FILOSOFISKA NOTISER

Årgång 2, Nr 3, December 2015

Daniel Rönnedal
Alethic-Deontic Logic: Deontic Accessibility
Defined in Terms of Alethic Accessibility

Daniel Rönnedal
Alethic-Deontic Logic and
the Alethic-Deontic Octagon

Daniel Rönnedal
Allmänna Normer och
Strukturen hos Normativa System: En Logisk Analys

John F. Peterson
Quine and Plato's Beard Revisited

William Simkulet
On the Signpost Principle of Alternate Possibilities: Why Contemporary
Frankfurt-Style Cases are Irrelevant to the Free Will Debate

ISSN: 2002-0198

Hemsida: www.filosofiskanotiser.com

Alethic-Deontic Logic: Deontic Accessibility Defined in Terms of Alethic Accessibility

Daniel Rönnedal

Abstract
According to many normative theories, to say that something ought to be, or ought to be done, is to state that the being or doing of this thing is in some sense a necessary condition (requirement) of something else. In this paper, I explore the consequences of such a view. I consider what kind of alethic-deontic logic is appropriate for theories of this sort. Alethic-deontic logic is a kind of bimodal logic that combines ordinary alethic (modal) logic and deontic logic. Ordinary alethic logic is a branch of logic that deals with modal concepts, such as necessity and possibility, modal sentences, arguments and systems. Deontic logic is the logic of norms. It deals with normative words, such as "ought", "right" and "wrong", normative sentences, arguments and systems. I will define the so-called deontic accessibility relation in terms of the so-called alethic accessibility relation, and I will examine the consequences of this definition. It will turn out that a particular alethic-deontic system, *Strong alethic-deontic logic*, is plausible given this definition. By adding a certain frame-condition, *the accessibility condition*, we obtain a slightly stronger system, *Full alethic-deontic logic*. Some of the technical details of these systems are briefly described. Most of the systems mentioned in this paper are developed in more detail elsewhere.

1. Introduction

Georg Henrik von Wright has suggested that "[t]o say that something ought to be, or ought to be done, is to state that the being or doing of this thing is a necessary condition (requirement) of something else" (von Wright (1971, p. 161)). He goes on:

> [T]o say that something ought to be or ought to be done is to say that the being or doing of this thing is a necessary condition of a certain other thing which is taken for granted or presupposed in the context. An 'ought'-statement is typically an *elliptic* statement of a necessary requirement. … This suggestion seems to me, on the whole, acceptable. If we accept it, then we are always, when confronted with

Filosofiska Notiser, Årgång 2, Nr. 3, December 2015, 3–26

> an ‘ought’, entitled to raise the question ‘Why?’, *i.e.* to ask for the thing for which this or that is alleged to be a necessary requirement. (von Wright (1971 pp. 171–172))

Other philosophers have given similar analyses of some fundamental normative concepts. According to Alan Ross Anderson:

> The intimate connection between obligations and sanctions in normative systems suggests that we might profitably begin by considering some penalty or sanction *S*, and *define* the deontic or normative notions of obligation, etc. along the following lines: a state-of-affairs *p* is obligatory if the falsity of *p* entails the sanction *S*; *p* is forbidden if *p* entails the sanction *S*; and *p* is permitted if it is possible that *p* is true without the sanction *S* being true. (Anderson (1956. p. 170))

By adding these definitions to various systems of alethic modal logic, Anderson achieves a kind of “reduction” of monadic deontic logic to alethic modal logic. A similar analysis is offered by Stig Kanger (1957). The basic idea is that it ought to be the case that A iff A is a necessary condition for avoiding the sanction or for meeting some kind of demands (e.g. the demands of morality). (See also Anderson (1956), (1958), (1959), (1967) and Åqvist (1987), Chapter IV.)

In this paper, I will explore a set of normative theories that in some sense share this basic idea and consider what kind of alethic-deontic logic is appropriate for systems of this kind. Even though such systems are similar to those developed by von Wright, Anderson and Kanger, they differ from the latter in several important ways. Roughly, according to the theories we will focus on in this paper:

> It ought to be the case that A iff A is a necessary condition for creating (obtaining) a possible world that has property M, where M can be almost any property in which we are interested.

The possible world w can, for instance, have M iff w is good enough, meets the requirements of morality, is morally acceptable, has a total amount of value that is positive, above a certain threshold or maximal, is at least as good as every other (alethically accessible) world, doesn’t contain any violations of rights, or is a Kingdom of Ends, etc. According to a theory of this kind, one ought to perform an action iff performing this action is a necessary condition for creating (obtaining) a possible world that has property M. In other words, one ought to perform an action iff the state of affairs that consists in one's performing this action is a necessary condition for creating (obtaining) a possible world that has the property M. Or again, one ought to perform this action iff the state of affairs that consists in one's performing this action obtains in every possible world that has property M. More precisely, all the theories we focus on in this paper define our basic deontic concepts in the following way:

> "It ought to be the case that A" is true in the possible world w iff "A" is true in every possible world that is alethically accessible from w and that has property M.
>
> "It is permitted that A" is true in the possible world w iff "A" is true in at least one possible world that is alethically accessible from w and that has property M.
>
> "It is forbidden that A" is true in the possible world w iff "not-A" is true in every possible world that is alethically accessible from w and that has property M.

Almost every, and perhaps every plausible theory taking this form – and "defining" the alethic accessibility relation in the same way – has the same alethic-deontic logic, even though "M" may stand for many different properties. An important subclass of theories of this kind is "doing the best we can" theories. The basic idea behind these theories is that we ought to do our best, or that we ought to do the best we can. One theory of this kind has, for instance, been developed by Fred Feldman (see Feldman (1986)). According to Feldman, "all of our moral obligations boil down to one - we morally ought to do the best we can." And by this he means, "we morally ought to do what we do in the intrinsically best possible worlds still accessible to us" (Feldman (1986, xi)). He goes on to say: "As I see it… what a person ought to do as of a time is what he does in the intrinsically best worlds accessible to him as of that time" (Feldman (1986, p. 13)). According to a theory of this kind, we can, for instance, define the concept of ought in the following way:

> "It ought to be the case that A" is true in the possible world w iff "A" is true in every possible world w′ that is alethically accessible from w and that is such that there is no other possible world w″ that is alethically accessible from w that is better than w′.

This idea can (in principle) be combined with almost any value-theory and with almost any analysis of the relation "better than". Intuitively, the definition entails that one ought to do A iff one does A in all the best alethically accessible worlds. If some kind of hedonism is true, then the possible world w is better than the possible world w′ iff the total amount of well-being ("pleasure" over "pain") is higher in w than in w′. If something else has value, e.g. justice, freedom, virtue, knowledge, beauty, friendship, love etc., these values will influence the relative values of different possible worlds. We will not develop on this here. The important thing to note is that many normative theories seem to share the same basic, formal structure. We therefore have good reason to question what sort of alethic-deontic logic is appropriate for theories of this kind.

Alethic-deontic logic is a form of bimodal logic that combines ordinary alethic (modal) logic and deontic logic. Ordinary alethic logic is a branch of logic that deals with modal concepts, such as necessity and possibility, modal sentences, arguments and systems. For some introductions, see e.g. Chellas (1980), Blackburn, de Rijke & Venema (2001), Blackburn, van Benthem & Wolter (eds.) (2007), Fitting & Mendelsohn (1998), Gabbay (1976), Gabbay & Guenthner (2001), Kracht (1999), Garson (2006), Girle (2000), Lewis & Langford (1932), Popkorn (1994), Segerberg (1971), and Zeman (1973). Deontic logic is the logic of norms. It deals with normative words, such as "ought", "right" and "wrong", normative sentences, arguments and systems. Introductions to this branch of logic can be found in e.g. Gabbay, Horty, Parent, van der Meyden & van der Torre (eds.) (2013), Hilpinen (1971), (1981), Rönnedal (2010), and Åqvist (1987), (2002). Alethic-deontic logic contains both modal and normative concepts and can be used to study how the two interact. In the paper Rönnedal (2012) I say more about various bimodal systems and in Rönnedal (2015) I prove some interesting theorems in some alethic-deontic systems (see also Rönnedal (2012b) and (2015b)). Anderson was perhaps the first philosopher to combine alethic and deontic logic (see Anderson (1956)). Fine & Schurz (1996), Gabbay & Guenthner (2001), Gabbay, Kurucz, Wolter & Zakharyaschev (2003), Kracht (1999), and Kracht & Wolter (1991) offer more information about how to combine various logical systems.

In monadic deontic logic the truth-conditions for normative sentences are usually defined in terms of a primitive deontic accessibility relation. The truth-conditions for "obligation-sentences", for instance, are often defined in the following way: "It ought to be the case that A" is true in a possible world w iff "A" is true in every possible world that is deontically accessible from w.

In Rönnedal (2012) I use two primitive accessibility relations, one alethic and one deontic. In this paper, we will define the deontic accessibility relation in terms of the alethic accessibility relation and see what follows. According to this definition, the possible world w′ is deontically accessible from the possible world w iff w′ is alethically accessible from w and w′ has the property M. Given this definition of the deontic accessibility relation, it follows from the standard definition of the truth-conditions for "ought-sentences" that "it ought to be the case that A" is true in the possible world w iff "A" is true in every possible world that is alethically accessible from w and that has property M.

In this paper I only consider some alethic-deontic systems. I don't say anything about how various norms might be related to different moments in time. However, all the systems I describe can be embedded in a temporal dimension in a more or less straightforward way. For an idea about how this might be possible, see Rönnedal (2012c) (see also Rönnedal (2012b)).

The essay is divided into seven sections. In part 2 I describe the syntax of our systems and in part 3 I talk about their semantics. Part 4 deals with the proof theoretic characterization of our logics, while part 5 offers some examples of theorems in the various systems and an analysis of some arguments. Part 6 gives information about some deductively equivalent systems; and Part 7 details soundness and completeness theorems.

2. Syntax

Alphabet. (i) A denumerably infinite set Prop of proposition letters p, q, r, s, t, p_1, q_1, r_1, s_1, t_1, p_2, q_2, r_2, s_2, t_2…, (ii) the primitive truth-functional connectives $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction), $\supset$ (material implication), and $\equiv$ (material equivalence), (iii) the modal (alethic) operators $\Box$, $\Diamond$, and $\ominus$, (iv) the deontic operators O, P, and F, and (v) the brackets (,).

Language. The language L is the set of well-formed formulas (wffs) generated by the usual clauses for proposition letters and propositionally compound sentences, and the following clauses: (i) if A is a wff, then $\Box$A, $\Diamond$A and $\ominus$A are wffs, (ii) if A is a wff, then OA, PA and FA are wffs, and (iii) nothing else is a wff.

Definitions. KA = PA $\wedge$ P$\neg$A, and NA = (OA $\vee$ O$\neg$A). $\bot$ (falsum) and $\top$ (verum) are defined as usual.

Capital letters A, B, C … are used to represent arbitrary (not necessarily atomic) formulas of the object language. The upper case Greek letter Γ represents an arbitrary set of formulas. Outer brackets around sentences are usually dropped if the result is not ambiguous. We also use a, b, c, … as proposition letters.

The translationfunction t. To understand the intended interpretation of the formal language in this essay we can use the following translation function. t(¬A) = It is not the case that t(A). t(A ⊃ B) = If t(A), then t(B). And similarly for all other propositional connectives. t(□A) = It is necessary that t(A). t(◇A) = It is possible that t(A). t(⬙A) = It is impossible that t(A). t(OA) = It ought to be the case that (it is obligatory that) t(A). t(PA) = It is permitted that t(A). t(FA) = It is forbidden that t(A). t(KA) = It is optional (deontically contingent) that t(A). t(NA) = It is non-optional (deontically non-contingent) that t(A). If t(p) and t(q) are English sentences, we can use t to translate a formal sentence containing p and q into English. For instance, let t(p) be "You are honest" and t(q) be "You lie". Then the t-translation of "(Op ∧ □(p ⊃ ¬q)) ⊃ O¬q" is "If it ought to be the case that you are honest and it is necessary that if you are honest then it is not the case that you lie, then it ought to be the case that it is not the case that you lie".[1] This is an instance of the so-called means-end principle that says that every necessary consequence of what ought to be ought to be.

There seem to be several different kinds of necessity and possibility: logical, metaphysical, natural, historical etc. If not otherwise stated, we will usually mean "historical necessity" by "necessity" in this paper.

3. Semantics

3.1 Basic concepts

Alethic-deontic frame. An (alethic-deontic) frame F is a relational structure <W, R, S>, where W is a non-empty set of possible worlds, and R and S are two binary accessibility relations on W.

R "corresponds" to the operators □, ◇ and ⬙, and S to the operators O, P and F. If Rww′, we shall say that w′ is R-accessible or alethically accessible from w, and if Sww′, that w′ is S-accessible or deontically accessible from w.

Alethic-deontic model. An (alethic-deontic) model M is a pair <F, V> where: (i) F is an alethic-deontic frame; and (ii) V is a valuation or interpretation function, which assigns a truth-value T (true) or F (false) to every proposition letter p in each world w ∈ W.

[1] Of course, stylistically this is not a particularly "nice" sentence. Nevertheless, it makes a good job in conveying the informal meaning of the formal sentence.

When M = <F, V> we say that M is *based on* the frame F, or that F is the frame *underlying* M. To save space, we shall also use the following notation for a model: <W, R, S, V>, where W, R, S and V are interpreted as usual. "**F**" stands for a class of frames and "**M**" for a class of models.

Truth in a model. Let M be any model <F, V>, based on a frame F = <W, R, S>. Let w be any member of W and let A be a well-formed sentence in L. $\Vdash_{M, w}$ A abbreviates *A is true at or in the possible world w in the model M*. The truth conditions for proposition letters and sentences built by truth-functional connectives are the usual ones. The truth conditions for the remaining sentences in L are given by the following clauses: (i) $\Vdash_{M, w} \Box A$ iff for all w′ ∈ W such that Rww′: $\Vdash_{M, w'}$ A, (ii) $\Vdash_{M, w} \Diamond A$ iff for at least one w′ ∈ W such that Rww′: $\Vdash_{M, w'}$ A, (iii) $\Vdash_{M, w} \diamondminus A$ iff for all w′ ∈ W such that Rww′: $\Vdash_{M, w'} \neg A$, (iv) $\Vdash_{M, w}$ OA iff for all w′ ∈ W such that Sww′: $\Vdash_{M, w'}$ A, (v) $\Vdash_{M, w}$ PA iff for at least one w′ ∈ W such that Sww′: $\Vdash_{M, w'}$ A, and (vi) $\Vdash_{M, w}$ FA iff for all w′ ∈ W such that Sww′: $\Vdash_{M, w'} \neg A$.

Validity. A sentence A is valid on or in a class of frames **F** ($\Vdash_{\mathbf{F}}$ A) iff A is true at every world in every model based on some frame in this class.

Satisfiability. A set of sentences Γ is satisfiable in a class of frames **F** iff at some world in some model based on some frame in **F** every sentence in Γ is true. Γ is satisfiable in a model iff at some possible world in the model all sentences in Γ are true.

Logical consequence. A sentence B is a logical consequence of a set of sentences Γ on or in a class of frames **F** ($\Gamma \Vdash_{\mathbf{F}}$ B) iff B is true at every world in every model based on a frame in **F** at which all members of Γ are true.

3.2 Conditions on a frame

We will begin this section with exploring several different conditions on our frames. These conditions are divided into three classes. The first class tells us something about the formal properties of the relation R, the second about the formal properties of the relation S, and the third about how S and R are related to each other in a frame. Then we will go one and define the deontic accessibility relation in terms of the alethic accessibility relation and consider the consequences of this definition.

The variables 'x', 'y', 'z' and 'w' in tables 1, 2 and 3 are taken to range over possible worlds in W, and the symbols ∧, ⊃, ∀ and ∃ are used as metalogical symbols in the standard way. Let F = <W, R, S> be a bimodal frame and M = <W, R, S, V> be a bimodal model. If S is serial in W, i.e. if ∀x∃ySxy, we say that S satisfies or fulfils condition C-dD and also that F and

M satisfy or fulfil condition C-dD and similarly in all other cases. C-dD is called "C-dD" because the tableau rule T-dD "corresponds" to C-dD and the sentence dD is valid on the class of all frames that satisfies condition C-dD and similarly in all other cases. Let C be any of the conditions in table 1, 2 or 3. Then a C-frame is a frame that satisfies condition C and a C-model is a model that satisfies C.

3.2.1 Conditions on the relation R

Condition	Formalization of Condition
C-aT	$\forall x Rxx$
C-aD	$\forall x \exists y Rxy$
C-aB	$\forall x \forall y (Rxy \supset Ryx)$
C-a4	$\forall x \forall y \forall z ((Rxy \wedge Ryz) \supset Rxz)$
C-a5	$\forall x \forall y \forall z ((Rxy \wedge Rxz) \supset Ryz)$

Table 1

3.2.2 Conditions on the relation S

Condition	Formalization of Condition
C-dD	$\forall x \exists y Sxy$
C-d4	$\forall x \forall y \forall z ((Sxy \wedge Syz) \supset Sxz)$
C-d5	$\forall x \forall y \forall z ((Sxy \wedge Sxz) \supset Syz)$
C-dT′	$\forall x \forall y (Sxy \supset Syy)$
C-dB′	$\forall x \forall y \forall z ((Sxy \wedge Syz) \supset Szy)$

Table 2

3.2.3 Mixed conditions on alethic-deontic frames

Condition	Formalization of Condition
C-MO	$\forall x \forall y (Sxy \supset Rxy)$
C-OC	$\forall x \exists y (Sxy \wedge Rxy)$
C-OC′	$\forall x \forall y (Sxy \supset \exists z (Ryz \wedge Syz))$
C-MO′	$\forall x \forall y \forall z ((Sxy \wedge Syz) \supset Ryz)$
C-ad4	$\forall x \forall y \forall z ((Rxy \wedge Syz) \supset Sxz)$
C-ad5	$\forall x \forall y \forall z ((Rxy \wedge Sxz) \supset Syz)$
C-PMP	$\forall x \forall y \forall z ((Sxy \wedge Rxz) \supset \exists w (Ryw \wedge Szw))$
C-OMP	$\forall x \forall y \forall z ((Rxy \wedge Syz) \supset \exists w (Sxw \wedge Rwz))$
C-MOP	$\forall x \forall y \forall z ((Sxy \wedge Ryz) \supset \exists w (Rxw \wedge Swz))$

Table 3

3.3 Definition of the deontic accessibility relation in terms of the alethic accessibility relation

Rönnedal (2012) gives information about some of the relationships between the conditions introduced above. The appendix in Rönnedal (2012b) offers more information. In this section we will see what happens if we define the deontic accessibility relation in terms of the alethic accessibility relation in a certain way. Here is our definition:

Def(S) $\forall x \forall y(Sxy \equiv (Rxy \wedge My))$. The possible world y is deontically accessible from x iff y is alethically accessible from x and y has the property M.

In our theorems below we treat M as an ordinary monadic predicate. But it can be replaced by almost any predicate and the proofs will go through anyway. It follows that, as we mentioned in the introduction, almost every, and perhaps every plausible theory taking this form – and "defining" the alethic accessibility relation in the same way – has the same alethic-deontic logic, even though "M" may stand for many different properties. As we also mentioned in the introduction, an important subclass of theories of this kind is "doing the best we can" theories. According to these theories, we ought to do our best, or the best we can (see the introduction). For theories of this kind, we can replace "My" in Def(S) by "$\neg\exists z((\neg z=y \wedge Rxz) \wedge Bzy)$", where Bzy is read "z is better than y". According to these theories, the deontic accessibility relation is defined in the following way: $\forall x \forall y(Sxy \equiv (Rxy \wedge \neg\exists z((\neg z=y \wedge Rxz) \wedge Bzy)))$, which says that the possible world y is deontically accessible from the possible world x iff y is alethically accessible from x and there is no other possible world z alethically accessible from x that is better than y.

Before we introduce our theorems, we will consider one more frame- and model-condition.

C-adD $\forall x \exists y(Rxy \wedge My)$

According to this condition, every possible world x can see at least one possible world y that has the property M. We will also call C-adD *the accessibility condition*.

We are now in a position to establish some theorems that tell us something about the consequences of Def(S).

Theorem 1. R is an equivalence relation (i) iff R is reflexive (C-aT), symmetric (C-aB) and transitive (C-a4), (ii) iff R is reflexive (C-aT) and Euclidean (C-a5), (iii) iff R is serial (C-aD), symmetric (C-aB) and transitive (C-a4), (iv) iff R is serial (C-aD), symmetric (C-aB) and Euclidean (C-a5).

Proof. Straightforward.

It is reasonable to assume that the alethic accessibility relation is an equivalence relation given almost any interpretation of our alethic concepts, for instance if we think about necessity, possibility and impossibility as historical, nomological, metaphysical or logical. If we assume this, our alethic operators will behave as S5-operators.

Theorem 2. (i) Def(S) and C-adD entail C-dD and C-OC. (ii) Def(S) entails C-MO and C-MO′. (iii) Def(S) and C-aT entail C-dT′ and C-OC′. (iv) Def(S) and C-aB entail C-dB′. (v) Def(S) and C-a4 entail C-d4 and C-ad4. (vi) Def(S), C-aB and C-a4 entail C-d5 and C-ad5. (vii) Def(S), C-aT and C-a4 entail C-OMP. (viii) Def(S), C-aT, C-aB and C-a4 entail C-PMP.

Proof. Left to the reader.

Theorem 3. (i) If Def(S) is true and R is an equivalence relation in a model M, then M satisfies C-d4, C-d5, C-dT′, C-dB′, C-OC′, C-MO, C-MO′, C-ad4, C-ad5, C-PMP and C-OMP, but not necessarily C-dD, C-OC and C-MOP. (ii) If we add the condition C-adD $\forall x \exists y(Rxy \wedge My)$ (i.e. for every world x there is a world y that is alethically accessible from x and that has property M), then M also satisfies C-dD and C-OC (but not necessarily C-MOP).

Proof. This follows from theorem 1 and theorem 2.

3.4 Classification of frame classes and the logic of a class of frames

The conditions on our frames listed in tables 1, 2 and 3 can be used to obtain a categorization of the set of all frames into various kinds. We shall say that $\mathbf{F}(C_1, \ldots, C_n)$ is the class of (all) frames that satisfies the conditions $C_1, \ldots, C_n$. E.g. **F**(C-dD, C-aT, C-MO) is the class of all frames that satisfies C-dD, C-aT and C-MO. $\mathbf{F}_s$ is the set of all frames where the deontic accessibility relation is defined in terms of the alethic accessibility relation, i.e. that satisfies Def(S); and an $\mathbf{F}_s$-frame is a frame that satisfies Def(S). $\mathbf{F}_s$(Eq) is the class of all $\mathbf{F}_s$-frames where R is an equivalence relation; and $\mathbf{F}_s$(Eq, C-adD) or $\mathbf{F}_s$(Eq, adD) is the class of all $\mathbf{F}_s$-frames that satisfies C-adD (and where R is an equivalence relation).

