Paraconsistency and Deontic Logic: Formal Systems for Reasoning with Normative Conflicts

Casey McGinnis

December 17, 2007

Contents

Preface

> Rushing into action, you fail.
> Trying to grasp things, you lose them.
> Forcing a project to completion,
> you ruin what was almost ripe.
>
> —*Tao Te Ching*[1]

I began this project in 2002, after reading *In Contradiction: A Study of the Transconsistent* [129] by Graham Priest. In Chapter 13 of that intriguing and thought-provoking book, Priest argues for the possibility of normative conflicts and presents a paraconsistent deontic logic for dealing with them. The intersection of paraconsistent logic and deontic logic seemed like a promising area to explore further, and soon I had some ideas—some of them apparently new—on how to undertake such an exploration.

The project took longer than expected to complete. Why? First, the inevitable "real world" distractions, detours, and debacles, which I will not go into here. Second, the nature of the topic together with the nature of my academic background. I am an analytic philosopher with a fair amount of training in (classical) logic and metalogic, but, unfortunately, little formal background in mathematics. Moreover, I have never taken a course in modal or non-classical logic, two fields in which I had to become reasonably competent in order to write this dissertation. (Priest's excellent *Introduction to Non-Classical Logic* [127] was the starting point for my self-training, and the influence of this book will be apparent throughout the present work.) Third, I scrapped the whole thing and started over from scratch numerous times (after going down numerous roads that were either dead ends or would have taken me to far afield).

[1] Trans. by Stephen Mitchell (New York: HarperPerennial, 1988).

The fourth and final reason that the project took longer than expected is that I had to make *three* people happy (or, at least, not overly *un*happy): my official advisors, William Hanson and Geoffrey Hellman of the University of Minnesota, Twin Cities, and my "unofficial" advisor, Lou Goble of Willamette University. I sometimes wonder why I chose to make things so difficult for myself. In the long run, however, it paid off: receiving extensive comments and criticism from three people, instead of the more typical one or two, made this a much better dissertation than it would have been otherwise (which, admittedly, may not be saying much). I am extremely grateful for my advisors' consistent encouragement and advice, without which I might have given up on this (sometimes frustrating) project a long time ago. If there are any non-trivial insights in this dissertation, I am certain that a significant percentage of them are due to my advisors. (In a few cases, when it seems particularly appropriate, I explicitly give them credit (in footnotes) for certain observations and arguments. There are many other cases in which their influence is not explicitly acknowledged.)

Work on this project was supported by a Swenson-Kierkegaard fellowship (spring semester, 2002) and a University of Minnesota Graduate School doctoral dissertation fellowship (2002-2003 academic year).

Portions of the material in this dissertation were presented at the following meetings:

- Australasian Association for Logic (Canberra, Australia, December 2002)
- Society for Exact Philosophy (Vancouver, Canada, May 2003)
- World Congress on Paraconsistency III (Toulouse, France, July 2003)
- First World Congress on Universal Logic (Montreux, Switzerland, March 2005)
- Workshop on Logic, Language, Information, and Computation (Florianópolis, Brazil, July 2005)
- LOGICA (Hejnice, Czech Republic, June 2006)
- Philosophy Department Colloquium (University of Minnesota, Duluth, March 2007)

I am grateful to the participants in these gatherings for their insightful feedback and discussion—especially Diderik Batens, Jean-Yves Béziau, Katalin Bimbó, Ross

Brady, Manuel Bremer, Bryson Brown, Walter Carnielli, David Cole, Alexandre Costa-Leite, Max Cresswell, Itala D'Ottaviano, Marcelo Finger, Melvin Fitting, Jason Ford, John Kearns, Nils Kurbis, João Marcos, Ed Mares, Ana Teresa Martins, Bob Meyer, Fred Michael, Gillman Payette, Graham Priest, Greg Restall, Peter Schotch, Hartley Slater, Heinrich Wansing, Wen-Fang Wang, Paul Wong, and, of course, anyone I may have accidentally omitted. I have also benefited from informal discussions and email exchanges with many of these people.

I would like to thank Richard L. Epstein and Graham Priest for providing detailed critical comments on early drafts. Manuel Bremer, Nils Kurbis, and Jack Stecher also provided some helpful feedback on such drafts. Obviously, any mistakes or other infelicities herein are my responsibility alone.

Thanks to MacKichan software for its wonderful program, Scientific Word, which helped ease me into the initially intimidating world of LaTeX typesetting. Lou Goble, Steve Lelchuk, and João Marcos also helped be with some LaTeX-related issues.

Thanks to my logic teachers: Don Moor (Professor Emeritus of Philosophy at Portland State University), William Hanson, Michael Kac, Geoffrey Hellman, and Stephen Donaho (the best teaching assistant in the history of the universe).

Finally, I wish to thank my parents, Neil and Joan McGinnis, for their unfailing encouragement and support.

Introduction

> Generally speaking, a man is supposed to walk to the left of a woman and also keep himself between her and the curb. Of course, it is frequently impossible to do both. But the great thinkers of all ages have been unanimous in their admiration of paradoxes.
>
> —P. J. O'Rourke, *Modern Manners*[2]

> "You may keep one of the children," he repeated. "The other one will have to go. Which one will you keep?"
>
> "You mean, I have to choose?"
>
> "You're a Polack, not a Yid. That gives you a privilege—a choice."
>
> Her thought process dwindled, ceased. Then she felt her legs crumple. "I can't choose! I can't choose!" She began to scream. Oh, how she recalled her own screams! Tormented angels never screeched so loudly above hell's pandemonium. *"Ich kann nicht wählen!"* she screamed.
>
> The doctor was aware of unwanted attention. "Shut up!" he ordered. "Hurry now and choose. Choose, goddamnit, or I'll send them both over there. Quick!"
>
> —William Styron, *Sophie's Choice*[3]

As the above quotations illustrate, normative conflicts range from the trivial and amusing to the heart-wrenching and deadly serious. Standard systems of deontic logic—the logic of obligation, prohibition, and permission—break down in the presence of normative conflicts: in them, the existence of a single normative conflict entails that *everything* is obligatory (as well as forbidden). This is bad. First of all,

[2] *Modern Manners: An Etiquette Guide for Rude People* (New York: Atlantic Monthly Press, 1983), p. 27.

[3] New York: Bantam, 1979, p. 589.

it just seems wrong. Second, it renders standard deontic logic completely useless in precisely those cases in which it is needed most, namely, those in which one is faced with a (real or apparent) clash of norms. Third, even if one thinks that normative conflicts are conceptually impossible, there are still cases in which it is important to avoid inferring arbitrary propositions from the assumption that a normative conflict obtains. Suppose, for example, that we want to program artificial intelligence systems with deontic logic in order to allow them to reason about what they may and may not do—something computer scientists have recently proposed.[4] We do not want a robot who has somehow acquired the beliefs *I am obliged to open the door* and *I am forbidden to open the door* to derive from these an additional belief, *I am obliged to kill everyone in sight* (though, admittedly, this might make for a good science fiction/horror movie).

This dissertation presents and investigates some formal systems for reasoning about normative conflicts in a sane and discriminating manner. Clearly this requires departing from standard deontic logic. Our main approach will be to develop conflict-tolerant deontic logics by constructing them on the basis of a *paraconsistent* logic—roughly, one that tolerates contradictions without "exploding" into triviality. The adequacy of paraconsistent logic as a replacement for classical logic will not be addressed in any detail. Rather, the focus will be on whether and how paraconsistency can be used to create a satisfactory conflict-tolerant deontic logic.

Chapter 1 gives a more precise account of what normative conflicts are, argues that they are (logically) possible, and addresses a number of objections. In the process, it introduces some important concepts and distinctions. While, as noted above, our investigation is worthwhile even if normative conflicts are impossible, the assumption that they are provides additional significance and motivation. A main point of the chapter is that, at most, arguments against the possibility of normative conflicts show that there is a *tradeoff* involved in allowing for their possibility: one must give up one or more principles that are *prima facie* plausible, but hardly immune to rational challenge. Which of these principles, if any, *should* be given up in order to allow for normative conflicts is a question that is addressed later in the dissertation (specifically, in Chapter 6), after conflict-tolerant logics have been developed in some detail. With certain facts about such systems at our disposal, we will be in a better position to give a well-informed, defensible answer to the question.

[4] See, e.g., Arkoudas et al. [11].

Chapter 2 presents the "standard" system of deontic logic, along with some closely related systems, and explains precisely how and why these systems break down in the presence of normative conflicts. These logics (which I call 'the D systems') are presented both semantically and proof-theoretically (using tableaus). Two alternative (but still classically-based) systems of deontic logic are then considered, and these too are rejected on the grounds that they are too weak in certain senses and too strong in others. I conclude that the classical principle of explosion, according to which "everything follows from a contradiction," is ultimately to blame for the failings of these systems, and that it is thus wise to investigate deontic logics that are constructed on the basis of logics that reject explosion, i.e., paraconsistent logics.

Chapter 3 introduces two basic paraconsistent logics, P4 and P3, that will serve as foundations for the conflict-tolerant deontic logics that are constructed in subsequent chapters. An attempt is made to show that P4 and P3 are independently motivated by certain intuitive considerations, and not merely *ad hoc* mathematical constructions with certain desirable formal properties. The systems are characterized both semantically and tableau-theoretically, and these characterizations are shown to be equivalent. P3, as it turns out, is equivalent to a number of systems existing in the literature, most notably the system J_3 introduced by D'Ottaviano and da Costa [50] in 1970. The tableau-style proof theories given here are new, though they are similar to the one given by Carnielli and Marcos [38] for their system LFI1 (which is also equivalent to J_3) and make liberal use of the techniques in Priest's *Introduction to Non-Classical Logic* [127].

I begin Chapter 4 by reviewing the existing literature on paraconsistent deontic logic in order to place the current approach in perspective. I then introduce a small family of paraconsistent deontic logics, the PD systems, that are based on P3 and P4 in the same way that the D systems are based on classical propositional logic (CPL). As usual, these systems are specified in terms of both semantic and proof-theoretic consequence relations, and the latter are shown to be sound and complete with respect to the former. The PD systems are new, not appearing anywhere in the literature, though they are similar in some ways to the systems of paraconsistent deontic logic introduced by da Costa and Carnielli [47] in 1986, as well as some of the systems in Goble's "Paraconsistent Modal Logic" [67] and my "Tableau Systems for Some Paraconsistent Modal Logics" [111]. The PD systems are compared and contrasted with the standard systems discussed in Chapter 2, and are found to be superior in most ways. However, these systems are subject to two particular objections, which are noted at the end of the chapter.

In Chapter 5, I consider an alternative approach to constructing a conflict-tolerant deontic logic—a sort of compromise between standard deontic logic and paraconsistent deontic logic which I call "semi-paraconsistent" deontic logic. This is an attempt to address the concerns raised at the end of Chapter 4. This approach is, to my knowledge, completely new; nothing like it has appeared in the literature (aside from my [115]). A small family of semi-paraconsistent deontic logics, the SPD systems, is defined. After demonstrating the usual soundness and completeness results, I highlight some interesting features of these systems, the most notable of which is that they tolerate normative conflicts while preserving *all* of classical propositional logic. I finish the chapter by considering and responding to some objections to semi-paraconsistent deontic logic (as I have presented it).

In Chapter 6, the concluding chapter, I revisit the "tradeoff" discussed in Chapter 1, and summarize what has been accomplished in the dissertation.

There are two appendices. Appendix A is a guide to the notation used in the dissertation. Appendix B presents (in concise form) the tableau rules for the logical systems that are investigated in detail in the body of the dissertation.

I conclude this Introduction with some miscellaneous remarks on scope, methodology, and style:

I assume that the reader is familiar with the essentials of classical first-order logic and set theory. No further background is presupposed, though familiarity with Priest's *Introduction to Non-Classical Logic* [127] (or a similar work such as Beall and Van Fraassen's *Possibilities and Paradox* [23]) would be very helpful.

I consider only propositional logics, i.e., logics without quantification, since the logical features in which I am primarily interested are present at the propositional level. Extending the propositional systems to include quantifiers ought to be fairly straightforward.

All proof theories in this dissertation are tableau-style. I favor tableaus over axioms, natural deduction, etc. for two main reasons. First, tableau proofs require no ingenuity to construct (and as such, are easier on humans and computers alike[5]).

[5] I have found that computer scientists are much more amenable to tableau-style proof theories than philosophers are. The reason is fairly obvious: tableaus do not require the mysterious element of *human intelligence*.

Second, they automatically generate counterexamples to invalid inferences—a feature that I find quite convenient. I write 'tableaus' rather than 'tableaux', since the latter strikes me as a bit pretentious.

Like most logicians, I deliberately sacrifice some precision and strict accuracy for the sake of succinctness, readability, and liveliness. For example, when expressing schemas[6] in formal languages, I generally leave corner quotes implicit. I *do* occasionally use corner quotes, but mainly just to express certain schemas in natural language, e.g. $\ulcorner$It is obligatory that $A\urcorner$. (It would be inappropriate to use regular quotation marks here, since A is a placeholder that is not actually being quoted. Leaving the corner quotes implicit would result in something unclear and apparently ungrammatical.) To observe the use/mention distinction I sometimes use single quotation marks (e.g. "the word 'ought' has five letters"), and at other times use italics ("the word *ought* has five letters"). I often fudge the distinctions between sentences and propositions, implication and entailment, 'it is obligatory that' and 'it ought to be that', etc. I would rather be guilty of some relatively inconsequential sloppiness than of boring the reader to tears with pointless pedantry. I assume that the reader will forgive this laxness and interpret my *modes of expression* (as opposed to my substantive claims and arguments) as charitably as possible.

When a generic pronoun is called for, I use 'she', 'he', 'he or she' or 'they', according to whim (or whatever sounds best to me in the context in question).

[6] Pretentious people will say: 'schemata'.

Chapter 1

Normative conflicts

> A system of norms that is impossible to obey is, of course, 'unreasonable'; and a lawgiver, for instance, that issued such a system of norms could be morally blamed. But to say that such a system exists is not a logical contradiction.
>
> —Erik Stenius [148, p. 254]

> For 'I am under an obligation to do A' and 'I am under an obligation to do not-A' are not *formally* inconsistent; nor [...] can a contradiction be deduced from them without begging the very question at issue; so it is not necessarily *illogical* to retain both, together with the principle(s) from which they come.
>
> —E. J. Lemmon [95, p. 47]

A normative conflict is a situation in which two or more norms conflict (or, perhaps, a norm conflicts with itself). Some *prima facie* examples:

- Val asks her husband Newton, "Do I look fat in this dress?" As a matter of fact, she *does* look fat. Newton ought to tell her the truth, since it is wrong to lie. Yet Newton also ought to lie, since it is wrong to needlessly hurt his wife's feelings. (Note that refusing to answer is tantamount to saying "yes"!)

- Alan promises his boss, Itala, to let her know immediately if a certain client calls. When the client calls, however, Itala is in a very important meeting which she insisted should not be interrupted for any reason.

- Little Bobby's mother told him earlier that he was forbidden to play video games after school. But his father left him a note explicitly *permitting* him to play video games after school.

- As a bus driver, Graham is obliged to drive safely and keep on schedule. But due to severe traffic problems today, he cannot do both.[1]

- High school student Carlos confides in his guidance counselor, Georg, that he is having thoughts of suicide. Georg promised beforehand to keep the conversation private, but he is now seriously concerned that Carlos may act on his thoughts, and feels an obligation to inform Carlos's parents.

It is a testament to the quotidianness of such conundrums that we have all sorts of clichés for describing them: 'between a rock and a hard place', 'between the hammer and the anvil', 'between Scylla and Charybdis', 'Catch-22', 'Sophie's choice', 'damned if you do, damned if you don't', 'you can't win', etc. Normative conflicts are the stuff of drama (and often comedy). Nearly every good work of fiction contains at least one intriguing (or hilarious) normative conflict. They are a big part of what makes life difficult—and interesting.

But are normative conflicts real, or merely apparent? Presumably everyone will concede that norms sometimes *seem* to conflict. More controversial is the claim that in some cases (whether actual or merely possible), norms *really do* conflict, whether we think they do or not. It is this latter claim—the claim that there are *ontological* as opposed to merely *epistemic* (or *doxastic*) normative conflicts—that I wish to defend in this chapter. The upshot will be that since a proper conceptualization (or ontology) of norms should not rule out normative conflicts *a priori*, a proper logic of norms should not either. The establishment of this claim will motivate the development of formal systems that can "tolerate" normative conflicts—a task to which the remainder of this dissertation will largely be dedicated.

1.1 What is a norm?

By 'norm' I mean: a rule or standard by which actions or states of affairs are evaluated (i.e. judged to be right or wrong, better or worse, acceptable or unacceptable, etc.). Some examples:

- Thou shalt not lie.

- Federal taxes must be paid by April 15.

[1] Cf. Hill [77, p. 240].

- Items may be returned within ninety days of purchase.
- Employees must wash their hands after using the restroom.
- It's rude to belch in public.
- If a player's king is in check then the player must make a move that eliminates the threat of capture.
- From A and $A \to B$ one may derive B.
- One should use charity when interpreting another person's argument.
- No dogs allowed.

As these examples indicate, norms come in many varieties. These include morals, laws, regulations, policies, principles of etiquette, rules of games, rules of deductive systems, rules of linguistic interchange, and so on. By 'norm' I mean to include *all* of these. This is important to emphasize, since most of the philosophical discussion of normative conflicts focuses on particular types of norms, such as *moral* norms or *legal* norms. Indeed, little has been written on whether norms *in general* can conflict. And of course, one may hold that certain specific types of norms, e.g. moral norms, cannot conflict, while norms *in general* can. That is, one may hold that certain specific types of normative conflict are impossible, but that other types of normative conflict are possible. (This is, incidentally, a position with which I am quite sympathetic.)

I now want to draw some important distinctions pertaining to norms.

1.1.1 General vs. particular norms

We should distinguish between general norms, such as *If you make a promise, you ought to keep it*, and specific norms, such as *Smith ought to keep the promise that he made to his wife on June 20, 2006 to mow the lawn on June 21, 2006.* The former is a universally quantified norm (or "rule") of the form[2], *Provided X obtains, it is obligatory that Y obtain.* The latter is an *instantiation* of such a general norm—the result of applying it to a specific case. (Strictly speaking, the relevant instantiation of the general norm would be: *If Smith promised his wife that he would mow the lawn on June 21, then he ought to mow the lawn on June 21.* From this and the fact that

[2] Here, obviously, I am using the word 'form' a bit loosely.

he *did* make the relevant promise we can derive that Smith ought to mow the lawn on June 21.)

1.1.2 Compulsory vs. elective norms

Some norms say what *should* (or *must*, or *ought to*) be done, or what *should* (or *must*, or *ought to*) be the case (e.g., *You must remain seated*). Others specify what *may* be done, or what *may* be the case (e.g., *You may be seated*). The first type of norm I will call *compulsory*; the second type I will call *elective*.[3] Compulsory norms are typically expressed with words like *ought*, *must*, *obligatory*, *obliged*, *required*, *mandatory*, *duty*, *forbidden*, *verboten*, *prohibited*, *wrong*, *illegal*, and *taboo*. (Note that, at least *prima facie*, a prohibition against A is simply an obligation to see to it that not-A. By the same token, an obligation to see to it that A is simply a prohibition against not-A.) Elective norms are typically expressed with words like *may*, *can*, *allowed*, *permitted*, *permissible*, *okay*, *legal*, *optional*, *acceptable*, *gratuitous*, *kosher*, and even, in contemporary slang, *cool* (as in "Is it cool to smoke in here?").

A compulsory norm states that the *failure* of a certain act to be performed (or of a certain state of affairs to obtain) is inconsistent with (i.e., in violation of) relevant or applicable normative standards. (For example, *You must remain seated* states that failing to remain seated would violate some relevant norm.) An elective norm states that an act's being performed (or a state of affairs's obtaining) is consistent with (i.e., not in violation of) relevant/applicable normative standards. (For example, *You may be seated* states that sitting down does not violate any relevant norms.)

Note that modal words like *must* and *can* also have non-normative (e.g., metaphysical or ability-related) meanings, as in 'The angles of a triangle must add up to 180 degrees' and 'Penguins can't fly'. In an episode of *The Simpsons*,[4] this ambiguity

[3] Philosophers use various terminology to capture the distinction between what I am calling *compulsory* and *elective* norms. Von Wright [155, p. 71] calls them *obligation-norms* and *permissive norms*, respectively. Munzer [120, p. 1141] calls them *duty-imposing* and *permission-granting*, resp. Hamblin [74, p. 75] calls them *prescriptive* and *permissive*, resp. Hill [77, p. 230] calls them *imperative* and *permissory*, resp. I prefer my terminology because it seems less likely to suggest that obligation is somehow more basic than prohibition, or that permission is somehow more basic than gratuity (= "permission-that-not"). While I *do*, in the present work, take obligation and permission as primitive (defining prohibition and gratuity in terms of these and negation), it ought to be acknowledged that this is merely a *convention* (much like taking 'and' rather than 'or', or 'all' rather than 'some', as primitive in a system of logic). It would be equally legitimate, in my view, to take prohibition and gratuity as primitive.

[4] "Homer the Great," Season 6.

is exploited for nefarious purposes:

MARGE:	Kids can be so cruel.
BART:	We can? Thanks, Mom!
LISA:	Ow! Cut it out, Bart!

As we shall see in the next chapter, the close analogy between metaphysical and normative (deontic) interpretations of *must* and *can* (and their cognates) is exploited in the standard "possible worlds" semantics of deontic logic.

1.1.3 Positive vs. negative norms

Consider the norm, *Smoking is forbidden*. On the one hand, this might be taken to be equivalent to *It is obligatory that one not smoke*, or *One has an obligation to refrain from smoking*. On the other hand, it may be taken to be equivalent to *It is not permitted that one smoke*, or *One does not have permission to smoke*. In the first case, we are talking about the *existence* of a *compulsory* norm. In the second, we are talking about the *non-existence* (i.e., *lack*) of an *elective* norm. *Prima facie*, the lack of a (certain) elective norm is *not* the same thing as the existence of a (certain) compulsory norm. Hence it seems reasonable to distinguish between *positive* and *negative* norms. A positive compulsory norm requires that something be done or that something be the case (e.g. *You must extinguish your cigarette*). A negative compulsory norm is the lack of a permission to do something (e.g. *You are not allowed to smoke here*). A positive elective norm permits something to be done (e.g. *You are permitted to smoke in the designated area*). A negative elective norm is the lack of an obligation to do something (*It is not obligatory that you extinguish your cigarette*). Thus we can speak of positive or negative obligation, prohibition, and permission.[5]

[5] Von Wright [155, p. 72] also distinguishes between what he calls "positive" and "negative" norms; however, his distinction is a different one—and one that is, in my opinion, incoherent (or at least extremely superficial). For Von Wright, a positive norm is one whose content (i.e. the thing that is commanded, prohibited, or permitted) is an *act*, while a negative norm is one whose content is a *forebearance*. I find this distinction incoherent (or at least superficial) because, to me, every act is a forebearance, and vice versa. For example, Von Wright interprets *The door may be left open* as expressing a negative permission—apparently because he interprets it as something like *You are permitted not to close the door*. However, it could just as well be interpreted as a positive permission—the permission to *perform the act* of leaving the door open. (To be fair, Von Wright probably would not conceive of leaving the door open as an act, *per se*, but rather as the forebearance of an act.) At any rate, I do not see *Von Wright's* distinction between positive and negative norms as an important one.

It would be quite reasonable to argue that a so-called "negative norm" is not really a *norm* at all; it is merely the *lack* of a norm. However, I think that we do often conceive of utterances like "You're not allowed to exceed the speed limit" as expressing norms (in *some* sense), and not (or not *merely*) the non-existence of certain norms. *Perhaps* this is evidence that non-permission really is the same thing as obligation-that-not, that non-prohibition really is the same thing as permission, and so on. However, I do not think this is something we should take for granted—first, because it is not entirely obvious from an intuitive point of view; and second, because it has important formal-logical implications that will be brought out in later chapters (especially Chapter 5). With these things in mind, I will continue to speak of negative norms as being a type of norm, even though we may ultimately decide that we do not wish to take this kind of talk literally.

1.1.4 Strong vs. weak permission

Philosophers sometimes speak of the distinction between *strong* (sometimes "positive") permission and *weak* (sometimes "negative") permission.[6] Strong permission is a positive norm to the effect that something is allowed. Weak permission is merely the absence of a positive norm to the effect that something is prohibited. What is puzzling is that philosophers rarely if ever (to my knowledge) speak of the equally coherent and important distinctions between strong and weak *prohibition*, and between strong and weak *obligation*. Strong prohibition is a positive norm to the effect that something is forbidden, while weak prohibition is the absence of a positive norm to the effect that something is allowed. Similarly, strong obligation is a positive norm to the effect that something is required, while weak obligation is the absence of a positive norm to the effect that something is gratuitous (i.e., permitted *not* to obtain). My distinction between positive and negative norms captures the same idea as the common distinction between strong and weak permission; however, it is more general, as it applies to compulsory as well as elective norms.

[6] See, e.g., [155, p. 86], [146], [5, p. 117].

1.1.5 "Ought to do" vs. "ought to be"

Another important distinction is between the "ought to do" (*Tun-Sollen*) and the "ought to be" (*Sein-Sollen*).[7] (Apparently German is more sensitive to this distinction than English.) This distinction is illustrated by the difference between *It is obligatory that you lock the door* and *You are obligated to lock the door*. In the first case we conceive of obligation as a property of a proposition (namely, *that you lock the door*); in the second, as a relation between an agent (you) and a type of act (locking the door). I will assume that any "ought to do" statement can be equivalently expressed as an "ought to be" statement, but not vice versa. For example, 'You ought to help John' is equivalent to 'It ought to be that you help John', but 'It ought to be that someone helps John' is not equivalent to 'There is someone who ought to help John'. Therefore, as a matter of generality as well as convenience, I will deal exclusively with the "ought to be," or *Sein-Sollen*.

1.1.6 Binding vs. non-binding norms

We sometimes refer to certain norms as "binding" to indicate that they are *genuine*, *real*, or *legitimate*, and to contrast them with "non-binding" norms, e.g., those that are promulgated by an agent lacking the proper authority to do so (for example, if I were to stipulate that eating French fries is against the law in the United States). Some might favor the view that "bindingness" is part of the very concept or essence of being a norm, and therefore 'non-binding norm' is, strictly speaking, a contradiction in terms: a so-called "non-binding norm" is no more a norm than imitation vanilla is vanilla, or a paper airplane is an airplane. In my opinion this is just a matter of which linguistic convention we wish to adopt with respect to the word 'norm'. I do not care which convention is adopted, provided it is understood that when *I* claim that normative conflicts are possible, I am claiming that *binding* norms can conflict. (I think almost anyone would concede that a binding norm can conflict with a non-binding norm, or that two or more non-binding norms can conflict with each other. These are not particularly interesting or controversial claims.)

[7] See, e.g., [157, p. 29].

1.1.7 Norms vs. normative propositions

Logic, as we normally conceive of it, deals with sentences, statements, or propositions—bearers of truth values. Logical consequence is normally defined as *truth-preservation*: if the premises of a logically valid argument are all true, then so must be the conclusion. Thus philosophers have long found it puzzling that norms seem to stand in logical relations even though they do not seem capable of being true or false. For example, it is plausible that 'You ought to help John' logically implies 'You ought to help someone'. It seems, then, that we are forced to either concede that there can be no logic of norms, or radically change our notion of what logic is. This is known as *Jørgensen's Dilemma*, after the Danish philosopher Jørgen Jørgensen, who first articulated it.[8]

Some philosophers, such as Von Wright [155, pp. 104-106] and Alchourrón and Bulygin [6], attempt to escape between the horns of this dilemma by distinguishing between *norms* and *normative propositions*. The idea is that to each norm there corresponds a normative proposition that asserts the *existence*, or "*bindingness*," of that norm. The very same sentence may express a norm or the corresponding normative proposition, depending on the context, speaker's intention, etc. For example, 'You must turn in your paper by Thursday' might express a norm when the instructor says it, and a normative proposition when a student says it (say, to another student, as a reminder). Norms cannot be true or false, but their corresponding normative propositions can. Thus, while a logic of norms may require a radical departure from the usual conception of logic, a logic of normative propositions does not.

I, for one, have some trouble making sense of the notion of a *logic of norms* (where 'norm' is understood in the way just specified). (Perhaps I am overly accustomed to thinking of logic as being about truth-preservation.) Thus my concern in the present work is with normative propositions. When I speak of normative conflicts, I am speaking of conflicts between (true) normative propositions rather than conflicts between norms. However, it should be noted that the latter are conceptually prior to the former; you can't have a conflict between (true) normative propositions without a conflict between norms.

In passing I will note that there *is* a promising way to think about a logic of norms, as opposed to merely a logic of norm-propositions. If we think of logical consequence

[8] See [88], [135].

as being essentially about *meaning (or information) containment*[9] rather than truth-preservation, then it seems we have a way of making sense of logical relations between norms (since norms are, presumably, meaningful and information-laden, even if they are not the sorts of things that can be true or false). I will not explore this option here, however, since (for present purposes, at least) I see no urgent need to develop a genuine logic of norms as opposed to a logic of normative propositions.

1.2 What is a normative conflict?

I propose to use 'normative conflict' to mean: *any set of normative propositions the elements of which are in tension or disharmony when considered jointly.* Anyone promulgating—in a single breath, so to speak—the norms corresponding to the elements of such a set would be guilty of at least a *pragmatic*, if not logical, contradiction.[10] (I do not think that our initial definition needs to be any more precise than this.)

Let us use A and B as placeholders for arbitrary statements or propositions. Some general forms of normative conflict are:

1. It is obligatory that A. It is forbidden A.
2. It is forbidden that A. It is permitted that A.
3. It is obligatory that A. It is permitted that not-A.
4. It is forbidden that A. It is obligatory that either A or B. It is forbidden that B.
5. It is forbidden that B. It is obligatory that if A then B. It is obligatory that A.
6. It is permissible that A. It is obligatory that if A then B. It is forbidden that B.

[9] See, e.g., Hanson [76], Brady [30, §2.2] for explications of this idea.

[10] A classic example of a pragmatic contradiction is "Moore's Paradox," named after G. E. Moore. (See "Moore's Paradox" in [119].) Suppose Alice asserts: "Bob is wealthy, but I don't believe it." The quoted sentence is clearly not a *logical* contradiction (for the proposition it expresses—*Bob is wealthy and Alice does not believe that Bob is wealthy*—could very well be true); nevertheless, Alice "contradicts herself" (in some sense) by asserting it.

7. It is obligatory that both A and not-A.

8. It is forbidden that either A or not-A.

(Obviously we could go on listing such conflict types indefinitely.) The set of all such conflict-types can be partitioned into two broad categories:

1.2.1 Escapable vs. inescapable conflicts

Conflicts like (1), (4), (5), (7), and (8) above are *inescapable*, in that, if one of them is instantiated, there is no way to avoid violating some norm or other. Consider (1), the situation in which A (whatever that may be) is both obligatory and forbidden. No matter how things turn out, either A or not-A will obtain (assuming that the law of excluded middle holds). Either way, a norm has been violated.

Conflicts like (2), (3), and (6), in contrast, are *escapable*: even when they are instantiated, it is possible to avoid any violation or "wrongdoing." This is because, in these cases, at least one of the conflicting norms is *elective* in nature. Thus, one can avoid violating any norm by simply refraining from doing what the elective norm allows or licenses. For example, suppose that I am in a situation in which I am both permitted and forbidden to smoke (but not obligated to smoke). I can avoid violating any norm by simply refraining from what I am permitted (but not obligated) to do—namely, smoke. On the other hand, suppose that I am (for some strange reason) both obligated and forbidden to smoke. In that case, I will have violated a norm whether I smoke or not.

Unfortunately, most of the literature on normative conflicts focuses exclusively on the inescapable kind (probably because these make for better Styronesque drama). However, at least a few theorists, notably Munzer [120], Hamblin [74], and Hill [77], have drawn attention to the distinction I am making here. (Hamblin refers to inescapable conflicts as *quandaries*.)

1.2.2 Canonical conflicts

We have looked at a few examples of forms that a normative conflict can take. In principle, there are infinitely many such forms. Fortunately, we can focus our discussion on two basic or "canonical" forms of normative conflict, since every normative conflict is, in a specific sense, reducible to one of these. Let O, F, and P stand for *it* *is obligatory that* (or *it ought to be that*),[11] *it is forbidden that*, and *it is permissible that*, respectively. Let A stand for an arbitrary sentence or proposition. Our canonical *inescapable* conflict is $\{OA, FA\}$, and our canonical *escapable* conflict is $\{FA, PA\}$.

Now, an escapable conflict is precisely one in which some proposition is both forbidden and permissible (or so it seems reasonable to stipulate). But an inescapable conflict is one in which there is some set of propositions the elements of which are individually obligatory but jointly impossible (or, again, so it seems reasonable to stipulate). We need to show that every such situation is one in which some proposition is both obligatory and forbidden.

Call the result of prefixing 'it ought to be that' to a sentence the *oughtification* of that sentence (e.g. the oughtification of 'I leave' is 'It ought to be that I leave', or, less stiltedly, 'I ought to leave'). Call the result of oughtifying all the elements of a set of sentences the "oughtification" of that set. Now consider the following "oughtification" principle:

> If a set of sentences, Γ, entails[12] a sentence, A, then the oughtification of Γ entails the oughtification of A.

This seems plausible. For example, it implies that since *modus ponens* (if A then B; A; therefore, B) is a valid inference, so is the following:

> It ought to be that if A then B. It ought to be that A. Therefore, it ought to be that B.

Now consider any unavoidable conflict. Such a conflict is characterized by the fact that there is some set Γ of propositions such that the elements of Γ are individually obligatory but jointly impossible. Pick any element of Γ and call it A. Let $\Gamma - A$ be the result of removing A from Γ. Since by assumption the elements of Γ cannot all be true, $\Gamma - A$ entails not-A. Thus, by the oughtification principle, not-A is obligatory, i.e. A is forbidden. Since by assumption A is also obligatory, we have a canonical inescapable conflict.

[11] The locutions *it is obligatory that* and *it ought to be that* are not strictly interchangeable in ordinary language. In particular, the former seems to imply the latter but not vice versa. However, for our purposes the two expressions will be treated as equivalent.

[12] For now, when I say that a set of sentences Γ *entails* a sentence A, I just mean that it is necessary that if all of the elements of Γ are true, then so is A.

1.3 The case for normative conflicts

In this section I will set out and develop three arguments for the claim that normative conflicts are logically possible: the *"say-so" argument*, the *argument from examples*, and the *argument from moral residue*.

1.3.1 The "say-so" argument

Here is a version of (what I will call) the "say-so" argument. Imagine a society that is ruled by two ruthless (and perhaps not too intelligent) co-dictators, Tweedledee and Tweedledum. In this society, whatever Tweedledee or Tweedledum explicitly commands is *ipso facto* (legally) obligatory, and whatever one of them explicitly deems okay is *ipso facto* (legally) permissible. In other words, the law is determined by the Tweedle brothers' "say-so." We may assume that 'Dee does not always remember or keep track of what 'Dum commands or permits (and vice versa). Given this, Tweedledee might explicitly command that p even though Tweedledum has explicitly commanded that not-p. In that case, p would be both obligatory and forbidden—an inescapable conflict. Moreover, Tweedledum might explicitly command that not-q even though Tweedledee has explicitly permitted that q. In that case, q would be both forbidden and permitted—an escapable conflict. While admittedly implausible, the society we are imagining is *possible*, as are the imagined behaviors of Tweedledee and Tweedledum in that society. Hence, both escapable and inescapable normative conflicts are possible.

Let us make this argument a little more abstract and explicit, so that it may be more easily assessed. First, let us define a *say-so society* as a society, $\mathcal{S}$, in which for some set of authority agents, $\alpha(\mathcal{S})$, anything explicitly commanded by a member of $a(\mathcal{S})$ is *ipso facto* (legally) obligatory and anything explicitly permitted by a member of $\alpha(\mathcal{S})$ is *ipso facto* (legally) permissible. Here, then, is the "say-so" argument:

1. It is possible that there is a say-so society $\mathcal{S}$ and proposition p such that some element of $\alpha(\mathcal{S})$ (= the lawmaking authorities of $\mathcal{S}$) explicitly commands p and some element of $\alpha(\mathcal{S})$ explicitly commands not-p. (assumption)

2. It is possible that for some proposition p, p is both obligatory and forbidden. (from 1, def. of say-so society)

3. Inescapable normative conflicts are possible. (from 2)

4. It is possible that there is a say-so society $\mathcal{S}$ and proposition p such that some element of $\alpha(\mathcal{S})$ explicitly commands not-p and some element of $\alpha(\mathcal{S})$ explicitly permits p. (assumption)

5. It is possible that for some proposition p, p is both forbidden and permitted. (from 4, def. of say-so society)

6. Escapable normative conflicts are possible. (from 5)

7. Both escapable and inescapable normative conflicts are possible. (from 3, 6)

All of the inferences in this argument are valid.[13] Thus, anyone wishing to deny the conclusion must deny one of the assumptions (1 or 4). Now, it seems obvious that *if* a say-so society is possible, then an authority agent of that society could command p while another commands not-p. By the same token, it seems obvious that an authority figure of that society could command not-q while another permits q. I conclude that the only reasonable way to challenge the argument above is to deny the possibility of a say-so society.

How might one argue against the possibility of a say-so society? Here is one possible argument:

> *A society controlled by a group with no limits on its power would be a totalitarian dictatorship or oligarchy. Hence, it would have no legitimacy. Hence, the "laws" it promulgated would not be binding—i.e., would not be* laws *at all. Such a society would incapable of generating legal obligations or permissions, and thus could not satisfy the definition of a say-so society.*

The problem with this argument is that it confuses moral obligation with legal obligation. Just as something may be morally obligatory without being legally obligatory, something may be legally obligatory without being morally obligatory. Indeed, something may be legally obligatory while being morally *forbidden*. There is an obvious sense in which, e.g., the legal system of Nazi Germany was illegitimate—it was inherently immoral. However, this does not change the fact that certain things were

[13] These inferences appeal to the widely-accepted principle if A implies B, then ⌜it is possible that A⌝ entails ⌜it is possible that B⌝. The somewhat awkward format of the argument above is an attempt to forestall the potential objection that I am committing the fallacy of thinking that because A and B jointly entail C, ⌜it is possible that A⌝ and ⌜it is possible that B⌝ jointly entail ⌜It is possible that C⌝.

truly and genuinely legal (i.e., legally permitted) and illegal (i.e., legally forbidden) in Nazi Germany. Indeed, we criticize Nazi Germany on precisely the grounds that it made certain immoral actions legal and certain moral actions illegal. The point is that *bad laws are still laws.* The 19th Century legal theorist John Austin made this point well:

> The existence of law is one thing; its merit or demerit is another. Whether it be or be not is one enquiry; whether it be or be not conformable to an assumed standard, is a different enquiry. A law, which actually exists, is a law, though we happen to dislike it, or though it vary from the text, by which we regulate our approbation and disapprobation. This truth, when formally announced as an abstract proposition, is so simple and glaring that it seems idle to insist upon it. But simple and glaring as it is, when enunciated in abstract expressions the enumeration of the instances in which it has been forgotten would fill a volume. [14, p. 184]

I conclude that the "argument from illegitimacy" (as we might call it) fails.

Are there any *other* arguments against the possibility of a say-so society? Perhaps I lack imagination, but I can only think of one, which I don't find very convincing:

> *If a say-so society were possible, then normative conflicts would be possible. But normative conflicts are not possible. Hence, there could not be a say-so society.*

We seem to have reached a stalemate. I say: "A say-so society is possible; hence, normative conflicts are possible." My opponent says: "Normative conflicts are impossible; hence, a say-so society is impossible." (As the old adage goes, one person's *modus ponens* is another's *modus tollens.*)

Can this stalemate be broken? Perhaps. So far, the debate has been very abstract; the opponent of normative conflicts has not been given any *concrete, plausible examples* of normative conflicts. Providing some such examples may just sway him or her to the other side of the fence.

1.3.2 The argument from examples

The general form of the argument from examples is this:

1. There is at least one example of a possible scenario in which a normative conflict obtains.

2. Therefore, normative conflicts are possible (i.e., it is possible for there to be at least one normative conflict).

(Obviously this could be broken into two more specific arguments, one for the possibility of *escapable* conflicts and one for the possibility of *inescapable* conflicts.) The argument is clearly valid; the only trick is providing convincing support for the premise. My approach will be to provide a plethora of examples, in the hope that my opponent will find at least one of them at least moderately convincing (that is, convincing as a *possible*, though not necessarily *plausible*, scenario). Some of my examples are real; others are made up. Some are original; others are borrowed or adapted. I'll start with some examples of *inescapable* conflicts.

Examples of inescapable conflicts

- A friend has lent me his car, on the condition that I return it on demand. He shows up at my doorstep, roaring drunk, and demands his keys. On the one hand, I ought to give him the keys, as I promised (sans qualification) to return them on demand. On the other hand, I ought *not* to give him the keys, as I would be facilitating the dangerous and potentially deadly crime of driving while highly intoxicated. (Cf. Lemmon [94, p. 105])

- I contract with Alice to be at a certain location at a certain time. I contract with Bob *not* to be at that location at that time. (Perhaps I am engaging in some kind of skullduggery; on the other hand, I may just be absent-minded.) Hence, I have an obligation both to be and not to be at said location at said time. (Cf. Priest [129, p. 182])

- My son, a wanted fugitive, asks me to take him in. On the one hand, I ought to say yes, since parents have an obligation to take care of their children. On the other hand, I ought to say no, since it is against the law to harbor a fugitive. (Cf. Garson [58, p. 48])

- I work for an organization that has the following policies in effect: (i) any document containing secret information must be kept in a safe when it is not being used; (ii) any document that has not been used in more than five years is to be thrown away. Now consider a document that contains secret information and has not been used in more than five years. It is both obligatory and forbidden that this document be kept in the safe. (Cf. Cholvy and Cuppens [44, p. 247])

- The civil code of Louisiana contained, at one time, the following two principles:[14] (i) minors must obtain the consent of their parents to marry; (ii) a marriage cannot be annulled on the ground that it was contracted without the consent of the parents. Suppose that minors Claire and Dave were married without the consent of their parents. According to Louisiana law, it is both obligatory and forbidden that their marriage be annulled.

- As Commander-in-Chief of the U.S. military, President Bush ought to attend at least some of the funerals of soldiers killed in Iraq, to show his respect and appreciation for the sacrifices they have made. On the other hand, he ought not to attend any such funerals, since he cannot realistically attend *all* of them, and it would be unfair to the relatives of those soldiers whose funerals he inevitably wouldn't attend.

Here is a more detailed, real-life example of an inescapable conflict:

The *Riggs v. Palmer* case[15] has been widely discussed among legal scholars.[16] This is the case of Elmer, a young man who has murdered his grandfather with the aim of inheriting the old man's considerable fortune. The judges of the New York Court of Appeals—in effect, the "supreme court" of New York State—are in agreement that Elmer must serve a long jail sentence for his crime. But should he receive his inheritance? Judge Earl, in the majority opinion, argues that he should not:

> What could be more unreasonable than to suppose that it was the legislative intention in the general laws passed for the orderly, peaceable, and just devolution of property that they should have operation in favor of one who murdered his ancestor that he might speedily come into the possession of his estate? Such an intention is inconveivable. . . . Besides, all laws, as well as all contracts, may be controlled in their operation and

[14] See Alchourrón [4, p. 338].

[15] *Riggs v. Palmer*, 115 N.Y. 506, 22 N.E. 188 (1889).

[16] See e.g. Dworkin [53].

> effect by general, fundamental maxims of the common law. No one shall be permitted to profit by his own fraud, or to take advantage of his own wrong, or to found any claim upon his own iniquity, or to acquire property by his own crime. [2, p. 178]

Judge Gray, in the dissenting opinion, argues that Elmer is entitled to his inheritance:

> The question we are dealing with is whether a testamentary disposition can be altered, or a will revoked, after the testator's death, through an appeal to the courts, when the legislature has by its own enactments prescribed exactly when and how wills may be made, altered, and revoked, and apparently, as it seems to me, when they have been fully complied with, has left no room for the exercise of an equitable jurisdiction by courts over such matters. ... [T]o concede the appellants' views would involve the imposition of an additional punishment or penalty upon the respondent. ... The law has punished him for his crime, and we may not say that it was an insufficient punishment. [2, pp. 179-80]

Here we have a sophisticated, rich, real-life case in which two well-established and extremely compelling legal principles clash. One way of interpreting the disagreement between the judges is to assume that there is a "right" answer somehow already encoded in the law, and that the judges are arguing about what that right answer is. I am inclined to interpret the disagreement differently. The law, as it currently stands, is simply *inconsistent* on this matter.[17] What the judges are arguing about is: (a) what should be done about the case at hand, given that Elmer cannot both receive and not receive his inheritance; but, perhaps more importantly, (b) what new precedent should be set, i.e., *how the current law should be revised* so as to deal justly and consistently with such cases in the future. In short, I believe that the judges are faced with a genuine *ontological* normative conflict—one that shows that the law, as it currently stands, is flawed and in need of modification.

Examples of escapable conflicts

- An amusement park has the following policies in effect. First, anyone eighteen or older is permitted to ride a certain roller coaster. Second, anyone under

[17] I am using 'inconsistent' in a colloquial sense here, not the logician's technical sense.

four feet tall is forbidden to ride that coaster. Billy is nineteen years old, and forty-seven inches tall. By the amusement park's rules, Billy is both permitted and forbidden to ride the roller coaster.

- A government passes laws to the following effect. First, anyone convicted of a felony is forbidden to vote. Second, any elected politician is permitted to vote. I am an elected politician and a convicted felon. Hence, I am both forbidden and permitted to vote.

- Mary's credit card company has the following policies. First, anyone with poor credit (say, with a credit score below a certain level) is required to pay off her full balance each month. Second, anyone with a Gold Card is permitted to refrain from paying off her full balance each month. Mary is a Gold Card holder with bad credit. Hence, Mary is required to pay off my full balance this month, though she is also permitted to refrain from doing so.

- An office has the following policies: (i) on Fridays, workers are permitted to dress casually (this entails men not having to wear ties); (ii) on days when important clients visit the office, workers are required to dress formally (this entails men wearing ties). Suppose important clients visit on a "casual Friday," and that I am a male worker at the office. According to office rules, I am both forbidden and permitted to refrain from wearing a tie.

- I specify a formal logic with the following derivation rules (among others): (i) it is permissible to infer $A \vee B$ from A; (ii) it is permissible to infer B from $A \vee B$ and $\neg A$; (iii) it is permissible to infer A from the set Γ if A can be derived from Γ via a series of permissible steps; (iv) it is forbidden to infer A from Γ if A does not share an atomic formula with some element of Γ. Within this system, it is both permissible and forbidden to infer (e.g.) q from $\{p, \neg p\}$.[18]

Here is a slightly more detailed, real-life example of an escapable conflict:

In the U.S. Supreme Court case *NLRB v. Bildisco and Bildisco, Inc.*,[19] the following situation was enountered. According to the National Labor Relations Act, the management of a company must bargain in good faith with the representatives of a certified union to which its workers belong. According to the Federal Bankruptcy Act,

[18] If you're not sure why this is the case, read Chapter 3!

[19] U.S., 104 S. CT. 1188, 79 L. Ed. 2d 482 (1983). Cited in Hill [77, p. 241].

insolvent debtors are permitted to avoid executory contracts. Bildisco & Bildisco, Inc. was an insolvent company whose workers belonged to a certified union. Thus, when the company cancelled an executory union labor contract, it did something that was both legal (i.e. legally permitted) and illegal (i.e. legally forbidden). As with *Riggs v. Palmer*, the court settled the issue *not* by disovering that there wasn't really a conflict in the law, but rather by setting a *new* legal precedent (and thus *changing* the law) so that similar problems would (hopefully) not arise in the future.

I hope that at least some of the above examples have "stuck," so to speak—or at least given my hypothetical opponent second thoughts about refusing to recognize the possibility of normative conflicts. If not, I have one last trick up my sleeve. . .

1.3.3 The argument from moral residue

The argument from "moral residue"[20] purports to show that *moral dilemmas*—i.e., inescapable conflicts of moral obligation—are possible (and, indeed, common). The argument can be formulated as follows. The first premise is that there are situations in which: (a) one cannot do both X and Y; (b) if one does X, it will be appropriate to feel regret/remorse/guilt[21] for not having done Y; and (c) if one does Y, it will be appropriate to feel regret/remorse/guilt for not having done X. The second premise is that it is appropriate to feel regret/remorse/guilt *only* when one has done something wrong. From these premises it follows that there are situations in which one will have done wrong (i.e., violated a moral obligation) no matter what one does.

I, for one, don't find this argument particularly compelling. (I find both of its premises somewhat dubious.) In fact, as I have mentioned before (see p. 8), I am quite sympathetic to the view that moral dilemmas (i.e., inescapable conflicts of moral obligation) are impossible. I mention the argument from moral residue only because some people *do* find it compelling (and if the reader happens to be in this class, so much the better for me!).[22]

---

[20] An argument along these lines is advanced by, e.g., Bernard Williams [158, §5] and Ruth Barcan Marcus [102, §II]. As far as I know, Marcus was the first to use the term 'residue' in this context: "to insist that there is in every case a solution without residue is false to the moral facts" [102, p. 198 (Gowans)].

[21] One could, of course, draw some subtle distinctions between regret, remorse, and guilt; but I don't think such conceptual fine-tuning is necessary in the present context.

[22] A nice discussion of the argument from moral residue can be found in McConnell [107] [108].

1.4 The case against normative conflicts

I have presented three arguments *for* the possibility of normative conflicts, in the hope that the reader will find them at least somewhat convincing. I now want to consider and rebut a number of arguments *against* the possibility of normative conflicts. Many of the arguments are quite similar; however, I have tried to tease out as many subtly distinct arguments as possible, with the aim of being careful and thorough. (Note: In the subsections titled 'Argument' below I assume the point of view of the hypothetical person who is arguing against the possibility of normative conflicts.)

1.4.1 The argument from interdefinability

Argument

Suppose there were a situation in which something is both forbidden and permitted (i.e., an escapable conflict). *Forbidden* means *not permitted*. Thus we would have a situation in which something is both permitted and not permitted, which is clearly impossible (a flat-out contradiction). Therefore, *escapable* conflicts, at least, are impossible.

Reply

There are two points to be made:

First, as I have said before, it is not clear that 'forbidden' and 'not permitted' mean the same thing: *It is forbidden that A* seems to say that a certain norm (namely, the prohibition of A) exists (or is binding), while *It is not permitted that A* seems to say that a certain norm (namely, the permission of A) *does not* exist (or *is not* binding). It is not at all obvious that the existence of a prohibitive norm is the same thing as the non-existence of a permissive norm with the same content. (By the *content* of a norm I mean: the proposition that is permitted, forbidden, etc.)

Second, even if the argument is sound, it only shows that *escapable* conflicts are impossible. It does not show that normative conflicts *in general* are impossible. My main claim in this chapter is that normative conflicts *of some kind* are possible. While I do in fact hold that both escapable and inescapable conflcits are possible, this claim is ultimately inessential to my main point.

1.4.2 The argument from "ought implies may"

Argument

Suppose there were a situation in which some proposition, say A, is both obligatory and forbidden (i.e., an inescapable conflict). Whatever is obligatory is permissible (*ought* implies *may*). Therefore, in this situation it is permissible that A. Therefore, in this situation it is both forbidden and permissible that A. Following the logic of the argument from interdefinability above, we arrive at a contradiction: A is both permitted and not permitted. Hence, inescapable conflicts (too) are impossible. Hence normative conflicts *in general* are impossible.

Reply

First, this argument appeals to the logic of the argument from interdefinability, which I have already called into question above.

Second, the principle that *ought* implies *may* is far from obvious (even if it is usually taken to be). Indeed, on certain assumptions (which are plausible, but which I do not necessarily accept), it is *equivalent to* the assumption that inescapable conflicts are impossible. In particular, it is equivalent if we assume that (a) *permissible* means the same as *not forbidden* (an assumption that the proponent of the argument above presumably accepts), and (b) $\ulcorner A$ implies not-$B\urcorner$ means the same as $\ulcorner$it is not the case that both A and $B\urcorner$ (an assumption that is built into classical logic). Suppose that $\ulcorner$it ought to be that $A\urcorner$ implies $\ulcorner$it is permitted that $A\urcorner$. By assumption (a), this is equivalent to the claim that $\ulcorner$it ought to be that $A\urcorner$ implies $\ulcorner$it is not forbidden that $A\urcorner$. By assumption (b), this is equivalent to the claim that it is not the case that both (i) it is obligatory that A and (ii) it is forbidden that A. So to assume that 'ought' implies 'may' *without an independent argument* is to simply beg the question against the possibility of inescapable normative conflicts. And indeed, the principle that 'ought' implies 'may' is rarely if ever supported with any independent argument; it is usually taken to be "obvious" (though in my view it is anything but).

1.4.3 The argument from aggregation/"ought implies can"

Argument

Suppose there were a situation in which both A and not-A are obligatory (i.e., an inescapable conflict). The *principle of aggregation* (for *ought*) states that if multiple propositions are *individually* obligatory, then they are *jointly* obligatory: that is, if OB and OC, then $O(B$ and $C)$. Since this is a correct principle, we have a situation in which $\ulcorner A$ and not-$A\urcorner$ is obligatory. Now, as Kant pointed out,[23] *ought* implies *can*: it cannot be obligatory to do the impossible. In other words, ⌜It is obligatory that $A\urcorner$ implies ⌜It is possible that $A\urcorner$.[24] Thus, we have a situation in which it is possible that both A and not-A (are true). But this is absurd: contradictions are *never* possible. (It is not even *possible* that they are possible.) Hence, inescapable conflicts are impossible.

Reply

This argument relies on two principles which have been called into question on *independent* grounds.

First, consider the principle of *aggregation* (sometimes called *agglomeration*). I, for one, am strongly inclined to accept this principle, so it would be disingenuous of me to argue against it. However, it may be worthwhile to quote Bernard Williams at some length here:

> Now there are certainly many characterisations of actions in the general field of evaluation for which agglomeration does not hold, and for which what holds of each action separately does not hold for both taken together: thus it may be *desirable*, or *advisable*, or *sensible*, or *prudent*, to do a, and again desirable or advisable, etc., to do b, but not desirable, etc., to do

[23] Actually, to my knowledge, Kant never says anything *quite* so blunt as "*ought* implies *can*." He does, however, strongly suggest the principle in numerous passages. (See Stern [149, §IV] for a collection of such passages.). For example, in *Religion Within the Boundaries of Mere Reason*, Kant writes that "duty commands nothing but what we can do" ([90, p. 92]; cited in Stern [149, p. 54]).

[24] Here we are glossing over the fact that the intended notion of *can* in "ought implies can" is likely to be a more specific type of possibility (e.g. *personal ability*) rather than bare *metaphysical* possibility. For present purposes this doesn't matter. For *whatever* more specific notion of possibility is relevant here, it surely *entails* bare metaphysical possibility. Therefore, if I can provide justification for doubting the principle that obligation implies *metaphysical* possibility, I have given justification for doubting that it implies whatever more specialized notion of possibility is actually relevant here.

> both *a* and *b*. The same holds, obviously enough, for what a man wants; thus marrying Susan and marrying Joan may be things each of which Tom wants to do, but he certainly does not want to do both. Now the mere existence of such cases is obviously not enough to persuade anyone to give up agglomeration for *ought*, since he might reasonably argue that *ought* is different in this respect; though it is worth noting that anyone who is disposed to say that the sorts of characterisations of actions that I just mentioned are evaluative *because they entail 'ought'-statements* will be under some pressure to reconsider the agglomerative properties of *ought*. I do not want to claim, however, that I have some knock-down disproof of the agglomeration principle; I want to claim only that it is not a self-evident datum of the logic of *ought*, and that if a more realistic picture of moral thought emerges from abandoning it, we should have no qualms abandoning it. [158, p. 132 (Gowans)]

I agree with Williams that the considerations he brings to bear do not obviously *refute* the aggregation principle; it *does*, however, seem to me that Williams has given substantial justification for eyeing the principle with at least mild suspicion.

Let us now turn to the (in)famous principle that *ought* implies *can* (sometimes referred to as "Kant's Principle" or "Kant's Rule"). This principle seems to have a problem similar to that of "ought implies may": namely, it seems to *beg the question* against the possibility of inescapable normative conflicts. Arguably, to say that *ought* implies *can* is to simply rule out (inescapable) normative conflicts from the get-go. However, the principle cannot be dismissed so easily; for unlike "ought implies may," "ought implies can" is not (at least in any obvious sense) *equivalent* to the claim that there are, or can be, no normative conflicts. Thus it can legitimately be used as a premise in a non-circular argument against the possibility of normative conflicts. Moreover, significant and formidable *independent* arguments have been given in support of the principle (whereas to my knowledge they have not in the case of "ought implies may"). The two arguments for "ought implies can" that I find most plausible and worthy of consideration are what I will call the *argument from blame* and the *argument from pointlessness*.[25]

One way of formulating the argument from blame is as follows:

[25] These arguments (among others) are discussed by Stern [149, §III]. What I am calling the argument from *pointlessness* Stern calls the argument from *anti-utopianism*.

> When someone does something that is *wrong*, we *blame* them for doing that thing. When someone does something that they *couldn't help*, we *do not* (for that very reason) blame them for doing that thing. Therefore, it cannot be wrong to do something that one cannot help doing. Now suppose, contrary to what we wish to show, that there could be a situation in which it is obligatory to do X, and it is obligatory to do Y, but it is impossible to do both X and Y. To put things another way: in the situation we are hypothesizing it is wrong to do not-X (i.e. refrain from doing X), and it is wrong to do not-Y (i.e. refrain from doing Y), but it is necessary that one do either not-X or not-Y. In this situation it would be wrong to do either not-X or not-Y, even though one *cannot help* but do either not-X or not-Y. But we have already pointed out that this is impossible. Hence there could not be a situation in which X and Y are both obligatory, but cannot both be done. An inescapable normative conflict is just such a situation. Therefore, there could not be an inescapable normative conflict.

Robert Stern attacks this argument on the ground that the initial claim is false: we *do not* always or necessarily hold people blameworthy when they perform an act that is wrong. As he puts it, "there is a distinction to be drawn between *agent* evaluation and *act* evaluation. That is, I can say that you are not to be criticized for doing or believing A because you were unable to do otherwise, while still holding that what you did or believed was wrong" [149, p. 47].[26]

Stern's criticism seems reasonable to me. However, I want to challenge the argument on a different point, namely that even if it succeeds, it establishes at most that inescapable *moral* conflicts are impossible. For blame is (I submit) an essentially *moral* concept. Perhaps it is true that when someone does something *morally* wrong, we always hold (or should hold) them morally blameworthy. But so what? It does not follow that when someone violates an obligation of *any* kind, we necessarily hold (or should hold) them blameworthy in any sense. Indeed, there are many real-life cases in which we do not hold agents blameworthy for violating obligations. For example, as drivers we have a (legal/non-moral) obligation not to run red lights. But we would not hold someone blameworthy for violating this obligation *if* they did so in order

[26] Walter Sinnott-Armstrong makes the same point in a slightly different way: "It is possible that an act ought to be done even though the agent would not be blameworthy for failing to do it" [144, p. 250].

to (say) save someone's life. Indeed, we would likely *praise* them for violating this obligation. (Anyone who would rather let an innocent person die than risk getting a traffic ticket is surely a reprobate!)

Let us turn, then, to the argument from pointlessness, which runs as follows:

> The function of a normative system is to prove a set of guidelines for how to behave, for how things ought to be—an ideal to which people should strive to conform. If a normative system is to fulfill this function, it must provide guidelines that are *possible* to follow, an ideal to which it is *possible* to conform. Otherwise the normative system is dysfunctional and, worse, *pointless*. A normative system that allowed for the possibility of normative conflicts—at least of the inescapable variety—could not fulfill its function, since it would be impossible for an agent to adhere to it.

This argument, unlike the previous one, does not suffer from the defect of being limited to moral norms. It makes a claim about *all* normative systems: moral systems, legal systems, systems of etiquette, etc. However, the argument is flawed on at least three other grounds:

First, as Stern [149, p. 50] points out, a normative standard (or ideal) to which it is impossible to conform would not necessarily be pointless, or even dysfunctional. A normative standard can serve as a source of *inspiration*, *purpose*, or *direction* even if it is quite literally impossible to satisfy. For example, it is (for all *practical* purposes, at least) impossible for a driver to obey all speed limits at all times. (I would be willing to bet that there is virtually no one who drives on a regular basis and has *never* broken a speed limit.) Nevertheless, speed limits serve an important purpose: they give us an ideal *toward* which to strive, even if it is unreasonable to expect that we will always live up to the standard that they set. Moral rules are the same way: they are often unrealistic, but are nevertheless valuable in that they at least *orient* us in the right direction.

Second, from the fact that a normative system is capable of generating normative conflicts, it does not follow that it is impossible to adhere perfectly to that normative system. Ruth Barcan Marcus makes this point well in her influential paper "Moral Dilemmas and Consistency" [102]. She ask us to imagine a "silly two-person card game" that works as follows. The deck is split between the two players. A round consists of each player drawing a card from the top of his stack. There are two rules:

black cards trump red cards, and higher-valued cards (aces high) trump lower-valued cards. The person winning the most rounds when all cards have been drawn is the winner. Now, it is obvious that this game *can* generate conflicts.[27] Suppose that one player draws a red ace and the other draws a black deuce. Then, according to the rules, each player's card trumps the other—a blatant (and apparently inescapable) conflict. On the other hand, it is quite possible for an entire game (or series of games) to be played without encountering any such conflict. Systems of morality, law, etc. are the same way. From the fact that a normative conflict—even an inescapable one—*can* arise under a system, it does not follow that it is impossible to adhere perfectly to that system.

Third, and (in my view) most importantly, even if it is true that a normative system that is impossible to follow is pointless and/or unable to serve its intended function, it simply does not follow that such a system is incapable of generating real, binding obligations. To argue otherwise would be akin to saying that a broken-down car is not really a car because it cannot (at least in its current state) fulfill its intended purpose. At most, the argument from pointlessness establishes that normative systems *ought* not to require people to do the impossible—that is, that 'ought' *ought* to imply 'can'. This *prescriptive* claim is unobjectionable, but it is a far cry from the much less plausible *descriptive* claim that 'ought' *does*, as a matter of fact, imply 'can'.

Of course, I haven't considered *every* argument that has been (or could be) given for the principle that *ought* implies *can*, as doing so would be beyond the scope of the present work. I have only considered the two arguments that seem most plausible to me. However, I hope that I have given some reasonable justification for thinking that Kant's principle should, at the very least, not be regarded as self-evident or non-negotiable.

For what it's worth, here is some independent evidence *against* the "ought implies can" principle. Consider the following passage from Jules Verne's *Around the World in Eighty Days* [153, p. 199]:

> "Sir," said Mr. Fogg to the captain, "three passengers have disappeared."

---

[27] I would argue that these are *normative* conflicts, since rules of games are norms. I have found, however, that not everyone shares this intuition. Fortunately, for present purposes it does not matter whether the conflicts in question are normative in nature.

> "Dead?" asked the captain.
>
> "Dead or prisoners; that is the uncertainty which must be ended. Do you propose to pursue the Sioux?"
>
> "That's a serious thing to do, sir," returned the captain. "These Indians may retreat beyond the Arkansas, and I cannot leave the fort unprotected."
>
> "The lives of three men are at stake, sir."
>
> "Doubtless; but can I risk the lives of fifty men to save three?"
>
> "I don't know whether you can, sir; but you ought to do so."

Clearly Phileas Fogg, for one, does not believe that 'ought' implies 'can'. Moreover, as one who has read the novel may recall, Mr. Fogg is an extremely logical fellow. Now, it is not clear in this passage whether the 'can' in question is the 'can' of *personal ability* or of *permissibility*; but either way the passage calls into question a principle that has been invoked in arguments against the possibility of normative conflicts (either "ought implies can" or "ought implies may").

1.4.4 The argument from explosion/deontic inheritance

Argument

According to the principle of *explosion* (more traditionally referred to as *ex contradictione sequitur quodlibet*[28]), A and not-A jointly entail B, where B may be any proposition whatsoever.[29] The principle of *deontic inheritance* says that obligations are closed under entailment—that if A entails B, then OA entails OB. More generally if A is a logical consequence of a set of formulas Γ, then the "oughtification" of A (i.e., $\ulcorner$It ought to be that $A\urcorner$) is a logical consequence of the oughtification of Γ (i.e. the result of prefixing each element of Γ with 'it ought to be that'). Combining the principle of explosion with the principle of deontic inheritance yields the principle of *deontic explosion*: if A and not-A are both obligatory, then B is obligatory, where B may be any proposition whatsoever. Thus, if there were an inescapable normative conflict, then *everything* would be obligatory (as well as forbidden!). For example, if

[28] "From a contradiction, anything follows."

[29] The inference from A and not-A to B is *truth-preserving*, for, since it is impossible for A and not-A to be (jointly) true, it is *a fortiori* impossible for A and not-A to be jointly true *while B is false.*

it were both obligatory and forbidden that I brush my teeth, then it would be obligatory that the Earth be destroyed, forbidden that I breathe, etc. Since inescapable normative conflicts would have this absurd consequence, they are impossible.

Reply

This argument, like the last one, relies on two assumptions that can be, and have been, challenged on independent grounds.

Let us first consider the principle of explosion. Quite understandably, many people have found the notion that, e.g., *I am tall and not tall* entails *Pigs fly* to be extremely counterintuitive. The standard response, of course, is that the inference is *valid* (since it is impossible for the premise to be true and the conclusion false) but never *sound* (since the premise cannot be true).

Not everyone is satisfied with this response, however. Generally, there are two different reasons for this dissatisfaction. First, one may have a strong intuition that the premises of a valid argument must *have something to do with*—i.e., be *relevant to*—the conclusion. (What does my stature have to do with the aerial locomotion of swine?) Thus, while truth-preservation may be *necessary* for logical validity, it is not *sufficient*: a further condition of *relevance* must be satisifed. Second, one may accept the identification of logical consequence with truth-preservation but reject the assumption that a sentence and its negation cannot be jointly true—or, at least, the assumption that logic should not *take into account* situations or scenarios—perhaps impossible ones—in which A and not-A are both true.

I do not wish to pursue this matter any further at this point. The issues just raised will be investigated in much more detail in later chapters (especially Chapter 3). For now it will be sufficient to note that the principle of explosion is *highly controversial*, and not something that everyone is simply going to accept without question or protest.

The second assumption on which the above argument relies is the principle of *deontic inheritance*. Interestingly (and, in the present context, somewhat ironically), this assumption has been challenged on the ground that *certain types of normative conflict can arise if it is accepted.* This is easily seen with the so-called "Gentle Murder" paradox first noted by James Forrester [56]. Consider the following statements:

1. Alice ought not to murder Bob.

2. If Alice *does* murder Bob, she ought to murder him gently.

These seem reasonable. At least, we can imagine a normative system that would include principles that would imply both of these statements. Now, suppose that Alice decides to violate the first norm and murder Bob anyway. Then according to the second (conditional) norm, she ought to (at least) murder him *gently*. Note, however, that *Alice murders Bob gently* entails *Alice murders Bob*. Hence, by the principle of deontic inheritance, Alice ought to murder Bob! Hence Alice both *ought* and *ought not* to murder Bob—an inescapable normative conflict. Some philosophers (e.g. Forrester [56] and Goble [60]) have taken this argument to show that the principle of deontic inheritance should be abandoned.[30] Of course, it would be rather hypocritical of *me* to take this view. I cannot (and would not want to) argue that a principle must be wrong *because* it entails that normative conflicts are possible—i.e., because it supports my thesis! Indeed, I am inclined to simply accept the conclusion (i.e. that Alice both ought and ought not to murder Bob) as another example of a possible normative conflict (though I would argue that a normative system entailing (1) and (2) is deeply flawed and in dire need of revision). Nevertheless, the considerations just raised provide another source of misgiving about the argument from explosion/deontic inheritance.

1.4.5 The argument from *prima facie* obligation

Argument

In *The Right and the Good* [136], W. D. Ross famously distinguishes between what he calls *prima facie* ("at first appearance") duties and duties *sans phrase* ("without qualification"). He characterizes the distinction as follows:

> When I am in a situation, as perhaps I always am, in which more than one [*prima facie* duty] is incumbent upon me, what I have to do is study the situation as fully as I can until I form a considered opinion (it is never more) that in the circumstances one of them is more incumbent than any other; then I am bound to think that to do this *prima facie* duty is my duty *sans phrase* in the situation ... I suggest '*prima facie* duty' or

[30] Goble [60, p. 226] suggests that it should be replaced with the principle that if A entails B, then OA entails $O(A \text{ and } B)$. As he notes, this requires giving up the widely-accepted principle that $O(A \text{ and } B)$ entails OA and OB. (To me, this is far too high a price.)

> 'conditional duty' as a brief way of referring to the characteristic (quite distinct from that of being a duty proper) 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. [137, p. 86]

Ross is speaking of *duty* rather than *obligation*, but for our purposes there is no difference. In a more recent work, David Brink characterizes the distinction between prima facie and "all-things-considered" (or *sans phrase*) obligations as follows:

> A prima facie obligation to do x means that there is a moral reason to do x or that x possesses a right-making characteristic. But prima facie obligations can be, and often are, defeated by other, weightier obligations, individually or in concert. A prima facie obligation to do x that is superior to all others constitutes an all-things-considered obligation to do x. An all-things-considered obligation to do x means that on balance, or in view of all morally relevant factors, x is what one ought to do or that x is supported by the strongest moral reasons. [32, p. 103]

Brink is of course talking about *moral* obligations here, but his point extends quite naturally to all types of obligation.

Now that we have the crucial distinction between *prima facie* and *all-things-considered* obligations at our disposal, it is easy to see why normative conflicts are impossible. *All* norms are prima facie, at least until they come to bear in a particular situation. In a particular situation there may be many prima facie norms in play, but the *weightiest* or *most incumbent* of these norms triumphs; *only it* is an all-things-considered obligation in that situation.

Now, we have been speaking of prima facie and all-things-considered *obligations*, but it makes equal sense to speak of prima facie and all-things-considered *permissions*. When, for example, a prima facie permission conflicts with a prima facie prohibition, the weightier norm (be it prohibitive or permissive) triumphs. As a consequence, both inescapable *and* escapable normative conflicts are impossible.

Reply

The appeal to prima facie obligation contains an important germ of truth: when faced with a conflict of obligation (or *any* kind of normative conflict), one should, of course, consider and weigh all relevant factors in order to make a sensible decision as to what, overall, is the best or most proper thing to do. This is an obvious point, which I would never dream of denying. However, the argument from prima facie obligation fails to establish what it purports to establish, namely, that normative conflicts are impossible.

One problem with the argument can be set in the form of a dilemma: so-called prima facie obligations are either (real, genuine) obligations or they are not. If they are, then the argument concedes that conflicts of obligation are possible—a result with which I am quite happy. So we must for the moment go with the assumption that prima facie obligations are not really obligations—any more than paper airplanes are airplanes, or fool's gold is gold. But, as John Searle [141] has noted, this view has absurd consequences. Suppose, for example, that I promise to do x. On the account we are considering, this promise creates a *prima facie* obligation, but not necessarily a *real* obligation, to do x. Now suppose that this prima facie obligation is overridden by another prima facie obligation (say, the obligation to do y), so that when it comes down to deciding whether to do x or y (I can't do both, we're assuming), I have an all-things-considered obligation to do y but not x. Then it turns out that my promise to do x counted for nothing. As Searle observes, "It is exactly as if I had never made a promise at all" [141, p. 82].

A second problem with the appeal to prima facie obligation is that there doesn't seem to be anything *conceptually incoherent* in the notion of a normative system that explicitly stipulates that certain of its norms are *always* to be regarded as all-things-considered (or *sans phrase*) norms; these can *never* be overridden by other norms. (In principle, a normative system could stipulate that *all* of its norms are to be interpreted in this way.) Now, I am not suggesting that it would be *wise* for a normative system to stipulate this sort of thing; I am only suggesting that it is *possible*. And if I am right about this, then we cannot rule out normative conflicts *a priori* by appealing to the distinction between prima facie and all-things-considered norms.

