Leslie Stevenson. *Ethics and Language.* Yale University Press, 1947.

[168] Michael Stocker. *Plural and Conflicting Values.* Clarendon Press, Oxford, 1990.

[169] Audun Stolpe. A theory of permission based on the notion of derogation. *Journal of Applied Logic*, 8(1):97–113, 2010.

[170] Christian Straßer. An adaptive logic for rational closure. In Marcello Coniglio and Itala Loffredo D'Ottaviano, editors, *The Many Sides of Logic*, pages 47–67. College Publications, London, 2009.

[171] Christian Straßer. An adaptive logic framework for conditional obligations and deontic dilemmas. *Logic and Logical Philosophy*, 19(1-2):95–128, 2010.

[172] Christian Straßer. *Defeasible Reasoning With Applications in Argumentation, Normative Reasoning and Default Reasoning.* PhD thesis, Ghent University, 2010.

[173] Christian Straßer. A deontic logic framework allowing for factual detachment. *Journal of Applied Logic*, 9(1):61–80, 2011.

[174] Christian Straßer and Mathieu Beirlaen. Towards more conflict-tolerant deontic logics by relaxing the interdefinability between obligations and permissions. Unpublished paper. Preprint available at `http://logica.ugent.be/centrum/writings/`.

[175] Christian Straßer, Mathieu Beirlaen, and Joke Meheus. Tolerating deontic conflicts by adaptively restricting inheritance. *Logique & Analyse.* In print. Preprint available at `http://logica.ugent.be/centrum/writings/`.

[176] 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(2):133–156, 2011.

[177] Judith Jarvis Thomson. Killing, letting die, and the trolley problem. *The Monist*, 59:204–217, 1976.

[178] Judith Jarvis Thomson. The trolley problem. *Yale Law Journal*, 94:1395–1415, 1985.

[179] Peter Unger. *Living High and Letting Die.* Oxford University Press, Oxford, 1996.

[180] Peter Vallentyne. Prohibition dilemmas and deontic logic. *Logique & Analyse*, 117-118:113–122, 1987.

[181] Frederik Van De Putte. *Generic Formats for Prioritized Adaptive Logics. With Applications in Deontic Logic, Abduction and Belief Revision.* PhD thesis, Ghent University, 2012.

[182] Frederik Van De Putte and Christian Straßer. A logic for prioritized normative reasoning. *Journal of Logic and Computation.* In print. DOI: 10.1093/logcom/exs008.

[183] Frederik Van De Putte and Peter Verdée. The dynamics of relevance: Adaptive belief revision. *Synthese*, to appear, 2012. Preprint available at `http://logica.UGent.be/centrum/writings/`.

[184] Jeroen van den Hoven and Gert-Jan Lokhorst. Deontic logic and computer-supported computer ethics. In J.H. Moor and T.W. Bynum, editors, *Cyberphilosophy: The Intersection of Computing and Philosophy*, pages 280–289.

Blackwell, Malden (MA), 2002.

[185] Leon van der Torre and Yao Hua Tan. Two-phase deontic logic. *Logique et Analyse*, (171-172):411–456, 2000.

[186] J.A. van Eck. A system of temporally relative modal and deontic predicate logic and its philosophical applications. *Logique & Analyse*, 99:249–290, 1982.

[187] Peter Verdée. Adaptive logics using the minimal abnormality strategy are π^1_1-complex. *Synthese*, 167:93–104, 2009.

[188] Peter Verdée and Stephan van der Waart van Gulik. A generic framework for adaptive vague logics. *Studia Logica*, 90:385–405, 2008.

[189] Liza Verhoeven. Proof theories for some prioritized consequence relations. *Logique et Analyse*, 183–184:325–344, 2003. Appeared 2005.

[190] Georg Henrik von Wright. Deontic logic. *Mind*, 60:1–15, 1951.

[191] Georg Henrik von Wright. *Logical Studies.* Routledge and Kegan, London, 1957.

[192] Georg Henrik von Wright. *Norm and Action. A Logical Enquiry.* Routledge and Kegan Paul, London, 1963.

[193] Georg Henrik von Wright. *An Essay in Deontic Logic and the General Theory of Action.* North-Holland, Amsterdam, 1968.

[194] Georg Henrik von Wright. Deontic logic: a personal view. *Ratio Juris*, 12(1):26–38, 1999.

[195] Peter Vranas. In defense of imperative inference. *Journal of Philosophical Logic*, 39:59–72, 2010.

[196] Wendell Wallach and Colin Allen. *Moral Machines. Teaching Robots Right from Wrong.* Oxford University Press, Oxford, 2009.

[197] Heinrich Wansing. Nested deontic modalities: another view of parking on highways. *Erkenntnis*, 49:185–199, 1998.

[198] Zach Weber. On paraconsistent ethics. *South African Journal of Philosophy*, 26(2):240–245, 2007.

[199] Anders Wedberg. Some problems in the logical analysis of legal science. *Theoria*, 17:246–275, 1951.

[200] Ota Weinberger. *Alternative Action Theory.* Kluwer Academic Publishers, Dordrecht/Boston/London, 1998.

[201] Gregory Wheeler and Marco Alberti. NO revision and NO contraction. *Minds and Machines*, 21(3):411–430, 2011.

[202] Vincent Wiegel, Jeroen van den Hoven, and Gert-Jan Lokhorst. Privacy, deontic epistemic action logic and software agents. *Ethics and Information Technology*, 7(4):251–264, 2005.

[203] Bernard Williams. Ethical consistency. *Proceedings of the Aristotelian Society (Supplementary Volumes)*, 39:103–124, 1965. Reprinted in [72, 156, 204].

[204] Bernard Williams. *Problems of the Self: Philosophical Papers 1956-1972.* Cambridge University Press, 1973.

[205] Jan Woleński. Deontic logic and possible worlds semantics: a historical sketch. *Studia Logica*, 49(2):273–282, 1990.

ISBN 978-90-7083-082-3
9 789070 830823