The set of all sentences (in L) that are valid in a class of frames **F** is called the logical system of (the system of or the logic of) **F**, in symbols S(**F**) = {A

∈ L: $\Vdash_{\mathbf{F}}$ A}. E.g. S(**F**(C-dD, C-aT, C-MO)) is the set of all sentences that are valid on the class of all frames that satisfies C-dD, C-aT and C-MO.

By using this classification of frame classes we can define a large set of systems. In the next section we will develop semantic tableau systems that exactly correspond to these logics. We will see that $\mathbf{F}_s$(Eq) corresponds to Strong alethic-deontic logic and $\mathbf{F}_s$(Eq, C-adD) to Full alethic-deontic logic.

4. Proof theory

4.1 Semantic tableaux

We use a kind of indexed semantic tableau systems in this paper. A similar technical apparatus can be found in e.g. Priest (2008). The propositional part is basically the same as in Smullyan (1968) and Jeffrey (1967).

The concepts of semantic tableau, branch, open and closed branch etc. are defined as in Priest (2008) and Rönnedal (2012b, p. 131). For more on semantic tableaux, see D'Agostino, Gabbay, Hähnle & Posegga (1999), Fitting (1983), and Fitting & Mendelsohn (1998).

4.2 Tableau rules

4.2.1 Propositional rules

We use the same propositional rules as in Priest (2008) and Rönnedal (2012b). These rules are interpreted exactly as in monomodal systems.

4.2.2 Basic a-Rules

□	◇	⟠
□A, i	◇A, i	⟠A, i
irj	↓	↓
↓	irj	□¬A, i
A, j	A, j	
	where j is new	
¬□	¬◇	⟠
¬□A, i	¬◇A, i	¬⟠A, i
↓	↓	↓
◇¬A, i	□¬A, i	◇A, i

Table 4

4.2.3 Basic d-Rules

The basic d-Rules look exactly like the basic a-Rules, except that □ is replaced by O, ◇ by P, ⬲ by F, and r by s. We give them similar names.

4.2.4 Accessibility rules (a-Rules)

T-aD	T-aT	T-aB	T-a4	T-a5
i	i	irj	irj	irj
↓	↓	↓	jrk	irk
irj	iri	jri	↓	↓
where j is new			irk	jrk

Table 5

4.2.5 Accessibility rules (d-Rules)

T-dD	T-d4	T-d5	T-dT′	T-dB′
i	isj	isj	isj	isj
↓	jsk	isk	↓	jsk
isj	↓	↓	jsj	↓
where j is new	isk	jsk		ksj

Table 6

4.2.6 Accessibility rules (ad-Rules)

T-MO	T-MO′	T-OC	T-OC′	
isj	isj	i	isj	
↓	jsk	↓	↓	
irj	↓	isj	jrk	
	jrk	irj	jsk	
		where j is new	where k is new	
T-ad4	**T-ad5**	**T-PMP**	**T-OMP**	**T-MOP**
irj	irj	isj	irj	isj
jsk	isk	irk	jsk	jrk
↓	↓	↓	↓	↓
isk	jsk	jrl	isl	irl
		ksl	lrk	lsk
		where l is new	where l is new	where l is new

Table 7

4.3 Tableau systems

A tableau system is a set of tableau rules. A (normal) alethic-deontic tableau system includes all propositional rules and all basic a- and d-Rules (sections 4.2.1 to 4.2.3 and table 4). The minimal (normal) bimodal tableau system is called “T”. By adding any subset of the accessibility rules introduced in sections 4.2.4 to 4.2.6 (tables 5, 6 and 7), we obtain an extension of T. Some of these are deductively equivalent, i.e. contain exactly the same set of theorems. We use the following conventions for naming systems. We write “$aA_1 \ldots A_n dB_1 \ldots B_n adC_1 \ldots C_n$”, where $A_1 \ldots A_n$ is a list (possibly empty) of (non-basic) a-Rules, $B_1 \ldots B_n$ is a list (possibly empty) of (non-basic) d-Rules, and $C_1 \ldots C_n$ is a list (possibly empty) of (non-basic) ad-Rules. We abbreviate by omitting the initial “a” in the names of the a-Rules after the first occurrence and similarly for the d- and ad-Rules. Also, the initial “T-” in every rule is usually omitted. If a system doesn’t include any (non-basic) a-Rules, we may also omit the initial “a”. The same goes for systems with no (non-basic) d- or ad-Rules. We will sometimes add “TS-” in the beginning of a name of a system to indicate that it is a tableau system we are talking about.

E.g. aDTB45dD45T′B′adOCMOOC′MO′45PMPOMP is the normal, alethic-deontic tableau system that includes the rules T-aD, T-aT, T-aB, T-a4, T-a5, T-dD, T-d4, T-d5, T-dT′, T-dB′, T-OC, T-MO, T-OC′, T-MO′, T-ad4, T-ad5, T-PMP and T-OMP. This system, which includes several redundant rules, will also be called T_s(Eq, adD) (since it corresponds to $\mathbf{F}_s$(Eq, adD)) or *Full alethic-deontic logic* (FADL). If we drop T-OC, and T-dD from this system, we obtain a system we will call T_s(Eq) (since it corresponds to $\mathbf{F}_s$(Eq)) or *Strong alethic-deontic logic* (StADL). There are many different systems that are equivalent to FADL and StADL (see section 6).

4.4 Some proof theoretical concepts and the logic of a tableau system

The concepts of proof, theorem, derivation, consistency, inconsistency in a system etc. can be defined in the usual way. $\vdash_S A$ says that A is a theorem in the system S and $\Gamma \vdash_S A$ says that A is derivable from Γ in S.

Let S be a tableau system. Then the logic (or the (logical) system) of S, L(S), is the set of all sentences (in L) that are provable in S, in symbols L(S) = $\{A \in L: \vdash_S A\}$. E.g. L(aTdDadMO) is the set of all sentences that are provable in the system aTdDadMO, i.e. in the system that includes the basic rules and the (non-basic) rules T-aT, T-dD and T-MO.

5. Examples of theorems and arguments

5.1 Examples of theorems

Theorem 4. The sentences in tables 8 to 16 are theorems (or more precisely theorem schemas) in the indicated systems.

Proof. Left to the reader.

Name	Theorem	System
aK	□(A ⊃ B) ⊃ (□A ⊃ □B)	T
aT	□A ⊃ A	TS-aT
aD	□A ⊃ ◇A	TS-aD
aB	A ⊃ □◇A	TS-aB
a4	□A ⊃ □□A	TS-a4
a5	◇A ⊃ □◇A	TS-a5

Table 8

Name	Theorem	System
dK	O(A ⊃ B) ⊃ (OA ⊃ OB)	T
dD	OA ⊃ PA	TS-dD
d4	OA ⊃ OOA	TS-d4
d5	PA ⊃ OPA	TS-d5
dT′	O(OA ⊃ A)	TS-dT′
dB′	O(POA ⊃ A)	TS-dB′

Table 9

Name	Theorem	System
MO	□A ⊃ OA	TS-MO
OC	OA ⊃ ◇A	TS-OC
OC′	O(OA ⊃ ◇A)	TS-OC′
MO′	O(□A ⊃ OA)	TS-MO′
ad4	OA ⊃ □OA	TS-ad4
ad5	PA ⊃ □PA	TS-ad5
PMP	P□A ⊃ □PA	TS-PMP
OMP	O□A ⊃ □OA	TS-OMP
MOP	□OA ⊃ O□A	TS-MOP

Table 10

Theorem	Sys	Theorem	Sys
FA ⊃ □FA	ad4	◇OA ⊃ OA	ad5
◇PA ⊃ PA	ad4	FA ∨ □PA	ad5
⟠PA ∨ PA	ad4	⟠OA ∨ OA	ad5

Table 11. Theorems in some systems (Sys = System)

Theorem	Sys	Theorem	Sys	Theorem	Sys
□A ⊃ PA	OC	⟠A ⊃ FA	MO	◇OA ⊃ O◇A	PMP
FA ⊃ ¬□A	OC	PA ⊃ ◇A	MO	P⟠A ⊃ ⟠OA	PMP
¬(OA ∧ ⟠A)	OC	¬(P¬A∧□A)	MO	◇PA ⊃ P◇A	OMP
¬(FA ∧ □A)	OC	¬(PA ∧ ⟠A)	MO	O⟠A ⊃ □FA	OMP
PA ∨ ◇¬A	OC	FA ∨ ◇A	MO	P◇A ⊃ ◇PA	MOP
P¬A ∨ ◇A	OC	OA ∨ ◇¬A	MO	□FA ⊃ O⟠A	MOP

Table 12. Theorems in some systems (Sys = System)

□(A ∧ B) ⊃ (OA ∧ OB)	(PA ∧ □(A ⊃ B)) ⊃ PB
(□A ∨ □B) ⊃ O(A ∨ B)	□(A ⊃ B) ⊃ (PA ⊃ PB)
(□A ∧ □B) ⊃ O(A ∧ B)	PA ⊃ (□(A ⊃ B) ⊃ PB)
P(A ∧ B) ⊃ (◇A ∧ ◇B)	(FB ∧ □(A ⊃ B)) ⊃ FA
P(A ∨ B) ⊃ (◇A ∨ ◇B)	□(A ⊃ B) ⊃ (FB ⊃ FA)
(PA ∨ PB) ⊃ ◇(A ∨ B)	FB ⊃ (□(A ⊃ B) ⊃ FA)
⟠(A ∨ B) ⊃ (FA ∧ FB)	(PA ∧ □(A ⊃ B)) ⊃ ◇B
(⟠A ∨ ⟠B) ⊃ F(A ∧ B)	□(A ⊃ B) ⊃ (PA ⊃ ◇B)
(⟠A ∧ ⟠B) ⊃ F(A ∨ B)	PA ⊃ (□(A ⊃ B) ⊃ ◇B)
□(A ≡ B) ⊃ (OA ≡ OB)	(KA ∧ □(A ⊃ B)) ⊃ PB
□(A ≡ B) ⊃ (PA ≡ PB)	□(A ⊃ B) ⊃ (KA ⊃ PB)
□(A ≡ B) ⊃ (FA ≡ FB)	KA ⊃ (□(A ⊃ B) ⊃ PB)
□(A ≡ B) ⊃ (¬OA ≡ ¬OB)	(KA ∧ □(A ⊃ B)) ⊃ ◇B
□(A ≡ B) ⊃ (KA ≡ KB)	□(A ⊃ B) ⊃ (KA ⊃ ◇B)
□(A ≡ B) ⊃ (NA ≡ NB)	KA ⊃ (□(A ⊃ B) ⊃ ◇B)
(OA ∧ □(A ⊃ B)) ⊃ OB	(¬OB ∧ □(A ⊃ B)) ⊃ ¬OA
□(A ⊃ B) ⊃ (OA ⊃ OB)	□(A ⊃ B) ⊃ (¬OB ⊃ ¬OA)
OA ⊃ (□(A ⊃ B) ⊃ OB)	¬OB ⊃ (□(A ⊃ B) ⊃ ¬OA)

Table 13. Theorems in TS-MO

□(A ⊃ B) ⊃ (□A ⊃ OB)
□(A ⊃ B) ⊃ (PA ⊃ ◇B)
□(A ⊃ B) ⊃ (⟠B ⊃ FA)

(O(A ∨ B) ∧ ⟠B) ⊃ OA
☐((A ∨ B) ⊃ C) ⊃ ((OA ∨ OB) ⊃ OC)
☐((A ∨ B) ⊃ C) ⊃ ((PA ∨ PB) ⊃ PC)
☐((A ∨ B) ⊃ C) ⊃ (FC ⊃ (FA ∧ FB))
☐(A ⊃ (B ∨ C)) ⊃ (PA ⊃ (PB ∨ PC))
☐(A ⊃ (B ∨ C)) ⊃ ((FB ∧ FC) ⊃ FA)
☐((A ∧ B) ⊃ C) ⊃ ((OA ∧ OB) ⊃ OC)
☐(A ⊃ (B ∧ C)) ⊃ (OA ⊃ (OB ∧ OC))
☐(A ⊃ (B ∧ C)) ⊃ (PA ⊃ (PB ∧ PC))
☐(A ⊃ (B ∧ C)) ⊃ ((FB ∨ FC) ⊃ FA)
(O(A ∨ B) ∧ (☐(A ⊃ C) ∧ ☐(B ⊃ C))) ⊃ OC
(O(A ∨ B) ∧ (☐(A ⊃ C) ∧ ☐(B ⊃ D))) ⊃ O(C ∨ D)
(OA ∧ (☐(A ⊃ B) ∧ ☐(A ⊃ C))) ⊃ (OB ∧ OC)
(O(A ∧ B) ∧ (☐(A ⊃ C) ∨ ☐(B ⊃ D))) ⊃ O(C ∨ D)
(OA ∧ (☐(A ⊃ B) ∨ ☐(A ⊃ C))) ⊃ O(B ∨ C)
(O(A ∧ B) ∧ (☐(A ⊃ C) ∧ ☐(B ⊃ D))) ⊃ (OC ∧ OD)

Table 14. Theorems in TS-MO

(OA ∧ OB) ⊃ ◇(A ∧ B)	☐(A ⊃ B) ⊃ (☐A ⊃ PB)
(☐A ∧ ☐B) ⊃ P(A ∧ B)	(☐A ∧ ☐(A ⊃ B)) ⊃ PB
(OA ∨ OB) ⊃ ◇(A ∨ B)	☐A ⊃ (☐(A ⊃ B) ⊃ PB)
(☐A ∨ ☐B) ⊃ P(A ∨ B)	☐(A ⊃ B) ⊃ (FB ⊃ ¬☐A)
O(A ∧ B) ⊃ (◇A ∧ ◇B)	(FB ∧ ☐(A ⊃ B)) ⊃ ¬☐A
☐(A ∧ B) ⊃ (PA ∧ PB)	FB ⊃ (☐(A ⊃ B) ⊃ ¬☐A)
☐(A ⊃ B) ⊃ (OA ⊃ ◇B)	☐(A ⊃ B) ⊃ (⟠B ⊃ ¬OA)
(OA ∧ ☐(A ⊃ B)) ⊃ ◇B	(⟠B ∧ ☐(A ⊃ B)) ⊃ ¬OA
OA ⊃ (☐(A ⊃ B) ⊃ ◇B)	⟠B ⊃ (☐(A ⊃ B) ⊃ ¬OA)

Table 15. Theorems in TS-OC

☐(A ⊃ B) ⊃ (OA ⊃ PB)
(OA ∧ ☐(A ⊃ B)) ⊃ PB
☐(A ⊃ B) ⊃ (FB ⊃ ¬OA)
(FB ∧ ☐(A ⊃ B)) ⊃ ¬OA
¬(O(A ∨ B) ∧ (⟠A ∧ ⟠B))
☐((A ∨ B) ⊃ C) ⊃ ((OA ∨ OB) ⊃ PC)
☐((A ∨ B) ⊃ C) ⊃ (FC ⊃ (¬OA ∧ ¬OB))
☐(A ⊃ (B ∨ C)) ⊃ (OA ⊃ (PB ∨ PC))
☐(A ⊃ (B ∨ C)) ⊃ ((FB ∧ FC) ⊃ ¬OA)

$\Box((A \wedge B) \supset C) \supset ((OA \wedge OB) \supset PC)$
$\Box((A \wedge B) \supset C) \supset (FC \supset (\neg OA \vee \neg OB))$
$\Box((A \wedge B) \supset C) \supset (FC \supset (P\neg A \vee P\neg B))$
$\Box(A \supset (B \wedge C)) \supset (OA \supset (PB \wedge PC))$
$\Box(A \supset (B \wedge C)) \supset ((FB \vee FC) \supset \neg OA)$
$\Box(A \supset (B \wedge C)) \supset ((\neg PB \vee \neg PC) \supset \neg OA)$
$(O(A \vee B) \wedge (\Box(A \supset C) \wedge \Box(B \supset C))) \supset PC$
$(O(A \vee B) \wedge (\Box(A \supset C) \wedge \Box(B \supset D))) \supset (PC \vee PD)$
$(OA \wedge (\Box(A \supset B) \wedge \Box(A \supset C))) \supset (PB \wedge PC)$
$(O(A \wedge B) \wedge (\Box(A \supset C) \vee \Box(B \supset D))) \supset (PC \vee PD)$
$(OA \wedge (\Box(A \supset B) \vee \Box(A \supset C))) \supset (PB \vee PC)$
$(O(A \wedge B) \wedge (\Box(A \supset C) \wedge \Box(B \supset D))) \supset (PC \wedge PD)$

Table 16. Theorems in TS-OC

Theorem 5. (i) All sentences in tables 8 – 16 except the “dD”, “OC” and “MOP”-sentences are theorems in Strong alethic-deontic logic (T_s(Eq)). (ii) All sentences in tables 8 – 16 except the “MOP”-sentences are theorems in Full alethic-deontic logic (T_s(Eq, adD)).

Proof. Left to the reader.

Theorem 6. (i) In Full alethic-deontic logic (T_s(Eq, adD)) the set of all sentences can be partitioned into the following, mutually exclusive, exhaustive subsets: □A ∧ OA, OA ∧ ¬□A, PA ∧ P¬A, FA ∧ ¬⬙A, and FA ∧ ⬙A. (ii) In Full alethic-deontic logic (T_s(Eq, adD)) the following is true: $\vdash \Box(A \equiv B) \supset (*A \equiv *B)$, where * = O, P, F, K and N. (iii) In Full alethic-deontic logic (T_s(Eq, adD)) there are exactly ten distinct modalities: A, ¬A, ◇A, □A, PA, OA, ¬◇A/⬙A, ¬□A, ¬PA/FA and ¬OA.

Proof. See Rönnedal (2015).

5.2 Examples of arguments

In this section we will illustrate how the systems we describe in this essay can be used to analyze some arguments formulated in English. Then we will show how we can prove that an argument is valid or invalid.

In every system that includes T-OC, Op ⊃ ◇p is a theorem. This is one version of the so-called ought-implies-can principle (Kant’s law), which says that if it ought to be the case that p then it is possible that p, i.e. only something possible is obligatory. The contraposition of this theorem, ⬙p ⊃ ¬Op, is also provable. This theorem says that nothing impossible is obligatory.

Consider the following argument.

Argument 1
It is not possible that you stop and help this injured man and keep your promise to your friend.
Hence, it is not the case that you (all-things considered) ought to stop and help this injured man and that you (all-things considered) ought to keep your promise to your friend.

This argument seems valid, it seems impossible that the premise could be true and the conclusion false, or – in other words – that it is necessary that the conclusion is true if the premise is true. And, in fact, we can prove that it is (syntactically) valid in every alethic-deontic system that includes the tableau rule T-OC. Argument 1 can be formalized in our systems in the following way: $\neg\Diamond(h \wedge k) : \neg(Oh \wedge Ok)$, where h = You stop and help this man, and k = you keep your promise to your friend.

(1) $\neg\Diamond(h \wedge k)$, 0
(2) $\neg\neg(Oh \wedge Ok)$, 0
(3) $Oh \wedge Ok$, 0 [2, ¬¬]
(4) Oh, 0 [3, ∧]
(5) Ok, 0 [3, ∧]
(6) $\Box\neg(h \wedge k)$, 0 [1, ¬◇]
(7) 0s1 [T-OC]
(8) 0r1 [T-OC]
(9) h, 1 [4, 7, O]
(10) k, 1 [5, 7, O]
(11) $\neg(h \wedge k)$, 1 [6, 8, □]
↙ ↘
(12) ¬h, 1 [11, ¬∧] | (13) ¬k, 1 [11, ¬∧]
(14) * [9, 12] | (15) * [10, 13]

Both branches in this tree are closed. Hence, the tree is closed. It follows that the tableau constitutes a derivation of the conclusion from the premise in every system that includes T-OC. Since, these systems are sound with respect to the class of all frames that satisfies C-OC, the conclusion is a consequence of the premise in the class of all frames that satisfies this condition.

Systems of this kind rule out moral dilemmas of the following form: $OA \wedge OB \wedge \neg\Diamond(A \wedge B)$. $\neg((Op \wedge Oq) \wedge \neg\Diamond(p \wedge q))$ is a theorem. This seems to me to be a plausible view. (See Rönnedal (2012b, pp. 75–96) for more on moral dilemmas.)

Now, consider the following argument.

Argument 2

You ought to be completely honest.

It is necessary that if you are completely honest, then you do not lie.

Hence, it is forbidden that you lie.

Argument 2 is also intuitively valid; it seems necessary that if the premises are true then the conclusion is true too. We can show that the conclusion is derivable from the premises in every tableau system that includes the tableau rule T-MO. Here is a symbolization of argument 2: $Oh, \Box(h \supset \neg l) : Fl$, where h = You are completely honest, and l = You lie.

(1) $Oh, 0$
(2) $\Box(h \supset \neg l), 0$
(3) $\neg Fl, 0$
(4) $Pl, 0$ [3, ¬F]
(5) $0s1$ [4, P]
(6) $l, 1$ [4, P]
(7) $h, 1$ [1, 5, O]
(8) $0r1$ [5, T-MO]
(9) $h \supset \neg l, 1$ [2, 8, □]

↙ ↘

(10) $\neg h, 1$ [9, ⊃]	(11) $\neg l, 1$ [9, ⊃]
(12) * [7, 10]	(13) * [6, 11]

Both branches in this tree are closed. So, the tree itself is closed. This shows that the conclusion is derivable from the premises in every tableau system that includes T-MO. Since systems of this kind are sound with respect to the class of all frames that satisfies C-MO, the conclusion follows from the premises in all C-MO-frames.

This seems to be intuitively reasonable. It is a kind of means-end reasoning. In fact, $(OA \wedge \Box(A \supset B)) \supset OB$ is derivable in every system that includes T-MO. This is a version of the so-called, means-end principle that says that every necessary consequence of an obligation is obligatory.

We will now show how our systems can be used to establish that an argument is not valid. Consider the following argument.

Argument 3
You should give money to some charity.
It is necessary that if you give money to Oxfam, then you give money to some charity.
Hence, you ought to give money to Oxfam.

This argument is similar to argument 2, and it might seem to be valid. Doesn't it involve a kind of means-end reasoning that is plausible? However, on closer examination, we see that the second premise says that giving money to Oxfam is a *sufficient* condition for giving money to some charity, not a *necessary* means or consequence. There are many ways of giving money to some charity and perhaps some other way is better. Therefore, we cannot exclude the possibility that the premises are true while the conclusion is false. Of course, it might be true that you ought to give money to some charity and also true that you ought to give money to Oxfam, but this doesn't entail that the conclusion follows from the premises.