1.4.6 The argument from implicit qualification

Argument

All norms contain *implicit qualifications* that rule out normative conflicts. For example, the moral norm *It is wrong to lie* must be understood as including many—perhaps *infinitely* many—unstated exceptions: It is wrong to lie, *except when not lying would result in an innocent person being killed, the world being destroyed, etc., etc.* Similarly, the legal norm *Drivers must stop at red lights* should be understood as including implicit qualification along the lines of: *except when stopping at the red light would prevent an ambulance from passing, a police officer has directed the driver not to stop, etc., etc.* It would be unrealistic to require that all such exceptions be made explicit. Alan Donagan makes this point well:

> If a man accepts an invitation to dinner, it would be absurd for his host to understand him as having promised not to prevent a serious accident, or not to bring relief to victims of one, if to do these things would prevent him from dining. It is a promiser's duty to express any condition to his promise which the promisee might misunderstand; but there would be no misunderstanding in such a case, and to demand that the promiser stipulate all the emergencies on which his obligation would be anulled would be vexatious as well as superfluous. [48, p. 93]

On this account, when I promise to attend dinner, this *does not* generate a "prima facie" obligation that may be overriden by other "prima facie" obligations; rather, it generates a real, full-blooded, all-things-considered obligation, albeit one with many tacit conditions. Indeed, *all* norms are laden with such conditions. Such implicit qualifications ensure that normative conflicts will never, and can never, arise.

Reply

I see two main problems with this argument:

First, if it is true that all norms include implicit exceptions/qualifications, then we ought—at least in principle—to be able to make those exceptions/qualifications *explicit* by, e.g., writing them down as a (perhaps denumerably infinite) list. But I, for one, am not at all confident that I could list every possible circumstance in which (say) a given promise that I have made might be anulled. Nor am I confident that *everyone in the world*, working as a committee and with the rest of eternity to get the job done, could do so.[31]

My second objection is more powerful, I think: even if we *could*, in principle, list every qualification on every norm, we would not necessarily be able to ascertain that these qualifications rule out the possibility of conflict.

In light of these two problems, I conclude that the argument from implicit qualification fails.

1.4.7 The argument from metanorms

Argument

Norms are often subject to "metanorms" which come into play when conflicts arise, and determine which of the competing norms is more/most important and therefore should prevail. For example, legal systems typically have metanorms to the effect that later laws override earlier laws (*lex posterior derogat priori*), particular laws override general laws (*lex specialis derogat generali*), laws made by higher authorities override laws made by lower authorities (*lex superior derogat inferiori*), etc. Alternatively, there may be rules in place that give certain agents (e.g. judges, elected leaders) the authority to resolve conflicts by fiat. Such metanorms ensure that normative conflicts will never arise.

Reply

I do not deny the existence or usefulness of such metanorms. However, there are three reasons why the argument just given fails to establish that normative conflicts are impossible. First, it is surely not a *conceptual truth* about normative systems that they always have such conflict-resolving metanorms in place. Second, even if it were, we could have no guarantee that the metanorms *themselves* would never conflict. (Postulating an infinite hierarchy of metanorms, metametanorms, metametametanorms, etc. is completely *ad hoc*, and would not help anyway.) Third, and most important, the fact that the purpose of such norms is to resolve normative conflicts *presupposes*

[31] One might object that *God* could make such a list, but I find it hard to take this suggestion seriously, for it seems to imply that *only God* is preventing normative conflicts from being possible (which seems silly). Moreover, I don't think that the opponent of normative conflicts, theist or not, is going to want his argument(s) to depend on the existence of God; he is going to want to convince atheists and agnostics, too.

that there are, or at least could be, normative conflicts to resolve in the first place. If normative conflicts were impossible, there would be no need to develop means for dealing with them. But there *is* (clearly) a need to develop means for dealing with them. Therefore, normative conflicts are possible.

1.4.8 The argument from neutralization

Argument

Conflicting norms *neutralize* each other. As soon as one effective or binding norm conflicts with another effective or binding norm, both are immediately cancelled or nullified. Thus, normative conflicts can never arise—at least not for longer than an instant (the instant at which they neutralize each other).

Reply

There are three problems with this argument:

First, if conflicting norms really neutralized each other, it would be too easy to unburden oneself of unwanted obligations. Suppose, for example, that I have promised a friend to help him move into a new apartment. When moving day rolls around, I don't feel like helping him (say, I'm too lazy). So, I call up my brother and arrange to go to a baseball game with him at the same time that I am supposed to be helping my friend move—thereby incurring a conflicting obligation. Problem solved! Moreover, I don't even have the obligation to go to the game, since that obligation was cancelled out, too, as part of the bargain. (That's good, because I don't like baseball anyway; I just agreed to go to the game so I could get out of my other obligation.)

It might be objected that while I may *try* to incur a conflicting obligation, I will not succeed. Why not? If I call up my brother and promise to go to the game with him, haven't I incurred a genuine obligation? If/when I don't show up, won't he be justified in criticizing me for breaking my promise? No, the objector says. For I didn't really incur an obligation, so there was none to break. What my brother *would* be justified in being mad about is the fact that I *pretended* to incur an obligation—I *deceived* him into thinking that I had really incurred an obligation to go to the game with him. In doing so, I violated *another* obligation—the obligation not to decieve.

I am somewhat sympathetic to the objection just considered. I think it shows (or at least points out) that it is a *good policy* for a normative system to have this sort of principle—i.e., the principle that one cannot incur an obligation that conflicts with a prior obligation—built into it. What I reject is the idea is that it is *essential to the very concept* of a normative system that such a principle be built into it—the idea that normative systems incorporate such principles *by definition*. I don't see anything *incoherent* or *contradictory* in the notion of a normative system that explicitly stipulates, say, that *all* promises incur obligations, *regardless* of whether they conflict with other promises.

A second problem with the argument from neutralization is this. Suppose that a prohibition (*FA*) conflicts with a permission (*PA*). On the neutralization account, both norms are nullified. Thus, presumably, it becomes the case that not-*FA* and not-PA. Now, suppose—for the moment, at least—that *FA* is definable as not-*PA*, as many believe. Then we have a flat-out contradiction: FA and not-FA. (Of course, I am skeptical about the claim that FA is definable as not-PA. I am merely pointing out that *if* one accepts this principle (as my hypothetical opponent very well might), *then* one must reject the argument from neutralization.)

A third problem is that, even if normative conflicts arise only for a single instant (i.e., an infinitesimal period of time), this is enough to establish my point. All I am claiming is that normative conflicts are possible. I am not claiming that "lasting" normative conflicts (i.e., conflicts of any significant duration) are possible (though I am happy to believe that they are).

1.4.9 The argument from disjunctive duty

Alan Donagan [49] and David Brink [32] have argued that when faced with two conflicting obligations (where neither overrides or outweighs the other), an agent merely has the obligation to fulfill *one or the other* obligation—a "disjunctive duty," as it were.

Let me back up slightly. In support of her claim that moral dilemmas are possible, Marcus asks us to imagine that "[t]he lives of identical twins are in jeapordy, and, through force of circumstances, I am in a position to save only one. Make the situation as symmetrical as you please" [102, p. 192 (Gowans)]. Her point is that there is no possible justification for saying that one has an obligation to save the first twin, but not the second (or vice versa). Moreover, it would be absurd to say that one has no obligation to save either of them. Hence it can only be that one has an obligation to save *both*, even though this is impossible.

In response to Marcus's argument Donagan writes:

> Where the lives of identical twins are in jeopardy and I can save one but only one, every serious rationalist moral system lays down that, whatever I do, I must save one of them. By postulating that the situation is symmetrical, Marcus herself implies that there are no grounds, moral or nonmoral, for saving either as opposed to the other. Why, then, does she not see that, as a practical question, Which am I to save? has no rational answer except "It does not matter," and as a moral question none except "There is no moral question?" Certainly there is no moral conflict: from the fact that I have a duty to save either *a* or *b*, it does not follow that I have a duty to save *a* and a duty to save *b*. Can it be seriously held that a fireman, who has rescued as many as he possibly could of a group trapped in a burning building, should blame himself for the deaths of those left behind, whose lives could have been saved only if he had not rescued some of those he did? [49, pp. 286-7 (Gowans)]

Brink makes the point more formally. Using 'o' to stand for prima facie obligation, 'O' to stand for all-things-considered obligation, and '>' to indicate that one prima facie norm outweighs or overrides another, he writes:[32]

> In an insoluble conflict of undefeated prima facie obligations, the following claims seem true.
>
> $o(A)$
>
> $o(B)$
>
> $\neg(o(A) > o(B))$
>
> $\neg(o(B) > o(A))$
>
> $O(A \vee B)$
>
> $\neg O(A)$
>
> $\neg O(B)$
>
> If so, the only all-things-considered obligation in an insoluble conflict is this disjunctive obligation. [32, p. 115]

[32] Brink is of course using '$\neg$' for negation and '$\vee$' for disjunction.

The argument from disjunctive duty is subject to two serious objections:

First, as with the argument from neutralization, it makes it too easy to evade unwanted obligations. Let us return to the example of my friend who is moving. I can still get out of helping him move by arranging to go to a baseball game with my brother. Once I do that, I only have the disjunctive obligation of *either* helping my friend move *or* going to the game with my brother. So while I can't get out of both (as I could if the obligations neutralized each other), I *can* decide which obligation I would prefer to fulfill, and walk away from the situation with a clear conscience.

The second objection is this. Suppose I am faced with a standard inescapable normative conflict: a situation in which I am obliged to see to it that A, and also obliged to see to it that not-A. On the disjunctive account, I merely have the *vacuous* obligation to see to it that A or not-A! No obligation could be easier to fulfill: no matter how things turn out, I am in the clear! This seems absurd. For all I would have to do to weasel out of an obligation that I don't like is to incur an obligation to do *precisely the opposite*. Then I am "home free" no matter what happens. Again, this would make life too easy.

1.4.10 The argument from relativity

Argument

Contradictory sentences can be true *according to*, or *relative to*, a body of information (e.g. a scientific theory, belief system, database, or work of fiction), but it would be a mistake to infer from this that contradictions can be true *simpliciter*, or absolutely. Similarly, norms can conflict *according to*, or *relative to*, a normative system, but it is a mistake to infer from this that norms can conflict *simpliciter*, or absolutely.

Reply

Just as I cannot make sense of the concept of something's being *known* absolutely (i.e. without at least implicit reference to some epistemic agent), I cannot make sense of something's being *obligatory* (or *permissible*) absolutely (i.e. without at least implicit reference to some normative system). Nevertheless, I am perfectly willing to grant, if only for the sake of argument, that "absolute" normative conflicts are impossible. For this does not contradict my central claim, which is that *normative conflicts (of some kind or other) are possible.* If it helps, one can think of expressions like 'it is obligatory that' and 'it is permissible that' as being implicitly parameterized, i.e. shorthand for something like 'it is obligatory in (according to, relative to) normative system S that' and 'it is permissible in normative system S that'. In fact, given my skepticism about the very idea of an "absolute" norm, I *encourage* this interpretation. To put my point bluntly: the fact that normative conflicts are relative doesn't make them any less real.

One might object that if I am going to take this stance, I am forced to concede that *true contradictions* (or dialetheias) are possible, too, since contradictory statements can be true *relative to* a theory, belief system, work of fiction, etc. This doesn't follow at all, however; for I think that the notion of absolute (non-relative) truth, unlike the notions of absolute knowledge or absolute obligation, is perfectly coherent. I *do* concede that contradictions can be true *in a story, model, etc.*; but I do not concede that they can be true in any absolute sense. Similarly, I believe that normative conflicts can hold or obtain according to a normative system, but I do not believe that they can hold or obtain in any absolute sense. The difference is that I reject absolute normative conflicts because I find the very notion of an absolute norm (and hence an absolute normative conflict) incoherent, whereas I reject absolute true contradictions because I think that they are ruled out by the nature of (absolute) truth.

1.4.11 The argument from impotence

Argument

Perhaps a normative system can generate conflicts of the sort you have discussed. But any system capable of generating conflicts is *ipso facto* impotent, i.e., incapable of generating binding obligations, etc. For, in the presence of normative conflicts, all normative distinctions would break down, and it would therefore be impossible to determine what one is obliged, forbidden, or permitted to do.

Reply

The premise of this argument is simply false. Take, for example, the case (mentioned in Section 1.3.2) of the document that on the one hand must be kept in a safe, and on the other hand must be thrown away. In this case there is a dilemma or quandary as to what one should do with the document in question, but there is no reason to think that any—let alone all—other norms are affected. It is still quite clear in such a situation that, e.g., one is still not obliged to *murder the organization's president*, that one is still not forbidden to *wear clothes to work*, etc. One may attempt to show via argument that, despite appearances to the contrary, conflicting norms really do obliterate all normative distinctions. But, as we saw in our discussion of the argument from explosion/deontic inheritance (Section 1.4.4), the assumptions invoked in such an argument are far from self-evident.

One can take the argument from impotence in another way, namely as a *refusal*: "I hereby refuse to acknowledge as binding or legitimate any normative system (or norm) that generates, or is capable of generating, a normative conflict." As such, it is not the sort of thing that can be refuted (one can't, after all, refute a speech act). I will, however, note that this seems an unrealistically high *ontological* standard to which to hold normative systems (or individual norms), especially systems of law (or individual laws). As Carlos Alchourrón and Eugenio Bulygin point out,

> it is extremely important to realize that inconsistent normative systems are perfectly possible and their occurrence, at least in certain areas like law, is rather frequent. The reason for this fact is fairly clear. The selection of the propositions that form the basis of the system [...] is based on certain empirical facts: the acts of commanding or promulgating. Now, there is nothing extravagant about the idea that an authority commands that p while another authority commands that $\neg p$. Even one and the same authority may command that p and that $\neg p$ 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. [5, pp. 112-3][33]

I think Alchourrón and Bulygin would agree that it is fine to take freedom-from-conflict as a *normative* standard (i.e., a criterion, among others, for evaluating the *goodness* or *adequacy* of normative systems), but to take it as an *ontological* standard (i.e., a necessary condition for *existence* as a legitimate normative system, i.e. one that is capable of generative effective/binding norms) is unrealistic and naive. Once

[33] I have replaced the authors' '$\sim$' with '$\neg$'.

again it would be wise to heed the words of John Austin: "The *existence* of law is one thing; its *merit or demerit* is another" [14, p. 184 (my emphasis)].

1.4.12 The argument from "common sense"

Argument

Human beings are not mindless automata who must blindly adhere to the literal interpretations of norms. When faced with an apparent normative conflict, we can simply use *common sense* to determine what to do. For example, if one has made conflicting promises, common sense would dictate that one ought to explain to both promisees that one has made a mistake, and work out a reasonable compromise that is as fair as possible to both parties. Similarly, with respect to the question of what to do with the five-year-old secret document, common sense would dictate that one should do whatever is likely to be more in accord with the general goals of the organization, whatever one would have an easier time justifying to one's superiors, etc.

Reply

Of course we often use common sense to figure out what to do in the face of a normative conflict, as we should. But this does not cancel out the fact that *there was a normative conflict that needed to be dealt with in the first place.* Indeed, the fact that we use common sense to deal with normative conflicts supports my case. For if normative conflicts were impossible, we would not need to use common sense to deal with them.

1.4.13 The argument from change of subject

Argument

Anyone who uses the terms *obligatory* (O), *forbidden* (F), and *permissible* (P) in such a way that both OA and FA, or both FA and PA, could be true, either does not know what she is talking about, or is using the terms in a new and unorthodox way. In either case, she cannot have shown that normative conflict is possible; for she has merely *changed the subject.* At most she has shown that some *other* kind of conflict ("pseudo-normative" conflict?) is possible. But that may be readily conceded. Abraham Lincoln is supposed to have once asked "If you call a tail a leg, how many legs does a dog have?" The answer, of course, is *four*: calling something a leg does not make it a leg. By the same token, calling something a norm does not make it a norm.

Reply

If I were to claim that, say, by definition *everything* is obligatory (i.e. for all A, OA), or that by definition whatever is obligatory is true (i.e. for all A, if OA then A), there might be some reason to think that I was using the term *obligatory* in a non-standard or eccentric way (or that I simply don't know what I'm talking about). This is because no reasonable person would find these claims even remotely plausible, so they would naturally assume that I was using the term differently than they do. However, many people, including many philosophers who have considered the issue very carefully, believe that normative conflicts are possible.[34] So it will not do to dismiss the claim that normative conflicts are possible as a mere "change of subject."

Of course, one is welcome to *stipulate* that she intends to use the words *obligation*, *permission*, etc. in such a way that, by definition, OA and FA, etc., cannot be true. But this is irrelevant to the question of whether normative conflicts are in fact possible (just as my decision to call a tail a leg has nothing to do with how many legs a dog actually has). When I claim that normative conflicts are possible, I am claiming that *the best semantic account or model* of our normal, everyday usage of the terms *obligatory*, *forbidden*, etc.—in other words, the best account of *whatever it is that we mean* by *obligatory*, etc.—does not rule out the possibility of PA and FA, etc. being jointly true. One can certainly argue that this claim is *wrong*, but one cannot legitimately say that I am somehow changing the subject, that I am not really talking about obligation, permission, etc., but rather about something else entirely. *That* would be nothing but a cop-out.

1.5 The tradeoff

I would like to think that in this chapter I have made a convincing case for the possibility of normative conflicts. But perhaps this is an overly ambitious hope. In all honesty, I will be happy if I have convinced the reader that we should, for the time

[34] Searle even refers to the existence of moral conflicts as "an obvious fact about human experience" [141, p. 86].

being at least, keep an *open mind* about whether normative conflicts are possible. We have seen that there is a *tradeoff* of sorts: if we accept the possibility of normative conflicts, then we are forced to abandon certain other principles that are at least *prima facie* plausible—either in and of themselves, or in light of arguments that can be given in their support. The principles in question are these:

1. *Interdefinability.* FA is equivalent to not-PA (and PA is equivalent to not-FA).
2. *Aggregation.* If two propositions are individually obligatory, then they are jointly obligatory.
3. *Ought implies may.* If something is obligatory, then it is permissible.
4. *Ought implies can.* If something is obligatory, then it is possible.
5. *Deontic inheritance.* If A entails B, then OA entails OB.
6. *Explosion.* Everything follows from a contradiction.

If we accept the possibility of normative conflicts, we must give up some appropriate combination of the above principles. Is it worth it? That is, does the evidence in support of normative conflicts outweigh the evidence in support of principles (1)-(6)? I submit that if we are to answer this question in an informed way, we must first investigate formal logical systems which allow for the possibility of normative conflicts, and thus reject some or all of the principles (1)-(6). If such systems can be developed in a satisfactory way—that is, if they are on the whole acceptable and not obviously inferior to their alternatives—then we will, I think, have all the more reason to accept that normative conflicts are possible, as well as a satisfactory way of reasoning in their presence. If not, then we will have good reason to think that, despite the arguments of this chapter, normative conflicts are indeed impossible, and thus there is no real need to develop means of reasoning with them.

Accordingly, we now turn to the formal investigation of deontic logics (logics of obligation, permission, etc.). We will return to the tradeoff in Chapter 6, where we will be in a better position to evaluate it.

Chapter 2

Standard deontic logic

> Standard deontic logic suffers badly from explosion. Since in classical logic $\alpha, \neg\alpha \models \beta$ it follows that $O\alpha, O\neg\alpha \models O\beta$: if you have inconsistent obligations you are obliged to do everything. This is surely absurd. People incur inconsistent obligations; this may give rise to legal or moral dilemmas, but hardly to legal or moral anarchy.
>
> —Graham Priest [128, p. 343]

> Deontic logic, born in its modern form in the early 1950's, has remained something of a problem child in the family of logical theories.
>
> —G. H. von Wright [156, p. 1]

In this chapter I begin considering formal logical systems that treat 'obligatory', 'permissible', and 'forbidden' as *logical* terms—that is, terms whose meanings remain constant across all (logical) interpretations, or models, of the language. Such systems provide, or purport to provide, mathematically precise characterizations of the contributions that 'obligatory', 'permissible', and 'forbidden' make to what follows from what.[1] For example, they tell us which of the following inferences are correct:[2]

You should eat.
You should drink.
You should be merry.
∴ You should eat, drink, and be merry.

[1] Cf. McNamara [116, ¶1].
[2] I use the symbol '∴' as shorthand for 'therefore'.

I am permitted to drink.
I am permitted to drive.
∴ I am permitted to drink and drive.

Ed is forbidden to jaywalk.
∴ It ought to be that if Ed jaywalks, he gets the death penalty.[3]

Walter ought to either visit his grandmother or call her.
It ought to be that if he visits, he calls.
∴ Walter ought to call his grandmother.[4]

You may kiss the bride.
∴ You may kiss the bride or kill the mother-in-law.

One should not smoke while praying.
∴ One should not pray while smoking.[5]

Traditionally such formal systems are referred to as *deontic logics*, and the study, or general class, of such systems is referred to as *deontic logic*.[6]

In what follows I set out the "standard" system of deontic logic, along with some variations on this system. (Collectively, I call these the D systems.) I show that the D systems, which are based on classical propositional logic, cannot deal adequately with normative conflicts. Next, I consider two alternatives to the D systems: *non-aggregative deontic logic*s, and *deontic logics of permitted inheritance*. These systems, which are also classically-based, have been marketed as being well-suited for reasoning with normative conflicts. However, as I will argue, they are each too strong in certain respects and too weak in others. Finally, I pinpoint the classical principle of explosion ("everything follows from a contradiction") as the culprit in these failings, and suggest that a solution may be found by constructing and investigating systems of deontic logic that are based on *paraconsistent logic*, a type of non-classical logic in which the principle of explosion fails.

[3] Cf. Mares [103, p. 5].
[4] Cf. Bonevac [29, p. 40].
[5] Based on an old joke. A man asks his priest, "May I smoke while I pray?" The priest replies, "Absolutely not!" The man goes home and ponders this. The next day he asks the same priest: "May I pray while I smoke?" The priest replies, "Why, of course!"
[6] For an excellent introduction to deontic logic, see McNamara [116]. Åqvist [10] and Hilpinen [80] are also good.

2.1 A note on 'deontic'

The term 'deontic logic' was first used by G. H. von Wright in his seminal paper of the same name [154], published in *Mind* in 1951. 'Deontic' is derived from the Greek $\delta\epsilon\acute{o}\nu\tau\omega\varsigma$, meaning *duly* or *as it should be* [155, p. 130]. In my view, a better term for the systems we will be discussing (at least, given the way we will be interpreting them) would be 'logic of normative propositions'.[7] Unfortunately, this expression is somewhat cumbersome. Thus, I will use the term 'deontic logic'. For the purposes of this dissertation, this can be thought of as an abbreviation of 'logic of normative propositions'. It should be noted, however, that some authors take the term to encompass, or even exclusively refer to, the logic of *norms*.[8]

2.2 Historical background

The main idea behind deontic logic (as it is usually construed) is that the deontic expressions 'obligatory', 'permissible', and 'forbidden' (and, perhaps, 'gratuitous') express *modes* of truth, making deontic logic a branch of modal logic. This idea was first hinted at by Leibniz,[9] who observed that:

- What is obligatory (*debitum*) is what is *necessary* for a good person to do.
- What is permitted (*licitum*) is what is *possible* for a good person to do.
- What is forbidden (*illicitum*) is what is *impossible* for a good person to do.

(He might have added that what is gratuitous is what is possible for a good person to refrain from doing.) Thus the deontic modalities appear to form a familiar "square of opposition":

OA	FA
PA	GA

[7] Alchourrón and Bulygin [3] [5] [6] long advocated the use of this term, but few have adopted it.

[8] For example, Alchourrón [3] uses 'deontic logic' to refer to the logic of norms and 'normative logic' to refer to the logic of normative propositions. (On his account, the latter presupposes the former.)

[9] See Von Wright [157].

Here the formulas on the top row are supposed to be *contraries* (i.e. they can't both be true), the formulas on the bottom row are supposed to be *subcontraries* (i.e. they can't both be false), and the formulas on the diagonals are supposed to be *contradictories* (i.e. they can't both be true and they can't both be false). OA is supposed to imply PA, and FA ($= O\neg A$) is supposed to imply GA ($= P\neg A$). (I repeatedly use the word 'supposed' because, as we shall see, I am skeptical about all of these claims. I concede, however, that they do have some *prima facie* plausibility.)

In the 1920s, Ernst Mally, a student of Alexius Meinong, developed the first formal system of (what would come to be known as) deontic logic.[10] Mally wrote '!' for 'it is obligatory that', 'f' for '*fordert*' (German for 'requires' or 'demands'), '$\cup$' for 'the unconditionally obligatory', and '$\cap$' for 'the unconditionally forbidden'. He then laid down some intuitive axioms:[11]

1. $AfB \wedge (B \rightarrow C) \rightarrow AfC$ (if A requires B and B implies C, then A requires C)

2. $AfB \wedge AfC \rightarrow Af(B \wedge C)$ (if A requires B and A requires C, then A requires both B and C)

3. $AfB \leftrightarrow !(A \rightarrow B)$ (A requires B iff it is obligatory that A implies B)

4. $\exists\cup!\cup$ (there is an unconditionally obligatory state of affairs)

5. $\neg(\cup f \cap)$ (the unconditionally obligatory does not require the unconditionally forbidden)

Mally's system has a certain *prima facie* plausibility, but unfortunately, as Karl Menger [118] observed, $!A \leftrightarrow A$ can be derived in it: that is, the system has it that what is true is precisely what *ought* to be true! I will refer to this "law"—and sometimes one or the other of its halves—as the *Panglossian principle*.

Von Wright's 1951 paper "Deontic Logic" (mentioned above) contains the first formal system of deontic logic that can be regarded as a plausible formalization of (certain aspects of) normative discourse in ordinary language. Von Wright took *permission* as the primitive deontic concept, construing it as a property of acts (technically, *types* of act), such as stealing and swimming. Thus von Wright's system can be thought of as formalizing the notion of the "ought to do" (*Tun-Sollen*) rather

[10] See Mally [100], Lockhorst [98], Lockhorst and Goble [99].

[11] For the purposes of this chapter, $\rightarrow$ represents a generic conditional while $\supset$ represents the classical material conditional.

than the "ought to be" (*Sein-Sollen*). He used the usual sentential connectives in a somewhat non-standard but easily interpretable way: if a is an act, $\neg a$ is the act of not peforming a, $a \vee b$ is the act of performing either a or b, etc. He wrote Pa for $\ulcorner a$ is permitted$\urcorner$, and Oa (defined as $\neg P \neg a$) for $\ulcorner a$ is obligatory$\urcorner$. The fundamental principles von Wright's system are the Principle of Deontic Distribution and the Principle of Permission, which are, respectively:

1. $P(a \vee b) \leftrightarrow Pa \vee Pb$ (performing a or b is permitted iff either a is permitted or b is permitted)

2. $Pa \vee P\neg a$ (either it is permissible to perform a or it is permissible to refrain from performing a)

Von Wright devised a rather curious semi-truth-functional semantics for his deontic logic. This semantics is best illustrated with an example: the truth value of $P(a \vee b)$ is a function of the truth values of $P(a \wedge b)$, $P(a \wedge \neg b)$, $P(\neg a \wedge b)$, and $P(\neg a \wedge \neg b)$. For example, if these last four formulas are *true*, *false*, *true*, and *false*, respectively, then $P(a \vee b)$ is *true*. This semantics is severely lacking in elegance and intuitive appeal, which is probably why it is not well known or widely used today.

In 1963, Saul Kripke published his important paper "Semantical Analysis of Modal Logic I" [92], in which he gives possible-worlds semantics for various "normal" modal logics. In a brief remark at the end of the paper, Kripke suggests that one could develop modal logics in which the box operator, $\Box$, could be interpreted as *deontic* necessity (i.e. obligation) by altering his semantics so as to render $\Box A \supset A$ invalid. Kripke's suggestion was taken up by William H. Hanson in his 1965 paper "Semantics for Deontic Logic" [75]. One of the deontic logics discussed in that paper, the system D, has come to be known as "standard deontic logic" (often "SDL").[12] Interestingly, from a purely proof-theoretic point of view, D is exactly the logic we get by adding to von Wright's original system the principle that tautologies (or tautological acts, e.g. either talking or not talking) are obligatory—a principle which, incidentally, von Wright explicitly considered and rejected. The main difference is that in D, the deontic operators apply to propositions rather than act-types. Thus the ascension of D to the status of "standard deontic logic" marked a move from the "ought to do" to

[12] D is also known as KD [43], OK^+ [10], and $\mathsf{K}\eta$ [127]. To complicate things further, I will refer to it as Ds. (See Section 2.3 for an explanation/apology.)

the "ought to be," and the latter has dominated deontic logic ever since.[13]

In a moment I will set out a family of deontic logics which I call the D systems. Among these is the system D, which I call Ds. But first:

2.3 A note on system names

In a work like this, one faces a dilemma (dare I say: *normative conflict*?) when it comes to choosing names for logical systems. On the one hand, such names should be compositional and informative (so that one can easily determine the basic, defining properties of the system, and its relations with other systems, from the name alone). On the other hand, such names should be as terse as possible; we want to avoid systems with names like $\mathsf{STW}_4\mathsf{D}^2\mathsf{ca.1x}^+\mathsf{GMv}^{\copyright}_{\rightarrow}$. Moreover, it is desirable to choose names that fit nicely with other system names that have been used in the literature. The best solution, I think, is a compromise. This results in system names that may not *always* be as compositional and informative as one would like, and do not always mesh as well as one would like with other names in the literature, but at least have succinct, pronounceable names. So, for example, I apologize for using 'D' in a way that flies in the face of the way it is normally used in the literature on deontic logic; but please realize that I have good reasons for doing so. (These reasons will become apparent in later chapters.)

2.4 Notational conventions

At this stage I want to establish some notational conventions to which I will adhere throughout this dissertation. The reader may consult Appendix A for a summary of the notation used herein.

- The set of atomic formulas, At, is $\{p_i : i \in \mathbb{N}\}$, i.e. $\{p_0, p_1, \cdots\}$.
- For mnemonic/aesthetic reasons I will often write atomic formulas as unsubscripted lowercase Roman letters (s, c, p, etc.) or strings thereof (*snap*, *crackle*, *pop*, etc.). I will also sometimes use p as a variable ranging over elements of At.

[13] It should be noted, though, that Mally's system was a formalization of the *Sein-Sollen*. It is also worth mentioning that in recent years, authors such as Horty [83] and Belnap, Perloff, and Xu [24] have revived the *Tun-Sollen* interpretation via the *stit* ('sees to it that') operator.

- A set of formulas Γ is *closed under* a unary connective, $*$ just in case $*A$ is in Γ whenever A is. Γ is closed under a binary connective $*$ just in case $(A * B)$ is in Γ whenever A and B are.

- If $*_1, \cdots, *_n$ are unary or binary connectives, $L(*_1, \cdots, *_n)$ is the smallest superset of At that is closed under $*_1, \cdots, *_n$.

- If S is a logical system, L_{S} is the language of that logical system.

- Except where otherwise specified or implied, uppercase Roman letters ($A, B, \cdots$) range over formulas and uppercase Greek letters ($\Gamma, \Delta, \cdots$) range over *finite* sets of formulas.[14]

- Outer parentheses of formulas may be omitted. For example, $(A \wedge B)$ may be written as $A \wedge B$.

- $\vee$ and $\wedge$ bind more tightly than $\supset$, $\equiv$, $\rightarrow$,and $\leftrightarrow$. Thus, for example, $A \supset A \vee B$ is interpreted as $A \supset (A \vee B)$.

- Γ / A is the inference whose premises are the elements of Γ and whose conclusion is A. (Technically, we may identify Γ / A with the ordered pair $\langle \Gamma, A \rangle$.) Set brackets may be omitted on the left side of the slash. For example, $\{A \supset B, A\} / B$ may be written as $A \supset B, A / B$.

- $\models$ represents semantic consequence. $\Vdash$ represents proof-theoretic consequence. $\vdash$ is used to generalize over $\models$ and $\Vdash$. Set brackets may be omitted on the left side of any of these "turnstile" symbols. For example, $\{A \supset B, A\} \vdash B$ may

[14] This is purely for convenience. Allowing infinite premise sets would introduce some complications that have little or no payoff, at least in the context of this dissertation. The goal here is to develop formal systems that can adequately model the correctness/incorrectness of *real-world, practical inferences* involving 'ought's and 'may's. It seems to me that all—or, at least, the vast majority—of inferences (especially *deontic* inferences) one encounters in the real world involve only finitely many premises.

Of course, it should be acknowledged that some axiomatic theories (e.g. first-order arithmetic) have infinitely many axioms, and that only a strong completeness theorem (which states that any semantic consequence of a set of formulas, whether finite or infinite, is derivable from that set) will guarantee that every sentence that is a consequence of the axioms is provable from the axioms. My completeness theorems establish only weak completeness, i.e., they show only that if a sentence is a consequence of a finite premise set, then it is provable from that set. It is conceivable that someone might wish to set up a theory with infinitely many axioms in one of the formal languages specified in this dissertation. In that case, absent a separate compactness theorem, the weak completeness results given here would not show that each consequence of the axioms is derivable from some finite subset of the axioms via tableaus. But it doesn't seem terribly important to address this issue here.

be written as $A \supset B, A \vdash B$. $\varnothing \vdash A$ may be abbreviated as $\vdash A$. $A \dashv\vdash B$ is shorthand for ($A \vdash B$ and $B \vdash A$).

- Commas are sometimes used to represent set union, e.g. writing $\Gamma \cup \Delta$ and $\Gamma \cup \{A\}$ as Γ, Δ and Γ, A, respectively. (This is primarily done on the left side of a turnstile or slash.)
- Where f is a unary function I may write $f(x)$ as f_x. Where f is a binary function I may write $f(x, y)$ as $f_y(x)$.
- I use $\Rightarrow$ for (metalinguistic) material implication, and $\Leftrightarrow$ for (metalinguistic) material equivalence.

2.5 The D systems

The language of the D systems is $L(\neg, \vee, O)$. We have the following defined connectives:

- $A \supset B =_{df} \neg A \vee B$
- $A \wedge B =_{df} \neg(\neg A \vee \neg B)$
- $A \equiv B =_{df} (A \supset B) \wedge (B \supset A)$
- $FA =_{df} O\neg A$
- $PA =_{df} \neg O\neg A$
- $GA =_{df} \neg OA$

2.5.1 Semantics

I will give the formal definitions first, then comment on the intuitions underlying them.

Definition 1 (D model) A model for D is a triple $\langle W, R, v\rangle$, where W is a non-empty set, $R \subseteq W^2$, and $v : At \times W \longrightarrow \{\mathbf{t}, \mathbf{f}\}$. v is extented to $\bar{v} : L(\neg, \vee, O) \times W \longrightarrow \{\mathbf{t}, \mathbf{f}\}$ via the following clauses. For all $w \in W$, $p \in At$, and $A, B \in L_{\mathsf{D}}$:

$$\begin{array}{lcl}
\bar{v}_w(p) & = & v_w(p) \\
\bar{v}_w(\neg A) & = & \begin{cases} \mathbf{t} \text{ if } \bar{v}_w(A) = \mathbf{f} \\ \mathbf{f} \text{ otherwise} \end{cases} \\
\bar{v}_w(A \vee B) & = & \begin{cases} \mathbf{t} \text{ if } \bar{v}_w(A) = \mathbf{t} \text{ or } \bar{v}_w(B) = \mathbf{t} \\ \mathbf{f} \text{ otherwise} \end{cases} \\
\bar{v}_w(OA) & = & \begin{cases} \mathbf{t} \text{ if } (\forall u \in W)(wRu \Rightarrow \bar{v}_u(A) = \mathbf{t}) \\ \mathbf{f} \text{ otherwise} \end{cases}
\end{array}$$

Remark 1 It is easy to check that the following hold:

$$\begin{array}{lcl}
\bar{v}_w(A \wedge B) & = & \begin{cases} \mathbf{t} \text{ if } \bar{v}_w(A) = \bar{v}_w(B) = \mathbf{t} \\ \mathbf{f} \text{ otherwise} \end{cases} \\
\bar{v}_w(A \supset B) & = & \begin{cases} \mathbf{t} \text{ if } (\bar{v}_w(A) = \mathbf{t} \Rightarrow \bar{v}_w(B) = \mathbf{t}) \\ \mathbf{f} \text{ otherwise} \end{cases} \\
\bar{v}_w(A \equiv B) & = & \begin{cases} \mathbf{t} \text{ if } \bar{v}_w(A) = \bar{v}_w(B) \\ \mathbf{f} \text{ otherwise} \end{cases} \\
\bar{v}_w(FA) & = & \begin{cases} \mathbf{t} \text{ if } \forall u(wRu \Rightarrow \bar{v}_u(A) = \mathbf{f}) \\ \mathbf{f} \text{ otherwise} \end{cases} \\
\bar{v}_w(PA) & = & \begin{cases} \mathbf{t} \text{ if } \exists u(wRu \text{ and } \bar{v}_u(A) = \mathbf{t}) \\ \mathbf{f} \text{ otherwise} \end{cases} \\
\bar{v}_w(GA) & = & \begin{cases} \mathbf{t} \text{ if } \exists u(wRu \text{ and } \bar{v}_u(A) = \mathbf{f}) \\ \mathbf{f} \text{ otherwise} \end{cases}
\end{array}$$

Remark 2 It is also easy to check that $\bar{v}_w(\neg A) = \mathbf{t}$ iff $\bar{v}_w(A) \neq \mathbf{t}$, and $\bar{v}_w(\neg A) = \mathbf{f}$ iff $\bar{v}_w(A) \neq \mathbf{f}$.

Definition 2 (Ds model) A Ds model is a D model $\langle W, R, v\rangle$ such that R is *serial*, i.e. $\forall w \exists u\ wRu$.

Definition 3 (Dh model) A Dh model is a D model $\langle W, R, v\rangle$ such that R is *shift-reflexive*, i.e. $\forall w \forall u(wRu \Rightarrow uRu)$.

Definition 4 (Dsh model) A Dsh model is a D model $\langle W, R, v\rangle$ such that R is both serial and shift-reflexive.

Notation 1 I will use D+ as a variable ranging over all four D systems, namely D, Ds, Dh, and Dsh.

Definition 5 (semantic consequence for $\mathsf{D}+$) A is a semantic consequence of Γ in $\mathsf{D}+$ (in symbols, $\Gamma \models_{\mathsf{D}+} A$) just in case for every $\mathsf{D}+$ model $\langle W, R, v\rangle$ and $w \in W$, if $\bar{v}_w(B) = \mathbf{t}$ for all $B \in \Gamma$, then $\bar{v}_w(A) = \mathbf{t}$.

Remark 3 The system D is just a deontic version of the well-known modal logic K. (Or, if you prefer, it just *is* K.)

Intuitively, W is a set of *possible worlds*, which I like to construe as *complete specifications of ways that the actual world might have been* (feel free to substitute your own preferred conception, if you have one). Alternatively, we can think of the elements of W as points, states, situations, "ways,"[15] or what-have-you; strictly speaking, they can be anything one wishes. R is the deontic accessibility relation: wRu may be read in any or all of the following ways (which I take to be roughly equivalent):

- everything that *ought* to be the case at w *is* the case at u
- everything that holds at u is permissible at w
- u is deontically perfect or ideal relative to w
- u is permissible from the perspective of w
- u is a deontic alternative to w
- u is w-permissible
- w "sees" or "permits" u

v is a valuation function, assigning truth values to atomic formulas at worlds. (Obviously, $\mathbf{t}$ and $\mathbf{f}$ represent the truth values TRUE and FALSE, respectively.[16]) $\bar{v}$ is the extention of v to all formulas of the langauge. The clauses for $\neg$ and $\vee$ are just the usual truth-functional clauses, relativized to worlds. OA can be thought of as saying something like ⌜A holds in every permissible world⌝ or ⌜in a perfect world, A would be true⌝.

Semantic consequence is defined in a standard way, as truth preservation at every world in every model. That is, A is a semantic consequence of Γ just in case whenever

[15] As in "ways things could be" and "ways things couldn't be."

[16] Later we will have reason to define $\mathbf{t}$ and $\mathbf{f}$ as $\{1\}$ and $\{0\}$, respectively, but this is not important now.

every element of Γ is true (at a world, in a model), A is also true (at that world, in that model).

The seriality condition says that every world has at least one deontic alternative—i.e., there are no "dead-end" worlds. Very roughly, this corresponds to the idea that in any situation, there is at least one permissible course of action. Seriality validates the so-called "D" schema, $OA \supset PA$, which says that whatever is obligatory is permissible. Many take $OA \supset PA$ to be the fundamental or characteristic axiom of deontic logic, the deontic counterpart to the "T" schema, $\Box A \supset A$.

The shift-reflexivity condition says that every seen world sees itself. Very roughly, this corresponds to the idea that one cannot violate an obligation by doing something that is permissible. Shift-reflexivity validates the so-called "M" (or "U") schema, $O(OA \supset A)$, which says that it is obligatory that obligations be fulfilled.[17] Some take this to be an obvious truth of deontic logic. Anderson, for example, writes:

> Of course, it is notoriously false, under *any* understanding of "if ... then —," that
>
> if Op then p.
>
> But wouldn't it be nice if it were true? In fact, it *ought* to be true, and would be, if this were the best of all possible worlds. [8, p. 356]

Others, such as Goble [63], have comparably strong intuitions against the principle, however.

Example 1 (inferential D schema) Consider the inference $OA \diagup PA$ ($\ulcorner$it is obligatory that A; therefore, it is permitted that $A\urcorner$). To see that this is valid in Ds (i.e., $OA \models_{\mathsf{Ds}} PA$), suppose that OA holds at an arbitrary world w. Then, by the clause for O, A holds at every world permitted by w. By seriality, there exists such a world; call it \$. Since (as one can easily verify), all worlds are consistent, $\neg A$ is not true at \$. Thus it is not the case that $\neg A$ is true at every world permitted by w. Thus $\neg O \neg A$ ($=_{df}$ PA) holds at w. Without seriality, however, this argument does not go through. Consider any D model in which $W = \{w_0\}$ and $R = \varnothing$. (Note that any such model is also, vacuously, a Dh model.) It is easy to check that in this model $\bar{v}_{w_0}(OA) = \mathbf{t}$ but $\bar{v}_{w_0}(PA) = \mathbf{f}$. Thus $OA \not\models_{\mathsf{D}} PA$ (and $OA \not\models_{\mathsf{Dh}} PA$).

[17] I will sometimes refer to this as the "obey all rules" principle. The reason is this. I recently encountered a reference to a television character (played by Don Knotts) saying "The first rule is: obey all rules." This statement, with its self-referential character, reminded me very much of $O(OA \supset A)$.

2.5.2 Proof theory

I now present tableau-style proof theories for the D systems. While this is not strictly necessary for my exposition in this chapter, it will provide a gentle introduction to some concepts and techniques that will be important in subsequent chapters. To make the dissertation reasonably self-contained, I will assume that the reader has little or no familiarity with tableau-style proof theories.

In general, a tableau looks something like this:

$$
\begin{tikzcd}
 & \cdot \arrow[d] & & \\
 & \cdot \arrow[dl] \arrow[dr] & & \\
\cdot \arrow[d] & & \cdot \arrow[dl] \arrow[dr] & \\
\cdot & \cdot & & \cdot
\end{tikzcd}
$$

The dots are called *nodes*. The node at the top is called the *root*. The nodes at the bottom are called *tips*. Any path from the root down a series of arrows as far as you can go is called a *branch*.[18]

Now for some definitions:

Definition 6 (initial list) Let $\Gamma = \{B_0, \cdots, B_n\}$. If $\neg A \notin \Gamma$, the initial list for Γ / A is:

$$
\begin{array}{c}
B_0\ 0 \\
\vdots \\
B_n\ 0 \\
\neg A\ 0
\end{array}
$$

To ensure that the initial list contains no redundant nodes, we stipulate that if $\neg A \in \Gamma$, then the initial list for Γ / A is simply:

$$
\begin{array}{c}
B_0\ 0 \\
\vdots \\
B_n\ 0
\end{array}
$$

[18] This paragraph is lifted, with some minor alterations, from Priest [127, p. 4].

Example 2 The initial list for $O\neg p, O\neg q \,/\, O\neg(p \vee q)$ is:

$$O\neg p \ 0$$
$$O\neg q \ 0$$
$$\neg O\neg(p \vee q) \ 0$$

Example 3 The initial list for $Op, \neg Op \,/\, Op$ is:

$$Op \ 0$$
$$\neg Op \ 0$$

Definition 7 (closed branch, open branch) A branch is *closed* iff nodes of the forms $A\ x$ and $\neg A\ x$ occur on it; otherwise it is *open*.

We now specify the tableau rules for the D systems. (For easy reference, these rules are collected in undiluted form in Appendix B.)

In general, there are two types of tableau rules: *non-branching* (or conjunctive) rules, and *branching* (or disjunctive) rules.

Definition 8 (non-branching tableau rule) A non-branching rule is of the form:

$$\begin{array}{c} \mathcal{P}_1 \\ \vdots \\ \mathcal{P}_n \\ \downarrow \\ \mathcal{S}_1 \\ \vdots \\ \mathcal{S}_m \end{array}$$

$\mathcal{P}_1, \ldots, \mathcal{P}_n$ are the *predecessor* nodes. $\mathcal{S}_1, \ldots, \mathcal{S}_m$ are the *successor* nodes. Such a rule says that if nodes of the forms $\mathcal{P}_1, \ldots, \mathcal{P}_n$ occur on a branch, then all nodes of the forms $\mathcal{S}_1, \ldots, \mathcal{S}_m$ *that do not already occur on that branch* may be added to the tip of that branch. (The stipulation that only nodes that don't already occur on the branch may be added ensures that branches will always never be redundant, i.e., no node will ever occur on a branch more than once.)

Definition 9 (branching tableau rule) A branching tableau rule is of the form:

$$\begin{array}{ccc} & \mathcal{P}_1 & \\ & \vdots & \\ & \mathcal{P}_n & \\ \swarrow & & \searrow \\ \mathcal{L}_1 & & \mathcal{R}_1 \\ \vdots & & \vdots \\ \mathcal{L}_m & & \mathcal{R}_k \end{array}$$

$\mathcal{P}_1, \ldots, \mathcal{P}_n$ are the predecessor nodes. $\mathcal{L}_1, \ldots, \mathcal{L}_m$ are the *left* successor nodes. $\mathcal{R}_1, \ldots, \mathcal{R}_k$ are the *right* successor nodes. Such a rule says that if nodes of the forms $\mathcal{P}_1, \ldots, \mathcal{P}_n$ occur on a branch, then the branch may be split at the tip, with all nodes of the forms $\mathcal{L}_1, \ldots, \mathcal{L}_m$ that do not already occur on the (newly generated) left branch being added to the (newly generated) left branch, and all nodes of the forms $\mathcal{R}_1, \ldots, \mathcal{R}_k$ that do not already occur on the (newly generated) right branch being added to the (newly generated) right branch. (Again, the stipulation that only nodes that don't already occur on the branch may be added ensures that branches will always never be redundant.)

A tableau rule may be applied to a given *set of predecessor nodes* only once. (Note: this is not to say that a given rule may not be applied to a given *node* more than once—see, for example, the $[O]$ rule below.) Moreover, we do not allow "null" applications of tableau rules—i.e., applications that do not result in any change to the tableau.

Definition 10 The double negation rule, $[\neg\neg]$, is as follows:

$$\begin{array}{c} {\scriptstyle [\neg\neg]} \\ \neg\neg A\ x \\ \downarrow \\ A\ x \end{array}$$

This rule says that if a node of the form $\neg\neg A\ x$ occurs on an *open* branch (where A is any formula of L_{D} and x is any natural number), and a node of the form $A\ x$ does not already occur on that branch, then a node of the form $A\ x$ may be added to the tip of that branch. For example, if $\neg\neg(p \vee q)\ 0$ occurs on a branch, and $p \vee q\ 0$ does not already occur on that branch, then $p \vee q\ 0$ may be added to the tip of that branch.

Definition 11 The rule for disjunction, $[\vee]$, is as follows:

$$
\begin{array}{c}
\scriptstyle [\vee] \\
A \vee B\ x \\
\swarrow \searrow \\
A\ x \qquad B\ x
\end{array}
$$

This rule says that if a node of the form $A \vee B\ x$ occurs on an open branch (where A and B are any formulas of L_{D} and x is any natural number), then the branch may be split at the tip, with $A\ x$ being added to one side (if it does not already occur on the branch) and $B\ x$ being added to the other (if it does not already occur on the branch). For example, if $(p \wedge q) \vee \neg r\ 1$ occurs on a branch, and neither $p \wedge q\ 1$ nor $\neg r\ 1$ already occurs on the branch, then the branch may be split at the tip, with $p \wedge q\ 1$ being added to the left side and $\neg r\ 1$ to the right branch. Another example: if $(p \wedge q) \vee \neg r\ 1$ occurs on a branch, and $p \wedge q\ 1$ occurs on that branch but $\neg r\ 1$ does not occurs on the branch, then the branch may be split at the tip, with no node being being added to the left branch and $\neg r\ 1$ bveing added to the right branch.

Definition 12 The rule for negated disjunction, $[\neg\vee]$, is as follows:

$$
\begin{array}{c}
\scriptstyle [\neg\vee] \\
\neg(A \vee B)\ x \\
\downarrow \\
\neg A\ x \\
\neg B\ x
\end{array}
$$

This rule says that if a node of the form $\neg(A \vee B)\ x$ occurs on an open branch, then a nodes of the forms $\neg A\ x$ and $\neg B\ x$ may be added to the tip of that branch, provided they do not already occur on the branch. For example, if $\neg(p \vee \neg q)\ 2$ occurs on a branch and neither $\neg p\ 2$ nor $\neg\neg q\ 2$ occurs on the branch, then $\neg p\ 2$ and $\neg\neg q\ 2$ may be added to the tip of that branch.

Definition 13 The rule for obligation, $[O]$, is as follows:

$$
\begin{array}{c}
\scriptstyle [O] \\
OA\ x \\
x \triangleright y \\
\downarrow \\
A\ y
\end{array}
$$

This rule says that if nodes of the forms $OA\ x$ and $x \triangleright y$ occur on an open branch (where A is any formula of L_{D} and x and y are any natural numbers), then a node of the form $A\ y$ may be added to the the tip of that branch (provided it does not already occur on that branch). For example, if $O(p \vee q)\ 0$ and $0 \triangleright 1$ occur on a branch, and $p \vee q\ 1$ does not already occur on that branch, then $p \vee q\ 1$ may be added to the tip of the branch. Unlike the previous rules, this rule may be applied to a given node more than once (and thus nodes cannot be checked off as "done" when this rule is applied to them). For example, if $Op\ 1$, $1 \triangleright 2$, and $1 \triangleright 3$ occur on a branch then we can apply $[O]$ to $Op\ 1$ twice: first to $Op\ 1$ and $1 \triangleright 2$ to yield $p\ 2$; then to $Op\ 1$ and $1 \triangleright 3$ to yield $p\ 3$.

Definition 14 The rule for negated O is as follows:

$$\begin{array}{c} {\scriptstyle [\neg O]} \\ \neg OA\ x \\ \downarrow \\ x \triangleright i \\ \neg A\ i \end{array}$$

This rule says that if a node of the form $\neg OA\ x$ occurs on an open branch, then nodes of the forms $x \triangleright i$ and $\neg A\ i$ may be added to the tip of that branch, provided that i is a "fresh" natural number, i.e. one not already occurring on the branch. For example, if $\neg O \neg q\ 0$ occurs on a branch, and 1 does not occur on that branch, then $0 \triangleright 1$ and $\neg\neg q\ 1$ may be added to the tip of that branch. This rule may be applied to a given node only once.

The system D includes just the five tableau rules specified above. Ds, Dh, and Dsh contain additional rules.

Definition 15 Ds and Dsh contain the rule for seriality:

$$\begin{array}{c} {\scriptstyle [s]} \\ OA\ x \\ \boxed{x \triangleright y} \\ \downarrow \\ x \triangleright i \\ A\ i \end{array}$$

This rule says that if a node of the form $OA\ x$ occurs on a branch, and no node of the form $x \rhd y$ occurs on the branch (a $\boxed{\text{box}}$ indicates a node's *not* occurring on the branch), then nodes of the form $x \rhd i$ and $A\ i$ may be added to the tip of the branch, provided i is fresh. For example, if $OA\ 1$ occurs on a branch, no node of the form $1 \rhd x$ occurs on the branch, and the natural number 2 does not occur on the branch, then $1 \rhd 2$ and $A\ 2$ may be added to the tip of that branch.

Definition 16 Dh and Dsh contain the rule for shift-reflexivity:

$$\begin{array}{c} {\scriptstyle [h]} \\ x \rhd y \\ \downarrow \\ y \rhd y \end{array}$$

This rule says that if a node of the form $x \rhd y$ occurs on an open branch, then a node of the form $y \rhd y$ may be added to the tip of that branch (provided it does not already occur on the branch). For example, if $0 \rhd 1$ occurs on a branch, then $1 \rhd 1$ may be added to the tip of that branch.

Remark 4 (order of application) The tableau rules may be applied in any order. However, it is a good idea to apply non-branching rules first, in order to keep tableaus as compact as possible.

Definition 17 (tableau) A D+ tableau for the inference Γ / A is any tableau that results from 0 or more applications of D+ tableau rules to the initial list for Γ / A.

Definition 18 (redundant branch/tableau) A branch of a tableau is *redundant* iff some node occurs on it more than once. (For the purposes of this definition, nodes are individuated solely by their typographical content, and not their location on the branch. Otherwise, it would be trivial that no branch is redundant.) A tableau is redundant iff some of its branches are.

Example 4 The following is an example of a redundant branch:[19]

$$\begin{array}{c} p\ 0 \\ q\ 0 \\ p\ 0 \end{array}$$

[19] Note that, technically, this is not a branch of any D+ tableau.

Example 5 The following is an example of a non-redundant branch:

$$p\ 0$$
$$q\ 0$$
$$p\ 1$$

Lemma 1 (non-redundancy) No D+ tableau is redundant.

Proof. A simple induction on the complexity of tableaus. Let Γ / A be any inference. First, observe that the initial list for Γ / A is non-redundant. (This is ensured by Definition 6, p. 59.) Next, observe that application of a tableau rule to a branch will never create a redundant extension of that branch. (This is ensured by Definitions 8 and 9, p. 60.) It follows, by Definition 17, that a D+ tableau cannot be redundant. ■

Definition 19 (complete branch/tableau) A branch of a D+ tableau is *complete* just in case it is closed or each of the following holds:

1. For each node of the form $\neg\neg A\ x$ on the branch, a node of the form $A\ x$ is on the branch.
2. For each node of the form $A \vee B\ x$ on the branch, a node of the form $A\ x$ or $B\ x$ is on the branch.
3. For each node of the form $\neg(A \vee B)\ x$ on the branch, nodes of the forms $\neg A\ x$ and $\neg B\ x$ are on the branch.
4. For each pair of nodes of the forms $OA\ x$ and $x \triangleright y$ on the branch, a node of the form $A\ y$ is on the branch.
5. For each node of the form $\neg OA\ x$ on the branch, nodes of the forms $x \triangleright i$ and $\neg A\ i$ are on the branch.
6. If D+ is a serial system, for each node of the form $OA\ x$ on the branch, there are nodes of the forms $x \triangleright i$ and $A\ i$ on the branch.
7. If D+ is a shift-reflexive system, for each node of the form $x \triangleright y$ $(x \neq y)$ on the branch, a node of the form $y \triangleright y$ is on the branch.

A tableau is complete iff all of its branches are complete.

Remark 5 Loosely speaking, a branch is complete just in case no further rules can be applied to it.

Definition 20 (closed/open tableau) A tableau is closed just in case all of its branches are closed; otherwise it is open.

Remark 6 We will later prove the "all or nothing" principle (Theorem 3, p. 82), which says that if there is an open, complete tableau for an inference, then there is no closed tableau for that inference.

Definition 21 (tableau-theoretic consequence) A is a tableau-theoretic consequence of Γ in $\mathsf{D}+$ (in symbols, $\Gamma \Vdash_{\mathsf{D}+} A$) just in case there is a closed $\mathsf{D}+$ tableau for $\Gamma \diagup A$.

At this point, it may be helpful to provide an intuitive explanation of our tableaus. A tableau is essentially a method of trying every possible way of jointly satisfying a set of formulas. Each branch represents an attempt to construct a model in which the formulas in the initial list are jointly true. If all of these attempts fail (i.e. all branches close), then we have shown that that the initial list is unsatisfiable. If the initial list was $\Gamma \cup \{\neg A\}$, this means that $\Gamma \models A$. The natural numbers on a branch represent the worlds in the model we are attempting to construct on that branch. A node of the form $A\ x$ indicates that (on the branch we are assuming) A is true at w_x, the world corresponding to x. A node of the form $x \triangleright y$ indicates that the world corresponding to x sees the world corresponding to y, i.e. w_x "sees" w_y. If we have applied all rules that we can to a branch and it has not closed, then we have succeeded in constructing a model satisfying the formulas in our initial list. We can verify that we have succeeded by checking the counterexample model that is "induced" by this open branch. (This will be clarified shortly.)

It will be convenient to have the following derived rules at our disposal. (Note that these are not officially included in our proof theories.)

$$\begin{array}{c} [\wedge] \\ A \wedge B\ x \\ \downarrow \\ A\ x \\ B\ x \end{array} \qquad \begin{array}{c} [\neg\wedge] \\ \neg(A \wedge B)\ x \\ \downarrow \\ \neg A \vee \neg B\ x \end{array} \qquad \begin{array}{c} [\neg\vee]' \\ \neg(A \vee B)\ x \\ \downarrow \\ \neg A \wedge \neg B\ x \end{array} \qquad \begin{array}{c} [\supset] \\ A \supset B\ x \\ \swarrow \quad \searrow \\ \neg A\ x \qquad B\ x \end{array}$$

$$\begin{array}{cccc} [\neg\supset] & [P] & [\neg P] & [\neg O]' \\ \neg(A \supset B)\ x & PA\ x & \neg PA\ x & \neg OA\ x \\ \downarrow & \downarrow & \downarrow & \downarrow \\ A\ x & x \triangleright i & O\neg A\ x & P\neg A\ x \\ \neg B\ x & A\ i & & \end{array}$$

It is easy to check that these rules really are superfluous. The rule $[\wedge]$, for example, tells us that if $A \wedge B\ x$ occurs on a branch, then $A\ x$ and $B\ x$ may be added to the tip of that branch. This is just what we would *eventually* get were we to decompose $\neg(\neg A \vee \neg B)\ x$ using the official rules above:

$$\begin{array}{c} \neg(\neg A \vee \neg B)\ x \\ \neg\neg A\ x \\ \neg\neg B\ x \\ A\ x \\ B\ x \end{array}$$

Note that by omitting $\neg\neg A\ x$ and $\neg\neg B\ x$ we do not miss any opportunities to close the branch.

2.5.3 Example proofs

I now provide some examples illustrating our tableau theories in practice. When I am done with a node, I check it off with '✓'.

Example 6 (deontic explosion) Here is a proof that $OA, FA \Vdash_{\mathsf{D}} OB$, using only primitive rules:

1.	$OA\ 0$	initial list
2.	$O\neg A\ 0$	"
3.	$\neg OB\ 0$✓	"
4.	$0 \triangleright 1$	from 3 by $[\neg O]$
5.	$\neg B\ 1$	"
6.	$A\ 1$	from 1 and 4 by $[O]$
7.	$\neg A\ 1$	from 2 and 4 by $[O]$
	$*$	from 6 and 7

Remark 7 I have decorated the tableau to explain what's going on. The "official" tableau looks like this:

$$OA\ 0$$
$$O\neg A\ 0$$
$$\neg OB\ 0$$
$$0 \rhd 1$$
$$\neg B\ 1$$
$$A\ 1$$
$$\neg A\ 1$$

Example 7 (O-aggregation) Here is a proof that $OA, OB \Vdash_{\mathsf{D}} O(A \wedge B)$, using only primitive rules:

1.	$OA\ 0$		initial list
2.	$OB\ 0$		"
3.	$\neg O \neg(\neg A \vee \neg B)\ 0✓$		"
4.	$0 \rhd 1$		from 3 by $[\neg O]$
5.	$\neg\neg(\neg A \vee \neg B)\ 1✓$		"
6.	$\neg A \vee \neg B\ 1✓$		from 5 by $[\neg\neg]$
7.	$A\ 1$		from 1 and 4 by $[O]$
8.	$B\ 1$		from 2 and 4 by $[O]$
9.	$\neg A\ 1$	$\neg B\ 1$	from 6 by $[\vee]$
	$*$	$*$	(both branches close)

Example 8 (ought implies may) Here is a proof that $\Vdash_{\mathsf{Ds}} OA \supset PA$, using only primitive rules:

1.	$\neg(\neg OA \vee \neg O \neg A)\ 0✓$	initial list
2.	$\neg\neg OA\ 0✓$	from 1 by $[\neg\vee]$
3.	$\neg\neg O \neg A\ 0✓$	"
4.	$OA\ 0$	from 2 by $[\neg\neg]$
5.	$O\neg A\ 0$	from 3 by $[\neg\neg]$
6.	$0 \rhd 1$	from 4 by $[s]$
7.	$A\ 1$	"
8.	$\neg A\ 1$	from 5 and 6 by $[O]$
	$*$	from 7 and 8

Example 9 (ought implies is) Here is a proof that $\nvdash_{\mathsf{Dsh}} Op \supset p$, using primitive and derived rules:

$$
\begin{array}{lcl}
1. & \neg(Op \supset p)\ 0\checkmark & \text{initial list} \\
2. & Op\ 0 & \text{from 2 by } [\neg \supset] \\
3. & \neg p\ 0 & \text{"} \\
4. & 0 \rhd 1 & \text{from 2 by } [s] \\
5. & p\ 1 & \text{"} \\
6. & 1 \rhd 1 & \text{from 4 by } [h] \\
 & \uparrow &
\end{array}
$$

We can read a (partially specified) countermodel off of the (infinite) open branch as follows. (This procedure is made more explicit in Definition 24 on p. 73.) Let $W = \{w_i : i \text{ is a natural number occurring on the branch}\} = \{w_0, w_1\}$. Let $R = \{\langle w_i, w_j\rangle : i \rhd j \text{ is on the branch}\} = \{\langle w_0, w_1\rangle, \langle w_1, w_1\rangle\}$. Since $\neg p$ 0 is on the branch, let $v_{w_0}(p) = \mathbf{f}$. Since p 1 is on the branch, let $v_{w_1}(p) = \mathbf{t}$. This model can be depicted as follows:

$$
\begin{tikzcd}
\boxed{w_0} \arrow[r] & \boxed{w_1} \arrow[loop above] \\
\neg p & p
\end{tikzcd}
$$

It is easy to check that on this model, $\bar{v}_{w_0}(Op \supset p) = \mathbf{f}$. Thus $\not\models_{\mathsf{Dsh}} Op \supset p$.

Example 10 (free choice permission) Here is a proof that $P(p \vee q) \nvdash_{\mathsf{D}} Pp \wedge Pq$, using primitive and derived rules:

$$
\begin{array}{c}
P(p \vee q)\ 0\checkmark \\
\neg(Pp \wedge Pq)\ 0\checkmark \\
\neg Pp \vee \neg Pq\ 0\checkmark \\
0 \rhd 1 \\
p \vee q\ 1\checkmark \\
\swarrow \quad \searrow \\
\begin{array}{cc}
\begin{array}{c}
\neg Pp\ 0\checkmark \\
O\neg p\ 0 \\
\neg p\ 1 \\
\swarrow \quad \searrow \\
\begin{array}{cc} p\ 1 & q\ 1 \\ * & \uparrow \end{array}
\end{array}
&
\begin{array}{c}
\neg Pq\ 0\checkmark \\
O\neg q\ 0 \\
\neg q\ 1 \\
\swarrow \quad \searrow \\
\begin{array}{cc} p\ 1 & q\ 1 \\ \uparrow & * \end{array}
\end{array}
\end{array}
\end{array}
$$

There are two open branches. Let's read a counterexample off of the one on the left. Let $W = \{w_0, w_1\}$, $R = \{\langle w_0, w_1\rangle\}$, $v_{w_1}(p) = \mathbf{f}$, $v_{w_1}(q) = \mathbf{t}$. It is easy to check that on this model, $\bar{v}_{w_0}(P(p \vee q)) = \mathbf{t}$ and $\bar{v}_{w_0}(Pp \wedge Pq) = \mathbf{f}$. Thus $P(p \vee q) \not\models_{\mathsf{D}} Pp \wedge Pq$.

2.5.4 Soundness and completeness

In this section I prove that our tableau theories are sound and complete with respect to their corresponding semantics.

Definition 22 (faithful) Let b be a branch of a $\mathsf{D}+$ tableau. A $\mathsf{D}+$ model $\mathfrak{M} = \langle W, R, v\rangle$ is *faithful* to b iff there is a "worldification" function $w : \mathbb{N} \longrightarrow W$ such that:[20]

- if $A\ x$ is on b, then $\bar{v}_{w_x}(A) = \mathbf{t}$
- if $x \triangleright y$ is on b, then $w_x R w_y$.

We will say that w *shows* $\mathfrak{M}$ to be faithful to b.

Lemma 2 (faith lemma) If a branch, b, of a $\mathsf{D}+$ tableau is closed, then no $\mathsf{D}+$ model is faithful to b.

Proof. Suppose that, contrary to what we wish to show, some $\mathsf{D}+$ model $\langle W, R, v\rangle$ is faithful to b . Since b is closed, nodes of the forms $B\ x$ and $\neg B\ x$ must occur on it. Thus, by Definition 22, $\bar{v}_{w_x}(B) = \bar{v}_{w_x}(\neg B) = \mathbf{t}$, which is impossible. ■

Notation 2 If b is a branch of a tableau, $b(\mathcal{N}_1, \cdots, \mathcal{N}_i)$ is the branch that results from adding the nodes $\mathcal{N}_1, \cdots, \mathcal{N}_i$ to the tip of b.

Lemma 3 (soundness lemma) If a $\mathsf{D}+$ model $\mathfrak{M} = \langle W, R, v\rangle$ is faithful to a branch of a $\mathsf{D}+$ tableau, b, and a $\mathsf{D}+$ tableau rule is applied to b, then $\mathfrak{M}$ is faithful to at least one of the branches thereby generated.

Proof. The proof is by cases. There are seven cases—one for each tableau rule.

- **Case 1** ($[\neg\neg]$). Suppose $\neg\neg A\ x$ is on b, and $[\neg\neg]$ is applied. Then $b(A\ x)$ is generated. Since $\mathfrak{M}$ is faithful to b, $\bar{v}_{w_x}(\neg\neg A) = \mathbf{t}$ (for some worldification function w) Thus $\bar{v}_{w_x}(\neg A) = \mathbf{f}$. Thus $\bar{v}_{w_x}(A) = \mathbf{t}$. Thus w shows $\mathfrak{M}$ to be faithful to $b(A\ x)$.

[20] I write $w(x)$ as w_x.

- **Case 2** ($[\vee]$). Suppose $A \vee B\ x$ is on b, and $[\vee]$ is applied. Then $b(A\ x)$ and $b(B\ x)$ are generated. Since $\mathfrak{M}$ is faithful to b, $\bar{v}_{w_x}(A \vee B) = \mathbf{t}$ (for some w) Thus either $\bar{v}_{w_x}(A) = \mathbf{t}$ or $\bar{v}_{w_x}(B) = \mathbf{t}$. Thus w shows $\mathfrak{M}$ to be faithful to either $b(A\ x)$ or $b(B\ x)$.

- **Case 3** ($[\neg\vee]$). Suppose $\neg(A \vee B)\ x$ is on b, and $[\neg\vee]$ is applied. Then $b(\neg A\ x, \neg B\ x)$ is generated. Since $\mathfrak{M}$ is faithful to b, $\bar{v}_{w_x}(\neg(A \vee B)) = \mathbf{t}$ (for some w) Thus $\bar{v}_{w_x}(A \vee B) = \mathbf{f}$. Thus $\bar{v}_{w_x}(A) = \bar{v}_{w_x}(B) = \mathbf{f}$. Thus $\bar{v}_{w_x}(\neg A) = \bar{v}_{w_x}(\neg B) = \mathbf{t}$. Thus w shows $\mathfrak{M}$ to be faithful to $b(\neg A\ x, \neg B\ x)$.

- **Case 4** ($[O]$). Suppose $OA\ x$ and $x \triangleright y$ are on b, and $[O]$ is applied. Then $b(A\ y)$ is generated. Since $\mathfrak{M}$ is faithful to b, $\bar{v}_{w_x}(OA) = \mathbf{t}$ and $w_x R w_y$ (for some w). Thus $\forall u(w_x R u \Rightarrow \bar{v}_u(A) = \mathbf{t})$. Thus $v_{w_y}(A) = \mathbf{t}$. Thus w shows $\mathfrak{M}$ to be faithful to $b(A\ y)$.

- **Case 5** ($[\neg O]$). Suppose $\neg OA\ x$ is on b, and $[\neg O]$ is applied. Then $b(x \triangleright i, \neg A\ i)$ is generated. Since $\mathfrak{M}$ is faithful to b, $\bar{v}_{w_x}(\neg OA) = \mathbf{t}$ (for some w). Thus $\bar{v}_{w_x}(OA) = \mathbf{f}$. Thus there is a world, call it \$, such that $w_x R\$$ and $\bar{v}_{\$}(A) = \mathbf{f}$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Moreover, $\hat{w}_x R \hat{w}_i$ and $\bar{v}_{\hat{w}_i}(A) = \mathbf{f}$. Thus $\bar{v}_{\hat{w}_i}(\neg A) = \mathbf{t}$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, \neg A\ i)$.

- **Case 6** ($[s]$). Suppose $OA\ x$ occurs on b, no node of the form $x \triangleright y$ occurs on b, and $[s]$ is applied. Then $b(x \triangleright i, A\ i)$ is generated. Since $[s]$ is a rule only for our serial systems, R is serial. Thus there is a world, call it \$, such that $w_x R\$$ (for some w). Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Thus $\bar{v}_{\hat{w}_x}(OA) = \mathbf{t}$. Thus $\forall u(\hat{w}_x R u \Rightarrow \bar{v}_u(A) = \mathbf{t})$. Thus, since $\hat{w}_x R \hat{w}_i$, $v_{\hat{w}_i}(A) = \mathbf{t}$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, A\ i)$.

- **Case 7** ($[h]$). Suppose $x \triangleright y$ occurs on b, and $[h]$ is applied. Then $b(y \triangleright y)$ is generated. Since $\mathfrak{M}$ is faithful to b, $w_x R w_y$ (for some w). Since $[h]$ is a rule only for our shift-reflextive systems, R is shift-reflexive. Thus $w_y R w_y$. Thus w shows $\mathfrak{M}$ to be faithful to $b(y \triangleright y)$.

■

Theorem 1 (soundness) If $\Gamma \Vdash_{\mathsf{D+}} A$, then $\Gamma \models_{\mathsf{D+}} A$.