Argument 3 can be symbolized in our systems in the following way: Og, $\Box(o \supset g)$: Oo, where g = You give money to some charity, and o = You give money to Oxfam. We can show that this deduction isn't derivable in any of our systems and that the conclusion doesn't follow from the premises in any class of frames we have described. First we will show that the conclusion isn't derivable from the premises in the weakest system T.

(1) Og, 0
(2) $\Box(o \supset g)$, 0
(3) $\neg$Oo, 0
(4) P$\neg$o, 0 [3, $\neg$O]
(5) 0s1 [4, P]
(6) $\neg$o, 1 [4, P]
(7) g, 1 [1, 5, O]

At this stage the tableau is complete and open, i.e. we have applied every T-rule we can. We can use the open branch to read off a countermodel. W =

$\{w_0, w_1\}$, S = $\{<w_0, w_1>\}$, R is empty and g is true and o false in w_1. Since g is true in w_1 and w_1 is the only deontically accessible world from w_0, Og is true in w_0. $\Box(o \supset g)$ is vacuously true in w_0 since no possible world is alethically accessible from w_0. However, Oo is false in w_0. For o is false in w_1 and w_1 is deontically accessible from w_0. So, all premises are true in w_0, while the conclusion is false. Hence, this model shows that the argument isn't valid in the class of all alethic-deontic frames. However, it doesn't establish that the conclusion doesn't follow from the premises in some subset of this class. Nevertheless, we can show that the conclusion doesn't follow from the premises in any class of frames we describe in this essay. To do this we extend our countermodel with the following information: Sw_1w_1, Rw_0w_0, Rw_1w_1, Rw_0w_1, Rw_1w_0, o is false in w_0. It follows that the conclusion isn't derivable from the premises in any tableau system we consider in this paper.

These examples illustrate the usefulness of our alethic-deontic systems.

6. Deductively equivalent systems

We have mentioned two special alethic-deontic systems: Strong alethic-deontic logic and Full alethic-deontic logic. Full alethic-deontic logic is the system aTDB45dDT′B′45adMOOCMO′OC′45OMPPMP, and Strong alethic-deontic logic is the system aTDB45dT′B′45adMOMO′OC′45OMPPMP. So, FADL includes all tableau rules we have introduced in this essay except T-MOP, and StADL includes all tableau rules except T-dD, T-OC and T-MOP. For our purposes, FADL and StADL are especially interesting since they correspond to the class of all frames where the deontic accessibility relation is defined in terms of the alethic accessibility relation (according to Def(S)), and where the alethic accessibility relation is an equivalence relation. In the case of FADL, we also assume condition C-adD ($\forall x \exists y(Rxy \wedge My)$). There are many "weaker" systems, i.e. systems with fewer tableau rules, that are deductively equivalent, i.e. contain exactly the same theorems, with FADL or StADL. The following theorem mentions some of these.

Theorem 7. (i) The following systems are deductively equivalent with FADL: aB4dDadMO4, aB4dDadMO5, aB5dDadMO4, aB5dDadMO5, aT5dDadMO4, aT5dDadMO5, aB4dadMOOC4, aB4dadMOOC5, aB5dadMOOC4, aB5dadMOOC5, aT5dadMOOC4, and aT5dadMOOC5. (ii) The following systems are deductively equivalent with StADL: aTB4adMO4, and aTB4adMO5.

Proof. Left to the reader. The appendix in Rönnedal (2012b) may be useful.

7. Soundness and completeness theorems

Let S = $aA_1...A_ndB_1...B_nadC_1...C_n$ be a normal alethic-deontic tableau system, where $A_1...A_n$ is some subclass of our (non-basic) a-Rules, $B_1...B_n$ is some subclass of our (non-basic) d-Rules and $C_1...C_n$ is some subclass of our (non-basic) ad-Rules. Then we shall say that the class of frames, **F**, corresponds to S just in case **F** = **F**(C-A_1, …, C-A_n, C-B_1, …, C-B_n, C-C_1, …, C-C_n).

S is strongly sound with respect to **F** iff $\Gamma \vdash_S A$ entails $\Gamma \Vdash_{\mathbf{F}} A$. S is strongly complete with respect to **F** just in case $\Gamma \Vdash_{\mathbf{F}} A$ entails $\Gamma \vdash_S A$.

Theorem 8 (Soundness theorem). Let S be any of our normal alethic-deontic tableau systems and let **F** be the class of frames that corresponds to S. Then S is strongly sound with respect to **F**.

Proof. See Rönnedal (2012) and/or Rönnedal (2012b). ■

Theorem 9 (Completeness theorem). Let S be any of our normal alethic-deontic tableau systems and let **F** be the class of frames that corresponds to S. Then S is strongly complete with respect to **F**.

Proof. See Rönnedal (2012) and/or Rönnedal (2012b). ■

From the soundness and completeness theorems and theorems 1–3 it follows that Strong alethic-deontic logic is the system that is appropriate for $\mathbf{F}_s$(Eq) and that Full alethic-deontic logic is the system that is appropriate for $\mathbf{F}_s$(Eq, C-adD).

References

Anderson, A. R. (1956). The formal analysis of normative systems. In N. Rescher (ed.). *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press, 1967, pp. 147–213.

Anderson, A. R. (1958). A reduction of deontic logic to alethic modal logic. *Mind*, Vol. 67, No. 265, pp. 100–103.

Anderson, A. R. (1959). On the logic of commitment. *Philosophical Studies* 10, pp. 23–27.

Anderson, A. R. (1967). Some Nasty Problems in the Formal Logic of Ethics. *Noûs*, Vol. 1, No. 4, pp. 345–360.

Blackburn, P., de Rijke, M. & Venema, Y. (2001). *Modal Logic*. Cambridge University Press.

Blackburn, P., van Benthem, J. & Wolter, F. (eds.). (2007). *Handbook of Modal Logic*. Elsevier.

Chellas, B. F. (1980). *Modal Logic: An Introduction*. Cambridge: Cambridge University Press.

Feldman, F. (1986). *Doing The Best We Can: An Essay in Informal Deontic Logic*. Dordrecht: D. Reidel Publishing Company.

Fine, K. & Schurz, G. (1996). Transfer Theorems for Multimodal Logics. In J. Copeland (ed.). (1996). *Logic and Reality. Essays in Pure and Applied Logic. In Memory of Arthur Prior*. Oxford University Press, Oxford, pp. 169–213.

Fitting, M. & Mendelsohn, R. L. (1998). *First-Order Modal Logic*. Kluwer Academic Publishers.

Gabbay, D. M. (1976). *Investigations in Modal and Tense Logics with Applications to Problems in Philosophy and Linguistics*. Dordrecht: D. Reidel Publishing Company.

Gabbay, D., Horty, J., Parent, X., van der Meyden, E. & van der Torre, L. (eds.). (2013). *Handbook of Deontic Logic and Normative Systems*. College Publications.

Gabbay, D. M. & Guenthner, F. (eds.). (2001). *Handbook of Philosophical Logic 2nd Edition*, Vol. 3, Dordrecht: Kluwer Academic Publishers.

Gabbay, D. M., Kurucz, A., Wolter, F. & Zakharyaschev, M. (2003). *Many-Dimensional Modal Logics: Theory and Applications*. Amsterdam: Elsevier.

Garson, J. W. (2006). *Modal Logic for Philosophers*. New York: Cambridge University Press.

Girle, R. (2000). *Modal Logics and Philosophy*. McGill-Queen's University Press.

Hilpinen, R. (ed.). (1971). *Deontic Logic: Introductory and Systematic Readings*. Dordrecht: D. Reidel Publishing Company.

Hilpinen, R. (ed.). (1981). *New Studies in Deontic Logic Norms, Actions, and the Foundation of Ethics*. Dordrecht: D. Reidel Publishing Company.

Kanger, S. (1957). New Foundations for Ethical Theory. Stockholm. Reprinted in *Hilpinen* (ed.) (1971), pp. 36–58.

Kracht, M. (1999). *Tools and Techniques in Modal Logic*. Amsterdam: Elsevier.

Kracht, M. & Wolter, F. (1991). Properties of Independently Axiomatizable Bimodal Logics. *The Journal of Symbolic Logic*, vol. 56, no. 4, pp. 1469–1485.

Lewis, C. I. & Langford, C. H. (1932). *Symbolic Logic*. New York: Dover Publications. Second edition 1959.

Popkorn, S. (1994). *First Steps in Modal Logic*. Cambridge University Press.

Rescher, N. (ed.). (1967). *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press.

Segerberg, K. (1971). *An Essay in Classical Modal Logic*. 3 vols. Uppsala: University of Uppsala.

Zeman, J. J. (1973). *Modal Logic: The Lewis-Modal Systems*. Oxford: Clarendon Press.

Rönnedal, D. (2010). *An Introduction to Deontic Logic*. Charleston, SC.

Rönnedal, D. (2012). Bimodal Logic. *Polish Journal of Philosophy*. Vol. VI, No. 2, pp. 71–93.

Rönnedal, D. (2012b). *Extensions of Deontic Logic: An Investigation into some Multi-Modal Systems*. Department of Philosophy, Stockholm University.

Rönnedal, D. (2012c). Temporal alethic-deontic logic and semantic tableaux. *Journal of Applied Logic*, 10, pp. 219–237.

Rönnedal, D. (2015). Alethic-Deontic Logic: Some Theorems. *Filosofiska Notiser*, Årgång 2, Nr. 1, pp. 61–77.

Rönnedal, D. (2015b). Alethic-Deontic Logic and the Alethic-Deontic Octagon. *Filosofiska Notiser*, Årgång 2, Nr. 3, pp. 27–68.

von Wright, G. H. (1971). Deontic Logic and the Theory of Conditions. In *Hilpinen* (ed.). (1971), pp. 159–177.

Åqvist, L. (1987). *Introduction to Deontic Logic and the Theory of Normative Systems*. Naples: Bibliopolis.

Åqvist, L. (2002). Deontic Logic. In Gabbay & Guenthner (eds.). *Handbook of Philosophical Logic 2nd Edition*. Vol. 8, Dordrecht: Kluwer Academic Publishers, pp. 147–264.

Daniel Rönnedal
Department of Philosophy
Stockholm University
daniel.ronnedal@philosophy.su.se

Alethic-Deontic Logic and the Alethic-Deontic Octagon

Daniel Rönnedal

Abstract

This paper will introduce and explore a set of alethic-deontic systems. Alethic-deontic logic is a form of logic that combines ordinary (alethic) modal logic, which deals with modal concepts such as necessity, possibility and impossibility, and deontic logic, which investigates normative expressions such as "ought", "right" and "wrong". I describe all the systems axiomatically. I say something about their properties and prove some theorems in and about them. We will be especially interested in how the different deontic and modal concepts are related to each other in various systems. We will map these relationships in an alethic-deontic octagon, a figure similar to the classical so-called square of opposition.

1. Introduction

In this paper I introduce and explore a set of alethic-deontic systems. Alethic-deontic logic is a kind of bimodal logic that combines ordinary (alethic) modal logic and deontic logic. Introductions to ordinary (alethic) modal logic can be found in e.g. Chellas (1980), Blackburn, de Rijke, & Venema (2001), Blackburn, van Benthem & Wolter (eds.) (2007), Fitting & Mendelsohn (1998), Gabbay (1976), Gabbay & Guenthner (2001), Kracht (1999), Garson (2006), Girle (2000), Lewis & Langford (1932), Popkorn (1994), Segerberg (1971), and Zeman (1973). This branch of logic deals with modal concepts, such as necessity, possibility and impossibility, modal sentences, arguments and systems. Introductions to deontic logic can be found in e.g. Gabbay, Horty, Parent, van der Meyden & van der Torre (eds.) (2013), Hilpinen (1971), (1981), Rönnedal (2010), and Åqvist (1987), (2002). Deontic logic deals with normative words, such as "ought", "right" and "wrong", normative sentences, arguments and systems. For more information about bimodal systems in general and alethic-deontic logics in particular, see e.g. Rönnedal (2012), (2012b), (2015), (2015b). Alethic-deontic logic combines ordinary alethic modal logic and deontic logic. Every axiomatic system in this paper is sound and complete with respect to its semantics (see Rönnedal (2012) and (2012b) for a proof). The present paper includes more information about these systems; I prove several theorems in and about them. We will be especially interested in the relationships between the different modal and normative concepts in various systems. We will use an alethic-deontic octagon, a figure similar to the classical so-called square of opposition, to map these relationships.[1]

The paper is divided into five sections. Section 2 is about syntax and semantics and section 3 about proof theory. Section 4 is the main part of the paper, in which I describe a set of normal alethic-deontic systems. Finally, section 5 includes information about the relationships between the systems I describe.

2. Syntax and semantics

We use the same kind of syntax and semantics as in Rönnedal (2015). However, we introduce a new deontic operator, U (unobligatory), defined in the following way: $Up \leftrightarrow \neg Op$. Furthermore, we use slightly different symbols and treat O and $\Box$ as primitive in this essay; all other operators are defined in terms of O and $\Box$ in a standard way. $\triangledown p$ (alethic contingency) = $\neg\Box p \wedge \neg\Box\neg p$; $\triangle p$ (alethic non-contingency) = $\Box p \vee \Box\neg p$; $\boxminus p$ (unnecessary) = $\neg\Box p$. $\top$ (Verum) = e.g. $p \vee \neg p$, $\bot$ (Falsum) = e.g. $\neg\top$.

Without further ado, let us turn to proof theory.

3. Proof theory

3.1 Systems of alethic-deontic logic

In this paper a system is usually identified with a set of sentences, not a set of theorems together with a deductive apparatus. The concept of a theorem is defined in the standard way (see e.g. Rönnedal (2010)).

Definition 1 (Alethic-deontic system). A set of sentences S is *a system of alethic-deontic logic* or simply an *alethic-deontic logic* or an *alethic-deontic system* (“ad” for short) if and only if:

(i) it contains all propositional tautologies,

[1] Anderson was perhaps the first philosopher to combine alethic and deontic logic (see Anderson (1956)). Fine & Schurz (1996), Gabbay & Guenthner (2001), Gabbay, Kurucz, Wolter & Zakharyaschev (2003), Kracht (1999), and Kracht & Wolter (1991) offer more information about how to combine various logical systems.

(ii) it is closed under modus ponens (MP) (if A is in S and $A \to B$ is in S, then so is B), and

(iii) it is closed under uniform substitution (if A belongs to S, then every (immediate) substitution instance of A is in S).

The concept of a substitution instance of A is defined in the usual way (see e.g. Rönnedal (2010)). "PL" (as in "propositional logic") contains every sentence that is valid due to its truth-functional nature. When we are talking about ad systems we presuppose that we are using a language that includes both deontic and alethic operators and not just alethic or just deontic terms. So, PL will include sentences that are not theorems in ordinary propositional logic or in pure deontic or alethic systems. For example PL contains not just $\neg(p \wedge \neg p)$ and $p \vee \neg p$, but also for instance $\neg(\Box Op \wedge \neg \Box Op)$ and $\Diamond Pp \vee \neg \Diamond Pp$. In a proof, "PL" may also indicate that the step is propositionally correct.

If it is clear from the context that we are speaking of alethic-deontic systems and alethic-deontic logics we will sometimes drop the word "alethic-deontic" and speak only of logics and of systems.

Example 2 (ad systems). (i) The inconsistent system, i.e. the set of *all* sentences is an alethic-deontic logic. This system is the largest alethic-deontic system, since every logic is included in it. (ii) Let **L** be a collection of alethic-deontic systems. Then the intersection of **L** is an alethic-deontic system too, where the intersection of **L** is defined in the standard way. (iii) The logic of any alethic-deontic frame is an alethic-deontic system. (iv) This is also true for logics of classes of alethic-deontic frames. (iv) PL ("propositional logic") is an alethic-deontic system. Since PL is a subset of every alethic-deontic logic, PL is the smallest alethic-deontic system.

We shall say that an *alethic-deontic system* S is *generated by* a set of sentences G iff S is the smallest alethic-deontic logic containing every sentence in G. PL, the set of all "tautologies" is generated by the empty set.

Definition 3 (Normal alethic-deontic system). An alethic-deontic system is *normal* if and only if:

(i) it contains the sentences $\Box(p \to q) \to (\Box p \to \Box q)$, $\Diamond p \leftrightarrow \neg\Box\neg p$, $\diamondminus p \leftrightarrow \Box\neg p$, $\boxminus p \leftrightarrow \neg\Box p$, $\triangledown p \leftrightarrow (\neg\Box p \wedge \neg\Box\neg p)$, $\triangle p \leftrightarrow (\Box p \vee \Box\neg p)$, $O(p \to q) \to (Op \to Oq)$, $Pp \leftrightarrow \neg O\neg p$, $Fp \leftrightarrow O\neg p$, $Up \leftrightarrow \neg Op$, $Kp \leftrightarrow (\neg Op \wedge \neg O\neg p)$, $Np \leftrightarrow (Op \vee O\neg p)$, and

(ii) it is closed under the rules of $\Box$-necessitation and O-necessitation (i.e if $\vdash A$, then $\vdash \Box A$, and if $\vdash A$, then $\vdash OA$).

Example 4 (Normal ad systems). (i) The inconsistent system is a normal aletic-deontic logic. (ii) PL is not a normal ad system. However, PL is included in every normal ad system. (iii) Let **L** be a collection of normal alethic-deontic systems. Then the intersection of **L** is a normal alethic-deontic system too. (iv) The logic of any alethic-deontic frame is a normal alethic-deontic system. (v) This is true also for the logic of every class of alethic-deontic frames. (vi) The pure deontic system **dK** (= **OK**) (Rönnedal (2010)) is not a normal alethic-deontic logic. Neither is the pure alethic system **aK** (Chellas (1980)). However, it follows from the definition that every normal ad system includes the minimal normal alethic logic **aK** and the minimal normal deontic logic **dK**.

The smallest normal ad system will be called "minimal alethic-deontic logic" (**MADL**) or **aKdK**.

When we speak of alethic-deontic systems in this essay, it is usually normal alethic-deontic systems we mean.

3.2 Normal alethic-deontic systems

3.2.1 Axioms

A normal alethic-deontic system can be represented by adding axioms to the minimal alethic-deontic logic **MADL**. We will consider three different kinds of axioms in this essay: pure deontic axioms, pure (alethic) modal axioms and bimodal (alethic) modal deontic axioms. And we will use these axioms to construct some normal alethic-deontic systems. The (alethic) modal axioms include aK, aT, aD, a4, aB and a5 (see table 1), well known from ordinary modal logic. The deontic axioms include dK, dD, d4, dT′, dB′ and d5 (see table 2), well know from pure deontic logic. We also consider nine bimodal axioms, i.e. axioms that contain both deontic and (alethic) modal operators, namely, MO, OC, OC′, MO′, ad4, ad5, PMP, OMP, MOP (see table 3). aK and dK are theorems in every normal alethic-deontic system. However, no other axiom is a theorem in **MADL**. Accordingly, we obtain a whole range of normal alethic-deontic systems by adding any subset of these to **MADL**. A system that fuses two monomodal systems, without any bimodal axioms, will be called an alethic-deontic combination (fusion) or ad combination (fusion) for short. See section 3.2.3 below.[2]

All in all we describe 21 different axioms, 19 besides aK and dK. Every ad system we consider will contain aK and dK and zero or more of the other

[2] All systems in this paper are generated from various axioms, rules of inference and the rule of substitution. An alternative is to use axiom schemas and dispense with the substitution rule. Both "methods" generate the same systems.

19 axioms. In fact, we will focus on the 16 systems that can be constructed from the axioms aD, dD, OC and MO. Some of these are deductively equivalent (see section 5).

3.2.1.1 Pure a-axioms

	a-axiom	Corresponding condition on R
aK	$\Box(p\to q)\to(\Box p\to\Box q)$	-
aT	$\Box p\to p$	$\forall x xRx$
aD	$\Box p\to\Diamond p$	$\forall x\exists y xRy$
aB	$p\to\Box\Diamond p$	$\forall x\forall y(xRy\to yRx)$
a4	$\Box p\to\Box\Box p$	$\forall x\forall y\forall z((xRy\wedge yRz)\to xRz)$
a5	$\Diamond p\to\Box\Diamond p$	$\forall x\forall y\forall z((xRy\wedge xRz)\to yRz)$

Table 1

3.2.1.2 Pure d-axioms

	d-axiom	Corresponding condition on S
dK	$O(p\to q)\to(Op\to Oq)$	-
dD	$Op\to Pp$	$\forall x\exists y xSy$
d4	$Op\to OOp$	$\forall x\forall y\forall z((xSy\wedge ySz)\to xSz)$
d5	$Pp\to OPp$	$\forall x\forall y\forall z((xSy\wedge xSz)\to ySz)$
dT′	$O(Op\to p)$	$\forall x\forall y(xSy\to ySy)$
dB′	$O(POp\to p)$	$\forall x\forall y\forall z((xSy\wedge ySz)\to zSy)$

Table 2

3.2.1.3 Mixed ad-axioms

	ad-axiom	Corresponding semantic condition
MO	$\Box p\to Op$	$\forall x\forall y(xSy\to xRy)$
OC	$Op\to\Diamond p$	$\forall x\exists y(xSy\wedge xRy)$
OC′	$O(Op\to\Diamond p)$	$\forall x\forall y(xSy\to\exists z(yRz\wedge ySz))$
MO′	$O(\Box p\to Op)$	$\forall x\forall y\forall z((xSy\wedge ySz)\to yRz)$
ad4	$Op\to\Box Op$	$\forall x\forall y\forall z((xRy\wedge ySz)\to xSz)$
ad5	$Pp\to\Box Pp$	$\forall x\forall y\forall z((xRy\wedge xSz)\to ySz)$
PMP	$P\Box p\to\Box Pp$	$\forall x\forall y\forall z((xSy\wedge xRz)\to\exists w(yRw\wedge zSw))$
OMP	$O\Box p\to\Box Op$	$\forall x\forall y\forall z((xRy\wedge ySz)\to\exists w(xSw\wedge wRz))$
MOP	$\Box Op\to O\Box p$	$\forall x\forall y\forall z((xSy\wedge yRz)\to\exists w(xRw\wedge wSz))$

Table 3

3.2.2 Axiomatic systems

We are now in a position to say something more systematic about alethic-deontic systems.

We have seen that **MADL** is the smallest normal alethic-deontic logic. This means that **MADL** is included in every other normal ad system. By adding axioms to the axiomatic basis it is possible to extend this logic.

As usual, we shall say that the normal alethic-deontic logic *generated* or *represented* by a set of sentences **Γ** is the smallest normal alethic-deontic logic that includes all sentences in **Γ**. **MADL** is represented by the empty set, extensions of this system by some non-empty set.

Let “**S**” be the name of a normal alethic-deontic system and “**Γ**” the name of a set of axioms. Then **S** + **Γ** is the smallest normal ad system that includes both **S** and every sentence in **Γ**. A special type of ad systems is called ad combinations (fusions). The name of an ad combination will often have the following form: “**aXdY**”, where **X** is a set of alethic axioms and **Y** is a set of deontic axioms (see below). More generally, we shall often write **aXdYadZ** for a normal ad system that can be represented by a set **X** of alethic axioms, a set **Y** of deontic axioms and a set **Z** of bimodal axioms (axioms that include both alethic and deontic operators). The ad combination **aXdY** = **aXdYad∅**. Sometimes we will replace **X**, **Y** and **Z** by names of alethic, deontic or alethic-deontic axioms or systems, respectively.

Example 5. a∅d∅ad∅ = **MADL**. Let **X** = {aT}, **Y** = {dD} and **Z** = {MO, OC}. Then **aXdYadZ** = **aTdDadMOOC** = **MADL** + {aT} ∪ {dD} ∪ {MO, OC} = **MADL** + {aT, dD, MO, OC} = **MADL** + {$\Box p \rightarrow p$, $Op \rightarrow Pp$, $\Box p \rightarrow Op$, $Op \rightarrow \Diamond p$}. Let **X** = {aT, aB, a4}, **Y** = {} and **Z** = {OC′}. Then **aXdYadZ = aTB4d∅adOC′** = **MADL** + {aT, aB, a4} ∪ {} ∪ {OC′} = **MADL** + {aT, aB, a4, OC′} = **MADL** + {$\Box p \rightarrow p$, $p \rightarrow \Box\Diamond p$, $\Box p \rightarrow \Box\Box p$, $O(Op \rightarrow \Diamond p)$}. **a∅dSDLad∅** = **MADL** + {dD} = **MADL** + {$Op \rightarrow Pp$}. **aS5dOS5+adMOOC** = **MADL** + {aT, aD, aB, a4, a5, dD, dT′, dB′, d4, d5, MO, OC}. Since aK and dK are included in every normal ad system, it is not necessary to mention them in the name of a system. E.g. the following identities hold: **aKdKad∅** = **a∅d∅ad∅, aKTdK5adOC** = **aTd5adOC** and **aK45dKad∅** = **a45d∅ad∅.**

3.2.3 ad combinations (fusions)

Let us say something more about ad combinations (fusions). A (normal) alethic-deontic system **adS** is called the combination (fusion) of a (normal) alethic modal system **aS** and a (normal) deontic system **dS**, written **aS**+**dS**, if and only if **adS** is the smallest (normal) ad system that includes both **aS** and **dS**. If **aS** is representable as **aX** and **dS** as **dY**, then **aS** + **dS** is representable as **aXdY**, where **X** and **Y** is a set of alethic and a set of deontic axioms, respectively. Hence, **aX** + **dY** = **aXdY**.

Note that **aXdY** ≠ **aX** ∪ **dY**, i.e. the ad combination of **aX** and **dY** is not identical to the union of the pure alethic system **aX** and the pure deontic system **dY**. For **aXdY** contains sentences that are not included in **aX** ∪ **dY**. Every normal ad system contains O-nec and □-nec. So, both □O(p → p) and O□(p → p) are, for instance, elements in **aXdY**, but not in **aX** ∪ **dY**. Other examples are the following sentences: □O(p → q) → (□Op → □Oq), ◇O(p ∧ q) ↔ ◇(Op ∧ Oq) and (□Fr ∧ □O((p ∨ q) → r)) → (□Fp ∧ □Fq). Furthermore, additional axioms together with one or more rules of inference may generate sentences that are theorems of the combination of **aS** and **dS** that are not theorems in the union of **aS** and **dS**. E.g. suppose that □p → p ∈ **X**, then □Pp → Pq ∈ **aXdY** but not □Pp → Pq ∈ **aX** ∪ **dY** (since □Pp → Pq is neither an element in **aX** nor in **dY**), or that Op → Pp ∈ **Y**, then □Op → □Pp ∈ **aXdY**, but not □Op → □Pp ∈ **aX** ∪ **dY** (since □Op → □Pp is neither an element in **aX** nor in **dY**). However, the union of the pure alethic system **aX** and the pure deontic system **dY** is of course a subset of the combination of **aX** and **dY**, **aX** ∪ **dY** ⊆ **aXdY**, i.e. everything included in **aX** ∪ **dY** is also included in **aXdY**. It follows that **aX** ∪ **dY** ⊂ **aXdY**.

4. Some normal alethic-deontic systems

I will now consider some normal alethic-deontic systems and I will prove some theorems in and about these systems.

4.1 Minimal alethic-deontic logic

Minimal alethic-deontic logic (**MADL**, **aKdK**, **aKdKad∅** or **a∅d∅ad∅**) is the smallest normal alethic-deontic logic. We will also call this system **S1**. Since it is a *normal* alethic-deontic system **MADL** includes PL, the axioms aK and dK, the usual definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation. Since it is the *smallest* normal alethic-deontic logic it contains no other axioms or rules of inference. A normal **aKdKad∅**-system is any normal alethic-deontic extension of **aKdKad∅**, i.e. every normal alethic-deontic system is a normal **aKdKad∅**-system, or, in other words, every normal ad system is an extension of **MADL**. This is true by definition and trivial.

MADL is an ad combination of the purely deontic system **dK** (= **OK)** and the purely alethic system **aK**. Hence, we can also call this system **aKdK** or simply **a∅d∅**. Recall that an ad combination of two systems is not the same as the union of these systems (section 3.2.3). So, **aKdK** ≠ **aK** ∪ **dK**. **aKdK** has theorems that are not elements in **aK** ∪ **dK** (e.g. □O(p → q) → □(Pp → Pq)). On the other hand, every sentence that belongs to either **aK** or **dK** is an element in **aKdK,** i.e. if s ∈ **aK** ∪ **dK**, then s ∈ **aKdK**, for any sentence s. It follows that if any formula is a theorem in either **aK** or **dK** it is also a theorem in every normal alethic-deontic logic.

I will now prove some theorems in and about **MADL**. Since **MADL** is included in every normal ad logic, these theorems hold in every ad system we consider in this essay.

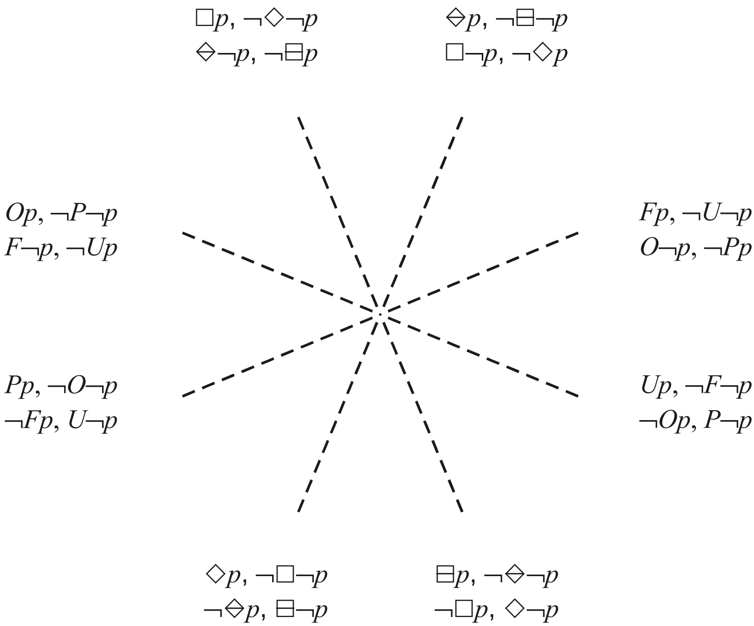

Figure 1. The Alethic-Deontic Octagon, MADL (S1).

4.1.1 The alethic-deontic octagon

It is possible to display some important logical relationships between O-, P-, F- and U-sentences in various deontic systems in a deontic square of opposition (see Rönnedal (2010)) and some important logical relationships between □-, ◇-, ⟠- and ⊟-sentences in various alethic systems in a similar alethic square of opposition. If we combine these figures we get something I will call the alethic-deontic octagon. This is a figure that can be used to represent some of the most important logical relationships between all primary deontic and alethic sentences, i.e. all of the following formulas: Op, Pp, Fp, Up, □p, ◇p, ⟠p and ⊟p. These relationships will vary from one ad system to another.

Figure 1 shows what the ad octagon looks like in **MADL**. All sentences that occur at a "node" in the figure are equivalent (e.g. $Op \leftrightarrow F\neg p$ and $\Diamond\neg p \leftrightarrow \neg\Box p$). Sentences that are connected via dashed lines are contradictories (e.g. $\neg(\Box p \leftrightarrow \Diamond\neg p)$, $\Box p \leftrightarrow \neg\Diamond\neg p$, $\neg(Pp \leftrightarrow Fp)$ and $Pp \leftrightarrow \neg Fp$ are theorems). Since **MADL** is the smallest ad system, these relationships hold in every ad system. However, all of the relationships displayed in this figure also hold in the union of **aK** and **dK**. So, the figure is perhaps more important for what it does not, than for what it does contain. The ad octagon will become more interesting in extensions of **MADL**.

4.1.2 The rule of replacement

The rule of replacement and the rule of simultaneous replacement hold in every normal ad system. The following section proves this. In our proofs we use the following derived rules: (OEQ) If $\vdash A \leftrightarrow B$, then $\vdash OA \leftrightarrow OB$, and (□EQ) If $\vdash A \leftrightarrow B$, then $\vdash \Box A \leftrightarrow \Box B$ (see Rönnedal (2010) for a proof of the first, the second can be established in a similar way). These rules are derivable in every normal ad system.

The rule of replacement (Rep). (i) If $\vdash A \leftrightarrow B$, then $\vdash C \leftrightarrow$ [B//A](C) (if A is equivalent to B is a theorem, then C is equivalent to [B//A](C) is a theorem), where [B//A](C) is like C except that zero or more occurrences of A are replaced by B (see Rönnedal (2010) for more information about the concept of replacement).

(ii) If $\vdash A \leftrightarrow B$ and $\vdash C$, then $\vdash$ [B//A](C) (if A is equivalent to B is a theorem and C is a theorem, then [B//A](C) is a theorem), where C and [B//A](C) are as in part (i).

(iii) If $\vdash A \leftrightarrow B$ and $\vdash$ [B//A](C), then $\vdash C$ (if A is equivalent to B is a theorem and [B//A](C) is a theorem, then C is a theorem), where C and [B//A](C) are as in part (i).

Proof: **Part (i)**. Suppose that the replacement of B for A is at zero places. Then [B//A](C) and C are identical and the result is trivial ($\vdash C \leftrightarrow$

[B//A](C), where [B//A](C) = C). Suppose that A and C is the same sentence and that C i.e. A is replaced by B. Then [B//A](C) is B. Hence (i) holds in this case too ($\vdash$ C $\leftrightarrow$ [B//A](C), where C = A and [B//A](C) = B). So, from now on we assume that A and C are distinct and that at least one occurrence of A is replaced by B in C. The rest of the proof is by induction on the length of A. Given $\vdash$ A $\leftrightarrow$ B.

Basis: C is atomic. Since C and A are distinct and C is atomic [B//A](C) = C. Hence, $\vdash$ C $\leftrightarrow$ [B//A](C), where [B//A](C) = C. Consequently the theorem holds when C is atomic.

Induction step. We want to show that if it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ C $\leftrightarrow$ [B//A](C) for C of any complexity, then it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ f(C) $\leftrightarrow$ f([B//A](C)), where f(C) is $\neg$C, D $\wedge$ C, C $\wedge$ D, D $\vee$ C, C $\vee$ D, D $\rightarrow$ C, C $\rightarrow$ D, D $\leftrightarrow$ C, C $\leftrightarrow$ D, OC, PC, FC, UC, KC, NC, $\Box$C, $\Diamond$C, ⬖C, $\boxminus$C, ∇C or $\triangle$C, and likewise for [B//A](C). Since conjunction, disjunction and equivalence are commutative, since equivalence, implication and conjunction can be expressed in terms of negation and disjunction, since P, F, U, K and N are definable in terms of O, and since $\Diamond$C, ⬖C, $\boxminus$C, ∇C and $\triangle$C are definable in terms of $\Box$, it is sufficient to consider four cases.

Case (i). $\neg$C. Suppose that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ C $\leftrightarrow$ [B//A](C). From the hypothesis $\vdash$ A $\leftrightarrow$ B. Hence, $\vdash$ C $\leftrightarrow$ [B//A](C). (C $\leftrightarrow$ [B//A](C)) $\leftrightarrow$ ($\neg$C $\leftrightarrow$ $\neg$[B//A](C)) is a tautology. Accordingly, $\vdash$ $\neg$C $\leftrightarrow$ $\neg$[B//A](C) by PL. It follows that if it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ C $\leftrightarrow$ [B//A](C), then it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ $\neg$C $\leftrightarrow$ $\neg$[B//A](C).

Case (ii). C $\vee$ D. Assume that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ C $\leftrightarrow$ [B//A](C). By the hypothesis $\vdash$ A $\leftrightarrow$ B. Thus, $\vdash$ C $\leftrightarrow$ [B//A](C). (C $\leftrightarrow$ [B//A](C)) $\leftrightarrow$ ((C $\vee$ D) $\leftrightarrow$ ([B//A](C) $\vee$ D)) is logically true in propositional logic. Hence, $\vdash$ (C $\vee$ D) $\leftrightarrow$ ([B//A](C) $\vee$ D) by PL. Consequently, if it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ C $\leftrightarrow$ [B//A](C), then it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ (C $\vee$ D) $\leftrightarrow$ ([B//A](C) $\vee$ D).

Case (iii). OC. Suppose that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ C $\leftrightarrow$ [B//A](C). By the hypothesis $\vdash$ A $\leftrightarrow$ B. Hence, $\vdash$ C $\leftrightarrow$ [B//A](C) and so, $\vdash$ OC $\leftrightarrow$ O[B//A](C) by (OEQ). In consequence, if it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ C $\leftrightarrow$ [B//A](C), then it is the case that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ OC $\leftrightarrow$ O[B//A](C).

Case (iv). $\Box$C. As in case (iii).

Conclusion. We have now shown that the rule of replacement holds where there are no connectives or operators outside A and B and that if it holds where there are n such logical connectives or operators it holds for n + 1. We conclude that the theorem holds in general.

Part (ii). Assume that (1) $\vdash$ A$\leftrightarrow$B and $\vdash$ C (A is equivalent to B is a theorem and C is a theorem) and that C and [B//A](C) are as in part (i). Then both (2) $\vdash$ A$\leftrightarrow$B and (3) $\vdash$ C [from (1)]. From (2) we obtain (4) $\vdash$ C$\leftrightarrow$ [B//A](C) [by PL and part (i)]. Hence, (5) $\vdash$ [B//A](C) [from 3 and 4 by PL]. Consequently, (6) if $\vdash$ A$\leftrightarrow$B and $\vdash$ C, then $\vdash$ [B//A](C) (if A is equivalent to B is a theorem and C is a theorem, then [B//A](C) is a theorem), where C and [B//A](C) are as in part (i) [from 1–5 by conditional proof discharging the assumption].

Part (iii). As in part (ii). Details are left to the reader. ■

The rule of simultaneous replacement. (i) If $\vdash$ $A_1 \leftrightarrow B_1$ and … and $\vdash$ $A_n \leftrightarrow B_n$ then $\vdash$ C$\leftrightarrow$[B_1//A_1, …, B_n//A_n](C) (if A_1 is equivalent to B_1 and … and A_n is equivalent to B_n are theorems, then C is equivalent to [B_1//A_1, …, B_n//A_n](C) is a theorem), where [B_1//A_1, …, B_n//A_n](C) is the result of replacing zero or more occurrences of A_1 in C by B_1 and … and replacing zero or more occurrences of A_n in C by B_n.

(ii) If $\vdash$ $A_1 \leftrightarrow B_1$ and … and $\vdash$ $A_n \leftrightarrow B_n$ and $\vdash$ C, then $\vdash$ [B_1//A_1, …, B_n//A_n](C) (if A_1 is equivalent to B_1 and … and A_n is equivalent to B_n are theorems and C is a theorem, then [B_1//A_1, …, B_n//A_n](C) is a theorem), where C and [B_1//A_1, …, B_n//A_n](C) are as in part (i).

(iii) If $\vdash$ $A_1 \leftrightarrow B_1$ and … and $\vdash$ $A_n \leftrightarrow B_n$ and $\vdash$ C$\leftrightarrow$[B_1//A_1, …, B_n//A_n](C), then $\vdash$ C (if A_1 is equivalent to B_1 and … and A_n is equivalent to B_n are theorems and [B_1//A_1, …, B_n//A_n](C) is a theorem, then C is a theorem), where C and [B_1//A_1, …, B_n//A_n](C) are as in part (i).

Proof. The proof is more or less obvious, simply use the rule of replacement repeatedly in crucial steps. ■

4.1.3 Interchange and duality theorems

Let us prove some interchange and duality theorems that can be used to quickly prove and recognize new theorems in **MADL** and other ad systems.

Theorem 6 (The ad interchange theorem (adIT)). Let $\otimes_1$ ... $\otimes_n$ be a sequence of deontic and alethic operators in a sentence such that each $\otimes_i$ is O, P, F, U, □, ◇, ⬘ or ⊟. If $\otimes_i$ = O, let $\otimes_i$’ = P, and vice versa, if ⊗ = F, let ⊗’ = U and vice versa, if ⊗ = □, let ⊗’ = ◇ and vice versa and if ⊗ = ⬘, let ⊗’ = ⊟ and vice versa, for every $\otimes_i$. Then, (i) $\vdash \otimes_1 \ldots \otimes_n A \leftrightarrow \neg\otimes_1' \ldots \otimes_n'\neg A$, (ii) $\vdash \neg\otimes_1 \ldots \otimes_n A \leftrightarrow \otimes_1' \ldots \otimes_n'\neg A$, and (iii) $\vdash \otimes_1 \ldots \otimes_n\neg A \leftrightarrow \neg\otimes_1' \ldots \otimes_n' A$, for any A.

Proof. **Part (i).** Let $\otimes_1 \dots \otimes_n$ be a sequence of deontic and alethic operators of the kind mentioned in the theorem. By PL $\vdash \otimes_1 \dots \otimes_n A \leftrightarrow \otimes_1 \dots \otimes_n A$, for any A. Now replace O by $\neg P \neg$, P by $\neg O \neg$, F by $\neg U \neg$, U by $\neg F \neg$, $\Box$ by $\neg \Diamond \neg$, $\Diamond$ by $\neg \Box \neg$, $\diamondminus$ by $\neg \boxminus \neg$ and $\boxminus$ by $\neg \diamondminus \neg$ throughout in the right hand side of this equivalence. Then we get the following theorem $\vdash \otimes_1 \dots \otimes_n A \leftrightarrow \neg \otimes_1 ' \neg \neg \otimes_2 ' \neg \dots \neg \otimes_{n-1} ' \neg \neg \otimes_n ' \neg A$, for any A. Use PL and replacement to delete all double negations. It follows that $\vdash \otimes_1 \dots \otimes_n A \leftrightarrow \neg \otimes_1 ' \dots \otimes_n ' \neg A$, for any A. This proves part (i).

Part (ii). PL and part (i) gives us $\vdash \neg \otimes_1 \dots \otimes_n A \leftrightarrow \neg \neg \otimes_1 ' \dots \otimes_n ' \neg A$, which again by PL immediately proves that $\vdash \neg \otimes_1 \dots \otimes_n A \leftrightarrow \otimes_1 ' \dots \otimes_n ' \neg A$, for any A.

Part (iii). By replacing A by $\neg A$ in part (i) we obtain $\vdash \otimes_1 \dots \otimes_n \neg A \leftrightarrow \neg \otimes_1 ' \dots \otimes_n ' \neg \neg A$. By PL and replacement it follows that $\vdash \otimes_1 \dots \otimes_n \neg A \leftrightarrow \neg \otimes_1 ' \dots \otimes_n ' A$, for any A. ■

Example 7. When n = 0 part (i) reduces to $\vdash A \leftrightarrow \neg \neg A$, part (ii) to $\vdash \neg A \leftrightarrow \neg A$ and part (iii) to $\vdash \neg A \leftrightarrow \neg A$. Let n = 1. Then the following schemas are examples of instances of the theorem $\vdash OA \leftrightarrow \neg P \neg A$, $\vdash PA \leftrightarrow \neg O \neg A$, $\vdash \Box A \leftrightarrow \neg \Diamond \neg A$, $\vdash \Diamond A \leftrightarrow \neg \Box \neg A$ [part (i)], $\vdash \neg OA \leftrightarrow P \neg A$, $\vdash \neg PA \leftrightarrow O \neg A$, $\vdash \neg \boxminus A \leftrightarrow \diamondminus \neg A$, $\vdash \neg \diamondminus A \leftrightarrow \boxminus \neg A$ [part (ii)], and $\vdash O \neg A \leftrightarrow \neg PA$, $\vdash P \neg A \leftrightarrow \neg OA$ [part (iii)]. In fact, many equivalences in figure 1 can be seen as special cases of adIT. Here are some more complex examples: $\vdash \Box O(p \rightarrow p) \leftrightarrow \neg \Diamond P \neg (p \rightarrow p)$, $\vdash \neg \Box P(p \wedge q) \leftrightarrow \Diamond O \neg (p \wedge q)$, $\vdash \boxminus \diamondminus F \neg (p \rightarrow (Pq \wedge Pr)) \leftrightarrow \neg \diamondminus \boxminus U(p \rightarrow (Pq \wedge Pr))$.

Theorem 8 (The ad interchange rule (adIR)). All of the following rules are derived in **MADL**. Let $\otimes$ and $\otimes'$ be as in adIT. Then, (i) $\vdash \otimes_1 \dots \otimes_n A$ iff $\vdash \neg \otimes_1 ' \dots \otimes_n ' \neg A$, (ii) $\vdash \neg \otimes_1 \dots \otimes_n A$ iff $\vdash \otimes_1 ' \dots \otimes_n ' \neg A$, and (iii) $\vdash \otimes_1 \dots \otimes_n \neg A$ iff $\vdash \neg \otimes_1 ' \dots \otimes_n ' A$, for any A.

Proof. The proofs are easy and are left to the reader (use the interchange theorem). ■

We can in fact prove something slightly stronger. The interchange theorem does not hold for modalities that include embedded negation signs. But our next theorem does.

Before turning to the duality theorem, we must first introduce the concept of duality.

Definition 9 (Duality). (i) Let L be a language that contains $\neg$, $\wedge$ and $\vee$ as the only propositional connectives. In addition, let L contain Verum and Falsum, and all normal alethic and deontic operators, i.e. O, P, F, U, $\Box$, $\Diamond$, $\diamondminus$, and $\boxminus$. Then *the dual of a sentence A* (in L), in symbols d(A), is defined as the result of replacing every atomic sentence by its negation and interchanging all occurrences of Verum and Falsum, $\wedge$ and $\vee$, $\Box$ and $\Diamond$, O and P, $\diamondminus$ and $\boxminus$, and F and U in A. (ii) Let L by a language that contains $\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$, Verum, Falsum and every normal alethic and deontic operator. Then if A is a sentence (in L), then *the dual of A*, in symbols d(A), is defined in the following manner.