Proof. We prove the contrapositive. Suppose $\Gamma \not\models_{\mathsf{D+}} A$. Then there is a $\mathsf{D+}$ model $\mathfrak{M} = \langle W, R, v\rangle$ and $w \in W$ such that $\bar{v}_w(B) = \mathbf{t}$ for all $B \in \Gamma$, but $\bar{v}_w(A) \neq \mathbf{t}$. Thus $\bar{v}_w(\neg A) = \mathbf{t}$. Thus $\mathfrak{M}$ is faithful to the initial list for $\Gamma \,/\, A$. Moreover, by the Soundness Lemma, each subsequent application of a $\mathsf{D+}$ tableau rule will yield at least one branch to which $\mathfrak{M}$ is faithful. Thus every $\mathsf{D+}$ tableau for $\Gamma \,/\, A$ has at least one branch to which $\mathfrak{M}$ is faithful. Let T be a $\mathsf{D+}$ tableau for $\Gamma \,/\, A$, and let b be one of the branches of T to which $\mathfrak{M}$ is faithful. By the Faith Lemma, b cannot be closed. Thus T is open. But T was an arbitrarily chosen $\mathsf{D+}$ tableau for $\Gamma \,/\, A$. Thus there is no closed $\mathsf{D+}$ tableau for $\Gamma \,/\, A$. Thus $\Gamma \not\vdash_{\mathsf{D+}} A$. ■

Definition 23 (induced model for non-serial systems) Let $\mathsf{D+} \in \{\mathsf{D}, \mathsf{Dh}\}$. Let b be an open, complete branch of a $\mathsf{D+}$ tableau. The $\mathsf{D+}$ model *induced* by b is the $\mathsf{D+}$ model $\langle W, R, v\rangle$ such that:

1. $W = \{w_i : i \text{ is a natural number on } b\}$
2. $R = \{\langle w_i, w_j\rangle : i \rhd j \text{ is on } b\}$
3. For all $p \in At$:
   (a) if $p\ x$ is on b, then $v_{w_x}(p) = \mathbf{t}$
   (b) if $\neg p\ x$ is on b, then $v_{w_x}(p) = \mathbf{f}$
   (c) if neither $p\ x$ nor $\neg p\ x$ is on b, let $v_{w_x}(p) = \mathbf{f}$.[21]

We need to show that $\langle W, R, v\rangle$, so defined, really is a $\mathsf{D+}$ model. In particular, we need to verify that if $\mathsf{D+}$ is a *shift-reflexive* system, then R is indeed shift-reflexive. Since b is complete, for each node of the form $x \rhd y$ $(x \neq y)$ occurring on b, there is a node of the form $y \rhd y$ on b. Thus for each w_x and $w_y \in W$ such that w_xRw_y, w_yRw_y.; that is, R is shift-reflexive. Next, we need to confirm that for all $p \in At$, if $p\ x$ is on b, then $v_{w_x}(p) = \mathbf{t}$, and if $\neg p\ x$ is on b, then $v_{w_x}(p) = \mathbf{f}$. This is ensured by the way we have defined our structure. Finally, note that v really is a *function.* Our definition will never require (*per impossible*) that $v_{w_x}(p) = \mathbf{t}$ and $v_{w_x}(p) = \mathbf{f}$, since if $p\ x$ and $\neg p\ x$ were both on b, then b would be closed (contrary to what we are assuming).

[21] This part of the assignment is arbitrary. (We could just as well have assigned such atomic formulas the truth value $\mathbf{t}$.) We include it only so that we can refer to "*the*" model induced by a branch.

Definition 24 (induced model for serial systems) Let $\mathsf{D+} \in \{\mathsf{Ds}, \mathsf{Dsh}\}$. Let b be an open, complete branch of a $\mathsf{D+}$ tableau. The $\mathsf{D+}$ model *induced* by b is the $\mathsf{D+}$ model $\langle W, R, v\rangle$ such that:

1. $W = \{w_i : i \text{ is a natural number on } b\}$
2. $R = \{\langle w_i, w_j\rangle : i \triangleright j \text{ is on } b\} \cup \{\langle w_i, w_i\rangle : i$ is on b but no node of the form $i \triangleright j$ is on $b\}$
3. For all $p \in At$:
   (a) if $p\ x$ is on b, then $v_{w_x}(p) = \mathbf{t}$
   (b) if $\neg p\ x$ is on b, then $v_{w_x}(p) = \mathbf{f}$
   (c) if neither $p\ x$ nor $\neg p\ x$ is on b, let $v_{w_x}(p) = \mathbf{f}$.

We need to show that $\langle W, R, v\rangle$, so defined, really is a $\mathsf{D+}$ model. First, we need to verify that R is indeed serial. That is, we need to show that for all $w \in W$, there is a $u \in W$ such that wRu. By clause 1 above, $W = \{w_i : i$ is a natural number on $b\}$. Thus we need to show that for each natural number i occurring on b, there is a $u \in W$ such that w_iRu. Now, for each i occurring on b, there are just two possible cases:

- *Case 1.* A node of the form $i \triangleright j$ occurs on b. Then, by (the first part of) clause 2, w_iRw_j.
- *Case 2.* No node of the form $i \triangleright j$ occurs on b. Then, by (the second part of) clause 2, w_iRw_i.

In both cases, there is a $u \in W$ such that w_iRu. Thus for all $w \in W$, there is a $u \in W$ such that wRu; i.e., R is serial. Second, we need to verify that if $\mathsf{D+} = \mathsf{Dsh}$, then R is shift-reflexive. That is, we need to show that for all natural numbers i, j occurring on b, if w_iRw_j, then w_jRw_j. If $i = j$, the result follows immediately. So suppose $i \neq j$. Then $i \triangleright j$ must occur on the branch. (For suppose it doesn't. Then, by clause 2, $i = j$, contadicting our assumption.) Thus, since b is complete, $j \triangleright j$ occurs on the branch. Thus, by clause 2, w_jRw_j, as required. Next, we need to confirm that for all $p \in At$, if $p\ x$ is on b, then $v_{w_x}(p) = \mathbf{t}$, and if $\neg p\ x$ is on b, then $v_{w_x}(p) = \mathbf{f}$. This is ensured by the way we have defined our structure. Finally, note that v really is a *function*. Our definition will never require (*per impossible*) that $v_{w_x}(p) = \mathbf{t}$ and $v_{w_x}(p) = \mathbf{f}$, since if $p\ x$ and $\neg p\ x$ were both on b, then b would be closed (contrary to what we are assuming).

Lemma 4 (completeness lemma for non-serial systems) Let $\mathsf{D+} \in \{\mathsf{D}, \mathsf{Dh}\}$. Let b be an open, complete branch of a $\mathsf{D+}$ tableau, and let $\langle W, R, v\rangle$ be the $\mathsf{D+}$ model induced by b. Then for all $A \in L_{\mathsf{D+}}$:

1. if $A\ x$ is on b, then $\bar{v}_{w_x}(A) = \mathbf{t}$;

2. if $\neg A\ x$ is on b, then $\bar{v}_{w_x}(A) = \mathbf{f}$.

Proof. An induction on the length of A. If A is atomic, the result follows immediately. The result needs to be shown for $A = \neg B$, $A = (B \vee C)$, and $A = OB$.

Negation. Suppose $\neg B\ x$ is on b. By the inductive hypothesis (IH), $\bar{v}_{w_x}(B) = \mathbf{f}$. Thus $\bar{v}_{w_x}(\neg B) = \mathbf{t}$. Now suppose $\neg\neg B\ x$ is on b. Since b is complete, $B\ x$ is on b. Thus, by the IH, $\bar{v}_{w_x}(B) = \mathbf{t}$. Thus $\bar{v}_{w_x}(\neg B) = \mathbf{f}$.

Disjunction. Suppose $B \vee C\ x$ is on b. Since b is complete, either $B\ x$ or $C\ x$ is on b. Thus, by the IH, either $\bar{v}_{w_x}(B) = \mathbf{t}$ or $\bar{v}_{w_x}(C) = \mathbf{t}$. Thus $\bar{v}_{w_x}(B \vee C) = \mathbf{t}$. Now suppose $\neg(B \vee C)\ x$ is on b. Since b is complete, $\neg B\ x$ and $\neg C\ x$ are on b. Thus, by the IH, $\bar{v}_{w_x}(B) = \mathbf{f}$ and $\bar{v}_{w_x}(C) = \mathbf{f}$. Thus $\bar{v}_{w_x}(B \vee C) = \mathbf{f}$.

Obligation. Suppose $OB\ x$ is on b. Since b is complete, $B\ y$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $\bar{v}_{w_y}(B) = \mathbf{t}$ for all y such that $x \triangleright y$ is on b. Thus, by Def. 23, $\bar{v}_u(B) = \mathbf{t}$ for all u such that $w_x R u$. Thus $\bar{v}_{w_x}(OB) = \mathbf{t}$. Now suppose $\neg OB$ x is on b. Since b is complete, $x \triangleright i$ and $\neg B\ i$ are on b. Thus, by the IH, $\bar{v}_{w_i}(B) = \mathbf{f}$. Thus, by Def. 23, $\bar{v}_u(B) = \mathbf{f}$ for some u such that $w_x R u$. Thus $\bar{v}_{w_x}(OB) = \mathbf{f}$. ■

Lemma 5 (completeness lemma for serial systems) Let $\mathsf{D+} \in \{\mathsf{Ds}, \mathsf{Dsh}\}$. Let b be an open, complete branch of a $\mathsf{D+}$ tableau, and let $\langle W, R, v\rangle$ be the $\mathsf{D+}$ model induced by b. Then for all $A \in L_{\mathsf{D+}}$

1. if $A\ x$ is on b, then $\bar{v}_{w_x}(A) = \mathbf{t}$;

2. if $\neg A\ x$ is on b, then $\bar{v}_{w_x}(A) = \mathbf{f}$.

Proof. An induction on the length of A. If A is atomic, the result follows immediately. The result needs to be shown for $A = \neg B$, $A = (B \vee C)$, and $A = OB$. The proof for negation and disjunction is the same as in the previous lemma. The proof for obligation is a little different:

Obligation. Suppose $OB\ x$ is on b. Since b is complete, $B\ y$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $\bar{v}_{w_y}(B) = \mathbf{t}$ for all y such that $x \triangleright y$ is on b. Moreover, since b is complete, it cannot be the case that no node of the form $x \triangleright y$ is on b. Thus, by Def. 24, $\{w_i : w_x R w_i\} = \{w_i : x \triangleright i \text{ is on } b\}$. Thus, by Def. 24, $\bar{v}_u(B) = \mathbf{t}$ for all u such that $w_x R u$. Thus $\bar{v}_{w_x}(OB) = \mathbf{t}$. Now suppose $\neg OB\ x$ is on b. Since b is complete, $x \triangleright i$ and $\neg B\ i$ are on b. Thus, by the IH, $\bar{v}_{w_i}(B) = \mathbf{f}$. Thus, by Def. 23, $\bar{v}_u(B) = \mathbf{f}$ for some u such that $w_x R u$. Thus $\bar{v}_{w_x}(OB) = \mathbf{f}$. ■

Definition 25 (finitely generated tableau) A tableau is *finitely generated* iff each node on the tableau has only finitely many immediate successors (where the "immediate successors" of a node are the ones immediately beneath it on the tableau).

Fact 1 Each D+ tableau is finitely generated.

Proof. Obvious, since each application of a rule yields at most two immediate successors to any given node. ■

Definition 26 (finite tableau) A tableau is *finite* iff it contains finitely many nodes; otherwise, it is *infinite*.

Definition 27 (finite branch) A branch of a tableau is *finite* iff it contains finitely many nodes; otherwise, it is *infinite*.

Lemma 6 (König's Lemma) Any infinite, finitely generated tableau must have an infinite branch.

Proof. See, e.g., Fitting [55, pp. 406-7] or Smullyan [145, p. 32]. ■

Definition 28 (depth of an index) If i is an index (natural number) occurring on a branch, b, then the *depth* of i (with respect to b, where $i > 0$), written $depth(i, b)$, is the least number n such that there are nodes of the forms $0 \triangleright x_1, x_1 \triangleright x_2, ..., x_{n-1} \triangleright i$ on b. (Intuitively, these nodes form a "chain" leading from 0 to i. So we can think of $depth(i, b)$ as the minimum number of "accessibility steps" needed to get from i back to 0 on b.) We stipulate that the depth of 0 is always 0 (since, intuitively, it takes 0 steps to get from 0 to itself).

Example 11 Consider the following D tableau for the inference $\neg OOp, \neg OOq /$ $\neg OO(p \wedge q)$:

1.	$\neg OOp\ 0\checkmark$	initial list
2.	$\neg OOq\ 0\checkmark$	"
3.	$\neg\neg OO(p \wedge q)\ 0\checkmark$	"
4.	$OO(p \wedge q)\ 0$	from 3 by $[\neg\neg]$
5.	$0 \rhd 1$	from 1 by $[\neg O]$
6.	$\neg Op\ 1\checkmark$	"
7.	$0 \rhd 2$	from 2 by $[\neg O]$
8.	$\neg Oq\ 2\checkmark$	"
9.	$1 \rhd 3$	from 6 by $[\neg O]$
10.	$\neg p\ 3$	"
11.	$2 \rhd 4$	from 8 by $[\neg O]$
12.	$\neg q\ 4$	"
13.	$O(p \wedge q)\ 1$	from 4, 5 by $[O]$
14.	$O(p \wedge q)\ 2$	from 4, 7 by $[O]$
15.	$p \wedge q\ 3$	from 9, 13 by $[O]$
16.	$p\ 3$	from 15 by $[\wedge]$
17.	$q\ 3$	"
	$*$	(10,16)

Let b the sole branch of this tableau. Then, using fairly obvious notation,[22] $depth(0, b)$ $= 0$, $depth(1, b) = depth(2, b) = 1$, and $depth(3, b) = depth(4, b) = 2$.

Definition 29 (pseudo-subformula) B is a *pseudo-subformula* of A iff B is a subformula of A or the negation of a subformula of A. For example, the pseudo-subformulas of $p \vee q$ are $p \vee q$ itself, p, q, $\neg(p \vee q)$, $\neg p$,and $\neg q$.

Lemma 7 If a node of the form $B\ x$ occurs on a $\mathsf{D}+$ tableau for Γ / A, then B is a pseudo-subformula of one of the elements of $\Gamma \cup \{\neg A\}$.

Proof. An induction on the complexity of tableaus. The result holds trivially for the initial list for Γ / A. We now need to show that if the result holds for a branch b, and a $\mathsf{D}+$ tableau rule is applied to b, then the result holds for each of the branches thereby generated. There are seven cases to consider:

[22] $depth(n, b)$ denotes the depth of the index n on the branch b.

- **Case 1** ($[\neg\neg]$). Suppose $\neg\neg C\ x$ is on b, and $[\neg\neg]$ is applied. Then $b(C\ x)$ is generated. By the IH, $\neg\neg C$ is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$—let us say D. Since $\neg\neg C$ is a pseudo-subformula of D, either $\neg C$ or $\neg\neg C$ is a subformula of D. Either way, C is a subformula (and hence a pseudo-subformula) of D. Thus C is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$, namely D. Thus the result holds for $b(C\ x)$.

- **Case 2** ($[\vee]$). Suppose $C \vee D\ x$ is on b, and $[\vee]$ is applied. Then $b(C\ x)$ and $b(D\ x)$ are generated. By the IH, $C \vee D$ is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$—let us say E. Since $C \vee D$ is a pseudo-subformula of E, and since it is not the negation of any formula, it must be a subformula of E. Thus C and D are both subformulas (and hence pseudo-subformulas) of some element of $\Gamma \cup \{\neg A\}$, namely E. Thus the result holds for both $b(C\ x)$ and $b(D\ x)$.

- **Case 3** ($[\neg\vee]$). Suppose $\neg(C \vee D)\ x$ is on b, and $[\neg\vee]$ is applied. Then $b(\neg C\ x, \neg D\ x)$ is generated. By the IH, $\neg(C \vee D)$ is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$—let us say E. Since $\neg(C \vee D)$ is a pseudo-subformula of E, either $C \vee D$ or $\neg(C \vee D)$ is a subformula of E. Thus C and D are both subformulas of E. Thus $\neg C$ and $\neg D$ are pseudo-subformulas of E. Thus $\neg C$ and $\neg D$ are both pseudo-subformulas of some element of $\Gamma \cup \{\neg A\}$, namely E. Thus the result holds for $b(\neg C\ x, \neg D\ x)$.

- **Case 4** ($[O]$). Suppose $OC\ x$ and $x \triangleright y$ are on b, and $[O]$ is applied. Then $b(C\ y)$ is generated. By the IH, OC is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$—let us say D. Since OC is a pseudo-subformula of D, and since it is not the negation of any formula, it must be a subformula of D. Thus C is a subformula (and hence pseudo-subformula) of D. Thus C is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$, namely D. Thus the result holds for $b(C\ y)$.

- **Case 5** ($[\neg O]$). Suppose $\neg OC\ x$ is on b, and $[\neg O]$ is applied. Then $b(x \triangleright i, \neg C\ i)$ is generated. By the IH, $\neg OC$ is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$—let us say D. Since $\neg OC$ is a pseudo-subformula of D, either OC or $\neg OC$ is a subformula of D. Either way, C is a subformula of D. Thus $\neg C$ is a pseudo-subformula of D. Thus $\neg C$ is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$, namely D. Thus the result holds for $b(x \triangleright i, \neg C\ i)$.

- **Case 6** ($[s]$). Suppose $OC\ x$ occurs on b, no node of the form $x \triangleright y$ occurs on b, and $[s]$ is applied. Then $b(x \triangleright i, C\ i)$ is generated. By the IH, OC is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$—let us say D. Since OC is a pseudo-subformula of D, and since it is not the negation of any formula, it must be a subformula of D. Thus C is a subformula (and hence pseudo-subformula) of D. Thus C is a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$, namely D. Thus the result holds for $b(x \triangleright i, C\ i)$.

- **Case 7** ($[h]$). Suppose $x \triangleright y$ occurs on b, and $[h]$ is applied. Then $b(y \triangleright y)$ is generated. By the IH, the result holds for b. Thus, trivially, it also holds for $b(y \triangleright y)$.

■

Fact 2 Each formula of $L_{\mathsf{D}+}$ has only finitely many pseudo-subformulas.

Proof. Obvious. ■

Definition 30 (O-degree) The *O-degree* of a formula is simply the number of occurrences of 'O' in it. For example, the O-degree of $O(Op \wedge O\neg q)$ is 3.

Definition 31 (index-generating rule) An *index-generating* tableau rule is simply one that introduces a new index (natural number) to a branch.

Remark 8 Note that $[\neg O]$ and $[s]$ are the only index-generating rules in the D-systems.

Lemma 8 Let b be any branch of a D+ tableau. For all x, y, the node $x \triangleright y$ occurs on b only if nodes of the forms $C\ x$ and $D\ y$ occur on b.

Proof. If $x \triangleright y$ occurs on b, then it can only have been introduced via an application of $[\neg O]$, $[s]$, or $[h]$. If it was introduced via $[\neg O]$, then nodes of the forms $\neg OA\ x$ and $A\ y$ must occur on b. If it was introduced via $[s]$, then nodes of the forms $OA\ x$ and $A\ y$ must occur on b. If it was introduced via $[h]$, then $x = y$ and a node of the form $z \triangleright x$, where $z \neq x$, occurs on b. Since $z \neq x$, $z \triangleright x$ must have been introduced via $[\neg O]$ or $[s]$. Either way, a node of $A\ y$ must occur on b. Hence, since $x = y$, a node of the form $C\ x$ occurs on b and a node of the form $D\ y$ occurs on b. In each of the three possible cases, nodes of the forms $C\ x$ and $D\ y$ occur on b. ■

Lemma 9 (finiteness lemma) Every $\mathsf{D}+$ tableau for Γ/A is finite.

Proof. Suppose, for reductio, that some $\mathsf{D}+$ tableau for Γ/A is infinite. By König's Lemma, this tableau must have an infinite branch—say, b. There must be, on b, infinitely many nodes. By Lemma 1 (p. 65), all of these nodes must be different. If an index x occurs on b, then it can only occur in a node of the form $B\ x$, $x \triangleright y$, or $y \triangleright x$. By Lemma 7, if $B\ x$ is on b, then B must be a pseudo-subformula of some element of $\Gamma \cup \{\neg A\}$. By Fact 2, there are only finitely many of these. So there are only finitely many nodes of the form $B\ x$ on b. Moreover, by Lemma 8, $x \triangleright y$ occurs on b only if nodes of the forms $C\ x$ and $D\ y$ occur on b. We have already established that there are only finitely many nodes of the form $C\ x$ or $D\ y$ on b. Thus, since b is non-redundant, there can only be finitely many nodes of the form $x \triangleright y$ on b. By parity of reasoning, there can only be finitely many nodes of the form $y \triangleright x$ on b. Thus each index can occur on b only finitely many times. Thus, since b is infinite, infinitely many indices must occur on b. There are, then, two cases:

- **Case 1.** For some number, d, there are infinitely many indices on b of depth d. Let n be the smallest such number. The only index of depth 0 occurring on b is 0, and this occurs only finitely often. Hence $n > 0$. Further, the only way an index of depth $n > 0$ can be introduced to b is by an index-generating rule being applied to a node whose formula has an index of depth $n - 1$. And an index-generating rule can be applied to such a node only once. So if there are infinitely many indices of depth n on b, there must be infinitely many indices of depth $n - 1$ on b, contradicting our initial assumption about n.

- **Case 2.** For each number, d, there are only finitely many indices on b of depth d. Then, since infinitely many indices occur on b, infinitely many index depths must be represented on b. Now, suppose the index y occurs on b, where $y > 0$. This index must occur in a node of the form $C\ y$, $y \triangleright z$, or $z \triangleright y$. But by Lemma 8, $y \triangleright z$, or $z \triangleright y$ occurs on b only if a node of the form $C\ y$ occurs on b. So a node of the form $C\ y$ must occur on b. Now, this node must have been added by applying either $[\neg O]$ or $[s]$ to a node of the form $D\ x$, where D is of higher O-degree than C, or else it must follow by the non-deontic rules (i.e. $[\neg\neg]$, $[\vee]$, and $[\neg\vee]$) from such a formula. And non-deontic rules do not increase O-degree. Thus the maximum O-degree of all formulas on b with the index y is lower than the maximum O-degree of all formulas on b with index

x. (Intuitively, as the indices get "deeper," the O-degrees of their associated formulas get smaller.) Thus since the maximum O-degree of formulas with index 0 is finite, there is an index n such that no index of depth greater than n can have been introduced on b, contradicting our assumption that infinitely many index depths are represented on b.

Since both cases lead to a contradiction, we have shown that every D+ tableau for Γ / A is finite. ■

Definition 32 (extension of a tableau) A D+ tableau T' is an *extension* of a D+ tableau T iff T' can be constructed by applying a series of D+ tableau rules to T.

Lemma 10 (extension lemma) Suppose there is no closed D+ tableau for Γ / A. Then any D+ tableau for Γ / A has an extension (also a D+ tableau for Γ / A) containing an open, complete branch.

Proof. Suppose there is no closed D+ tableau for Γ / A. Now take any tableau for Γ / A, and use the following procedure to construct an extension of this tableau (which extension is also a D+ tableau for Γ / A) that contains an open, complete branch:

Step A. Pick an open branch of the tableau (there must be one). If this branch is complete, then stop. If not, go to Step B.

Step B. Since the branch is not complete, one of the following must hold:

1. $\neg\neg A\ x$ is on the branch, but $A\ x$ is not on the branch.

2. $A \vee B\ x$ is on the branch, but neither $A\ x$ nor $B\ x$ is on the branch.

3. $\neg(A \vee B)\ x$ is on the branch, but $\neg A\ x$ and $\neg B\ x$ are not both on the branch.

4. $OA\ x$ and $x \triangleright y$ are on the branch, but $A\ y$ is not on the branch.

5. $\neg OA\ x$ is on the branch, but $x \triangleright i$ and $\neg A\ i$ are not both on the branch.

6. D+ is a serial system, and $OA\ x$ is on the branch, but $x \triangleright i$ and $A\ i$ are not both on the branch.

7. D+ is a shift-reflexive system, and $x \triangleright y$ $(x \neq y)$ is on the branch, but $y \triangleright y$ is not on the branch.

If (1) does not hold, go to Step C1. If (1) holds, apply [$\neg\neg$] to the appropriate node (the first such node, if there is more than one). If the resulting branch is closed, return to Step A. If it open and incomplete, return to Step B. If it is open and complete, then stop.

Step C1. If (2) does not hold, go to Step D. If (2) holds, apply [$\vee$] to the appropriate node (the first such node, if there is more than one). Consider the leftmost of the two resulting branches. If this branch is closed or incomplete, go to Step C2. If it is open and complete, then stop.

Step C2. Consider the rightmost of the two resulting branches. If this branch is closed or incomplete, return to Step A. If it is open and complete, then stop.

Step D. If (3) does not hold, go to Step E. If (3) holds, apply [$\neg\vee$] to the appropriate node (the first such node, if there is more than one). If the resulting branch is closed, return to Step A. If it open and incomplete, return to Step B. If it is open and complete, then stop.

Step E. If (4) does not hold, go to Step F. If (4) holds, apply [O] to the appropriate pair of nodes (the first such pair of nodes, if there is more than one). If the resulting branch is closed, return to Step A. If it open and incomplete, return to Step B. If it is open and complete, then stop.

Step F. If (5) does not hold, go to Step G. If (5) holds, apply [$\neg O$] to the appropriate node (the first such node, if there is more than one). If the resulting branch is closed, return to Step A. If it open and incomplete, return to Step B. If it is open and complete, then stop.

Step G. If (6) does not hold, go to Step H. If (6) holds, apply [s] to the appropriate node (the first such node, if there is more than one). If the resulting branch is closed, return to Step A. If it open and incomplete, return to Step B. If it is open and complete, then stop.

Step H. If this step has been reached, (7) must hold. Apply [h] to the appropriate node (the first such node, if there is more than one). If the resulting branch is closed, return to Step A. If it open and incomplete, return to Step B. If it is open and complete, then stop.

This completes the specification of the procedure.

After a finite number of steps, this procedure must result in an open, complete branch. For suppose it doesn't. Since, by Lemma 9, each tableau is finite, the construction of the tableau must be finished at some point, and after it is finished there will be at least one open branch in it that is incomplete. Thus one of (1)-(7) must hold, which means that the tableau wasn't finished after all. Contradiction. ■

Theorem 2 (completeness) If $\Gamma \models_{\mathsf{D+}} A$, then $\Gamma \Vdash_{\mathsf{D+}} A$.

Proof. Suppose $\Gamma \nVdash_{\mathsf{D+}} A$. Then, by Definition 21 (p. 66), there is no closed D+ tableau for Γ / A. Consider any D+ tableau for Γ / A. By Lemma 10, this tableau has an extension (also a D+ tableau for Γ / A) containing at least one open, complete branch. Let $\langle W, R, v\rangle$ be the D+ model induced by this branch. By the relevant Completeness Lemma (either Lemma 4 or Lemma 5, depending on whether D+ is serial), $\bar{v}_{w_0}(B) = \mathbf{t}$ for all $B \in \Gamma \cup \{\neg A\}$. Thus $\Gamma \not\models_{\mathsf{D+}} A$. ■

Theorem 3 ("all or nothing") If there is an open, complete D+ tableau for Γ / A then there is no closed D+ tableau for Γ / A.

Proof. Suppose there is an open, complete D+ tableau for Γ / A. Pick an open branch, b, of this tableau, and let $\langle W, R, v\rangle$ be the D+ model induced by b. By the Completeness Lemma, $\bar{v}_{w_0}(B) = \mathbf{t}$ for all $B \in \Gamma \cup \{\neg A\}$. Thus $\Gamma \not\models_{\mathsf{D+}} A$. Thus, by the Soundness Theorem, $\Gamma \nVdash_{\mathsf{D+}} A$. Thus there is no closed D+ tableau for Γ / A. ■

Theorem 4 (decidability) The D+ systems are *decidable*; that is, there is an effective procedure which, when applied to any inference Γ / A, determines, in a finite number of steps, whether $\Gamma \vdash_{\mathsf{D+}} A$.

Proof. One such decision procedure is as follows. Start with the initial list for Γ / A. Begin applying D+ tableau rules, in any order. By Lemma 9, the tableau will terminate after a finite number of steps. If it is closed, then $\Gamma \vdash_{\mathsf{D+}} A$. If it is open, then by the "all-or-nothing" theorem, $\Gamma \nvdash_{\mathsf{D+}} A$. ■

2.6 Notable features of the D systems

In this section I highlight some notable features of the D systems. I will not try to give an exhaustive list of such features; rather, I will focus on a few representative ones. My reason for highlighting these features here is to give the reader (and myself) a "feel" for what holds and does not hold in (and of) standard deontic logics. This will provide a basis for comparing, contrasting, and evaluating different systems of deontic logic in the chapters that follow.

Recall that I use '$\mathsf{D+}$' as a variable ranging over these systems, and $\vdash$ to generalize over $\models$ and $\Vdash$ (which is legitimate given the soundness and completeness results established above). Also, I will write $A \dashv\vdash B$ as shorthand for ($A \vdash B$ and $B \vdash A$). I will often omit "system name" subscripts when context permits.

2.6.1 Relations between systems

Fact 3 The D systems are extensions of classical propositional logic (CPL). That is, if $\Gamma \vdash_{\mathsf{CPL}} A$ then $\Gamma \vdash_{\mathsf{D+}} A$.

Fact 4 The following subsumption relations hold between the D systems ($\mathsf{X} \to \mathsf{Y}$ means that if $\Gamma \vdash_{\mathsf{X}} A$ then $\Gamma \vdash_{\mathsf{Y}} A$):

$$\begin{array}{ccc} \mathsf{D} & \longrightarrow & \mathsf{Ds} \\ \downarrow & & \downarrow \\ \mathsf{Dh} & \longrightarrow & \mathsf{Dsh} \end{array}$$

2.6.2 Basic logical properties

Fact 5 The D systems enjoy the basic "Tarskian"[23] properties:

1. $A \in \Gamma \Rightarrow \Gamma \vdash_{\mathsf{D+}} A$ [reflexivity]

2. $\Gamma \vdash_{\mathsf{D+}} A \Rightarrow \Gamma, \Delta \vdash_{\mathsf{D+}} A$ [monotonicity]

3. $(\Gamma \vdash_{\mathsf{D+}} A$ and $A \vdash_{\mathsf{D+}} B) \Rightarrow \Gamma \vdash_{\mathsf{D+}} B$ [transitivity]

Fact 6 (deduction theorem) $\Gamma, A \vdash_{\mathsf{D+}} B$ iff $\Gamma \vdash_{\mathsf{D+}} A \supset B$.

Definition 33 $A[B/B']$ is the result of replacing every occurrence of B in A with B'.[24] For example, $(p \wedge q)[q/r] = (p \wedge r)$.

Fact 7 (replacement theorem) In $\mathsf{D+}$, if $B \dashv\vdash B'$ then $A \dashv\vdash A[B/B']$.

Proof. See, e.g., Chellas [43, §4.2]. ■

[23] So-called because Alfred Tarski famously highlighted these as (some of) the fundamental features of a logical consequence operator/relation. (See, e.g., his "On Some Fundamental Concepts of Metamathematics" in [150].)

[24] Chellas [43, p. 125] defines $A[B/B']$ as the result of replacing *zero or more* occurrences of B in A with B'. As such, his version of the replacement theorem states something a little stronger than mine. However, I prefer my definition of $A[B/B']$, since it makes $A[B/B']$ a genuine *function* of A, B, and B'.

2.6.3 Deontic inheritance

Definition 34 (oughtification) $O\Gamma =_{df} \{OA : A \in \Gamma\}$.

Fact 8 (inferential inheritance) If $\Gamma \vdash_{\mathsf{D+}} A$ then $O\Gamma \vdash_{\mathsf{D+}} OA$.

Proof. Suppose $O\Gamma \nvdash_{\mathsf{D+}} OA$. Then there is a $\mathsf{D+}$ model $\langle W, R, v\rangle$ and $w \in W$ such that $\bar{v}_w(B) = \mathbf{t}$ for all $B \in O\Gamma$ and $\bar{v}_w(OA) \neq \mathbf{t}$. Thus, by the truth condition for O, there is a $u \in W$ such that wRu and $\bar{v}_u(A) \neq \mathbf{t}$. Thus, by the truth condition for O again, $\bar{v}_u(B) = \mathbf{t}$ for all $B \in \Gamma$. Thus $\Gamma \nvdash_{\mathsf{D+}} A$. ■

Remark 9 Note that the following principle *does not* hold: If $\Gamma \vdash_{\mathsf{D+}} A$ then $P\Gamma \vdash_{\mathsf{D+}} PA$. This seems quite appropriate. Suppose, for example, that both A and B are permitted. It certainly does not follow that $A \wedge B$ is permitted. (Consider the "drinking and driving" example from before.) However, it is true that if $A \vdash_{\mathsf{D+}} B$ then $PA \vdash_{\mathsf{D+}} PB$.

Fact 9 (implicational inheritance) If $\vdash_{\mathsf{D+}} A \supset B$ then $\vdash_{\mathsf{D+}} OA \supset OB$.

Proof. Follows easily from the Deduction Theorem (Fact 6) and the Inferential Inheritance Principle (Fact 8). ■

Fact 10 Here are a few more noteworthy, easily verified facts related to deontic inheritance:

1. If $\vdash_{\mathsf{D+}} A \equiv B$ then $\vdash_{\mathsf{D+}} OA \equiv OB$.
2. If $\vdash_{\mathsf{D+}} A \supset B$ then $\vdash_{\mathsf{D+}} PA \supset PB$.
3. If $\vdash_{\mathsf{D+}} A \equiv B$ then $\vdash_{\mathsf{D+}} PA \equiv PB$.
4. If $\vdash_{\mathsf{D+}} A \supset B$ then $\vdash_{\mathsf{D+}} FB \supset FA$.
5. If $\vdash_{\mathsf{D+}} A \equiv B$ then $\vdash_{\mathsf{D+}} FA \equiv FB$.

2.6.4 Miscellaneous validities/invalidities

Here I list a number of notable validities and invalidities. I include "nicknames" for many of these principles/inferences, since many of them are notable enough to merit such monikers. (Some of the nicknames are self-explanatory or have already been explained; others will be explained later.)

Each of the following statements applies to *all* D systems except where otherwise specified. (When I *do* specify the system, this should be taken to conversationally imply—à la Grice [73]—that the statement holds only for that system and its extensions.)

1. $PA \dashv\vdash \neg FA$ [weak permission]
2. $OA, OB \vdash O(A \wedge B)$ [O-aggregation][25]
3. $O(A \wedge B) \vdash OA$ and $O(A \wedge B) \vdash OB$ [O-conjunction elimination]
4. $P(A \vee B) \dashv\vdash PA \vee PB$ [principle of deontic distribution]
5. $OA \vdash O(A \vee B)$ [Ross's paradox 1]
6. $PA \vdash P(A \vee B)$ [Ross's paradox 2]
7. $OA \vdash O(B \supset A)$ [Prior's paradox 1]
8. $OA \vdash O(\neg A \supset B)$ [Prior's paradox 2]
9. $P(A \vee B) \nvdash PA \wedge PB$ [free choice permission][26]
10. $O(A \supset B), OA \vdash OB$ [deontic detachment]
11. $O(A \supset B), PA \vdash PB$ [permissive deontic detachment]
12. $O(A \supset B), FB \vdash FA$ [deontic modus tollens]
13. $O(A \supset B), A \nvdash OB$ [deviant deontic detachment 1]
14. $A \supset OB, OA \nvdash OB$ [deviant deontic detachment 2]
15. $A \supset OB \nvdash O(A \supset B)$
16. $\vdash_{\mathsf{Ds}} OA \supset PA$ [ought implies may][27]
17. $\vdash_{\mathsf{Dh}} O(OA \supset A)$ [U schema]

[25] See Example 7.
[26] See Example 10.
[27] See Example 8.

18. $\vdash_{\mathsf{Dh}} OOA \supset OA$ [X schema]

19. $\nvdash OA \supset A$ [Panglossian principle 1][28]

20. $\nvdash A \supset OA$ [Panglossian principle 2]

21. $\vdash_{\mathsf{Ds}} PA \vee P\neg A$ [principle of permission]

22. $\vdash FA \vee PA$ [deontic covering principle]

23. $O(A \vee B), FA \vdash OB$ [deontic disjunctive syllogism]

24. $FA \vdash F(A \wedge B)$ [penitent's paradox]

25. $\vdash O(A \vee \neg A)$ [tautologies are obligatory]

26. $\vdash F(A \wedge \neg A)$ [contradictions are forbidden]

27. $\vdash_{\mathsf{Ds}} P(A \vee \neg A)$ [tautologies are permitted]

2.6.5 Ross's paradox and free choice permission

As noted above, $OA \supset O(A \vee B)$ is valid in D+. Thus, for example, if I ought to mail the letter, then I ought to either mail it or burn it. This is commonly referred to as *Ross's paradox*, after Alf Ross [135]. The schema is considered "paradoxical" because $O(mail \vee burn)$ seems to imply that burning the letter is an acceptable way for me to discharge my obligation. Of course, it implies no such thing: $O(A \vee B) \supset PB$ is not valid in any of the D systems, nor should it be.

There is a simple explanation of the "paradox" in terms of Grice's theory of conversational implicature [73]. To assert $O(mail \vee burn)$ when the stronger $O(mail)$ holds (assuming the asserter *knows* that it holds) would be to flout the maxim of Quantity—in particular, the principle that one should make one's contribution *as informative as required* for the current purposes of the exchange. In other words, a well-intentioned speaker would never *assert* that I ought to mail or burn the letter if he knew the more informative fact that I simply ought to mail it. But appropriate *assertion* conditions are one thing; *truth* conditions are another. The implication[29] in question bears the hallmark of a conversational implicature (as opposed to a logical consequence)—it is *defeasible*:

[28] See Example 9.

[29] I am using 'implication' in a loose, colloquial sense here.

> **Q:** Is it true that I ought to either mail or burn the letter?
>
> **A:** Yes, but that is quite misleading, because you ought to mail it and you ought not to burn it.

The schema $PA \supset P(A \vee B)$ is also valid in $\mathsf{D+}$. Thus, if I may smoke, then I may smoke or kill. This version of Ross's paradox is slightly trickier, because it is tangled up with the notion of *free choice permission* [89]—the concept of permission under which, e.g., 'You may have coffee or tea' implies 'You may have coffee and you may have tea'. If we construe P as free choice permission, then accepting Ross's second paradox would allow us to derive $PA \wedge PB$ from PA, which is clearly wrong. But again, Grice's theory comes to the rescue. If I know that you are permitted to have coffee but not permitted to have tea, I would be flouting the maxim of Quantity were I to assert that you may have coffee or tea. Again the implication is defeasible:

> **Q:** Is it true that I am permitted to have coffee or tea?
>
> **A:** Yes, but that is quite misleading, because you are are permitted to have coffee but you are not permitted to have tea.

Another way to quell the concerns raised by Ross's paradox(es) is to simply think more carefully about what O and P actually mean: OA says, roughly, *In a perfect world, A would be true*. But a world in which A is true is obviously a world in which $A \vee B$ is true. It follows naturally that in a perfect world, $A \vee B$ would be true, i.e. $O(A \vee B)$ holds. PA says, roughly, *There is a perfect world in which A is true*. Again, it clearly follows that there is a perfect world in which $A \vee B$ is true, i.e. $P(A \vee B)$ holds.

2.6.6 Prior's paradox and conditional obligation

As noted above, $OA \supset O(B \supset A)$ and $OA \supset O(\neg A \supset B)$ are valid in $\mathsf{D+}$. Thus if you ought to keep your promise, you ought to keep it if it kills you and you ought to be killed if you don't keep it! This is known as *Prior's paradox*, after A. N. Prior [132]. This paradox and others have led some deontic logicians to conclude that conditional obligation involves a "special, already normative-laden kind of 'if'" [117, p. 3]. Accordingly, these logicians have introduced the dyadic deontic operator $O(A|B)$, which can be read as ⌜it ought to be that A, given B⌝. Unconditional obligation, OA, is then defined as $O(A|\top)$, where $\top$ is an arbitrary tautology. There are various ways of providing semantics for $O(\,|\,)$, most of which are fairly complicated and, in my opinion, not very intuitive.[30]

It is no accident that $O(A|B)$ looks a lot like the conditional probability operator, $\Pr(A|B)$, which denotes the real number between 0 and 1 that is the probability of A, given (only!) the information that B. Just as the value of $\Pr(A|B)$ is independent of $\Pr(A|B \wedge C)$ in probability theory, the truth value of $O(A|B)$ is independent of the truth value of $O(A|B \wedge C)$ in dyadic deontic logic. The analogy with conditional probability is a bad one, however, because obligation, unlike probability, is not subjective.[31] The obligatoriness of a proposition, unlike the probability of a proposition, does not depend on what the agent considering that proposition happens to know or believe: gaining new knowledge can make something more or less probable, but it cannot make something more or less obligatory.[32] For this reason, I believe that dyadic deontic logic is misguided. Like Bonevac [29], I hold that a proper account of conditional obligation ought to (and will) be a natural consequence of a proper account of conditionals and a proper account of obligation.

But I digress. The important thing to note here is that my objections to the D systems (and to classically-based deontic logics in general) apply with equal force to the popular dyadic extensions of those systems. 'Extension' is the operative word here: since I will be claiming that the D systems are too strong in a specific sense, it will follow that the dyadic systems that are based on them are also too strong, since the latter include the former.

Returning to Prior's paradoxes: given that $A \supset B$ is equivalent to $\neg A \vee B$, Prior's paradoxes are clearly just notational variants on Ross's paradoxes, and thus can be explained away in the same manner. Of course, one may wish to challenge the dubious assumption that ⌜if A then B⌝ is equivalent to ⌜either not-A or B⌝. (Indeed, I will implicitly challenge this assumption in the next chapter.) My point is just that there is nothing objectionable about Prior's paradoxes beyond what may be objectionable about the material conditional *per se*.

Finally, note that in the D systems $O(A \supset B)$ is not equivalent to $A \supset OB$ (neither entails the other). Moreover, both $O(A \supset B), A / OB$ and $A \supset OB, OA / OB$ are

[30] See Lewis [96] for a survey.

[31] Yes, I am being a dogmatic Bayesian here!

[32] A possible exception to this claim might be norms that make reference to epistemic states, such as *If you know that a person is in danger, you should help that person.*

invalid. Some have the intuition that $O(A \supset B)$ and $A \supset OB$ ought to be treated as equivalent, and that the aforementioned inferences ought to be regarded as valid. (See, e.g., Castañeda [42].) For example, the following inferences ought to be considered correct:

It ought to be that if you promise to help, you help.
You promised to help.
$\therefore$ You ought to help.

If you promise to help, you ought to help.
You ought to promise to help.
$\therefore$ You ought to help.

Let us consider what condition could be imposed on the semantics of the D systems to make $O(A \supset B)$ logically equivalent to $A \supset OB$:

$$\begin{aligned}
\bar{v}_w(O(A \supset B)) = \mathbf{t} \quad &\text{iff} \quad \bar{v}_w(O(\neg A \vee B)) = \mathbf{t} \\
&\text{iff} \quad \forall u(wRu \Rightarrow \bar{v}_u(\neg A \vee B) = \mathbf{t}) \\
&\text{iff} \quad \forall u(wRu \Rightarrow \bar{v}_u(A) = \mathbf{f} \text{ or } \bar{v}_u(B) = \mathbf{t}).
\end{aligned}$$

$$\begin{aligned}
\bar{v}_w(A \supset OB) = \mathbf{t} \quad &\text{iff} \quad \bar{v}_w(\neg A \vee OB) = \mathbf{t} \\
&\text{iff} \quad \bar{v}_w(\neg A) = \mathbf{t} \text{ or } \bar{v}_w(OB) = \mathbf{t} \\
&\text{iff} \quad \bar{v}_w(A) = \mathbf{f} \text{ or } \forall u(wRu \Rightarrow \bar{v}_u(B) = \mathbf{t}).
\end{aligned}$$

Thus to get the equivalence we would need to impose the restriction that

$$\begin{gathered}
\forall u(wRu \Rightarrow \bar{v}_u(A) = \mathbf{f} \text{ or } \bar{v}_u(B) = \mathbf{t}) \\
\text{iff} \\
\bar{v}_w(A) = \mathbf{f} \text{ or } \forall u(wRu \Rightarrow \bar{v}_u(B) = \mathbf{t})
\end{gathered}$$

In English this restriction says:

At all worlds seen by w, either A fails or B holds
iff
either A fails at w or B holds at all worlds seen by w.

It's not clear to me that there is any intuitive justification for this condition. Moreover, it seems possible to construct counterexamples to the purported equivalence, in both directions. Suppose, for example, that I believe the income tax ought to be eliminated. Then I will accept

It ought to be that if I earn income, I don't pay taxes on it

but I may not accept

If I earn income, I ought not to pay taxes on it.

After all, I might believe that as a good citizen I ought to obey the law as it stands, even though I disagree with it and think it should be changed. For the same reason, I may accept that

If I earn income, I ought to pay taxes on it

while rejecting

It ought to be that if I earn income, I pay taxes on it.

I conclude that it is quite appropriate for $O(A \supset B)$ and $A \supset OB$ to be logically independent of one another.

2.6.7 Principles related to conflict-(in)tolerance

I now list a number of features that are at least *prima facie* problematic, given the assumption (defended in the previous chapter) that normative conflicts are possible. Specifically, each of these features seems to indicate that some or all of the D systems are inadequate for dealing with normative conflicts (note: if $\vdash$ has no subscript, the subscript should be understood to be $\mathsf{D}+$):

1. $\vdash \neg(FA \wedge PA)$ [no escapable conflicts]
2. $\vdash_{\mathsf{Ds}} \neg(OA \wedge FA)$ [no inescapable conflicts]
3. $FA, PA \vdash B$ [escapable conflicts "explode"]
4. $OA, FA \vdash OB$ [inescapable conflicts "deontically explode"][33]
5. $OA, FA \vdash_{\mathsf{Ds}} B$ [inescapable conflicts (just plain) "explode"]

[33]See Example 6, p. 67.

(1) rules out escapable conflicts. (2) rules out inescapable ones (though only in systems with seriality; at a "dead end" world OA and FA can both be true). It gets worse from there. (3) says that something's being forbidden and permitted entails *everything*, while (4) says that something's being obligatory and forbidden entails—well, if not everything, then at least that everything is obligatory (as well as forbidden). Thus, for example, *all* of the D systems deem the following inferences valid:

You are forbidden to smoke.
You are permitted to smoke.
∴ The world has come to an end.

You should put the document in the safe.
You should not put the document in the safe.
∴ The world should be destroyed.

This is clearly unacceptable—on the assumption that normative conflicts are possible. Facing a normative conflict may be *confusing* and *unpleasant*, but it is not the end of the world! (Nor *should* it be!)

2.7 Classical alternatives to the D systems

In this section I consider two classically-based alternatives to the D systems that have been promoted—falsely, I believe—as being better equipped to deal with normative conflicts.

2.7.1 Non-aggregative deontic logic

In "Non-Kripkean Deontic Logic" [140], Peter Schotch and Ray Jennings propose a system similar to Ds except that the single accessiblility relation, R, is replaced with a set, $\mathcal{R}$, of such relations.[34] The elements of $\mathcal{R}$ can be thought of as corresponding to various, possibly conflicting, normative systems. The clause for O is formulated as follows:

$$\bar{v}_w(OA) \quad = \quad \begin{cases} \mathbf{t} \text{ if } (\exists R \in \mathcal{R})(\forall u)(wRu \Rightarrow \bar{v}_u(A) = \mathbf{t}) \\ \mathbf{f} \text{ otherwise} \end{cases}$$

[34] See also Chellas [43, Exercise 3.14]. This system is investigated more thoroughly in Goble [62], which also contains an alternative, "preference-based" semantics for the system.

Thus, loosely speaking, OA is true just in case some normative system requires that A be true. I will call Schotch and Jennings's system 'NA' (for 'non-aggregative'; this will be explained in a moment).

$OA, FA \mathbin{/} OB$ is invalid in NA. Just consider the following (partially specified) model:

- $W = \{w_0, w_1\}$
- $\mathcal{R} = \{R_0, R_1\}$
- $R_0 = \{\langle w_0, w_0\rangle, \langle w_1, w_1\rangle\}$
- $R_1 = \{\langle w_0, w_1\rangle, \langle w_1, w_0\rangle\}$
- $v_{w_0}(p) = \mathbf{t}$
- $v_{w_0}(q) = v_{w_1}(p) = v_{w_1}(q) = \mathbf{f}$

This model can be depicted as follows:

$$\begin{tikzcd} \boxed{w_0} \arrow[loop above, "0"] \arrow[r, shift left, "1"] & \boxed{w_1} \arrow[loop above, "0"] \arrow[l, shift left, "1"] \\ p & \neg p \\ \neg q & \neg q \end{tikzcd}$$

It is easy to check that in this model Op and $O\neg p$ are true at w_0 but Oq is not.

There are two serious (and closely related) problems with NA, in my view. First while it invalidates $OA, FA \mathbin{/} OB$, it still validates $O(A \wedge \neg A) \mathbin{/} OB$. Indeed, since $O(A \wedge \neg A)$ is unsatisfiable, it validates the stronger $O(A \wedge \neg A) \mathbin{/} B$. Now, some have the intuition that while a normative conflict of the form $OA \wedge O\neg A$ may be possible, one of the form $O(A \wedge \neg A)$ is definitely not. Risto Hilpinen, for example, writes: "a person may perhaps be under *conflicting* obligations, but he can hardly be subject to a self-contradictory (or impossible) obligation" [79, p. 192]. However, $O(A \wedge \neg A)$ follows from $OA \wedge O\neg A$ by the intuitively correct principle of aggregation, $OA \wedge OB \supset O(A \wedge B)$. Of course, this principle fails in NA (hence the name). The model above is a counterexample: while Op and $O\neg p$ hold at w_0, $O(p \wedge \neg p)$ does not. Thus NA invalidates, e.g., the following intuitively correct inference:

George ought to get a haircut.
George ought to get a real job.
$\therefore$ George ought to get a haircut and get a real job.

My second objection to NA, then, is simply that it *lives up to its name*, invalidating a deontic inference that I take to be clearly correct. Note that NA also invalidates the dual of aggregation, $P(A \vee B) \supset PA \vee PA$, which is of course one half of von Wright's Principle of Deontic Distribution (see Section 2.2).

NA achieves conflict tolerance only by abandoning one of the most fundamental and intuitively correct principles of deontic logic. Moreover, just like the D systems, it breaks down in the presence of self-conflicting obligations, i.e., those of the form $O(A \wedge \neg A)$. And finally, just like the D systems, it collapses in the presence of escapable conflicts, i.e., those of the form $FA \wedge PA$. In short, NA achieves only superficial conflict-tolerance, and at a higher cost than one should be willing to pay even for real, robust conflict-tolerance.

2.7.2 Deontic logic with permitted inheritance

In "A Logic for Deontic Dilemmas"[35] [66], Lou Goble specifies a classically-based deontic logic, DPM, which he claims is suitable for dealing with normative conflicts of the form $\{OA, O\neg A\}$ (he calls such conflicts *deontic dilemmas*).[36] Goble gives a minimal model (or "neighborhood") semantics for DPM.[37] The deontic accessibility relation, R, is replaced with a function, $\mathcal{O} : W \longrightarrow \wp\wp W$. If A is a formula, $|A| =_{df} \{w : \bar{v}_w(A) = \mathbf{t}\}$ is the set of worlds at which A is true, or the proposition expressed by A. Thus $\mathcal{O}$ can be thought of as taking a world to the set of all propositions that are obligatory at that world.

The clause for O is given, quite simply, as follows:

$$\bar{v}_w(OA) \quad = \quad \begin{cases} \mathbf{t} \text{ if } |A| \in \mathcal{O}_w \\ \mathbf{f} \text{ otherwise.} \end{cases}$$

That is, OA holds at w just in case the proposition expressed by A is in the set of propositions that are obligatory at w. The following restrictions are placed on $\mathcal{O}$:

[35] See also Goble [65].

[36] Actually Goble specifies two slightly different systems, DPM.1 and DPM.2; I will focus on the former, since it validates aggregation and is therefore less objectionable.

[37] For a good introduction to minimal model semantics, see Chellas [43, chap. 7].

- $W \in \mathcal{O}_w$
- if $X \in \mathcal{O}_w$ and $Y \in \mathcal{O}_w$, then $X \cap Y \in \mathcal{O}_w$
- if $X \subseteq Y$ and $X \in \mathcal{O}_w$ and $\bar{X} \notin \mathcal{O}_w$, then $Y \in \mathcal{O}_w$

F and P (and G, if one likes) are defined as usual.

Goble calls DPM a deontic logic of "permitted inheritance", since it rejects the usual principle of inheritance,

$$\text{if } \vdash A \supset B \text{ then } \vdash OA \supset OB, \tag{RM}$$

in favor of the weaker principle of *permitted* inheritance,

$$\text{if } \vdash A \supset B \text{ then } \vdash PA \supset (OA \supset OB). \tag{RPM}$$

The rule of deontic equivalence,

$$\text{if } \vdash A \equiv B \text{ then } \vdash OA \equiv OB, \tag{RE}$$

also holds for DPM.

Like NA, DPM invalidates $OA, FA / OB$. Unlike NA, however, it also invalidates $O(A \wedge \neg A) / OB$ and preserves the validity of $OA, OB / O(A \wedge B)$. However, DPM still validates $FA, PA / B$. It might be thought that DPM is at least adequate for dealing with inescapable normative conflicts. It is not, however. For DPM is seriously flawed in another way. Since $A \wedge \neg A \equiv B \wedge \neg B$ is a classical tautology, by (RPM) $O(A \wedge \neg A) \equiv O(B \wedge \neg B)$ is valid in DPM. Thus $O(A \wedge \neg A) \supset O(B \wedge \neg B)$ is also valid. Yet $O(A \wedge \neg A) \supset OB$ is not valid in DPM. This is possible because DPM invalidates $O(A \wedge B) \supset OA$ $[OB]$, which we might label 'O-conjunction elimination'.[38]

Surely it would be hard to imagine a more intuitively correct deontic inference than O-conjunction elimination! It is hard to give much credibility to a deontic logic that tells us that the following inference, for example, is invalid:

George ought to get a haircut and get a real job.
$\therefore$ George ought to get a haircut.

[38] I am grateful to Goble himself for bringing this point to my attention!

Of course, a restriction could be placed on $\mathcal{O}$ in order to make this inference valid, namely:

$$\text{if } X \cap Y \in \mathcal{O}_w \text{ then } X \in \mathcal{O}_w.$$

But then $O(A \wedge \neg A) \,/\, OB$ would become valid, undermining even the claim that the system is robustly tolerant of *inescapable* conflicts.

Note that DPM also invalidates the dual of O-conjunction elimination, $PA \vee PB \supset P(A \vee B)$, i.e., the *other* half of von Wright's Principle of Deontic Distribution (see §2.2). (Those who are (still) hung up on Ross's paradox may find this feature reassuring, but I, for one, do not.)

Like NA, DPM achieves conflict tolerance only by invalidating one of the most fundamental and intuitively correct principles of deontic logic, namely von Wright's Principle of Deontic Distribution. (NA rejects one half of it; DPM rejects the other.) Moreover, just like D+ and NA, DPM collapses into triviality in the presence of escapable conflicts. Thus while DPM is, at least, robustly tolerant of *inescapable* conflicts, it achieves this tolerance at far too high a cost.

2.8 Blame explosion

It is becoming clear that trying to develop a robustly conflict-tolerant deontic logic based on classical logic is a futile endeavor. Ultimately, the problems with these systems all trace back to the dubious principle of explosion. It is time we tried abandoning classical logic altogether, and constructing our deontic logic upon a paraconsistent foundation. The next chapter is dedicated to making sure that this foundation is solid.

Chapter 3

Basic paraconsistent logic

> Paraconsistent logic arises as an attempt to deal more constructively with inconsistent premise sets than classical logic does. Some of those who take an interest in it are fairly conservative thinkers who are merely sensitive to our epistemic limits, and prepared to acknowledge the need to reason, from time to time, with premises that cannot all be true, without having them explode in a puff of logic.
>
> —Bryson Brown [33, p. 490].

> That paraconsistent logic admits of some inconsistencies should not mean that you can be as incoherent as you want, if you work in this area. That you tolerate inconsistencies should not mean that you eagerly expect for them to be found.
>
> —João Marcos [101, p. xlvii]

A paraconsistent logic is one that rejects the principle of explosion (a.k.a. *ex contradictione sequitur quodlibet*), according to which "everything follows from a contradiction."[1] In particular, a logical consequence relation, $\vdash$, is paraconsistent just in case $A, \neg A \nvdash B$, i.e., it is not the case that $\forall A \forall B(A, \neg A \vdash B)$.

The aim of this chapter is to build a basic paraconsistent logic that will serve as a well-motivated and technically sound foundation for a robustly conflict-tolerant deontic logic. After laying down some adequacy conditions for a paraconsistent logic, I will walk through the instructive process of devising a logic satisfying these conditions.

[1] For an exceptionally clear and broad introduction to paraconsistent logic, see Priest [128]. Brown [34], Priest and Tanaka [131], and Bremer [31] are also good.

As it turns out, the logic I will construct, which I call 'P3' (for 'paraconsistent 3-valued logic'), is equivalent to the system J_3 proposed by D'Ottaviano and da Costa [50] in a somewhat obscure 1970 paper (written in French).[2] (Alternatively, P3 and J_3 can be thought of as alternative presentations of the same logic.[3]) I provide a tableau-style proof theory for P3 that is similar to the one Carnielli and Marcos [38] have given for their paraconsistent logic LFI1 (which is also, incidentally, equivalent to J_3 and therefore P3).[4]

For reasons that will become clear later in the dissertation, I will also define a very similar but slightly weaker system, P4 ('paraconsistent 4-valued logic').

I want to begin the chapter, however, by providing a bit of historical background and attempting to dispel a common myth about paraconsistent logic.

3.1 Historical background

The first paraconsistent logician was, arguably, the first logician, Aristotle. Aristotle's syllogistic does not license drawing arbitrary conclusions from contradictory premise sets such as {Some swans are black, No swans are black}. However, formal systems of paraconsistent logic that utilized modern mathematical techniques and were explicitly *designed* to be paraconsistent did not emerge until the mid-twentieth century. While the Russians and Orlov and the famous Polish logician Łukasiewicz at least hinted at paraconsistent logics early in the century, the first widely known, fully specified paraconsistent logic, D_2 (for 'two-valued discursive sentential calculus') was presented by Stanisław Jaśkowski, a former student of Łukasiewicz, in 1948.[5] Independently, South American logicians Newton da Costa and F. G. Asenjo developed their own brand(s) of paraconsistent logic in the 1950s and '60s.[6] Other important figures in the early development of paraconsistent logic were the U. S. logicians Alan Ross Anderson and Nuel Belnap, whose concern with avoiding fallacies of (ir)relevance led them to construct "relevance logics" (a.k.a. "relevant logics") which are *de facto* paraconsistent,

[2] A nice presentation of J_3 is also given by Epstein and D'Ottaviano in [54]. (Actually, the logic presented there is closer to our system P3 than to the original formulation of J_3.)

[3] The question of when a logic L_1 is identical to a logic L_2 is far from straightforward.

[4] See Batens and De Clerq [20] for yet another system that is equivalent to J_3. Again, depending on one's criteria for identity between logics, all of these systems may be regarded as distinct but equivalent, or they may be regarded as simply different presentations of the same logic.

[5] See Jaśkowski [86].

[6] See Asenjo [12], da Costa [45].

even though paraconsistency *per se* was not their primary concern (or was only one aspect of their primary concern).[7] The term 'paraconsistent'—apparently meaning *beyond (or beside) the consistent*—was coined in 1976 by Peruvian philosopher Francisco Miró Quesada. (For a more extensive history of paraconsistent logic, see da Costa et al. [46].)

Today there are hundreds of scholars working in the field of paraconsistent logic, and many different approaches to constructing paraconsistent logics. Paraconsistent logic has found theoretical applications in philosophy and mathematics, as well as practical applications in computer science.[8] In 1990, *Mathematical Reviews* added a subject category for paraconsistent logic. The first World Congress on Paraconsistency took place in Belgium in 1997, and subsequent congresses were held in Brazil (2000) and France (2003).[9]

3.2 Three views on inconsistency

It is important to distinguish between three very different views that are often conflated:

- *Contradictionism*: The view that there is, in general, nothing wrong with contradictions, and that one should feel free to be as inconsistent as one likes.
- *Dialetheism*: The view that there are, or could be, true contradictions ("dialetheias"), and therefore it is, or could be, rational to believe certain contradictions.
- *Paraconsistentism*: The view that contradictions need not be regarded as entailing everything—i.e., an inconsistent theory is not necessarily trivial.

The contradictionist's point of view is nicely expressed by Walt Whitman in *Leaves of Grass*:

---

[7] See Anderson and Belnap [9].

[8] See e.g. Grant and Subramanian [72], Bertossi et al. [25].

[9] Selected papers from these meetings can be found in Batens et al. [19], Carnielli et al. [36], and Béziau, Carnielli, and Gabbay [27]. The fourth World Congress on Paraconsistency is scheduled to take place in Australia in the summer of 2008.

> Do I contradict myself?
> Very well then; I contradict myself.
> (I am large; I contain multitudes.)[10]

I have encountered a few self-proclaimed "postmodernists" and "deconstructionists" who apparently delight in advocating self-contradictory views (e.g. "there are no facts, only interpretations"). But I have never met a *serious* thinker who was a contradictionist, and I doubt that there has ever been such a person.[11]

The most prominent contemporary proponent of dialetheism is Graham Priest, who argues that certain antinomies (e.g. the liar paradox, Russell's paradox, certain paradoxes of change and motion) force us, rationally, to accept certain contradictions as being true.[12] But Priest is quick to add that

> ... inconsistency is a rational black mark. If we have views that are inconsistent then we are probably incorrect. We should go back and examine why we hold such a view, and what the alternatives are. We may find that we would be better off going a different way. But we may find that there are no better ways to go. In which case, we may just have to conclude that the improbable is the case. After all, the improbable happens sometimes. [126, p. 424]

Dialetheism is certainly a *radical* view (at least in some ways); unlike contradictionism, however, it is a serious philosophical position, substantiated with formidable arguments.

Paraconsistentism, in stark contrast with both contradictionism and dialethism, is inherently *conservative*, in that it merely implores us to be more *cautious* in the inferences we draw in the presence of inconsistent information.[13] This view is nicely

---

[10] Book III ("Song of Myself"), §51.

[11] Heraclitus, Hegel, and Marx are often cited as examples of philosophers who not only tolerated but celebrated contradictions. (See e.g. [46, p. 113].) Textual support for this claim is rarely if ever provided, however. One could point to, e.g., Heraclitan fragments such as:

> Conjunctions: wholes and not wholes, the converging the diverging, the consonant the dissonant, from all things one, and from one all things. (Fragment 10, cited in [85, p. 45])

But even setting translation issues aside, this fragment is, in itself, far too vague and detached from context to be regarded as a ringing endorsement of contradictionism.

[12] See Priest [129], [130].

[13] Compare the following quotation from David Nelson ([121], quoted in [101, p. xxvi]):

captured in the words of Ralph Waldo Emerson: "Suppose you should contradict yourself; what then?"[14] Contradictions are bad, but you need (and should) not go berserk when you encounter one. Paraconsistentism shifts the emphasis from whether contradictions can be true to the question of what *follows* from them. It is worth emphasizing that paraconsistent logics are in general *weaker* than classical logic, in that they invalidate certain inferences (such as explosion) that classical logic validates, and do not validate any inferences that classical logic invalidates. In particular, contrary to popular myth, paraconsistent logics generally do not validate any contradictions; in fact, as we shall see, many of them even validate (one version of) the law of *non*-contradiction, $\neg(A \wedge \neg A)$!

Suppose a civil engineer maintains that bridges ought to be designed with certain features that prevent them from collapsing when they develop cracks. Does it follow that this engineer holds that it is *perfectly acceptable* for bridges to have cracks in them, and that we should not take steps to prevent such cracks from developing, or to fix them when they do develop? (Or worse, that we should gleefully attack bridges with sledgehammers, aiming to produce as many cracks as we can?) Of course not. Rather the engineer, recognizing that despite our best efforts some bridges will inevitably develop cracks, simply wants to ensure that such cracks will not lead to catastrophe—i.e., physical collapse. Similarly, the paraconsistent logician, recognizing that despite our best efforts some theories (databases, belief systems, works of fiction, etc.) will inevitably turn out to be inconsistent, wants to ensure that such inconsistencies do not lead to catastrophe—i.e., triviality. *The main point of paraconsistent logic is not to make contradictions respectable, but rather to prevent them from doing too much harm.*

> In both the intuitionistic and classical logic all contradictions are equivalent. This makes it impossible to consider such entities at all in mathematics. It is not clear to me that such a *radical* position regarding contradiction is necessary. I feel that it may be possible to conceive a logic which does more justice to the uncertainty of the empirical situation insofar as negation is concerned. [emphasis mine—CM]

[14] From Emerson's essay "Self-Reliance."

3.3 The road to paraconsistency

3.3.1 Adequacy criteria

I will now lay down, somewhat dogmatically, three adequacy criteria for a paraconsistent logic. In addition to invalidating explosion, which it must do by definition, our paraconsistent logic should:

1. *have a clear, intuitive semantics.* The semantics should be independently defensible and not merely an *ad hoc* construction motivated solely by the desire for a system with certain formal characteristics.

2. *diverge as little as possible from classical logic.* I like to think of this as the "principle of minimal mutilation." Among other things, it entails that our paraconsistent logic should not reject a classically valid inference unless doing so is required to achieve paraconsistency (or there is some other very good reason for doing so). It also entails that our logical consequence relation should have the usual "Tarskian" properties of reflexivity ($\Gamma \subseteq Cn(\Gamma)$), monotonicity ($\Gamma \subseteq \Delta \Rightarrow Cn(\Gamma) \subseteq Cn(\Delta)$), and transitivity ($Cn(Cn(\Gamma)) \subseteq Cn(\Gamma)$) unless there is a compelling reason to give them up.[15]

3. *not validate any trivial variations on explosion*, e.g. $A \wedge \neg A / B$ or $\neg A \supset (A \supset B)$. That is, our logic should be "robustly" paraconsistent.

3.3.2 Invalidating explosion

Explosion is valid in classical logic because there are no models in which A and $\neg A$ are true, and thus, *a fortiori*, no models in which they are true *and B is not true.* Hence a natural way to render explosion invalid is to simply allow for models in which A and $\neg A$ are both true (and B is not true). And a natural way to do *this*, in turn, is to allow for models in which A is both true and false.

This approach will inevitably strike some readers as absurd. David Lewis undoubtedly expresses the mainstream point of view when he writes, in response to this proposal: "Nothing is, and nothing could be, literally both true and false. This we know for certain, and *a priori*, without any exception for especially perplexing subject matters" [97, p. 434].

[15] $Cn(\Gamma) =_{df} \{A : \Gamma \vdash A\}$.

But we need not construe A's being both true and false *in a model* as A's *really being* both true and false (say, in some possible world). We can interpret A's being true and false in a model as A's being true and false in some (perhaps impossible) context or according to some (perhaps inconsistent) source or body of information (e.g. a theory or knowledge base). J. Michael Dunn, an early advocate of the type of semantics under consideration, makes the point this way:

> Do not get me wrong—I am not claiming that there are sentences which are in fact both true and false. I am merely pointing out that there are plenty of situations where we suppose, assert, believe, etc., contradictory sentences to be true, and we therefore need a semantics which expresses the truth conditions of contradictions in terms of the truth values that the ingredient sentences would have to take in order for the contradictions to be true. [51, p. 157]

The system we are about to define was first presented by F. G. Asenjo [12] in 1966, though it was popularized as LP ('logic of paradox') by Graham Priest [124] beginning in 1979. (As we will see, P3, the paraconsistent logic I favor, is obtained by simply adding a certain special connective to LP.)

A formula is now assigned a *set* of truth values: $\{1\}$ for exclusively true, $\{0\}$ for exclusively false, or $\{1, 0\}$ for both true and false. (For convenience I will often write these as **t**, **f**, and **b**, respectively.[16]) We could also allow formulas to be assigned the empty set, $\varnothing$, representing 'neither true nor false'. This is the approach favored by, for example, Dunn [51] [52]. We will not avail ourselves of this option at the moment, however, as it results in a weaker logic and is unnecessary for achieving paraconsistency. (Recall the principle of minimal mutilation.)

The truth and falsity conditions for negation, disjunction, and conjunction can be given quite naturally as follows:

- $\neg A$ is true iff A is false
- $\neg A$ is false iff A is true
- $A \vee B$ is true iff A is true or B is true

[16] The **t** and **f** we are talking about here can and should be identified with the **t** and **f** of Chapter 2.

- $A \vee B$ is false iff A is false and B is false
- $A \wedge B$ is true iff A is true and B is true
- $A \wedge B$ is false iff A is false or B is false

More formally, an LP valuation is a function $v : At \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}\}$ that is extended to $\bar{v} : L(\neg, \vee, \wedge) \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}\}$ via the following clauses:

$$\begin{array}{lcl}
\bar{v}(p) & = & v(p) \\
1 \in \bar{v}(\neg A) & \Leftrightarrow & 0 \in \bar{v}(A) \\
0 \in \bar{v}(\neg A) & \Leftrightarrow & 1 \in \bar{v}(A) \\
1 \in \bar{v}(A \vee B) & \Leftrightarrow & 1 \in \bar{v}(A) \text{ or } 1 \in \bar{v}(B) \\
0 \in \bar{v}(A \vee B) & \Leftrightarrow & 0 \in \bar{v}(A) \text{ and } 0 \in \bar{v}(B) \\
1 \in \bar{v}(A \wedge B) & \Leftrightarrow & 1 \in \bar{v}(A) \text{ and } 1 \in \bar{v}(B) \\
0 \in \bar{v}(A \wedge B) & \Leftrightarrow & 0 \in \bar{v}(A) \text{ or } 0 \in \bar{v}(B)
\end{array}$$

Note that we have to adopt this "double-entry bookkeeping" system due to the fact that truth and falsity or no longer mutually exclusive. Our clauses yield the following truth tables:

	$\neg$
t	**f**
f	**t**
b	**b**

$\vee$	**t**	**f**	**b**
t	**t**	**t**	**t**
f	**t**	**f**	**b**
b	**t**	**b**	**b**

$\wedge$	**t**	**f**	**b**
t	**t**	**f**	**b**
f	**f**	**f**	**f**
b	**b**	**f**	**b**

It is important to note that these tables agree with the classical truth tables with respect to the "classical" or "well-behaved" truth values, **t** and **f**. In particular, every classical valuation is an LP valuation (but not vice versa). It is also worth noting that, *modulo* notation, these are also the truth tables for Kleene's strong three-valued logic [91, §64]. (The only difference is that in Kleene's logic **b** is non-designated.)

We define semantic consequence as truth-preservation: A is a semantic consequence of Γ (in symbols, $\Gamma \models A$) just in case every valuation that makes all the elements of Γ true makes A true . (It is important to keep in mind here that 'true' means: either *exclusively true* (**t**) or *both true and false* (**b**).)

Let us use 'CPL' to denote classical propositional logic, as characterized by a function $v : At \longrightarrow \{\mathbf{t}, \mathbf{f}\}$ that is extended to all formulas via the usual classical clauses or truth tables. Here are two notable theorems about the relation between LP and CPL (cf. Priest [124]):

Theorem 5 (subclassicality of LP) If $\Gamma \models_{\mathsf{LP}} A$ then $\Gamma \models_{\mathsf{CPL}} A$.

Proof. Suppose $\Gamma \not\models_{\mathsf{CPL}} A$. Then there is a CPL valuation verifying (i.e. making true) every element of Γ but not A. Thus, since all CPL valuations are LP valuations, there is an LP valuation verifying every element of Γ but not A. Thus $\Gamma \not\models_{\mathsf{LP}} A$. ■

Lemma 11 Let v be an LP valuation, and let v_{CPL} be the CPL valuation that is just like v except that for all $p \in At$, if $v(p) = \mathbf{b}$, $v_{\mathsf{CPL}}(p) = \mathbf{t}$. Then:

- if $\bar{v}_{\mathsf{CPL}}(A) = \mathbf{t}$, then $1 \in \bar{v}(A)$
- if $\bar{v}_{\mathsf{CPL}}(A) = \mathbf{f}$, then $0 \in \bar{v}(A)$.

Proof. A simple induction on the length of A. If A is atomic, the result is immediate. Here is the inductive step for $A = B \wedge C$. First suppose that $\bar{v}_{\mathsf{CPL}}(B \wedge C) = \mathbf{t}$. Then $\bar{v}_{\mathsf{CPL}}(B) = \bar{v}_{\mathsf{CPL}}(C) = \mathbf{t}$. Thus, by the inductive hypothesis, $1 \in \bar{v}(B)$ and $1 \in \bar{v}(C)$, whence $1 \in \bar{v}(B \wedge C)$. Now suppose $\bar{v}_{\mathsf{CPL}}(B \wedge C) = \mathbf{f}$. Then $\bar{v}_{\mathsf{CPL}}(B) = \mathbf{f}$ or $\bar{v}_{\mathsf{CPL}}(C) = \mathbf{f}$. Thus, by the inductive hypothesis, $0 \in \bar{v}(B)$ or $0 \in \bar{v}(C)$, whence $0 \in \bar{v}(B \wedge C)$. ■

Theorem 6 (LP preserves classical tautologies) If $\models_{\mathsf{CPL}} A$ then $\models_{\mathsf{LP}} A$.