1 $d(p) = \neg p$, when p is atomic	6 $d(A \vee B) = (d(A) \wedge d(B))$	11 $d(OA) = Pd(A)$
2 $d(\text{Verum}) = \text{Falsum}$	7 $d(A \rightarrow B) = (\neg d(A) \wedge d(B))$	12 $d(PA) = Od(A)$
3 $d(\text{Falsum}) = \text{Verum}$	8 $d(A \leftrightarrow B) = (d(A) \leftrightarrow \neg d(B))$	13 $d(\diamondminus A) = \boxminus d(A)$
4 $d(\neg A) = \neg d(A)$	9 $d(\Box A) = \Diamond d(A)$	14 $d(\boxminus A) = \diamondminus d(A)$
5 $d(A \wedge B) = (d(A) \vee d(B))$	10 $d(\Diamond A) = \Box d(A)$	15 $d(FA) = Ud(A)$
		16 $d(UA) = Fd(A)$

Example 10. (i) $d(O\top) = P\bot$, (ii) $d(p \wedge q) = (\neg p \vee \neg q)$ (iii) $d(Op \rightarrow Pp) = (\neg P \neg p \wedge O \neg p)$, (iv) $d(Op \rightarrow \Diamond p) = (\neg P \neg p \wedge \Box \neg p)$, (v) $d(Pp \leftrightarrow \neg Fp) = (O \neg p \leftrightarrow \neg\neg U \neg p)$, (vi) $d(\Box(p \rightarrow q) \rightarrow (Op \rightarrow Oq)) = (\neg \Diamond(\neg\neg p \wedge \neg q) \wedge (\neg P \neg p \wedge P \neg q))$, (vii) $d(P(p \wedge q) \rightarrow \Diamond(p \vee q)) = (\neg O(\neg p \vee \neg q) \wedge \Box(\neg p \wedge \neg q))$, (viii) $d(((Fq \wedge Fr) \wedge \Box(p \rightarrow (q \vee r))) \rightarrow Fp) = (\neg((U \neg q \vee U \neg r) \vee \Diamond(\neg\neg p \wedge (\neg q \wedge \neg r))) \wedge U \neg p)$, (ix) $d(Up \rightarrow \boxminus p) = (\neg F \neg p \wedge \diamondminus \neg p)$, (x) $d((O(p \vee q) \wedge \diamondminus p) \rightarrow Pq) = (\neg(P(\neg p \wedge \neg q) \vee \boxminus \neg p) \wedge O \neg q)$.

Proof. We prove part (vi) and (vii) and leave the rest to the reader.

Part (vi). $d(\Box(p \rightarrow q) \rightarrow (Op \rightarrow Oq)) = (\neg d(\Box(p \rightarrow q)) \wedge d(Op \rightarrow Oq))$ [part 7] $= (\neg \Diamond d(p \rightarrow q) \wedge d(Op \rightarrow Oq))$ [part 9] $= (\neg \Diamond(\neg d(p) \wedge d(q)) \wedge d(Op \rightarrow Oq))$ [part 7] $= (\neg \Diamond(\neg d(p) \wedge d(q)) \wedge (\neg d(Op) \wedge d(Oq)))$ [part 7] $= (\neg \Diamond(\neg d(p) \wedge d(q)) \wedge (\neg Pd(p) \wedge Pd(q)))$ [part 11] $= (\neg \Diamond(\neg\neg p \wedge \neg q) \wedge (\neg P \neg p \wedge P \neg q))$ [part 1].

Part (vii). $d(P(p \wedge q) \rightarrow \Diamond(p \vee q)) = (\neg d(P(p \wedge q)) \wedge d(\Diamond(p \vee q)))$ [part 7] $= (\neg Od(p \wedge q) \wedge d(\Diamond(p \vee q)))$ [part 12] $= (\neg O(d(p) \vee d(q)) \wedge d(\Diamond(p \vee q)))$ [part 5] $= (\neg O(d(p) \vee d(q)) \wedge \Box d(p \vee q))$ [part 10] $= (\neg O(d(p) \vee d(q)) \wedge \Box(d(p) \wedge d(q))$ [part 6] $= (\neg O(\neg p \vee \neg q) \wedge \Box(\neg p \wedge \neg q))$ [part 1]. ■

Theorem 11 (The duality theorem (DUAL)). Let **S** be a normal alethic-deontic system. Then **S** has the following theorems and rules of inference.

(i) $\vdash_S A \leftrightarrow \neg d(A)$.
(ii) $\vdash_S \neg A \leftrightarrow d(A)$.
(iii) if $\vdash_S A$, then $\vdash_S \neg d(A)$.
(iv) if $\vdash_S \neg A$, then $\vdash_S d(A)$.
(v) if $\vdash_S A \rightarrow B$, then $\vdash_S d(B) \rightarrow d(A)$.
(vi) if $\vdash_S A \leftrightarrow B$, then $\vdash_S d(A) \leftrightarrow d(B)$.

Proof. Assume throughout that **S** is a normal alethic-deontic system.

Part (i) $\vdash_S A \leftrightarrow \neg d(A)$. Part (i) says that A and the negation of the duality of A are equivalent in S. We want to show that the theorem holds for any sentence regardless of complexity. We prove this by induction on the lenght of A.

Basis. A is atomic. (1) $\vdash p \leftrightarrow \neg\neg p$, for every atomic sentence p [by PL]. Hence, (2) $\vdash p \leftrightarrow \neg d(p)$, for every atomic sentence p [from 1 and the definition of duality part 1]. Consequently, the theorem holds when A is atomic.

Induction step. We want to show that if the theorem holds for a sentence A of given complexity, it holds for every sentence of next higher degree of complexity. Induction hypotheses: the theorem holds for every sentence B and C shorter than A, i.e. $\vdash B \leftrightarrow \neg d(B)$ and $\vdash C \leftrightarrow \neg d(C)$. $A = B \wedge C$, $A = B \rightarrow C$, $A = B \leftrightarrow C$, $A = \boxminus B$, and $A = UB$. Left as exercise.

$A = \neg B$. $\vdash \neg B \leftrightarrow \neg\neg d(B)$ [by the induction hypothesis and PL]. Consequently, $\vdash \neg B \leftrightarrow \neg d(\neg B)$ [by the definition of duality part 4].

$A = B \vee C$. (1) $\vdash (B \vee C) \leftrightarrow (\neg d(B) \vee \neg d(C))$ [induction hypothesis and replacement]. (2) $\vdash (\neg d(B) \vee \neg d(C)) \leftrightarrow \neg(d(B) \wedge d(C))$ [by PL]. (3) $\vdash \neg(d(B) \wedge d(C)) \leftrightarrow \neg d(B \vee C)$ [by the definition of duality part 6 and replacement]. (4) $\vdash (B \vee C) \leftrightarrow \neg d(B \vee C)$ [from 1, 2 and 3 by PL].

$A = \Box B$. This is exactly as in the case $A = OB$ (see below), just replace every occurrence of O by $\Box$ throughout and replace the justification for step (3) by "the definition of duality part 9".

$A = \Diamond B$. (1) $\vdash \Diamond B \leftrightarrow \Diamond\neg d(B)$ [induction hypothesis, replacement]. (2) $\vdash \Diamond\neg d(B) \leftrightarrow \neg\Box d(B)$ [definition of $\Diamond$ and replacement]. (3) $\vdash \neg\Box d(B) \leftrightarrow \neg d(\Diamond B)$ [by the definition of duality part 10]. It follows that $\vdash \Diamond B \leftrightarrow \neg d(\Diamond B)$ [from 1, 2 and 3 by PL].

$A = \diamondminus B$. Similar to the case where $A = FB$ (see below).

$A = OB$. (1) $\vdash OB \leftrightarrow O\neg d(B)$ [induction hypothesis, replacement]. (2) $\vdash O\neg d(B) \leftrightarrow \neg Pd(B)$ [definition of P, PL]. (3) $\vdash \neg Pd(B) \leftrightarrow \neg d(OB)$ [by the definition of duality part 11]. Thus, (4) $\vdash OB \leftrightarrow \neg d(OB)$ [from 1, 2 and 3 by PL].

$A = PB$. This is exactly as in the case $A = \Diamond B$ (see above), just replace every occurrence of $\Diamond$ by P throughout and replace the justification for step (3) by "the definition of duality part 9".

$A = FB$. (1) $\vdash FB \leftrightarrow F\neg d(B)$ [induction hypothesis, replacement]. (2) $\vdash F\neg d(B) \leftrightarrow \neg Ud(B)$ [interchange]. (3) $\vdash \neg Ud(B) \leftrightarrow \neg d(FB)$ [the definition of duality part 15]. Thus, (4) $\vdash$ FB $\leftrightarrow$ ¬d(FB) [from 1, 2 and 3 by PL].

Part (ii) $\vdash_S$ ¬A $\leftrightarrow$ d(A). Part (ii) follows immediately from part (i) by PL. The interpretation is similar.

Part (iii) if $\vdash_S$ A, then $\vdash_S$ ¬d(A). According to part (iii) the negation of the duality of A is a theorem in S, if A is a theorem in S. Suppose that $\vdash$ A. Then by part (i) and (MP) it follows that $\vdash$ ¬d(A). So, if $\vdash$ A, then $\vdash$ ¬d(A) [by conditional proof].

Part (iv) if $\vdash_S$ ¬A, then $\vdash_S$ d(A). This part is proved as part (iii), but use part (ii) instead of part (i) in the proof. It is interpreted similarly.

Part (v) if $\vdash_S$ A $\rightarrow$ B, then $\vdash_S$ d(B) $\rightarrow$ d(A). Part (v) says that if A implies B is a theorem, then the duality of B implies the duality of A is a theorem. Suppose (1) $\vdash$ A $\rightarrow$ B. Then by part (i) and replacement we get (2) $\vdash$ ¬d(A) $\rightarrow$ ¬d(B) [from 1]. Accordingly, (3) $\vdash$ d(B) $\rightarrow$ d(A) [from 2 by PL]. It follows that if $\vdash$ A $\rightarrow$ B, then $\vdash$ d(B) $\rightarrow$ d(A) [by conditional proof from 1 – 3].

Part (vi) if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ d(A) $\leftrightarrow$ d(B). If it is a theorem that A is equivalent to B, then it is a theorem that the duality of A is equivalent to the duality of B, according to this part. Suppose $\vdash$ A $\leftrightarrow$ B. Then (2) $\vdash$ A $\rightarrow$ B and (3) $\vdash$ B $\rightarrow$ A [from 1 by PL]. (4) $\vdash$ d(B) $\rightarrow$ d(A) [from 2 and part (v)]. (5) $\vdash$ d(A) $\rightarrow$ d(B) [from 3 and part (v)]. Hence, (6) $\vdash$ d(A) $\leftrightarrow$ d(B) [from 4 and 5 by PL]. It follows that if $\vdash$ A $\leftrightarrow$ B, then $\vdash$ d(A) $\leftrightarrow$ d(B) [by conditional proof from 1 – 5]. ■

Theorem 12 (The duality corollary (Dual)). The *dual of an alethic-deontic modality* M, D(M), is the modality that is obtained from M by interchanging O and P, F and U, □ and ◇, and ⊟ and ⬙, respectively, throughout. Let M and N be alethic-deontic OPFU□◇⊟⬙ modalities and D(M) and D(N) be the dual of M and N respectively. Then:

Part (i) $\vdash$ MA$\leftrightarrow$¬D(M)¬A.
Part (ii) $\vdash$ MA iff $\vdash$ ¬D(M)¬A.
Part (iii) $\vdash$ MA$\rightarrow$NA iff $\vdash$ D(N)A$\rightarrow$D(M)A.
Part (iv) $\vdash$ MA$\leftrightarrow$NA iff $\vdash$ D(M)A$\leftrightarrow$D(N)A.

Proof. **Part (i).** $\vdash$ MA$\leftrightarrow$¬D(M)¬A. (1) $\vdash$ MA$\leftrightarrow$¬d(MA) [Dual]. (2) $\vdash$ ¬d(MA)$\leftrightarrow$¬D(M)dA [PL, the definition of duality]. (3) $\vdash$ ¬D(M)dA $\leftrightarrow$¬D(M)¬A [PL, Dual, replacement]. (4) $\vdash$ MA$\leftrightarrow$¬D(M)¬A [from 1, 2 and 3 by PL].

Part (ii). (1) $\vdash$ MA $\rightarrow$¬D(M)¬A [from (i) and PL]. Suppose (2) $\vdash$ MA. Then (3) $\vdash$ ¬D(M)¬A [from 1 and 2 by (MP)]. Hence, (4) if $\vdash$ MA, then $\vdash \neg D(M)\neg A$ [from 2 – 3 by conditional proof]. (5) $\vdash \neg D(M)\neg A \rightarrow$ MA [from (i) and PL]. Suppose (6) $\vdash \neg D(M)\neg A$. Then (7) $\vdash$ MA [from 5 and 6 by (MP)]. So, (8) if $\vdash \neg D(M)\neg A$, then $\vdash$ MA [from 6 – 7 by conditional proof]. It follows that $\vdash$ MA iff $\vdash \neg D(M)\neg A$ [from 4 and 8 by classical logic].

Part (iii). (1) $\vdash MA \rightarrow NA$ iff (2) $\vdash \neg D(M)\neg A \rightarrow \neg D(N)\neg A$ [from 1 by replacement] iff (3) $\vdash D(N)\neg A \rightarrow D(M)\neg A$ [from 2 by PL] iff (4) $\vdash D(N)A \rightarrow D(M)A$ [from 3 by PL and replacement]; in conclusion, (5) $\vdash MA \rightarrow NA$ iff $\vdash D(N)A \rightarrow D(M)A$ [from 1 – 4 by classical logic].

Part (iv). Suppose (1) $\vdash MA \leftrightarrow NA$. Then (2) $\vdash MA \rightarrow NA$ [from 1 by PL] and (3) $\vdash NA \rightarrow MA$ [from 1 by PL]. (4) $\vdash D(N)A \rightarrow D(M)A$ [from 2 and part (iii)]. (5) $\vdash D(M)A \rightarrow D(N)A$ [from 3 and part (iii)]. (6) $\vdash D(M)A \leftrightarrow D(N)A$ [from 4 and 5 by PL]. Consequently, (7) if $\vdash MA \leftrightarrow NA$ then $\vdash D(M)A \leftrightarrow D(N)A$ [by conditional proof from 1 – 6]. Suppose (8) $\vdash D(M)A \leftrightarrow D(N)A$. Then (9) $\vdash D(M)A \rightarrow D(N)A$ [from 8 by PL] and (10) $\vdash D(N)A \rightarrow D(M)A$ [from 8 by PL]. (11) $\vdash NA \rightarrow MA$ [from 9 and part (iii)]. (12) $\vdash MA \rightarrow NA$ [from 10 and part (iii)]. (13) $\vdash MA \leftrightarrow NA$ [11, 12, PL]. Consequently, (14) if $\vdash D(M)A \leftrightarrow D(N)A$, then $\vdash MA \leftrightarrow NA$ [from 8 – 13 by conditional proof]. It follows that (15) $\vdash MA \leftrightarrow NA$ iff $\vdash D(M)A \leftrightarrow D(N)A$ [from 7 and 14 by classical logic]. ■

Comment 13. Note that both the duality theorem and the duality corollary are abbreviated "Dual". When any theorem or any rule that is part of one of these propositions is used, we will indicate this by writing "Dual" in the justificatory entry.

The following theorem illustrates how the duality corollary can be used.

Theorem	**"Dual" theorem**	**Theorem**	**"Dual" theorem**
1 $Op \rightarrow OOp$	T(1) $PPp \rightarrow Pp$	7 $\Box p \rightarrow Op$	T(7) $Pp \rightarrow \Diamond p$
2 $Pp \rightarrow OPp$	T(2) $POp \rightarrow Op$	8 $Op \rightarrow \Diamond p$	T(8) $\Box p \rightarrow Pp$
3 $\Box p \rightarrow p$	T(3) $p \rightarrow \Diamond p$	9 $Op \rightarrow \Box Op$	T(9) $\Diamond Pp \rightarrow Pp$
4 $\Box p \rightarrow \Box\Box p$	T(4) $\Diamond\Diamond p \rightarrow \Diamond p$	10 $Pp \rightarrow \Box Pp$	T(10) $\Diamond Op \rightarrow Op$
5 $p \rightarrow \Box\Diamond p$	T(5) $\Diamond\Box p \rightarrow p$	11 $O\Box p \rightarrow \Box Op$	T(11) $\Diamond Pp \rightarrow P\Diamond p$
6 $\Diamond p \rightarrow \Box\Diamond p$	T(6) $\Diamond\Box p \rightarrow \Box p$	12 $\Box Op \rightarrow O\Box p$	T(12) $P\Diamond p \rightarrow \Diamond Pp$

Table 4

Theorem 14. *Let* **S** *be a normal alethic-deontic system. Then* n *is a theorem in* **S** *if and only if* T(n) *is a theorem in* **S** (*for* $1 \leq n \leq 12$ *in table 4*).

Proof. This follows immediately from the duality corollary part (iii). In every case n has the form $MA \rightarrow NA$ and T(n) the form $D(N)A \rightarrow D(M)A$. ■

Comment 15. In the table 4 I have called T(n) "dual" theorems since T(n) may be derived from n (for $1 \leq n \leq 12$) by Dual. However, in a strict sense T(n) is of course not the dual of n.

4.1.4 More rules of inference in MADL

I will end this section by proving a set of new derived rules that are admissible in every ad system (theorem 16). First, we will introduce some new concepts.

All rules that can be derived in **dK** and in **aK** also hold in **MADL**. Some of these rules have a similar form as is easy to see. Let ⊗ be any of the following operators: O, P, □ or ◇. Then every rule of the following kind holds in **MADL**: if $\vdash A \rightarrow B$, then $\vdash \otimes A \rightarrow \otimes B$. Let us call a rule of this kind a *monotic rule of type I* (a MI rule). Let ⊗ be any of the following operators: F, U, ⬖ or ⊟. Then every rule of the following kind holds in **MADL**: if $\vdash A \rightarrow B$, then $\vdash \otimes B \rightarrow \otimes A$. We shall say that a rule of this kind is a *monotonic rule of type II* (or a converse monotonic rule) (a MII rule). Both type I and type II rules are called *monotonic*.

Theorem 16 (The inference rule theorem I). Let **S** be a normal ad system and let M and N be affirmative OP□◇ modalities. By an *affirmative* OP□◇ *modality* we mean a modality, i.e. a finite sequence, possibly empty, of the operators ¬, O, P, □ and ◇, in which ¬ occurs an even number of times (including zero). Then the following sentence is a theorem in **S**: A = MA → NA if and only if **S** has any of the following theorems or rules of inference: A' = D(N) → D(M)A, (R1) if $\vdash_S A \rightarrow B$ then $\vdash_S MA \rightarrow NB$, or (R2) if $\vdash_S A \rightarrow B$, then $\vdash_S D(N)A \rightarrow D(M)B$.

Proof. We assume throughout that **S** is a normal ad system. To prove this theorem it is sufficient to establish how to obtain (i) A' from A, (ii) A from A', (iii) (R1) from A, (iv) A from (R1), (v) (R2) from (R1), and (vi) (R1) from (R2), in **S**. (ConP = Conditional Proof.)

Part (i) and **part (ii)** follow directly from the duality corollary.

Part (iii). From A to (R1). We assume that **S** includes MA → NA and then show (R1): if $\vdash_S A \rightarrow B$, then $\vdash_S MA \rightarrow NB$.

1.	$\vdash_S A \rightarrow B$	[Assumption]
2.	$\vdash_S NA \rightarrow NB$	[1, Repeated applications of MI rules]
3.	$\vdash_S MA \rightarrow NA$	[Given]
4.	$\vdash_S MA \rightarrow NB$	[2, 3, PL]
5.	If $\vdash_S A \rightarrow B$, then $\vdash_S MA \rightarrow NB$.	[ConP 1–4]

Part (iv). From (R1) to A. We assume that (R1), if $\vdash_S A \rightarrow B$ then $\vdash_S MA \rightarrow NB$, is a rule of inference in **S** and then prove A: $MA \rightarrow NA$.

1. $\vdash_S A \rightarrow A$ [PL]
2. $\vdash_S MA \rightarrow NA$ [1, R1]

Part (v). From (R1) to (R2). Suppose that (R1), if $\vdash_S A \rightarrow B$ then $\vdash_S MA \rightarrow NB$, is a rule of inference in **S**. We must prove that (R2), if $\vdash_S A \rightarrow B$ then $\vdash_S D(N)A \rightarrow D(M)B$, is a rule of inference in **S** too.

1. $\vdash_S A \rightarrow B$ [Assumption]
2. $\vdash_S A \rightarrow A$ [PL]
3. $\vdash_S MA \rightarrow NA$ [2, R1]
4. $\vdash_S D(N)A \rightarrow D(M)A$ [3, Dual corollary]
5. $\vdash_S D(M)A \rightarrow D(M)B$ [1, Repeated applications of MI rules]
6. $\vdash_S D(N)A \rightarrow D(M)B$ [4, 5, PL]
7. If $\vdash_S A \rightarrow B$, then $\vdash_S D(N)A \rightarrow D(M)B$ [1–6, ConP]

Part (vi). From (R2) to (R1). We suppose that **S** includes (R2), if $\vdash_S A \rightarrow B$ then $\vdash_S D(N)A \rightarrow D(M)B$, and then show that (R1), if $\vdash_S A \rightarrow B$ then $\vdash_S MA \rightarrow NB$, is included in **S** too.

1. $\vdash_S A \rightarrow B$ [Assumption]
2. $\vdash_S A \rightarrow A$ [PL]
3. $\vdash_S D(N)A \rightarrow D(M)A$ [2, R2]
4. $\vdash_S MA \rightarrow NA$ [3, Dual corollary]
5. $\vdash_S NA \rightarrow NB$ [1, Repeated applications of MI rules]
6. $\vdash_S MA \rightarrow NB$ [4, 5, PL]
7. If $\vdash_S A \rightarrow B$, then $\vdash_S MA \rightarrow NB$ [1–6, ConP] ■

Example 17. The following examples are consequences or instances of theorem 16. (i) If **S** includes $OA \rightarrow \Diamond A$, then: if $\vdash_S A \rightarrow B$, then $\vdash_S OA \rightarrow \Diamond B$. (ii) If **S** includes $\Box A \rightarrow OA$, then: if $\vdash_S A \rightarrow B$, then $\vdash_S \Box A \rightarrow OB$. (iii) If **S** includes $OA \rightarrow \Diamond A$, then: if $\vdash_S A \rightarrow B$, then $\vdash_S \Box A \rightarrow PB$. (iv) If **S** includes $\Box A \rightarrow OA$, then: if $\vdash_S A \rightarrow B$, then $\vdash_S PA \rightarrow \Diamond B$. (See sections 4.4.3 and 4.5.3.) (v) If **S** includes $OA \rightarrow \Box OA$, then: if $\vdash_S A \rightarrow B$, then $\vdash_S OA \rightarrow \Box OB$. (vi) If **S** includes $PA \rightarrow \Box PA$, then: if $\vdash_S A \rightarrow B$, then $\vdash_S PA \rightarrow \Box PB$.

We will now begin to consider some extensions of **MADL**.

4.2 aKDdKad∅

aKDdKad∅, the smallest normal alethic-deontic logic that includes the axiom aD (i.e. the sentence $\Box p \rightarrow \Diamond p$), is the same system as **MADL** + {aD}. We will also call this system **S2**. Like every normal alethic-deontic system **aKDdKad∅** includes PL, the axioms aK and dK, the usual definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation. Every normal alethic-deontic system that includes aD is a normal **aKDdKad∅**-system. In other words, any normal alethic-deontic system that is an extension of **aKDdKad∅** is a normal **aKDdKad∅**-system. Since **aKDdKad∅** does not contain any mixed axioms, any axioms that contain both alethic and deontic operators, it is an ad combination. More precisely, it is an ad combination of the purely alethic system **aKD** and the smallest normal deontic system **OK**.

We will now consider what the ad octagon looks like in this system.

4.2.1 The alethic-deontic octagon

Figure 2 is a picture of the alethic-deontic octagon in **aKDdKad∅**. It is interpreted in the same way as the ad octagon in section 4.1.

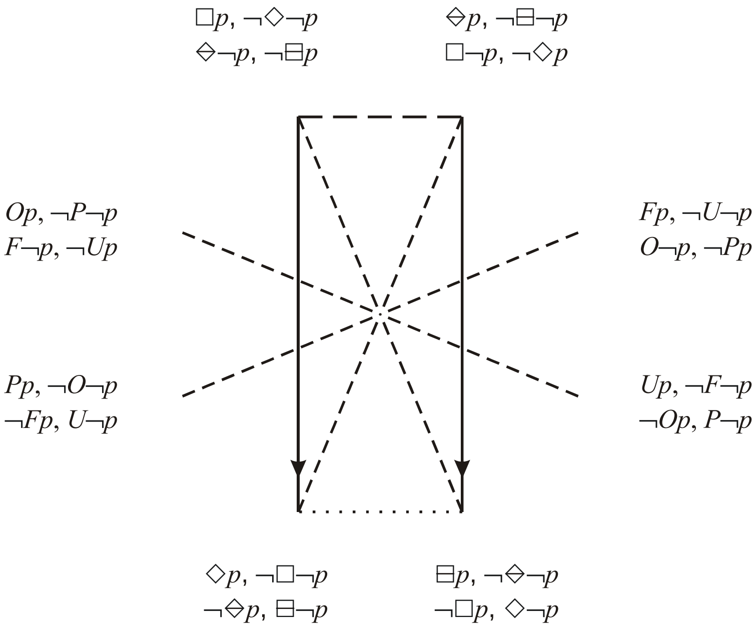

Figure 2. The Alethic-Deontic Octagon, MADL + {aD} **(S2).**

A dashed line connects sentences that are contradictories, e.g. Op and P¬p, and ⟠p and ◇p. An arrow from a sentence, A, to another sentence, B, indicates that A implies B; e.g. □p→¬□¬p and ⟠p→◇¬p are theorems in **MADL** + {aD}. A dotted line between two sentences, A and B, represents the fact that A and B are subcontraries, e.g. we can prove that ◇p∨◇¬p and ¬⟠p∨¬□p are theorems in the current system. Finally, a dotted line with long dots between two sentences, A and B, indicates that A and B are contraries, for instance □p and ⟠p, and ⟠¬p and ¬◇p; i.e. ¬(□p∧⟠p) and ¬(⟠¬p∧¬◇p) are theorems in **MADL** + {aD} (see Rönnedal (2010) for more on these concepts). We state this result formally.

Theorem 18. *All of the relationships displayed in figure 2 hold in every normal alethic-deontic* **aKDdKad∅***-system.*

Proof. This follows immediately from the fact that **aKDdKad∅** includes **OK** and **aKD**. ■

Remark 19. Note that a **aKDdKad∅**-system that is a proper extension of **aKDdKad∅** may contain more relations than those displayed in figure 2. The system **aKDdKDad∅** is, for instance, a proper extension of **aKDdKad∅**. The ad octagon for this system is displayed in figure 7. As can be seen, this system includes e.g. the sentences Op → Pp and Fp → P¬p, which are not theorems in **aKDdKad∅**. However, no **aKDdKad∅**-system lacks any of the theorems indicated in figure 2. Similar remarks apply to several other theorems involving the ad octagon stated in this paper.

4.3 aKdKDad∅

aKdKDad∅ is another example of an ad combination. It is identical to **aKdSDLad∅**, i.e. to the ad combination of the purely alethic system **aK** (see Chellas (1980)) and the purely deontic system Standard deontic logic (**SDL**) (see Rönnedal (2010)). In other words, **aKdKDad∅** is the smallest normal alethic-deontic logic that includes the axiom dD, i.e. the sentence Op→Pp. Accordingly, **aKdKDad∅** = **MADL** + {dD}. We will also call this system **S3**. Since it is a normal alethic-deontic system **aKdKDad∅** includes PL, the axioms aK and dK, the usual definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation. A normal **aKdKDad∅**-system is any normal alethic-deontic system that includes dD, or in other words, any normal alethic-deontic system that is an extension of **aKdKDad∅**.

Let us consider some properties of this system.

4.3.1 The alethic-deontic octagon

The alethic-deontic octagon in **aKdKDad∅** is similar to the ad octagon in **aKDdKad∅**. The differences are due to the fact that **aKdKDad∅** includes dD while **aKDdKad∅** includes aD. The similarities depend upon the formal similarities between these axioms.

Theorem 20. *All of the relationships displayed in figure 3 hold in every normal alethic-deontic* **aKdKDad∅***-system.*

Proof. This follows immediately from the fact that **aKdKDad∅** includes **SDL** and **aK**. ■

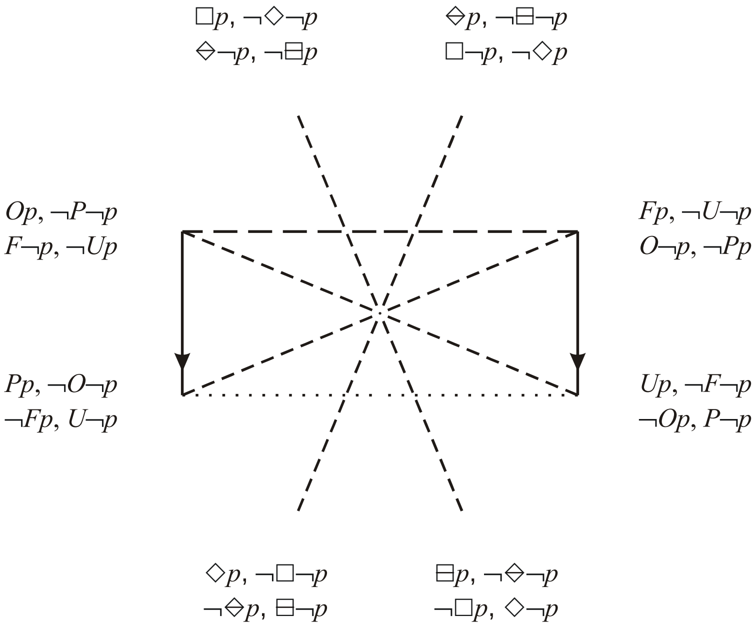

Figure 3. The Alethic-Deontic Octagon, MADL + {dD} **(S3).**

Next we turn to two ad systems that include mixed axioms: **aKdKadOC** and **aKdKadMO**. OC and MO are two of the most interesting mixed axioms we will consider.

4.4 aKdKadOC

aKdKadOC is the smallest normal alethic-deontic logic that contains the axiom OC, i.e. the sentence Op → ◇p. Therefore, **aKdKadOC** = **MADL** + {OC}. We will also call this system **S7**. Since it is a normal alethic-deontic system **aKdKadOC** includes PL, the axioms aK and dK, the usual definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation. A normal **aKdKadOC**-system is any normal alethic-deontic system that includes OC, or in other words, any normal alethic-deontic system that is an extension of **aKdKadOC**.

Let us consider some properties of this system.

First of all we note that O(Op → ◇p) is a theorem in **aKdKadOC**. OC′ follows immediately from OC by O-necessitation. Accordingly, OC′ is a theorem in every **aKdKadOC**-system.

Next we turn to the alethic-deontic octagon in **aKdKadOC**.

4.4.1 The alethic-deontic octagon in aKdKadOC

Every system considered so far has been an ad combination, i.e. a combination of a purely alethic and a purely deontic system (see above). However, **aKdKadOC** is not a system of this kind, since OC includes both alethic and deontic operators. This is an example of a mixed axiom. When OC is added to **MADL** several interesting theorems follow. Figure 4 displays the relationships between primary alethic and deontic sentences in **aKdKadOC**.

Theorem 21. *All of the relationships displayed in figure 4 hold in every normal alethic-deontic* **aKdKadOC**-*system.*

Proof. Most of the proofs are quite easy and are left to the reader. (Table 5 includes a list of some of the theorems that are displayed in figure 4.) ■

	Theorems	
□p → Pp	¬(□p ∧ Fp)	Pp ∨ ◇¬p
Fp → ◇¬p	¬(Op ∧ ⟠p)	P¬p ∨ ◇p
Fp → ¬□p	¬(Op ∧ □¬p)	Pp ∨ ¬□p
⟠p → ¬Op	¬(O¬p ∧ □p)	¬Op ∨ ◇p
⟠p → P¬p	¬(O¬p ∧ ⟠¬p)	¬O¬p ∨ ◇¬p
Fp → ⊟p	¬(F¬p ∧ □¬p)	¬Fp ∨ ◇¬p
⟠p → Up	¬(Fp ∧ ⟠¬p)	¬F¬p ∨ ◇p

Table 5

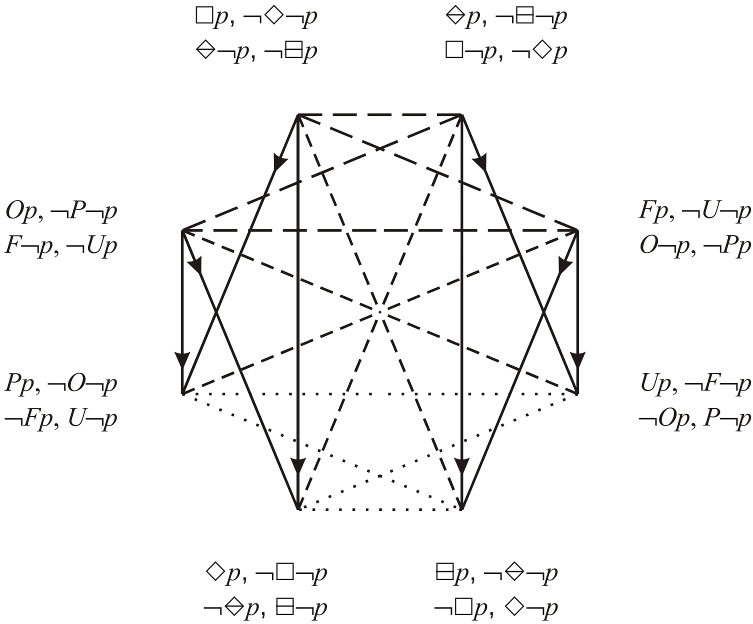

Figure 4. The Alethic-Deontic Octagon, MADL + {OC} **(S7).**

Note that all of the sentences in table 5 are equivalent in the system **aKdKadOC**. So, **MADL** + any sentence in table 5 is deductively equivalent with **aKdKadOC**. All sentences in table 5 stand or fall together. If we accept one of them we should accept all the others and if we reject one, we should reject all the others.

Since OC is one interpretation of the so-called ought-can principle, **aKdKadOC** tells us something about what follows by accepting this principle.

4.4.2 Some theorems including necessary implications

I will soon establish some derived rules in **aKdKadOC**. But first I will consider some theorems that include necessary implications.

	Theorem	Intuitive reading
(i)	□(p → q) → (Op → ◇q)	If it is necessary that p implies q, then it is possible that q if it is obligatory that p.

(ii)	(Op ∧ □(p → q)) → ◇q	If it is obligatory that p and it is necessary that p implies q, then it is possible that q.
(iii)	Op → (□(p → q) → ◇q)	If it is obligatory that p, then it is possible that q if it is necessary that p implies q.
(iv)	□(p → q) → (□p → Pq)	If it is necessary that p implies q, then it is permitted that q if it is necessary that p.
(v)	(□p ∧ □(p → q)) → Pq	If it is necessary that p and it is necessary that p implies q, then it is permitted that q.
(vi)	□p → (□(p → q) → Pq)	If it is necessary that p, then it is permitted that q if it is necessary that p implies q.
(vii)	□(p → q) → (Fq → ⊟p)	If it is necessary that p implies q, then it is unnecessary that p if it is forbidden that q.
(viii)	(Fq ∧ □(p → q)) → ⊟p	If it is forbidden that q and it is necessary that p implies q, then it is unnecessary that p.
(ix)	Fq → (□(p → q) → ⊟p)	If it is forbidden that q, then it is unecessary that p if it is necessary that p implies q.
(x)	□(p → q) → (⟠q → Up)	If it is necessary that p implies q, then it is unobligatory that p if it is impossible that q.
(xi)	(⟠q ∧ □(p → q)) → Up	If it is impossible that q and it is necessary that p implies q, then it is unobligatory that p.
(xii)	⟠q → (□(p → q) → Up)	If it is impossible that q, then it is unobligatory that p if it is necessary that p implies q.

Table 6

Theorem 22. *All sentences in table 6 are theorems in* **aKdKadOC**.

Proof. Part (ii) and part (iii) follow immediately from part (i) by PL. Likewise part (v) and part (vi) follow from part (iv), part (viii) and part (ix) from part (vii), and part (xi) and part (xii) from part (x), all by PL. So, we only have to show part (i), part (iv), part (vii) and part (x). I will leave part (vii) and part (x) to the reader and prove the rest.

Part (i). □(p → q) → (Op → ◇q)

1.	□(p → q) → (◇p → ◇q)	[**aK**]
2.	Op → ◇p	[OC]
3.	□(p → q) → (Op → ◇q)	[1, 2, PL]

Step (1) is a theorem in the minimal alethic modal system **aK**. So, it is a theorem in every normal alethic and alethic-deontic system. We have indicated this by writing **aK** in the justificatory entry. Part (i) says that if it is necessary that p implies q, then if it is obligatory that p then it is possible that q.

Part (iv). □(p → q) → (□p →Pq)

1.	□(p → q) → (□p →□q)	[aK]
2.	□q → Pq	[T21 q/p]
3.	□(p → q) → (□p →Pq)	[1, 2, PL]

Part (iv) claims that if it is necessary that p implies q, then if it is necessary that p then it is permitted that q. Step (1) is the axiom aK, step (2) is obtained from theorem 21 by substituting q for p and step (3) is deduced from (1) and (2) by PL. ■

We are now in a position to prove some derived rules.

4.4.3 Some derived rules in aKdKadOC

	Derived Rules	Intuitive reading
(i)	If ⊢ A → B, then ⊢ OA → ◇B	If A implies B is a theorem, then OA implies ◇B is a theorem.
(ii)	If ⊢ A → B, then ⊢ □A → PB	If A implies B is a theorem, then □A implies PB is a theorem.
(iii)	If ⊢ A → B, then ⊢ FB → ⊟A	If A implies B is a theorem, then FB implies ⊟B is a theorem.
(iv)	If ⊢ A → B, then ⊢ ⟠B → UA	If A implies B is a theorem, then ⟠B implies UA is a theorem.

Table 7

Theorem 23. *All rules in table 7 are derived rules in* **aKdKadOC**.

Proof. I will prove part (i) and part (ii), part (iii) and part (iv) are left to the reader.

Derived rule (i). If ⊢ A → B, then ⊢ OA → ◇B. If A implies B is a theorem, then OA implies ◇B is a theorem.

Proof. Suppose (1) ⊢ A → B. Then, (2) ⊢ □(A → B) [from 1 by □-necessitation]. Hence, (3) ⊢ OA → ◇B [from 2 and theorem 22]. It follows that (4) if ⊢ A → B, then ⊢ OA → ◇B [by conditional proof from 1–3 discharging the assumption].

Derived rule (ii). If ⊢ A → B, then ⊢ □A → PB. If A implies B is a theorem, then □A implies PB.

Proof. Suppose (1) ⊢ A → B. Then (2) ⊢ □(A → B) [from 1 by □-necessitation]. (3) ⊢ □(A → B) → (□A → PB) [by theorem 22]. So, (4) ⊢ □A → PB [from 2 and 3 by MP]. Consequently, (5) if ⊢ A → B, then ⊢ □A → PB [by conditional proof from 1–4 discharging the assumption]. ■

4.4.4 Conjunctive and disjunctive obligations, permissions, necessities and possibilities

Let us consider some theorems that include conjunctive and disjunctive obligations, permissions, necessities and possibilities.

	Theorem	Informal reading
(i)	$(Op \wedge Oq) \rightarrow \Diamond(p \wedge q)$	If it is obligatory that p and it is obligatory that q, then it is possible that p and q.
(ii)	$(\Box p \wedge \Box q) \rightarrow P(p \wedge q)$	If both p and q are necessary, then it is permitted that p and q.
(iii)	$(Op \vee Oq) \rightarrow \Diamond(p \vee q)$	If it is obligatory that p or it is obligatory that q, then it is possible that p or q.
(iv)	$(\Box p \vee \Box q) \rightarrow P(p \vee q)$	If either p or q is necessary, then it is permitted that p or q.
(v)	$O(p \wedge q) \rightarrow (\Diamond p \wedge \Diamond q)$	If it is obligatory that p and q, then both p and q are possible.
(vi)	$\Box(p \wedge q) \rightarrow (Pp \wedge Pq)$	If it is necessary that p and q, then it is permitted that p and it is permitted that q.

Table 8

Theorem 24. *Every sentences in table 8 is a theorem in* **aKdKadOC**.

Proof. Straightforward. Rönnedal (2010) may be helpful. ■

In section 4.4.6 we will see how to generalise this theorem.

4.4.5 More rules

Let us consider some more derived rules.

	Derived Rule	
(i)	If $\vdash (A_1 \vee ... \vee A_n) \rightarrow A$, then $\vdash (OA_1 \vee ... \vee OA_n) \rightarrow \Diamond A$	(for $n \geq 0$)
(ii)	If $\vdash (A_1 \vee ... \vee A_n) \rightarrow A$, then $\vdash ⟠A \rightarrow (UA_1 \wedge ... \wedge UA_n)$	(for $n \geq 0$)
(iii)	If $\vdash (A_1 \vee ... \vee A_n) \rightarrow A$, then $\vdash (\Box A_1 \vee ... \vee \Box A_n) \rightarrow PA$	(for $n \geq 0$)
(iv)	If $\vdash (A_1 \vee ... \vee A_n) \rightarrow A$, then $\vdash FA \rightarrow (\boxminus A_1 \wedge ... \wedge \boxminus A_n)$	(for $n \geq 0$)
(v)	If $\vdash (A_1 \wedge ... \wedge A_n) \rightarrow A$, then $\vdash (OA_1 \wedge ... \wedge OA_n) \rightarrow \Diamond A$	(for $n \geq 0$)
(vi)	If $\vdash (A_1 \wedge ... \wedge A_n) \rightarrow A$, then $\vdash ⟠A \rightarrow (UA_1 \vee ... \vee UA_n)$	(for $n \geq 0$)
(vii)	If $\vdash (A_1 \wedge ... \wedge A_n) \rightarrow A$, then $\vdash (\Box A_1 \wedge ... \wedge \Box A_n) \rightarrow PA$	(for $n \geq 0$)
(viii)	If $\vdash (A_1 \wedge ... \wedge A_n) \rightarrow A$, then $\vdash FA \rightarrow (\boxminus A_1 \vee ... \vee \boxminus A_n)$	(for $n \geq 0$)
(ix)	If $\vdash A \rightarrow (A_1 \vee ... \vee A_n)$, then $\vdash OA \rightarrow (\Diamond A_1 \vee ... \vee \Diamond A_n)$	(for $n \geq 0$)
(x)	If $\vdash A \rightarrow (A_1 \vee ... \vee A_n)$, then $\vdash (⟠A_1 \wedge ... \wedge ⟠A_n) \rightarrow UA$	(for $n \geq 0$)
(xi)	If $\vdash A \rightarrow (A_1 \vee ... \vee A_n)$, then $\vdash \Box A \rightarrow (PA_1 \vee ... \vee PA_n)$	(for $n \geq 0$)

(xii)	If $\vdash A \rightarrow (A_1 \vee ... \vee A_n)$, then $\vdash (FA_1 \wedge ... \wedge FA_n) \rightarrow \boxminus A$	(for $n \geq 0$)
(xiii)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash OA \rightarrow (\Diamond A_1 \wedge ... \wedge \Diamond A_n)$	(for $n \geq 0$)
(xiv)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash (\diamondminus A_1 \vee ... \vee \diamondminus A_n) \rightarrow UA$	(for $n \geq 0$)
(xv)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash \Box A \rightarrow (PA_1 \wedge ... \wedge PA_n)$	(for $n \geq 0$)
(xvi)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash (FA_1 \vee ... \vee FA_n) \rightarrow \boxminus A$	(for $n \geq 0$)

Table 9

Theorem 25. *All rules in table 9 are derived rules in* **aKdKadOC**.

Proof. Left to the reader. Rönnedal (2010) may be helpful. ■

4.4.6 Generalisations of distribution theorems

	Theorem		Theorem
(i)	$(Op_1 \wedge ... \wedge Op_n) \rightarrow \Diamond(p_1 \wedge ... \wedge p_n)$	(iv)	$(\Box p_1 \vee ... \vee \Box p_n) \rightarrow P(p_1 \vee ... \vee p_n)$
(ii)	$(\Box p_1 \wedge ... \wedge \Box p_n) \rightarrow P(p_1 \wedge ... \wedge p_n)$	(v)	$O(p_1 \wedge ... \wedge p_n) \rightarrow (\Diamond p_1 \wedge ... \wedge \Diamond p_n)$
(iii)	$(Op_1 \vee ... \vee Op_n) \rightarrow \Diamond(p_1 \vee ... \vee p_n)$	(vi)	$\Box(p_1 \wedge ... \wedge p_n) \rightarrow (Pp_1 \wedge ... \wedge Pp_n)$

Table 10

Theorem 26. *All sentences of the forms in table 10 are theorems in* **aKdKadOC**.