Proof. Suppose $\not\models_{\mathsf{LP}} A$. Then there is an LP valuation, v, such that $1 \notin \bar{v}(A)$. Thus, by Lemma 11, $\bar{v}_{\mathsf{CPL}}(A) \neq \mathbf{t}$, whence $\not\models_{\mathsf{CPL}} A$. ■

This last theorem may be somewhat surprising, as it entails that the law of non-contradiction holds in LP, i.e. $\models_{\mathsf{LP}} \neg(A \wedge \neg A)$. The explanation for this somewhat ironic fact is that while $A \wedge \neg A$ is sometimes true, it is also always false.

Many classically valid inference schemas are valid in LP. For example, in LP we have:

1. $A, B \models A \wedge B$ (conjunction introduction)
2. $A \wedge B \models A$ $[B]$ (conjunction elimination)
3. A $[B] \models A \vee B$ (disjunction introduction)
4. if $A \models C$ and $B \models C$ then $A \vee B \models C$ (disjunction elimination[17])

[17]This is, of course, but one version of disjunction elimination. Another version, $A \supset C, B \supset C, A \vee B \vdash C$, does not hold in LP (if $A \supset B$ is defined as $\neg A \vee B$).

5. $A \models \neg\neg A$ and $\neg\neg A \models A$ (double negation)

6. $\neg(A \wedge B) \models \neg A \vee \neg B$ [etc.] (De Morgan's rules)

Notably, the following *do not* hold for LP:

1. $A, \neg A \models B$ (explosion)

2. $A \vee B, \neg A \models B$ and $\neg A \vee B, A \models B$ (disjunctive syllogism)

3. $A \supset B, A \models B$ (modus ponens)

That explosion fails in LP is, of course, no surprise. That disjunctive syllogism fails may be seen as a serious flaw, as DS is a commonly used and intuitively correct pattern of inference. William Hanson makes the point as follows:

> Imagine someone who, under oath, admits to having received a message, but denies that it conveyed the information that B. If it were later revealed that the message in question was $\{A, \neg A \vee B\}$, perjury charges would be in order. [76, p. 666][18]

I will have more to say about DS a bit later in the chapter.

Given that $A \supset B$ is defined as $\neg A \vee B$, modus ponens falls along with disjunctive syllogism in LP. This is clearly intolerable. Surely a conditional connective that does not even satisfy MP is no conditional connective at all. I turn to this problem in the next section.

3.3.3 Defining a detachable conditional

The truth table for the material conditional in LP is:

$\supset$	$\{1\}$	$\{0\}$	$\{1,0\}$
$\{1\}$	$\{1\}$	$\{0\}$	$\{1,0\}$
$\{0\}$	$\{1\}$	$\{1\}$	$\{1\}$
$\{1,0\}$	$\{1\}$	$\{\mathbf{1},\mathbf{0}\}$	$\{1,0\}$

[18] I have changed the author's '$-$' to '$\neg$'.

Note that the only case that is problematic—i.e., the only one that constitutes a counterexample to MP—is the case in which the antecedent is assigned $\{1, 0\}$ and the consequent is assigned $\{0\}$. The principle of minimal mutilation dictates that we change the truth table as little as is necessary to render MP valid. That means simply letting $A \supset B$ take $\{0\}$ instead of $\{1,0\}$ when A takes $\{1,0\}$ and B takes $\{0\}$. That is, we change the truth table to:

$\supset$	$\{1\}$	$\{0\}$	$\{1,0\}$
$\{1\}$	$\{1\}$	$\{0\}$	$\{1,0\}$
$\{0\}$	$\{1\}$	$\{1\}$	$\{1\}$
$\{1,0\}$	$\{1\}$	$\{\mathbf{0}\}$	$\{1,0\}$

Now, if the PMM were the *only* justification for this new truth table, we would be open to the charge that we are "cheating"—i.e., violating adequacy condition number three (from §3.3.1). However, the revised truth table happens to be exactly the one determined by the following quite intuitive truth and falsity conditions (which, incidentally, determine the classical truth table for $\supset$ when applied to the classical truth values):

- $A \supset B$ is true just in case either A is not true or B is true
- $A \supset B$ is false just in case A is true and B is false

The system resulting from adding $\supset$ as a primitive connective with the truth table above is known in the literature as PI^{s}, $\mathsf{RM}_3^{\supset}$, and Pac.[19] I will call it Pac. In addition to modus ponens, Pac validates some other important inferences involving $\supset$ which are invalid in LP, such as:

- $A \supset B, B \supset C \,/\, A \supset C$ (hypothetical syllogism)
- $A \vee B, A \supset C, B \supset C \,/\, C$ (disjunction elimination)

Notably, Pac validates neither of the following:

- $A \supset B, \neg B \,/\, \neg A$ (modus tollens)
- $A \supset B \,/\, \neg B \supset \neg A$ (contraposition)

[19] See Batens [17] and Avron [15] [16].

However, a strengthened, "contraposible" conditional can be defined as follows:

$$A \to B =_{df} (A \supset B) \wedge (\neg B \supset \neg A).$$

In Pac we have:

- $A \to B, \neg B \models \neg A$ (modus tollens)
- $A \to B \models \neg B \to \neg A$ (contraposition)

$\to$ has the following truth table:

$\to$	$\{1\}$	$\{0\}$	$\{1,0\}$
$\{1\}$	$\{1\}$	$\{0\}$	$\{0\}$
$\{0\}$	$\{1\}$	$\{1\}$	$\{1\}$
$\{1,0\}$	$\{1\}$	$\{0\}$	$\{1,0\}$

The logic that results from taking $\to$ as primitive instead of $\supset$ is known in the literature as RM_3.[20] It is worth noting that $A \supset B$ can be defined in RM_3 as $(A \to B) \vee B$; hence Pac and RM_3 are, in essence, the same logic.[21]

3.3.4 Boolean negation

As noted before, disjunctive syllogism, $A \vee B, \neg A / B$, is invalid in LP. It also, of course, fails in Pac. However, DS fails to preserve truth only when A is "ill-behaved," i.e. both true and false. Thus, it seems that the inference ought to go through if we add a premise to the effect that A is *well*-behaved, i.e. *not* both true and false. What should this premise be? An obvious candidate is $\neg(A \wedge \neg A)$. This does not work, however, for $\neg(A \wedge \neg A)$ does not really say that A is well-behaved; it is, after all, true in *all* valuations, even those in which A is *ill*-behaved. In fact, there is no formula $\Phi(A)$ in Pac that is true in exactly those valuations in which A is well-behaved. (For suppose there is, and consider the valuation in which every atomic formula is both true and false. The truth tables ensure that this universal ill-behavedness extends to all formulas. Thus all formulas, including $\Phi(A)$ (whatever it may be), are true (as

[20] See e.g. Priest [127, §7.4].
[21] See Avron [15, p. 287].

well as false) in this valuation.) This is a serious and objectionable limitation in the expressive power of Pac (and, of course, LP).[22]

Logicians sometimes distinguish between the so-called "De Morgan" and "Boolean" truth conditions for negation (I use '$\sim$' as a generic negation connective):

De Morgan:	$\sim A$ is true iff A is false
Boolean:	$\sim A$ is true iff A is not true

This distinction is often ignored—quite understandably, since it collapses when truth and falsity are treated as mutually exclusive and exhaustive, as in classical logic. It is an important distinction in paraconsistent logic (and certain other non-classical logics), however. The $\neg$ connective of LP, and Pac expresses De Morgan negation. Boolean negation is simply inexpressible in LP and Pac. So let us introduce a new connective, #, such that $\#A$ is uniquely true if A is not true, and uniquely false otherwise. This connective has the following truth table:

	#
t	**f**
f	**t**
b	**f**

I will use 'P3' to denote the logic that results from adding # to Pac (or to LP; as we will see, the results are equivalent). We can now define a well-behavedness or classicality connective, ©, as follows: $©A = \#(A \wedge \neg A)$.[23] This connective has the following truth table:

	©
t	**t**
f	**t**
b	**f**

Intuitively, $©A$ says that A is well-behaved, i.e. A takes a classical truth value. Adding the premise $©A$ to DS yields a truth-preserving inference: $A \vee B, \neg A, ©A \,/\, B$. In fact, *all* classically valid inferences go through when embellished with appropriate consistency assumptions. (This is proved later, in Theorem 11.)

[22] For a particularly cogent argument to this effect, see Batens [18]. See also Parsons [122] and my [113].

[23] If we allow truth-value *gaps* as well as gluts, then we must define $©A$ differently, e.g. as $(A \vee \neg A) \wedge \#(A \wedge \neg A)$, which says that A as either true or false, but is not both.

Note that we can now define $A \supset B$ as $\#A \vee B$ (check the truth tables).

As I mentioned earlier, P3 is equivalent to the logic J_3 introduced by D'Ottaviano and da Costa [50] in 1970. The main difference is that instead of $\#$, J_3 has the connective $\triangledown$, with the following truth table:

	$\triangledown$
t	**t**
f	**f**
b	**t**

In J_3, the Boolean negation of A can be expressed with the formula $\neg\triangledown A$, and the well-behavedness of A can be expressed with the formula $\neg\triangledown(A \wedge \neg A)$.

3.3.5 Multiple conclusions

Until now we have taken for granted that logical consequence is a relation between a set of sentences and a single sentence, and that an inference is an ordered pair consisting of a set of sentences and a single sentence. At this point it may be helpful to generalize these notions. For the moment, let us think of logical consequence as a relation between a set of sentences and another (or perhaps the same) set of sentences. $\Gamma \vdash \Delta$ means that the set Δ is a consequence of the set Γ. Similarly, Γ / Δ is the inference having the elements of Γ as its premises and the elements of Δ as its conclusions.

The notion of truth-preservation is now understood as follows. The inference from Γ to Δ is truth-preserving just in case the truth of *every* element of Γ guarantees the truth of *some* element of Δ. Or, contrapositively, the untruth of every element of Δ guarantees the untruth of some element of Γ. Clearly a single-conclusion consequence relation is just a special type of multiple-conclusion consequence relation—one in which the conclusion set is required to be a singleton.

One may wonder why we do not say that an inference is truth-preserving just in case *all* of its conclusions are true whenever all of its premises are true. That is, why don't we interpret the conclusion set *conjunctively* (like the premise set) rather than disjunctively? Basically, the reason is that the disjunctive interpretation of the conclusion set captures the *duality* of conjunction and disjunction in the appropriate way, while the conjunctive interpretation does not. Consider, for example, the introduction and elimination rules for conjunction. The introduction rule is $\ulcorner$from A, B

infer $A \wedge B$⌝. The elimination rules are ⌜from $A \wedge B$ infer A⌝ and ⌜from $A \wedge B$ infer B⌝. Now consider disjunction. The introduction rules are ⌜from A infer $A \vee B$⌝ and ⌜from B infer $A \vee B$⌝. Treating disjunction as the "mirror image" of conjunction, the natural elimination rule would be ⌜from $A \vee B$ infer A, B⌝. Obviously, this rule is truth-preserving only if we interpret the conclusion set *disjunctively*. (Cf. Shoesmith and Smiley [143, pp. 2-3].)

While this multiple-conclusion conception of logical consequence and inference may seem unnatural (especially to those of us accustomed to the single-conclusion framework), it has significant advantages. As JC Beall and Greg Restall observe, "[a] profound symmetry emerges when one considers arguments involving multiple premises and multiple conclusions" [22, p. 14]. We have already seen (in the previous paragraph) how the disjunction elimination rule can be formulated more elegantly within the multiple-conclusion framework. A more pertinent example (in the present context) is that we can now say that a logical consequence relation, $\vdash$, is *paraconsistent* just in case

$$A, \neg A \nvdash \varnothing$$

and *paracomplete* just in case

$$\varnothing \nvdash A, \neg A.$$

Classical logic is of course neither paraconsistent nor paracomplete. LP, Pac, and P3 are each paraconsistent but not paracomplete. Intuitionistic logic is an example of a logic that is paracomplete but not paraconsistent.[24] An example of a logic that is both paraconsistent and paracomplete is FDE (for 'first degree entailment'), which is just like LP except that formulas are (also) allowed to be neither true nor false.[25] Elsewhere [114] I have referred to logics that are either paraconsistent or paracomplete (or both) as *paralogics*. Marcos [101] has used the delightful term *paranormal* to describe logics that are both paraconsistent and paracomplete.

While the multiple-conclusion framework is more elegant than the single-conclusion framework in some ways, it also introduces some complications that are tangential to present concerns. Hence, I will stick with the single-conclusion framework throughout most of this dissertation, though I will occasionally remark on the

[24] For an interesting study of the paraconsistent logic that is the exact dual of intuitionistic logic, see Urbas [152].

[25] See, e.g., Priest [127, chap. 8]. An interesting fact about FDE is that it has no theorems (i.e. no valid inferences with no premises). Every formula is untrue in the model that makes each atomic formula neither true nor false.

multiple-conclusion versions of the systems discussed.

3.4 The P systems (P3 and P4)

In this section I set out the semantics of P3 more formally, and provide a tableau-style proof theory that is sound and complete with respect to this semantics. Actually, I am going to start by defining a slightly weaker system, P4, that is *paracomplete* as well as paraconsistent, as it allows truth-value gaps (i.e. formulas that are neither true nor false). P3 is easily obtained by placing a straightforward restriction on P4 (namely, "no gaps"). While we haven't enountered any strong motivation for "going four-valued" *yet*, we *will* encounter such motivation in later chapters. I will refer to P3 and P4 as the "P systems". I will use '$\mathsf{P}\frac{3}{4}$' to generalize over P3 and P4 (so that when I assert something about $\mathsf{P}\frac{3}{4}$ I am really asserting it about both P3 and P4).

The language of the P systems is $L_{\mathsf{P}\frac{3}{4}} = L(\neg, \#, \vee)$. We have the following defined connectives:

$$\begin{array}{lll} A \wedge B & =_{df} & \neg(\neg A \vee \neg B) \\ A \supset B & =_{df} & \#A \vee B \\ A \equiv B & =_{df} & (A \supset B) \wedge (B \supset A) \\ A \to B & =_{df} & (A \supset B) \wedge (\neg B \supset \neg A) \\ A \leftrightarrow B & =_{df} & (A \to B) \wedge (B \to A) \\ \copyright A & =_{df} & (A \vee \neg A) \wedge \#(A \wedge \neg A) \end{array}$$

As usual, Γ, Δ, etc. always range over *finite* sets of sentences.

3.4.1 Semantics

Notation 3 (truth values) I use $\mathbf{t}$, $\mathbf{f}$, $\mathbf{b}$, and $\mathbf{n}$ ('true', 'false', 'both', and 'neither') as shorthand for $\{1\}$, $\{0\}$, $\{1, 0\}$, and $\varnothing$, respectively.

Definition 35 (P4 valuation) A P4 valuation is a function $v : At \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ that is extended to $\bar{v} : L_{\mathsf{P}\frac{3}{4}} \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ via the following clauses:

$$
\begin{array}{lcl}
\bar{v}(p) & = & v(p) \\
1 \in \bar{v}(\neg A) & \Leftrightarrow & 0 \in \bar{v}(A) \\
0 \in \bar{v}(\neg A) & \Leftrightarrow & 1 \in \bar{v}(A) \\
1 \in \bar{v}(\# A) & \Leftrightarrow & 1 \notin \bar{v}(A) \\
0 \in \bar{v}(\# A) & \Leftrightarrow & 1 \in \bar{v}(A) \\
1 \in \bar{v}(A \vee B) & \Leftrightarrow & 1 \in \bar{v}(A) \text{ or } 1 \in \bar{v}(B) \\
0 \in \bar{v}(A \vee B) & \Leftrightarrow & 0 \in \bar{v}(A) \text{ and } 0 \in \bar{v}(B)
\end{array}
$$

Definition 36 (P3 valuation) A P3 valuation is a P4 valuation, v, such that for all $p \in At, v(p) \neq \mathbf{n}$.

Fact 11 (exhaustion) For any P3 valuation v and $A \in L_{\mathsf{P}\frac{3}{4}}$, either $1 \in \bar{v}(A)$ or $0 \in \bar{v}(A)$.

Proof. A straightforward induction on the length of A. ∎

The reader can check that the following hold in all P4 valuations (and hence in all P3 valuations):

$$
\begin{array}{lcl}
1 \in \bar{v}(A \wedge B) & \Leftrightarrow & 1 \in \bar{v}(A) \text{ and } 1 \in \bar{v}(B) \\
0 \in \bar{v}(A \wedge B) & \Leftrightarrow & 0 \in \bar{v}(A) \text{ or } 0 \in \bar{v}(B) \\
1 \in \bar{v}(A \supset B) & \Leftrightarrow & 1 \in \bar{v}(A) \Rightarrow 1 \in \bar{v}(B) \\
0 \in \bar{v}(A \supset B) & \Leftrightarrow & 1 \in \bar{v}(A) \text{ and } 0 \in \bar{v}(B) \\
1 \in \bar{v}(A \equiv B) & \Leftrightarrow & 1 \in \bar{v}(A) \Leftrightarrow 1 \in \bar{v}(B) \\
0 \in \bar{v}(A \equiv B) & \Leftrightarrow & (1 \in \bar{v}(A) \text{ and } 0 \in \bar{v}(B)) \text{ or} \\
 & & (1 \in \bar{v}(B) \text{ and } 0 \in \bar{v}(A)) \\
1 \in \bar{v}(A \rightarrow B) & \Leftrightarrow & (1 \in \bar{v}(A) \Rightarrow 1 \in \bar{v}(B)) \text{ and} \\
 & & (0 \in \bar{v}(B) \Rightarrow 0 \in \bar{v}(A)) \\
0 \in \bar{v}(A \rightarrow B) & \Leftrightarrow & 1 \in \bar{v}(A) \text{ and } 0 \in \bar{v}(B) \\
1 \in \bar{v}(A \leftrightarrow B) & \Leftrightarrow & \bar{v}(A) = \bar{v}(B) \\
0 \in \bar{v}(A \leftrightarrow B) & \Leftrightarrow & (1 \in \bar{v}(A) \text{ and } 0 \in \bar{v}(B)) \text{ or} \\
 & & (1 \in \bar{v}(B) \text{ and } 0 \in \bar{v}(B)) \\
1 \in \bar{v}(©A) & \Leftrightarrow & \bar{v}(A) \in \{\mathbf{t}, \mathbf{f}\} \\
0 \in \bar{v}(©A) & \Leftrightarrow & \bar{v}(A) \in \{\mathbf{b}, \mathbf{n}\}
\end{array}
$$

For easy reference we display the truth tables here:

A	$\neg A$	$\#A$	©A
t	**f**	**f**	**t**
f	**t**	**t**	**t**
b	**b**	**f**	**f**
n	**n**	**t**	**f**

A	B	$A \vee B$	$A \wedge B$	$A \supset B$	$A \equiv B$	$A \to B$	$A \leftrightarrow B$
t	**t**	**t**	**t**	**t**	**t**	**t**	**t**
t	**f**	**t**	**f**	**f**	**f**	**f**	**f**
t	**b**	**t**	**b**	**b**	**b**	**f**	**f**
t	**n**	**t**	**n**	**n**	**n**	**n**	**n**
f	**t**	**t**	**f**	**t**	**f**	**t**	**f**
f	**f**	**f**	**f**	**t**	**t**	**t**	**t**
f	**b**	**b**	**f**	**t**	**f**	**t**	**f**
f	**n**	**n**	**f**	**t**	**t**	**t**	**n**
b	**t**	**t**	**b**	**t**	**b**	**t**	**f**
b	**f**	**b**	**f**	**f**	**f**	**f**	**f**
b	**b**	**b**	**b**	**b**	**b**	**b**	**b**
b	**n**	**t**	**f**	**n**	**n**	**n**	**n**
n	**t**	**t**	**n**	**t**	**n**	**t**	**n**
n	**f**	**n**	**f**	**t**	**t**	**n**	**n**
n	**b**	**t**	**f**	**t**	**n**	**n**	**n**
n	**n**	**n**	**n**	**t**	**t**	**t**	**t**

Definition 37 (verification) A valuation v *verifies* a formula, A, just in case $1 \in \bar{v}(A)$, and verifies a set of formulas, Γ, just in case $1 \in \bar{v}(A)$ for all $A \in \Gamma$.

Definition 38 (semantic consequence) $\Gamma \models_{\mathsf{P}\frac{3}{4}} A$ just in case every $\mathsf{P}\frac{3}{4}$ valuation that verifies Γ verifies A.

Remark 10 Obviously, P3 and P4 are decidable via the matrix (truth table) method.

3.4.2 Proof theory

Definition 39 (initial list) The initial list for Γ / A, where $\Gamma = \{B_0, \ldots, B_n\}$, is

$$B_0 +$$
$$\vdots$$
$$B_n +$$
$$A -$$

Remark 11 A node of the form $A+$ indicates that A is true. A node of the form $A-$ indicates that A is not true. Thus the initial list corresponds to the assumption that all of the elements of Γ are true and A is not true. If we show that this assumption cannot hold, then we have shown that $\Gamma \vdash_{\mathsf{P}\frac{3}{4}} A$.

Remark 12 It is easy enough to define a proof theory for multiple-conclusion versions of our systems. The initial list for Γ / Δ, where $\Gamma = \{A_0, \ldots, A_n\}$ and $\Delta = \{B_0, \ldots, B_m\}$, is

$$A_0 +$$
$$\vdots$$
$$A_n +$$
$$B_0 -$$
$$\vdots$$
$$B_m -$$

The rest of the tableau theory is the same.

Definition 40 (closed branch, open branch) In the tableau theory for P4, a branch is *closed* iff nodes of the forms A + and A − occur on it. In P3 we have the additional closure condition: A − and $\neg A$ −. If a branch is not closed, then it is *open.*

The tableau rules for $\mathsf{P}\frac{3}{4}$ are as follows. (These rules are presented in undiluted form in Appendix B.)

Remark 13 The tableau rules for $\mathsf{P}\frac{3}{4}$ are to be understood in the same way as the tableau rules for the D systems. In particular, redundant nodes may not be added to a branch, a rule may be applied to a given set of predecessor nodes only once, and "null" applications of rules are now allowed. (Cf. Section 2.5.2.)

Remark 14 P3 and P4 have exactly the same tableau rules. The only difference in their respective tableau theories is that P3 has the additional closure condition mentioned in Definition 40.

Definition 41 The P^3_4 rule for double negation is:

$$\begin{array}{c} \scriptstyle[\neg\neg] \\ \neg\neg A \pm \\ \downarrow \\ A \pm \end{array}$$

This rule says that if a node of the form $\neg\neg A +$ ($\neg\neg A -$) occurs on a branch, then a node of the form $A +$ ($A -$) may be added to the top of that branch (provided it does not already occur thereon). For example, if $\neg\neg(p \vee q) -$ is on a branch, and $p \vee q -$ does not already occur on that branch, then $p \vee q -$ may be added to the tip of that branch.

Definition 42 The P^3_4 rule for true disjunction is:

$$\begin{array}{ccc} & \scriptstyle[\vee+] & \\ & A \vee B + & \\ \swarrow & & \searrow \\ A + & & B + \end{array}$$

This rule says that if a node of the form $A \vee B +$ occurs on a branch, then the branch may be split at the tip, with $A +$ being added to one side (if it is not already on the branch) and $B +$ to the other (if it is not already on the branch). For example, if $\neg q \vee p +$ occurs on a branch, and neither $\neg q +$ nor $p +$ already occurs on the branch, then the branch may be split at the tip, with $\neg q +$ being added to one side and $p +$ to the other.

Definition 43 The rule for *untrue* disjunction (not to be confused with *false* disjunction!—see the next definition) is:

$$\begin{array}{c} \scriptstyle[\vee-] \\ A \vee B - \\ \downarrow \\ A - \\ B - \end{array}$$

This rule says that if a node of the form $A \vee B -$ occurs on a branch, then nodes of the forms $A -$ and $B -$ may be added to the tip of that branch (provided they do not already occur on the branch). For example, if $(p \wedge q) \vee ©r$ $-$ is on a branch, then $p \wedge q$ $-$ and $©r$ $-$ may be added to the tip of that branch (provided they do not already occur on the branch).

Definition 44 The rule for false disjunction is:

$$\begin{array}{c} {\scriptstyle [\neg\vee+]} \\ \neg(A \vee B)\ + \\ \downarrow \\ \neg A\ + \\ \neg B\ + \end{array}$$

Definition 45 The rule for unfalse disjunction is:

$$\begin{array}{ccc} & {\scriptstyle [\neg\vee-]} & \\ & \neg(A \vee B)\ - & \\ \swarrow & & \searrow \\ \neg A\ - & & \neg B\ - \end{array}$$

Definition 46 The rule for Boolean negation are:

$$\begin{array}{c} {\scriptstyle [\#]} \\ \#A\ \pm \\ \downarrow \\ A\ \mp \end{array}$$

Remark 15 Note that the $\pm$ sign flips to a $\mp$ sign in this rule. This rule says that if a node of the form $\#A$ $+$ ($\#A$ $-$) occurs on a branch, then a node of the form A $-$ (A $+$) may be added to the tip of that branch. For example, if $\#(p \vee q)$ $+$ occurs on a branch, then $p \vee q$ $-$ may be added to the tip of that branch.

Definition 47 The rule for (De Morgan) negated Boolean negation is:

$$\begin{array}{c} {\scriptstyle [\neg\#]} \\ \neg\#A\ \pm \\ \downarrow \\ A\ \pm \end{array}$$

Remark 16 As usual, although it is not required, it is advisable to always apply non-branching rules before branching rules, to prevent unnecessary proliferation of branches.

Definition 48 (tableau) A P^3_4 tableau for the inference Γ / A is any tableau that results from 0 or more applications of P^3_4 tableau rules to the initial list for Γ / A.

Definition 49 (complete branch/tableau) A branch of a P^3_4 tableau is *complete* just in case it is closed or each of the following holds:

1. For each node of the form $\neg\neg A +$ on the branch, a node of the form $A +$ is on the branch.

2. For each node of the form $\neg\neg A -$ on the branch, a node of the form $A -$ is on the branch.

3. For each node of the form $A \vee B +$ on the branch, a node of the form $A +$ or $B +$ is on the branch.

4. For each node of the form $A \vee B -$ on the branch, nodes of the forms $A -$ and $B -$ are on the branch.

5. For each node of the form $\neg(A \vee B) +$ on the branch, nodes of the forms $\neg A +$ and $\neg B +$ are on the branch.

6. For each node of the form $\neg(A \vee B) -$ on the branch, a node of the form $\neg A -$ or $\neg B -$ is on the branch.

7. For each node of the form $\#A +$ on the branch, a node of the form $A -$ is on the branch.

8. For each node of the form $\#A -$ on the branch, a node of the form $A +$ is on the branch.

9. For each node of the form $\neg\#A +$ on the branch, a node of the form $A +$ is on the branch.

10. For each node of the form $\neg\#A -$ on the branch, a node of the form $A -$ is on the branch.

A tableau is complete iff all of its branches are complete.

Definition 50 (closed/open tableau) A tableau is closed just in case all of its branches are closed; otherwise it is open.

Remark 17 We will later prove the "all or nothing" principle, which says that if there is an open, complete $\mathsf{P}\frac{3}{4}$ tableau for an inference, then there is no closed P tableau for that inference.

Definition 51 (tableau-theoretic consequence) $\Gamma \Vdash_{\mathsf{P}\frac{3}{4}} A$ iff there is a closed $\mathsf{P}\frac{3}{4}$ tableau for $\Gamma \,/\, A$.

For convenience (only), we have the following derived rules:

$$\begin{array}{c} [\wedge+] \\ A \wedge B + \\ \downarrow \\ A + \\ B + \end{array} \qquad \begin{array}{c} [\wedge-] \\ A \wedge B - \\ \swarrow \searrow \\ A - \quad B - \end{array} \qquad \begin{array}{c} [\neg\wedge] \\ \neg(A \wedge B) \pm \\ \downarrow \\ \neg A \vee \neg B \pm \end{array} \qquad \begin{array}{c} [\neg\vee] \\ \neg(A \vee B) \pm \\ \downarrow \\ \neg A \wedge \neg B \pm \end{array}$$

$$\begin{array}{c} [\supset+] \\ A \supset B + \\ \swarrow \searrow \\ A - \quad B + \end{array} \qquad \begin{array}{c} [\supset-] \\ A \supset B - \\ \downarrow \\ A + \\ B - \end{array} \qquad \begin{array}{c} [\neg\supset] \\ \neg(A \supset B) \pm \\ \downarrow \\ A \wedge \neg B \pm \end{array} \qquad \begin{array}{c} [\equiv+] \\ A \equiv B + \\ \swarrow \searrow \\ \begin{array}{cc} A - & A + \\ B - & B + \end{array} \end{array}$$

$$\begin{array}{c} [\equiv-] \\ A \equiv B - \\ \swarrow \searrow \\ \begin{array}{cc} A + & A - \\ B - & B + \end{array} \end{array} \qquad \begin{array}{c} [\neg\equiv+] \\ \neg(A \equiv B) + \\ \swarrow \searrow \\ \begin{array}{cc} A + & B + \\ \neg B + & \neg A + \end{array} \end{array}$$

$$\begin{array}{c} [\neg\equiv-] \\ \neg(A \equiv B) - \\ \swarrow \qquad \searrow \\ \begin{array}{cc} \begin{array}{c} A - \\ \swarrow \searrow \\ B - \quad \neg A - \end{array} & \begin{array}{c} \neg B - \\ \swarrow \searrow \\ B - \quad \neg A - \end{array} \end{array} \end{array}$$

$$
\begin{array}{c}
[\to +] \\
A \to B\ + \\
\swarrow \qquad \searrow \\
\begin{array}{cc}
\begin{array}{c} A\ - \\ \swarrow\searrow \\ \neg B\ - \qquad \neg A\ + \end{array} &
\begin{array}{c} B+ \\ \swarrow\searrow \\ \neg B- \qquad \neg A\ + \end{array}
\end{array}
\end{array}
$$

$$
\begin{array}{ccc}
\begin{array}{c}
[\to -] \\
A \to B\ + \\
\swarrow\searrow \\
\begin{array}{cc} A\ + & \neg B\ + \\ B\ - & \neg A\ - \end{array}
\end{array}
&
\begin{array}{c}
[\neg\!\to +] \\
\neg(A \to B)\ + \\
\downarrow \\
A\ + \\
\neg B\ +
\end{array}
&
\begin{array}{c}
[\neg\!\to -] \\
\neg(A \to B)\ - \\
\swarrow\searrow \\
A\ - \qquad \neg B\ -
\end{array}
\end{array}
$$

$$
\begin{array}{c}
[\leftrightarrow +] \\
A \leftrightarrow B\ + \\
\swarrow \qquad \searrow \\
\begin{array}{cc}
\begin{array}{c} A\ - \\ \swarrow\searrow \\ \begin{array}{cc} \neg B\ - & \neg A\ + \\ B- & B- \\ \neg A- & \neg B+ \end{array} \end{array} &
\begin{array}{c} B+ \\ \swarrow\searrow \\ \begin{array}{cc} \neg B- & \neg A\ + \\ A+ & A+ \\ \neg A- & \neg B+ \end{array} \end{array}
\end{array}
\end{array}
$$

$$
\begin{array}{c}
[\leftrightarrow -] \\
A \leftrightarrow B\ + \\
\swarrow \qquad \swarrow \qquad \searrow \qquad \searrow \\
\begin{array}{cccc} A\ + & \neg B\ + & B+ & \neg A\ + \\ B- & \neg A- & A- & \neg B- \end{array}
\end{array}
$$

$$
\begin{array}{cc}
\begin{array}{c}
[\neg\!\to +] \\
\neg(A \leftrightarrow B)\ + \\
\swarrow\searrow \\
\begin{array}{cc} A\ + & B\ + \\ \neg B\ + & \neg A\ + \end{array}
\end{array}
&
\begin{array}{c}
[\neg\!\leftrightarrow -] \\
\neg(A \leftrightarrow B)\ + \\
\swarrow \qquad \searrow \\
\begin{array}{cc}
\begin{array}{c} A\ - \\ \swarrow\searrow \\ B\ - \qquad \neg A\ - \end{array} &
\begin{array}{c} \neg B- \\ \swarrow\searrow \\ B- \qquad \neg A\ - \end{array}
\end{array}
\end{array}
\end{array}
$$

$$
\begin{array}{cc}
\multicolumn{2}{c}{[\copyright+]} \\
\multicolumn{2}{c}{\copyright A +} \\
\multicolumn{2}{c}{\swarrow \searrow} \\
A + & A - \\
\neg A - & \neg A +
\end{array}
\qquad
\begin{array}{cc}
\multicolumn{2}{c}{[\copyright-]} \\
\multicolumn{2}{c}{\copyright A -} \\
\multicolumn{2}{c}{\swarrow \searrow} \\
A + & A - \\
\neg A + & \neg A -
\end{array}
\qquad
\begin{array}{c}
[\neg\copyright] \\
\neg\copyright A \pm \\
\downarrow \\
\copyright A \mp
\end{array}
$$

As usual, it must be emphasized that these rules, while handy, are completely superfluous, and are not included in the "official" proof theory.

3.4.3 Example proofs

Example 12 (explosion) Here is a proof that $p, \neg p \nVdash_{\mathsf{P3}} q$:

$$
\begin{array}{rl}
p + & \text{initial list} \\
\neg p + & \text{"} \\
q - & \text{"} \\
\uparrow & \text{(no rules can be applied)}
\end{array}
$$

We can read the obvious counterexample off the open branch: let $v(p) = \mathbf{b}$, $v(q) = \mathbf{f}$. (The procedure for reading counterexamples off open branches is made more explicit in Definition 54, p. 124.)

Example 13 (modus ponens) Here is a proof that $A \supset B, A \Vdash_{\mathsf{P4}} B$, using the derived rule for true $\supset$:

$$
\begin{array}{cc}
\multicolumn{2}{c}{A \supset B + \checkmark} \\
\multicolumn{2}{c}{A +} \\
\multicolumn{2}{c}{B -} \\
\multicolumn{2}{c}{\swarrow \searrow} \\
A - & B + \\
* & *
\end{array}
$$

Example 14 (modus tollens for $\rightarrow$) Here is a proof that $A \rightarrow B, \neg B \Vdash_{\mathsf{P4}} \neg A$, using the derived rule for true $\rightarrow$:

$$
\begin{array}{c}
A \to B + \checkmark \\
\neg B + \\
\neg A - \\
\swarrow \searrow \\
\begin{array}{cc}
\begin{array}{c} A - \\ \swarrow \searrow \\ \begin{array}{cc} \neg B - & \neg A + \\ * & * \end{array} \end{array}
&
\begin{array}{c} B + \\ \swarrow \searrow \\ \begin{array}{cc} \neg B - & \neg A + \\ * & * \end{array} \end{array}
\end{array}
\end{array}
$$

Example 15 (De Morgan non-contradiction) Here is a proof that $\nVdash_{\mathsf{P4}} \neg(p \wedge \neg p)$, using primitive and derived rules:

1.	$\neg(p \wedge \neg p) - \checkmark$	initial list
2.	$\neg p \vee \neg\neg p - \checkmark$	from 1 by $[\neg\wedge]$
3.	$\neg p -$	from 2 by $[\vee-]$
4.	$\neg\neg p - \checkmark$	from 2 by $[\vee-]$
5.	$p -$	from 4 by $[\neg\neg]$
	$\uparrow$	(open and complete)

Reading the obvious counterexample off the open branch, we let $v(p) = \mathbf{n}$. (Note that the formula is valid in $\mathsf{P3}$, as $\mathsf{P3}$ precludes truth-value gaps.)

Example 16 (Boolean non-contradiction) Here is a proof that $\Vdash_{\mathsf{P4}} \#(A \wedge \#A)$, using primitive and derived rules:

1.	$\#(A \wedge \#A) - \checkmark$	initial list
2.	$A \wedge \#A + \checkmark$	from 1 by $[\#-]$
3.	$A +$	from 2 by $[\wedge+]$
4.	$\#A + \checkmark$	from 2 by $[\wedge+]$
5.	$A -$	from 4 by $[\#+]$
	$*$	from 3 and 5

3.4.4 Soundness and completeness

I now prove the soundness and completeness of the proof theory with respect to the semantics.

Definition 52 (faithful) Let b be a branch of a P^{3}_{4} tableau. A P^{3}_{4} valuation v is *faithful* to b just in case:

- if $A+$ is on b, then $1 \in \bar{v}(A)$
- if $A-$ is on b, then $1 \notin \bar{v}(A)$.

Lemma 12 (faith lemma) If a branch, b, of a $\mathsf{P}\frac{3}{4}$ tableau is closed, then no P valuation is faithful to b.

Proof. For reductio, suppose that the $\mathsf{P}\frac{3}{4}$ valuation v is faithful to b . There are two cases:

1. $\mathsf{P}\frac{3}{4} = \mathsf{P4}$. Then, since b is closed, nodes of the forms $A+$ and $A-$ occur on it. Thus $1 \in \bar{v}(A)$ and $1 \notin \bar{v}(A)$, which is impossible.
2. $\mathsf{P}\frac{3}{4} = \mathsf{P3}$. Then, since b is closed, nodes of the forms $A+$ and $A-$, or $A-$ and $\neg A-$, occur on it. Thus either $1 \in \bar{v}(A)$ and $1 \notin \bar{v}(A)$, which is impossible, or $1 \notin \bar{v}(A)$ and $1 \notin \bar{v}(\neg A)$, which is also (by the Exhaustion Lemma) impossible.

■

Notation 4 Recall that if b is a branch of a tableau, $b(\mathcal{N}_1, \cdots, \mathcal{N}_i)$ is the branch that results from adding the nodes $\mathcal{N}_1, \cdots, \mathcal{N}_i$ to the tip of b.

Lemma 13 (soundness lemma) If a $\mathsf{P}\frac{3}{4}$ valuation, v, is faithful to a branch of a $\mathsf{P}\frac{3}{4}$ tableau, b, and a $\mathsf{P}\frac{3}{4}$ tableau rule is applied to b, then v is faithful to at least one of the branches thereby generated.

Proof. The proof is by cases. There are 7 cases to consider (one for each tableau rule). I prove only selected cases here; the remaining cases are similar.

- **Case 1** ($[\neg\neg]$). Suppose $\neg\neg A+$ is on b, and $[\neg\neg]$ is applied. Then $b(A+)$ is generated. Since v is faithful to b, $1 \in \bar{v}(\neg\neg A)$. Thus $0 \in \bar{v}(\neg A)$. Thus $1 \in \bar{v}(A)$. Thus v is faithful to $b(A+)$. Now suppose $\neg\neg A-$ is on b, and $[\neg\neg]$ is applied. Then $b(A-)$ is generated. Since v is faithful to b, $1 \notin \bar{v}(\neg\neg A)$. Thus $0 \notin \bar{v}(\neg A)$. Thus $1 \notin \bar{v}(A)$. Thus v is faithful to $b(A-)$.
- **Case 2** ($[\vee+]$). Suppose $A \vee B+$ is on b, and $[\vee+]$ is applied. Then $b(A+)$ and $b(B+)$ are generated. Since v is faithful to b, $1 \in \bar{v}(A \vee B)$. Thus $1 \in \bar{v}(A)$ or $1 \in \bar{v}(B)$. Thus v is faithful to either $b(A+)$ or $b(B+)$.

- **Case 6** ([#]). Suppose $\#A\ +$ is on b, and [#] is applied. Then $b(A\ -)$ is generated. Since v is faithful to b, $1 \in \bar{v}(\#A)$. Thus $1 \notin \bar{v}(A)$ Thus v is faithful to $b(A\ -)$. Now suppose that $\#A\ -$ is on b, and [#] is applied. Then $b(A\ +)$ is generated. Since v is faithful to b, $1 \notin \bar{v}(\#A)$. Thus $1 \in \bar{v}(A)$ Thus v is faithful to $b(A\ +)$.

- **Case 7** ([¬#]). Suppose $\neg\#A\ +$ is on b, and [¬#] is applied. Then $b(A\ +)$ is generated. Since v is faithful to b, $1 \in \bar{v}(\neg\#A)$. Thus $0 \in \bar{v}(\#A)$. Thus $1 \in \bar{v}(A)$ Thus v is faithful to $b(A\ +)$. Now suppose that $\neg\#A\ -$ is on b, and [¬#] is applied. Then $b(A\ -)$ is generated. Since v is faithful to b, $1 \notin \bar{v}(\neg\#A)$. Thus $0 \notin \bar{v}(\#A)$. Thus $1 \notin \bar{v}(A)$ Thus v is faithful to $b(A\ -)$.

■

Theorem 7 (soundness) If $\Gamma \Vdash_{\mathsf{P}^{\frac{3}{4}}} A$ then $\Gamma \models_{\mathsf{P}^{\frac{3}{4}}} A$.

Proof. We prove the contrapositive. Suppose $\Gamma \not\models_{\mathsf{P}^{\frac{3}{4}}} A$. Then there is a $\mathsf{P}^{\frac{3}{4}}$ valuation v such that $1 \in \bar{v}(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}(A)$ v is faithful to the initial list for Γ / A. Moreover, by the Soundness Lemma (Lemma 13), each subsequent application of a $\mathsf{P}^{\frac{3}{4}}$ tableau rule will yield at least one branch to which v is faithful. Thus every $\mathsf{P}^{\frac{3}{4}}$ tableau for Γ / A has at least one branch to which v is faithful. Let T be a $\mathsf{P}^{\frac{3}{4}}$ tableau for Γ / A, and let b be one of the branches of T to which v is faithful. By the Faith Lemma (Lemma 12), b is open. Thus T is open. But T was an arbitrarily chosen $\mathsf{P}^{\frac{3}{4}}$ tableau for Γ / A. Thus there is no closed $\mathsf{P}^{\frac{3}{4}}$ tableau for Γ / A. Thus $\Gamma \nVdash_{\mathsf{P}^{\frac{3}{4}}} A$. ■

Definition 53 (length of a formula) The *length* of a formula is simply the number of symbols in it (including parentheses). For example, the length of the formula $\neg((p \wedge \#q) \supset \neg r)$ is 12.

Lemma 14 (completion lemma) It is possible to specify a *complete* $\mathsf{P}^{\frac{3}{4}}$ tableau for any inference Γ / A.

Proof. This is true because each $\mathsf{P}^{\frac{3}{4}}$ tableau for Γ / A, no matter how it is constructed, will terminate (i.e. be complete) after a finite number of steps. Each $\mathsf{P}^{\frac{3}{4}}$ tableau begins with a finite list of nodes, each of which contains a formula of finite length. The tableau is expanded by repeatedly choosing an unchecked node on an open branch, checking off that node, and adding to the tableau some finite number of nodes, each of which contains a formula of *shorter* length than the formula in the node that was checked. Eventually, then, a point must be reached where all formulas on unchecked nodes on open branches have length 1 (i.e. an atomic formula) or 2 (i.e. a (De Morgan) negated atomic formula). At that point, no further rules can be applied, so the tableau is complete.[26] ■

Definition 54 (induced valuation) Let b be an open, complete branch of a P^3_4 tableau. The P^3_4 valuation *induced* by b is the valuation v determined by the following chart:

<table>
<tr><th>$p+$</th><th>$p-$</th><th>$\neg p+$</th><th>$\neg p-$</th><th>P4</th><th>P3</th></tr>
<tr><td>Y</td><td>Y</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>Y</td><td>N</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>N</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>N</td><td>N</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>N</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>N</td><td>Y</td><td>N</td><td colspan="2">$v(p) = \mathbf{b}$</td></tr>
<tr><td>Y</td><td>N</td><td>N</td><td>Y</td><td colspan="2">$v(p) = \mathbf{t}$</td></tr>
<tr><td>Y</td><td>N</td><td>N</td><td>N</td><td colspan="2">$v(p) = \mathbf{t}$</td></tr>
<tr><td>N</td><td>Y</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>N</td><td>Y</td><td>Y</td><td>N</td><td colspan="2">$v(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>Y</td><td>N</td><td>Y</td><td>$v(p) = \mathbf{n}$</td><td>$*$</td></tr>
<tr><td>N</td><td>Y</td><td>N</td><td>N</td><td colspan="2">$v(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>N</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>N</td><td>N</td><td>Y</td><td>N</td><td colspan="2">$v(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>N</td><td>N</td><td>Y</td><td colspan="2">$v(p) = \mathbf{t}$</td></tr>
<tr><td>N</td><td>N</td><td>N</td><td>N</td><td colspan="2">$v(p) = \mathbf{t}$</td></tr>
</table>

The chart is to be read as follows. For all $p \in At$, find the row in the chart that accurately describes p's status on the branch. (For example, if $p-$ and $\neg p+$ are on the branch, but neither $p+$ nor $\neg p-$ is on the branch, then the relevant row would be N/Y/Y/N.) If P^3_4 is P4, assign p the truth value specified in the P4 column. If P^3_4 is P3, assign p the truth value specified in the P3 column. An asterisk ('$*$') indicates that the combination in question is impossible, as the branch would be closed.

[26] This proof is similar to the one given by Richard Jeffrey (for truth trees for classical propositional logic) in [87, pp. 32-3].

Remark 18 Note that the P^3_4 valuation induced by an open branch is always well-defined.

Lemma 15 (completeness lemma) Let b be an open, complete branch of a P^3_4 tableau. Let v be the P^3_4 valuation induced by b. Then for all $A \in L_{\mathsf{P}^3_4}$:

1. if $A+$ is on b, then $1 \in \bar{v}(A)$
2. if $A-$ is on b, then $1 \notin \bar{v}(A)$
3. if $\neg A+$ is on b, then $0 \in \bar{v}(A)$
4. if $\neg A-$ is on b, then $0 \notin \bar{v}(A)$.

Proof. An induction on the length of A. If A is atomic, then the result holds by the definition of an induced model. (Check the chart.)

Inductive step for $A = \neg B$:

1. Suppose $\neg B+$ is on b. Then, by the inductive hypothesis (IH), $0 \in \bar{v}(B)$. Thus $1 \in \bar{v}(\neg B)$.
2. Suppose $\neg B-$ is on b. Then, by the IH, $0 \notin \bar{v}(B)$. Thus $1 \notin \bar{v}(\neg B)$.
3. Suppose $\neg\neg B+$ is on b. Then $B+$ is on b. Thus, by the IH, $1 \in \bar{v}(B)$. Thus $0 \in \bar{v}(\neg B)$.
4. Suppose $\neg\neg B-$ is on b. Then $B-$ is on b. Thus, by the IH, $1 \notin \bar{v}(B)$. Thus $0 \notin \bar{v}(\neg B)$.

Inductive step for $A = \#B$:

1. Suppose $\#B+$ is on b. Then $B-$ is on b. Thus, by the IH, $1 \notin \bar{v}(B)$. Thus $1 \in \bar{v}(\#B)$.
2. Suppose $\#B-$ is on b. Then $B+$ is on b. Thus, by the IH, $1 \in \bar{v}(B)$. Thus $1 \notin \bar{v}(\#B)$.
3. Suppose $\neg\#B+$ is on b. Then $B+$ is on b. Thus, by the IH, $1 \in \bar{v}(B)$. Thus $0 \in \bar{v}(\#B)$.

4. Suppose $\neg \# B\ -$ is on b. Then $B\ -$ is on b. Thus, by the IH, $1 \notin \bar{v}(B)$. Thus $0 \notin \bar{v}(\# B)$.

Inductive step for $A = B \vee C$:

1. Suppose $B \vee C\ +$ is on b. Then either $B\ +$ or $C\ +$ is on b. Thus, by the IH, either $1 \in \bar{v}(B)$ or $1 \in \bar{v}(C)$. Thus $1 \in \bar{v}(B \vee C)$.

2. Suppose $B \vee C\ -$ is on b. Then $B\ -$ and $C\ -$ are on b. Thus, by the IH, $1 \notin \bar{v}(B)$ and $1 \notin \bar{v}(C)$. Thus $1 \notin \bar{v}(B \vee C)$.

3. Suppose $\neg(B \vee C)\ +$ is on b. Then $\neg B\ +$ and $\neg C\ +$ are on b. Thus, by the IH, $0 \in \bar{v}(B)$ and $0 \in \bar{v}(C)$. Thus $0 \in \bar{v}(B \vee C)$.

4. Suppose $\neg(B \vee C)\ -$ is on b. Then either $\neg B\ -$ or $\neg C\ -$ is on b. Thus, by the IH, either $0 \notin \bar{v}(B)$ or $0 \notin \bar{v}(C)$. Thus $0 \notin \bar{v}(B \vee C)$.

■

Theorem 8 (completeness) If $\Gamma \models_{\mathsf{P}\frac{3}{4}} A$ then $\Gamma \Vdash_{\mathsf{P}\frac{3}{4}} A$.

Proof. Suppose $\Gamma \nVdash_{\mathsf{P}\frac{3}{4}} A$. Then there is no closed $\mathsf{P}\frac{3}{4}$ tableau for $\Gamma \diagup A$. By the Completion Lemma (Lemma 14), we know that there is an open, complete $\mathsf{P}\frac{3}{4}$ tableau for $\Gamma \diagup A$. Let T be such a tableau, and choose some open, complete branch b of T. Let v be the $\mathsf{P}\frac{3}{4}$ valuation induced by b. By the Completeness Lemma (Lemma 15), $1 \in \bar{v}(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}(A)$. Thus $\Gamma \not\models_{\mathsf{P}\frac{3}{4}} A$. ■

Corollary 1 (all or nothing) If there is an open, complete $\mathsf{P}\frac{3}{4}$ tableau for $\Gamma \diagup A$, then there is no closed $\mathsf{P}\frac{3}{4}$ tableau for $\Gamma \diagup A$.

Proof. Suppose there is an open, complete $\mathsf{P}\frac{3}{4}$ tableau for $\Gamma \diagup A$. Pick an open branch, b, of this tableau, and let v be the $\mathsf{P}\frac{3}{4}$ valuation induced by b. By the Completeness Lemma (Lemma 15), $1 \in \bar{v}(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}(A)$. Thus $\Gamma \not\models_{\mathsf{P}\frac{3}{4}} A$. Thus, by the Soundness Theorem (Theorem 7), $\Gamma \nVdash_{\mathsf{P}\frac{3}{4}} A$. Thus there is no closed $\mathsf{P}\frac{3}{4}$ tableau for $\Gamma \diagup A$. ■

3.5 Notable features of the P systems

In this section I highlight some significant features of the P systems ($\mathsf{P3}$ and $\mathsf{P4}$). Remember that I use $\vdash$ to generalize over $\models$ and $\Vdash$.

3.5.1 Relations between systems

Fact 12 $\mathsf{P3}$ is a proper extension of $\mathsf{P4}$.

Proof. That $\mathsf{P3}$ is an extension of $\mathsf{P4}$ follows from the fact that every $\mathsf{P3}$ model is a $\mathsf{P4}$ model (so that whenever there is a $\mathsf{P3}$ counterexample for an inference, there is a $\mathsf{P4}$ counterexample for that inference). That $\mathsf{P3}$ is a *proper* extension of $\mathsf{P4}$ follows from the fact that (e.g.) $\vdash_{\mathsf{P3}} A \vee \neg A$ but $\nvdash_{\mathsf{P4}} A \vee \neg A$. ■

Fact 13 CPL (classical propositional logic) is a proper extension of $\mathsf{P3}$ (with respect to the language $L_{\mathsf{CPL}} \cap L_{\mathsf{P}\frac{3}{4}} = L(\neg, \vee)$).

Proof. That CPL is an extension of $\mathsf{P3}$ follows from the fact that every CPL model is a $\mathsf{P3}$ model. That CPL is a *proper* extension of $\mathsf{P3}$ follows from the fact that (e.g.) $A, \neg A \vdash_{\mathsf{CPL}} B$ but $A, \neg A \nvdash_{\mathsf{P3}} B$. ■

Remark 19 Note that, unlike LP, $\mathsf{P3}$ does not validate all classical tautologies. For example, $A \supset (\neg A \supset B)$ is invalid in $\mathsf{P3}$. (LP either doesn't include $\supset$ in its language, or defines it as $A \supset B =_{df} \neg A \vee B$.)

Remark 20 Note that, unlike FDE, $\mathsf{P4}$ *does* have some theorems. An example is $A \supset A$. (Like LP, FDE either doesn't include $\supset$ in its language, or defines it as $A \supset B =_{df} \neg A \vee B$.)

3.5.2 Basic logical properties

Fact 14 $\mathsf{P}\frac{3}{4}$ enjoys the basic "Tarskian" properties:[27]

1. $A \in \Gamma \Rightarrow \Gamma \vdash_{\mathsf{D+}} A$ [reflexivity]
2. $\Gamma \vdash_{\mathsf{D+}} A \Rightarrow \Gamma, \Delta \vdash_{\mathsf{D+}} A$ [monotonicity]
3. $(\Gamma \vdash_{\mathsf{D+}} A$ and $A \vdash_{\mathsf{D+}} B) \Rightarrow \Gamma \vdash_{\mathsf{D+}} B$ [transitivity]

Theorem 9 (deduction theorem for $\supset$) $\Gamma, A \vdash_{\mathsf{P}\frac{3}{4}} B$ iff $\Gamma \vdash_{\mathsf{P}\frac{3}{4}} A \supset B$.

Proof. For the right-to-left half, suppose $\Gamma, A \nvdash_{\mathsf{P}\frac{3}{4}} B$. Then there is a $\mathsf{P}\frac{3}{4}$ valuation, v, that verifies $\Gamma \cup \{A\}$ but not B. Thus v verifies A but not B. Thus v does not verify $A \supset B$. Thus v verifies Γ but not $A \supset B$. Thus $\Gamma \nvdash_{\mathsf{P}\frac{3}{4}} A \supset B$.

[27] Recall from Definition **??** that $Cn(\Gamma) = \{A : \Gamma \vdash A\}$.

For the left-to-right half, suppose $\Gamma \nvdash_{\mathsf{P}\frac{3}{4}} A \supset B$. Then there is a $\mathsf{P}\frac{3}{4}$ valuation, v, that verifies Γ but not $A \supset B$. Since v fails to verify $A \supset B$, v verifies A but not B. Thus v verifies $\Gamma \cup \{A\}$ but not B. Thus $\Gamma, A \nvdash_{\mathsf{P}\frac{3}{4}} B$. ■

Fact 15 The deduction theorem fails for $\rightarrow$.

Proof. Just observe that $p, q \vdash_{\mathsf{P}\frac{3}{4}} p$ but $p \nvdash_{\mathsf{P}\frac{3}{4}} q \rightarrow p$. (Counterexample: $v(p) = \mathbf{b}$, $v(q) = \mathbf{t}$.) ■

Fact 16 In $\mathsf{P}\frac{3}{4}$ we have:

1. $A \dashv\vdash B$ iff $\vdash A \equiv B$
2. if $\vdash A \leftrightarrow B$ then $\vdash A \equiv B$

Fact 17 In $\mathsf{P}\frac{3}{4}$ it is not the case that if $A \dashv\vdash B$ then $\vdash A \leftrightarrow B$.

Proof. Observe that in $\mathsf{P}\frac{3}{4}$ we have $p \vee \neg p \dashv\vdash q \vee \neg q$ but $\nvdash p \vee \neg p \leftrightarrow q \vee \neg q$. ■

Fact 18 (failure of replacement) In $\mathsf{P}\frac{3}{4}$ it is not the case that if $B \dashv\vdash B'$ then $A \dashv\vdash A[B/B']$.[28]

Proof. Observe that $p \vee \neg p \dashv\vdash q \vee \neg q$ but $\neg(p \vee \neg p) \nvdash \neg(q \vee \neg q)$. ■

Theorem 10 (alternative replacement) If $\vdash_{\mathsf{P}\frac{3}{4}} B \leftrightarrow B'$ then $\vdash_{\mathsf{P}\frac{3}{4}} A \leftrightarrow A[B/B']$.

Proof. (Cf. Chellas [43, pp. 125-6].) We first consider the possibility that $A = B$. In that case, $A[B/B'] = B'$. We need to show that $\vdash A \leftrightarrow A[B/B']$, i.e., $\vdash B \leftrightarrow B'$. But this is precisely what we are assuming. So the result holds when $A = B$. Thus we may assume henceforth that $A \neq B$.

We now proceed by induction on the length of A. Suppose A is atomic. Then, since $A \neq B$, $A[B/B'] = A$. We need to show that $\vdash A \leftrightarrow A[B/B']$, i.e., $\vdash A \leftrightarrow A$. But this holds trivially, since (as one can easily check) for *all* A, $\vdash A \leftrightarrow A$.

Now, assuming that the result holds for C and D, we need to show that it holds for $\neg C$, $C \vee D$, and ©C.

Suppose $A = \neg C$. By the inductive hypothesis (IH), $\vdash C \leftrightarrow C[B/B']$. Thus $C \leftrightarrow C[B/B']$ is verified by all $\mathsf{P}\frac{3}{4}$ valuations. Thus (by the truth condition for

[28] Recall from Definition 33 that $A[B/B']$ is the result of replacing each occurrence of B in A with B'.

$\leftrightarrow$) in all P^{3}_{4} valuations, $\bar{v}(C) = \bar{v}(C[B/B'])$. Thus (since $\neg$ is truth-functional) $\bar{v}(\neg C) = \bar{v}(\neg C[B/B'])$ in all P^{3}_{4} valuations. Thus (by the truth condition for $\leftrightarrow$ again), $\neg C \leftrightarrow \neg C[B/B']$ is verified by all P valuations. Thus $\vdash \neg C \leftrightarrow \neg C[B/B']$.

Suppose $A = C \vee D$. By the IH, $\vdash C \leftrightarrow C[B/B']$ and $\vdash D \leftrightarrow D[B/B']$. Thus (by the truth condition for $\leftrightarrow$), $\bar{v}(C) = \bar{v}(C[B/B'])$ and $\bar{v}(D) = \bar{v}(D[B/B'])$ in all P valuations. Thus (since $\vee$ is truth-functional), $\bar{v}(C \vee D) = \bar{v}(C[B/B'] \vee D[B/B'])$ in all P^{3}_{4} valuations. But note that $C[B/B'] \vee D[B/B'] = (C \vee D)[B/B']$. Thus $\bar{v}(C \vee D) = \bar{v}((C \vee D)[B/B'])$ in all P^{3}_{4} valuations. Thus (by the truth condition for $\leftrightarrow$ again), $\vdash (C \vee D) \leftrightarrow (C \vee D)[B/B']$.

Finally, suppose $A = \#C$. By the IH, $\vdash C \leftrightarrow C[B/B']$. Thus (by the truth condition for $\leftrightarrow$), $\bar{v}(C) = \bar{v}(C[B/B'])$ in all P^{3}_{4} valuations. Thus, since $\#$ is truth-functional, $\bar{v}(\#C) = \bar{v}(\#C[B/B'])$ in all P^{3}_{4} valuations. Thus (by the truth condition for $\leftrightarrow$ again) $\vdash \#C \leftrightarrow \#C[B/B']$. ■

Definition 55 (relevant) If $\Rrightarrow$ is a conditional connective of a logic L, then L is *relevant* with respect to $\Rrightarrow$ just in case $\vdash_{\mathsf{L}} A \Rrightarrow B$ only if A and B share an atomic formula.

Fact 19 P^{3}_{4} is not relevant with respect to $\supset$ or $\rightarrow$.

Proof. Use the truth table or tableau method to verify that (e.g.) $\vdash_{\mathsf{P}^{3}_{4}} \neg(p \supset p) \rightarrow (q \supset q)$, and hence $\vdash_{\mathsf{P}^{3}_{4}} \neg(p \supset p) \supset (q \supset q)$. ■

3.5.3 Miscellaneous validities/invalidities

Notation 5 I use $\Rrightarrow$ to generalize over $\supset$ and $\rightarrow$.

Fact 20 (positive principles) In P^{3}_{4} we have:

1. $A, B \vdash A \wedge B$ [conjunction introduction]
2. $A \wedge B \vdash A$ and $A \wedge B \vdash B$ [conjunction elimination]
3. $A \vdash A \vee B$ and $B \vdash A \vee B$ [disjunction introduction]
4. if $A \vdash C$ and $B \vdash C$, then $A \vee B \vdash C$ [disjunction elimination][29]

[29]Note that in the multiple-conclusion framework we can formulate this principle more elegantly as $A \vee B \vdash A, B$.

5. $A \Rrightarrow A$ [identity]

6. $A \Rrightarrow B, A \vdash B$ [modus ponens][30]

7. $A \Rrightarrow B, B \Rrightarrow C \vdash A \Rrightarrow C$ [hypothetical syllogism]

8. $B \vdash A \supset B$ but $B \nvdash A \rightarrow B$ [positive paradox]

9. $\vdash (A \supset B) \vee (B \supset A)$ but $\nvdash (A \rightarrow B) \vee (B \rightarrow A)$ [disjunctive paradox]

Fact 21 (LEM and LNC) The following hold:

1. $\vdash_{\mathsf{P3}} A \vee \neg A$ [law of excluded middle]

2. $\nvdash_{\mathsf{P4}} A \vee \neg A$

3. $\vdash_{\mathsf{P}\frac{3}{4}} A \vee \# A$ [law of excluded middle for Boolean negation]

4. $\vdash_{\mathsf{P3}} \neg(A \wedge \neg A)$ [law of non-contradiction]

5. $\nvdash_{\mathsf{P4}} \neg(A \wedge \neg A)$ [see Example 15]

6. $\vdash_{\mathsf{P}\frac{3}{4}} \#(A \wedge \# A)$ [law of non-contradiction for Boolean negation][31]

Fact 22 (double negation) In $\mathsf{P}\frac{3}{4}$ we have:

1. $\vdash A \leftrightarrow \neg\neg A$

2. $\vdash \#\# A \equiv A$

3. $\vdash A \rightarrow \#\# A$

4. $\nvdash \#\# p \rightarrow p$ (let $v(p) = \mathbf{b}$)

5. $\neg\neg A \dashv\vdash \neg \# A$

6. $\# \neg A \dashv\vdash \#\# A$

7. $\neg\neg A \dashv\vdash \#\# A$

Fact 23 (De Morgan laws) In $\mathsf{P}\frac{3}{4}$ we have:

---

[30] See Example 13.

[31] See Example 16.

1. $\vdash \neg(A \wedge B) \leftrightarrow \neg A \vee \neg B$
2. $\vdash \neg(A \vee B) \leftrightarrow \neg A \wedge \neg B$
3. $\vdash \#(A \wedge B) \leftrightarrow \#A \vee \#B$
4. $\vdash \#(A \vee B) \leftrightarrow \#A \wedge \#B$

Fact 24 Each of the following holds:

1. $\neg A \nvdash_{\mathsf{P}\frac{3}{4}} \#A$
2. $\#A \vdash_{\mathsf{P3}} \neg A$
3. $\#A \nvdash_{\mathsf{P4}} \neg A$

Fact 25 (material conditional) Each of the following holds:

1. $A \supset B \vdash_{\mathsf{P3}} \neg A \vee B$
2. $A \supset B \nvdash_{\mathsf{P4}} \neg A \vee B$
3. $\neg A \vee B \nvdash_{\mathsf{P}\frac{3}{4}} A \supset B$.

Fact 26 (disjunctive syllogism) In $\mathsf{P}\frac{3}{4}$ we have:

1. $A \vee B, \neg A \nvdash B$
2. $\neg A \vee B, A \nvdash B$
3. $A \vee B, \#A \vdash B$
4. $\#A \vee B, A \vdash B$

Fact 27 (modus tollens/contraposition) In $\mathsf{P}\frac{3}{4}$ we have:

1. $A \supset B, \neg B \nvdash \neg A$
2. $A \supset B \nvdash \neg B \supset \neg A$
3. $A \rightarrow B, \neg B \vdash \neg A$ [see Example 14]
4. $A \rightarrow B \vdash \neg B \rightarrow \neg A$

5. $A \supset B, \#B \vdash \#A$

6. $A \supset B \vdash \#B \supset \#A$

Fact 28 (reductio principles) In $\mathsf{P}\frac{3}{4}$ we have:

1. $A \Rrightarrow \neg A \vdash \neg A$ [simple reductio]

2. $A \supset B \wedge \neg B \nvdash \neg A$ [complex reductio]

3. $A \rightarrow B \wedge \neg B \vdash \neg A$

4. $A \supset B \wedge \#B \vdash \#A$

Fact 29 (explosion) Each of the following holds:

1. $A, \neg A \nvdash_{\mathsf{P}\frac{3}{4}} B$ [see Example 12]

2. $A \wedge \neg A \nvdash_{\mathsf{P}\frac{3}{4}} B$

3. $\nvdash_{\mathsf{P}\frac{3}{4}} A \supset (\neg A \supset B)$

4. $\nvdash_{\mathsf{P}\frac{3}{4}} A \rightarrow (\neg A \rightarrow B)$

5. $\nvdash_{\mathsf{P}\frac{3}{4}} A \wedge \neg A \supset B$

6. $A, \#A \vdash_{\mathsf{P}\frac{3}{4}} B$

7. $A \wedge \#A \vdash_{\mathsf{P}\frac{3}{4}} B$

8. $\vdash_{\mathsf{P}\frac{3}{4}} A \supset (\#A \supset B)$

9. $\vdash_{\mathsf{P}\frac{3}{4}} A \rightarrow (\#A \rightarrow B)$

10. $\vdash_{\mathsf{P}\frac{3}{4}} A \wedge \#A \supset B$

11. $\vdash_{\mathsf{P}\frac{3}{4}} A \wedge \#A \rightarrow B$

Fact 30 (gentle principles) The following hold for P:

1. $A, \neg A, ©A \vdash B$ [gentle explosion]

2. $A \vee B, \neg A, ©A \vdash B$ [gentle disjunctive syllogism]

3. $A \supset B, \neg B, \copyright B \vdash \neg A$ [gentle modus tollens]

4. $A \supset B \wedge \neg B, \copyright B \vdash \neg A$ [gentle complex reductio]

Remark 21 (P is an LFI) Carnielli, Coniglio, and Marcos [37] define a *logic of formal inconsistency*, or LFI, as any logic in which explosion is not valid, but *gentle* explosion is. (Their characterization of "gentle explosion" is somewhat more general than the version given above, allowing for other formulas to play the role of $\copyright A$.) Clearly P3 and P4 are LFIs.

As promised, we now show that, in a certain precise sense, all of classical logic can be "recaptured" in $\mathsf{P}\frac{3}{4}$.

3.5.4 The classical recapture

Definition 56 Let $\Gamma^{\copyright} = \{\copyright p : p \in At \text{ and } p \text{ occurs in an element of } \Gamma\}$.

Definition 57 (well-behaved consequence) $\Gamma \vdash^{\copyright}_{\mathsf{P}\frac{3}{4}} A$ iff $\Gamma, (\Gamma \cup \{A\})^{\copyright} \vdash_{\mathsf{P}\frac{3}{4}} A$.

Theorem 11 (classical recapture) If $\Gamma \vdash_{\mathsf{CPL}} A$ then $\Gamma \vdash^{\copyright}_{\mathsf{P}\frac{3}{4}} A$.

Proof. Suppose $\Gamma \nvdash^{\copyright}_{\mathsf{P}\frac{3}{4}} A$. Then $\Gamma, (\Gamma \cup \{A\})^{\copyright} \nvdash_{\mathsf{P}} A$. Thus there is a $\mathsf{P}\frac{3}{4}$ valuation, v, such that $1 \in \bar{v}(B)$ for all $B \in \Gamma \cup (\Gamma \cup \{A\})^{\copyright}$ and $1 \notin \bar{v}(A)$. Thus v assigns a classical truth value ($\mathbf{t}$ or $\mathbf{f}$) to each atomic formula occurring in a formula in $\Gamma \cup \{A\}$. Thus, by the truth tables, v assigns a classical truth value to each formula in $\Gamma \cup \{A\}$. Thus $\bar{v}(B) = \mathbf{t}$ for all $B \in \Gamma$ and $\bar{v}(A) = \mathbf{f}$. Let v_{CPL} be any classical valuation that agrees with v with respect to the elements of $(\Gamma \cup \{A\})^{\copyright}$. Then, since the matrices for CPL agree with the matrices for $\mathsf{P}\frac{3}{4}$ with respect to the classical truth values, $\bar{v}_{\mathsf{CPL}}(B) = \mathbf{t}$ for all $B \in \Gamma$ and $\bar{v}_{\mathsf{CPL}}(A) = \mathbf{f}$. Thus $\Gamma \nvdash_{\mathsf{CPL}} A$. ■

In plain English, the above theorem says that in P3 or P4 you are permitted to reason classically, provided that you explicitly state your assumption that all of the atomic formulas with which you are working are *well-behaved*.

3.6 The adequacy of the P systems

We now have at our disposal a well-motivated and technically sound paraconsistent logic on which to construct a conflict-tolerant deontic logic. P3, I submit, satisfies all of the adequacy criteria set out at the beginning of the chapter. Let us recall these criteria:

First, we specified that our paraconsisistent logic should have a clear, intuitive semantics. I believe that the semantics of P3 is quite clear and intuitive. The only major hurdle is accepting the idea of a formula's being both true and false. This does not require accepting that formulas can *really be* true and false (though one *may* wish to accept this); it only requires broadening our conception of the "set-ups" or "ways" that a logic must take into account to include certain contradictory, but still *logically* coherent, states of affairs. (Compare the quote from Dunn in Section 3.3.2.) Once this hurdle is overcome, P3 emerges naturally via some familiar—one might even say "classical"—principles and intuitions.

Second, we specified that our paraconsistent logic should diverge as little as possible from classical logic. Now, it is well known that giving up explosion comes at a cost. Consider the following simple derivation:

1.	A	assumption
2.	$\neg A$	assumption
3	$A \vee B$	from 1 by disjunction introduction
4.	B	from 2,3 by disjunctive syllogism

This shows that giving up explosion requires one to give up either disjunction introduction or disjunctive syllogism (or transitivity of derivability). Of these, P3 gives up *only* disjunctive syllogism. Now, it cannot be denied that we use (natural language versions of) disjunctive syllogism all the time, and treat it as a correct form of reasoning. So doesn't the fact that P3 rejects disjunctive syllogism count as a knockdown argument against it? No. For what *appear* to be straightforward applications of disjunctive syllogism may actually be *enthymemic* applications of an inference that is valid in P, namely $A \vee B, \neg A, ©A \vdash B$ or $\neg A \vee B, A, ©A \vdash B$. If I receive the message $\{A, \neg A \vee B\}$ and have no good reason to think that A is both true and false, it is reasonable (and in some cases, e.g. those in which someone's life is at stake, *obligatory*) for me to bring in the assumption that $©A$ and thereby infer B. The fact that B is not, strictly speaking, *entailed* by the content of the message is no reason to exonerate me from perjury, for messages convey information not just via their content, but also via the background assumptions that are brought to bear in interpreting that content.

To draw an analogy: suppose I am pulled over for driving 80 in a 55 mile per hour zone. The officer says to me: "Didn't you see the sign that says 'Speed limit 55'?" I reply: "Yes, but that sign does not convey the information that *the speed limit on this particular stretch of road is 55 miles per hour*. For it could be interpreted in a numerous other ways, for example *the speed limit on some other stretch of road is at least 55 kilometers per second*, *the speed limit on this particular stretch of road was enacted in 1955*, *there are 55 reasons you should obey the speed limit*, etc." Now, it is undeniably true that the words 'speed limit 55' do not, *in and of themselves*, convey the information that *the speed limit on this particular stretch of road is 55 miles per hour*. Still, the officer would be right to dismiss my protest as ridiculous, as I should be expected to bring my background knowledge to bear in interpreting the sign.

We can give a precise sense to the notion that $\mathsf{P3}$ diverges from classical logic as little as possible. A logic $\mathsf{L1}$ is said to be *maximal* with respect to a logic $\mathsf{L2}$ just in case $\mathsf{L2}$ is an extension of $\mathsf{L1}$, and given any theorem A of $\mathsf{L2}$ that is not a theorem of $\mathsf{L1}$, the result of adding A as an axiom to a sound and complete axiomatization of $\mathsf{L1}$ is such that all theorems of $\mathsf{L2}$ are provable in it. It has been shown that $\mathsf{LFI1}$ is maximal relative to classical logic (see [40]). Thus since P is equivalent to $\mathsf{LFI1}$, the result holds for P as well.