Proof. **Part (i)**. $(Op_1 \wedge ... \wedge Op_n) \rightarrow \Diamond(p_1 \wedge ... \wedge p_n)$.

1. $(p_1 \wedge ... \wedge p_n) \rightarrow (p_1 \wedge ... \wedge p_n)$ [PL]
2. $(Op_1 \wedge ... \wedge Op_n) \rightarrow \Diamond(p_1 \wedge ... \wedge p_n)$ [1, T25(v)]

Part (i) says that if it is obligatory that p_1 and … and it is obligatory that p_n, then it is possible that p_1 and … and p_n. So, a conjunction is possible if each conjunct is obligatory. The proof uses T25(v): if $\vdash (A_1 \wedge ... \wedge A_n) \rightarrow A$, then $\vdash (OA_1 \wedge ... \wedge OA_n) \rightarrow \Diamond A$. Note that the sentence on line (1) is of the form $(A_1 \wedge ... \wedge A_n) \rightarrow A$. An alternative proof uses O-distribution and OC like this.

1. $O(p_1 \wedge ... \wedge p_n) \rightarrow \Diamond(p_1 \wedge ... \wedge p_n)$ [OC, $p_1 \wedge ... \wedge p_n/p$]
2. $(Op_1 \wedge ... \wedge Op_n) \rightarrow \Diamond(p_1 \wedge ... \wedge p_n)$ [1, Dist]

See Rönnedal (2010) for more on how various deontic operators distribute.

Part (ii). $(\Box p_1 \wedge ... \wedge \Box p_n) \rightarrow P(p_1 \wedge ... \wedge p_n)$.

Part (ii) claims that a conjunction is permitted if each conjunct is necessary. In other words, if it is necessary that p_1 and … and it is necessary that p_n, then it is permitted that p_1 and … and p_n. The proof is similar to the proof of part (i), but this time use T25(vii): if $\vdash (A_1 \wedge ... \wedge A_n) \rightarrow A$, then $\vdash (\Box A_1 \wedge ... \wedge \Box A_n) \rightarrow PA$.

Part (iii). $(Op_1 \vee ... \vee Op_n) \rightarrow \Diamond(p_1 \vee ... \vee p_n)$.

According to part (iii) a disjunction is possible if any disjunct is obligatory. That is, if it is obligatory that p_1 or … or it is obligatory that p_n, then it is possible that p_1 or … or p_n. The proof is similar to the proof of part (i) but uses T25(i): if $\vdash (A_1 \lor ... \lor A_n) \to A$, then $\vdash (OA_1 \lor ... \lor OA_n) \to \Diamond A$.

Part (iv). $(\Box p_1 \lor ... \lor \Box p_n) \to P(p_1 \lor ... \lor p_n)$.

Part (iv) asserts that if it is necessary that p_1 or … or it is necessary that p_n, then it is permitted that p_1 or … or p_n. So, a disjunction is permitted if any disjunct is necessary. The sentence follows directly from PL and T25(iii): if $\vdash (A_1 \lor ... \lor A_n) \to A$, then $\vdash (\Box A_1 \lor ... \lor \Box A_n) \to PA$.

Part (v). $O(p_1 \land ... \land p_n) \to (\Diamond p_1 \land ... \land \Diamond p_n)$.

Part (v) follows immediately from PL and T25(xiii): if $\vdash A \to (A_1 \land ... \land A_n)$, then $\vdash OA \to (\Diamond A_1 \land ... \land \Diamond A_n)$. According to the sentence each conjunct is possible if a conjunction is obligatory. I.e. if it is obligatory that p_1 and … and p_n, then it is possible that p_1 and … and it is possible that p_n.

Part (vi). $\Box(p_1 \land ... \land p_n) \to (Pp_1 \land ... \land Pp_n)$.

Part (vi) says that it is permitted that p_1 and … and it is permitted that p_n if it is necessary that p_1 and … and p_n. So, if a conjunction is necessary, then each conjunct is permitted. The proof is similar to the proof of part (i) but uses T25(xv): if $\vdash A \to (A_1 \land ... \land A_n)$, then $\vdash \Box A \to (PA_1 \land ... \land PA_n)$. ■

Theorem 27. (i) **MADL** + OC *includes* OC′, $\Box p \to \Diamond p$ *and* $Op \to Pp$. (ii) *All of the following systems are deductively equivalent:* **aKdKadOC**, **aKDdKadOC**, **aKdKDadOC** *and* **aKDdKDadOC**.

Proof. Left to the reader. ■

4.5 aKdKadMO

The smallest normal alethic-deontic logic that includes the axiom MO, i.e. the sentence $\Box p \to Op$, is **aKdKadMO**. Accordingly, **aKdKadMO** = **MADL** + {MO}. We will also call this system **S4**. Since it is a normal alethic-deontic system **aKdKadMO** includes PL, the axioms aK and dK, the usual definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation. A normal **aKdKadMO**-system is any normal alethic-deontic system that includes MO, or in other words, any normal alethic-deontic system that is an extension of **aKdKadMO**.

Let us consider some properties of this system.

4.5.1 The alethic-deontic octagon

Figure 5 displays the alethic-deontic octagon in the system **aKdKadMO**. The octagon is interpreted as usual.

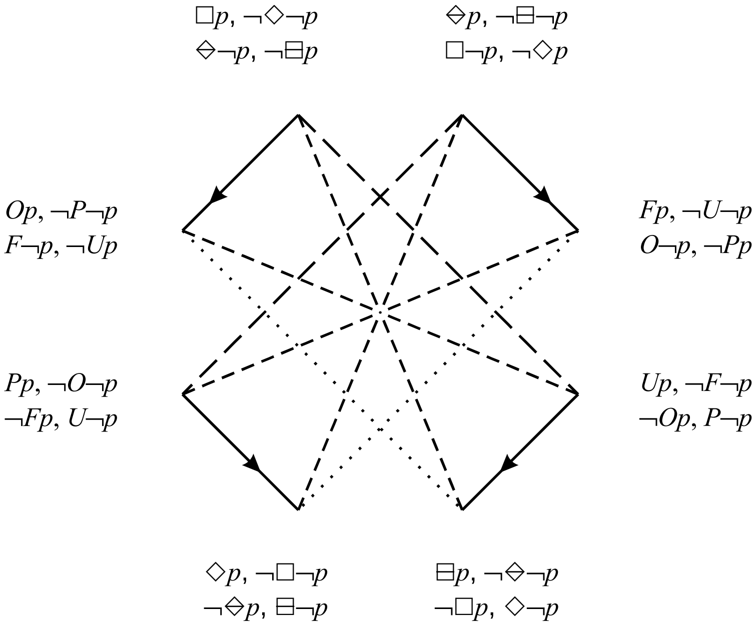

Figure 5. The Alethic-Deontic Octagon, MADL + {MO} **(S4).**

Theorem 28. *All of the relationships displayed in figure 5 hold in every normal alethic-deontic* **aKdKadMO**-*system.*

Proof. Most of the proofs are quite easy and are left to the reader. Table 11 includes a list of some theorems displayed in figure 5. Note that all sentences in this table are equivalent in **aKdKadMO**. **MADL** + any sentence in table 11 is deductively equivalent with **MADL** + {MO}. ■

	Theorems	
Pp → ◇p	¬(□p ∧ ¬Op)	Fp ∨ ◇p
⬙p → Fp	¬(Pp ∧ □¬p)	Op ∨ ◇¬p
¬Op → ¬□p	¬(□p ∧ P¬p)	O¬p ∨ ◇p
⬙p → ¬Pp	¬(Pp ∧ ⬙p)	Op ∨ ¬□p
□p → F¬p	¬(⬙¬p ∧ ¬Op)	F¬p ∨ ◇¬p
Pp→¬⬙p	¬(¬Fp ∧ ⬙p)	Fp ∨ ¬□¬p
Up → ⊟p	¬(□p ∧ ¬F¬p)	¬Pp ∨ ◇p

Table 11

4.5.2 The means-end principle

Perhaps the most interesting feature of the system **aKdKadMO** is that the so-called means-end principle is a theorem in it. According to the means-end principle every necessary consequence of what ought to be ought to be. Table 12 includes this principle and several similar theorems.

	Theorem	**Informal reading**
(i)	$\Box(p \to q) \to (Op \to Oq)$	If it is necessary that p implies q, then if it ought to be that p then it ought to be that q.
(ii)	$(Op \land \Box(p \to q)) \to Oq$	If it ought to be that p and it is necessary that if p then q, then it ought to be that q.
(iii)	$Op \to (\Box(p \to q) \to Oq)$	If it ought to be that p, then if it is necessary that p implies q then it ought to be that q.
(iv)	$\Box(p \to q) \to (Pp \to Pq)$	If it is necessary that p implies q, then if it is permitted that p it is permitted that q.
(v)	$(Pp \land \Box(p \to q)) \to Pq$	If it is permitted that p and it is necessary that p implies q, then q is permitted.
(vi)	$Pp \to (\Box(p \to q) \to Pq)$	If it is permitted that p, then if it is necessary that p implies q then q is permitted.
(vii)	$\Box(p \to q) \to (Fq \to Fp)$	If it is necessary that p implies q, then if it is forbidden that q then it is forbidden that p.
(viii)	$(Fq \land \Box(p \to q)) \to Fp$	If it is forbidden that q and it is necessary that p implies q, then it is forbidden that p.
(ix)	$Fq \to (\Box(p \to q) \to Fp)$	If it is forbidden that q, then if it is necessary that p implies q then it is forbidden that p.
(x)	$\Box(p \to q) \to (\Box p \to Oq)$	If it is necessary that p implies q, then it is obligatory that q if it is necessary that p.
(xi)	$(\Box p \land \Box(p \to q)) \to Oq$	If it is necessary that p and it is necessary that p implies q, then it is obligatory that q.
(xii)	$\Box p \to (\Box(p \to q) \to Oq)$	If it is necessary that p, then if it is necessary that p implies q it is obligatory that q.
(xiii)	$\Box(p \to q) \to (Pp \to \Diamond q)$	If it is necessary that p implies q, then it is possible that q if it is permitted that p.
(xiv)	$(Pp \land \Box(p \to q)) \to \Diamond q$	If it is permitted that p and it is necessary that p implies q, then it is possible that q.
(xv)	$Pp \to (\Box(p \to q) \to \Diamond q)$	If it is permitted that p, then if it is necessary that p implies q it is possible that q.
(xvi)	$\Box(p \to q) \to (\not\Diamond q \to Fp)$	If it is necessary that p implies q, then it is forbidden that p if it is impossible that q.

(xvii)	$(\ominus q \wedge \Box(p \rightarrow q)) \rightarrow Fp$	If it is impossible that q and it is necessary that p implies q, then it is forbidden that p.
(xviii)	$\ominus q \rightarrow (\Box(p \rightarrow q) \rightarrow Fp)$	If it is impossible that q, then if it is necessary that p implies q it is forbiden that p.
(xix)	$\Box(p \rightarrow q) \rightarrow (Uq \rightarrow \boxminus p)$	If it is necessary that p implies q, then it is unnecessary that p if it is unobligatory that q.
(xx)	$(Uq \wedge \Box(p \rightarrow q)) \rightarrow \boxminus p$	If it is unobligatory that q and it is necessary that p implies q, then it is unnecessary that p.
(xxi)	$Uq \rightarrow (\Box(p \rightarrow q) \rightarrow \boxminus p)$	If it is unobligatory that q, then if it is necessary that p implies q it is unnecessary that p.

Table 12

Theorem 29. *All sentences in table 12 are theorems in* **aKdKadMO**.

Proof. I will prove part (i), part (vii), part (xiii) and part (xvi) to illustrate the method, and leave the rest to the reader. The philosophically most interesting parts are perhaps part (i) – (ix). These theorems can be used to derive obligations from obligations, permissions from permissions and prohibitions from prohibitions, with the help of necessary implications.

Part (i). $\Box(p \rightarrow q) \rightarrow (Op \rightarrow Oq)$

1. $O(p \rightarrow q) \rightarrow (Op \rightarrow Oq)$ [dK]
2. $\Box(p \rightarrow q) \rightarrow O(p \rightarrow q)$ [MO, p→q/p]
3. $\Box(p \rightarrow q) \rightarrow (Op \rightarrow Oq)$ [1, 2, PL]

Part (i) is one version of the means-end principle. Part (ii) and (iii) are similar versions of this principle. It is easy to derive part (ii) and part (iii) from part (i). The means-end principle is intuitively plausible and can be very useful when deriving "new" obligations from "old" obligations. Suppose for instance that it ought to be that everyone is honest. Then it follows that you ought to be honest. For it is necessary that if everyone is honest then you are honest.

Part (vii). $\Box(p \rightarrow q) \rightarrow (Fq \rightarrow Fp)$

1. $O(p \rightarrow q) \rightarrow (Fq \rightarrow Fp)$ [**OK**]
2. $\Box(p \rightarrow q) \rightarrow O(p \rightarrow q)$ [MO, p→q/p]
3. $\Box(p \rightarrow q) \rightarrow (Fq \rightarrow Fp)$ [1, 2, PL]

Note that the sentence at step (1) is provable in the deontic system **OK**; and since every normal ad system includes **OK**, this sentence is a theorem in **MADL** + MO too. Part (vii) is also a quite useful principle. Suppose it is forbidden that you smoke in this restaurant. Then it follows that it is forbidden that you smoke a cigar in this restaurant. For it is necessary that if you smoke a cigar in this restaurant you smoke in this restaurant.

Part (xiii). $\Box(p \to q) \to (Pp \to \Diamond q)$

1.	$\Box(p \to q) \to O(p \to q)$	[MO, $p \to q/p$]
2.	$O(p \to q) \to (Pp \to Pq)$	[**OK**]
3.	$Pq \to \Diamond q$	[T28, q/p]
4.	$\Box(p \to q) \to (Pp \to \Diamond q)$	[1, 2, 3, PL]

Note that the sentence at step (2) is provable in the deontic system **OK**; and since every normal ad system includes **OK**, this sentence is a theorem in **MADL** + MO too.

Part (xvi). $\Box(p \to q) \to (⟠q \to Fp)$

1.	$\Box(p \to q) \to (Pp \to \Diamond q)$	[(xiii)]
2.	$(Pp \to \Diamond q) \leftrightarrow (⟠q \to Fp)$	[PL, adIT etc.]
3.	$\Box(p \to q) \to (⟠q \to Fp)$	[1, 2, PL] ■

4.5.3 Some derived rules in aKdKadMO

	Derived Rules	Informal reading
(i)	If $\vdash A \to B$, then $\vdash \Box A \to OB$	If A implies B is a theorem, then $\Box$A implies OB is a theorem.
(ii)	If $\vdash A \to B$, then $\vdash PA \to \Diamond B$	If A implies B is a theorem, then PA implies $\Diamond$B is a theorem.
(iii)	If $\vdash A \to B$, then $\vdash ⟠B \to FA$	If A implies B is a theorem, then ⟠B implies FA is a theorem.
(iv)	If $\vdash A \to B$, then $\vdash UB \to \boxminus A$	If A implies B is a theorem, then UB implies $\boxminus$A is a theorem.

Table 13

Theorem 30. *All rules in table 13 are derived rules in* **aKdKadMO**.

Proof.

Derived rule (i). If $\vdash A \to B$, then $\vdash \Box A \to OB$. If A implies B is a theorem, then $\Box$A implies OB is a theorem.

Proof. Suppose that (1) $\vdash A \to B$. Then (2) $\vdash \Box(A \to B)$ [by $\Box$-necessitation from 1]. (3) $\vdash \Box(A \to B) \to (\Box A \to OB)$ [by theorem 29(x)]. Hence, (4) $\vdash \Box A \to OB$ [from 2 and 3 by modus ponens]. It follows that (5) if $\vdash A \to B$, then $\vdash \Box A \to OB$ [by conditional proof from 1–4 discharging the assumption].

Derived rule (ii). If $\vdash A \to B$, then $\vdash PA \to \Diamond B$. If A implies B is a theorem, then PA implies $\Diamond$B is a theorem.

Proof. Suppose (1) $\vdash A \to B$. Then (2) $\vdash \Box(A \to B)$ [from 1 by $\Box$-necessitation]. Hence, (3) $\vdash PA \to \Diamond B$ [from 2 and theorem 29(xiii)]. It follows that (4) if ⊢ A → B, then ⊢ PA → ◇B [by conditional proof from 1–3 discharging the assumption].

Derived rule (iii). If ⊢ A → B, then ⊢ ⟠B → FA. If A implies B is a theorem, then ⟠B implies FA is a theorem.

Proof. Suppose that (1) ⊢A → B. Then (2) ⊢ □(A → B) [from 1 by □-necessitation]. Hence, (3) ⊢ ⟠B → FA [from 2 and theorem 29(xvi)]. It follows that (4) if ⊢ A → B, then ⊢ ⟠B → FA [by conditional proof from 1–3 discharging the assumption].

Derived rule (iv). If ⊢ A → B, then ⊢ UB → ⊟A. If A implies B is a theorem, then UB implies ⊟A is a theorem. Proof is left to the reader. ■

4.5.4 Conjunctive and disjunctive obligations, permissions, necessities and possibilities

Let us consider some theorems that include conjunctive and disjunctive obligations, permissions, necessities and possibilities.

	Theorem	Intuitive reading
(i)	(□p ∧ □q) → O(p ∧ q)	If both p and q are necessary, then it is obligatory that p and q.
(ii)	□(p ∧ q) → (Op ∧ Oq)	If it is necessary that p and q, then it is obligatory that p and it is obligatory that q.
(iii)	P(p ∧ q) → (◇p ∧ ◇q)	If it is permitted that p and q, then it is possible that p and it is possible that q.
(iv)	(⟠p ∨ ⟠q) → F(p ∧ q)	If it is impossible that p or it is impossible that q, then it is fobidden that p and q.
(v)	⟠(p ∨ q) → (Fp ∧ Fq)	If it is impossible that p or q, then both p and q are forbidden.
(vi)	(□p ∨ □q) → O(p ∨ q)	If it is necessary that p or it is necessary that q, then it is obligatory that p or q.
(vii)	(Pp ∨ Pq) → ◇(p ∨ q)	If it is permitted that p or it is permitted that q, then it is possible that p or q.
(viii)	P(p ∨ q) → (◇p ∨ ◇q)	If it is permitted that p or q, then either p or q is possible.
(ix)	(⟠p ∧ ⟠q) → F(p ∨ q)	If it is impossible that p and it is impossible that q, then it is forbidden that p or q.

Table 14

Theorem 31. *Every sentence in table 14 is a theorem in* **aKdKadMO**.

Proof. Straightforward. ■

4.5.5 The contingency octagon in aKdKadMO

It is possible to construct an alethic-deontic contingency octagon that displays the relationships between the alethic and deontic “contingency” operators. Figure 6 shows us how these concepts are related in **aKdKadMO**.

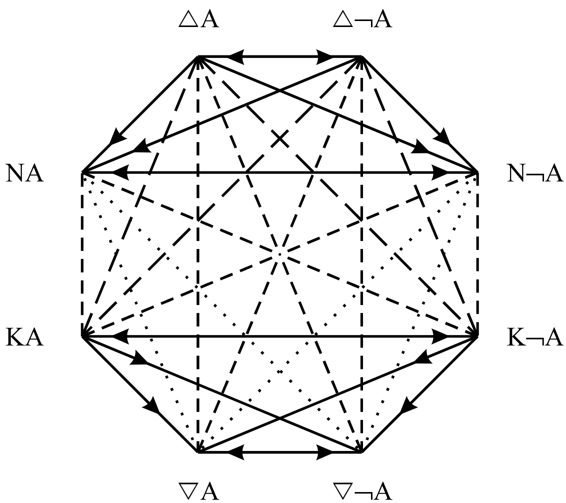

Figure 6. The Contingency Octagon in aKdKadMO.

4.5.6 More rules in aKdKadMO

	Derived Rules	
(i)	If ⊢ $(A_1 \vee ... \vee A_n) \rightarrow A$, then ⊢ $(□A_1 \vee ... \vee □A_n) \rightarrow OA$	(for $n \geq 0$)
(ii)	If ⊢ $(A_1 \vee ... \vee A_n) \rightarrow A$, then ⊢ $(PA_1 \vee ... \vee PA_n) \rightarrow$ ◇A	(for $n \geq 0$)
(iii)	If ⊢ $(A_1 \vee ... \vee A_n) \rightarrow A$, then ⊢ ⟠$A \rightarrow (FA_1 \wedge ... \wedge FA_n)$	(for $n \geq 0$)
(iv)	If ⊢ $(A_1 \vee ... \vee A_n) \rightarrow A$, then ⊢ $UA \rightarrow$ (⊟$A_1 \wedge ... \wedge$ ⊟A_n)	(for $n \geq 0$)
(v)	If ⊢ $(A_1 \wedge ... \wedge A_n) \rightarrow A$, then ⊢ $(□A_1 \wedge ... \wedge □A_n) \rightarrow OA$	(for $n \geq 0$)
(vi)	If ⊢ $(A_1 \wedge ... \wedge A_n) \rightarrow A$, then ⊢ $UA \rightarrow$ (⊟$A_1 \vee ... \vee$ ⊟A_n)	(for $n \geq 0$)
(vii)	If ⊢ $A \rightarrow (A_1 \vee ... \vee A_n)$, then ⊢ $PA \rightarrow$ (◇$A_1 \vee ... \vee$ ◇A_n)	(for $n \geq 0$)
(viii)	If ⊢ $A \rightarrow (A_1 \vee ... \vee A_n)$, then ⊢ (⟠$A_1 \wedge ... \wedge$ ⟠A_n) $\rightarrow FA$	(for $n \geq 0$)

(ix)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash \Box A \rightarrow (OA_1 \wedge ... \wedge OA_n)$	(for $n \geq 0$)
(x)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash PA \rightarrow (\Diamond A_1 \wedge ... \wedge \Diamond A_n)$	(for $n \geq 0$)
(xi)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash$ (⟠$A_1 \vee ... \vee$ ⟠$A_n) \rightarrow FA$	(for $n \geq 0$)
(xii)	If $\vdash A \rightarrow (A_1 \wedge ... \wedge A_n)$, then $\vdash (UA_1 \vee ... \vee UA_n) \rightarrow \boxminus A$	(for $n \geq 0$)

Table 15

Theorem 32. *All rules in table 15 are derived rules in* **aKdKadMO**.
Proof. Left to the reader. ■

	Theorem		Theorem
(i)	$(\Box p_1 \wedge ... \wedge \Box p_n) \rightarrow O(p_1 \wedge ... \wedge p_n)$	(v)	⟠$(p_1 \vee ... \vee p_n) \rightarrow (Fp_1 \wedge ... \wedge Fp_n)$
(ii)	$\Box(p_1 \wedge ... \wedge p_n) \rightarrow (Op_1 \wedge ... \wedge Op_n)$	(vi)	$(\Box p_1 \vee ... \vee \Box p_n) \rightarrow O(p_1 \vee ... \vee p_n)$
(iii)	$P(p_1 \wedge ... \wedge p_n) \rightarrow (\Diamond p_1 \wedge ... \wedge \Diamond p_n)$	(vii)	$(Pp_1 \vee ... \vee Pp_n) \rightarrow \Diamond(p_1 \vee ... \vee p_n)$
(iv)	(⟠$p_1 \vee ... \vee$ ⟠$p_n) \rightarrow F(p_1 \wedge ... \wedge p_n)$	(viii)	$P(p_1 \vee ... \vee p_n) \rightarrow (\Diamond p_1 \vee ... \vee \Diamond p_n)$
		(ix)	(⟠$p_1 \wedge ... \wedge$ ⟠$p_n) \rightarrow F(p_1 \vee ... \vee p_n)$

Table 16

Theorem 33. *Every sentence in table 16 is a theorem in* **aKdKadMO**.
Proof. Straightforward. ■

4.6 aKDdKDad∅

The smallest normal alethic-deontic logic that includes the axioms aD and dD, i.e. the sentences $Op \rightarrow Pp$ and $\Box p \rightarrow \Diamond p$, is **aKDdKDad∅**. Consequently, **aKDdKDad∅** = **MADL** + {aD, dD}. It is our first example of an ad system that contains more than one additional axiom. Nevertheless, the system is an ad combination of the pure alethic logic **aKD** and the pure deontic logic **dKD** (**SDL**), since it doesn't contain any mixed axioms, in contrast to our two previous systems **aKdKadOC** and **aKdKadMO**. We will also call this system **S5**. **aKDdKDad∅** includes PL, the axioms aK and dK, the ordinary definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation, like every normal alethic-deontic system. We shall say that any normal alethic-deontic system that is an extension of **aKDdKDad∅**, i.e. any normal alethic-deontic system that includes aD and dD, is a normal **aKDdKDad∅**-system.