Third, we specified that our paraconsistent logic should not validate any trivial variations on explosion, e.g. $A \wedge \neg A \mathbin{/} B$ or $A \supset (\neg A \supset B)$. $\mathsf{P3}$ satisfies this condition, as shown by Fact 29 (p. 132). A possible objection is that this condition is violated, since P validates a trival variation on explosion, namely "Boolean explosion," $A, \# A \mathbin{/} B$. This is not a trivial variant on explosion, however, for $\#$ is a very different connective from $\neg$. We can acknowledge the possibility of a formula being both true and false while maintaining that a formula cannot be both true and untrue.

I now want to explore the consequences of using $\mathsf{P3}$ and $\mathsf{P4}$ as bases for conflict-tolerant deontic logics. We turn to this task in the next chapter.

Chapter 4

Paraconsistent deontic logic

> A good deal of effort has already been expended in trying to save mainstream deontic theory from the worst of the difficulties it encounters, through its modality and consistency requirements. Most of the repairs attempted amount however to epicycling: trying to tack theory-saving devices onto the initial theory, rather than questioning any fundamental features of it.
>
> —Richard Routley and Val Plumwood [138, p. 657]

> Most systems of deontic logic are built within a classical frame. But there are alternative ways of building them. Within, for example, a paraconsistent deontic logic the notion of doability is seen under a different aspect. And seeing it thus may actually answer to a way in which we sometimes see it in 'real life'.
>
> —G. H. von Wright [156, p. 386]

In this chapter I construct a small family of paraconsistent deontic logics that are "robustly" tolerant of normative conflicts, yet share all (or at least most) of the good features of the D systems presented in Chapter 2. I call these the PD ('paraconsistent deontic') systems. My approach is quite straightforward: I construct the PD systems on the basis of P3 and P4 in just the manner that the D systems are constructed on the basis of classical propositional logic (CPL).

I am not the first to construct a paraconsistent deontic logic; a number of such systems have appeared in the literature. Thus, before delving into the details of the PD systems, I will briefly review some previous efforts in the area of paraconsistent deontic logic, in order to show how the current approach relates to (and, in particular, differs from) these.

4.1 Previous efforts

I believe that the existing paraconsistent deontic logics can be grouped into three basic families: relevant deontic logics, deontic LFIs, and LP-based deontic logics.[1] I will look at each of these in turn (in roughly the chronological order in which they were introduced).

4.1.1 Relevant deontic logics

To my knowledge, the first (published) system of paraconsistent deontic logic appears in Alan Ross Anderson's 1967 paper, "Some Nasty Problems in the Formal Logic of Ethics" [8]. In this paper, Anderson constructs a deontic logic on the basis of the (now well-known) system R of relevant implication, which he, along with Nuel Belnap and others, popularized.[2] The most significant feature of R is that, unlike classical logic, it satisfies the following "relevance" condition (cf. Def. 55): $A \to B$ is provable only if A and B share a non-logical constant (e.g. an atomic formula).[3] (Loosely speaking, this means that a proposition must be *relevant* to, i.e. have something to do with, what it logically implies.) R is thus weaker than classical propositional logic: while it contains many classical theorems such as $A \wedge B \to A$ and $\neg\neg A \to A$, it invalidates alleged fallacies of relevance such as $A \wedge \neg A \to B$ and $A \to (B \to B)$.

Anderson's appoach to deontic logic was based on an idea that he had been advocating for some time, namely that

> [i]t is analytic of the notion of "obligation" that if obligations (in any of the various legal, or moral, or ethical, or game-like senses of the term) are not fulfilled, then something goes haywire. I take this to be a logical

[1] There are (at least) two rather idiosyncratic systems that I have not investigated closely and am not sure how to classify, namely those found in [1] and [13].

[2] *Relevant* logics are also known as *relevance* logics. Anderson's paper, being intended for a general philosophical audience, is rather light on formal details and contains a fair amount of hand-waving. A more rigorous presentation of the system is given by Goble in "The Iteration of Deontic Modalities" [59], which, in Goble's words, is "something of a sequel (despite its earlier date)" [61, p. 331].

[3] Here and throughout this section I use $\to$ as a generic conditional connective.

> point, the substantive matters having to do with law, morals, ethics, or games, being dependent on what it means to "go haywire." [8, p. 359]

Anderson creates a deontic logic by simply adding the special constant V to R, and defining the deontic operators as follows:

- $OA =_{df} \neg A \rightarrow V$
- $FA =_{df} O\neg A$
- $PA =_{df} \neg O \neg A$

(Defining the deontic operators in this way is often characterized as the "Andersonian reduction.") Intuitively, V means "a violation has occurred," or (in Anderson's breezy parlance) "something goes haywire." The axiom $\neg(\neg V \rightarrow V)$ is added, to indicate that V is "avoidable."[4] In the resulting deontic logic certain dubious formulas such as $O(A \wedge \neg A) \rightarrow OB$ are not provable.

An intriguing feature of R is that, unlike the D (and PD) systems, it contains the theorem $O(A \rightarrow B) \leftrightarrow (A \rightarrow OB)$, i.e. it treats the two standard options for formalizing conditional obligations as equivalent. This may be seen as an asset or a liability, depending on one's intuitions. (In my view it is a liability; see the discussion in Section 2.6.6 above.) The logic also contains some plausible deontic principles such as $OA \rightarrow PA$ and $O(OA \rightarrow A)$. However, it has some extremely dubious theorems that are not provable in the D (or PD) systems, such as:

- $A \rightarrow OPA$
- $A \rightarrow FFA$
- $OOA \rightarrow A$
- $POA \rightarrow A$
- $(A \rightarrow OB) \rightarrow (\neg B \rightarrow FA)$

[4] Actually, Anderson [8] does not explicitly state how to formalize the principle that V is avoidable, though he does mention and appeal to this principle. The formulation I give here is from McArthur [106, p. 148].

Anderson comments on some of these, admitting that they *appear* to be "bad guys," but attempting to explain away their *prima facie* implausibility with some "fancy foot-work." Anderson's 1967 system is discussed in detail in papers by Goble [59] [63] and McArthur [106], both of whom comment on the problematic theorems above.

A closely related approach, investigated by Mares [103] and also Goble [63], is to define O using the (alethic) necessity operator: $OA =_{df} \Box(\neg A \rightarrow V)$. Thus A is obligatory just in case it is *necessary* that the negation of A relevantly implies V. This system has some advantages over Anderson's 1967 system. In particular, the dubious schemas listed above are not provable in it. However, Anderson explicitly rejected this definition of O (which he, it should be noted, first proposed in [7]) on the ground that norms are at least sometimes *contingent*: for example, the fact that if I do not pay property tax then a violation has occurred is surely not a *necessary* truth. Anderson's point seems to me correct. Thus while this system may be more satisfactory in certain *formal* respects, it seems to be less satisfactory in certain *philosophical* ones.

Speaking of contingent facts, there does not seem to be any *necessary* connection between the so-called "Andersonian reduction" and the relevance logic R (or relevance logic or paraconsistent logic in general). Just as one can perform the "Andersonian reduction" using classical logic (indeed, Anderson did just that in his [7]), one can construct a relevant deontic logic in a more conventional way, without the Andersonian reduction. This approach is investigated by the Routleys [139] [138] and Goble [61].[5]

4.1.2 Deontic LFIs

In their 1986 paper "On Paraconsistent Deontic Logic," da Costa and Carnielli construct a deontic logic on the basis of da Costa's well-known paraconsistent propositional logic C_1.[6] Recall (from Remark 21) that a logic of formal inconsistency, or LFI, is a logic that invalidates explosion but validates some form of gentle explosion, e.g. $A, \neg A, ©A / B$. C_1, like $\mathsf{P}\frac{3}{4}$, is an LFI. It is quite different from $\mathsf{P}\frac{3}{4}$, however. Some of its notable features are:

1. it has a two-valued semantics;

[5]Stelzner [147] also appears to fall into this category, though I have not studied this paper very carefully.

[6]For presentations of C_1, see e.g. [45], [39], [37], [101].

2. its negation connective is not truth-functional;

3. it does not validate $\neg(A \wedge \neg A)$;

4. the consistency of A can be expressed with the formula $\neg(A \wedge \neg A)$.

Note that C_1 differs from our system $\mathsf{P3}$ in all four of these respects, and from $\mathsf{P4}$ in all respects but (3).

Da Costa and Carnielli construct their deontic logic on the basis of C_1 in a standard, predictable way, using Kripke-style world semantics. With respect to deontic formulas, the resulting system is very similar to D, except that it invalidates deontic explosion, deontic disjunctive syllogism, and certain related schemas. In fact, as the authors note, C_1 is but the first of an inifinite hierarchy of paraconsistent "C-systems": C_1, C_2, etc., each of which can serve as the basis for a paraconsistent deontic logic. Thus the authors have, in effect, provided a template for constructing an infinite hierarchy of paraconsistent deontic logics. (I have never been clear on what is supposed to be *interesting* about such hierarchies, but that is another matter.)

Deontic LFIs are further discussed and investigated in works by da Costa, Carnielli, and Puga [47] [133] and Grana [70] [71].

4.1.3 LP-based deontic logics

In his influential book *In Contradiction* [129], first published in 1987, Graham Priest constructs a deontic logic on the basis of his paraconsistent logic Δ, which is essentially LP embellished with a non-truth-functional (but non-relevant) conditional. As in LP, the consistency of a formula is not expressible in Δ. Thus Δ is neither a relevant logic nor an LFI.

Priest uses neither the Andersonian reduction nor the standard possible-worlds semantics for the deontic operators; instead, he introduces a rather idiosyncratic appoach that is somewhat akin to a neighborhood semantics (which we touched on in §2.7.2). At each world w, O has an extension, $\omega^+(w)$ and an anti-extention, $\omega^-(w)$.[7] The former is the set of sentences of which obligatoriness may be truly predicated at w, while the latter is the set of sentences of which it may be falsely predicated at w. The truth and falsity conditions for O are then stated, quite simply, as follows:

[7] Worlds are involved here only because they are used in the semantics of the conditional connective.

- OA is true at w iff $A \in \omega^+(w)$
- OA is false at w iff $A \in \omega^-(w)$

(F and P are defined in the usual ways.) $\omega^+(w)$ and $\omega^-(w)$ must be mutually exhaustive (though they need not be mutually exclusive). Also, ω^+ is closed under entailment (if A entails B and $A \in \omega^+(w)$, then $B \in \omega^+(w)$) and conjunction (if $A \in \omega^+(w)$ and $B \in \omega^+(w)$, then $A \wedge B \in \omega^+(w)$), while ω^- is closed dually in the opposite direction: if A entails B and $B \in \omega^-(w)$, then $A \in \omega^-(w)$, and if $A \wedge B \in \omega^-(w)$ then $A \in \omega^-(w)$ or $B \in \omega^-(w)$. In the resulting system, deontic explosion and its ilk are rejected while certain desirable principles such as $O(A \to B) \to (OA \to OB)$ and $OA \wedge OB \to O(A \wedge B)$ are preserved. $OA \to PA$ is not valid in Priest's system, but only because Priest does not think that it is correct; it could easily be validated by requiring that $\neg A \in \omega^-(w)$ whenever $A \in \omega^+(w)$.

In the second, revised edition of *In Contradiction* (published in 2006), Priest remarks that he "would ... now prefer an appropriate world-semantics for O" [129, p. 282], and proceeds to provide a fairly standard one which results in essentially the same logic.

4.1.4 The current approach in perspective

My approach differs from the relevant deontic logic approach in the obvious way that it is based on a non-relevant logic: as I noted in the previous chapter, P3 and P4 validate certain alleged "fallacies of (ir)relevance" such as $\neg(p \supset p) \supset (q \supset q)$ I have nothing against relevant logics (or relevant deontic logics) *per se*; however, the existing semantics for relevance logics tend to be rather complex and not very intuitive. Moreover, relevance logics are not very amenable to tableau-style proof theories, which I obviously favor. Finally, I do not share—at least, not strongly—the intuition that a proposition must be *relevant* to what it implies or entails. I say these things not to disparage relevance logic, which I consider to be an interesting and worthwhile field of investigation, but simply to give an indication why I do not adopt the relevant approach here.

The systems I develop in the current chapter fall under both of the other categories —indeed, I believe they are the first systems of which this can be said. They are deontic LFIs, in that P3 and P4, the systems on which they are based, are LFIs. And some of them are LP-based deontic logics, in that they are based on P3, which is in turn based on LP. The PD systems are unique in that they are (to my knowledge, at least) the first paraconsistent deontic logics to include both an internal consistency operator and a truth-functional negation. I believe that, as such, they combine the best aspects of the second and third approaches to paraconsistent deontic logic.

Now that the present approach has been placed in perspective, let us move on to the details.

4.2 The PD systems

The language of the PD systems is $L(\neg, \#, \vee, O)$. The non-deontic connectives are defined as in the P systems. The deontic connectives are defined as in D.

4.2.1 Semantics

Definition 58 (PD4 model) A PD4 model is a triple $\langle W, R, v\rangle$ where W is a non-empty set, $R \subseteq W^2$, and $v : At \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ that is extended to $\bar{v} : L_{\mathsf{PD}} \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ via the following clauses:[8]

$$
\begin{array}{lcl}
\bar{v}_w(p) & = & v_w(p) \\
1 \in \bar{v}_w(\neg A) & \Leftrightarrow & 0 \in \bar{v}_w(A) \\
0 \in \bar{v}_w(\neg A) & \Leftrightarrow & 1 \in \bar{v}_w(A) \\
1 \in \bar{v}_w(\# A) & \Leftrightarrow & 1 \notin \bar{v}_w(A) \\
0 \in \bar{v}_w(\# A) & \Leftrightarrow & 1 \in \bar{v}_w(A) \\
1 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ or } 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 0 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(OA) & \Leftrightarrow & \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(OA) & \Leftrightarrow & \exists u(wRu \text{ and } 0 \in \bar{v}_u(A))
\end{array}
$$

Definition 59 (PD3 model) A PD3 model is a PD4 model $\langle W, R, v\rangle$ such that for all $p \in At$ and $w \in W$, either $1 \in v_w(p)$ or $0 \in v_w(p)$.

Notation 6 I will use $\mathsf{PD}\frac{3}{4}$ to generalize over PD4 and PD3.

Lemma 16 (exhaustion lemma) For every PD3 model $\langle W, R, v\rangle$, $w \in W$, and $A \in L_{\mathsf{PD}}$, either $1 \in \bar{v}_w(A)$ or $0 \in \bar{v}_w(A)$.

---

[8] I write $v(p, w)$ as $v_w(p)$, etc.

Proof. A simple induction on the length of A. ■

The reader can check that the following hold:

$$\begin{array}{lcl}
1 \in \bar{v}_w(A \wedge B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ and } 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \wedge B) & \Leftrightarrow & 0 \in \bar{v}_w(A) \text{ or } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \Rightarrow 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(FA) & \Leftrightarrow & \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(FA) & \Leftrightarrow & \exists u(wRu \text{ and } 1 \in \bar{v}_u(A)) \\
1 \in \bar{v}_w(PA) & \Leftrightarrow & \exists u(wRu \text{ and } 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(PA) & \Leftrightarrow & \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A))
\end{array}$$

Definition 60 ($\mathsf{PD}\frac{3}{4}\mathsf{s}$ model) A $\mathsf{PD}\frac{3}{4}\mathsf{s}$ model is a $\mathsf{PD}\frac{3}{4}$ model $\langle W, R, v\rangle$ such that R is *serial*, i.e. $\forall w \exists u\ wRu$.

Definition 61 ($\mathsf{PD}\frac{3}{4}\mathsf{h}$ model) A $\mathsf{PD}\frac{3}{4}\mathsf{h}$ model is a $\mathsf{PD}\frac{3}{4}$ model $\langle W, R, v\rangle$ such that R is *shift-reflexive*, i.e. $\forall w \forall u(wRu \Rightarrow wRu)$.

Definition 62 ($\mathsf{PD}\frac{3}{4}\mathsf{sh}$ model) A $\mathsf{PD}\frac{3}{4}\mathsf{sh}$ model is a $\mathsf{PD}\frac{3}{4}$ model $\langle W, R, v\rangle$ such that R is both serial and shift-reflexive.

Notation 7 I will use '$\mathsf{PD}\frac{3}{4}+$' as a variable ranging over all eight PD systems, namely PD4, PD3, PD4s, PD3s, PD4h, PD3h, PD4sh, PD3sh.

Definition 63 (semantic consequence for PD+) $\Gamma \models_{\mathsf{PD}\frac{3}{4}+} A$ iff for all $\mathsf{PD}\frac{3}{4}+$ models $\langle W, R, v\rangle$ and $w \in W$, if $1 \in \bar{v}_w(B)$ for all $B \in \Gamma$, then $1 \in \bar{v}_w(A)$.

4.2.2 Proof theory

Definition 64 (initial list) The initial list for Γ / A, where $\Gamma = \{B_0, \ldots, B_n\}$, is

$$\begin{array}{c}
B_0\ 0+ \\
\vdots \\
B_n\ 0+ \\
A\ 0-
\end{array}$$

Remark 22 A node of the form $A\ x+$ indicates that (on the branch we are assuming that) A is true at w_x (the world represented by x), and a node of the form $A\ x-$ indicates that A is not true at w_x. Thus the initial list corresponds to the assumption that each of the elements of Γ is true at w_0 but A is not. If we show that this assumption cannot hold, then we have shown that $\Gamma \vdash A$.

Definition 65 (closure conditions) A branch of a $\mathsf{P4}+$ tableau is *closed* iff nodes of the forms $A\ x+$ and $A\ x-$ occur on it. In $\mathsf{P3}+$ there is an additional closure condition: $A\ x-$ and $\neg A\ x-$.

The tableau rules for $\mathsf{PD}\frac{3}{4}+$ are as follows. (These rules are presented in undiluted form in Appendix B.) I will not provide much explanation of these rules, assuming the reader can figure out what the rules "say" based on his or her familiarity with the tableau systems presented in Sections 2.5.2 and 3.4.2.

Remark 23 As usual, redundant nodes may not be added to a branch, a rule may be applied to a given set of predecessor nodes only once, and "null" applications of rules are now allowed. (Cf. Section 2.5.2.)

Definition 66 (double negation) The $\mathsf{PD}\frac{3}{4}+$ rule for double negation is:

$$\begin{array}{c} [\neg\neg] \\ \neg\neg A\ x\pm \\ \downarrow \\ A\ x\pm \end{array}$$

Definition 67 (disjunction rules) The $\mathsf{PD}\frac{3}{4}+$ rules for disjunction are:

$$\begin{array}{ccc} [\vee+] & & \\ A \vee B\ x+ & & \\ \swarrow \quad \searrow & & \\ A\ x+ \qquad B\ x+ & & \end{array} \qquad \begin{array}{c} [\vee-] \\ A \vee B\ x- \\ \downarrow \\ A\ x- \\ B\ x- \end{array}$$

Definition 68 (negated disjunction rules) The $\mathsf{PD}\frac{3}{4}+$ rules for negated disjunction are:

$$\begin{array}{c} [\neg\vee+] \\ \neg(A \vee B)\ x+ \\ \downarrow \\ \neg A\ x+ \\ \neg B\ x+ \end{array} \qquad \begin{array}{c} [\neg\vee-] \\ \neg(A \vee B)\ x- \\ \swarrow \quad \searrow \\ \neg A\ x- \qquad \neg B\ x- \end{array}$$

Definition 69 (Boolean negation) The rule for Boolean negation is:

$$\begin{array}{c} [\#] \\ \#A\ x\pm \\ \downarrow \\ A\ x\mp \end{array}$$

Definition 70 (negated Boolean negation) The rule for (De Morgan) negated Boolean negation is:

$$\begin{array}{c} [\neg\#] \\ \neg\#A\ x\pm \\ \downarrow \\ A\ x\pm \end{array}$$

Definition 71 (obligation rules) The $\mathsf{PD}\frac{3}{4}+$ rules for obligation are:

$$\begin{array}{c} [O+] \\ OA\ x+ \\ x \triangleright y \\ \downarrow \\ A\ y+ \end{array} \qquad \begin{array}{c} [O-] \\ OA\ x- \\ \downarrow \\ x \triangleright i \\ A\ i- \end{array}$$

Definition 72 (negated obligation rules) The $\mathsf{PD}\frac{3}{4}+$ rules for negated obligation are:

$$\begin{array}{c} [\neg O+] \\ \neg OA\ x+ \\ \downarrow \\ x \triangleright i \\ \neg A\ i+ \end{array} \qquad \begin{array}{c} [\neg O-] \\ \neg OA\ x- \\ x \triangleright y \\ \downarrow \\ \neg A\ y- \end{array}$$

Definition 73 (seriality rules) $\mathsf{PD}\frac{3}{4}\mathsf{s}$ and $\mathsf{PD}\frac{3}{4}\mathsf{sh}$ have the seriality rules:

$$\begin{array}{cc} [s+] & [s-] \\ OA\ x+ & \neg OA\ x- \\ \boxed{x \rhd y} & \boxed{x \rhd y} \\ \downarrow & \downarrow \\ x \rhd i & x \rhd i \\ A\ i+ & \neg A\ i- \end{array}$$

Recall from Definition 15 (p. 63) that a $\boxed{\text{box}}$ indicates that the node in question *does not* occur on the branch. So, e.g., the rule $[s+]$ says that if a node of the form OA $x+$ occurs on a branch, and no node of the form $x \rhd y$ occurs on that branch, then nodes of the forms $x \rhd i$ and $A\ i+$ may be added to the tip of the branch.

Definition 74 (shift-reflexivity rule) $\mathsf{PD}\frac{3}{4}\mathsf{h}$ and $\mathsf{PD}\frac{3}{4}\mathsf{sh}$ have the shift-reflexivity rule:

$$\begin{array}{c} [h] \\ x \rhd y \\ \downarrow \\ y \rhd y \end{array}$$

Remark 24 These rules are interpreted just as one would expect, given familiarity with the tableau theories for $\mathsf{D}+$ and $\mathsf{P}\frac{3}{4}$. The rules for the extensional connectives are just the $\mathsf{P}\frac{3}{4}$ rules relativized to worlds. The rules for O are similar to those in $\mathsf{D}+$, but adjustments are made to accommodate our "double-entry bookkeeping" semantics. In particular, we have four rules for O (corresponding to 'true', 'not true', 'false', and 'not false') instead of just two (for 'true'/'not false' and 'false'/'not true').

Definition 75 (tableau) A $\mathsf{PD}\frac{3}{4}+$ tableau for the inference Γ / A is any tableau that results from 0 or more applications of $\mathsf{PD}\frac{3}{4}+$ tableau rules to the initial list for Γ / A.

Lemma 17 (non-redundancy) No $\mathsf{PD}\frac{3}{4}+$ tableau is redundant.

Proof. A simple induction on the complexity of tableaus. Similar to the proof of Lemma 1 (p. 65). ■

Definition 76 (complete branch/tableau) A branch of a $\mathsf{PD}\frac{3}{4}+$ tableau is *complete* just in case it is closed or each of the following holds:

1. For each node of the form $\neg\neg A\ x+$ on the branch, a node of the form $A\ x+$ is on the branch.

2. For each node of the form $\neg\neg A\ x-$ on the branch, a node of the form $A\ x-$ is on the branch.

3. For each node of the form $A \vee B\ x+$ on the branch, a node of the form $A\ x+$ or $B\ x+$ is on the branch.

4. For each node of the form $A \vee B\ x-$ on the branch, nodes of the forms $A\ x-$ and $B\ x-$ are on the branch.

5. For each node of the form $\neg(A \vee B)\ x+$ on the branch, nodes of the forms $\neg A\ x+$ and $\neg B\ x+$ are on the branch.

6. For each node of the form $\neg(A \vee B)\ x-$ on the branch, a node of the form $\neg A\ x-$ or $\neg B\ x-$ is on the branch.

7. For each node of the form $\#A\ x+$ on the branch, a node of the form $A\ x-$ is on the branch.

8. For each node of the form $\#A\ x-$ on the branch, a node of the form $A\ x+$ is on the branch.

9. For each node of the form $\neg\#A\ x+$ on the branch, a node of the form $A\ x+$ is on the branch.

10. For each node of the form $\neg\#A\ x-$ on the branch, a node of the form $A\ x-$ is on the branch.

11. For each pair of nodes of the forms $OA\ x+$ and $x \triangleright y$ on the branch, a node of the form $A\ y+$ is on the branch.

12. For each node of the form $OA\ x-$ on the branch, nodes of the forms $x \triangleright i$ and $A\ i-$ are on the branch.

13. For each node of the form $\neg OA\ x+$ on the branch, nodes of the forms $x \triangleright i$ and $\neg A\ i+$ are on the branch.

14. For each pair of nodes of the forms $\neg OA\ x-$ and $x \triangleright y$ on the branch, a node of the form $\neg A\ y-$ is on the branch.

15. If $\mathsf{PD}\frac{3}{4}+$ is a serial system, for each node of the form $OA\ x$ on the branch, there are nodes of the forms $x \rhd i$ and $A\ i+$ on the branch.

16. If $\mathsf{PD}\frac{3}{4}+$ is a serial system, for each node of the form $\neg OA\ x-$ on the branch, there are nodes of the forms $x \rhd i$ and $\neg A\ i-$ on the branch.

17. If $\mathsf{PD}\frac{3}{4}+$ is a shift-reflexive system, for each node of the form $x \rhd y$ $(x \neq y)$ on the branch, a node of the form $y \rhd y$ is on the branch.

A tableau is complete iff all of its branches are complete.

Definition 77 (closed/open tableau) A tableau is closed just in case all of its branches are closed; otherwise it is open.

Definition 78 (tableau-theoretic consequence) $\Gamma \Vdash_{\mathsf{PD}\frac{3}{4}+} A$ iff there is a closed $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A.

We have the following shortcut/convenience rules:

$$\begin{array}{c} [\wedge+] \\ A \wedge B\ x+ \\ \downarrow \\ A\ x+ \\ B\ x+ \end{array} \qquad \begin{array}{c} [\wedge-] \\ A \wedge B\ x- \\ \swarrow \quad \searrow \\ A\ x- \qquad B\ x- \end{array} \qquad \begin{array}{c} [\neg\wedge] \\ \neg(A \wedge B)\ x\pm \\ \downarrow \\ \neg A \vee \neg B\ x\pm \end{array} \qquad \begin{array}{c} [\neg\vee] \\ \neg(A \vee B)\ x\pm \\ \downarrow \\ \neg A \wedge \neg B\ x\pm \end{array}$$

$$\begin{array}{c} [\supset+] \\ A \supset B\ x+ \\ \swarrow \quad \searrow \\ A\ x- \qquad B\ x+ \end{array} \qquad \begin{array}{c} [\supset-] \\ A \supset B\ x- \\ \downarrow \\ A\ x+ \\ B\ x- \end{array} \qquad \begin{array}{c} [\neg\supset+] \\ \neg(A \supset B)\ x\pm \\ \downarrow \\ A \wedge \neg B\ x\pm \end{array}$$

$$\begin{array}{c} [P+] \\ PA\ x+ \\ \downarrow \\ x \rhd i \\ A\ i+ \end{array} \qquad \begin{array}{c} [P-] \\ PA\ x- \\ x \rhd y \\ \downarrow \\ A\ y- \end{array} \qquad \begin{array}{c} [\neg O] \\ \neg OA\ x\pm \\ \downarrow \\ P\neg A\ x\pm \end{array} \qquad \begin{array}{c} [\neg P] \\ \neg PA\ x\pm \\ \downarrow \\ O\neg A\ x\pm \end{array}$$

$$
\begin{array}{c}
[\copyright+] \\
\copyright A\ x+ \\
\swarrow \searrow \\
\begin{array}{cc} A\ x+ & A\ x- \\ \neg A\ x- & \neg A\ x+ \end{array}
\end{array}
\qquad
\begin{array}{c}
[\copyright+] \\
\copyright A\ x- \\
\swarrow \searrow \\
\begin{array}{cc} A\ x+ & A\ x- \\ \neg A\ x+ & \neg A\ x- \end{array}
\end{array}
\qquad
\begin{array}{c}
[\neg\copyright] \\
\neg\copyright A\ x\pm \\
\downarrow \\
\copyright A\ x\mp
\end{array}
$$

As usual, these rules are superfluous and not included in the "official" proof theory.

4.2.3 Example proofs

I now provide some examples illustrating our tableau theory in practice.

Example 17 (escapable deontic explosion) Here is a proof that $Fp, Pp \nVdash_{\mathsf{PD3}} Oq$, using primitive and derived rules:

1.	$O\neg p\ 0+$	initial list
2.	$Pp\ 0+$	"
3.	$Oq\ 0-$	"
4.	$0 \rhd 1$	from 2 by $[P+]$
5.	$p\ 1+$	"
6.	$0 \rhd 2$	from 3 by $[O-]$
7.	$q\ 2-$	"
8.	$\neg p\ 1+$	from 1 and 4 by $[O+]$
9.	$\neg p\ 2+$	from 1 and 6 by $[O+]$
	$\uparrow$	(open and complete)

We can read a countermodel off the open branch as follows. (The procedure is made more explicit in Definition 80 below.) Let $W = \{w_0, w_1, w_2\}$, $R = \{\langle w_0, w_1\rangle, \langle w_0, w_2\rangle\}$, $v_{w_1}(p) = \mathbf{b}$, $v_{w_2}(q) = v_{w_2}(p) = \mathbf{f}$. The reader can check that $1 \in \bar{v}_{w_0}(Fp)$ and $1 \in \bar{v}_{w_0}(Pp)$ but $1 \notin \bar{v}_{w_0}(Oq)$. Hence $Fp, Pp \not\models_{\mathsf{PD3}} Oq$.

Example 18 (no inescapable conflicts) Here is a proof that $\Vdash_{\mathsf{PD3s}} \neg(OA \wedge FA)$, using primitive and derived rules:

1.	$\neg(OA \wedge O\neg A)$ 0−	initial list
2.	$\neg OA \vee \neg O\neg A$ 0−	from 1 by $[\neg\wedge]$
3.	$\neg OA$ 0−	from 2 by $[\vee-]$
4.	$\neg O\neg A$ 0−	"
5.	$0 \triangleright 1$	by $[s]$
6.	$\neg A$ 1−	from 3 and 5 by $[\neg O-]$
7.	$\neg\neg A$ 1−	from 4 and 5 by $[\neg O-]$
8.	A 1−	from 7 by $[\neg\neg]$
	$*$	from 6 and 8

Remark 25 Note that the branch doesn't close in PD4s, which allows gaps. Hence $\not\Vdash_{\mathsf{PD4s}} \neg(OA \wedge FA)$.

Example 19 Here is a proof that $\not\Vdash_{\mathsf{PD3s}} \#(Op \wedge Fp)$:

1.	$\#(Op \wedge O\neg p)$ 0−	initial list
2.	$Op \wedge O\neg p$ 0+	from 1 by $[\#+]$
3.	Op 0+	from 2 by $[\wedge+]$
4.	$O\neg p$ 0+	"
5.	$0 \triangleright 1$	from 3 by $[s]$
6.	p 1+	"
7.	$\neg p$ 1+	from 4 and 5 by $[O+]$
	$\uparrow$	

The open branch suggests the following simple countermodel: $W = \{w_0, w_1\}$, $R = \{\langle w_0, w_1\rangle, \langle w_1, w_1\rangle\}$, $v_{w_1}(p) = \mathbf{b}$. In this model $\bar{v}_{w_0}(Op \wedge O\neg p) = \mathbf{b}$. Hence $\bar{v}_{w_0}(\#(Op \wedge O\neg p)) = \mathbf{f}$. Hence $\not\models_{\mathsf{PD3s}} \#(Op \wedge Fp)$.

Example 20 (obey all rules) Here is a proof that $\Vdash_{\mathsf{PD4h}} O(OA \supset A)$, using primitive and derived rules:

1.	$O(OA \supset A)$ 0−	initial list
2.	$0 \triangleright 1$	from 1 by $[O-]$
3.	$OA \supset A$ 1−	from 1 by $[O-]$
4.	OA 1+	from 3 by $[\supset -]$
5.	A 1−	from 3 by $[\supset -]$
6.	$1 \triangleright 1$	from 2 by $[h]$
7.	A 1+	from 4 and 6 by $[O+]$
	$*$	from 5 and 7

Example 21 (covering principle) Here is a proof that $\Vdash_{\mathsf{PD3s}} FA \vee PA$, using primitive and derived rules:

1.	$O\neg A \vee PA\ 0-$	initial list
2.	$O\neg A 0-$	from 1 by $[\vee-]$
3.	$PA\ 0-$	"
4.	$0 \rhd 1$	"
5.	$\neg A\ 1-$	from 2 by $[O-]$
6.	$A\ 1-$	from 3 and 4 by $[P-]$
	$*$	from 5 and 6

Remark 26 Note that the branch doesn't close in PD4s; hence $\nVdash_{\mathsf{PD4s}} FA \vee PA$.

4.2.4 Soundness and completeness

I now prove the soundness and completeness of the proof theory with respect to the semantics.

Definition 79 (faithful) Let b be a branch of a $\mathsf{PD}\frac{3}{4}+$ tableau. A $\mathsf{PD}\frac{3}{4}+$ model $\mathfrak{M} = \langle W, R, v\rangle$ is *faithful* to b just in case there is a function $w : \mathbb{N} \longrightarrow W$ such that:[9]

- if $x \rhd y$ is on b, then $w_x R w_y$
- if $A\ x+$ is on b, then $1 \in \bar{v}_{w_x}(A)$
- if $A\ x-$ is on b, then $1 \notin \bar{v}_{w_x}(A)$.

We say that w *shows* $\mathfrak{M}$ to be faithful to b.

Lemma 18 (faith lemma) If a branch, b, of a $\mathsf{PD}\frac{3}{4}+$ tableau is closed, then no $\mathsf{PD}\frac{3}{4}+$ model is faithful to b.

Proof. For reductio, suppose that the $\mathsf{PD}\frac{3}{4}$ model $\langle W, R, v\rangle$ is faithful to b . There are two cases:

1. $\mathsf{PD}\frac{3}{4}+$ is a four-valued system. Then, since b is closed, nodes of the forms $A\ x+$ and $A\ x-$ occur on it. Thus $1 \in \bar{v}_{w_x}(A)$ and $1 \notin \bar{v}_{w_x}(A)$, which is impossible.

[9] I write $w(x)$ as w_x, etc.

2. $\mathsf{PD}\frac{3}{4}+$ is a three-valued system. Then, since b is closed, nodes of the forms $A\ +$ and $A\ -$, or $A\ -$ and $\neg A\ -$, occur on it. Thus either $1 \in \bar{v}_{w_x}(A)$ and $1 \notin \bar{v}_{w_x}(A)$, which is impossible, or $1 \notin \bar{v}_{w_x}(A)$ and $1 \notin \bar{v}_{w_x}(\neg A)$, which is also (by the Exhaustion Lemma) impossible.

■

Lemma 19 (soundness lemma) If a $\mathsf{PD}\frac{3}{4}+$ model $\mathfrak{M} = \langle W, R, v\rangle$ is faithful to a branch, b, and a $\mathsf{PD}\frac{3}{4}+$ tableau rule is applied to b, then $\mathfrak{M}$ is faithful to at least one of the branches thereby generated.

Proof. The proof is by cases. There are 14 cases to consider (one for each tableau rule). The cases for the truth-functional connectives ($\neg$, $\vee$, and $\#$) are the same as in $\mathsf{P}\frac{3}{4}$, only relativized to worlds in an obvious way. Thus I will only prove the cases for the "deontic" rules.

- **Case 8** ($[O+]$). Suppose $OA\ x+$ and $x \triangleright y$ are on b, and $[O+]$ is applied. Then $b(A\ y+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(OA)$ and $w_x R w_y$. Thus $\forall u(w_x R u \Rightarrow 1 \in \bar{v}_u(A))$. Thus $1 \in v_{w_y}(A)$. Thus $\mathfrak{M}$ is faithful to $b(A\ y+)$.

- **Case 9** ($[O-]$). Suppose $OA\ x-$ is on b, and $[O-]$ is applied. Then $b(x \triangleright i, A\ i-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(OA)$. Thus there is a world, call it \$, such that $w_x R\$$ and $1 \notin \bar{v}_{\$}(A)$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Moreover, $\hat{w}_x R \hat{w}_i$ and $1 \notin \bar{v}_{\hat{w}_i}(A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, A\ i-)$.

- **Case 10** ($[\neg O+]$). Suppose $\neg OA\ x+$ is on b, and $[\neg O+]$ is applied. Then $b(x \triangleright i, \neg A\ i+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(\neg OA)$, whence $0 \in \bar{v}_{w_x}(OA)$. Thus there is a world, call it \$, such that $w_x R\$$ and $0 \in \bar{v}_{\$}(A)$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Moreover, $\hat{w}_x R \hat{w}_i$ and $0 \in \bar{v}_{\hat{w}_i}(A)$. Thus $1 \in \bar{v}_{\hat{w}_i}(\neg A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, \neg A\ i+)$.

- **Case 11** ($[\neg O-]$). Suppose $\neg OA\ x-$ and $x \triangleright y$ are on b, and $[\neg O-]$ is applied. Then $b(\neg A\ y-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(\neg OA)$ and $w_x R w_y$. Thus $0 \notin \bar{v}_{w_x}(OA)$. Thus $0 \notin \bar{v}_u(A)$ for all u such that $w_x R u$. Thus, since $w_x R w_y$, $0 \notin \bar{v}_{w_y}(A)$. Thus $1 \notin \bar{v}_{w_y}(\neg A)$. Thus $\mathfrak{M}$ is faithful to $b(\neg A\ y-)$.

- **Case 12** ($[s+]$). Suppose $OA\ x+$ occurs on b, no node of the form $x \triangleright y$ occurs on b, and $[s+]$ is applied. Then $b(x \triangleright i, A\ i+)$ is generated. Since $[s+]$ is a rule only for our serial systems, R is serial. Thus there is a world, call it \$, such that $w_x R\$$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Thus $1 \in \bar{v}_{\hat{w}_x}(OA)$. Thus $\forall u(\hat{w}_x Ru \Rightarrow 1 \in \bar{v}_u(A))$. Thus, since $\hat{w}_x R\hat{w}_i$, $1 \in v_{\hat{w}_i}(A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, A\ i+)$.

- **Case 13** ($[s-]$). Suppose $\neg OA\ x-$ occurs on b, no node of the form $x \triangleright y$ occurs on b, and $[s-]$ is applied. Then $b(x \triangleright i, \neg A\ i-)$ is generated. Since $[s-]$ is a rule only for our serial systems, R is serial. Thus there is a world, call it \$, such that $w_x R\$$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Thus $1 \notin \bar{v}_{\hat{w}_x}(\neg OA)$. Thus $0 \notin \bar{v}_{\hat{w}_x}(OA)$. Thus $\forall u(\hat{w}_x Ru \Rightarrow 0 \notin \bar{v}_u(A))$. Thus, since $\hat{w}_x R\hat{w}_i$, $0 \notin v_{\hat{w}_i}(A)$. Thus $1 \notin v_{\hat{w}_i}(\neg A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, \neg A\ i-)$.

- **Case 14** ($[h]$). Suppose $x \triangleright y$ occurs on b, and $[h]$ is applied. Then $b(y \triangleright y)$ is generated. Since $\mathfrak{M}$ is faithful to b, $w_x R w_y$. Since $[h]$ is a rule only for our shift-reflextive systems, R is shift-reflexive. Thus $w_y R w_y$. Thus $\mathfrak{M}$ is faithful to $b(y \triangleright y)$.

■

Theorem 12 (soundness) If $\Gamma \Vdash_{\mathsf{PD}\frac{3}{4}+} A$, then $\Gamma \models_{\mathsf{PD}\frac{3}{4}+} A$.

Proof. Suppose $\Gamma \not\models_{\mathsf{PD}\frac{3}{4}+} A$. Then there is a $\mathsf{PD}\frac{3}{4}+$ model $\mathfrak{M} = \langle W, R, v\rangle$ and $w \in W$ such that $1 \in \bar{v}_w(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}_w(A)$. $\mathfrak{M}$ is faithful to the initial list for Γ / A. Moreover, by the Soundness Lemma (Lemma 19), each subsequent application of a $\mathsf{PD}\frac{3}{4}+$ tableau rule will yield at least one branch to which $\mathfrak{M}$ is faithful. Thus every $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A has at least one branch to which $\mathfrak{M}$ is faithful. Let T be a $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A, and let b be one of the branches of T to which $\mathfrak{M}$ is faithful. By the Faith Lemma (Lemma 18), b cannot be closed. Thus T is open. But T was an arbitrarily chosen $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. Thus there is no closed $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. Thus $\Gamma \nVdash_{\mathsf{PD}\frac{3}{4}+} A$. ■

Definition 80 (induced model for non-serial systems) Let $\mathsf{PD}\frac{3}{4}+$ be a non-serial PD system. Let b be an open, complete branch of a $\mathsf{PD}\frac{3}{4}+$ tableau. The

$\mathsf{PD}\frac{3}{4}+$ model *induced* by b is the model $\langle W, R, v\rangle$ such that $W = \{w_i : i$ is a natural number occurring on $b\}$, $R = \{\langle w_i, w_j\rangle : i \triangleright j$ occurs on $b\}$, and v is determined by the following chart:

<table>
<tr><th>$p\ x+$</th><th>$p\ x-$</th><th>$\neg p\ x+$</th><th>$\neg p\ x-$</th><th>PD4+</th><th>PD3+</th></tr>
<tr><td>Y</td><td>Y</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>Y</td><td>N</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>N</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>N</td><td>N</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>N</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>N</td><td>Y</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{b}$</td></tr>
<tr><td>Y</td><td>N</td><td>N</td><td>Y</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
<tr><td>Y</td><td>N</td><td>N</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
<tr><td>N</td><td>Y</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>N</td><td>Y</td><td>Y</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>Y</td><td>N</td><td>Y</td><td>$v_{w_x}(p) = \mathbf{n}$</td><td>$*$</td></tr>
<tr><td>N</td><td>Y</td><td>N</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>N</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>N</td><td>N</td><td>Y</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>N</td><td>N</td><td>Y</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
<tr><td>N</td><td>N</td><td>N</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
</table>

We need to show that $\langle W, R, v\rangle$, so defined, really is a $\mathsf{PD}\frac{3}{4}+$ model. First, we need to verify that if $\mathsf{PD}\frac{3}{4}+$ is a *shift-reflexive* system, then R is indeed shift-reflexive. Since b is complete, for each node of the form $x \triangleright y$ occurring on b, there is a node of the form $y \triangleright y$ on b. Thus for each w_x and $w_y \in W$ such that $w_x R w_y$, $w_y R w_y$.; that is, R is shift-reflexive. Next, we need to verify that v is a well-defined function mapping world/atomic formula pairs into $\{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ for the PD4 systems and $\{\mathbf{t}, \mathbf{f}, \mathbf{b}\}$ for the PD3 systems. It is easy to see that the chart guarantees this.

Definition 81 (induced model for serial systems) Let $\mathsf{PD}\frac{3}{4}+$ be a serial PD system. Let b be an open, complete branch of a $\mathsf{PD}\frac{3}{4}+$ tableau. The $\mathsf{PD}\frac{3}{4}+$ model *induced* by b is the model $\langle W, R, v\rangle$ such that (i) $W = \{w_i : i$ is a natural number on $b\}$; (ii) $R = \{\langle w_i, w_j\rangle : i \triangleright j$ is on $b\} \cup \{\langle w_i, w_i\rangle : i$ is on b but no node of the form $i \triangleright j$ is on $b\}$; and (iii) v is determined by the same chart used in the previous definition.

We need to show that $\langle W, R, v\rangle$, so defined, really is a $\mathsf{PD}\frac{3}{4}+$ model. First, we need to verify that R is indeed serial. That is, we need to show that for all $w \in W$, there is a $u \in W$ such that wRu. By clause (i) above, $W = \{w_i : i$ is a natural number on $b\}$. Thus we need to show that for each natural number i occurring on b, there is a $u \in W$ such that w_iRu. Now, for each i occurring on b, there are just two possible cases:

- *Case 1.* A node of the form $i \triangleright j$ occurs on b. Then, by clause (ii), w_iRw_j.
- *Case 2.* No node of the form $i \triangleright j$ occurs on b. Then, by clause (ii), w_iRw_i.

In both cases, there is a $u \in W$ such that w_iRu. Thus for all $w \in W$, there is a $u \in W$ such that wRu; i.e., R is serial. Second, we need to verify that if $\mathsf{PD}\frac{3}{4}+$ is a shift-reflexive system, then R is shift-reflexive. That is, we need to show that for all natural numbers i, j occurring on b, if w_iRw_j, then w_jRw_j. If $i = j$, the result follows immediately. So suppose $i \neq j$. Then $i \triangleright j$ must occur on the branch. (For suppose it doesn't. Then, by clause (ii), $i = j$, contadicting our assumption.) Thus, since b is complete, $j \triangleright j$ occurs on the branch. Thus, by clause (ii), w_jRw_j, as required. Finally, we need to verify that v is a well-defined function mapping world/atomic formula pairs into $\{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ for the $\mathsf{PD4}$ systems and $\{\mathbf{t}, \mathbf{f}, \mathbf{b}\}$ for the $\mathsf{PD3}$ systems. It is easy to see that the chart guarantees this.

Lemma 20 (completeness lemma for non-serial systems) Let $\mathsf{PD}\frac{3}{4}+$ be a non-serial $\mathsf{PD}\frac{3}{4}$ system. Let b be an open, complete branch of a $\mathsf{PD}\frac{3}{4}+$ tableau. Let $\langle W, R, v\rangle$ be the $\mathsf{PD}\frac{3}{4}+$ model induced by b. Then for all A :

1. if $A\ x+$ is on b, $1 \in \bar{v}_{w_x}(A)$
2. if $A\ x-$ is on b, $1 \notin \bar{v}_{w_x}(A)$
3. if $\neg A\ x+$ is on b, $0 \in \bar{v}_{w_x}(A)$
4. if $\neg A\ x-$ is on b, $0 \notin \bar{v}_{w_x}(A)$

Proof. An induction on the length of A. If A is atomic, then the result is (almost) immediate. The inductive steps for the truth-functional connectives $\neg$, $\vee$, and $\#$ are the same as in the proof of the completeness lemma for $\mathsf{P}\frac{3}{4}$ (Lemma 15), only relativized to worlds in an obvious way. Thus I will prove only the inductive step for $A = OB$:

1. Suppose $OB\ x+$ is on b. Then, since b is complete, $B\ y+$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the inductive hypothesis (IH), $1 \in \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of the model induced by b, $\bar{v}_{w_y}(B)$ for all w_y such that $w_x R w_y$. Thus $1 \in \bar{v}_{w_x}(OB)$.

2. Suppose $OB\ x-$ is on b. Then, since b is complete, $x \triangleright i$ and $B\ i-$ are on b. By the definition of the model induced by b, $w_x R w_i$. By the IH, $1 \notin \bar{v}_{w_i}(B)$. Thus $1 \notin \bar{v}_{w_x}(OB)$.

3. Suppose $\neg OB\ x+$ is on b. Then, since b is complete, $x \triangleright i$ and $\neg B\ i+$ are on b. By the definition of the model induced by b, $w_x R w_i$. By the IH, $0 \in \bar{v}_{w_i}(B)$. Thus $0 \in \bar{v}_{w_x}(OB)$.

4. Suppose $\neg OB\ x-$ is on b. Then, since b is complete, $\neg B\ y-$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $0 \notin \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of an induced model, $0 \notin \bar{v}_{w_y}(B)$ for all w_y such that $w_x R w_y$. Thus $0 \notin \bar{v}_{w_x}(OB)$.

■

Lemma 21 (completeness lemma for serial systems) Let $\mathsf{PD}\frac{3}{4}+$ be a serial $\mathsf{PD}\frac{3}{4}$ system. Let b be an open, complete branch of a $\mathsf{PD}\frac{3}{4}+$ tableau. Let $\langle W, R, v\rangle$ be the $\mathsf{PD}\frac{3}{4}+$ model induced by b. Then for all A :

1. if $A\ x+$ is on b, $1 \in \bar{v}_{w_x}(A)$
2. if $A\ x-$ is on b, $1 \notin \bar{v}_{w_x}(A)$
3. if $\neg A\ x+$ is on b, $0 \in \bar{v}_{w_x}(A)$
4. if $\neg A\ x-$ is on b, $0 \notin \bar{v}_{w_x}(A)$

Proof. An induction on the length of A. If A is atomic, then the result is (almost) immediate. The inductive steps for the truth-functional connectives $\neg$, $\vee$, and $\#$ are the same as in the proof of the completeness lemma for $\mathsf{P}\frac{3}{4}$ (Lemma 15), only relativized to worlds in an obvious way. Thus I will prove only the inductive step for $A = OB$:

1. Suppose $OB\ x+$ is on b. Then, since b is complete, $B\ y+$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the inductive hypothesis (IH), $1 \in \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Moreover, since b is complete, $[s+]$ has been applied to it wherever possible; so, since by assumption $OB\ x+$ is on b, it cannot be the case that no node of the form $x \triangleright y$ is on b. Thus, by Def. 81, $\{w_i : w_x R w_i\} = \{w_i : x \triangleright i \text{ is on } b\}$. Thus, by Def. 81, $1 \in \bar{v}_u(B)$ for all u such that $w_x R u$. Thus $1 \in \bar{v}_{w_x}(OB)$.

2. Suppose $OB\ x-$ is on b. Then, since b is complete, $x \triangleright i$ and $B\ i-$ are on b. By the definition of the model induced by b, $w_x R w_i$. By the IH, $1 \notin \bar{v}_{w_i}(B)$. Thus $1 \notin \bar{v}_{w_x}(OB)$.

3. Suppose $\neg OB\ x+$ is on b. Then, since b is complete, $x \triangleright i$ and $\neg B\ i+$ are on b. By the definition of the model induced by b, $w_x R w_i$. By the IH, $0 \in \bar{v}_{w_i}(B)$. Thus $0 \in \bar{v}_{w_x}(OB)$.

4. Suppose $\neg OB\ x-$ is on b. Then, since b is complete, $\neg B\ y-$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the inductive hypothesis (IH), $0 \notin \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Moreover, since b is complete, $[s-]$ has been applied to it wherever possible; so, since by assumption $\neg OB\ x-$ is on b, it cannot be the case that no node of the form $x \triangleright y$ is on b. Thus, by Def. 81, $\{w_i : w_x R w_i\} = \{w_i : x \triangleright i \text{ is on } b\}$. Thus, by Def. 81, $0 \notin \bar{v}_u(B)$ for all u such that $w_x R u$. Thus $0 \notin \bar{v}_{w_x}(OB)$.

■

Fact 31 Each $\mathsf{PD}\frac{3}{4}+$ tableau is finitely generated (see Def. 25, p. 75).

Proof. Obvious, since each application of a rule yields at most two immediate successors to any given node. ■

Lemma 22 If a node of the form $B\ x\pm$ occurs on a $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A, then B is a pseudo-subformula (see Def. 29, p. 76) of one of the elements of $\Gamma \cup \{A\}$.

Proof. A straightforward induction on the complexity of tableaus. Similar to the proof of Lemma 7, p. 76. ■

Fact 32 Each formula of $L_{\mathsf{PD}\frac{3}{4}+}$ has only finitely many pseudo-subformulas.

Proof. Obvious. ■

Remark 27 Note that $[O-]$, $[\neg O+]$, $[s+]$, and $[s-]$ are the only index-generating rules in the $\mathsf{PD}\frac{3}{4}$-systems.

Lemma 23 Let b be any non-redundant branch of a $\mathsf{PD}\frac{3}{4}+$ tableau. For all x, y, the node $x \triangleright y$ occurs on b only if nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b.

Proof. If $x \triangleright y$ occurs on b, then it can only have been introduced via an application of $[O-]$, $[\neg O+]$, $[s+]$, $[s-]$, or $[h]$. Let us consider each case in turn:

- Case 1 ($[O-]$). Then nodes of the forms $OA\ x-$ and $A\ y-$ must occur on b. Thus nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b.
- Case 2 ($[\neg O+]$). Then nodes of the forms $\neg OA\ x+$ and $\neg A\ y+$ must occur on b. Thus nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b.
- Case 3 ($[s+]$). Then nodes of the forms $OA\ x+$ and $A\ y+$ must occur on b. Thus nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b.
- Case 4 ($[s-]$). Then nodes of the forms $\neg OA\ x-$ and $\neg A\ y-$ must occur on b. Thus nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b.
- Case 5 ($[h]$). Then $x = y$ and a node of the form $z \triangleright x$, where $z \neq x$, occurs on b. Since $z \neq x$, $z \triangleright x$ must have been introduced via $[O-]$, $[\neg O+]$, $[s+]$, or $[s-]$. In any case, a node of the form $A\ y$ must occur on b. Hence, since $x = y$, a node of the form $C\ x\pm$ occurs on b and a node of the form $D\ y\pm$ occurs on b.

In each of the five possible cases, nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b. ■

Lemma 24 (finiteness lemma) Every $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A is finite.

Proof. Suppose, for reductio, that some $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A is infinite. By König's Lemma, this tableau must have an infinite branch—say, b. There must be, on b, infinitely many nodes. By Lemma 17 (p. 146), all of these nodes must be different. If an index x occurs on b, then it can only occur in a node of the form $B\ x+$, $B\ x-$, $x \triangleright y$, or $y \triangleright x$. By Lemma 22, if $B\ x+$ or $B\ x-$ is on b, then B must be a pseudo-subformula of some element of $\Gamma \cup \{A\}$. By Fact 32, there are only finitely many of these. So there are only finitely many nodes of the form $B\ x\pm$ on b. Moreover, by Lemma 23, $x \triangleright y$ occurs on b only if nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b. We have already established that there are only finitely many nodes of the form $C\ x\pm$ or $D\ y\pm$ on b. Thus, since b is non-redundant, there can only be finitely many nodes of the form $x \triangleright y$ on b. By parity of reasoning, there can only be finitely many nodes of the form $y \triangleright x$ on b. Thus each index can occur on b only finitely many times. Thus, since b is infinite, infinitely many indices must occur on b. There are, then, two cases:

- **Case 1.** For some number, d, there are infinitely many indices on b of depth d. Let n be the smallest such number. The only index of depth 0 occurring on b is 0, and this occurs only finitely often. Hence $n > 0$. Further, the only way an index of depth $n > 0$ can be introduced to b is by an index-generating rule being applied to a node whose formula has an index of depth $n - 1$. And an index-generating rule can be applied to such a node only once. So if there are infinitely many indices of depth n on b, there must be infinitely many indices of depth $n - 1$ on b, contradicting our initial assumption about n.

- **Case 2.** For each number, d, there are only finitely many indices on b of depth d. Then, since infinitely many indices occur on b, infinitely many index depths must be represented on b. Now, suppose the index y occurs on b, where $y > 0$. This index must occur in a node of the form $C\ y\pm$, $y \triangleright z$, or $z \triangleright y$. But by Lemma 23, $y \triangleright z$, or $z \triangleright y$ occurs on b only if a node of the form $C\ y\pm$ occurs on b. So a node of the form $C\ y\pm$ must occur on b. Now, this node must have been added by applying either $[O-]$, $[\neg O+]$, $[s-]$, or $[s-]$ to a node of the form $D\ x\pm$, where D is of higher O-degree than C, or else it must follow by the non-deontic rules (i.e. $[\neg\neg]$, $[\vee+]$, $[\vee-]$, $[\neg \vee +]$, $[\neg \vee -]$, $[\supset +]$, $[\supset -]$, $[\neg \supset +]$, $[\neg \supset -]$,$[\#]$, $[\neg\#]$) from such a formula. And non-deontic rules do not increase O-degree. Thus the maximum O-degree of all formulas on b with the index y is lower than the maximum O-degree of all formulas on b with index x. (Intuitively, as the indices get "deeper," the O-degrees of their associated formulas get smaller.) Thus since the maximum O-degree of formulas with index 0 is finite, there is an index n such that no index of depth greater than n can have been introduced on b, contradicting our assumption that infinitely many index depths are represented on b.

Since both cases lead to a contradiction, we have shown that every $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A is finite. ■

Lemma 25 (extension lemma) Suppose there is no closed $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. Then any $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A has an extension (also a $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A) containing an open, complete branch.

Proof. Suppose there is no closed $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. Now take any $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. We can use a procedure similar to the one given in the proof of Lemma 10 (p. 80) to construct an extension of this tableau (which extension is also a $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A) that contains an open, complete branch. Details of this procedure are straightforward but tedious, and are thus omitted. ■

Theorem 13 (completeness) If $\Gamma \models_{\mathsf{PD}\frac{3}{4}+} A$, then $\Gamma \Vdash_{\mathsf{PD}\frac{3}{4}+} A$.

Proof. Suppose $\Gamma \nVdash_{\mathsf{PD}\frac{3}{4}+} A$. Then, by Definition 78 (p. 148), there is no closed $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. Consider any $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. By Lemma 25, this tableau has an extension (also a $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A) containing at least one open, complete branch. Let $\langle W, R, v\rangle$ be the $\mathsf{PD}\frac{3}{4}+$ model induced by this branch. By the relevant Completeness Lemma (either Lemma 20 or Lemma 21, depending on whether $\mathsf{PD}\frac{3}{4}+$ is serial), $1 \in \bar{v}_{w_0}(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}_{w_0}(A)$. Thus $\Gamma \not\models_{\mathsf{PD}\frac{3}{4}+} A$. ■

Theorem 14 ("all or nothing") If there is an open, complete $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A then there is no closed $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A.

Proof. Suppose there is an open, complete $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. Pick an open branch, b, of this tableau, and let $\langle W, R, v\rangle$ be the $\mathsf{PD}\frac{3}{4}+$ model induced by b. By the Completeness Lemma, $1 \in \bar{v}_{w_0}(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}_{w_0}(A)$. Thus $\Gamma \not\models_{\mathsf{D}+} A$. Thus, by the Soundness Theorem, $\Gamma \nVdash_{\mathsf{D}+} A$. Thus there is no closed $\mathsf{PD}\frac{3}{4}+$ tableau for Γ / A. ■

Theorem 15 (decidability) The $\mathsf{PD}\frac{3}{4}+$ systems are *decidable*; that is, there is an effective procedure which, when applied to any inference Γ / A, determines, in a finite number of steps, whether $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} A$.

Proof. One such decision procedure is as follows. Start with the initial list for Γ / A. Begin applying $\mathsf{PD}\frac{3}{4}+$ tableau rules, in any order. By Lemma 24 the tableau will terminate after a finite number of steps. If it is closed, then $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} A$. If it is open, then by the "all-or-nothing" theorem, $\Gamma \nvdash_{\mathsf{PD}\frac{3}{4}+} A$. ■

4.3 Notable features of the PD systems

In this section I highlight some notable features of the PD systems. Recall that I use '$\mathsf{PD}\frac{3}{4}+$' as a variable ranging over these systems.

4.3.1 Relations between systems

Fact 33 The $\mathsf{PD}\frac{3}{4}$ systems are extensions of $\mathsf{P}\frac{3}{4}$. That is, if $\Gamma \vdash_{\mathsf{P}\frac{3}{4}} A$ then $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} A$.

Fact 34 The following subsumption relations hold between the $\mathsf{PD}\frac{3}{4}$ systems:

$$\begin{array}{ccc} \mathsf{PD}\frac{3}{4} & \longrightarrow & \mathsf{PD}\frac{3}{4}\mathsf{s} \\ \downarrow & & \downarrow \\ \mathsf{PD}\frac{3}{4}\mathsf{h} & \longrightarrow & \mathsf{PD}\frac{3}{4}\mathsf{sh} \end{array}$$

Lemma 26 Every D (Ds, Dh, Dsh) model is a PD4 (PD4s, PD4h, PD4sh) model (for the language $L_{\mathsf{D}} \cap L_{\mathsf{PD}}$).

Proof. Straightforward. Just observe that the semantic clauses of a PD4 (PD4s, PD4h, PD4sh) model agree with the clauses of a D (Ds, Dh, Dsh) model with respect to the classical truth values (**t** and **f**). ■

Lemma 27 Every D (Ds, Dh, Dsh) model is a PD3 (PD3s, PD3h, PD3sh) model (for the language $L_{\mathsf{D}} \cap L_{\mathsf{PD}}$).

Proof. Follows directly from the previous lemma plus the fact that every PD3 (PD3s, PD3h, PD3sh) model is a PD4 (PD4s, PD4h, PD4sh) model. ■

Fact 35 Each of the following holds (where $\Gamma \cup \{A\} \subseteq L_{\mathsf{PD}} \cap L_{\mathsf{D}}$):

- $\Gamma \vdash_{\mathsf{PD3}} A \Rightarrow \Gamma \vdash_{\mathsf{D}} A$
- $\Gamma \vdash_{\mathsf{PD3s}} A \Rightarrow \Gamma \vdash_{\mathsf{Ds}} A$
- $\Gamma \vdash_{\mathsf{PD3h}} A \Rightarrow \Gamma \vdash_{\mathsf{Dh}} A$
- $\Gamma \vdash_{\mathsf{PD3sh}} A \Rightarrow \Gamma \vdash_{\mathsf{Dsh}} A$

Proof. Follows easily from the previous lemma. ■

4.3.2 Basic logical properties

Fact 36 The PD systems enjoy the basic "Tarskian" properties:

1. $A \in \Gamma \Rightarrow \Gamma \vdash_{\mathsf{D}+} A$ [reflexivity]
2. $\Gamma \vdash_{\mathsf{D}+} A \Rightarrow \Gamma, \Delta \vdash_{\mathsf{D}+} A$ [monotonicity]
3. $(\Gamma \vdash_{\mathsf{D}+} A$ and $A \vdash_{\mathsf{D}+} B) \Rightarrow \Gamma \vdash_{\mathsf{D}+} B$ [transitivity]

Theorem 16 (deduction theorem for $\supset$) $\Gamma, A \vdash_{\mathsf{PD}\frac{3}{4}+} B$ iff $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} A \supset B$.

Proof. Similar to the proof of Theorem 9. ■

Fact 37 The deduction theorem fails for $\rightarrow$ in $\mathsf{PD}\frac{3}{4}+$.

Proof. Just observe that $p, q \vdash_{\mathsf{PD}\frac{3}{4}+} p$ but $p \nvdash_{\mathsf{PD}\frac{3}{4}+} q \rightarrow p$. (Counterexample: $v_w(p) = \mathbf{b}$, $v_w(q) = \mathbf{t}$.) ■

Fact 38 In $\mathsf{PD}\frac{3}{4}+$ we have:

1. $A \dashv\vdash B$ iff $\vdash A \equiv B$
2. if $\vdash A \leftrightarrow B$ then $\vdash A \equiv B$

Fact 39 In $\mathsf{PD}\frac{3}{4}+$ it is not the case that if $A \dashv\vdash B$ then $\vdash A \leftrightarrow B$.

Proof. Observe that $p \vee \neg p \dashv\vdash q \vee \neg q$ but $\nvdash p \vee \neg p \leftrightarrow q \vee \neg q$. ■

Fact 40 (failure of replacement) In $\mathsf{PD}\frac{3}{4}+$ it is not the case that if $B \dashv\vdash B'$ then $A \dashv\vdash A[B/B']$.

Proof. Observe that $p \vee \neg p \dashv\vdash q \vee \neg q$ but $\neg(p \vee \neg p) \nvdash \neg(q \vee \neg q)$. ■

Theorem 17 (alt. replacement) If $\vdash_{\mathsf{PD}\frac{3}{4}+} B \leftrightarrow B'$ then $\vdash_{\mathsf{PD}\frac{3}{4}+} A \leftrightarrow A[B/B']$.

Proof. An extension of the proof of the alternative replacement theorem for P. We just need to show that the result holds when $A = OC$. By the inductive hypothesis, $\vdash_{\mathsf{PD}\frac{3}{4}+} C \leftrightarrow C[B/B']$. Thus by Fact 46 (below), $\vdash_{\mathsf{PD}\frac{3}{4}+} OC \leftrightarrow O(C)[B/B']$. But $O(C)[B/B'] = OC[B/B']$. Thus $\vdash_{\mathsf{PD}\frac{3}{4}+} OC \leftrightarrow OC[B/B']$. ■

Fact 41 $\mathsf{PD}\frac{3}{4}+$ is not relevant with respect to $\supset$ or $\rightarrow$.

Proof. Observe that $\vdash_{\mathsf{PD}\frac{3}{4}+} p \wedge \neg p \rightarrow q \vee \neg q$, and hence $\vdash_{\mathsf{PD}\frac{3}{4}+} p \wedge \neg p \supset q \vee \neg q$. ■

4.3.3 Deontic inheritance

Fact 42 (consequential inheritance) If $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} A$ then $O\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} OA$.[10]

Proof. Suppose $O\Gamma \nvdash_{\mathsf{PD}\frac{3}{4}+} OA$. Then there is a $\mathsf{PD}\frac{3}{4}+$ model $\langle W, R, v\rangle$ and $w \in W$ such that $1 \in \bar{v}_w(B)$ for all $B \in O\Gamma$ but $1 \notin \bar{v}_w(OA)$. Thus there is a $u \in W$ such that wRu and $1 \notin \bar{v}_u(A)$. Thus for all $C \in \Gamma$, $1 \in \bar{v}_u(C)$. Thus there is a $\mathsf{PD}\frac{3}{4}+$ model $\langle W, R, v\rangle$ and $u \in W$ such that $1 \in \bar{v}_u(C)$ for all $C \in \Gamma$ but $1 \notin \bar{v}_u(A)$. Thus $\Gamma \nvdash_{\mathsf{PD}\frac{3}{4}+} A$. ■

Remark 28 Given the last result, one would *think* that the more general principle of *multi-conclusion* consequential inheritance would hold. That is, one would think that if $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} \Delta$ then $O\Gamma \vdash_{\mathsf{PD}\frac{3}{4}+} O\Delta$. Indeed, I struggled for a while to prove this conjecture, but in vain. Finally I realized that couldn't prove it because it *isn't true*! Counterexample: $A, \#A \vdash_{\mathsf{PD}} \varnothing$ but $OA, O\#A \nvdash_{\mathsf{PD}} O\varnothing$. This is because $\{A, \#A\}$ is unsatisfiable in all eight PD systems, but $\{OA, O\#A\}$ is not. One might think that the result would hold on the condition that $\Delta \neq \varnothing$, or that $\mathsf{PD}\frac{3}{4}+$ be a system with seriality. But these qualified versions fail, too: for example, $\vdash_{\mathsf{PD3s}} A, \neg A$ but $\nvdash_{\mathsf{P3Ds}} OA, O\neg A$.Incidentally, it worth noting that the same sort of thing occurs in the multiple-conclusion versions of standard deontic logics, and for the same sort of reason. For example, $\vdash_{\mathsf{D+}} A, \neg A$ but $\nvdash_{\mathsf{DK+}} OA, O\neg A$. However, it is not too difficult to show that if $\vdash_{\mathsf{X}} \Delta$ then $\vdash_{\mathsf{X}} P\Delta$, for $\mathsf{X} \in \{\mathsf{Ds}, \mathsf{PD}\frac{3}{4}\mathsf{s}\}$.

Fact 43 (implicational inheritance) If $\vdash_{\mathsf{PD}\frac{3}{4}+} A \supset B$ then $\vdash_{\mathsf{PD}\frac{3}{4}+} OA \supset OB$.

Proof. Immediate from the Deduction Theorem (Theorem 16) and the consequential inheritance principle (Fact 42). ■

Fact 44 (strong implicational inheritance) If $\vdash_{\mathsf{PD}\frac{3}{4}+} A \to B$ then $\vdash_{\mathsf{PD}\frac{3}{4}+} OA \to OB$.

Proof. Suppose that $\not\vdash_{\mathsf{PD}\frac{3}{4}+} OA \to OB$. Then there is a $\mathsf{PD}\frac{3}{4}+$ model $\langle W, R, v\rangle$ and $w \in W$ such that $1 \notin \bar{v}_w(OA \to OB)$. There are two cases, then:

1. $1 \in \bar{v}_w(OA)$ and $1 \notin \bar{v}_w(OB)$. Thus there is a u such that wRu and $1 \notin \bar{v}_u(B)$. Thus $1 \in \bar{v}_u(A)$. Thus $1 \notin \bar{v}_u(A \supset B)$. Thus $1 \notin \bar{v}_u(A \to B)$. Thus $\not\vdash_{\mathsf{PD}\frac{3}{4}+} A \to B$.

[10] Recall from Definition 34 that $O\Gamma =_{df} \{OA : A \in \Gamma\}$.

2. $0 \in \bar{v}_w(OB)$ and $0 \notin \bar{v}_w(OA)$. Thus there is a u such that wRu and $0 \in \bar{v}_u(B)$. Thus $0 \notin \bar{v}_u(A)$. Thus $1 \in \bar{v}_u(\neg B)$ and $1 \notin \bar{v}_u(\neg A)$. Thus $1 \notin \bar{v}_u(\neg B \supset \neg A)$. Thus $1 \notin \bar{v}_u(A \to B)$. Thus $\not\vdash_{\mathsf{PD}\frac{3}{4}+} A \to B$.

In either case, $\not\vdash_{\mathsf{PD}\frac{3}{4}+} A \to B$. ■

Remark 29 The next two facts follow easily from the previous two (respectively), plus the easily verified fact that a conjunction is valid in $\mathsf{PD}\frac{3}{4}+$ iff both of its conjuncts are.

Fact 45 (deontic equivalence) If $\vdash_{\mathsf{PD}\frac{3}{4}+} A \equiv B$ then $\vdash_{\mathsf{PD}\frac{3}{4}+} OA \equiv OB$.

Fact 46 (strong deontic equivalence) If $\vdash_{\mathsf{PD}\frac{3}{4}+} A \leftrightarrow B$ then $\vdash_{\mathsf{PD}\frac{3}{4}+} OA \leftrightarrow OB$.

4.3.4 Principles related to conflict-(in)tolerance

Because of the following facts, it might be thought that the PD3 systems are not (fully or robustly) tolerant of normative conflicts:

1. $\vdash_{\mathsf{PD3}} \neg(FA \wedge PA)$

2. $\vdash_{\mathsf{PD3s}} \neg(OA \wedge FA)$ [see Example 18, p. 149]

(1) seems to rule out escapable conflicts, while (2) seems to rule out inescapable conflicts. And indeed, these were the correct interpretations of the respective validities in *standard* deontic logic (i.e., the D systems). In a paraconsistent framework, however, we must think about things differently. We are now working with *De Morgan* negation, as opposed to Boolean negation: the negation of a statement entails the *falsehood* of that statement, but does not entail its *non-truth*. Thus the fact that a statement is always false no longer rules out the possibility that it is also true. So, just as in P3 the fact that $B \wedge \neg B$ is always false does not rule out the possibility that it is also true, in PD3+ the fact that $FA \wedge PA$ is always false does not rule out the possibility that it is also true. (Indeed, $FA \wedge PA$ is an *instance* of $B \wedge \neg B$.) Similarly, in PD3s+ the fact that $OA \wedge FA$ is always false does not rule out the possibility that it is also true. We do, however, have the following:

1. $\nvdash_{\mathsf{PD3+}} \#(FA \wedge PA)$

2. $\nvdash_{\mathsf{PD3+}} \#(OA \wedge FA)$ [see Example 19, p. 150]

Thus it is not a logical truth that $FA \wedge PA$ or $OA \wedge FA$ is *not true*. This seems to justify the claim that the PD3 systems, unlike the D systems, do not rule out the possibility of normative conflicts.

Nevertheless, the PD3 systems (as defined so far) require that we accept escapable normative conflicts as full-fledged dialetheias (true contradictions). Any situation in which $FA \wedge PA$ is true is also a situation in which $\neg(FA \wedge PA)$ is true. Similarly, the PD3 systems with seriality require that we accept inescapable normative conflicts as dialetheias: any situation in which $OA \wedge FA$ is true is also one in which $\neg(OA \wedge FA)$ is true. These are requirements which one may, quite understandably, wish to resist. One may wish to say that, yes, there are situations in which normative conflicts obtain, but these are not cases in which something is both true and false. *That*, after all, is (still) impossible.

As it turns out, in the PD4 systems, which allow for truth value *gaps* (i.e., formulas that are neither true nor false) as well as gluts (formulas that are both true and false), both $\neg(FA \wedge PA)$ and $\neg(OA \wedge FA)$ are invalid (in all systems, regardless of the restrictions on R). Consider, for example, the following (partially specified) model:

- $W = \{w_0, w_1\}$
- $R = \{\langle w_0, w_1\rangle, \langle w_1, w_1\rangle\}$
- $v_{w_1}(p) = \varnothing$

This model can be depicted graphically as follows:

$$\begin{tikzcd} \boxed{w_0} \arrow[r] & \boxed{w_1} \arrow[loop above] \\ & p\ - \\ & \neg p\ - \end{tikzcd}$$

As one can easily check, both $\neg(Fp \wedge Pp)$ and $\neg(Op \wedge Fp)$ are untrue (more specifically, neither true nor false) at w_0 in this model. This was only possible because we allowed p to be neither true nor false at w_1 (something that was previously verboten).

Now, for the first time (in the present work) I believe, we have a real motivation for "going four-valued"—more specifically, for constructing our deontic logic on the basis of a propositional logic that is *paracomplete* as well as paraconsistent. Notably, each of the following is invalid in PD4+:

1. $\neg(FA \wedge PA)$ [no escapable conflicts]
2. $\neg(OA \wedge FA)$ [no inescapable conflicts][11]
3. $\neg O(A \wedge \neg A)$ [no contradictory obligations]
4. $PA \vee P\neg A$ [principle of permission]
5. $FA \vee PA$ [deontic covering principle][12]
6. $O(A \vee \neg A)$ [tautologies are obligatory]
7. $F(A \wedge \neg A)$ [contradictions are forbidden]
8. $P(A \vee \neg A)$ [tautologies are permitted]

We have already given motivation for rejecting 1 and 2. Rejecting 3 seems to go along naturally with rejecting 2. While there is some intuition supporting 4 and 5, many will say "good riddance" to at least 6-8.

In spite of what's just been said, I don't think that it is *essential* to go four-valued in order to achieve conflict-tolerance. After all, one *could* bite the bullet and accept that whenever $FA \wedge PA$ is true, it is also false. The same goes for $OA \wedge FA$.[13] At most, this approach requires one to accept a limited or specialized form of dialetheism, namely, the view that *statements involving norms* can be both true and false. Perhaps a more principled view would be that when convention, stipulation, "say-so" or other such social phenomena are involved, statements can be true and false. The distinction being appealed to here has been discussed by John Searle in his book, *The Construction of Social Reality*:

> Without implying that these are the only kinds of facts that exist in the world, we need to distinguish between *brute facts* such as the fact that the sun is ninety-three million miles from the earth and *institutional* facts such as the fact that Clinton is president. Brute facts exist independently

[11] See Remark 25 (p. 150).

[12] See Remark 26, p. 151.

[13] Interestingly, even Priest (never one to be overly squeamish about accepting true contradictions!) rejects the view that if $OA \wedge O\neg A$ is true, then it is also false. Priest writes that the principle that $O\neg A$ entails $\neg OA$ "serves to multiply contradictions by turning inconsistent obligations into dialetheias. Second, the principle seems to have no rationale. The orthodox rationale is certainly undercut by the existence of inconsistent obligations, and no other appears to take its place. Hence the principle would appear to multiply contradictions beyond necessity." [129, p. 195]

> of any human institutions; institutional facts can exist only within human institutions. [142, p. 27]

So one might hold that *institutional* facts can be both true and false, or (equivalently) that institutiontal facts can be contradictory. Since at least some normative statements are institutional in nature, this strongly suggests (even if it does not strictly entail[14]) that normative statements can be contradictory. Call this view *institutional dialetheism*. Institutional dialetheism is perfectly consistent with the claim that *brute* facts cannot be both true and false, or equivalently, that there cannot be any contradictory brute facts. As such, institutional dialetheism is a *relatively* conservative/benign (and, I think, at least somewhat plausible) form of dialetheism.

At any rate, the main point of this section is that neither escapable nor inescapable conflicts explode (or deontically explode) in any of the paraconsistent deontic logics we have been discussing. Nor do contradictory obligations explode (or deontically explode). That is, in all of our PD systems we have:

1. $FA, PA \nvdash B$
2. $OA, FA \nvdash B$
3. $FA, PA \nvdash OB$ [see Example 17]
4. $OA, FA \nvdash OB$
5. $O(A \wedge \neg A) \nvdash B$
6. $O(A \wedge \neg A) \nvdash OB$

Thus all of these systems can be regarded as robustly conflict-tolerant. The main difference between the PD3 systems and the weaker PD4 systems is that the former require us to regard normative conflicts as full-fledged dialetheias. In that sense, the more conservative advocate of normative conflicts may prefer the PD4 systems, which are consistent with the view that normative conflicts are not necessarily dialethic in nature.

[14] In case it is not obvious, I say this because the following inference is clearly not deductively valid: It is possible for an X to be Z; some Ys are Xs; therefore, it is possible for a Y to be Z.

4.3.5 Miscellaneous validities/invalidities

We have seen that the PD systems are, unlike the D systems, robustly conflict-tolerant. Here are some other noteworthy facts about the systems:

1. $PA \dashv\vdash \neg FA$ [weak permission]
2. $OA, OB \vdash O(A \wedge B)$ [aggregation]
3. $P(A \vee B) \dashv\vdash PA \vee PB$ [principle of deontic distribution]
4. $OA \vdash O(A \vee B)$ [Ross's paradox 1]
5. $PA \vdash P(A \vee B)$ [Ross's paradox 2]
6. $OA \vdash O(B \supset A)$ [Prior's paradox 1]
7. $FA \nvdash O(A \supset B)$ [Prior's paradox 2]
8. $P(A \vee B) \nvdash PA \wedge PB$ [free choice permission]
9. $O(A \supset B), OA \vdash OB$ [deontic detachment]
10. $O(A \supset B), PA \vdash PB$ [permissive deontic detachment]
11. $O(A \supset B), FB \nvdash FA$ [deontic modus tollens]
12. $O(A \to B), FB \vdash FA$ [deontic modus tollens for $\to$]
13. $O(A \supset B), A \nvdash OB$ [deviant deontic detachment 1]
14. $A \supset OB, OA \nvdash OB$ [deviant deontic detachment 2]
15. $A \supset OB \nvdash O(A \supset B)$
16. $OA \vdash_{\mathsf{PD}\frac{3}{4}\mathsf{s}} PA$ [inferential D schema]
17. $\vdash_{\mathsf{PD}\frac{3}{4}\mathsf{h}} O(OA \supset A)$ ["obey all rules"][15]
18. $OOA \vdash_{\mathsf{PD}\frac{3}{4}\mathsf{h}} OA$ [O collapse]
19. $OA \nvdash A$ [Panglossian principle 1]

[15] See Example 20, p. 150.

20. $A \nvdash OA$ [Panglossian principle 2]

21. $O(A \vee B), FA \nvdash OB$ [deontic disjunctive syllogism]

22. $F(A \wedge B), OA \nvdash FB$

23. $FA \vdash F(A \wedge B)$ [penitent's paradox]

The PD systems agree with the D systems on all of the above, except Prior's paradox (version 2), deontic modus tollens, and deontic disjunctive syllogism. All three of these fail for essentially the same reason: A can be both obligatory and forbidden. However, the PD systems do validate variations on each of the principles, namely:

1. $O\#A \vdash O(A \supset B)$

2. $O(A \supset B), O\#B \vdash O\#A$

3. $O(A \vee B), O\#A \vdash OB$

4.3.6 Recapturing standard deontic logic

Remark 30 Recall from Definition 56 that $\Gamma^{©} =_{df} \{©p : p \in At$ and p occurs in an element of $\Gamma\}$.

Definition 82 (well-behaved consequence) $\Gamma \vdash^{©}_{\mathsf{PD}\frac{3}{4}+} A$ iff $\Gamma, (\Gamma \cup \{A\})^{©} \vdash_{\mathsf{PD}\frac{3}{4}+} A$.

Theorem 18 (recapture) Each of the following holds:

- If $\Gamma \vdash_{\mathsf{D}} A$ then $\Gamma \vdash^{©}_{\mathsf{PD4}} A$.
- If $\Gamma \vdash_{\mathsf{Ds}} A$ then $\Gamma \vdash^{©}_{\mathsf{PD4s}} A$.
- If $\Gamma \vdash_{\mathsf{Dh}} A$ then $\Gamma \vdash^{©}_{\mathsf{PD4h}} A$.
- If $\Gamma \vdash_{\mathsf{Dsh}} A$ then $\Gamma \vdash^{©}_{\mathsf{PD4sh}} A$.

Proof. A fairly obvious modification of the proof of Theorem 11. ■

4.4 Objections to the PD systems

I now want to briefly note two possible objections that may be raised against the PD systems specified above.

4.4.1 Too much paraconsistency

One may object to the PD systems on the ground that they are *too* paraconsistent, so to speak. All we wanted was to get rid of the infamous deontic explosion (and its ilk); giving up ordinary explosion (and thus disjunctive syllogism along with it) was more than we bargained for. (We threw out the baby with the bathwater.) What we really want is a deontic logic that is paraconsistent only within the scope of the deontic operators, and classical otherwise—a "semi-paraconsistent" deontic logic, as it were.

4.4.2 Not enough paraconsistency

Another objection to the PD systems is that they are, in a sense, not paraconsistent enough. For there are still deontic sentences that "explode", rendering everything obligatory. In particular, the following principles hold in all the PD systems:

1. $OA, FA, O©A \vdash OB$
2. $OA, O\#A \vdash OB$
3. $FA, PA, O©A \vdash OB$
4. $O\#A, PA \vdash OB$

Why should we allow for the possibility that $OA \wedge FA$ or $FA \wedge PA$ is true (without everything being obligatory), but not allow for the possibility that $OA \wedge O\#A$ or $O\#A \wedge PA$ is true (without everything being obligatory)? It certainly seems possible that, for example, a certain body of law require that a certain proposition not only be false, but that it be *uniquely* false. (That it would probably never occur to lawmakers to add such a clause is a moot point, of course; we are only concerned with what is possible.)

In the next chapter I introduce a brand-new approach to deontic logic that addresses both of the above objections.

Chapter 5

Semi-paraconsistent deontic logic

[U]niverses of worlds as well as worlds themselves may be built in many ways.

—Nelson Goodman [68, p. 5]

'Mid pleasures and palaces though we may roam
Be it ever so humble, there's no place like home

—John Howard Payne [123]

In this chapter I follow through on the suggestion of creating a "semi-paraconsistent" deontic logic—one that is paraconsistent within the scope of the deontic operators but classical otherwise. (Actually, I will specify a small family of such logics.) I will start by bringing out the motivating intuitions for such a system.

5.1 Motivating intuitions

5.1.1 Possible worlds vs. permissible worlds

Consider the usual Kripke-style semantics for (alethic) modal logics (as presented in, e.g. Chellas [43] or Priest [127]). We have a set of objects, W, and a binary relation, R, relating elements of W to elements of W (i.e., $R \subseteq W \times W$). The elements of W are usually referred to as *possible worlds*, and R is referred to as an *accessibility relation*. wRu is interpreted as saying that u is possible from the perspective of w. Chellas, for example, interprets wRu as saying that "the world u is possible relative to—or is relevant to—the world w" [43, p. 68].[1] And Priest writes that "R is a

[1] For notational uniformity I have replaced Chellas's 'α' and 'β' with 'w' and 'u', respectively.

relation of relative possibility, so that uRv means that, relative to u, situation v is possible" [127, p. 21].

Now, in some modal systems, R has a property that we might call *inverse seriality* or *left-unboundedness*:[2] for all w there is a u such that uRw. (The system S5, in which all worlds see each other, is an example of such a system.) In *these* systems, it makes sense to refer to each element of W as a *possible* world, since each element of W is possible from the perspective of, or relative to, some world or other. (Compare the fact that a mother is a person who bears the *mother of* relation to someone.) But in systems in which R does not necessarily have this property (e.g. the basic modal logic K, in which no restrictions whatsoever are placed on R), it really does not make sense to refer to each element of W as *possible*. For there may be, in a given model, worlds that are not possible from the perspective of any world—i.e., worlds w such that for all $u \in W$ it is not the case that uRw. At most, we can refer to each element of w as a *potentially* possible world. So the word 'possible' is more applicable/relevant to the accessibility relation than to the worlds that it relates. We refer to worlds as *possible* worlds not so much because they are possible *per se*, but because of how we interpret R: we take wRu to say that u is *possible* from the perspective of w.

However, as we have already seen (cf. Section 2.5.1), in *deontic* logic wRu is not best interpreted as saying that u is *possible* from the perspective of w. A more appropriate interpretation is that u is *permissible* from the perspective of w. If wRu holds, then no norms that are in effect at w are violated at u: everything that *ought* to be true at w, *is* true at u. Now, when we say that a world u is *possible* from the perspective of a world w, it makes sense to assume that u is consistent—that is, that no formula is such that both it and its negation are true at u. After all, contradictions are impossible; they cannot be true; hence there is no *possible* world in which they are true. But when we say that u is *permissible* from the perspective of w, we need not, and should not, assume that u is consistent. After all, contradictions can be permissible. Suppose, for example, that OA, FA, and PA all hold at a world. Then, by the usual semantics for the deontic operators, there must be some (impossible) world at which A and $\neg A$ hold. (Of course, if we assume that R is serial, we get the same result if just OA and FA are true at a world.) Since we are assuming that something can be obligatory, forbidden, and permitted, we should *not* assume that permissible worlds (i.e., worlds that are permissible from the perspective of some

[2] The term *left-unboundedness* is used by Blackburn et al. [28, p. 207] . These authors use the term *right-unboundedness* for the property I have been calling *seriality*.

world) are consistent.

In short, we should assume that *possible* worlds are consistent, but we should not assume that *permissible* worlds are. It is obvious that not all possible worlds are permissible. (For example, there is a world in which I commit murder. This world is possible, but not permissible.) It should now be equally obvious that not all permissible worlds are possible. (For example, there might be a permissible but impossible world in which I both do and do not shut the door.) Thus in deontic logic it makes sense to let elements of W be inconsistent—with one exception. We do not want to define semantic consequence in terms of truth-preservation at *permissible* worlds; we want to define it in terms of truth-preservation at *possible* worlds. This is because a valid argument should guarantee that it is *impossible* (as opposed to *impermissible*) for the premises to be true while the conclusion is untrue. We could designate some subset of W as "possible" worlds and define semantic consequence as truth-preservation at each element of this set. (This might be appropriate, for example, if we wanted to introduce alethic modal operators, $\Box$ and $\Diamond$, that could interact with the deontic operators.) For present purposes, however, we can simplify things by simply designating just one world as possible, namely the "actual world." We will say that A is a semantic consequence of Γ iff every model is such that if all of the elements of Γ are true at the actual world, then A is true at the actual world. This captures the idea that a logically valid argument is such that it is impossible for the premises to be true without the conclusion being true.

5.1.2 The home world and the "double standard"

Let us call the home world '@', to indicate that it is the *actual* world, the "home world," or the world we are "*at*." (It will sometimes be convenient to refer to @ as w_0.) @ will be the only world that is required to be consistent; other worlds may be inconsistent. Semantic consequence will be defined as truth-preservation at the home world. Since the home world is a possible world, all formulas will take classical truth values at it. This should give us all of classical logic, while allowing for normative conflicts. Let us take a first stab at how the semantics will look.

Remark 31 We will construct our basic semi-paraconsistent deontic logics on the basis of $\mathsf{P4}$. It will be very easy to define corresponding three-valued versions, based on $\mathsf{P3}$. While we don't need to allow for truth-value gaps in order to render $\neg(FA \wedge$

PA) and $\neg(OA \wedge FA)$ invalid (as we did with the PD systems), we will see that there is other motivation for "going four-valued": in particular, it prevents negative permission ($\bar{P}$) from entailing positive permission (P), and negative prohibition ($\bar{F}$) from entailing positive prohibition (F).

Remark 32 We will not introduce the Boolean negation operator ('#') here. It would have little utility, for Boolean negation and De Morgan negation collapse (i.e. are equivalent) at the home world.

Let us define a model as a quadruple $\langle W, @, R, v\rangle$ where W is a set, $@ \in W$, $R \subseteq W^2$, and $v : At \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ such that $v_@(p) \in \{\mathbf{t}, \mathbf{f}\}$ for all $p \in At$. v is extended to $\bar{v} : L(\neg, \vee, \supset, O) \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ via the following (now-familiar) clauses:

$$
\begin{array}{lcl}
\bar{v}_w(p) & = & v_w(p) \\
1 \in \bar{v}_w(\neg A) & \Leftrightarrow & 0 \in \bar{v}_w(A) \\
0 \in \bar{v}_w(\neg A) & \Leftrightarrow & 1 \in \bar{v}_w(A) \\
1 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ or } 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 0 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \Rightarrow 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(OA) & \Leftrightarrow & \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(OA) & \Leftrightarrow & \exists u(wRu \text{ and } 0 \in \bar{v}_u(A))
\end{array}
$$

Semantic consequence is defined as truth-preservation at the *home world*: $\Gamma \models A$ iff for all models $\langle W, @, R, v\rangle$, if $1 \in \bar{v}_@(B)$ for all $B \in \Gamma$, then $1 \in \bar{v}_@(A)$.

We have *stipulated* that the home world is consistent with respect to the *atomic* formulas. What we need to check is that it is consistent with respect to *all* formulas. This is easy to confirm with respect to the basic extensional connectives ($\neg$, $\wedge$, and $\supset$). (Just check the truth tables and note that whenever all inputs are classical, the output is classical.) We run into a problem, however, when we consider the deontic operator O. Consider the following (partially specified) model:

- $W = \{@, w_1\}$
- $R = \{\langle @, w_1\rangle\}$
- $v_{w_1}(p) = \mathbf{b}$

On this model, Op is both true and false at @.

How can we ensure that *deontic* formulas (i.e., those with O as their main connective) always take a classical truth value at the home world? An obvious solution is to reformulate the clauses for O as follows:

$$\begin{array}{lcl} 1 \in v_w(OA) & \Leftrightarrow & \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A)) \\ 0 \in v_w(OA) & \Leftrightarrow & 1 \notin \bar{v}_w(OA) \end{array}$$

In other words, the *truth* condition for OA is as usual, but OA is now *false* just in case it is *not true*. Clearly this ensures that OA will always take a classical truth value at the home world. It has a rather unpleasant side effect, however: it validates the following special case of deontic explosion: $OOA, FOA \diagup OB$.[3] So if it is both obligatory and forbidden that something be obligatory, then everything is obligatory! This is unacceptable. The problem is that we have done more than what was necessary to ensure the consistency of the home world: we have applied the "Boolean" falsity condition (i.e. "false = not true") to *all* worlds, where it is only needed at @.

We can escape the problem of iterated deontic explosion we encountered by applying a kind of "double standard" to the falsity condition for OA. We "split the difference" between the first falsity condition and the second falsity condition by letting the latter apply at the home world while letting the former apply at all other worlds:

$$\begin{array}{lcl} 1 \in \bar{v}_w(OA) & \Leftrightarrow & \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A)) \\ 0 \in \bar{v}_w(OA) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(OA)) \text{ or} \\ & & (w \neq @ \text{ and } \exists u(wRu \text{ and } 0 \in \bar{v}_u(A))) \end{array}$$

This seems to give us the best of both worlds (no pun intended). We ensure that the home world remains consistent (and thus preserve all of classical propositional logic), while ensuring that "iterative deontic explosion" (i.e. $OOA, FOA \diagup OB$) is not valid. And of course, most importantly, we have ensured that ordinary deontic explosion ($OA, FA \diagup OB$) is invalid.

5.1.3 Permission as primitive

We are not out of the woods yet, however. For consider escapable conficts, i.e. those of the form $\{FA, PA\}$. If we define F and P in the usual ways, $FA \wedge PA$ is a flat-contradiction ($O\neg A \wedge \neg O\neg A$). Thus in our models FA and PA can never be jointly

[3] I am grateful to William H. Hanson for bringing this point to my attention.

true at @, and thus $\{FA, PA\}$ entails everything. But we want to allow for escapable conflicts as well as inescapable ones. We have already seen (cf. Section 1.1.3) that there is some *independent* justification for *not* defining compulsory norms in terms of elective norms, or vice versa. So let us continue to define FA as $O\neg A$, but let us take P as primitive. What should its semantic clauses look like? Using O as a model, we have:

$$\begin{array}{lcl} 1 \in \bar{v}_w(PA) & \Leftrightarrow & \exists u(wRu \text{ and } 1 \in \bar{v}_u(A)) \\ 0 \in \bar{v}_w(PA) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(PA)) \text{ or} \\ & & (w \neq @ \text{ and } \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A))) \end{array}$$

Now consider the following (partially specified) model:

- $W = \{@, w_1\}$
- $R = \{\langle @, @\rangle, \langle @, w_1\rangle\}$
- $v_@(p) = v_@(q) = \mathbf{f}$
- $v_{w_1}(p) = \mathbf{b}$

In this model we have Fp and Pp true at @, without Oq being true at @. Thus we have set up a system in which both escapable and inescapable conflicts can hold (at the home world) without "exploding."

In case we still have a use for "negative" or "weak" permission (cf. Section 1.1.4), we can define it as follows: $\bar{P}A =_{df} \neg FA$. Similarly we can define negative or weak prohibition as $\bar{F}A =_{df} \neg PA$. Finally, of course, we could define negative or weak obligation as $\bar{O}A =_{df} \neg P\neg A$.

It now appears that we are ready to specify our systems more formally.

5.2 The SPD systems

Our basic semi-paraconsistent deontic logic is called SPD4 ('semi-paraconsistent deontic logic with 4 truth values'). The slightly stronger system SPD3 is obtained from SPD4 by adding an exhaustion restriction (at each world, every atomic formula is either true or false). We will also define three extensions of SPD4 and three extensions of SPD3 by placing restrictions on R (seriality, shift-reflexivity, or both). The eight systems will be referred to collectively as "the SPD systems", and, following the already-established notational convention, I will use '$\mathsf{SPD}\frac{3}{4}+$' to generalize over them. The language of the SPD systems is $L(\neg, \vee, \supset, O, P)$. The connectives $\wedge$, $\equiv$, $\rightarrow$, $\leftrightarrow$, and the deontic operator F are defined as in the PD systems. Negative permission ($\bar{P}$) and negative prohibition ($\bar{F}$) are defined as follows:

- $\bar{P}A =_{df} \neg FA$
- $\bar{F}A =_{df} \neg PA$

5.2.1 Semantics

Definition 83 (SPD4 model) An SPD4 model is a quadruple $\langle W, @, R, v\rangle$ where W is a set, $@ \in W$, $R \subseteq W^2$, and $v : At \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ such that $v_@(p) \in \{\mathbf{t}, \mathbf{f}\}$ for all $p \in At$. v is extended to $\bar{v} : L_{\mathsf{SPD}} \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ via the following clauses:

$$
\begin{array}{lcl}
\bar{v}_w(p) & = & v_w(p) \\
1 \in \bar{v}_w(\neg A) & \Leftrightarrow & 0 \in \bar{v}_w(A) \\
0 \in \bar{v}_w(\neg A) & \Leftrightarrow & 1 \in \bar{v}_w(A) \\
1 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ or } 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 0 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \Rightarrow 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(OA) & \Leftrightarrow & \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(OA) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(OA)) \text{ or} \\
 & & (w \neq @ \text{ and } \exists u(wRu \text{ and } 0 \in \bar{v}_u(A))) \\
1 \in \bar{v}_w(PA) & \Leftrightarrow & \exists u(wRu \text{ and } 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(PA) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(PA)) \text{ or} \\
 & & (w \neq @ \text{ and } \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A)))
\end{array}
$$

The following is a crucial feature of this construction:

Lemma 28 (home world lemma) For all SPD4 models $\langle W, @, R, v\rangle$ and all $A \in L_{\mathsf{SPD}}$, $\bar{v}_@(A) \in \{\mathbf{t}, \mathbf{f}\}$. That is, every formula takes a classical truth value at the home world.

Proof. A simple induction on the length of A. ∎

The reader can check that the following hold:

$$
\begin{array}{lcl}
1 \in \bar{v}_w(FA) & \Leftrightarrow & \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(FA) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(FA)) \text{ or} \\
 & & (w \neq @ \text{ and } \exists u(wRu \text{ and } 1 \in \bar{v}_u(A))) \\
1 \in \bar{v}_w(\bar{P}A) & \Leftrightarrow & 0 \in \bar{v}_w(FA) \\
0 \in \bar{v}_w(\bar{P}A) & \Leftrightarrow & 1 \in \bar{v}_w(FA) \\
1 \in \bar{v}_w(\bar{F}A) & \Leftrightarrow & 0 \in \bar{v}_w(PA) \\
0 \in \bar{v}_w(\bar{F}A) & \Leftrightarrow & 1 \in \bar{v}_w(PA)
\end{array}
$$

Definition 84 (SPD3 model) An SPD3 model is an SPD4 model $\langle W, @, R, v\rangle$ such that for all $w \in W$ and all $p \in At$, $v_w(p) \in \{\mathbf{t}, \mathbf{f}, \mathbf{b}\}$.

Definition 85 ($\mathsf{SPD}\frac{3}{4}\mathsf{s}$ model) An $\mathsf{SPD}\frac{3}{4}\mathsf{s}$ model is an $\mathsf{SPD}\frac{3}{4}$ model $\langle W, @, R, v\rangle$ such that R is serial, i.e. $\forall w \exists u\ wRu$.

Definition 86 ($\mathsf{SPD}\frac{3}{4}\mathsf{h}$ model) An $\mathsf{SPD}\frac{3}{4}\mathsf{h}$ model is an $\mathsf{SPD}\frac{3}{4}$ model $\langle W, @, R, v\rangle$ such that R is shift-reflexive, i.e. $\forall w \forall u(wRu \Rightarrow wRu)$.

Definition 87 ($\mathsf{SPD}\frac{3}{4}\mathsf{sh}$ model) An $\mathsf{SPD}\frac{3}{4}\mathsf{sh}$ model is an $\mathsf{SPD}\frac{3}{4}$ model $\langle W, @, R, v\rangle$ such that R is both serial and shift-reflexive.

Lemma 29 (exhaustion lemma) For all SPD3+ models $\langle W, @, R, v\rangle$, all $w \in W$, and all $A \in L_{\mathsf{SPD}}$, either $1 \in \bar{v}_w(A)$ or $0 \in \bar{v}_w(A)$.

Proof. A simple induction on the length of A. ■

Remark 33 It is easy to check that the Home World Lemma holds for our seven extensions of SPD4.

Definition 88 (semantic consequence) $\Gamma \models_{\mathsf{SPD}\frac{3}{4}+} A$ iff for all $\mathsf{SPD}\frac{3}{4}+$ models $\langle W, @, R, v\rangle$, if $1 \in \bar{v}_@(B)$ for all $B \in \Gamma$, then $1 \in \bar{v}_@(A)$.

Remark 34 Note that, given the Home World Lemma (Lemma 28), we could equivalently define semantic consequence as follows: $\Gamma \models_{\mathsf{SPD}\frac{3}{4}+} A$ iff for all $\mathsf{SPD}\frac{3}{4}+$ models $\langle W, @, R, v\rangle$, if $\bar{v}_@(B) = \mathbf{t}$ for all $B \in \Gamma$, then $\bar{v}_@(A) = \mathbf{t}$.

5.2.2 Proof theory

Definition 89 (initial list) The initial list for Γ / A, where $\Gamma = \{B_0, \ldots, B_n\}$, is

$$\begin{array}{c} B_0\ 0+ \\ \vdots \\ B_n\ 0+ \\ A\ 0- \end{array}$$

Definition 90 (closed branch, open branch) In $\mathsf{SPD4}+$, a branch is *closed* iff nodes of the forms $A\ x+$ and $A\ x-$, $A\ 0-$ and $\neg A\ 0-$, or $A\ 0+$ and $\neg A\ 0+$ occur on it. In $\mathsf{SPD3}+$ we have the additional closure condition: $A\ x-$ and $\neg A\ x-$.

The tableau rules for $\mathsf{SPD}\frac{3}{4}+$ are as follows. (These rules are presented in undiluted form in Appendix B.) I will not provide much explanation of these rules, assuming the reader can figure out what the rules "say" based on his or her familiarity with the tableau systems presented in earlier chapters.

Remark 35 As usual, redundant nodes may not be added to a branch, a rule may be applied to a given set of predecessor nodes only once, and "null" applications of rules are now allowed.

Definition 91 (double negation) The $\mathsf{SPD}\frac{3}{4}+$ rule for double negation is:

$$\begin{array}{c} {\scriptstyle [\neg\neg]} \\ \neg\neg A\ x\pm \\ \downarrow \\ A\ x\pm \end{array}$$

Definition 92 (disjunction rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for disjunction are:

$$\begin{array}{ccc} {\scriptstyle [\vee+]} & & \\ & A \vee B\ x+ & \\ \swarrow & & \searrow \\ A\ x+ & & B\ x+ \end{array} \qquad \begin{array}{c} {\scriptstyle [\vee-]} \\ A \vee B\ x- \\ \downarrow \\ A\ x- \\ B\ x- \end{array}$$

Definition 93 (negated disjunction rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for negated disjunction are:

$$\begin{array}{c} [\neg\vee+] \\ \neg(A \vee B)\ x+ \\ \downarrow \\ \neg A\ x+ \\ \neg B\ x+ \end{array} \qquad \begin{array}{c} [\neg\vee-] \\ \neg(A \vee B)\ x- \\ \swarrow \quad \searrow \\ \neg A\ x- \qquad \neg B\ x- \end{array}$$

Definition 94 (implication rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for implication are:

$$\begin{array}{c} [\supset+] \\ A \supset B\ x+ \\ \swarrow \quad \searrow \\ A\ x- \qquad B\ x+ \end{array} \qquad \begin{array}{c} [\supset-] \\ A \supset B\ x- \\ \downarrow \\ A\ x+ \\ B\ x- \end{array}$$

Definition 95 (negated implication rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for negated implication are:

$$\begin{array}{c} [\neg\supset+] \\ \neg(A \supset B)\ x+ \\ \downarrow \\ A\ x+ \\ \neg B\ x+ \end{array} \qquad \begin{array}{c} [\neg\supset-] \\ \neg(A \supset B)\ x+ \\ \swarrow \quad \searrow \\ A\ x- \qquad \neg B\ x- \end{array}$$

Definition 96 (obligation rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for obligation are:

$$\begin{array}{c} [O+] \\ OA\ x+ \\ x \rhd y \\ \downarrow \\ A\ y+ \end{array} \qquad \begin{array}{c} [O-] \\ OA\ x- \\ \downarrow \\ x \rhd i \\ A\ i- \end{array}$$

Definition 97 (negated obligation rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for negated obligation are:

$$\begin{array}{c} [\neg O_0] \\ \neg OA\ 0\pm \\ \downarrow \\ OA\ 0\mp \end{array} \qquad \begin{array}{c} [\neg O+] \\ \neg OA\ z+ \\ \downarrow \\ z \rhd i \\ \neg A\ i+ \end{array} \qquad \begin{array}{c} [\neg O-] \\ \neg OA\ z- \\ z \rhd y \\ \downarrow \\ \neg A\ y- \end{array}$$

Remark 36 Here z is any *positive* natural number on the branch (i.e. any number other than 0). The rule $[\neg O_0]$ says that if $\neg OA\ 0+$ ($\neg OA\ 0-$) occurs on a branch, then $OA\ 0-$ ($OA\ 0+$) can be added to the tip of that branch. In these rules we see the proof-theoretic manifestation of the "double standard" in the falsity condition for O.

Definition 98 (permission rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for permission are:

$$\begin{array}{ccc} {\scriptstyle [P+]} & \quad & {\scriptstyle [P-]} \\ PA\ x+ & & PA\ x- \\ \downarrow & & x \rhd y \\ x \rhd i & & \downarrow \\ A\ i+ & & A\ y- \end{array}$$

Definition 99 (negated permission rules) The $\mathsf{SPD}\frac{3}{4}+$ rules for negated permission are:

$$\begin{array}{ccccc} {\scriptstyle [\neg P_0]} & \quad & {\scriptstyle [\neg P+]} & \quad & {\scriptstyle [\neg P-]} \\ \neg PA\ 0\pm & & \neg PA\ z+ & & \neg PA\ z- \\ \downarrow & & z \rhd y & & \downarrow \\ PA\ 0\mp & & \downarrow & & z \rhd i \\ & & \neg A\ y+ & & \neg A\ i- \end{array}$$

Definition 100 (seriality rules) $\mathsf{SPD}\frac{3}{4}\mathsf{s}$ and $\mathsf{SPD}\frac{3}{4}\mathsf{sh}$ have the seriality rules:

$$\begin{array}{ccccccc} {\scriptstyle [sO+]} & \quad & {\scriptstyle [sO-]} & \quad & {\scriptstyle [sP+]} & \quad & {\scriptstyle [sP-]} \\ OA\ x+ & & \neg OA\ z- & & \neg PA\ z+ & & PA\ x- \\ \boxed{x \rhd y} & & \boxed{z \rhd y} & & \boxed{z \rhd y} & & \boxed{x \rhd y} \\ \downarrow & & \downarrow & & \downarrow & & \downarrow \\ x \rhd i & & z \rhd i & & z \rhd i & & x \rhd i \\ A\ i+ & & \neg A\ i- & & \neg A\ i+ & & A\ i- \end{array}$$

Recall from Definition 15 (p. 63) that a $\boxed{\text{box}}$ indicates that the node in question *does not* occur on the branch. So, e.g., the rule $[sP+]$ says that if a node of the form $\neg PA$ $z-$ ($z \neq 0$) occurs on a branch, and no node of the form $z \rhd y$ occurs on that branch, then nodes of the forms $z \rhd i$ and $\neg A\ i+$ may be added to the tip of the branch.

Remark 37 Note that the seriality rules $[sO-]$ and $[sP+]$ can only be applied to nodes of the forms $\neg OA\ x-$ and $\neg PA\ x+$, respectively, when $x \neq 0$. For if $x = 0$, then $\neg OA\ x-$ will be converted to $OA\ x+$ via $[\neg O_0]$, and $\neg PA\ x+$ will be converted to $PA\ x-$ via $[\neg P_0]$.

Definition 101 (shift-reflexivity rule) $\mathsf{SPD}\frac{3}{4}\mathsf{h}$ and $\mathsf{SPD}\frac{3}{4}\mathsf{sh}$ have the shift-reflexivity rule:

$$\begin{array}{c} [h] \\ x \triangleright y \\ \downarrow \\ y \triangleright y \end{array}$$

Definition 102 (tableau) An $\mathsf{SPD}\frac{3}{4}+$ tableau for the inference $\Gamma \diagup A$ is any tableau that results from 0 or more applications of $\mathsf{SPD}\frac{3}{4}+$ tableau rules to the initial list for $\Gamma \diagup A$.

Lemma 30 (non-redundancy) No $\mathsf{PD}\frac{3}{4}+$ tableau is redundant.

Proof. A simple induction on the complexity of tableaus. Similar to the proof of Lemma 1 (p. 65). ■

Definition 103 (complete branch/tableau) A branch of an $\mathsf{SPD}\frac{3}{4}+$ tableau is *complete* just in case it is closed or each of the following holds:

1. For each node of the form $\neg\neg A\ x+$ on the branch, a node of the form $A\ x+$ is on the branch.

2. For each node of the form $\neg\neg A\ x-$ on the branch, a node of the form $A\ x-$ is on the branch.

3. For each node of the form $A \vee B\ x+$ on the branch, a node of the form $A\ x+$ or $B\ x+$ is on the branch.

4. For each node of the form $A \vee B\ x-$ on the branch, nodes of the forms $A\ x-$ and $B\ x-$ are on the branch.

5. For each node of the form $\neg(A \vee B)\ x+$ on the branch, nodes of the forms $\neg A$ $x+$ and $\neg B\ x+$ are on the branch.

6. For each node of the form $\neg(A \vee B)\ x-$ on the branch, a node of the form $\neg A\ x-$ or $\neg B\ x-$ is on the branch.

7. For each node of the form $A \supset B\ x+$ on the branch, a node of the form $A\ x-$ or $B\ x+$ is on the branch.

8. For each node of the form $A \supset B\ x-$ on the branch, nodes of the forms $A\ x+$ and $B\ x-$ are on the branch.

9. For each node of the form $\neg(A \supset B)\ x+$ on the branch, nodes of the forms $A\ x+$ and $\neg B\ x+$ are on the branch.

10. For each node of the form $\neg(A \supset B)\ x-$ on the branch, a node of the form $A\ x-$ or $\neg B\ x-$ is on the branch.

11. For each pair of nodes of the forms $OA\ x+$ and $x \triangleright y$ on the branch, a node of the form $A\ y+$ is on the branch.

12. For each node of the form $OA\ x-$ on the branch, nodes of the forms $x \triangleright i$ and $A\ i-$ are on the branch.

13. For each node of the form $\neg OA\ 0+$ on the branch, a node of the form $OA\ 0-$ is on the branch.

14. For each node of the form $\neg OA\ 0-$ on the branch, a node of the form $OA\ 0+$ is on the branch.

15. For each node of the form $\neg OA\ z+$ $(z \neq 0)$ on the branch, nodes of the forms $z \triangleright i$ and $\neg A\ i+$ are on the branch.

16. For each pair of nodes of the forms $\neg OA\ z-$ and $z \triangleright y$ $(z \neq 0)$ on the branch, a node of the form $\neg A\ y-$ is on the branch.

17. For each node of the form $PA\ x+$ on the branch, nodes of the forms $x \triangleright i$ and $A\ i+$ are on the branch.

18. For each pair of nodes of the forms $PA\ x-$ and $x \triangleright y$ on the branch, a node of the form $A\ y-$ is on the branch.

19. For each node of the form $\neg PA\ 0+$ on the branch, a node of the form $PA\ 0-$ is on the branch.

20. For each node of the form $\neg PA\ 0-$ on the branch, a node of the form $PA\ 0+$ is on the branch.

21. For each pair of nodes of the forms $\neg PA\ z+$ and $z \rhd y$ $(z \neq 0)$ on the branch, a node of the form $\neg A\ y+$ is on the branch.

22. For each node of the form $\neg PA\ z+$ $(z \neq 0)$ on the branch, nodes of the forms $z \rhd i$ and $\neg A\ i-$ are on the branch.

23. If $\mathsf{SPD}\frac{3}{4}+$ is a serial system, for each node of the form $OA\ x$ on the branch, there are nodes of the forms $x \rhd i$ and $A\ i+$ on the branch.

24. If $\mathsf{SPD}\frac{3}{4}+$ is a serial system, for each node of the form $\neg OA\ z-$ $(z \neq 0)$ on the branch, there are nodes of the forms $x \rhd i$ and $\neg A\ i-$ on the branch.

25. If $\mathsf{SPD}\frac{3}{4}+$ is a serial system, for each node of the form $\neg PA\ z+$ $(z \neq 0)$ on the branch, there are nodes of the forms $z \rhd i$ and $\neg A\ i+$ on the branch.

26. If $\mathsf{SPD}\frac{3}{4}+$ is a serial system, for each node of the form $PA\ x-$ on the branch, there are nodes of the forms $x \rhd i$ and $A\ i-$ on the branch.

27. If $\mathsf{SPD}\frac{3}{4}+$ is a shift-reflexive system, for each node of the form $x \rhd y$ $(x \neq y)$ on the branch, a node of the form $y \rhd y$ is on the branch.

A tableau is complete iff all of its branches are complete.

Definition 104 (closed/open tableau) A tableau is closed just in case all of its branches are closed; otherwise it is open.

Definition 105 (tableau-theoretic consequence) $\Gamma \Vdash_{\mathsf{SPD}\frac{3}{4}+} A$ iff there is a closed $\mathsf{SPD}\frac{3}{4}+$ tableau for $\Gamma \,/\, A$.

We have the following shortcut/convenience rules:

$$
\begin{array}{c}
[\wedge+] \\
A \wedge B\ x+ \\
\downarrow \\
A\ x+ \\
B\ x+
\end{array}
\qquad
\begin{array}{c}
[\wedge-] \\
A \wedge B\ x- \\
\swarrow \quad \searrow \\
A\ x- \qquad B\ x-
\end{array}
\qquad
\begin{array}{c}
[\neg\wedge] \\
\neg(A \wedge B)\ x\pm \\
\downarrow \\
\neg A \vee \neg B\ x\pm
\end{array}
\qquad
\begin{array}{c}
[\neg\vee] \\
\neg(A \vee B)\ x\pm \\
\downarrow \\
\neg A \wedge \neg B\ x\pm
\end{array}
$$

As usual, these rules are superfluous and not included in the "official" proof theory.

5.2.3 Example proofs

Example 22 (disjunctive syllogism) Here is a proof that $A \vee B, \neg A \Vdash_{\mathsf{SPD4}} B$, using just primitive rules:

$$\begin{array}{lccl}
1. & \multicolumn{2}{c}{A \vee B\ 0+\checkmark} & \text{initial list} \\
2. & \multicolumn{2}{c}{\neg A\ 0+} & \text{"} \\
3. & \multicolumn{2}{c}{B\ 0-} & \text{"} \\
 & \multicolumn{2}{c}{\swarrow\searrow} & \\
4. & A\ 0+ & B\ 0+ & \text{from 1 by } [\vee+] \\
 & * & * & \text{(both branches close)}
\end{array}$$

Example 23 (modus tollens) Here is a proof that $A \supset B, \neg B \Vdash_{\mathsf{SPD4}} A$, using just primitive rules:

$$\begin{array}{lccl}
1. & \multicolumn{2}{c}{A \supset B\ 0+\checkmark} & \text{initial list} \\
2. & \multicolumn{2}{c}{\neg B\ 0+} & \text{"} \\
3. & \multicolumn{2}{c}{\neg A\ 0-} & \text{"} \\
 & \multicolumn{2}{c}{\swarrow\searrow} & \\
4. & A\ 0- & B\ 0+ & \text{from 1 by } [\supset+] \\
 & * & * & \text{(both branches close)}
\end{array}$$

Example 24 (inescapable deontic explosion) Here is a proof that $Op, Fp \nVdash_{\mathsf{SPD3}} Oq$, using just primitive rules:

$$\begin{array}{lcl}
1. & Op\ 0+ & \text{initial list} \\
2. & O\neg p\ 0+ & \text{"} \\
3. & Oq\ 0-\checkmark & \text{"} \\
4. & 0 \rhd 1 & \text{from 3 by } [O-] \\
5. & q\ 1- & \text{"} \\
6. & p\ 1+ & \text{from 1 and 4 by } [O+] \\
7. & \neg p\ 1+ & \text{from 2 and 4 by } [O+] \\
 & \uparrow & \text{(open and complete)}
\end{array}$$

We can read the following counterexample off the open branch. (As usual, the procedure is made more explicit in the definition of induced model below (Definition 107, p. 107).) Let $W = \{@, w_1\}$, $@ = w_0$, $R = \{\langle @, w_1\rangle\}$, $v_{w_1}(p) = \mathbf{b}$, $v_{w_1}(q) = \mathbf{f}$. On this model we have $\bar{v}_@(Op) = \bar{v}_@(Fp) = \mathbf{t}$ but $\bar{v}_@(Oq) = \mathbf{f}$.

Example 25 (no escapable conflicts) Here is a proof that $\nVdash_{\mathsf{SPD3}} \neg(Fp \wedge Pp)$, using primitive and derived rules:

1.	$\neg(O\neg p \wedge Pp)\ 0 - \checkmark$	initial list
2.	$\neg O\neg p \vee \neg Pp\ 0 - \checkmark$	from 1 by $[\neg\wedge]$
3.	$\neg O\neg p\ 0 - \checkmark$	from 2 by $[\vee-]$
4.	$\neg Pp\ 0 - \checkmark$	"
5.	$O\neg p\ 0+$	from 3 by $[\neg O_0]$
6.	$Pp\ 0 + \checkmark$	from 4 by $[\neg P_0]$
7.	$0 \rhd 1$	from 6 by $[P+]$
8.	$p\ 1+$	"
9.	$\neg p\ 1+$	from 5 and 7 by $[O+]$
	$\uparrow$	

The counterexample used in the last example works in this case, too.

Example 26 (ought implies may) Here is a proof that $\Vdash_{\mathsf{SPD4s}} OA \supset PA$:

1.	$OA \supset PA\ 0 - \checkmark$	initial list
2.	$OA\ 0+$	from 1 by $[\supset -]$
3.	$PA\ 0-$	"
4.	$0 \rhd 1$	from 2 by $[sO+]$
5.	$A\ 1+$	"
6.	$A\ 1-$	from 3 and 4 by $[P-]$
	$*$	from 5 and 6

Example 27 (ought implies negative permission) Here is a proof that $\nVdash_{\mathsf{SPD3s}}$ $Op \supset \bar{P}p$:

1.	$Op \supset \neg O\neg p\ 0 - \checkmark$	initial list
2.	$Op\ 0+$	from 1 by $[\supset -]$
3.	$\neg O\neg p\ 0 - \checkmark$	"
4.	$O\neg p\ 0+$	from 3 by $[\neg O_0]$
5.	$0 \rhd 1$	from 2 by $[sO+]$
6.	$p\ 1+$	"
7.	$\neg p\ 1+$	from 4 and 5 by $[O+]$
	$\uparrow$	

Example 28 (obey all rules) Here is a proof that $\Vdash_{\mathsf{SPD4h}} O(OA \supset A)$:

$$\begin{array}{lcl}
1. & O(OA \supset A)\ 0 - \checkmark & \text{initial list} \\
2. & 0 \triangleright 1 & \text{from 1 by } [O-] \\
3. & OA \supset A\ 1 - \checkmark & " \\
4. & OA\ 1+ & \text{from 3 by } [\supset -] \\
5. & A\ 1- & " \\
6. & 1 \triangleright 1 & \text{by } [h] \\
7. & A\ 1+ & \text{from 4 and 6 by } [O+] \\
 & * & \text{from 5 and 7}
\end{array}$$

Example 29 (DDS for negative prohibition) Here is a proof that $O(A \vee B), \bar{F}A \Vdash_{\mathsf{SPD4}} OB$:

$$\begin{array}{lccl}
1. & \multicolumn{2}{c}{O(A \vee B)\ 0+} & \text{initial list} \\
2. & \multicolumn{2}{c}{\neg PA\ 0 + \checkmark} & " \\
3. & \multicolumn{2}{c}{OB\ 0 - \checkmark} & " \\
4. & \multicolumn{2}{c}{PA\ 0-} & \text{from 2 by } [\neg P_0] \\
5. & \multicolumn{2}{c}{0 \triangleright 1} & \text{from 3 by } [O-] \\
6. & \multicolumn{2}{c}{B\ 1-} & " \\
7. & \multicolumn{2}{c}{A\ 1-} & \text{from 4 and 5 by } [P-] \\
8. & \multicolumn{2}{c}{A \vee B\ 1 + \checkmark} & \text{from 1 and 5 by } [O+] \\
 & \swarrow & \searrow & \\
9. & A\ 1+ & B\ 1+ & \text{from 8 by } [\vee+] \\
 & * & * & \text{(both branches close)}
\end{array}$$

Example 30 (covering principle) Here is a proof that $\Vdash_{\mathsf{SPD3}} FA \vee PA$:

$$\begin{array}{lcl}
1. & O\neg A \vee PA\ 0 - \checkmark & \text{initial list} \\
2. & O\neg A\ 0 - \checkmark & \text{from 1 by } [\vee-] \\
3. & PA\ 0- & " \\
4. & 0 \triangleright 1 & \text{from 2 by } [O-] \\
5. & \neg A\ 1- & " \\
6. & A\ 1- & \text{from 3 and 4 by } [P-] \\
 & * & \text{from 5 and 6}
\end{array}$$

Remark 38 Note that the tableau doesn't close in SPD4, which allows gaps. Hence $\nVdash_{\mathsf{SPD4}} FA \vee PA$.

Example 31 (principle of permission) Here is a proof that $\Vdash_{\mathsf{SPD3s}} PA \vee P\neg A$:

1.	$PA \vee P\neg A\ 0- \checkmark$	initial list
2.	$PA\ 0-$	from 1 by $[\vee-]$
3.	$P\neg A\ 0-$	"
4.	$0 \triangleright 1$	from 2 by $[sP-]$
5.	$A\ 1-$	"
6.	$\neg A\ 1-$	from 3 and 4 by $[P-]$
	$*$	from 5 and 6

Remark 39 Note that both seriality and the exclusion of gaps are required in order to close the tableau. Thus $PA \vee P\neg A$ is *not* valid in SPD3 or SPD4s.

Example 32 (weird inference) Here is a proof that $P(A \wedge B), O(A \supset \neg B) \Vdash_{\mathsf{SPD4}} P(B \wedge \neg B)$:

1.	$P(A \wedge B)\ 0+$		initial list
2.	$O(A \supset \neg B)\ 0+$		"
3.	$P(B \wedge \neg B)\ 0-$		"
4.	$0 \triangleright 1$		from 1 by $[P+]$
5.	$A \wedge B\ 1+$		"
6.	$A\ 1+$		from 5 by $[\wedge+]$
7.	$B\ 1+$		"
8.	$A \supset \neg B\ 1+$		from 2 and 4 by $[O+]$
9.	$B \wedge \neg B\ 1-$		from 3 and 4 by $[P-]$
	$\swarrow$	$\searrow$	
10.	$A\ 1-$	$\neg B\ 1+$	from 8 by $[\supset +]$
	$*$	$\swarrow \searrow$	
11.		$B\ 1- \qquad \neg B\ 1-$	from 9 by $[\wedge-]$
		$* \qquad\qquad *$	

5.2.4 Soundness and completeness

I now prove the soundness and completeness of the proof theory with respect to the semantics.

Definition 106 (faithful) Let b be a branch of an $\mathsf{SPD}\frac{3}{4}+$ tableau. An $\mathsf{SPD}\frac{3}{4}+$ model $\mathfrak{M} = \langle W, @, R, v\rangle$ is *faithful* to b just in case there is a function $w : \mathbb{N} \longrightarrow W$ such that:[4]

- $w_x = @$ iff $x = 0$
- if $x \rhd y$ is on b, then $w_x R w_y$
- if $A\ x+$ is on b, then $1 \in \bar{v}_{w_x}(A)$
- if $A\ x-$ is on b, then $1 \notin \bar{v}_{w_x}(A)$.

We say that w *shows* $\mathfrak{M}$ to be faithful to b.

Lemma 31 (faith lemma) If a branch, b, of an $\mathsf{SPD}\frac{3}{4}+$ tableau is closed, then no $\mathsf{SPD}\frac{3}{4}+$ model is faithful to b.

Proof. For reductio, suppose that the $\mathsf{SPD}\frac{3}{4}+$ model $\mathfrak{M} = \langle W, @, R, v\rangle$ is faithful to b . Since b is closed, there are three cases:

1. Nodes of the forms $A\ x+$ and $A\ x-$ occur on b. Then, since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(A)$ and $1 \notin \bar{v}_{w_x}(A)$. (Contradiction.)
2. Nodes of the forms $A\ 0+$ and $\neg A\ 0+$ occur on it. Then, since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{@}(A)$ and $1 \in \bar{v}_{@}(\neg A)$, contradicting the Home World Lemma (Lemma 28).
3. Nodes of the forms $A\ 0-$ and $\neg A\ 0-$ occur on it. Then, since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{@}(A)$ and $1 \notin \bar{v}_{@}(\neg A)$, contradicting the Home World Lemma (Lemma 28).
4. $\mathsf{SPD}\frac{3}{4}+$ is an extension of $\mathsf{SPD3}$, and nodes of the forms, $A\ x-$ and $\neg A\ x-$ occur on b. Then, since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(A)$ and $1 \notin \bar{v}_{w_x}(\neg A)$, contradicting the Exhaustion Lemma (Lemma 29).

In each case, we have a contradiction. ■

Lemma 32 (soundness lemma) If an $\mathsf{SPD}\frac{3}{4}+$ model $\mathfrak{M} = \langle W, @, R, v\rangle$ is faithful to a branch, b, and an $\mathsf{SPD}\frac{3}{4}+$ tableau rule is applied to b, then $\mathfrak{M}$ is faithful to at least one of the branches thereby generated.

[4] I write $w(x)$ as w_x, etc.

Proof. The proof is by cases. There are 24 cases to consider (one for each tableau rule). The cases for $\neg$ and $\vee$ are the same as in $\mathsf{PD}\frac{3}{4}+$. Thus I will only prove the cases for $\supset$, O, and P.

- **Case 6** ($[\supset +]$). Suppose $A \supset B\ x+$ is on b, and $[\supset +]$ is applied. Then $b(A\ x-)$ and $b(B\ x+)$ are generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(A \supset B)$. Thus either $1 \notin \bar{v}_{w_x}(A)$ or $1 \in \bar{v}_{w_x}(B)$. Thus $\mathfrak{M}$ is faithful to either $b(A\ x-)$ or $b(B\ x+)$.

- **Case 7** ($[\supset -]$). Suppose $A \supset B\ x-$ is on b, and $[\supset -]$ is applied. Then $b(A\ x+, B\ x-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(A \supset B)$. Thus $1 \in \bar{v}_{w_x}(A)$ and $1 \notin \bar{v}_{w_x}(B)$. Thus $\mathfrak{M}$ is faithful to $b(A\ x+, B\ x-)$.

- **Case 8** ($[\neg \supset +]$). Suppose $\neg(A \supset B)\ x+$ is on b, and $[\neg \supset +]$ is applied. Then $b(A\ x+, \neg B\ x+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(\neg(A \supset B))$. Thus $0 \in \bar{v}_{w_x}(A \supset B)$. Thus $1 \in \bar{v}_{w_x}(A)$ and $0 \in \bar{v}_{w_x}(B)$. Thus $1 \in \bar{v}_{w_x}(\neg B)$. Thus $\mathfrak{M}$ is faithful to $b(A\ x+, \neg B\ x+)$.

- **Case 9** ($[\neg \supset -]$). Suppose $\neg(A \supset B)\ x-$ is on b, and $[\neg \supset -]$ is applied. Then $b(A\ x-)$ and $b(\neg B\ x-)$ are generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(\neg(A \supset B))$. Thus $0 \notin \bar{v}_{w_x}(A \supset B)$. Thus either $1 \notin \bar{v}_{w_x}(A)$ or $0 \notin \bar{v}_{w_x}(B)$. Thus either $1 \notin \bar{v}_{w_x}(A)$ or $1 \notin \bar{v}_{w_x}(\neg B)$. Thus $\mathfrak{M}$ is faithful to either $b(A\ x-)$ or $b(\neg B\ x-)$.

- **Case 10** ($[O+]$). Suppose $OA\ x+$ and $x \triangleright y$ are on b, and $[O+]$ is applied. Then $b(A\ y+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(OA)$ and $w_x R w_y$. Thus $1 \in \bar{v}_u(A)$ for all u such that $w_x R u$. Thus $1 \in \bar{v}_{w_y}(A)$. Thus $\mathfrak{M}$ is faithful to $b(A\ y+)$.

- **Case 11** ($[O-]$). Suppose $OA\ x-$ is on b, and $[O-]$ is applied. Then $b(x \triangleright i, A\ i-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(OA)$. Thus there is a world (call it \$) such that $w_x R\$$ and $1 \notin \bar{v}_{\$}(A)$. Let $\hat{w}$ be just like w except that $\hat{w}_i = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Moreover, $\hat{w}_x R \hat{w}_i$ and $1 \notin \bar{v}_{\hat{w}_i}(A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, A\ i-)$.

- **Case 12** ($[\neg O_0]$). Suppose $\neg OA\ 0+$ is on b, and $[\neg O_0]$ is applied. Then $b(OA\ 0-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(\neg OA)$. Thus $0 \in \bar{v}_{w_x}(OA)$. Thus, since $w_0 = @$, $1 \notin \bar{v}_{w_0}(OA)$. Thus $\mathfrak{M}$ is faithful to $b(OA\ 0-)$. Now suppose that $\neg OA\ 0-$ is on b, and $[\neg O_0]$ is applied. Then $b(OA\ 0+)$ is generated.

Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(\neg OA)$. Thus $0 \notin \bar{v}_{w_x}(OA)$. Thus, since $w_0 = @$, $1 \in \bar{v}_{w_0}(OA)$. Thus $\mathfrak{M}$ is faithful to $b(OA\ 0+)$.

- **Case 13** ($[\neg O+]$). Suppose $\neg OA\ z+$ is on b, and $[\neg O+]$ is applied. Then $b(z \triangleright i, \neg A\ i+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_z}(\neg OA)$. Thus $0 \in \bar{v}_{w_z}(OA)$. Thus, since $w_z \neq @$, there is a world (call it \$) such that $w_z R\$$ and $0 \in \bar{v}_{\$}(A)$. Let $\hat{w}$ be just like w except that $\hat{w}_i = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Moreover, $\hat{w}_z R \hat{w}_i$ and $1 \in \bar{v}_{\hat{w}_i}(\neg A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(z \triangleright i, \neg A\ i+)$.

- **Case 14** ($[\neg O-]$). Suppose $\neg OA\ z-$ and $z \triangleright y$ are on b, and $[\neg O-]$ is applied. Then $b(\neg A\ y-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_z}(\neg OA)$ and $w_z R w_y$. Thus $0 \notin \bar{v}_{w_z}(OA)$. Thus, since $w_z \neq @$, $0 \notin \bar{v}_u(A)$ for all u such that $w_z R u$. Thus $0 \notin \bar{v}_{w_y}(A)$. Thus $1 \notin \bar{v}_{w_y}(\neg A)$. Thus $\mathfrak{M}$ is faithful to $b(\neg A\ y-)$.

- **Case 15** ($[P+]$). Suppose $PA\ x+$ is on b, and $[P+]$ is applied. Then $b(x \triangleright i, A\ i+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(PA)$. Thus there is a world (call it \$) such that $w_x R\$$ and $1 \in \bar{v}_{\$}(A)$. Let $\hat{w}$ be just like w except that $\hat{w}_i = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Moreover, $\hat{w}_x R \hat{w}_i$ and $1 \in \bar{v}_{\hat{w}_i}(A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, A\ i+)$.

- **Case 16** ($[P-]$). Suppose $PA\ x-$ and $x \triangleright y$ are on b, and $[P-]$ is applied. Then $b(A\ y-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(PA)$ and $w_x R w_y$. Thus $1 \notin \bar{v}_u(A)$ for all u such that $w_x R u$. Thus $1 \notin \bar{v}_{w_y}(A)$. Thus $\mathfrak{M}$ is faithful to $b(A\ y-)$.

- **Case 17** ($[\neg P_0]$). Suppose $\neg PA\ 0+$ is on b, and $[\neg P_0]$ is applied. Then $b(PA\ 0-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_x}(\neg PA)$. Thus $0 \in \bar{v}_{w_x}(PA)$. Thus, since $w_0 = @$, $1 \notin \bar{v}_{w_0}(PA)$. Thus $\mathfrak{M}$ is faithful to $b(PA\ 0-)$. Now suppose that $\neg PA\ 0-$ is on b, and $[\neg P_0]$ is applied. Then $b(PA\ 0+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_x}(\neg PA)$. Thus $0 \notin \bar{v}_{w_x}(PA)$. Thus, since $w_0 = @$, $1 \in \bar{v}_{w_0}(PA)$. Thus $\mathfrak{M}$ is faithful to $b(PA\ 0+)$.

- **Case 18** ($[\neg P+]$). Suppose $\neg PA\ z+$ and $z \triangleright y$ are on b, and $[\neg P+]$ is applied. Then $b(\neg A\ y+)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \in \bar{v}_{w_z}(\neg PA)$ and $w_z R w_y$. Thus $0 \in \bar{v}_{w_z}(PA)$. Thus, since $w_z \neq @$, $0 \in \bar{v}_u(A)$ for all u such that $w_z R u$. Thus $0 \in \bar{v}_{w_y}(A)$. Thus $1 \in \bar{v}_{w_y}(\neg A)$. Thus $\mathfrak{M}$ is faithful to $b(\neg A\ y+)$.

- **Case 19** ($[\neg P-]$). Suppose $\neg PA\ z-$ is on b, and $[\neg P-]$ is applied. Then $b(z \rhd i, \neg A\ i-)$ is generated. Since $\mathfrak{M}$ is faithful to b, $1 \notin \bar{v}_{w_z}(\neg PA)$. Thus $0 \notin \bar{v}_{w_z}(PA)$. Thus, since $w_z \neq @$, there is a world (call it \$) such that $w_z R\$$ and $0 \notin \bar{v}_{\$}(A)$. Let $\hat{w}$ be just like w except that $\hat{w}_i = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Moreover, $\hat{w}_z R \hat{w}_i$ and $1 \notin \bar{v}_{\hat{w}_i}(\neg A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(z \rhd i, \neg A\ i-)$.

- **Case 20** ($[sO+]$). Suppose $OA\ x+$ occurs on b, no node of the form $x \rhd y$ occurs on b, and $[sO+]$ is applied. Then $b(x \rhd i, A\ i+)$ is generated. Since $[sO+]$ is a rule only for our serial systems, R is serial. Thus there is a world, call it \$, such that $w_x R\$$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Thus $1 \in \bar{v}_{\hat{w}_x}(OA)$. Thus $\forall u(\hat{w}_x Ru \Rightarrow 1 \in \bar{v}_u(A))$. Thus, since $\hat{w}_x R \hat{w}_i$, $1 \in v_{\hat{w}_i}(A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \rhd i, A\ i+)$.

- **Case 21** ($[sO-]$). Suppose $\neg OA\ z-$ occurs on b, no node of the form $z \rhd y$ occurs on b, and $[sO-]$ is applied. Then $b(z \rhd i, \neg A\ i-)$ is generated. Since $[sO-]$ is a rule only for our serial systems, R is serial. Thus there is a world, call it \$, such that $w_z R\$$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Thus $1 \notin \bar{v}_{\hat{w}_z}(\neg OA)$. Thus, since $z \neq 0$, $0 \notin \bar{v}_{\hat{w}_z}(OA)$. Thus $\forall u(\hat{w}_z Ru \Rightarrow 0 \notin \bar{v}_u(A))$. Thus, since $\hat{w}_z R \hat{w}_i$, $0 \notin v_{\hat{w}_i}(A)$. Thus $1 \notin v_{\hat{w}_i}(\neg A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(z \rhd i, \neg A\ i-)$.

- **Case 22** ($[sP+]$). Suppose $\neg PA\ z+$ occurs on b, no node of the form $z \rhd y$ occurs on b, and $[sP+]$ is applied. Then $b(z \rhd i, \neg A\ i+)$ is generated. Since $[sP+]$ is a rule only for our serial systems, R is serial. Thus there is a world, call it \$, such that $w_z R\$$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Thus $1 \in \bar{v}_{\hat{w}_z}(\neg PA)$. Thus, since $z \neq 0$, $0 \in \bar{v}_{\hat{w}_z}(PA)$. Thus $\forall u(\hat{w}_z Ru \Rightarrow 0 \in \bar{v}_u(A))$. Thus, since $\hat{w}_z R \hat{w}_i$, $0 \in v_{\hat{w}_i}(A)$. Thus $1 \in v_{\hat{w}_i}(\neg A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(z \rhd i, \neg A\ i+)$.

- **Case 23** ($[sP-]$). Suppose $PA\ x-$ occurs on b, no node of the form $x \rhd y$ occurs on b, and $[sP-]$ is applied. Then $b(x \rhd i, A\ i-)$ is generated. Since $[sP-]$ is a rule only for our serial systems, R is serial. Thus there is a world, call it \$, such that

$w_x R\$$. Let $\hat{w}$ be just like w except that $\hat{w}(i) = \$$. Since i does not occur on b, $\hat{w}$ shows $\mathfrak{M}$ to be faithful to b. Thus $1 \notin \bar{v}_{\hat{w}_x}(PA)$. Thus $\forall u(\hat{w}_x Ru \Rightarrow 1 \notin \bar{v}_u(A))$. Thus, since $\hat{w}_x R\hat{w}_i$, $1 \notin v_{\hat{w}_i}(A)$. Thus $\hat{w}$ shows $\mathfrak{M}$ to be faithful to $b(x \triangleright i, A\ i-)$.

- **Case 24** ($[h]$). Suppose $x \triangleright y$ occurs on b, and $[h]$ is applied. Then $b(y \triangleright y)$ is generated. Since $\mathfrak{M}$ is faithful to b, $w_x R w_y$. Since $[h]$ is a rule only for our shift-reflextive systems, R is shift-reflexive. Thus $w_y R w_y$. Thus $\mathfrak{M}$ is faithful to $b(y \triangleright y)$.

■

Theorem 19 (soundness) If $\Gamma \Vdash_{\mathsf{SPD}\frac{3}{4}+} A$, then $\Gamma \models_{\mathsf{SPD}\frac{3}{4}+} A$.

Proof. Suppose $\Gamma \not\models_{\mathsf{SPD}\frac{3}{4}+} A$. Then there is an $\mathsf{SPD}\frac{3}{4}+$ model $\mathfrak{M} = \langle W, @, R, v\rangle$ such that $1 \in \bar{v}_@(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}_@(A)$. $\mathfrak{M}$ is faithful to the initial list for Γ / A. Moreover, by the Soundness Lemma (Lemma 32), each subsequent application of an $\mathsf{SPD}\frac{3}{4}+$ tableau rule will yield at least one branch to which $\mathfrak{M}$ is faithful. Thus every $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A has at least one branch to which $\mathfrak{M}$ is faithful. Let T be an $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A, and let b be one of the branches of T to which $\mathfrak{M}$ is faithful. By the Faith Lemma (Lemma 31), b cannot be closed. Thus T is open. But T was an arbitrarily chosen $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A. Thus there is no closed $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A. Thus $\Gamma \nVdash_{\mathsf{SPD}\frac{3}{4}+} A$. ■

Definition 107 (induced model for non-serial systems) Let $\mathsf{SPD}\frac{3}{4}+$ be a non-serial SPD system. Let b be an open, complete branch of a $\mathsf{SPD}\frac{3}{4}+$ tableau. The $\mathsf{SPD}\frac{3}{4}+$ model *induced* by b is the model $\langle W, @, R, v\rangle$ such that $W = \{w_i : i$ is a natural number occurring on $b\}$, $w_x = @$ iff $x = 0$, $R = \{\langle w_i, w_j\rangle : i \triangleright j$ occurs on $b\}$, and v is determined by the following two charts (one for the case in which $x = 0$ and the other for the case in which $x \neq 0$):

$x = 0$				
$p\ x+$	$p\ x-$	$\neg p\ x+$	$\neg p\ x-$	$\mathsf{SPD}^{3}_{4}+$
Y	Y	Y	Y	$*$
Y	Y	Y	N	$*$
Y	Y	N	Y	$*$
Y	Y	N	N	$*$
Y	N	Y	Y	$*$
Y	N	Y	N	$*$
Y	N	N	Y	$v_{w_x}(p) = \mathbf{t}$
Y	N	N	N	$v_{w_x}(p) = \mathbf{t}$
N	Y	Y	Y	$*$
N	Y	Y	N	$v_{w_x}(p) = \mathbf{f}$
N	Y	N	Y	$*$
N	Y	N	N	$v_{w_x}(p) = \mathbf{f}$
N	N	Y	Y	$*$
N	N	Y	N	$v_{w_x}(p) = \mathbf{f}$
N	N	N	Y	$v_{w_x}(p) = \mathbf{t}$
N	N	N	N	$v_{w_x}(p) = \mathbf{t}$

<table>
<tr><th colspan="6">$x \neq 0$</th></tr>
<tr><th>p $x+$</th><th>p $x-$</th><th>$\neg p$ $x+$</th><th>$\neg p$ $x-$</th><th>SPD4+</th><th>SPD3+</th></tr>
<tr><td>Y</td><td>Y</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>Y</td><td>N</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>N</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>Y</td><td>N</td><td>N</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>N</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>Y</td><td>N</td><td>Y</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{b}$</td></tr>
<tr><td>Y</td><td>N</td><td>N</td><td>Y</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
<tr><td>Y</td><td>N</td><td>N</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
<tr><td>N</td><td>Y</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>N</td><td>Y</td><td>Y</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>Y</td><td>N</td><td>Y</td><td>$v_{w_x}(p) = \mathbf{n}$</td><td>$*$</td></tr>
<tr><td>N</td><td>Y</td><td>N</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>N</td><td>Y</td><td>Y</td><td colspan="2">$*$</td></tr>
<tr><td>N</td><td>N</td><td>Y</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{f}$</td></tr>
<tr><td>N</td><td>N</td><td>N</td><td>Y</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
<tr><td>N</td><td>N</td><td>N</td><td>N</td><td colspan="2">$v_{w_x}(p) = \mathbf{t}$</td></tr>
</table>

We need to show that $\langle W, @, R, v\rangle$, so defined, really is an $\mathsf{SPD}\frac{3}{4}+$ model. First, we need to verify that if $\mathsf{SPD}\frac{3}{4}+$ is a *shift-reflexive* system, then R is indeed shift-reflexive. Since b is complete, for each node of the form $x \triangleright y$ occurring on b, there is a node of the form $y \triangleright y$ on b. Thus for each w_x and $w_y \in W$ such that $w_x R w_y$, $w_y R w_y$.; that is, R is shift-reflexive. Second, we need to verify that (i) v is a well-defined function mapping world/atomic formula pairs into $\{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ (for SPD4+) and $\{\mathbf{t}, \mathbf{f}, \mathbf{b}\}$ (for SPD3+), and also that (ii) $v_@(p) \in \{\mathbf{t}, \mathbf{f}\}$ for all $p \in At$. It is easy to see that the charts above guarantee (i) and (ii).

Definition 108 (induced model for serial systems) Let $\mathsf{SPD}\frac{3}{4}+$ be a serial SPD system. Let b be an open, complete branch of a $\mathsf{SPD}\frac{3}{4}+$ tableau. The $\mathsf{SPD}\frac{3}{4}+$ model *induced* by b is the model $\langle W, @, R, v\rangle$ such that (i) $W = \{w_i : i$ is a natural number on $b\}$; (ii) $w_x = @$ iff $x = 0$; (iii) $R = \{\langle w_i, w_j\rangle : i \triangleright j$ is on $b\} \cup \{\langle w_i, w_i\rangle : i$ is on b but no node of the form $i \triangleright j$ is on $b\}$; and (iiii) v is determined by the same charts used in the previous definition.

We need to show that $\langle W, @, R, v\rangle$, so defined, really is an $\mathsf{SPD}\frac{3}{4}+$ model. First, we need to verify that R is indeed serial. That is, we need to show that for all $w \in W$, there is a $u \in W$ such that wRu. By clause (i) above, $W = \{w_i : i$ is a natural number on $b\}$. Thus we need to show that for each natural number i occurring on b, there is a $u \in W$ such that $w_i Ru$. Now, for each i occurring on b, there are just two possible cases:

- *Case 1.* A node of the form $i \triangleright j$ occurs on b. Then, by clause (iiI), $w_i R w_j$.
- *Case 2.* No node of the form $i \triangleright j$ occurs on b. Then, by clause (iIi), $w_i R w_i$.

In both cases, there is a $u \in W$ such that $w_i Ru$. Thus for all $w \in W$, there is a $u \in W$ such that wRu; i.e., R is serial. Second, we need to verify that if $\mathsf{SPD}\frac{3}{4}+$ is a shift-reflexive system, then R is shift-reflexive. That is, we need to show that for all natural numbers i, j occurring on b, if $w_i R w_j$, then $w_j R w_j$. If $i = j$, the result follows immediately. So suppose $i \neq j$. Then $i \triangleright j$ must occur on the branch. (For suppose it doesn't. Then, by clause (iii), $i = j$, contadicting our assumption.) Thus, since b is complete, $j \triangleright j$ occurs on the branch. Thus, by clause (iii), $w_j R w_j$, as required. Finally, we need to verify that v is a well-defined function mapping world/atomic formula pairs into $\{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ for the $\mathsf{SPD4}$ systems and $\{\mathbf{t}, \mathbf{f}, \mathbf{b}\}$ for the $\mathsf{SPD3}$ systems. It is easy to see that the chart guarantees this.

Lemma 33 (completeness lemma for non-serial systems) Let $\mathsf{SPD}\frac{3}{4}+$ be a non-serial SPD system. Let b be an open, complete branch of an $\mathsf{SPD}\frac{3}{4}+$ tableau. Let $\langle W, @, R, v\rangle$ be the $\mathsf{SPD}\frac{3}{4}+$ model induced by b. Then for all A :

1. if $A\ x+$ is on b, then $1 \in \bar{v}_{w_x}(A)$
2. if $A\ x-$ is on b, then $1 \notin \bar{v}_{w_x}(A)$
3. if $\neg A\ x+$ is on b, then $0 \in \bar{v}_{w_x}(A)$
4. if $\neg A\ x-$ is on b, then $0 \notin \bar{v}_{w_x}(A)$

Proof. An induction on the length of A. If A is atomic, then the result is (almost) immediate. (Just check the charts above.) The inductive steps for $\neg$ and $\vee$ are the same as in the proof of the Completeness Lemma for the non-serial $\mathsf{PD}\frac{3}{4}$ systems (i.e. Lemma 20). Thus I will prove only the inductive steps for implication ($\supset$), obligation (O), and permission (P).

Implication

1. Suppose $B \supset C\ x+$ is on b. Then, since b is complete, either $B\ x-$ or $C\ x+$ is on b. Thus, by the inductive hypothesis (IH), either $1 \notin \bar{v}_{w_x}(B)$ or $1 \in \bar{v}_{w_x}(C)$. Thus $1 \in \bar{v}_{w_x}(B \supset C)$.

2. Suppose $B \supset C\ x-$ is on b. Then, since b is complete, both $B\ x+$ and $C\ x-$ are on b. Thus, by the IH, $1 \in \bar{v}_{w_x}(B)$ and $1 \notin \bar{v}_{w_x}(C)$. Thus $1 \notin \bar{v}_{w_x}(B \supset C)$.

3. Suppose $\neg(B \supset C)\ x+$ is on b. Then, since b is complete, both $B\ x+$ and $\neg C\ x+$ are on b. Thus, by the IH, $1 \in \bar{v}_{w_x}(B)$ and $0 \in \bar{v}_{w_x}(C)$. Thus $0 \in \bar{v}_{w_x}(B \supset C)$.

4. Suppose $\neg(B \supset C)\ x-$ is on b. Then, since b is complete, either $B\ x-$ or $\neg C\ x-$ is on b. Thus, by the IH, either $1 \notin \bar{v}_{w_x}(B)$ or $0 \notin \bar{v}_{w_x}(C)$. Thus $0 \notin \bar{v}_{w_x}(B \supset C)$.

Obligation

1. Suppose $OB\ x+$ is on b. Then, since b is complete, $B\ y+$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $1 \in \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of the model induced by b (Definition **??**), $1 \in \bar{v}_{w_y}(B)$ for all w_y such that $w_x R w_y$. Thus $1 \in \bar{v}_{w_x}(OB)$.

2. Suppose $OB\ x-$ is on b. Then, since b is complete, $x \triangleright i$ and $B\ i-$ are on b. By the definition of the model induced by b, $w_x R w_i$. By the IH, $1 \notin \bar{v}_{w_i}(B)$. Thus it is not the case that for all $u \in W$, if $w_x R u$ then $1 \in \bar{v}_u(B)$. Thus $1 \notin \bar{v}_{w_x}(OB)$.

3. Suppose $\neg OB\ x+$ is on b. There are two cases:

   (a) $x = 0$. Then, since b is complete, $OB\ x-$ is on b. Thus, since b is complete, $x \triangleright i$ and $B\ i-$ are on b. By the definition of the model induced by b, $w_x R w_i$. By the IH, $1 \notin \bar{v}_{w_i}(B)$. Thus it is not the case that for all $u \in W$, if $w_x R u$ then $1 \in \bar{v}_u(B)$. Thus $1 \notin \bar{v}_{w_x}(OB)$. Thus, since $w_x = @$, $0 \in \bar{v}_{w_x}(OB)$.

   (b) $x \neq 0$. Then, since b is complete, $x \triangleright i$ and $\neg B\ i+$ are on b. Thus, by the definition of the model induced by b, $w_x R w_i$. By the IH, $0 \in \bar{v}_{w_i}(B)$. Thus there is a $u \in W$ such that $w_x R u$ and $0 \in \bar{v}_u(B)$. Thus, since $w_x \neq @$, $0 \in \bar{v}_{w_x}(OB)$.

4. Suppose $\neg OB\ x-$ is on b. There are two cases:

(a) $x = 0$. Then, since b is complete, $OB\ x+$ is on b. Thus, since b is complete, $B\ y+$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $1 \in \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of the model induced by b (Definition **??**), $\bar{v}_{w_y}(B)$ for all w_y such that $w_x R w_y$. Thus $1 \in \bar{v}_{w_x}(OB)$.Thus, since $w_x = @$, $0 \notin \bar{v}_{w_x}(OB)$.

(b) $x \neq 0$. Then, since b is complete, $\neg B\ y-$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $0 \notin \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of the model induced by b (Definition **??**), $0 \notin \bar{v}_{w_y}(B)$ for all w_y such that $w_x R w_y$. Thus $0 \notin \bar{v}_{w_x}(OB)$.

Permission

1. Suppose $PB\ x+$ is on b. Then, since b is complete, $x \triangleright i$ and $B\ i+$ are on b. Thus, by the definition of the model induced by b, $w_x R w_i$. By the IH, $1 \in \bar{v}_{w_i}(B)$. Thus there is a $u \in W$ such that $w_x R u$ and $1 \in \bar{v}_u(B)$. Thus $1 \in \bar{v}_{w_x}(PB)$.

2. Suppose $PB\ x-$ is on b. Then, since b is complete, $B\ y-$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $1 \notin \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of the model induced by b (Definition **??**), $1 \notin \bar{v}_{w_y}(B)$ for all u such that $w_x R u$. Thus $1 \notin \bar{v}_{w_x}(PB)$.

3. Suppose $\neg PB\ x+$ is on b. There are two cases:

   (a) $x = 0$. Then, since b is complete, $PB\ x-$ is on b. Thus, since b is complete, $B\ y-$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $1 \notin \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of the model induced by b (Definition **??**), $1 \notin \bar{v}_{w_y}(B)$ for all u such that $w_x R u$. Thus $1 \notin \bar{v}_{w_x}(PB)$.Thus, since $w_x = @$, $0 \in \bar{v}_{w_x}(PB)$.

   (b) $x \neq 0$.Then, since b is complete, $\neg B\ y+$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $0 \in \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Thus, by the definition of the model induced by b (Definition **??**), $0 \in \bar{v}_{w_y}(B)$ for all u such that $w_x R u$. Thus $0 \in \bar{v}_{w_x}(PB)$.

4. Suppose $\neg PB\ x-$ is on b. There are two cases:

   (a) $x = 0$. Then, since b is complete, $PB\ x+$ is on b. Thus, since b is complete, $x \triangleright i$ and $B\ i+$ are on b. Thus, by the definition of the model induced by

b, w_xRw_i. By the IH, $1 \in \bar{v}_{w_i}(B)$. Thus there is a $u \in W$ such that w_xRu and $1 \in \bar{v}_u(B)$. Thus $1 \in \bar{v}_{w_x}(PB)$.Thus, since $w_x = @$, $0 \notin \bar{v}_{w_x}(PB)$.

(b) $x \neq 0$. Then, since b is complete, $x \triangleright i$ and $\neg B\ i-$ are on b. Thus, by the definition of the model induced by b, w_xRw_i. By the IH, $0 \notin \bar{v}_{w_i}(B)$. Thus there is a $u \in W$ such that w_xRu and $0 \notin \bar{v}_u(B)$. Thus, since $w_x \neq @$, $0 \notin \bar{v}_{w_x}(PB)$.

■

Lemma 34 (completeness lemma for serial systems) Let $\mathsf{SPD}\frac{3}{4}+$ be a serial SPD system. Let b be an open, complete branch of an $\mathsf{SPD}\frac{3}{4}+$ tableau. Let $\langle W, @, R, v\rangle$ be the $\mathsf{SPD}\frac{3}{4}+$ model induced by b. Then for all A :

1. if $A\ x+$ is on b, then $1 \in \bar{v}_{w_x}(A)$
2. if $A\ x-$ is on b, then $1 \notin \bar{v}_{w_x}(A)$
3. if $\neg A\ x+$ is on b, then $0 \in \bar{v}_{w_x}(A)$
4. if $\neg A\ x-$ is on b, then $0 \notin \bar{v}_{w_x}(A)$

Proof. An induction on the length of A. If A is atomic, then the result is (almost) immediate. (Just check the charts above.) The inductive steps for $\neg$ and $\vee$ are the same as in the proof of the Completeness Lemma for the non-serial $\mathsf{PD}\frac{3}{4}$ systems (i.e. Lemma 20). Moreover, the inductive step for $\supset$ is the same as in the proof of the previous lemma. Thus I will prove only the inductive steps for obligation (O), and permission (P).

Obligation

1. Suppose $OB\ x+$ is on b. Then, since b is complete, $B\ y+$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the inductive hypothesis (IH), $1 \in \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Moreover, since b is complete, $[sO+]$ has been applied to it wherever possible; so, since by assumption $OB\ x+$ is on b, it cannot be the case that no node of the form $x \triangleright y$ is on b. Thus, by Definition 108, $\{w_i : w_xRw_i\} = \{w_i : x \triangleright i \text{ is on } b\}$. Thus, by Definition 108 again, $1 \in \bar{v}_u(B)$ for all u such that w_xRu. Thus $1 \in \bar{v}_{w_x}(OB)$.

2. Suppose $OB\ x-$ is on b. Then, since b is complete, $x \rhd i$ and $B\ i-$ are on b. By Definition 108, $w_x R w_i$. By the IH, $1 \notin \bar{v}_{w_i}(B)$. Thus it is not the case that for all $u \in W$, if $w_x R u$ then $1 \in \bar{v}_u(B)$. Thus $1 \notin \bar{v}_{w_x}(OB)$.

3. Suppose $\neg OB\ x+$ is on b. There are two cases:

   (a) $x = 0$. Then, since b is complete, $OB\ x-$ is on b. Thus, since b is complete, $x \rhd i$ and $B\ i-$ are on b. By Definition 108, $w_x R w_i$. By the IH, $1 \notin \bar{v}_{w_i}(B)$. Thus it is not the case that for all $u \in W$, if $w_x R u$ then $1 \in \bar{v}_u(B)$. Thus $1 \notin \bar{v}_{w_x}(OB)$. Thus, since $w_x = @$, $0 \in \bar{v}_{w_x}(OB)$.

   (b) $x \neq 0$. Then, since b is complete, $x \rhd i$ and $\neg B\ i+$ are on b. Thus, by Definition 108, $w_x R w_i$. By the IH, $0 \in \bar{v}_{w_i}(B)$. Thus there is a $u \in W$ such that $w_x R u$ and $0 \in \bar{v}_u(B)$. Thus, since $w_x \neq @$, $0 \in \bar{v}_{w_x}(OB)$.

4. Suppose $\neg OB\ x-$ is on b. There are two cases:

   (a) $x = 0$. Then, since b is complete, $OB\ x+$ is on b. Thus, since b is complete, $B\ y+$ is on b for all y such that $x \rhd y$ is on b. Thus, by the IH, $1 \in \bar{v}_{w_y}(B)$ for all y such that $x \rhd y$ is on b. Moreover, since b is complete, [$sO+$] has been applied to it wherever possible; so, since $OB\ x+$ is on b, it cannot be the case that no node of the form $x \rhd y$ is on b. Thus, by Def. 108, $\{w_i : w_x R w_i\} = \{w_i : x \rhd i \text{ is on } b\}$. Thus, by Def. 108, $1 \in \bar{v}_u(B)$ for all u such that $w_x R u$. Thus $1 \in \bar{v}_{w_x}(OB)$. Thus, since $w_x = @$, $0 \notin \bar{v}_{w_x}(OB)$.

   (b) $x \neq 0$. Then, since b is complete, $\neg B\ y-$ is on b for all y such that $x \rhd y$ is on b. Thus, by the IH, $0 \notin \bar{v}_{w_y}(B)$ for all y such that $x \rhd y$ is on b. Moreover, since b is complete, [$sO-$] has been applied to it wherever possible; so, since $\neg OB\ x-$ is on b, it cannot be the case that no node of the form $x \rhd y$ is on b. Thus, by Def. 108, $\{w_i : w_x R w_i\} = \{w_i : x \rhd i \text{ is on } b\}$. Thus, by Def. 108, $0 \notin \bar{v}_u(B)$ for all u such that $w_x R u$. Thus, since $w_x \neq @$ $0 \notin \bar{v}_{w_x}(OB)$.

Permission

1. Suppose $PB\ x+$ is on b. Then, since b is complete, $x \rhd i$ and $B\ i+$ are on b. By Definition 108, $w_x R w_i$. By the IH, $1 \in \bar{v}_{w_i}(B)$. Thus $1 \in \bar{v}_{w_x}(PB)$.

2. Suppose $PB\ x-$ is on b. Then, since b is complete, $B\ y-$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the inductive hypothesis (IH), $1 \notin \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Moreover, since b is complete, $[sP-]$ has been applied to it wherever possible; so, since by assumption $PB\ x-$ is on b, it cannot be the case that no node of the form $x \triangleright y$ is on b. Thus, by Definition 108, $\{w_i : w_x R w_i\} = \{w_i : x \triangleright i \text{ is on } b\}$. Thus, by Definition 108 again, $1 \notin \bar{v}_u(B)$ for all u such that $w_x R u$. Thus $1 \notin \bar{v}_{w_x}(PB)$.

3. Suppose $\neg PB\ x+$ is on b. There are two cases:

   (a) $x = 0$. Then, since b is complete, $PB\ x-$ is on b. Thus, since b is complete, $B\ y-$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $1 \notin \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Moreover, since b is complete, $[sP-]$ has been applied to it wherever possible; so, since $PB\ x-$ is on b, it cannot be the case that no node of the form $x \triangleright y$ is on b. Thus, by Def. 108, $\{w_i : w_x R w_i\} = \{w_i : x \triangleright i \text{ is on } b\}$. Thus, by Def. 108, $1 \notin \bar{v}_u(B)$ for all u such that $w_x R u$. Thus $1 \notin \bar{v}_{w_x}(PB)$. Thus, since $w_x = @$, $0 \in \bar{v}_{w_x}(PB)$.

   (b) $x \neq 0$. Then, since b is complete, $\neg B\ y+$ is on b for all y such that $x \triangleright y$ is on b. Thus, by the IH, $0 \in \bar{v}_{w_y}(B)$ for all y such that $x \triangleright y$ is on b. Moreover, since b is complete, $[sP+]$ has been applied to it wherever possible; so, since $\neg PB\ x+$ is on b (and $x \neq 0$), it cannot be the case that no node of the form $x \triangleright y$ is on b. Thus, by Def. 108, $\{w_i : w_x R w_i\} = \{w_i : x \triangleright i \text{ is on } b\}$. Thus, by Def. 108, $0 \in \bar{v}_u(B)$ for all u such that $w_x R u$. Thus, since $w_x \neq @$, $0 \in \bar{v}_{w_x}(PB)$.

4. Suppose $\neg PB\ x-$ is on b. There are two cases:

   (a) $x = 0$. Then, since b is complete, $PB\ x+$ is on b. Thus, since b is complete, $x \triangleright i$ and $B\ i+$ are on b. By Definition 108, $w_x R w_i$. By the IH, $1 \in \bar{v}_{w_i}(B)$. Thus $1 \in \bar{v}_{w_x}(PB)$. Thus, since $w_x = @$, $0 \notin \bar{v}_{w_x}(PB)$.

   (b) $x \neq 0$. Then, since b is complete, $x \triangleright i$ and $\neg B\ i-$ are on b. Thus, by Definition 108, $w_x R w_i$. By the IH, $0 \notin \bar{v}_{w_i}(B)$. Thus there is a $u \in W$ such that $w_x R u$ and $0 \notin \bar{v}_u(B)$. Thus, since $w_x \neq @$, $0 \notin \bar{v}_{w_x}(PB)$.

■

Fact 47 Each $\mathsf{SPD}\frac{3}{4}+$ tableau is finitely generated (see Def. 25, p. 75).

Proof. Obvious, since each application of a rule yields at most two immediate successors to any given node. ■

Lemma 35 If a node of the form $B\ x\pm$ occurs on an $\mathsf{SPD}\frac{3}{4}+$ tableau for $\Gamma \diagup A$, then B is a pseudo-subformula (see Def. 29, p. 76) of one of the elements of $\Gamma \cup \{A\}$.

Proof. A straightforward induction on the complexity of tableaus. Similar to the proof of Lemma 7, p. 76. ■

Fact 48 Each formula of $L_{\mathsf{PD}\frac{3}{4}+}$ has only finitely many pseudo-subformulas.

Proof. Obvious. ■

Remark 40 Note that $[O-]$, $[\neg O+]$, $[P+]$, $[\neg P-]$, $[sO+]$, $[sO-]$, $[sP+]$, and $[sP-]$ are the only index-generating rules in the $\mathsf{SPD}\frac{3}{4}$-systems.

Lemma 36 Let b be any branch of a $\mathsf{SPD}\frac{3}{4}+$ tableau. For all x, y, the node $x \triangleright y$ occurs on b only if nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b.

Proof. If $x \triangleright y$ occurs on b, then it can only have been introduced via an application of $[O-]$, $[\neg O+]$, $[P+]$, $[\neg P-]$, $[sO+]$, $[sO-]$, $[sP+]$, $[sP-]$, or $[h]$. Let us consider each case in turn:

- Case 1 ($[O-]$). Then nodes of the forms $OA\ x-$ and $A\ y-$ must occur on b.
- Case 2 ($[\neg O+]$). Then nodes of the forms $\neg OA\ x+$ and $\neg A\ y+$ must occur on b.
- Case 3 ($[P+]$). Then nodes of the forms $PA\ x+$ and $A\ y+$ must occur on b.
- Case 4 ($[\neg P-]$). Then nodes of the forms $\neg PA\ x-$ and $\neg A\ y-$ must occur on b.
- Case 5 ($[sO+]$). Then nodes of the forms $OA\ x+$ and $A\ y+$ must occur on b.
- Case 6 ($[sO-]$). Then nodes of the forms $\neg OA\ x-$ and $\neg A\ y-$ must occur on b.
- Case 7 ($[sP+]$).Then nodes of the forms $\neg PA\ x+$ and $\neg A\ y+$ must occur on b.

- Case 8 ($[sP-]$).Then nodes of the forms $PA\ x-$ and $A\ y-$ must occur on b.
- Case 9 ($[h]$). Then $x = y$ and a node of the form $z \rhd x$, where $z \neq x$, occurs on b. Since $z \neq x$, $z \rhd x$ must have been introduced via $[O-]$, $[\neg O+]$, $[s+]$, or $[s-]$. In any case, a node of the form $A\ y\pm$ must occur on b.

In each of the nine possible cases, nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b.

■

Definition 109 (OP-degree) The OP-degree of an $\mathsf{SPD}\frac{3}{4}+$ formula is simply the number of instances of O and P that occur in it. For example, the OP-degree of $OPO(OA \supset PA)$ is 5.

Lemma 37 (finiteness lemma) Every $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A is finite.

Proof. Suppose, for reductio, that some $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A is infinite. By König's Lemma, this tableau must have an infinite branch—say, b. There must be, on b, infinitely many nodes. By Lemma 30 (p. 182), all of these nodes must be different. If an index x occurs on b, then it can only occur in a node of the form $B\ x+$, $B\ x-$, $x \rhd y$, or $y \rhd x$. By Lemma 35, if $B\ x+$ or $B\ x-$ is on b, then B must be a pseudo-subformula of some element of $\Gamma \cup \{A\}$. By Fact 48, there are only finitely many of these. So there are only finitely many nodes of the form $B\ x\pm$ on b. Moreover, by Lemma 36, $x \rhd y$ occurs on b only if nodes of the forms $C\ x\pm$ and $D\ y\pm$ occur on b. We have already established that there are only finitely many nodes of the form $C\ x\pm$ or $D\ y\pm$ on b. Thus, since b is non-redundant, there can only be finitely many nodes of the form $x \rhd y$ on b. By parity of reasoning, there can only be finitely many nodes of the form $y \rhd x$ on b. Thus each index can occur on b only finitely many times. Thus, since b is infinite, infinitely many indices must occur on b. There are, then, two cases:

- **Case 1.** For some number, d, there are infinitely many indices on b of depth d. Let n be the smallest such number. The only index of depth 0 occurring on b is 0, and this occurs only finitely often. Hence $n > 0$. Further, the only way an index of depth $n > 0$ can be introduced to b is by an index-generating rule being applied to a node whose formula has an index of depth $n - 1$. And an index-generating rule can be applied to such a node only once. So if there are infinitely many indices of depth n on b, there must be infinitely many indices of depth $n - 1$ on b, contradicting our initial assumption about n.

- **Case 2.** For each number, d, there are only finitely many indices on b of depth d. Then, since infinitely many indices occur on b, infinitely many index depths must be represented on b. Now, suppose the index y occurs on b, where $y > 0$. This index must occur in a node of the form $C\ y\pm$, $y \triangleright z$, or $z \triangleright y$. But by Lemma 36, $y \triangleright z$, or $z \triangleright y$ occurs on b only if a node of the form $C\ y\pm$ occurs on b. So a node of the form $C\ y\pm$ must occur on b. Now, this node must have been added by applying either $[O-]$, $[\neg O+]$, $[P+]$, $[\neg P-]$, $[sO+]$, $[sO-]$, $[sP+]$, or $[sP-]$ to a node of the form $D\ x\pm$, where D is of higher OP-degree than C, or else it must follow by the non-deontic rules (i.e. $[\neg\neg]$, $[\vee+]$, $[\vee-]$, $[\neg\vee+]$, $[\neg\vee-]$, $[\supset+]$, $[\supset-]$, $[\neg\supset+]$, $[\neg\supset-]$) from such a formula. And non-deontic rules do not increase OP-degree. Thus the maximum OP-degree of all formulas on b with the index y is lower than the maximum OP-degree of all formulas on b with index x. (Intuitively, as the indices get "deeper," the OP-degrees of their associated formulas get smaller.) Thus since the maximum OP-degree of formulas with index 0 is finite, there is an index n such that no index of depth greater than n can have been introduced on b, contradicting our assumption that infinitely many index depths are represented on b.

Since both cases lead to a contradiction, we have shown that every $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A is finite. ■

Lemma 38 (extension lemma) Suppose there is no closed $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A. Then any $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A has an extension (also a $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A) containing an open, complete branch.

Proof. Suppose there is no closed $\mathsf{SPD}+$ tableau for Γ / A. Now take any tableau for Γ / A. We can use a procedure similar to the one given in the proof of Lemma 10 (p. 80) to construct an extension of this tableau (which extension is also a $\mathsf{SPD}+$ tableau for Γ / A) that contains an open, complete branch. Details of this procedure are straightforward but tedious, and are thus omitted. ■

Theorem 20 (completeness) If $\Gamma \models_{\mathsf{SPD}\frac{3}{4}+} A$, then $\Gamma \Vdash_{\mathsf{SPD}\frac{3}{4}+} A$.

Proof. Suppose $\Gamma \nVdash_{\mathsf{SPD}\frac{3}{4}+} A$. Then, by Definition 105 (p. 184), there is no closed $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A. Consider any $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A. By Lemma 38, this tableau has an extension (also a $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A) containing at least one open, complete branch. Let $\langle W, @, R, v\rangle$ be the $\mathsf{SPD}\frac{3}{4}+$ model induced by this branch.

By the relevant Completeness Lemma (either Lemma 33 or Lemma 34, depending on whether $\mathsf{SPD}\frac{3}{4}+$ is serial), $1 \in \bar{v}_{@}(B)$ for all $B \in \Gamma$ and $1 \notin \bar{v}_{@}(A)$. Thus $\Gamma \not\models_{\mathsf{SPD}\frac{3}{4}+} A$. ■

Theorem 21 ("all or nothing") If there is an open, complete $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A then there is no closed $\mathsf{SPD}\frac{3}{4}+$ tableau for Γ / A.

Proof. This follows from the Completeness Lemmas and the Soundness Theorem in the usual way. ■

Theorem 22 (decidability) The $\mathsf{SPD}\frac{3}{4}+$ systems are *decidable*; that is, there is an effective procedure which, when applied to any inference Γ / A, determines, in a finite number of steps, whether $\Gamma \vdash_{\mathsf{SPD}\frac{3}{4}+} A$.

Proof. One such decision procedure is as follows. Start with the initial list for Γ / A. Begin applying $\mathsf{SPD}\frac{3}{4}+$ tableau rules, in any order. By Lemma 37 the tableau will terminate after a finite number of steps. If it is closed, then $\Gamma \vdash_{\mathsf{SPD}\frac{3}{4}+} A$. If it is open, then by the "all-or-nothing" theorem, $\Gamma \nvdash_{\mathsf{SPD}\frac{3}{4}+} A$. ■

5.3 Notable features of the SPD systems

In this section I highlight some notable features of the SPD systems. Recall that I use '$\mathsf{SPD}\frac{3}{4}+$' as a variable ranging over these systems.

5.3.1 Relations between systems

Fact 49 The SPD systems are extensions of classical propositional logic (CPL). That is, if $\Gamma \vdash_{\mathsf{CPL}} A$ then $\Gamma \vdash_{\mathsf{SPD}\frac{3}{4}+} A$ (where $\Gamma \cup \{A\} \subseteq L_{\mathsf{CPL}} \cap L_{\mathsf{SPD}\frac{3}{4}+}$).

Proof. Suppose that $\Gamma \nvdash_{\mathsf{SPD}\frac{3}{4}+} A$. Then there is an $\mathsf{SPD}\frac{3}{4}+$ model $\langle W, @, R, v\rangle$ such that $\bar{v}_{@}(B) = \mathbf{t}$ for all $B \in \Gamma$ and $\bar{v}_{@}(A) \neq \mathbf{t}$. (Cf. Remark 34.) Define a CPL valuation, v_{CPL}, as follows. For all $p \in At$, let $v_{\mathsf{CPL}}(p) = v_{@}(p)$. A simple induction shows that for all $C \in L_{\mathsf{CPL}} \cap L_{\mathsf{SPD}\frac{3}{4}+}$, $\bar{v}_{\mathsf{CPL}}(C) = \mathbf{t}$ iff $\bar{v}_{@}(C) = \mathbf{t}$.[5] Thus $\bar{v}_{\mathsf{CPL}}(B) = \mathbf{t}$ for all $B \in \Gamma$ and $\bar{v}_{\mathsf{CPL}}(A) \neq \mathbf{t}$. Thus $\Gamma \nvdash_{\mathsf{CPL}} A$. ■

Fact 50 The following subsumption relations hold between the $\mathsf{SP}\frac{3}{4}\mathsf{D}$ systems:

[5] In case it is not obvious, $\bar{v}_{\mathsf{CPL}}$ is the result of extending v_{CPL} to all formulas in the usual (classical) way.

$$\begin{array}{ccc} \mathsf{SPD}\tfrac{3}{4} & \longrightarrow & \mathsf{SPD}\tfrac{3}{4}\mathsf{s} \\ \downarrow & & \downarrow \\ \mathsf{SPD}\tfrac{3}{4}\mathsf{h} & \longrightarrow & \mathsf{SPD}\tfrac{3}{4}\mathsf{sh} \end{array}$$

Fact 51 Each SPD3 system is a proper extension of its corresponding SPD4 system. (For example, SPD3s is a proper extension of SPD4s.)

Lemma 39 Every D (Ds, Dh, Dsh) model is an SPD3 (SPD3s, SPD3h, SPD3sh) model (for the language $L_{\mathsf{D}} \cap L_{\mathsf{SPD+}}$).

Proof. Let $\langle W, R, v\rangle$ by any D+ model. Let @ be any element of W. It is easy to check that $\langle W, @, R, v\rangle$ is an SPD3+ model. (Just observe that the semantic clauses of a SPD3 (SPD3s, SPD3h, SPD3sh) model agree with the clauses of a D (Ds, Dh, Dsh) model with respect to the classical truth values (**t** and **f**).) ■

Fact 52 Each of the following holds (where $\Gamma \cup \{A\} \subseteq L_{\mathsf{SPD}} \cap L_{\mathsf{D}}$):

- $\Gamma \vdash_{\mathsf{SPD3}} A \Rightarrow \Gamma \vdash_{\mathsf{D}} A$
- $\Gamma \vdash_{\mathsf{SPD3s}} A \Rightarrow \Gamma \vdash_{\mathsf{Ds}} A$
- $\Gamma \vdash_{\mathsf{SPD3h}} A \Rightarrow \Gamma \vdash_{\mathsf{Dh}} A$
- $\Gamma \vdash_{\mathsf{SPD3sh}} A \Rightarrow \Gamma \vdash_{\mathsf{Dsh}} A$

Proof. Follows easily from Lemma 39. ■

Fact 53 *None* of the following hold (where $\Gamma \cup \{A\} \subseteq L_{\mathsf{SPD}} \cap L_{\mathsf{PD}}$):

- $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}} A \Rightarrow \Gamma \vdash_{\mathsf{SPD}\frac{3}{4}} A$
- $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}\mathsf{s}} A \Rightarrow \Gamma \vdash_{\mathsf{SPD}\frac{3}{4}\mathsf{s}} A$
- $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}\mathsf{h}} A \Rightarrow \Gamma \vdash_{\mathsf{SPD}\frac{3}{4}\mathsf{h}} A$
- $\Gamma \vdash_{\mathsf{PD}\frac{3}{4}\mathsf{sh}} A \Rightarrow \Gamma \vdash_{\mathsf{SPD}\frac{3}{4}\mathsf{sh}} A$

Proof. One can check that $\neg OA \,/\, \neg O\neg(B \supset B)$ is valid in each of the PD systems but invalid in each of the SPD systems. ■

Remark 41 $\neg OA / \neg O \neg (B \supset B)$ is not an intuitively correct inference.[6] Thus the fact that it is valid in $\mathsf{PD}\frac{3}{4}+$ (as well as $\mathsf{D}+$) but not in $\mathsf{SPD}\frac{3}{4}+$ seems to count as a point in favor of the semi-paraconsistent systems.

Remark 42 An interesting and related fact is that $\neg(OA \wedge FA)$ is valid in $\mathsf{PD3s}+$ but not $\mathsf{SPD3s}+$. This too seems to count in favor of the semi-paraconsistent systems.

5.3.2 Basic logical properties

Fact 54 The SPD systems enjoy the basic "Tarskian" properties:

1. $A \in \Gamma \Rightarrow \Gamma \vdash_{\mathsf{D}+} A$ [reflexivity]
2. $\Gamma \vdash_{\mathsf{D}+} A \Rightarrow \Gamma, \Delta \vdash_{\mathsf{D}+} A$ [monotonicity]
3. $(\Gamma \vdash_{\mathsf{D}+} A$ and $A \vdash_{\mathsf{D}+} B) \Rightarrow \Gamma \vdash_{\mathsf{D}+} B$ [transitivity]

Remark 43 An anonymous referee once referred to the the SPD systems as being "non-monotonic." At first I was puzzled by this, given (2) above. I think what the referee was alluding to is the fact that deontic inheritance (if $\vdash A \supset B$ then $\vdash OA \supset OB$), sometimes referred to as the "monotonicity rule," fails in SPD. (See Fact 59 below.) In general, a unary function f is *monotonic* with respect to a relation R just in case whenever xRy, $f(x)Rf(y)$. (For example, the numerical successor function $(s(x) \mapsto x+1)$, is monotonic with respect to the greater-than relation $(>)$: whenever $x > y$, $s(x) > s(y)$.) Let f be the function that takes a formula A to the formula OA. Let R be the relation that holds between formulas A and B just in case $\vdash A \supset B$. Then in $\mathsf{SPD}\frac{3}{4}+$ it is *not* the case that whenever xRy, $f(x)Rf(y)$. So in *that* sense, the SPD systems are non-monotonic.

Theorem 23 (deduction theorem for $\supset$) $\Gamma, A \vdash_{\mathsf{SPD}\frac{3}{4}+} B$ iff $\Gamma \vdash_{\mathsf{SPD}\frac{3}{4}+} A \supset B$.

Proof. Straightforward. ■

Lemma 40 For all $\mathsf{SPD}\frac{3}{4}+$ models $\langle W, @, R, v\rangle$, $\bar{v}_@(A \supset B) = \bar{v}_@(A \to B)$.

[6] Consider a situation in which it is not obligatory that you twist $(\neg Ot)$, but it is both obligatory and forbidden that you shout (Os and $O\neg s$). By aggregation, it is obligatory that you both shout and not shout $(O(s \wedge \neg s))$. Thus (by the equivalence of $A \wedge \neg B$ and $\neg(A \supset B)$) it is obligatory that it not be the case that if you shout then you shout $(O\neg(s \supset s))$ Thus it is true that $\neg Ot$, but not true that $\neg O \neg (s \supset s)$.

Proof. By the Home World Lemma, either $\bar{v}_@(A \supset B) = \mathbf{t}$ or $\bar{v}_@(A \supset B) = \mathbf{f}$. If $\bar{v}_@(A \supset B) = \mathbf{t}$, then either $\bar{v}_@(A) = \mathbf{f}$ or $\bar{v}_@(B) = \mathbf{t}$, whence $\bar{v}_@(A \to B) = \mathbf{t}$. If $\bar{v}_@(A \supset B) = \mathbf{f}$, then $\bar{v}_@(A) = \mathbf{t}$ and $\bar{v}_@(B) = \mathbf{f}$, whence $\bar{v}_@(A \to B) = \mathbf{f}$. ■

Theorem 24 (deduction theorem for $\to$) $\Gamma, A \vdash_{\mathsf{SPD}\frac{3}{4}+} B$ iff $\Gamma \vdash_{\mathsf{SPD}\frac{3}{4}+} A \to B$.

Proof. Follows from Theorem 23 and Lemma 40. ■

Fact 55 In $\mathsf{SPD}\frac{3}{4}+$, $A \dashv\vdash B$ iff $\vdash A \equiv B$ iff $\vdash A \leftrightarrow B$.

Proof. Straightforward. ■

Fact 56 (failure of replacement) In $\mathsf{SPD}\frac{3}{4}+$ it is not the case that if $B \dashv\vdash B'$ then $A \dashv\vdash A[B/B']$.

Proof. Observe that in $\mathsf{SPD}\frac{3}{4}+$ we have $p \wedge \neg p \dashv\vdash q \wedge \neg q$ but $O(p \wedge \neg p) \nvdash O(q \wedge \neg q)$. ■

Fact 57 (failure of alt. replacement) In $\mathsf{SPD}\frac{3}{4}+$ it is not the case that if $\vdash B \leftrightarrow B'$ then $\vdash A \leftrightarrow A[B/B']$.

Proof. Immediate from Facts 55 and 56. ■

Remark 44 It makes perfect sense for the principle(s) of replacement to fail, given the intuitions motivating SPD. A and B are logically or semantically equivalent just in case they are true in exactly the same *possible* worlds.[7] It does not follow that they are true in exactly the same *permissible* worlds. Thus, from the fact that A and B are equivalent it does not follow that OA and OB are equivalent. For OA says that A is true at all permissible worlds, and OB says that B is true at all permissible worlds. The fact that A and B are true at the same *possible* worlds does not guarantee that they are true at the same *permissible* worlds.

5.3.3 Deontic inheritance

Fact 58 (failure of consequential inheritance) It is not the case that if $\Gamma \vdash_{\mathsf{SPD}\frac{3}{4}+} A$ then $O\Gamma \vdash_{\mathsf{SPD}\frac{3}{4}+} OA$.

[7] Note that in our semi-paraconsistent framework, @ is serving as a generic possible world—proxy for a set of such worlds. We only need one possible world to make the system work as intended. We would need a *set* of possible worlds if we wanted to introduce alethic modal operators ($\Box$ and $\Diamond$) into our system.

Proof. Just observe that $p, \neg p \vdash_{\mathsf{SPD}\frac{3}{4}+} q$ but $Op, O\neg p \nvdash_{\mathsf{SPD}\frac{3}{4}+} Oq$. ■

Fact 59 (failure of implicational inheritance) It is not the case that if $\vdash_{\mathsf{SPD}\frac{3}{4}+} A \supset B$ then $\vdash_{\mathsf{SPD}\frac{3}{4}+} OA \supset OB$.

Proof. Observe that $\vdash_{\mathsf{SPD}\frac{3}{4}+} p \wedge \neg p \supset q$ but $\nvdash_{\mathsf{SPD}\frac{3}{4}+} O(p \wedge \neg p) \supset Oq$. ■

Fact 60 (failure of strong implicational inheritance) It is not the case that if $\vdash_{\mathsf{SPD}\frac{3}{4}+} A \rightarrow B$ then $\vdash_{\mathsf{SPD}\frac{3}{4}+} OA \rightarrow OB$.

Proof. Immediate from Facts 40 and 59. ■

Fact 61 (failure of deontic equivalence) Neither of the following holds in $\mathsf{SPD}\frac{3}{4}+$:

1. If $\vdash_{\mathsf{SPD}\frac{3}{4}+} A \equiv B$ then $\vdash_{\mathsf{SPD}\frac{3}{4}+} OA \equiv OB$.

2. If $\vdash_{\mathsf{SPD}\frac{3}{4}+} A \leftrightarrow B$ then $\vdash_{\mathsf{SPD}\frac{3}{4}+} OA \leftrightarrow OB$.

Proof. Let $A = p \wedge \neg p$, $B = q \wedge \neg q$. ■

Remark 45 It makes sense for the inheritance principles to fail for essentially the same reason that it makes sense for the replacement principles to fail. (Cf. Remark 44.). Intuitively, A entails B iff every *possible* world at which A is true is a world at which B is true. It does not follow that every *permissible* world at which A is true is a world at which B is true. Thus from the fact that A entails B it does not follow that OA entails OB. After all, there could be *permissible* worlds at which A holds but B does not. Thus from the fact that A is true at all permissible worlds (OA), we cannot legitimately infer that B is true at all permissible worlds (OB).

5.3.4 Principles related to conflict-(in)tolerance

Here are some facts showing that the SPD systems are robustly conflict-tolerant:

1. $\nvdash_{\mathsf{SPD}\frac{3}{4}+} \neg(FA \wedge PA)$ [see Example 25, p. 186]

2. $\nvdash_{\mathsf{SPD}\frac{3}{4}+} \neg(OA \wedge FA)$

3. $\nvdash_{\mathsf{SPD}\frac{3}{4}+} \neg O(A \wedge \neg A)$

4. $FA, PA \nvdash_{\mathsf{SPD}\frac{3}{4}+} OB$

5. $OA, FA \nvdash_{\mathsf{SPD}\frac{3}{4}+} OB$ [see Example 24, p. 185]

6. $O(A \wedge \neg A) \nvdash_{\mathsf{SPD}\frac{3}{4}+} OB$

Thus, loosely speaking, in the SPD systems it is not a logical truth that normative conflicts don't occur; nor does the existence of a normative conflict entail that everything is obligatory (and, *a fortiori*, the existence of a normative conflict does not entail *everything*). The only other systems we have seen of which *all* of these things (i.e., 1-6 above) can be said are the PD4 systems. But those systems, unlike the SPD systems, reject some important and unobjectionable validities of classical logic, e.g. disjunctive syllogism ($A \vee B, \neg A \diagup B$) and the law of excluded middle ($A \vee \neg A$).

It should be noted that the SPD systems *do* validate some (superficially) similar inferences that might be considered "bad." For example:

1. $FA, \bar{P}A \vdash_{\mathsf{SPD}\frac{3}{4}+} B$

2. $\bar{F}A, PA \vdash_{\mathsf{SPD}\frac{3}{4}+} B$

3. $OA, \bar{F}A \vdash_{\mathsf{SPD}\frac{3}{4}\mathsf{s}+} B$

4. $OA, \bar{F}A \vdash_{\mathsf{SPD}\frac{3}{4}+} OB$

(1) and (2) are, of course, just instances of ordinary explosion ($C, \neg C \diagup D$), so they ought not to worry us. (3) and (4) cannot be explained away so easily, however. I will return to these schemas later in the chapter (in Sections 5.3.8 and 5.4.4).

5.3.5 Relations between permission and prohibition

Here are some facts about the relations between positive and negative permission and prohibition in the SPD systems:

1. $PA \nvdash_{\mathsf{SPD}\frac{3}{4}+} \bar{P}A$

2. $\bar{P}A \vdash_{\mathsf{SPD3}+} PA$

3. $\bar{P}A \nvdash_{\mathsf{SPD4}+} PA$

4. $FA \nvdash_{\mathsf{SPD}\frac{3}{4}+} \bar{F}A$

5. $\bar{F}A \vdash_{\mathsf{SPD3+}} FA$

6. $\bar{F}A \nvdash_{\mathsf{SPD4+}} FA$

7. $\vdash_{\mathsf{SPD3+}} FA \vee PA$ [see Example 30, p. 187]

8. $\nvdash_{\mathsf{SPD4+}} FA \vee PA$

9. $\nvdash_{\mathsf{SPD}\frac{3}{4}+} \bar{F}A \vee \bar{P}A$ [equivalent to $\neg(PA \wedge FA)$]

10. $\vdash_{\mathsf{SPD}\frac{3}{4}+} FA \vee \bar{P}A$ [an instance of excluded middle]

11. $\vdash_{\mathsf{SPD}\frac{3}{4}+} \bar{F}A \vee PA$ [an instance of excluded middle]

12. $\vdash_{\mathsf{SPD3+}} O(FA \vee PA)$

13. $\nvdash_{\mathsf{SPD4+}} O(FA \vee PA)$

It is quite appropriate for positive permission/prohibition not to entail negative permission/prohibition. For, as I have argued, there are (or at least can be) situations in which something (say, A) is both permitted and forbidden, in which case PA is true while $\neg FA$ is not (a counterexample to $PA \diagup \bar{P}A$), and FA is true but $\neg PA$ is not (counterexample to $FA \diagup \bar{F}A$).

What about the other direction? Should negative permission (prohibition) entail positive permission (prohibition)? One could argue that it should not: from the fact that there is no positive norm forbidding A it does not seem to follow that there is a positive norm permitting A. True, a normative system might include a meta-norm to this effect (e.g. "Anything that is not forbidden in this system should be understood to be permitted in this system"), but this does not seem like it should hold of *all* normative systems as a matter of *logic*. This gives us *some* reason to prefer the four-valued semi-paraconsistent systems ($\mathsf{SPD4+}$) over the three-valued systems ($\mathsf{SPD3+}$). However, it is important to note that, since $\{\neg FA, \neg PA\}$ is not a normative conflict (at least when permission is taken as primitive, as it is in the SPD systems), a deontic logic in which $\neg FA \wedge \neg PA$ cannot be true (and hence "explodes") is not necessarily conflict-intolerant.

Why do the four-valued systems invalidate $\bar{P}A \diagup PA$ and $\bar{F}A \diagup FA$, while the three-valued systems do not? Since the four-valued systems allow gaps at permissible worlds we can have situations like the following:

$$\begin{tikzcd} \boxed{@} \arrow[r] & \boxed{w_1} \arrow[loop above] \\ & p : \mathbf{n} \end{tikzcd}$$

In this model, $\neg Pp$ and $\neg Fp$ are both true at the home world, *precisely because* Pp and Fp are not. If we *don't* allow gaps, then $PA \vee FA$ is validated. For @ either accesses a world at which A holds, or it does not. If it does, then PA is true at @. If it doesn't, then A fails at every world seen by @, whence (since gaps are ruled out), $\neg A$ holds at every world seen by @, whence FA is true at @.

Not only do the four-valued semi-paraconsistent systems invalidate $\bar{P}A / PA$ and $\bar{F}A / FA$; they invalidate them *for the right reason*, namely, not everything is covered by a norm: some things are neither forbidden nor permitted (nor obligatory). If something (say, A) is not forbidden, then there must be a world (call it \$) at which $\neg A$ fails. If A is not permitted either, then there is no world at which A holds. Thus A fails at \$. Thus A is "gappy" at \$. Thus the possibility of something's being neither forbidden nor permitted provides independent motivation for allowing truth-value gaps at permissible worlds.

5.3.6 Deontic conjunction/disjunction principles

Here are some principles related to the interaction of the deontic operators with conjunction and disjunction. In $\mathsf{SPD}\frac{3}{4}+$ we have:

1. $OA, OB \vdash O(A \wedge B)$ [aggregation]

2. $P(A \vee B) \dashv\vdash PA \vee PB$ [principle of deontic distribution]

3. $\bar{P}(A \vee B) \dashv\vdash \bar{P}A \vee \bar{P}B$ [principle of deontic distribution for $\bar{P}$]

4. $OA \vdash O(A \vee B)$ [Ross's paradox 1]

5. $PA \vdash P(A \vee B)$ [Ross's paradox 2]

6. $\bar{P}A \vdash \bar{P}(A \vee B)$ [Ross's paradox 2 for $\bar{P}$]

7. $P(A \vee B) \nvdash PA \wedge PB$ [free choice permission]

8. $\bar{P}(A \vee B) \nvdash \bar{P}A \wedge \bar{P}B$ [free choice permission for $\bar{P}$]

9. $FA \vdash F(A \wedge B)$ [penitent's paradox]

10. $\bar{F}A \vdash \bar{F}(A \wedge B)$ [penitent's paradox for $\bar{F}$]

The SPD systems share *all* of these features with both the D systems and the PD systems. All of these features are natural outcomes of our interpretation of the deontic operators: OA is true at a world w iff A is true at *all* worlds permitted by w, and PA is true at w iff A is true at *some* worlds permitted by w. Once one accepts this interpretation, one is more or less forced to accept the principles above. Consider, for example, Ross's second paradox: $PA \vdash P(A \vee B)$. Suppose PA is true at a world, w. Then A is true at some world seen by w. But any world at which A is true is a world at which $A \vee B$ is true. Thus $A \vee B$ is true at some world seen by w. Thus $P(A \vee B)$.

Of course, if one finds any of the principles above counterintuitive, one might say that this provides reason for rejecting our current interpretation of the deontic operators. But I have already shown (see Section 2.6.5) that the alleged counterintuitiveness of things like Ross's "paradox", the penitent's "paradox" (if I am forbidden to lie, then I am forbidden to apologize for having lied), and (what we might call) the free choice permission "paradox" (*You can have soup or salad* does not entail *You can have soup and you can have salad*) can be explained away quite easily in terms of Grice's theory of conversational implicature [73]. Consider, for example, the so-called penitent's paradox. If I say that you are forbidden to lie and apologize for it, you might take that to imply that both the lying *and* the apologizing are forbidden. But this implication is defeasible:

> **Q:** Is it true that I am forbidden to lie and then apologize for it?
>
> **A:** Yes, but that is quite misleading, because it's only the *lying* part that's forbidden. If you *do* lie, then you are not only *not forbidden* to apologize for it; you are *obligated* to apologize for it.

Deontic disjunctive syllogism

Deontic disjunctive syllogism,

$$O(A \vee B), FA \vdash OB, \tag{DDS}$$

fails in the SPD systems, as it does in the PD systems. Some might see this as a liability in both families of systems, as DDS appears to be a valid inference. It is easy to see, however, that if inescapable normative conflicts are possible (which, as

I have argued, they are), then it is quite appropriate for DDS to fail. Consider, for example, a situation in which both Op and Fp hold, but Oq does not. Since Op holds in this situation, by Ross's (so-called) "paradox" we have it that $O(p \vee q)$ holds in the situation as well. Hence we have a counterexample to DDS.

A variation on deontic disjunctive syllogism,

$$F(A \wedge B), OA \vdash FB,$$

also fails in the SPD systems. Again, if inescapable normative conflicts are possible, then this is entirely appropriate. (Suppose Op and Fp hold, but Fq does not. Since Fp entails $F(p \wedge q)$, we have a counterexample.)

Interestingly, the following variation on DDS, in which positive prohibition is replaced with negative prohibition, is valid in each of the SPD systems:[8]

$$O(A \vee B), \bar{F}A \vdash OB. \tag{DDS$'$}$$

The reason is simple. The first premise, $O(A \vee B)$, says that $A \vee B$ is true at all permissible worlds. The second premise, $\bar{F}A$, says that A is not true at any permissible worlds. (Contrast this with FA, which says that A is false at all permissible worlds.) It follows that B must be true at all permissible worlds (since that's the only way $A \vee B$ could be true at those worlds). The fact that this inference is valid in the SPD systems softens the blow of having to give up DDS. Consider the following natural-language inference (cf Horty [84, p. 578]):

You ought to either fight in the army or perform alternative service.
You ought not to fight in the army.
$\therefore$ You ought to perform alternative service.

One could make the case that English words like 'forbidden', 'permitted', and even 'ought' are *ambiguous*, allowing both positive and negative interpretations. The inference above may be valid, but only when the second 'ought' is interpreted *negatively* (so that the statement is equivalent to 'You are not permitted to fight in the army'). The same point can be made about seemingly valid natural-language inferences such as:

[8]See Example 29, p. 187.

You ought not to drink and drive.
You ought to drive.
∴ You ought not to drink.

5.3.7 Principles related to conditional obligation

First, consider "Prior's paradox(es)" and some variations:

1. $OA \vdash_{\mathsf{SPD}\frac{3}{4}+} O(B \supset A)$ [Prior's paradox 1]

2. $OA \nvdash_{\mathsf{SPD}\frac{3}{4}+} O(B \rightarrow A)$ [Prior's paradox 1 for $\rightarrow$]

3. $FA \nvdash_{\mathsf{SPD}\frac{3}{4}+} O(A \supset B)$ [Prior's paradox 2]

4. $\bar{F}A \nvdash_{\mathsf{SPD}\frac{3}{4}+} O(A \supset B)$ [Prior's paradox 2 for $\bar{F}$]

5. $PA \vdash_{\mathsf{SPD}\frac{3}{4}+} P(B \supset A)$ [Prior's paradox 3]

6. $PA \nvdash_{\mathsf{SPD}\frac{3}{4}+} P(B \rightarrow A)$ [Prior's paradox 3 for $\rightarrow$]

7. $\bar{P}A \vdash_{\mathsf{SPD}\frac{3}{4}+} \bar{P}(B \supset A)$ [Prior's paradox 3 for $\bar{P}$]

Now, people tend to take one of two positions on Prior's so-called paradoxes. Either they regard them as counterintuitive and thus problematic for any system that validates them, or they believe they can be explained away, e.g. in terms of conversational implicature (much as I have explained away Ross's paradoxes). A curious fact about the SPD systems is that they validate *some* of the inferences in this family, but invalidate others. For example, both $OA/O(B \supset A)$ and $PA/P(B \supset A)$ are validated in the SPD systems, but $FA/O(A \supset B)$ is not. I want to argue that this is exactly how things *should* turn out, given our assumption that (inescapable) normative conflicts are possible. I start with the assumption that a conditional statement is true iff either its antecedent is untrue or its consequent is true. Now suppose we have a situation in which FA holds. Then A is *false* at all permissible worlds. Now, if truth-value gluts could not occur at permissible worlds, it would be legitimate to infer that A is *untrue* at all permissible worlds, whence $A \supset B$ is true at all permissible worlds, whence $O(A \supset B)$ holds. But gluts *can* occur at permissible worlds. For example, it might be true that OA as well as FA. In that case, A would be both true and false at all permissible worlds. This is compatible with B being untrue at some permissible worlds. Hence we cannot conclude that at every permissible world either A is untrue or B is true, i.e. we cannot conclude that $O(A \supset B)$ holds.

But let us set aside all talk of worlds for a moment. Intuitively, it is a good thing that $FA \diagup O(A \supset B)$ is invalidated. Consider, for example, the following natural language inference (which we encountered early in Chapter 2):[9]

Ed is forbidden to jaywalk.
$\therefore$ It ought to be that if Ed jaywalks, he gets the death penalty.

Even the most ardent defender of capital punishment would surely cringe at such a cruel and unusual inference!

What about the *other* version's of Prior's paradox—the ones that are *valid* in SPD+? Again, these are justifiable by appealing to the truth-condition for $A \supset B$. Suppose that OA is true. Then A is true at all permissible worlds. Thus, *a fortiori*, at all permissible worlds either B is untrue or A is true, i.e. $O(B \supset A)$ holds. A similar justification can be given for the validity of $PA \diagup P(B \supset A)$.

It is easy to see that, given the semantics of $\supset$, $OA \diagup O(B \supset A)$ and $PA \diagup P(B \supset A)$, unlike $FA \diagup O(A \supset B)$, are simply minor variations on Ross's paradoxes. As such, they should be regarded as correct, though potentially misleading, inferences. If the speaker knows that OA is true, then she would be acting in bad faith if she were to assert $O(B \supset A)$, though she would not be lying. The same goes for PA and $P(B \supset A)$.

Now let us consider some deontic-detachment-like principles. In $\mathsf{SPD}\frac{3}{4}+$ we have:

1. $O(A \supset B), OA \vdash OB$ [deontic detachment]
2. $O(A \supset B), PA \vdash PB$ [permissive deontic detachment]
3. $O(A \supset B), \bar{P}A \nvdash \bar{P}B$ [permissive deontic detachment for $\bar{P}$]
4. $O(A \supset B), FB \nvdash FA$ [deontic modus tollens]
5. $O(A \rightarrow B), FB \vdash FA$ [deontic modus tollens for $\rightarrow$]
6. $O(A \supset B), \bar{F}B \vdash \bar{F}A$ [deontic modus tollens for $\bar{F}$]
7. $O(A \supset B), A \nvdash OB$ [deviant deontic detachment 1]

[9] Cf. Mares [103, p. 5].

8. $A \supset OB, OA \nvdash OB$ [deviant deontic detachment 2]

9. $A \supset OB \nvdash O(A \supset B)$

Deontic detachment and permissive deontic detachment are pretty uncontroversial, I think. They are both valid in the SPD systems, as they are in the PD systems and D systems. It makes sense for permissive deontic detachment to fail for negative permission ($\bar{P}$). The first premise, $O(A \supset B)$, says that at every permissible world, either A fails or B holds. The second premise, $\bar{P}A$ ($=_{df}$ $\neg FA$), says that there is a permissible world at which $\neg A$ fails. What follows from these premises? Certainly not $\bar{P}B$ ($=_{df}$ $\neg FB$), which says that there is a permissible world at which $\neg P$ fails. Consider, for example, the following model:

$$\begin{tikzcd} \boxed{@} \arrow[r] & \boxed{w_1} \arrow[loop above] \\ & p : \mathbf{t} \\ & q : \mathbf{b} \end{tikzcd}$$

In this model $O(p \supset q)$ and Fq are true at @, while Fp is false. Hence (invoking the Home World Lemma), $O(p \supset q)$ and $\neg Fp$ ($=_{df}$ $\bar{P}p$) are true while $\neg Fq$ ($=_{df}$ $\bar{P}q$) is not true.

Deontic modus tollens ($O(A \supset B), FB / FA$) fails in SPD. It is easy to see why: just consider a situation in which Oq and Fq hold, but Fp does not. By one of Prior's *innocuous* "paradoxes," $O(p \supset q)$ holds in this situation. Hence we have a counterexample to the inference. Now, admittedly, we often regard natural language inferences such as the following as correct:

It ought to be that you drive only if you have a driver's license
You are forbidden to have a driver's license.
$\therefore$ You are forbidden to drive.

But once again we can appeal to the ambiguity of the deontic operators (between positive and negative interpretations) and/or the ambiguity of conditionals (between contraposible and non-contraposible interpretations) to account for the validity of the inference. As noted above, both $O(A \supset B), \bar{F}B / \bar{F}A$ and $O(A \to B), FB / FA$ are valid in the SPD systems.

As for the "deviant deontic detachment" principles, I have already argued (see Section 2.6.6) that neither of the elements of $\{O(A \supset B), A \supset OB\}$ should entail the other. Moreover, the SPD systems share these features with both the D systems and the PD systems (so a defender of one of *those* systems certainly can't complain here).

5.3.8 Some "add-on" deontic principles

There are two schemas that are commonly put forth as plausible deontic axioms: $OA \supset PA$ and $O(OA \supset A)$. $OA \supset PA$ is valid in our serial systems,[10] and $O(OA \supset A)$ (as well as its close relative, $O(A \supset PA)$) is valid in our shift-reflexive systems.[11] A side-effect of the validity of $O(OA \supset A)$ and $O(A \supset PA)$ is the validity of $OOA \supset OA$ and $PA \supset PPA$. (Suppose OOA holds. Then by $O(OA \supset A)$ and deontic detachment we can derive OA. Now suppose PA holds. Then by $O(A \supset PA)$ and permissive detachment[12] we have PPA.)

Intuitively, I see little reason to think that obligation entails permission. While a normative system could of course have an (explicit or implicit) meta-norm to the effect that it allows whatever it requires, it doesn't seem to be a *necessary truth* that all normative systems include such a meta-norm. The same goes for $O(OA \supset A)$. I conclude that there is little reason to require that the deontic accessibility be serial or shift-reflexive.

Note that the "ought implies negative permission" principle, $OA \supset \bar{P}A$, is invalid even in SPD3s. This is clearly a good thing, given that this schema is equivalent (in classical propositional logic, and thus in the SPD systems) to $\neg(OA \wedge FA)$, and thus would rule out inescapable normative conflicts. On the other hand, $O(A \supset \bar{P}A)$ (the "negative permission" version of $O(A \supset PA)$) *is* valid in $\mathsf{SPD}\frac{3}{4}\mathsf{h}$. This too seems appropriate. For $O(A \supset \bar{P}A)$, i.e. $O(A \supset \neg FA)$, says that every permissible world is such that if A holds, A is not forbidden. Assuming that every permissible world permits itself—an appropriate way of construing shift-reflexivity—this is quite natural. One may protest that even if every permissible world permits itself, there could be a permissible world at which a forbidden proposition holds, e.g. a permissible world at which $A \wedge FA$ holds. I agree, but, since such a world could be inconsistent, it does not follow that $\neg FA$ does not hold at that world. To put it another way, $P(A \wedge FA)$ is consistent with $O(A \supset \neg FA)$, though these do jointly entail $P(FA \wedge \neg FA)$.[13] But we have already allowed that contradictions can be permissible, so this

[10] See Example 26, p. 186.
[11] See Example 28, p. 187.
[12] I.e., the principle that $O(B \supset C), PB \vdash PC$.
[13] See Example 32, p. 188.

isn't particularly problematic.

5.4 Objections to the SPD systems

In this section I consider and respond to four potential objections to the SPD systems. In the subsections titled 'Objection' I will assume the point of view of the hypothetical objector.

5.4.1 Is the double standard *ad hoc*?

Objection

The "double standard" involved in the falsity conditions for the deontic operators is an *ad hoc* device, tacked on solely to avoid a problem. It has no independent motivation. Moreover, there is no principled reason for applying the double standard to the *falsity* conditions rather than the *truth* conditions. It would make just as much sense to formulate the clauses for O and P as follows:

$$
\begin{aligned}
1 \in \bar{v}_w(OA) \quad &\Leftrightarrow \quad (w = @ \text{ and } 0 \notin \bar{v}_w(OA)) \text{ or} \\
& \quad (w \neq @ \text{ and } \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A))) \\
0 \in \bar{v}_w(OA) \quad &\Leftrightarrow \quad \exists u(wRu \text{ and } 0 \in \bar{v}_u(A)) \\
1 \in \bar{v}_w(PA) \quad &\Leftrightarrow \quad (w = @ \text{ and } 0 \notin \bar{v}_w(PA)) \text{ or} \\
& \quad (w \neq @ \text{ and } \exists u(wRu \text{ and } 1 \in \bar{v}_u(A))) \\
0 \in \bar{v}_w(PA) \quad &\Leftrightarrow \quad \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A))
\end{aligned}
$$

In fact, it would seem that the most *principled* way of implementing the double standard would be to apply it to both the truth conditions *and* the falsity conditions for the deontic operators, like this:

$$
\begin{aligned}
1 \in \bar{v}_w(OA) \quad &\Leftrightarrow \quad (w = @ \text{ and } 0 \notin \bar{v}_w(OA)) \text{ or} \\
& \quad (w \neq @ \text{ and } \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A))) \\
0 \in \bar{v}_w(OA) \quad &\Leftrightarrow \quad (w = @ \text{ and } 1 \notin \bar{v}_w(OA)) \text{ or} \\
& \quad (w \neq @ \text{ and } \exists u(wRu \text{ and } 0 \in \bar{v}_u(A))) \\
1 \in \bar{v}_w(PA) \quad &\Leftrightarrow \quad (w = @ \text{ and } 1 \notin \bar{v}_w(PA)) \text{ or} \\
& \quad (w \neq @ \text{ and } \exists u(wRu \text{ and } 1 \in \bar{v}_u(A))) \\
0 \in \bar{v}_w(PA) \quad &\Leftrightarrow \quad (w = @ \text{ and } 1 \notin \bar{v}_w(PA)) \text{ or} \\
& \quad (w \neq @ \text{ and } \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A)))
\end{aligned}
$$

But on this semantics, OA and PA take *arbitrary* (albeit classical) truth values at the home world.[14] In other words, formulas of the forms OA and PA are treated as if they were atomic. This turns your semi-paraconsistent deontic logic into a mere "Rube Goldberg" version of classical propositional logic! Surely that was not your intent.

Response

There *is* independent motivation for the double standard. Given that the home world, @, is consistent and complete, negation behaves in a "Boolean" manner at the homeworld. That is, at the homeworld, $\neg A$ is true iff A is not true. Since $\neg A$ means that A is false, falsity inherits this Boolean property at the home world . Specifically, at the home world a formula is *false* just in case it is *not true.* At other worlds this is not (necessarily) the case. In light of this fact, the most principled way to formulate the semantic clauses for the SPD systems would be like this:

$$
\begin{array}{lcl}
\bar{v}_w(p) & = & v_w(p) \\
1 \in \bar{v}_w(\neg A) & \Leftrightarrow & 0 \in \bar{v}_w(A) \\
0 \in \bar{v}_w(\neg A) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(\neg A)) \text{ or} \\
 & & (w \neq @ \text{ and } 1 \in \bar{v}_w(A)) \\
1 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ or } 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \vee B) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(A \vee B)) \text{ or} \\
 & & (w \neq @ \text{ and } 0 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B)) \\
1 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \Rightarrow 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \supset B) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(A \supset B)) \text{ or} \\
 & & (w \neq @ \text{ and } 1 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B)) \\
1 \in \bar{v}_w(OA) & \Leftrightarrow & \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(OA) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(OA)) \text{ or} \\
 & & (w \neq @ \text{ and } \exists u(wRu \text{ and } 0 \in \bar{v}_u(A))) \\
1 \in \bar{v}_w(PA) & \Leftrightarrow & \exists u(wRu \text{ and } 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(PA) & \Leftrightarrow & (w = @ \text{ and } 1 \notin \bar{v}_w(PA)) \text{ or} \\
 & & (w \neq @ \text{ and } \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A)))
\end{array}
$$

But this is needlessly complex. Since the clauses for the extensional connectives ($\neg$, $\vee$, and $\supset$) do not "reach out" to other worlds (which are possibly inconsistent and/or

[14] Note that with these clauses $\bar{v}$ is no longer *determined* by v: for any given v, there are *many* functions $\bar{v}$ satisfying the semantic clauses.

incomplete), there is no risk that they (or rather, formulas of which they are the main connective) will take a non-classical truth value at the home world. Thus there is no need to make the double standard explicit for $\neg$, $\vee$, and $\supset$.

Not only is there independent motivation for the double standard, there is a principled reason for applying the double standard to the falsity conditions rather than the truth conditions. If the two double standards were equivalent, it would merely be a matter of convention which one we chose. They are not equivalent, however. Consider, for example, the following model:

$$\begin{tikzcd} \boxed{@} \arrow[r] & \boxed{w_1} \arrow[loop above] \\ & p : \mathbf{b} \end{tikzcd}$$

If the double standard is applied to the *falsity* condition for O (as it is in our "official" SPD systems), then Op is assigned $\mathbf{t}$ at @. If the double standard is applied to the *truth* condition for O, then Op is assigned $\mathbf{f}$ at @. Given that the two double standards are not equivalent, there can be a principled reason for choosing one over the other. And indeed, there is such a reason. We have seen that both $\neg(FA \wedge PA)$ and $\neg(OA \wedge FA)$ are invalid in *all* of our SPD systems. This is clearly a good thing, given our assumption that normative conflicts (of both types) are possible. If the double standard were applied to the *truth* conditions for the deontic operators, $\neg(FA \wedge PA)$ would be valid in (the newly defined) SPD3, and $\neg(OA \wedge FA)$ would be valid in (the newly defined) SPD3s.[15]

Of course, there is no free lunch: as one would expect, there are schemas that are valid in the SPD systems that would be rendered invalid if the double standard were "moved" to the truth conditions. An example is $PA \vee P\neg A$.[16] This schema is valid in SPD3s+ but would be invalid if we used the alternative double standard. But unlike $\neg(OA \wedge FA)$ and $\neg(FA \wedge PA)$, $PA \vee P\neg A$ is not clearly a "bad guy." In fact, many people view it as a "good guy."[17] So it is proper (or at least preferable) to apply the double standard to the *falsity* conditions for O and P.

[15] I am grateful to Geoffrey Hellman for suggesting the line of argument advanced in this paragraph.

[16] See Example 31, p. 188.

[17] For example, during the Q-and-A session of a talk I gave at a conference, a fairly well-known philosopher/logician referred to systems of deontic logic that invalidate $PA \vee P\neg A$ as being overly "harsh." It's easy to see why s/he said this. One can imagine a person saying "What do you mean I'm not permitted to speak and I'm not permitted to refrain from speaking? I've got not choice but to do one or the other! Boy, that's *harsh*."

It should be acknowledged, however, that this defense of applying the double standard to falsity conditions, rather than truth conditions, for the deontic operators is weaker than it might be. The reason for this is that the points made about the schemas $\neg(FA \wedge PA)$, $\neg(OA \wedge FA)$, and $PA \vee P\neg A$ in the preceding paragraphs apply only to the 3-valued ($\mathsf{SPD3+}$) systems, whereas I have argued (in Section 5.3.5 above) that there is some philosophical motivation for rejecting these in favor of the weaker 4-valued ($\mathsf{SPD4+}$) systems. Nevertheless, I have at least shown that the decision to apply the double standard to the falsity conditions rather than the truth conditions has *some* reasonable justification, since it makes an advantageous difference in at least *some* of the semi-paraconsistent deontic logics that we have been considering.

5.4.2 A slippery slope to dialetheism?

Objection

If we can stipulate what is *obligatory* (*permissible*, etc.), we must also be able to stipulate what is *not obligatory* (not permissible, etc.). (Indeed, if we can stipulate what is permissible, and 'permissible' just means 'not obligatory that not', it follows that we can stipulate what is not obligatory.) And if we can stipulate what is not obligatory, then we ought to allow for the possibility not just of normative conflicts, but of full-blown deontic dialetheias (e.g. of the form $OA \wedge \neg OA$) at the home world. We may still hold that non-deontic formulas (i.e. ones containing no deontic operators) are always consistent at the home world. If we limit our dialetheism to deontic formulas in the way being suggested (cf. the discussion of "institutional dialetheism" in Section 4.3.4), we can preserve, in qualified forms, certain inferences that would otherwise be categorically invalid. For example, disjunctive syllogism ($A \vee B, \neg A \,/\, B$) would be valid provided that A and B contain no deontic operators.

Response

Let me respond to this argument with an analogy. Making it obligatory that A is like placing A into an "obligation box"—a big box of propositions that magically closes its contents under (some variety of) logical consequence. A proposition A is obligatory just in case it is in the box. Clearly it is possible for both A and $\neg A$ to be in the box. (In the simplest case, we just put them directly in; in other cases, they appear in more complicated ways as a result of the box's magical closure property.)

However, it is not possible for A to be both in the box and not in the box. Thus it is not possible for A to be both obligatory and non-obligatory.

Similarly, making it permissible that A is like placing A into a "permission box." The permission box has a different closure property than the obligation box. While the obligation box closes its contents (taken as a whole) under logical consequence, the permission box merely closes each *individual* proposition in the box under logical consequence. For example, if A is in the box, then $\neg\neg A$ and $A \vee B$ are in the box. If A and B are in the box, it is not necessarily the case that $A \wedge B$ is in the box (since $A \wedge B$ doesn't follow from A alone, or from B alone). A given proposition, A, is either in the box or it is not. If it is, then PA is true and $\neg PA$ is not. If it is not, then $\neg PA$ is true and PA is untrue. Thus it is impossible for both PA and $\neg PA$ to be true.

Incidentally, if one holds that 'ought' implies 'may', one might picture the obligation box as being *inside* the permission box, so that everything that is in the former is also in the latter, whence OA entails PA. If one holds that $O(OA \supset A)$ is a truth of deontic logic, then one might imagine that every instance of $OA \supset A$ is in the obligation box.

All this talk of "obligation boxes" and "permission boxes" is somewhat fanciful, of course, but I believe that it provides us with a reasonable way of thinking about norms (and normative propositions) according to which normative conflicts (e.g. $OA \wedge O\neg A$) are possible but deontic dialethieas (e.g. $OA \wedge \neg OA$) are not. In other words, it gives us some traction that keeps us from sliding down the slipperly slope to full-blown dialetheism.

5.4.3 Another slippery slope argument

Objection

In the SPD systems certain laws of logic—e.g. explosion, disjunctive syllogism, non-contradiction, excluded middle—are allowed to "fail" at worlds other than the home world. The argument for this is that it can be, e.g., obligatory that A is both true and false, or permitted that A is neither true nor false. Why stop there? If *some* laws of logic can fail at permissible worlds, why couldn't *any* (or *all*) of them fail? For example, why couldn't there be permissible worlds at which *modus ponens*, or *conjunction elimination*, fails? If we allow such worlds, then inferences like $O(A \supset$

$B), OA/OB$ and $O(A \wedge B)/OA$ are invalidated. Indeed, we would have *no* valid inferences involving deontic operators essentially. In effect, this would be the death of deontic logic. There does not seem to be any justification for letting only *some* laws of logic fail at worlds other than the home world. Hence, either we shouldn't let any of them fail (in which case we should reject the SPD systems), or we should let *all* of them fail, in which case there is no deontic logic.

Response

In responding to this objection, I need to say what I think a deontic logic ought to do. To do this, I need to say what a logic is. Logics are tools. They are not right or wrong, true or false, correct or incorrect; they are only more or less *useful*. Just as a hammer is a tool that is used to pound nails, a logic is a tool that is used to regiment and evaluate natural langauge inferences. A well-thought-out formal logical system can resolve questions about whether certain natural language inferences are valid. Translate the inference into the formal language, then check whether it is valid. In my view, that is what logics are for.

So a deontic logic ought to capture what we take to be correct deontic reasoning in natural language. Clearly some deontic inferences are intuitively valid, and some are not. (Consider, e.g., the inferences listed at the beginning of Chapter 2. I think that most people would agree, for the most part, on which of these inferences are valid and which are not.) This seems to indicate that in our intuitive semantics for deontic language (if there is such a thing), we *do not* believe that all laws of logic could fail at permissible worlds. I believe that the SPD systems do a good job of capturing our intutions about deontic inferences—of accounting for the data, so to speak. They do this by expanding our notion of what a permissible world is, by allowing inconsistency and incompleteness at such worlds. Why not allow, e.g., modus ponens to fail at permissible worlds? Just because deontic detachment ($O(A \supset B), OA/OB$) is intuitively valid. That is reason enough. I would have no problem with allowing any or all of the laws of logic to fail at permissible worlds if it better captured our intuitions. But it would not. Instead, it would render deontic logic useless (or nonexistent, depending on how one looks at it).

5.4.4 What about Boolean normative conflicts?

Objection

It was arbitrary to exclude Boolean negation from the SPD systems. The excuse given was that Boolean and De Morgan negation "collapse" at the home world. While this is true, they do not collapse at other worlds. (If we add #, with the usual semantics, to the SPD systems we find that $O(\neg A \supset \# A)$ fails in all of the systems, and $O(\# A \supset \neg A)$ is valid only in the three-valued systems.) So there is no good reason to exclude Boolean negation from the language of the SPD systems. And if we do include it, then (as in the PD systems) we find that certain counterintuitive inferences like $OA, O\#A \,/\, OB$ are valid. If, as you have argued, we can (under certain circumstances) *stipulate* conflicts like $\{OA, O\neg A\}$ into existence, then surely we can stipulate conflicts like $\{OA, O\#A\}$ into existence as well. And if it is counterintuitive for $\{OA, O\neg A\}$ to "explode," then surely it is just as counterintuitive for $\{OA, O\#A\}$ to explode. Hence SPD tolerates only *some* kinds of normative conflicts. It can't handle what we might call *Boolean* normative conflicts. Boolean negation throws a monkey wrench into the works. You can't legitimately evade this problem by pretending that Boolean negation doesn't exist.

Response

This, I must confess, is a very powerful objection. Despite my affinity for the semi-paraconsistent deontic logics I have created, I fear that the objector may be right. Nevertheless, I want to suggest two possible responses to the objection. Unfortunately, I cannot bring myself to wholeheartedly endorse either of them.

The first response is to claim that while it is possible to stipulate that something ought to be *true*, or ought to be *false*, it is impossible to stipulate that something ought to be *untrue*, or ought to be *unfalse*. On this account, normative conflicts of the form $\{OA, O\#A\}$ are truly impossible, so it does not matter if they explode. We can embellish the obligation/permission box analogy as follows. Propositions involving Boolean negation cannot be put into the obligation box or the permission box; they are rejected. Trying to put, say, $\#A$ into the obligation box is like trying to put a Canadian quarter into an American vending machine: just as the quarter goes straight into the coin return, the proposition goes straight into the "proposition return". Just as we should not take the rejection of the Canadian quarter as implying that the coin is counterfeit or not legal tender, we should not take the rejection of the $\#A$ as implying that it is not a genuine, meaningful proposition. It is just that it is not the type of proposition that can be placed in the obligation box. Why? Perhaps because the designer of the boxes saw that they could *explode* if Boolean propositions were allowed to be put into them, and hence took precautions to ensure that this could never happen.

This response strikes me as rather *ad hoc*. I cannot think of any *independent* motivation for holding that Boolean normative conflicts are impossible. I believe that, under certain circumstances, social beings (e.g. the co-dictators Tweedledee and Tweedledum of Section 1.3.1) can stipulate norms into existence. Since social beings are prone to error, sometimes these norms conflict. While it might never *occur* to a legislator or other norm-maker to promulgate a norm that clearly and explicitly involves Boolean negation (e.g., "It ought to be untrue that anyone under twenty-one buys alcohol"), it does seems *possible*. Therefore, I cannot bring myself to endorse this response.

The second response is more interesting, more sophisticated, and more plausible. According to this response, there is *no such thing* as Boolean negation. Graham Priest, for example, has argued that Boolean negation is *incoherent*—that there simply is no concept of Boolean negation to be expressed. Now, Priest's motivation for denying the existence of Boolean negation is somewhat different from mine. (As he rightly acknowledges, his proposed dialetheic solution to the semantic paradoxes is undermined if the notion of Boolean negation is coherent.[18]) But if he is right, then we can simply deny the existence of Boolean normative conflicts; for if there is no such thing as Boolean negation, then clearly there is no such thing as a Boolean normative conflict. And if there is no such thing as a Boolean normative conflict, then the fact that the SPD systems cannot deal with them is no fault. In Chapter 5 of *Doubt Truth to Be a Liar* [130], and in his earlier paper "Boolean Negation and All That" [125], Priest argues that (a) the existence of a coherent notion of Boolean negation contradicts an intuitive principle about information which he calls the Augmentation Constraint; and (b) the standard semantic argument for the existence of a coherent notion of Boolean negation begs the question (i.e., assumes precisely what it sets out to prove). Let's consider Priest's arguments for these claims.

[18]The reason is that if Boolean negation is coherent, then we can formulate a liar sentence involving it and apply the unrestricted T-schema to derive a "Boolean" contradiction ($A \wedge \#A$) which explodes into triviality.

First, a point of clarification. Thus far I have been using the term 'Boolean negation' to denote a negation connective, #, such that $\#A$ is true iff A is untrue. It is not clear to me that Priest denies the existence or coherence of such a connective. What he *does* deny is that there is an *explosive* negation connective, i.e. a negation connective, #, such that $\{A, \#A\}$ logically implies everything. Of course, these two features are closely related: it is usually argued that Boolean negation explodes precisely *because* of its truth condition. In what follows, I will use the term 'Boolean negation' to mean a negation connective with *both* of the aforementioned features (i.e., a negation connective, #, such that $\#A$ is true iff A is untrue, *and* $\{A, \#A\}$ entails B). (Note: I take *ex falso quodlibet* ($\#A \vdash A \supset B$) to be a trivial variant on explosion ($\#A, A \vdash B$).)

The argument from augmentation (as I will call it) goes like this. Worlds can and should be thought of as *information states*. Information states are closed under logical consequence: if the elements of Γ hold in an information state and $\Gamma \vdash A$, then A holds in that information state. Moreover, an information state can be *incomplete*; that is, if w is an information state, then we can have neither A nor $\neg A$ holding at w. Now, suppose that # expresses Boolean negation (as characterized in the last paragraph). Consider an information state, w, in which neither A nor $\neg A$ holds. Then, by the truth condition for #, both $\#A$ and $\#\neg A$ hold in w. Suppose we add a single piece of information, A, to w, leaving *everything* else about w (including the fact that $\#A$ holds in w) untouched. Then both A and $\#A$ hold at w, and w is rendered trivial. An intuitive principle, which Priest calls the *Augmentation Contraint*, has it that if an information state w is incomplete with respect to a piece of information, A, then there is a state w' that is exactly like w except that it includes A. If Boolean negation is coherent, then the Augmentation Constraint is incorrect. But the Augmentation Constraint is correct. Therefore, Boolean negation is incoherent.

Now, as Priest acknowledges, one can evade the foregoing argument by simply denying the Augmentation Constraint. I, for one, am agnostic about the Augmentation Constraint. I find it somewhat plausible, but not obviously true. But the argument does show that, as Priest puts it, "the meaningfulness of Boolean negation is not to be taken for granted" [130, p. 90]. Accordingly, we need to ask what arguments can be given for the claim that there is a negation with the two properties I have associated with the term 'Boolean'. One such argument runs as follows:

1. A is a semantic consequence of Γ (in symbols, $\Gamma \models A$) iff for every interpreta-

tion[19] $\mathcal{I}$, if all of the elements of Γ are true in $\mathcal{I}$ then A is true in $\mathcal{I}$.

2. For any $\mathcal{I}$ and any A, not all of the elements of $\{A, \#A\}$ are true in $\mathcal{I}$.

3. Therefore, for any $\mathcal{I}$ and any A, if all of the elements of $\{A, \#A\}$ are true in $\mathcal{I}$, then B is true in $\mathcal{I}$. [from 2]

4. Therefore, $A, \#A \models B$. [from 1 and 3]

Priest points out that step 3 is illegitmate unless the 'not' in premise 2 is understood as *Boolean* negation (since De Morgan negation doesn't satisfy *ex falso quodlibet*). But the coherence of Boolean negation is precisely what the argument is supposed to establish. Priest concludes that the argument is viciously circular.

A fairly obvious response to Priest's argument is that semantic consequence can be (and often is) defined in a slightly different way, namely:

> $\Gamma \models A$ iff there is no interpretation $\mathcal{I}$ such that all of the elements of Γ are true in $\mathcal{I}$ and A is not true in $\mathcal{I}$.

If we use this as our definition, we need only assume that the 'not' in premise 2 is such that $\ulcorner$not $A\urcorner$ entails $\ulcorner$not both A and $B\urcorner$. Both Boolean and De Morgan negation satisfy this condition. Priest anticipates this response, and points out that the adequacy of the alternative definition requires justification. In particular, we need to make sure that it licenses us to conclude that A is true when we know that the elements of Γ are true, and that $\Gamma \models A$. The justification for this would look something like this:

> Suppose that all of the elements of Γ are true (in an interpretation, $\mathcal{I}$), and that $\Gamma \models A$. Then, by the definition of $\models$, it is not the case that both (i) all of the elements of Γ are true in $\mathcal{I}$ and (ii) A is not true in $\mathcal{I}$). Thus, since all of the elements of Γ are true in $\mathcal{I}$, A is true in $\mathcal{I}$.

The reasoning used here is of the form:

$$\text{Not both } A \text{ and } B;\ A \therefore \text{Not } B$$

[19] Here an interpretation could be a valuation, model, or what-have-you.

But this inference is simply a variant on disjunctive syllogism, and as such is valid only for Boolean negation, not De Morgan negation! So the justification of the definition relies on the assumption that the negation in question is Boolean, which again, is precisely what's at issue.

What are we to make of Priest's arguments? What they show, I think, is that it is not so obvious that Boolean negation is a meaningful concept, and that the usual argument for its meaningfulness begs the question. But neither of these points establish or imply that there is no such thing as Boolean negation. Priest's second argument is reminiscent of Quine's famous argument that one cannot derive the rules of logic from linguistic conventions without *using* those very rules.[20] For example, to derive an instance of modus ponens from the stipulated truth condition for $\ulcorner$if A then $B\urcorner$, one must *use* another instance of modus ponens, which in turn needs to be justified, and so on, ad infinitum, resulting in a vicious infinite regress. Quine did not take his argument to show that there is *no such thing* as a conditional connective satisfying modus ponens! Rather, he took it to show that logical truth cannot be the product of linguistic convention alone. Perhaps we should draw a similar conclusion here. Priest's second argument simply shows that we cannot derive explosion from the truth condition for Boolean negation without *using* explosion. It does not show that there is no negation connective satisfying explosion.

But suppose, for the moment, that I were to take the position that there is no such thing as Boolean negation. This would open the way for me to claim that the SPD systems are robustly conflict-tolerant (with respect to all types of normative conflict). But it also might just undermine the whole motivation for deviating from standard deontic logic in the first place. For if Boolean negation is incoherent, then "bad guys" like $FA, PA/B$ and $OA, FA/OB$ do not appear to be valid in the semantics of standard deontic logic! After all, the arguments purporting to establish their validity depend on the assumption that Boolean negation is coherent in just the same way that the argument for explosion depends on it. Thus it may turn out that standard deontic logic is conflict-tolerant after all! By the same token, it may turn out that (strange and paradoxical as it sounds) classical logic is actually paraconsistent! (More specifically, its *semantics* is paraconsistent. The proof theories that have been given for classical logic are therefore unsound, as they license deriving B from $\{A, \neg A\}$.)

[20] See his "Truth by Convention" [134], which was influenced by Lewis Carroll's "What the Tortoise Said to Achilles" [41].

Let us assume, then, that there really is such a thing as Boolean negation. Then, unfortunately, I cannot claim that the SPD systems are tolerant of all types of normative conflicts. What I *will* claim is that the SPD systems (especially SPD4) deal with normative conflicts more effectively than any other system that has been proposed to date. If I am right about that, then I have accomplished something significant in this dissertation.

One final point should be noted before concluding the chapter. During my Ph.D. final oral exam (November 30, 2007), Lou Goble brought to my attention the very interesting and somewhat disturbing fact that even though my SPD systems do not contain a connective expressing Boolean negation, a formula *equivalent* to $O\#A$ can be expressed in them, namely $\neg PA$. The equivalence can be verified as follows:

$$\begin{array}{rll} 1 \in \bar{v}_{@}(\neg PA) & \text{iff} & 0 \in \bar{v}_{@}(PA) \\ & \text{iff} & 1 \notin \bar{v}_{@}(PA) \\ & \text{iff} & \text{not: } (\exists u(@Ru \text{ and } 1 \in \bar{v}_u(A))) \\ & \text{iff} & \forall u(@Ru \Rightarrow 1 \notin \bar{v}_u(A)) \\ & \text{iff} & \forall u(@Ru \Rightarrow 1 \in \bar{v}_u(\#A)) \\ & \text{iff} & 1 \in \bar{v}_{@}(O\#A) \end{array}$$

Note that this equivalence is precisely why, as noted in Section cite, the set $\{OA, \bar{F}A\}$ deontically explodes in all of the SPD systems (and just plain explodes in the serial SPD systems). What are we to make of this equivalence? Using the obligation/permission box analogy (see Section cite), I have tried to argue (or at least provide a way of making sense of the suggestion) that so-called negative norms ($\neg Op, \neg Fp, \neg Pp$, etc.), unlike positive norms (Op, Fp, Pp, etc.), cannot be stipulated into existence. The equivalence of $\neg PA$ and $O\#A$ seems to suggest that taking P as primitive does not, as it first appeared, allow us to draw a clear conceptual distinction between positive and negative norms in the SPD systems: for certain negative norms are (at least if Boolean negation is legitimate), *equivalent to* certain positive norms. My tentative conclusion is this. Assuming that Boolean negation is meaningful, the fact that $\{OA, \bar{F}A\}$ deontically explodes in the SPD systems is exactly as worrisome as the fact that $\{OA, O\#A\}$ would deontically explode in the SPD systems (that is, if it were expressible). I have already conceded that the latter fact is a significant problem for semi-paraconsistent deontic logic (as I have characterized it). Thus I am forced to concede that the former fact is also a significant problem.

Chapter 6

Conclusion

> If a man starts with certain assumptions, he may be a good logician and a good citizen, a wise man, a successful figure. If he starts with certain other assumptions, he may be an equally good logician and a bankrupt, a criminal, a raving lunatic.
>
> —G.K. Chesterton[1]

> Jen had always believed philosophers ought to have their heads knocked repeatedly, lest they become trapped in the rhythms of their own if-thens.
>
> —David Brin, *Earth*[2]

In Section 1.5, I pointed out that if one wishes to accept the possibility of normative conflicts, then one must give up some appropriate combination of the following principles:

1. *Interdefinability.* FA is definable as not-PA, and PA is definable as not-FA.
2. *Aggregation.* If two propositions are individually obligatory, then they are jointly obligatory.
3. *Ought implies may.* If something is obligatory, then it is permissible.
4. *Ought implies can.* If something is obligatory, then it is possible.
5. *Deontic inheritance.* If A entails B, then OA entails OB.

[1] *Daily News* of London, Feb. 25, 1905.

[2] New York: Bantam, 1990, p. 516.

6. *Explosion.* Everything follows from a contradiction.

Given the developments of previous chapters, I believe we should give up *Interdefinability*, *Ought implies can*, and *Deontic inheritance*. (I also believe that we should give up *Ought implies may*, but this is optional, not required.) As I have noted, there are reasons for giving these up that are independent of the need to accommodate normative conflicts. I believe I have given additional support for rejecting them by developing some attractive (if imperfect) systems of semi-paraconsistent deontic logic in which they fail. Conversely, the fact that (2) and (6) hold in these systems gives us some (additional) reason to accept these principles. Let's consider each of the principles (1)-(6) in turn.

Interdefinability. In the SPD systems, FA is not definable as $\neg PA$; nor is PA definable as $\neg FA$. This is quite deliberate. As explained in Section 5.1.3, there is good reason to take P as primitive: if we define it as $\neg F$, then escapable conflicts of the form $FA \wedge PA$ become flat-out contradictions of the form $FA \wedge \neg FA$. Similarly, if we define F as $\neg P$, then $FA \wedge PA$ becomes $\neg PA \wedge PA$. As I argued in Chapter 1, one can and should hold that normative conflicts (both escapable and inescapable) are possible. However, one who takes this position need not (and should not) accept the much more radical position of dialetheism, the view that contradictions can be true. If F and P are interdefinable, then anyone allowing for the possibility of escapable normative conflicts is forced into dialetheism. Thus there is a strong motivation for rejecting the view that F and P are interdefinable. Moreover, there is a very intuitive explanation for *why* they are not interdefinable. FA says that there is a positive norm prohibiting A. Thus $\neg FA$ says that there is no such positive norm. PA says that there is a positive norm licensing A. Thus $\neg PA$ says that there is no such positive norm. Intuitively, the existence of a positive norm prohibiting A is not the same thing as the absence of a positive norm licensing A. Similarly, the existence of a positive norm licensing A is not the same thing as the absence of a positive norm prohibiting A. Thus, intuitively, FA does not entail $\neg PA$, PA does not entail $\neg FA$, $\neg PA$ does not entail FA, and $\neg FA$ does not entail PA. (Of course, as noted in Section 5.3.5, p. 210, the latter two entailments *do* hold in SPD3+. This, as I have acknowledged, provides us with *some* motivation for favoring the four-valued semi-paraconsistent systems, though it should also be recognized that a deontic logic in which these entailments hold can still be robustly conflict-tolerant.) Now, a normative system could include meta-norms to the effect that whatever is not forbidden is permitted, whatever is not permitted is forbidden, and so on, but it is not a necessary truth that all normative systems have such meta-norms in effect.

Aggregation. The principle of O-aggregation, $OA \wedge OB / O(A \wedge B)$, is valid in the SPD systems. As we saw in Section 2.7.1, there are systems of deontic logic which invalidate $OA \wedge O\neg A / OB$ but still validate $O(A \wedge \neg A) / OB$. This is possible because these systems reject aggregation, so that $OA \wedge O\neg A$ does not entail $O(A \wedge \neg A)$. Advocates of such systems will argue that while normative conflicts of the form $OA \wedge O\neg A$ are possible, conflicts of the form $O(A \wedge \neg A)$ are not, *precisely because* O-aggregation is an invalid principle. I agree, of course, that conflicts of the form $OA \wedge O\neg A$ are possible. However, my position is that conflicts of the form $O(A \wedge \neg A)$ *are* possible precisely because O-aggregation is a valid principle. Hence they should not explode, as they do in non-aggregative deontic logics. Perhaps what we have here is an irresolvable clash of intuitions. All I can do is put forward my reasons for accepting aggregation. These reasons are essentially two. First, it seems to me that natural langauge inferences such as the following are intuitively valid:

George ought to get a haircut.
George ought to get a real job.
$\therefore$ George ought to get a haircut and get a real job.

Contrast these with the following (an instance of P-aggregation), which is obviously invalid:

I am permitted to drink.
I am permitted to drive.
$\therefore$ I am permitted to drink and drive.

Second, I believe that OA should be interpreted as saying that A is true at all permissible worlds. One can reject this interpretation, of course. But I think one has to concede that it is a very compelling and useful interpretation. And as soon as one accepts this interpretation, one is forced to accept aggregation. For if A is true at all permissible worlds, and B is true at all permissible worlds, then clearly $A \wedge B$ is true at all permissible worlds.

Ought implies may. The principle that 'ought' implies 'may', $OA \supset PA$, is valid only in the SPD systems with serial R. As I pointed out in Section 5.3.8, a normative system may include a metanorm to the effect that whatever is obligatory is permitted, but it does not seem to be a necessary truth about normative systems that they *always* have such a meta-norm in place. Moreover, the principle that 'ought' implies 'may' seems to multiply normative conflicts beyond necessity. For if $OA \supset PA$ is valid, then so is $OA \wedge FA \supset PA \wedge FA$; that is if 'ought' implies 'may', then any situation in which an *inescapable* conflict obtains is also one in which an *escapable* conflict obtains. (This is not to say that all inescapable conflicts are escapable!) It's not clear to me why this should be the case. Thus, I reject the principle that 'ought' implies 'may', and accordingly prefer the SPD systems in which the accessibility relation is not required to be serial. However, I am not as strongly committed to the rejection of this principle as I am to the rejection of *Intedefinability*, *Ought implies can*, and *Deontic inheritance*. For the consequences of accepting *Ought implies may* are merely undesirable, while the consequences of accepting any of the other three principles are disastrous.

Ought implies can. We have a very good reason for rejecting the principle that 'ought' implies 'can' (a.k.a. "Kant's rule"), namely the fact that contradictions can be obligatory but cannot be true. This intuition is captured in the SPD systems. Of course, we don't have the machinery to actually express Kant's rule in the language of $\mathsf{SPD}\frac{3}{4}+$. (In particular, we have no way of expressing the relevant notion of 'can'.) However, we can make the point at the "meta" level by noting that it is *not* the case that if OA is true at @ in some $\mathsf{SPD}\frac{3}{4}+$ model, then A is true at @ in some $\mathsf{SPD}\frac{3}{4}+$ model. (Here, being true at @ in some $\mathsf{SPD}\frac{3}{4}+$ model is serving as a proxy for 'possibility'.) For example, there is an $\mathsf{SPD}\frac{3}{4}+$ model in which $O(p \wedge \neg p)$ holds at the home world, but there is no $\mathsf{SPD}\frac{3}{4}+$ model at which $p \wedge \neg p$ holds at the home word. Alternatively, we could expand the $\mathsf{SPD}\frac{3}{4}+$ systems to include alethic possibility operator $\Diamond$. ($\Box$ can be defined as $\neg\Diamond\neg$.). Consider the following definitions:

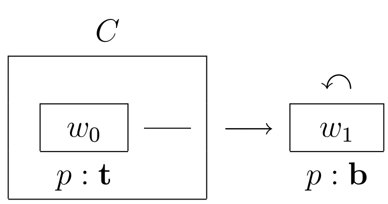
Definition 110 ($\mathsf{SPD4}\Diamond$ model) An $\mathsf{SPD4}\Diamond$ model is a quadruple $\langle W, C, R, v\rangle$, where W is a set, C is a non-empty subset of W, $R \subseteq W^2$, and $v : At \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ such that for all $w \in C$ and $p \in At$, $v_w(p) \in \{\mathbf{t}, \mathbf{f}\}$. v is extended to $\bar{v} : L_{\mathsf{SPD4}\Diamond} \times W \longrightarrow \{\mathbf{t}, \mathbf{f}, \mathbf{b}, \mathbf{n}\}$ via the following clauses:

$$
\begin{array}{lcl}
\bar{v}_w(p) & = & v_w(p) \\
1 \in \bar{v}_w(\neg A) & \Leftrightarrow & 0 \in \bar{v}_w(A) \\
0 \in \bar{v}_w(\neg A) & \Leftrightarrow & 1 \in \bar{v}_w(A) \\
1 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ or } 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \vee B) & \Leftrightarrow & 0 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \Rightarrow 1 \in \bar{v}_w(B) \\
0 \in \bar{v}_w(A \supset B) & \Leftrightarrow & 1 \in \bar{v}_w(A) \text{ and } 0 \in \bar{v}_w(B) \\
1 \in \bar{v}_w(\Diamond A) & \Leftrightarrow & \exists u(u \in C \text{ and } 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(\Diamond A) & \Leftrightarrow & \forall u(u \in C \Rightarrow 0 \in \bar{v}_u(A)) \\
1 \in \bar{v}_w(OA) & \Leftrightarrow & \forall u(wRu \Rightarrow 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(OA) & \Leftrightarrow & (w \in C \text{ and } 1 \notin \bar{v}_w(OA)) \text{ or} \\
 & & (w \notin C \text{ and } \exists u(wRu \text{ and } 0 \in \bar{v}_u(A))) \\
1 \in \bar{v}_w(PA) & \Leftrightarrow & \exists u(wRu \text{ and } 1 \in \bar{v}_u(A)) \\
0 \in \bar{v}_w(PA) & \Leftrightarrow & (w \in C \text{ and } 1 \notin \bar{v}_w(PA)) \text{ or} \\
 & & (w \notin C \text{ and } \forall u(wRu \Rightarrow 0 \in \bar{v}_u(A)))
\end{array}
$$

Definition 111 (semantic consequence) $\Gamma \models_{\mathsf{SDP4}\Diamond} A$ iff for all $\mathsf{SPD4}\Diamond$ models $\langle W, C, R, v \rangle$ and all $w \in C$, if $1 \in \bar{v}_w(B)$ for all $B \in \Gamma$, then $1 \in \bar{v}_w(A)$.

Instead of just one world that is required to be consistent and complete, we now have a non-empty set of such worlds (C). (A simple induction shows that the elements of C are consistent and complete with respect to all formulas, not just atomic ones.) Alethic possibility ($\Diamond$) is defined as truth at some element of C—intuitively, truth at some possible world. (Of course, we could specify certain extensions of $\mathsf{SPD4}\Diamond$ by excluding truth-value gaps, requiring that R be serial, etc.)

Consider the following $\mathsf{SPD4}\Diamond$ model:



In this model $O(p \wedge \neg p)$ holds at w_0 (an element of C) but $\Diamond(p \wedge \neg p)$ does not. Thus we have $\not\models_{\mathsf{SPD4}\Diamond} OA \supset \Diamond A$. It is also interesting to note (if only in passing) that both $\Box A \supset OA$ and $OA \wedge \Box(A \supset B) \supset OB$ fail in $\mathsf{SPD4}\Diamond$.

Note that the "ought implies can" principle faces problems quite independent of strictly logical considerations—for example, problems having to do with mental illness, addiction, and weakness of the will (or *akrasia*). The alcoholic ought not to drink herself into a stupor, but she can't help it. The obsessive-compulsive ought not to wash his hands two hundred times a day, but he can't help it. The copralaliac ought not to shout obscenties in public, but she can't help it. And so on. If, as such examples strongly suggest, obligation does not imply *psychological ability*, then it is not clear why it should imply *logical possibility.* At the very least, we can see that the following is a very weak argument for Kant's rule (since the first premise is probably false):

1. Obligation implies psychological ability.
2. Psychological ability implies logical possibility.
3. So, obligation implies logical possibility.

Explosion and Deontic Inheritance. The usual justification for explosion stands: since it cannot be the case that A and $\neg A$ are true, *a fortiori*, it cannot be the case that A and $\neg A$ are true and B is untrue. On the other hand, it can be the case that OA and $O\neg A$ are true and OB is untrue. In other words, explosion is valid but deontic explosion is not. Thus deontic inheritance must be abandoned. There is a simple, intuitive explanation for why it is invalid: certain inferences, such as explosion, disjunctive syllogism, and excluded middle, are truth-preserving at all *possible* worlds but not at all *permissible* worlds, since the latter can be inconsistent and/or incomplete. A pleasant side effect of abandoning the principle of deontic inheritance is that we resolve certain so-called paradoxes of deontic logic, such as the "Gentle Murder" paradox described in Section 1.4.4.

What has been accomplished in this dissertation? I have given an account of what normative conflicts are, and argued that they are possible. I have shown that standard deontic logic, as well as some classically-based alternatives to it, break down in the presence of normative conflicts. I have constructed a family of paraconsistent deontic logics that deal with normative conflicts much more effectively, but only at the significant cost of giving up part of classical propositional logic. Finally, I have proposed a compromise of sorts: a family of "semi-paraconsistent" deontic logics that handle normative conflicts at least as effectively as paraconsistent deontic logics, while preserving *all* of classical propositional logic. These systems are subject to a number of objections, most of which can be answered satisfactorily. However, there is one objection which seems to have some merit—namely, that the semi-paraconsistent systems cannot handle "Boolean" normative conflicts, e.g. $OA \wedge O\#A$. I have suggested two possible responses to this objection, but find that I cannot bring myself to confidently endorse either one.

So: I began with the fairly ambitious goal of developing a new type of deontic logic that could handle all normative conflicts effectively. What I claim to have accomplished is something more humble, but still significant: I have developed a new type of deontic logic that can handle normative conflicts more effectively than any other system proposed to date. More generally (and metaphorically), I have mapped a significant portion of the logical territory where paraconsistency and deontic logic intersect. I have laid out some of the consequences of adopting certain assumptions about norms, normative conflicts, and logic in general. I think that I have shown that a good logician can believe in the possibility of normative conflicts without ending up a bankrupt, a criminal, or a raving lunatic. I hope I have challenged some of the reader's assumptions in a positive way, suggesting new possibilities and ways of thinking—knocking some philosophers' heads (including my own), so to speak. Perhaps I have changed the reader's mind about certain points. Perhaps I have given the reader an opportunity to strengthen justifications for beliefs he or she already held. But most importantly, I hope I have laid the groundwork for further work in deontic logic. Perhaps someday someone will discover the key to developing a deontic logic with an intuitive and justifiable semantics that is robustly tolerant of all types of normative conflicts. Perhaps this will result from a creative modification of the SPD systems, such as constructing them on the basis of some other type of paraconsistent logic (e.g. dual-intuitionistic logic [152], non-adjunctive logic [86], or relevance logic [9]). On the other hand, it may result from adopting another approach entirely, such as the approach suggested in Section 1.4.10, wherein the deontic operators are explicitly relativized to specific normative systems. Naturally, I am inclined to think that the SPD systems are on the right track. But only time will tell.

Appendix A

Guide to notation

Basic logical connectives

$\neg$	(De Morgan) negation
$\wedge$	conjunction
$\vee$	disjunction
$\supset$	implication
$\rightarrow$	strong/contraposible implication
$\equiv$	equivalence
$\leftrightarrow$	strong/contraposible equivalence

LFI connectives

$\#$	untrue (Boolean negation)
©	well-behaved (classically-valued)

Deontic operators

O	obligatory
F	forbidden
P	permissible
G	gratuitous
$\bar{F}$	negative prohibition ($= \neg P$)
$\bar{P}$	negative permission ($= \neg F$)

Set theory

$\in$	element of
$\subseteq$	subset of
$\cup$	union
Γ, Δ	$\Gamma \cup \Delta$
Γ, A	$\Gamma \cup \{A\}$
$\cap$	intersection
$\varnothing$	empty set
$\wp$	powerset
$\times$	Cartesian product
$\langle \cdots \rangle$	ordered tuple
$\{x : \Phi(x)\}$	the set of x such that $\Phi(x)$

Logical consequence

$\models$	semantic
$\Vdash$	proof-theoretic
$\vdash$	generic
$\dashv\vdash$	$A \vdash B$ and $B \vdash A$
$Cn(\Gamma)$	set of A such that $\Gamma \vdash A$

Logical symbols (metalanguage)

$\Rightarrow$	material implication
$\Leftrightarrow$	material equivalence
$\forall$	for all
$\exists$	there exists
$\notin, \nvdash$, etc.	negation

Languages

At	atomic formulas ($= \{p_0, p_1, \cdots\}$)
$L(\cdots)$	At closed under the connectives $\cdots$
L_{S}	language of the system S
$p, q, \cdots$	atomic formulas
$A, B, \cdots$	formulas
$A[B/B']$	result of replacing each occurrence of B in A with B'
$\Gamma, \Delta, \cdots$	finite sets of formulas
$\Rrightarrow$	variable ranging over $\supset$, $\rightarrow$
$*$	variable ranging over binary connectives

Models

W	set of worlds (states, points, etc.)
R	(deontic) accessibility relation
v	valuation for atomic formulas only
$\bar{v}$	valuation for all formulas
@ $(= w_0)$	the actual (or "home") world

Tableaus

$+$	is true
$-$	is not true
$*$	closed branch
$\uparrow$	open, complete branch
$b(\cdots)$	result of adding $\cdots$ to b

Truth values

$\mathbf{t}$ $(=_{df} \{1\})$	true only, uniquely true
$\mathbf{f}$ $(=_{df} \{0\})$	false only, uniquely false
$\mathbf{b}$ $(=_{df} \{1, 0\})$	both true and false
$\mathbf{n}$ $(=_{df} \varnothing)$	neither true nor false

Miscellaneous

$\therefore$, $/$	therefore
$f : D \longrightarrow R$	f is a function with domain D, range R
$\mathbb{N}$	natural numbers ($= \{0, 1, 2, \cdots\}$)
IH	short for 'inductive hypothesis'

Appendix B

Tableau rules

B.1 The D systems

- D includes all rules except $[s]$ and $[h]$.
- Ds includes all rules except $[h]$.
- Dh includes all rules except $[s]$.
- Dsh includes all rules.
- In all four systems, a branch is closed iff nodes of the forms $A\ x$ and $\neg A\ x$ occur on it.

B.1.1 Primitive rules

$$\begin{array}{c} [\neg\neg] \\ \neg\neg A\ x \\ \downarrow \\ A\ x \end{array} \qquad \begin{array}{c} [\vee] \\ A \vee B\ x \\ \swarrow \quad \searrow \\ A\ x \qquad B\ x \end{array} \qquad \begin{array}{c} [\neg\vee] \\ \neg(A \vee B)\ x \\ \downarrow \\ \neg A\ x \\ \neg B\ x \end{array}$$

$$\begin{array}{c} [O] \\ OA\ x \\ x \rhd y \\ \downarrow \\ A\ y \end{array} \qquad \begin{array}{c} [\neg O] \\ \neg OA\ x \\ \downarrow \\ x \rhd i \\ \neg A\ i \end{array} \qquad \begin{array}{c} [s] \\ OA\ x \\ \boxed{x \rhd y} \\ \downarrow \\ x \rhd i \\ A\ i \end{array} \qquad \begin{array}{c} [h] \\ x \rhd y \\ \downarrow \\ y \rhd y \end{array}$$

B.1.2 Derived rules

$$\begin{array}{c} [\wedge] \\ A \wedge B\ x \\ \downarrow \\ A\ x \\ B\ x \end{array} \qquad \begin{array}{c} [\neg\wedge] \\ \neg(A \wedge B)\ x \\ \downarrow \\ \neg A \vee \neg B\ x \end{array} \qquad \begin{array}{c} [\neg\vee]' \\ \neg(A \vee B)\ x \\ \downarrow \\ \neg A \wedge \neg B\ x \end{array} \qquad \begin{array}{cc} \multicolumn{2}{c}{[\supset]} \\ \multicolumn{2}{c}{A \supset B\ x} \\ \swarrow & \searrow \\ \neg A\ x & B\ x \end{array}$$

$$\begin{array}{c} [\neg\supset] \\ \neg(A \supset B)\ x \\ \downarrow \\ A\ x \\ \neg B\ x \end{array} \qquad \begin{array}{c} [P] \\ PA\ x \\ \downarrow \\ x \rhd i \\ A\ i \end{array} \qquad \begin{array}{c} [\neg P] \\ \neg PA\ x \\ \downarrow \\ O\neg A\ x \end{array} \qquad \begin{array}{c} [\neg O]' \\ \neg OA\ x \\ \downarrow \\ P\neg A\ x \end{array}$$

B.2 P4 and P3

- Both P4 and P3 include all of the rules below.
- In P4, a branch is closed iff nodes of the forms $A+$ and $A-$ occur on it.
- P3 has the additional closure condition: $A-$ and $\neg A-$.

B.2.1 Primitive rules

$$\begin{array}{c} [\neg\neg] \\ \neg\neg A \pm \\ \downarrow \\ A \pm \end{array} \qquad \begin{array}{cc} \multicolumn{2}{c}{[\vee+]} \\ \multicolumn{2}{c}{A \vee B +} \\ \swarrow & \searrow \\ A + & B + \end{array} \qquad \begin{array}{c} [\vee-] \\ A \vee B - \\ \downarrow \\ A - \\ B - \end{array} \qquad \begin{array}{c} [\neg\vee+] \\ \neg(A \vee B) + \\ \downarrow \\ \neg A + \\ \neg B + \end{array}$$

$$\begin{array}{c}[\neg\vee-]\\ \neg(A\vee B)\ -\\ \swarrow\quad\searrow\\ \neg A\ - \qquad \neg B\ -\end{array}\qquad\qquad \begin{array}{c}[\#]\\ \#A\ \pm\\ \downarrow\\ A\ \mp\end{array}\qquad\qquad \begin{array}{c}[\neg\#]\\ \neg\#A\ \pm\\ \downarrow\\ A\ \pm\end{array}$$

B.2.2 Derived rules

$$\begin{array}{c}[\wedge+]\\ A\wedge B\ +\\ \downarrow\\ A\ +\\ B\ +\end{array}\qquad \begin{array}{c}[\wedge-]\\ A\wedge B\ -\\ \swarrow\quad\searrow\\ A\ - \qquad B\ -\end{array}\qquad \begin{array}{c}[\neg\wedge]\\ \neg(A\wedge B)\ \pm\\ \downarrow\\ \neg A\vee\neg B\ \pm\end{array}\qquad \begin{array}{c}[\neg\vee]\\ \neg(A\vee B)\ \pm\\ \downarrow\\ \neg A\wedge\neg B\ \pm\end{array}$$

$$\begin{array}{c}[\supset+]\\ A\supset B\ +\\ \swarrow\quad\searrow\\ A\ - \qquad B\ +\end{array}\qquad \begin{array}{c}[\supset-]\\ A\supset B\ -\\ \downarrow\\ A\ +\\ B\ -\end{array}\qquad \begin{array}{c}[\neg\supset]\\ \neg(A\supset B)\ \pm\\ \downarrow\\ A\wedge\neg B\ \pm\end{array}\qquad \begin{array}{c}[\equiv+]\\ A\equiv B\ +\\ \swarrow\quad\searrow\\ \begin{array}{cc}A\ - & A\ +\\ B\ - & B\ +\end{array}\end{array}$$

$$\begin{array}{c}[\equiv-]\\ A\equiv B\ -\\ \swarrow\quad\searrow\\ \begin{array}{cc}A\ + & A\ -\\ B\ - & B\ +\end{array}\end{array}\qquad \begin{array}{c}[\neg\equiv+]\\ \neg(A\equiv B)\ +\\ \swarrow\quad\searrow\\ \begin{array}{cc}A\ + & B\ +\\ \neg B\ + & \neg A\ +\end{array}\end{array}$$

$$\begin{array}{c}[\neg\equiv-]\\ \neg(A\equiv B)\ -\\ \swarrow\qquad\qquad\searrow\\ \begin{array}{cc}\begin{array}{c}A\ -\\ \swarrow\quad\searrow\\ B\ - \qquad \neg A\ -\end{array} & \begin{array}{c}\neg B-\\ \swarrow\quad\searrow\\ B- \qquad \neg A\ -\end{array}\end{array}\end{array}$$

[→+]

$$\begin{array}{c} A \to B + \\ \swarrow \qquad \searrow \\ \begin{array}{cc} A - & B+ \\ \swarrow\searrow & \swarrow\searrow \\ \neg B - \qquad \neg A + & \neg B- \qquad \neg A + \end{array} \end{array}$$

[→−]

$$\begin{array}{c} A \to B + \\ \swarrow\searrow \\ \begin{array}{cc} A + & \neg B + \\ B - & \neg A - \end{array} \end{array}$$

[¬→+]

$$\begin{array}{c} \neg(A \to B) + \\ \downarrow \\ A + \\ \neg B + \end{array}$$

[¬→−]

$$\begin{array}{c} \neg(A \to B) - \\ \swarrow\searrow \\ \begin{array}{cc} A - & \neg B - \end{array} \end{array}$$

[↔+]

$$\begin{array}{c} A \leftrightarrow B + \\ \swarrow \qquad \searrow \\ \begin{array}{cc} A - & B+ \\ \swarrow\searrow & \swarrow\searrow \\ \begin{array}{cc} \neg B - & \neg A + \\ B- & B- \\ \neg A- & \neg B+ \end{array} & \begin{array}{cc} \neg B- & \neg A + \\ A+ & A+ \\ \neg A- & \neg B+ \end{array} \end{array} \end{array}$$

[↔−]

$$\begin{array}{c} A \leftrightarrow B + \\ \swarrow \qquad \swarrow \qquad \searrow \qquad \searrow \\ \begin{array}{cccc} A + & \neg B + & B+ & \neg A + \\ B- & \neg A- & A- & \neg B- \end{array} \end{array}$$

[¬→+]

$$\begin{array}{c} \neg(A \leftrightarrow B) + \\ \swarrow\searrow \\ \begin{array}{cc} A + & B + \\ \neg B + & \neg A + \end{array} \end{array}$$

[¬↔−]

$$\begin{array}{c} \neg(A \leftrightarrow B) + \\ \swarrow \qquad \searrow \\ \begin{array}{cc} A - & \neg B- \\ \swarrow\searrow & \swarrow\searrow \\ B - \qquad \neg A - & B- \qquad \neg A - \end{array} \end{array}$$

$$\begin{array}{c} [\copyright+] \\ \copyright A + \\ \swarrow \quad \searrow \\ \begin{array}{cc} A + & A - \\ \neg A - & \neg A + \end{array} \end{array} \qquad \begin{array}{c} [\copyright-] \\ \copyright A - \\ \swarrow \quad \searrow \\ \begin{array}{cc} A + & A - \\ \neg A + & \neg A - \end{array} \end{array} \qquad \begin{array}{c} [\neg\copyright] \\ \neg\copyright A \pm \\ \downarrow \\ \copyright A \mp \end{array}$$

B.3 The PD systems

- $\mathsf{PD}\frac{3}{4}$ includes all rules except $[s+]$, $[s-]$, and $[h]$.
- $\mathsf{PD}\frac{3}{4}\mathsf{s}$ includes all rules except $[h]$.
- $\mathsf{PD}\frac{3}{4}\mathsf{h}$ includes all rules except $[s+]$ and $[s-]$.
- $\mathsf{PD}\frac{3}{4}\mathsf{sh}$ includes all rules.
- In $\mathsf{PD4}+$, a branch is closed iff nodes of the forms $A\ x+$ and $A\ x-$ occur on it.
- $\mathsf{PD3}+$ has the additional closure condition: $A\ x-$ and $\neg A\ x-$.

B.3.1 Primitive rules

$$\begin{array}{c} [\neg\neg] \\ \neg\neg A\ x\pm \\ \downarrow \\ A\ x\pm \end{array} \qquad \begin{array}{c} [\vee+] \\ A \vee B\ x+ \\ \swarrow \quad \searrow \\ A\ x+ \qquad B\ x+ \end{array} \qquad \begin{array}{c} [\vee-] \\ A \vee B\ x- \\ \downarrow \\ A\ x- \\ B\ x- \end{array} \qquad \begin{array}{c} [\neg\vee+] \\ \neg(A \vee B)\ x+ \\ \downarrow \\ \neg A\ x+ \\ \neg B\ x+ \end{array}$$

$$\begin{array}{c} [\neg\vee-] \\ \neg(A \vee B)\ x- \\ \swarrow \quad \searrow \\ \neg A\ x- \qquad \neg B\ x- \end{array} \qquad \begin{array}{c} [\#] \\ \#A\ x\pm \\ \downarrow \\ A\ x\mp \end{array} \qquad \begin{array}{c} [\neg\#] \\ \neg\#A\ x\pm \\ \downarrow \\ A\ x\pm \end{array}$$

$$\begin{array}{c} [O+] \\ OA\ x+ \\ x \rhd y \\ \downarrow \\ A\ y+ \end{array} \qquad \begin{array}{c} [O-] \\ OA\ x- \\ \downarrow \\ x \rhd i \\ A\ i- \end{array} \qquad \begin{array}{c} [\neg O+] \\ \neg OA\ x+ \\ \downarrow \\ x \rhd i \\ \neg A\ i+ \end{array} \qquad \begin{array}{c} [\neg O-] \\ \neg OA\ x- \\ x \rhd y \\ \downarrow \\ \neg A\ y- \end{array}$$

$$
\begin{array}{c}
[s+] \\
OA\ x+ \\
\boxed{x \rhd y} \\
\downarrow \\
x \rhd i \\
A\ i+
\end{array}
\qquad
\begin{array}{c}
[s-] \\
\neg OA\ x- \\
\boxed{x \rhd y} \\
\downarrow \\
x \rhd i \\
\neg A\ i-
\end{array}
\qquad
\begin{array}{c}
[h] \\
x \rhd y \\
\downarrow \\
y \rhd y
\end{array}
$$

B.3.2 Derived rules

$$
\begin{array}{c}
[\wedge+] \\
A \wedge B\ x+ \\
\downarrow \\
A\ x+ \\
B\ x+
\end{array}
\qquad
\begin{array}{c}
[\wedge-] \\
A \wedge B\ x- \\
\swarrow \searrow \\
A\ x- \qquad B\ x-
\end{array}
\qquad
\begin{array}{c}
[\neg\wedge] \\
\neg(A \wedge B)\ x\pm \\
\downarrow \\
\neg A \vee \neg B\ x\pm
\end{array}
\qquad
\begin{array}{c}
[\neg\vee] \\
\neg(A \vee B)\ x\pm \\
\downarrow \\
\neg A \wedge \neg B\ x\pm
\end{array}
$$

$$
\begin{array}{c}
[\supset+] \\
A \supset B\ x+ \\
\swarrow \searrow \\
A\ x- \qquad B\ x+
\end{array}
\qquad
\begin{array}{c}
[\supset-] \\
A \supset B\ x- \\
\downarrow \\
A\ x+ \\
B\ x-
\end{array}
\qquad
\begin{array}{c}
[\neg\supset+] \\
\neg(A \supset B)\ x\pm \\
\downarrow \\
A \wedge \neg B\ x\pm
\end{array}
$$

$$
\begin{array}{c}
[P+] \\
PA\ x+ \\
\downarrow \\
x \rhd i \\
A\ i+
\end{array}
\qquad
\begin{array}{c}
[P-] \\
PA\ x- \\
x \rhd y \\
\downarrow \\
A\ y-
\end{array}
\qquad
\begin{array}{c}
[\neg O] \\
\neg OA\ x\pm \\
\downarrow \\
P\neg A\ x\pm
\end{array}
\qquad
\begin{array}{c}
[\neg P] \\
\neg PA\ x\pm \\
\downarrow \\
O\neg A\ x\pm
\end{array}
$$

$$
\begin{array}{c}
[©+] \\
©A\ x+ \\
\swarrow \searrow \\
\begin{array}{cc}
A\ x+ & A\ x- \\
\neg A\ x- & \neg A\ x+
\end{array}
\end{array}
\qquad
\begin{array}{c}
[©+] \\
©A\ x- \\
\swarrow \searrow \\
\begin{array}{cc}
A\ x+ & A\ x- \\
\neg A\ x+ & \neg A\ x-
\end{array}
\end{array}
\qquad
\begin{array}{c}
[\neg©] \\
\neg©A\ x\pm \\
\downarrow \\
©A\ x\mp
\end{array}
$$

B.4 The SPD systems

- SPD^{3}_{4} includes all rules except $[sO+]$, $[sO-]$, $[sP+]$, $[sP-]$, and $[h]$.

- $\mathsf{SPD}\frac{3}{4}\mathsf{s}$ includes all rules except $[h]$.
- $\mathsf{SPD}\frac{3}{4}\mathsf{h}$ includes all rules except $[sO+]$, $[sO-]$, $[sP+]$, and $[sP-]$.
- $\mathsf{SPD}\frac{3}{4}\mathsf{sh}$ includes all rules.
- In $\mathsf{SPD4}+$, a branch is closed iff nodes of the forms $A\ x+$ and $A\ x-$, $A\ 0+$ and $\neg A\ 0+$, or $A\ 0-$ and $\neg A\ 0-$ occur on it.
- $\mathsf{SPD3}+$ has the additional closure condition: $A\ x-$ and $\neg A\ x-$.

B.4.1 Primitive rules

$$\begin{array}{c} [\neg\neg] \\ \neg\neg A\ x\pm \\ \downarrow \\ A\ x\pm \end{array} \qquad \begin{array}{cc} \multicolumn{2}{c}{[\vee+]} \\ \multicolumn{2}{c}{A \vee B\ x+} \\ \swarrow & \searrow \\ A\ x+ & B\ x+ \end{array} \qquad \begin{array}{c} [\vee-] \\ A \vee B\ x- \\ \downarrow \\ A\ x- \\ B\ x- \end{array} \qquad \begin{array}{c} [\neg\vee+] \\ \neg(A \vee B)\ x+ \\ \downarrow \\ \neg A\ x+ \\ \neg B\ x+ \end{array}$$

$$\begin{array}{cc} \multicolumn{2}{c}{[\neg\vee-]} \\ \multicolumn{2}{c}{\neg(A \vee B)\ x-} \\ \swarrow & \searrow \\ \neg A\ x- & \neg B\ x- \end{array} \qquad \begin{array}{cc} \multicolumn{2}{c}{[\supset+]} \\ \multicolumn{2}{c}{A \supset B\ x+} \\ \swarrow & \searrow \\ A\ x- & B\ x+ \end{array} \qquad \begin{array}{c} [\supset-] \\ A \supset B\ x- \\ \downarrow \\ A\ x+ \\ B\ x- \end{array} \qquad \begin{array}{c} [\neg\supset+] \\ \neg(A \supset B)\ x+ \\ \downarrow \\ A\ x+ \\ \neg B\ x+ \end{array}$$

$$\begin{array}{cc} \multicolumn{2}{c}{[\neg\supset-]} \\ \multicolumn{2}{c}{\neg(A \supset B)\ x+} \\ \swarrow & \searrow \\ A\ x- & \neg B\ x- \end{array} \qquad \begin{array}{c} [O+] \\ OA\ x+ \\ x \rhd y \\ \downarrow \\ A\ y+ \end{array} \qquad \begin{array}{c} [O-] \\ OA\ x- \\ \downarrow \\ x \rhd i \\ A\ i- \end{array} \qquad \begin{array}{c} [\neg O_0] \\ \neg OA\ 0\pm \\ \downarrow \\ OA\ 0\mp \end{array}$$

$$\begin{array}{c} [\neg O+] \\ \neg OA\ z+ \\ \downarrow \\ z \rhd i \\ \neg A\ i+ \end{array} \qquad \begin{array}{c} [\neg O-] \\ \neg OA\ z- \\ z \rhd y \\ \downarrow \\ \neg A\ y- \end{array} \qquad \begin{array}{c} [P+] \\ PA\ x+ \\ \downarrow \\ x \rhd i \\ A\ i+ \end{array} \qquad \begin{array}{c} [P-] \\ PA\ x- \\ x \rhd y \\ \downarrow \\ A\ y- \end{array}$$

$$
\begin{array}{c}
{\scriptstyle [\neg P_0]} \\
\neg PA\ 0\pm \\
\downarrow \\
PA\ 0\mp
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [\neg P+]} \\
\neg PA\ z+ \\
z \triangleright y \\
\downarrow \\
\neg A\ y+
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [\neg P-]} \\
\neg PA\ z- \\
\downarrow \\
z \triangleright i \\
\neg A\ i-
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [sO+]} \\
OA\ x+ \\
\boxed{x \triangleright y} \\
\downarrow \\
x \triangleright i \\
A\ i+
\end{array}
$$

$$
\begin{array}{c}
{\scriptstyle [sO-]} \\
\neg OA\ z- \\
\boxed{z \triangleright y} \\
\downarrow \\
z \triangleright i \\
\neg A\ i-
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [sP+]} \\
\neg PA\ z+ \\
\boxed{z \triangleright y} \\
\downarrow \\
z \triangleright i \\
\neg A\ i+
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [sP-]} \\
PA\ x- \\
\boxed{x \triangleright y} \\
\downarrow \\
x \triangleright i \\
A\ i-
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [h]} \\
x \triangleright y \\
\downarrow \\
y \triangleright y
\end{array}
$$

B.4.2 Derived rules

$$
\begin{array}{c}
{\scriptstyle [\wedge+]} \\
A \wedge B\ x+ \\
\downarrow \\
A\ x+ \\
B\ x+
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [\wedge-]} \\
A \wedge B\ x- \\
\swarrow \quad \searrow \\
A\ x- \qquad B\ x-
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [\neg\wedge]} \\
\neg(A \wedge B)\ x\pm \\
\downarrow \\
\neg A \vee \neg B\ x\pm
\end{array}
\qquad
\begin{array}{c}
{\scriptstyle [\neg\vee]} \\
\neg(A \vee B)\ x\pm \\
\downarrow \\
\neg A \wedge \neg B\ x\pm
\end{array}
$$

Bibliography

[1] Abe, J. M. and Pujatti, L. 2001. A meta-interpreter based on paraconsistent legal knowledge engineering. *Logic and Logical Philosophy* **9**: 129–36.

[2] Adams, D. M. 2000. *Philosophical Problems in the Law*. 3rd ed. Belmont, CA: Wadsworth.

[3] Alchourrón, C. E. 1969. Logic of norms and logic of normative propositions. *Logique et Analyse* **12**: 242–68.

[4] Alchourrón, C. E. 1996. On law and logic. *Ratio Juris* **9**(4): 331–48.

[5] Alchourrón, C. E. and Bulygin, E. 1982. The expressive conception of norms. In Hilpinen (ed.) [78].

[6] Alchourrón, C. E. and Bulygin, E. 1989. Von Wright on deontic logic and the philosophy of law. In P. A. Schilpp and L. E. Hahn (eds.), *The Philosophy of G. H. von Wright*. La Salle, IL: Open Court.

[7] Anderson, A. R. 1956. The formal analysis of normative systems. Reprinted in N. Rescher (ed.), *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press, 1967.

[8] Anderson, A. R. 1967. Some nasty problems in the formal logic of ethics. *Noûs* **1**: 345–60.

[9] Anderson, A. R. and Belnap, N. D. 1975. *Entailment: The Logic of Relevance and Necessity*. Vol. 1. Princeton: Princeton University Press.

[10] Åqvist, L. 2002. Deontic logic. In D. M. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, 2nd ed., Volume 8. Dordrecht: Kluwer Academic Publishers.

[11] Arkoudas, K., Bringsjord, S., and Bello, P. 2005. Toward ethical robots via mechanized deontic logic. AAAI Fall Symbosium, November 2005.

[12] Asenjo, F. G. 1966. A calculus of antinomies. *Notre Dame Journal of Formal Logic* **VII**(1): 103–5.

[13] Ausín, F. J. and Peña, L. 2000. Paraconsistent deontic logic with enforceable rights. In Batens et al. (eds.) [19].

[14] Austin, J. 1832. *The Province of Jurisprudence Determined*. Amherst, New York: Prometheus Books, 2000.

[15] Avron, A. 1986. On an implication connective of RM. *Notre Dame Journal of Formal Logic* **27**(2): 201–9.

[16] Avron, 1991. Natural 3-valued logics—characterization and proof theory. *Journal of Symbolic Logic* **56**: 276–94.

[17] Batens, D. 1980. Paraconsistent extensional propositional logics. *Logique et Analyse* **23**: 195–234.

[18] Batens, D. 1990. Against global paraconsistency. *Studies in Soviet Thought* **39**: 209–29.

[19] Batens, D., et al. (eds.) 2000. *Frontiers of Paraconsistent Logic*. Philadelphia: Research Studies Press.

[20] Batens, D. and De Clerq, K. 2004. A rich paraconsistent extension of full positive logic. *Logique et Analyse* **185–188**: 227–57.

[21] Baumrin, B. H. and Lupu, P. 1984. A common occurrence: conflicting duties. *Metaphilosophy* **15**(2): 77–90.

[22] Beall, J. and Restall, G. 2006. *Logical Pluralism*. Oxford: Oxford University Press.

[23] Beall, J. and Van Fraassen, B. C. 2003. *Possibilities and Paradox: An Introduction to Modal and Many-valued Logic*. Oxford: Oxford University Press.

[24] Belnap, N., Perloff, M., and Xu, M. 2001. *Facing the Future*. Oxford: Oxford University Press.

[25] Bertossi, L, Hunter, A., and Schaub, T. (eds). 2004. *Inconsistency Tolerance*. Berlin: Springer.

[26] Béziau, J.-Y. 2003. New light on the square of oppositions and its nameless corner. *Logical Investigations* **10**: 218–232.

[27] Béziau, J.-Y., Carnielli, W. A., and Gabbay, D. (eds.). In press. *Handbook of Paraconsistency and Negation*. Amserdam: Elsevier Science.

[28] Blackburn, P., de Rijke, M., and Venema, Y. 2001. *Modal Logic*. Cambridge: Cambridge University Press.

[29] Bonevac, D. 1998. Against conditional obligation. *Noûs* **32**: 37–53.

[30] Brady, R. T. 2006. *Universal Logic*. Stanford: CSLI Publications.

[31] Bremer, M. 2005. *An Introduction to Paraconsistent Logics*. Frankfurt: Peter Lang.

[32] Brink, D. O. 1996. Moral conflict and its structure. In Mason (ed.) [104].

[33] Brown, B. 1999. Yes, Virginia, there really are paraconsistent logics. *Journal of Philosophical Logic* **28**: 489–500.

[34] Brown, B. 2002. On paraconsistency. In D. Jacquette (ed.), *A Companion to Philosophical Logic*. Malden, MA: Blackwell.

[35] Carnap, R. 1943. *The Formalization of Logic*. Cambridge, MA: Harvard University Pres.

[36] Carnielli, W. A., Coniglio, M. E., and D'Ottaviano, I. M. L. (eds.). 2002. *Paraconsistency: The Logical Way to the Inconsistent*. New York: Marcel Dekker.

[37] Carnielli, W. A, Coniglio, M. E., and Marcos, J. In print. Logics of formal inconsistency. In D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, 2nd ed., Volume 12. Dordrecht: Kluwer Academic Publishers.

[38] Carnielli, W. A. and Marcos, J. 2001. Tableau systems for logics of formal inconsistency. (Extended abstract.) In H. R. Arabnia (ed.), *Proceedings of the 2001 International Conference on Artificial Intelligence*, Vol II. CSREA Press, 2001.

[39] Carnielli, W. A. and Marcos, J. 2002. A taxonomy of C-systems. In Carnielli et al. (eds.) [36].

[40] Carnielli, W. A. , Marcos, J., and de Amo, S. 2000. Formal inconsistency and evolutionary databases. *Logic and Logical Philosophy* **8**(2): 115–52.

[41] Carroll, L. 1905?. What the tortoise said to Achilles. *Mind* ??: ??–??.

[42] Castañeda, H.-N. 1981. The paradoxes of deontic logic: the simplest solution to all of them in one fell swoop. In Hilpinen (eds.) [78, pp. 37-85].

[43] Chellas, B. F. 1980. *Modal Logic*. Cambridge: Cambridge University Press.

[44] Cholvy, L. and Cuppens, F. 1999. Reasoning about norms provided by conflicting regulations. In McNamara and Prakken (eds.) [117].

[45] Da Costa, N.C.A. 1974. On the theory of inconsistent formal systems. *Notre Dame Journal of Formal Logic* **XV**(4): 497–510.

[46] Da Costa, N. C. A., Béziau, J.-Y., and Bueno, O. A. S. Paraconsistent logic in a historical perspective. *Logique et Analyse* **150–152**: 111–25.

[47] Da Costa, N.C.A. and Carnielli, W. A. 1986. On paraconsistent deontic logic. *Philosophia* **16**: 293–305.

[48] Donagan, A. 1977. *The Theory of Morality*. Chicago: University of Chicago Press.

[49] Donagan, A. 1984. Consistency in rationalist moral systems. *Journal of Philosophy* **81**: 291–309. Reprinted in [69, pp. 271-90].

[50] D'Ottaviano, I. M. L. and da Costa, N. C. A. 1970. Sur un problème de Jaśkowski. *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences de Paris* **270**: 1349–53.

[51] Dunn, J. M. 1976. Intuitive semantics for first-degree entailments and 'coupled trees'. *Philosophical Studies* **29**: 149–68.

[52] Dunn, J. M. 2000. Partiality and its dual. *Studia Logica* **65**: 5–40.

[53] Dworkin, R. 1986. *Law's Empire*. Cambridge, MA: Harvard University Press.

[54] Epstein, R. L. 2000. *Propositional Logics*. 2nd ed. Belmont, CA: Wadsworth.

[55] Fitting, M. 1983. *Proof Methods for Modal and Intuitionistic Logics*. Dordrecht: D. Reidel.

[56] Forrester, J. 1984. Gentle murder, or the adverbial samaritan. *Journal of Philosophy* **81**: 193–7.

[57] Forrester, J. 1995. Conflicts of obligation. *American Philosophical Quarterly* **32**(1): 31–44.

[58] Garson, J. W. 2006. *Modal Logic for Philosophers*. Cambridge: Cambridge University Press.

[59] Goble, L. 1966. The iteration of deontic modalities. *Logique et Analyse* **9**: 197–209.

[60] Goble, L. 1991. Murder most gentle: the paradox deepens. *Philosophical Studies* **64**: 217–27.

[61] Goble, L. 1999. Deontic logic with relevance. In McNamara and Prakken (eds.) [117].

[62] Goble, L. 2000. Multiplex semantics for deontic logic. *Nordic Journal of Philosophical Logic* **5**: 115–34.

[63] Goble, L. 2001. The Andersonian reduction and relevant deontic logic. In B. Brown and J. Woods (eds.), *New Studies in Exact Philosophy*. Oxford: Hermes Scientific Publishers.

[64] Goble, L. (ed.). 2001. *The Blackwell Guide to Philosophical Logic*. Malden, MA: Blackwell Publishers.

[65] Goble, L. 2004. A proposal for dealing with deontic dilemmas. In A. Lomuscio and D. Nute (eds.), *Lecture Notes in Computer Science Vol. 3065: Proceedings of the 7th International Workshop on Deontic Logic in Computer Science*. Berlin: Springer.

[66] Goble, L. 2005. A logic for deontic dilemmas. *Journal of Applied Logic* **3**(3–4): 461–83.

[67] Goble, L. 2006. Paraconsistent modal logic. *Logique et Analyse* **49**(193).

[68] Goodman, N. 1978. *Ways of Worldmaking*. Indianapolis: Hackett.

[69] Gowans, C. (ed.). 1987. *Moral Dilemmas*. New York: Oxford University Press.

[70] Grana, N. 1990. *Logica Deontica Paraconsistente*. Napoli: Liguori Editore.

[71] Grana, M. 1990. On a minimal non-alethic logic. *Bulletin of the Section of Logic* **19**(1): 25–9.

[72] Grant, J. and Subrahmanian, V. S. 2000. Applications of paraconsistency in data and knowledge bases. *Synthese* **125**: 121–32.

[73] Grice, H. P. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.

[74] Hamblin, C. L. 1972. Quandaries and the logic of rules. *Journal of Philosophical Logic* **1**: 74–85.

[75] Hanson, W. H. 1965. Semantics for deontic logic. *Logique et Analyse* **8**: 177–90.

[76] Hanson, W. H. 1980. First degree entailments and information. *Notre Dame Journal of Formal Logic* **21**(4): 659–71.

[77] Hill, H. H. 1987. A functional taxonomy of normative conflict. *Law and Philosophy* **6**: 227–47.

[78] Hilpinen, R. (ed.). 1981. *New Studies in Deontic Logic*. Dordrecht: Reidel.

[79] Hilpinen, R. 1985. Normative conflicts and legal reasoning. In E. Bulygin et al. (eds.), *Man, Law and Modern Forms of Life*. Doredrecht: D. Reidel.

[80] Hilpinen, R. 2001. Deontic logic. In Goble (ed.) [64].

[81] Holbo, J. 2002. Moral dilemmas and the logic of obligations. *American Philosophical Quarterly* **39**(3): 259–74.

[82] Horn, L. R. 2001. *A Natural History of Negation*. Stanford: CSLI Publications.

[83] Horty, J. F. 2001. *Agency and Deontic Logic*. Oxford: Oxford University Press.

[84] Horty, J. F. 2003. Reasoning with moral conflicts. *Noûs* **37**(4): 557–605.

[85] Hussey, E. 1972. *The Presocratics*. Indianapolis: Hackett.

[86] Jaskowski, S. 1969. Propositional calculus for contradictory deductive systems. *Studia Logica* **24**: 143–57.

[87] Jeffrey, R. 2006 [1967]. *Formal Logic: Its Scope and Limits*. 4th ed. Indianapolis: Hackett.

[88] Jørgensen, J. 1937–8. Imperatives and logic. *Erkenntnis* **7**: 288–96.

[89] Kamp, H. 1974. Free choice permission. *Proceedings of the Aristotelian Society* **74**: 57–74.

[90] Kant, I. 1793. *Religion within the Boundaries of Mere Reason*. Trans. George di Giovanni, in Immanuel Kant, *Religion and Rational Theology*, trans. and ed. Allen W. Wood and George di Giovanni. Cambridge: Cambridge University Press, 1996.

[91] Kleene, S. C. 1952. *Introduction to Metamathematics*. Amsterdam: North-Holland.

[92] Kripke, S. 1963. Semantics for normal modal logics I: normal propositional modal calculi. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik* **9**: 67–96.

[93] Lambert, K. 2001. Free logics. In Goble (ed.) [64].

[94] Lemmon, E. J. 1962. Moral dilemmas. *Philosophical Review* **70**: 139–43 and 148–58. Reprinted in Gowans (ed.) [69, pp. 101-14].

[95] Lemmon E.J. 1965. Deontic logic and the logic of imperatives. *Logique et Analyse* **VII**(2): 39–61.

[96] Lewis, D. 1974. Semantic analyses for dyadic deontic logic. In S. Stenlund (ed.), *Logical Theory and Semantic Analysis*. Dordrecht: Reidel. Reprinted in D. Lewis, *Papers in Ethics and Social Philosophy*. Cambridge: Cambridge University Press, 2000.

[97] Lewis, D. 1982. Logic for equivocators. *Noûs* **16**: 431–41. Reprinted in David Lewis, *Papers in Philosophical Logic*. Cambridge: Cambridge University Press, 1998, pp. 97-110.

[98] Lokhorst, G.-J. 2004. Mally's deontic logic. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2004 Edition.

[99] Lockhorst, G.-J. C. and Goble, L. 2004. Mally's deontic logic. *Grazer Philosophiche Studien* **27**: 37–57.

[100] Mally, E. 1971 [1926]. Grundgesetze des Sollens: Elemente der Logik des Willens. Reprinted in E. Mally, *Logische Schriften: Großes Logikfragment, Grundgesetze des Sollens*. Dordrecht: D. Reidel, 1971.

[101] Marcos, J. 2005. *Logics of Formal Inconsistency*. Ph.D. thesis, IFCH / Unicamp, Brazil; IST / UTL, Portugal.

[102] Marcus, R. M. 1980. Moral dilemmas and consistency. *Journal of Philosophy* **77**: 121–36. Reprinted in [69, pp. 188-204].

[103] Mares, E. 1992. Andersonian deontic logic. *Theoria* **58**: 3–20.

[104] Mason, H. E. (ed.) 1996. *Moral Dilemmas and Moral Theory*. New York: Oxford University Press.

[105] Mautner, T. 1973. Flaws in laws. *The Philosophical Review* **82**(1): 83–98.

[106] McArthur, R. P. 1981. Anderson's deontic logic and relevant implication. *Notre Dame Journal of Formal Logic* **22**(2): 145–54.

[107] McConnell, T. 1996. Moral residue and dilemmas. In Mason (ed.) [104, pp. 36-47].

[108] McConnell, T. 2002. Moral dilemmas. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Summer 2002 Edition.

[109] McGinnis, C. 2003. Some of the best worlds are impossible: semi-paraconsistent deontic logic and its motivation (abstract). *Bulletin of Symbolic Logic* **9**(2): 262.

[110] McGinnis, C. 2003. Review of articles in W. A. Carnielli et al. (eds.), *Paraconsistency: The Logical Way to the Inconsistent*. *Bulletin of Symbolic Logic* **9**(3): 410–2.

[111] McGinnis, C. 2005. Review of Manuel Bremer, *Introduction to Paraconsistent Logics*. *Bulletin of Symbolic Logic* **11**(3): 447–51.

[112] McGinnis, C. 2006. Tableau systems for some paraconsistent modal logics. *Electronic Notes in Theoretical Computer Science* **143**: 141–57.

[113] McGinnis, C. 2007. Paraconsistency and logical hypocrisy. In O. Tomala and R. Honzik (eds.), *The Logica Yearbook 2006*. Prague: Filosofia.

[114] McGinnis, C. 2007. Some multi-conclusion modal paralogics. *Logica Universalis* **1**(2): 335-353.

[115] McGinnis, C. In press. Semi-paraconsistent deontic logic. To appear in Béziau, Carnielli, and Gabbay (eds.) [27].

[116] McNamara, P. 2006. Deontic logic. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Spring 2006 Edition.

[117] McNamara, P. and Prakken, H. (eds.). 1999. *Norms, Logics and Information Systems*. IOS Press.

[118] Menger, K. 1939. A logic of the doubtful: on optative and imperative logic. In *Reports of a Mathematical Colloquium*. Notre Dame: Indiana University Press.

[119] Moore, G. E. 1993. *Selected Writings*. Edited by T. Baldwin. London: Routledge.

[120] Munzer, S. 1972. Validity and legal conflicts. *The Yale Law Journal* **82**(6): 1140–74.

[121] Nelson, D. 1959. Negation and separation of concepts in constructive systems. In A. Heyting (ed.), *Constructivity in Mathematics*. Amsterdam: North-Holland, 1959.

[122] Parsons, T. 1990. True contradictions. *Canadian Journal of Philosophy* **20**(3): 335–54.

[123] Payne, J. H. 1880. *Home, Sweet Home*. Boston: Lee and Shepard. (Retrieved using Google Books on May 24, 2007.)

[124] Priest, G. 1979. The logic of paradox. *Journal of Philosophical Logic* **8**: 219–41.

[125] Priest, G. 1990. Boolean negation and all that. *Journal of Philosophical Logic* **19**: 201–15.

[126] Priest, G. 1998. What is so bad about contradictions? *Journal of Philosophy*.**95**: 410–26.

[127] Priest, G. 2001. *An Introduction to Non-Classical Logic*. Cambridge: Cambridge University Press.

[128] Priest, G. 2002. Paraconsistent logic. In D. Gabbay and F. Guenthner (eds.), *Handbook of Philosophical Logic*, 2nd ed., Vol. 6. Dordrecht: Kluwer.

[129] Priest, G. 2006 [1987]. *In Contradiction: A Study of the Transconsistent*. 2nd ed. Oxford: Oxford University Press.

[130] Priest, G. 2006. *Doubt Truth to Be a Liar*. Oxford: Oxford University Press.

[131] Priest, G. and Tanaka, K. 1996. Paraconsistent logic. In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Winter 2004 Edition.

[132] Prior, A. N. 1954. The paradoxes of derived obligation. *Mind* **63**: 64–5.

[133] Puga, L., da Costa, N. C. A., and Carnielli, W. A. 1988. Kantian and non-Kantian logics. *Logique et Analyse* **31**(121–122): 3–9.

[134] Quine, W. V. 1936. Truth by convention. In O. H. Lee (ed.), *Philosophical Essays for A. N. Whitehead*. New York: Longmans, Green & Company, Inc.

[135] Ross, A. 1941. Imperatives and logic. *Theoria* **7**: 53–71.

[136] Ross, W. D. 1930. *The Right and the Good*. Oxford: Oxford University Press.

[137] Ross, W. D. 1987 [1930]. *Prima facie* duties. (A selection from Ross [136].) Reprinted in Gowans (ed.) [69, pp. 83-100].

[138] Routley, R. and Plumwood, V. 1989. Moral dilemmas and the logic of deontic notions. In G. Priest et al. (eds.), *Paraconsistent Logic: Essays on the Inconsistent*. Munich: Philosophia, 1989.

[139] Routley, R. and Routley, V. 1983. Semantical foundations for value theory. *Noûs* **7**: 441–56.

[140] Schotch, P. K. and Jennings, R. E. 1981. Non-Kripkean deontic logic. In Hilpinen (ed.) [78].

[141] Searle, J. R. 1978. *Prima facie* obligations. In J. Raz (ed.), *Practical Reasoning*. Oxford: Oxford University Press, 1978.

[142] Searle, J. R. 1995. *The Construction of Social Reality*. New York: The Free Press.

[143] Shoesmith, D. J. and Smiley, T. J. 1978. *Multiple-Conclusion Logic*. Cambridge: Cambridge University Press.

[144] Sinnott-Armstrong, W. 1984. 'Ought' conversationally implies 'can'. *Philosophical Review* **93**(2): 249–61.

[145] Smullyan, R. M. 1995 [1968]. *First-Order Logic*. 2nd ed. New York: Dover.

[146] Soeteman, A. 1987. Weak and strong permission in the law. In Frändberg and M. van Hoecke (eds.), *The Structure of Law*. Uppsula: Iustus Förlag.

[147] Stelzner, W. 1992. Relevant deontic logic. *Journal of Philosophical Logic* **21**: 193–216.

[148] Stenius, E. 1963. The principles of a logic of normative systems. *Acta Philosophica Fennica* **16**: 247–60.

[149] Stern, R. 2004. Does 'ought' imply 'can'? And did Kant think it does? *Utilitas* **16**(1): 42–61.

[150] Tarski, A. 1983 [1956]. *Logic, Semantics, Metamathematics*. Indianapolis: Hackett.

[151] Trigg, R. 1971. Moral conflict. *Mind* **80**(317): 41–55.

[152] Urbas, I. 1996. Dual-intuitionistic logic. *Notre Dame Journal of Formal Logic* **37**(3): 440–51.

[153] Verne, J. 1872. *Around the World in Eighty Days*. Translated by J. Rogers. New York: Signet, 1991.

[154] Von Wright, G. H. 1951. Deontic logic. *Mind* **60**: 1–15.

[155] Von Wright, G. H. 1963. *Norm and Action*. London: Routledge & Kegan Paul.

[156] Von Wright, G. H. 1991. Is there a logic of norms? *Ratio Juris* **4**: 265–83.

[157] Von Wright, G. H. 1999. Deontic logic: a personal view. *Ratio Juris* **12**(1): 26–38.

[158] Williams, B. 1973. Ethical consistency. In *Problems of the Self: Philosophical Papers 1956–1972*. Cambridge: Cambridge University Press. Reprinted in Gowans (ed.) [69].

Index