Let us consider some properties of this system.

4.6.1 The alethic-deontic octagon

Figure 7 displays the alethic-deontic octagon in the system **aKDdKDad∅**. The octagon is interpreted as usual.

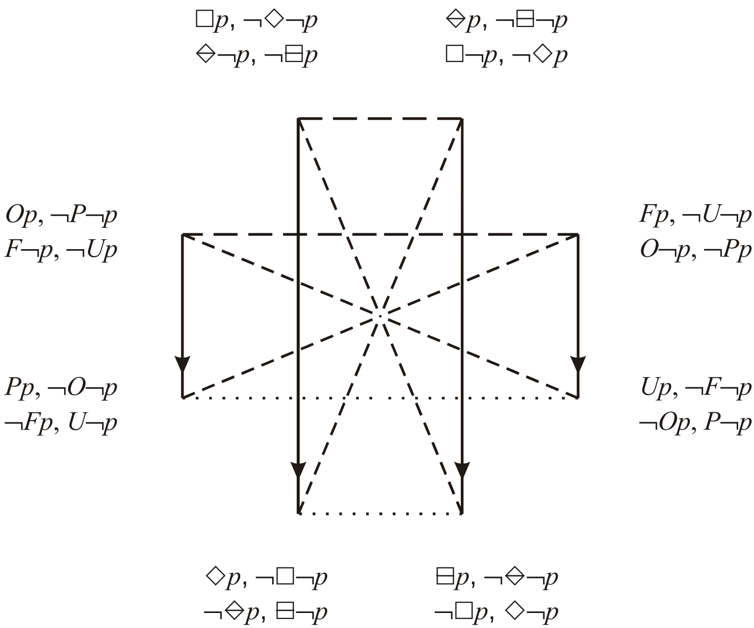

Figure 7. The Alethic-Deontic Octagon, MADL + {aD, dD} **(S5).**

Note that the figure essentially is a combination of the ad octagon for the system **aKDdKad∅** and the ad octagon for the system **aKdKDad∅**. Now, this should come as no surprise, since **aKDdKDad∅** includes every sentence in **aKDdKad∅** and in **aKdKDad∅**. Furthermore, since **MADL** + {aD, dD} contains these systems, it is a **aKDdKad∅**-system, as well as a **aKdKDad∅**- and a **aKDdKDad∅**-system. No mixed axioms are included in the system. Hence, no interesting relationships between deontic and alethic propositions are forthcoming.

The next system we consider includes both a pure additional alethic axiom and a mixed axiom.

4.7 aKDdKadMO

The smallest normal alethic-deontic logic that includes the axiom aD and MO, i.e. the sentences □p → ◇p and □p → Op, is **aKDdKadMO**. Accordingly, **aKDdKadMO** = **MADL** + {aD, MO}. We will also call this system **S6**. Since it is a normal alethic-deontic system **aKDdKadMO** includes PL, the axioms aK and dK, the usual definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation. A normal **aKDdKadMO**-system is any normal alethic-deontic system that includes aD and MO, or in other words, any normal alethic-deontic system that is an extension of **aKDdKadMO**.

Let us consider the alethic-deontic octagon in **aKDdKadMO**.

4.7.1 The alethic-deontic octagon

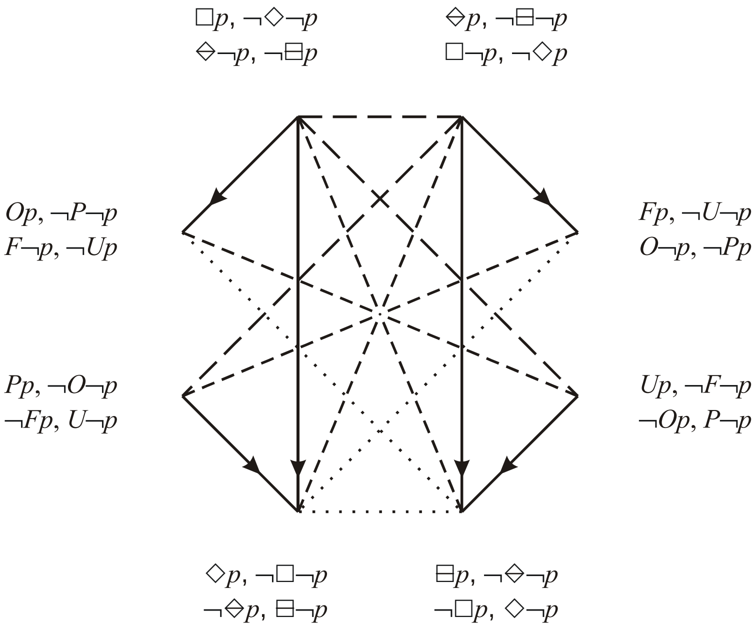

Figure 8. The Alethic-Deontic Octagon, MADL + {aD, MO} **(S6).**

4.8 SADL, aKdKadOCMO

aKdKadOCMO is the smallest normal alethic-deontic logic that includes the axioms OC and MO, i.e. the sentences $Op \rightarrow \Diamond p$ and $\Box p \rightarrow Op$. Accordingly, **aKdKadOCMO** = **MADL** + {OC, MO}. We will also call this system **S8** or Standard alethic-deontic logic **(SADL)**. Since it is a normal alethic-deontic system **aKdKadOCMO** includes PL, the axioms aK and dK, the usual definitions of the alethic and deontic operators, modus ponens, □-necessitation and O-necessitation. A normal **aKdKadOCMO**-system is any normal alethic-deontic system that includes OC and MO, or in other words, any normal alethic-deontic system that is an extension of **aKdKadOCMO**.

Let us consider some properties of this system.

4.8.1 The alethic-deontic octagon

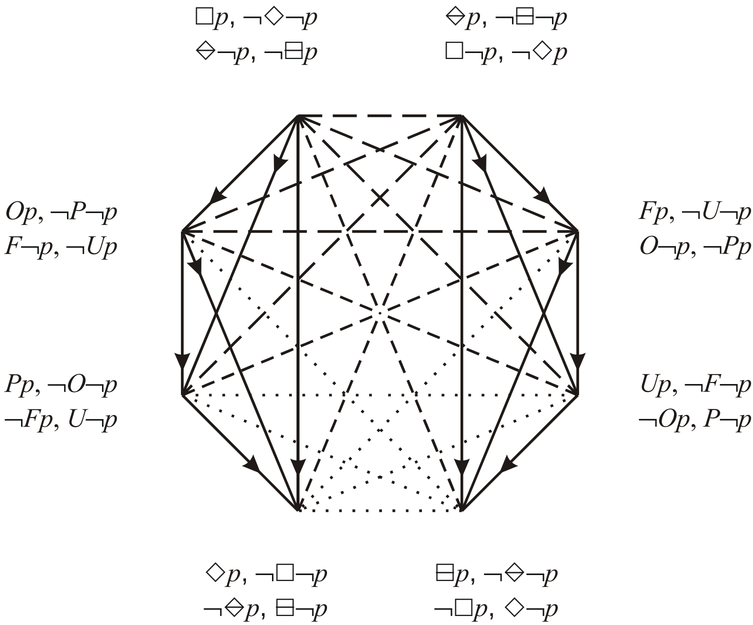

Figure 9. The Alethic-Deontic Octagon, aKdKadOCMO, SADL (S8).

4.8.2 Deductively equivalent systems

It is easy to see that all systems above are included in **aKdKadOCMO**. **MADL** is included since it is included in every normal ad system. Every extension of minimal alethic-deontic logic discussed so far in this paper is constructed by adding one or several of the axioms aD, dD, OC and MO to this system. The sentences aD and dD are theorems already in **aKdKadOC**. So, it is obvious that these sentences are provable also in **aKdKadOCMO**.

Hence, **aKdKadOCMO** includes **aKDdKad∅**, **aKdKDad∅** and **aKDdKDad∅**. Furthermore, it is also obvious that **aKdKadOC**, **aKdKadMO**, and **aKDdKadMO** are included in **aKdKadOCMO** since dD, OC and MO are theorems in **aKdKadOCMO**. Section 5 includes information about the relationships between all logics mentioned in this essay.

This completes our discussion of various alethic-deontic systems in this paper. We will end this article with some information about how the systems in this essay are related to each other.

5. Relationships between systems

Figure 10 displays the relationships between the systems we have discussed in this paper. Systems higher up in the diagram are stronger than systems lower down. So, S8 is the strongest system and S1 the weakest system. All other systems are included in S8 and S1 is included in all other systems.

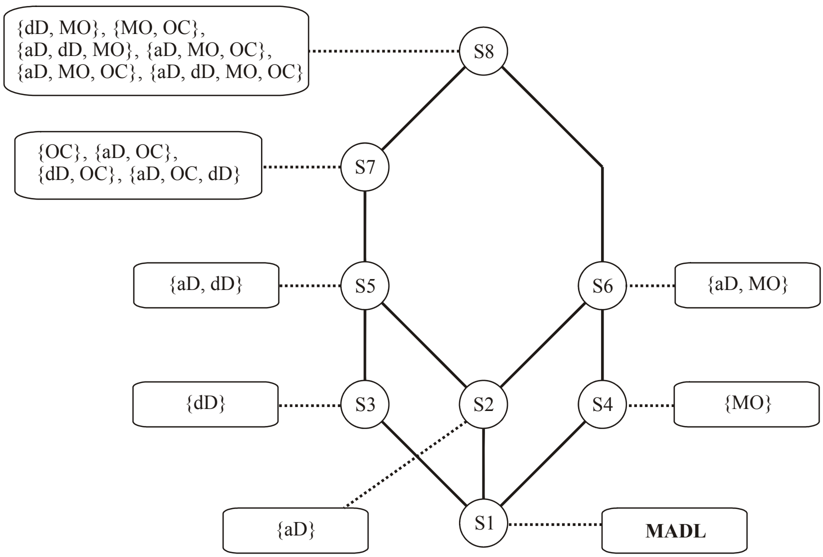

Figure 10. Relationships between some ad systems.

Nr	Systematic name	Extensions of MADL	Equivalent Systems
S1	**aKdKad∅**	**MADL**	
S2	**aKDdKad∅**	**MADL** + {aD}	
S3	**aKdKDad∅**	**MADL** + {dD}	
S4	**aKdKadMO**	**MADL** + {MO}	
S5	**aKDdKDad∅**	**MADL** + {aD, dD}	
S6	**aKDdKadMO**	**MADL** + {aD, MO}	
S7	**aKdKadOC**	**MADL** + {OC}	**MADL** + {aD, OC}**, MADL** + {dD, OC}, **MADL** + {aD, dD, OC}
S8	**aKdKadOCMO**	**MADL** + {OC, MO}	**MADL** + {aD, MO}**, MADL** + {aD, dD, MO}**, MADL** + {aD, OC, MO}**,** **MADL** + {aD, OC, MO}**, MADL** + {aD, dD, OC, MO}

Comment 34. In this paper I have described a set of alethic-deontic systems that include alethic and deontic operators that are used to symbolize various deontic and alethic modal concepts. But all systems have many possible informal readings. In Rönnedal (2012) I mention some interpretations of various bimodal systems. If we interpret □ as an epistemic operator and O as a doxastic operator, we obtain a set of epistemic-doxastic systems. If □ is read as "It is always the case that" or "It is and it is always going to be the case that" and O as "It is always going to be the case that", we obtain a set of bimodal temporal systems, etc. So, the results in this paper should be interesting not only to alethic-deontic logicians, but to any logician who wants to develop some kind of bimodal system.

References

Anderson, A. R. (1956). The formal analysis of normative systems. In N. Rescher (ed.). *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press, 1967, pp. 147–213.

Anderson, A. R. (1958). A reduction of deontic logic to alethic modal logic. *Mind*, Vol. 67, No. 265, pp. 100–103.

Anderson, A. R. (1959). On the logic of commitment. *Philosophical Studies* 10, pp. 23–27.

Anderson, A. R. (1967). Some Nasty Problems in the Formal Logic of Ethics. *Noûs*, Vol. 1, No. 4, pp. 345–360.

Blackburn, P., de Rijke, M. & Venema, Y. (2001). *Modal Logic*. Cambridge University Press.

Blackburn, P., van Benthem, J. & Wolter, F. (eds.). (2007). *Handbook of Modal Logic*. Elsevier.

Chellas, B. F. (1980). *Modal Logic: An Introduction*. Cambridge: Cambridge University Press.

Fine, K. & Schurz, G. (1996). Transfer Theorems for Multimodal Logics. In J. Copeland (ed.) (1996). *Logic and Reality. Essays in Pure and Applied Logic. In Memory of Arthur Prior*. Oxford University Press, Oxford, pp. 169–213.

Fitting, M. & Mendelsohn, R. L. (1998). *First-Order Modal Logic*. Kluwer Academic Publishers.

Gabbay, D. M. (1976). *Investigations in Modal and Tense Logics with Applications to Problems in Philosophy and Linguistics*. Dordrecht: D. Reidel Publishing Company.

Gabbay, D., Horty, J., Parent, X., van der Meyden, E. & van der Torre, L. (eds.). (2013). *Handbook of Deontic Logic and Normative Systems*. College Publications.

Gabbay, D. M. & Guenthner, F. (eds.). (2001). *Handbook of Philosophical Logic 2nd Edition*, Vol. 3, Dordrecht: Kluwer Academic Publishers.

Gabbay, D. M., Kurucz, A., Wolter, F. & Zakharyaschev, M. (2003). *Many-Dimensional Modal Logics: Theory and Applications*. Amsterdam: Elsevier.

Garson, J. W. (2006). *Modal Logic for Philosophers*. New York: Cambridge University Press.

Girle, R. (2000). *Modal Logics and Philosophy*. McGill-Queen's University Press.

Hilpinen, R. (ed.). (1971). *Deontic Logic: Introductory and Systematic Readings*. Dordrecht: D. Reidel Publishing Company.

Hilpinen, R. (ed.). (1981). *New Studies in Deontic Logic Norms, Actions, and the Foundation of Ethics*. Dordrecht: D. Reidel Publishing Company.

Kanger, S. (1957). New Foundations for Ethical Theory. Stockholm. Reprinted in *Hilpinen* (ed.) (1971), pp. 36–58.

Kracht, M. (1999). *Tools and Techniques in Modal Logic*. Amsterdam: Elsevier.

Kracht, M. & Wolter, F. (1991). Properties of Independently Axiomatizable Bimodal Logics. *The Journal of Symbolic Logic*, vol. 56, no. 4, pp. 1469–1485.

Lewis, C. I. & Langford, C. H. (1932). *Symbolic Logic*. New York: Dover Publications. Second edition 1959.

Popkorn, S. (1994). *First Steps in Modal Logic*. Cambridge University Press.

Rescher, N. (ed.). (1967). *The Logic of Decision and Action*. Pittsburgh: University of Pittsburgh Press.

Segerberg, K. (1971). *An Essay in Classical Modal Logic*. 3 vols. Uppsala: University of Uppsala.

Zeman, J. J. (1973). *Modal Logic: The Lewis-Modal Systems*. Oxford: Clarendon Press.

Rönnedal, D. (2010). *An Introduction to Deontic Logic*. Charleston, SC.

Rönnedal, D. (2012). Bimodal Logic. *Polish Journal of Philosophy*. Vol. VI, No. 2, pp. 71–93.

Rönnedal, D. (2012b). *Extensions of Deontic Logic: An Investigation into some Multi-Modal Systems*. Department of Philosophy, Stockholm University.

Rönnedal, D. (2015). Alethic-Deontic Logic: Some Theorems. *Filosofiska Notiser*, Årgång 2, Nr. 1, pp. 61–77.

Rönnedal, D. (2015b). Alethic-Deontic Logic: Deontic Accessibility Defined in Terms of Alethic Accessibility. *Filosofiska Notiser*, Årgång 2, Nr. 3, pp. 3–26.

Åqvist, L. (1987). *Introduction to Deontic Logic and the Theory of Normative Systems*. Naples: Bibliopolis.

Åqvist, L. (2002). Deontic Logic. In Gabbay & Guenthner (eds.). *Handbook of Philosophical Logic 2nd Edition*, Vol. 8, Dordrecht: Kluwer Academic Publishers, pp. 147–264.

Daniel Rönnedal
Department of Philosophy
Stockholm University
daniel.ronnedal@philosophy.su.se

Allmänna Normer och Strukturen hos Normativa System: En Logisk Analys

Daniel Rönnedal

Abstrakt
Den här uppsatsen handlar om *allmänna* eller *generella* eller *universella normer* och *strukturen hos normativa system*. Allmänna normer är normer som uttalar sig om *alla* entiteter eller individer eller fenomen av ett visst slag. Men vilken logisk form har de? Kan de användas för att härleda andra generella normer och normer som handlar om enskilda individer? Det tycks förekomma åtminstone två olika typer av föreskrifter av denna typ: normer där vi kvantifierar över *handlingar* eller *beteenden* och normer där vi kvantifierar över *personer*, *människor* eller *levande eller medvetna varelser*. I den här uppsatsen undersöker jag den logiska formen hos dessa. Jag koncentrerar mig på den senare typen och visar hur det är naturligt att använda en kvantifierad deontisk logik för att symbolisera föreskrifter av detta slag och förstå deras logiska form. Jag beskriver hur det är möjligt att använda allmänna normer för att härleda andra allmänna eller partikulära normer och hur man med hjälp av en eller flera allmänna föreskrifter kan bygga upp ett helt normativt system.

1. Introduktion

Den här uppsatsen handlar om *allmänna* eller *generella* eller *universella normer* och *strukturen hos normativa system*. Allmänna normer är normer som uttalar sig om *alla* entiteter eller individer eller fenomen av ett visst slag. Det tycks förekomma åtminstone två olika typer av allmänna normer: normer där vi kvantifierar över *handlingar* eller *beteenden* och normer där vi kvantifierar över *personer*, *människor* eller *levande eller medvetna varelser*.

Här följer några exempel på den första typen.

1. Alla lögner är förbjudna.
2. Inga löftesbrott är tillåtna.
3. Alla dråp är otillåtna.

Filosofiska Notiser, Årgång 2, Nr. 3, December 2015, 69–98

Satser som dessa tycks tala om handlingar av olika slag: lögner, löftesbrott och dråp. De säger något om vilka normativa egenskaper dessa handlingar har. (1) uttrycker t.ex. att alla lögner har egenskapen att vara förbjudna; eller med andra ord, varje handling som har egenskapen att vara en lögn har egenskapen att vara förbjuden. Enskilda, partikulära handlingar kan tillhöra olika kategorier och olika kategorier av handlingar kan ha olika normativa egenskaper. För att symbolisera normativa satser av detta slag kan vi använda oss av vanlig predikatlogik. Vi kan t.ex. formalisera (1) – (3) på följande sätt: (1) $\forall x(Lx \rightarrow Fx)$, (2) $\neg\exists x(Bx \wedge Tx)$, och (3) $\forall x(Dx \rightarrow \neg Tx)$, där Lx läses ”x är en lögn”, Fx ”x är förbjuden”, Bx ”x är ett löftesbrott”, Dx ”x är ett dråp”, och Tx ”x är tillåten”.

Här följer några exempel på allmänna normer av det andra slaget, där vi kvantifierar över personer, människor eller levande eller medvetna varelser.

4. Alla bör vara ärliga.
5. Ingen får utsättas för tortyr eller omänsklig bestraffning.
6. Alla bör hålla sina löften.
7. Ingen får våldta någon.
8. Alla bör göra något för att hjälpa människor i nöd.
9. Det är fel om någon utsätts för mobbning.

När vi talar om ”alla” och ”ingen” i dessa satser, så tycks vi kvantifiera över människor, personer eller levande eller medvetna varelser. Det är naturligt att tolka ”alla” som ”alla människor (personer, levande eller medvetna varelser)” och ”ingen” som ”ingen människa (person, levande eller medveten varelse)”; detsamma gäller interpretationen av ”någon”. Om man vill vara noggrann när man formulerar olika allmänna normer av detta slag, bör man explicit ange vad man kvantifierar över, om det är alla människor, personer, levande eller medvetna varelser eller något annat. Ofta framgår det av kontexten vad en viss talare, författare eller avsändare avser. Vissa generella normer kan vara begränsade på olika sätt. Betrakta t.ex. följande tillåtelse:

10. Alla svenska medborgare får rösta i valet.

Denna norm ger inte alla människor rätt att rösta i valet utan endast alla svenska medborgare. Satsen uttrycker likväl en allmän norm, eftersom den talar om *alla* individer av ett visst slag.

I den här uppsatsen undersöker jag den logiska formen hos allmänna normer. Jag koncentrerar mig på den senare typen och visar hur det är naturligt att använda en kvantifierad deontisk logik för att symbolisera sådana normer och förstå deras logiska form.

Allmänna normer kan också kallas för ”moraliska regler” eller ”moraliska principer”. Dessa begrepp kan dock nyttjas i flera olika betydelser. Så, låt mig säga lite mer om hur uttrycken ifråga används i den här uppsatsen.[1]

Ibland antas moraliska regler vara nödvändiga och gälla för alltid, vid alla tidpunkter. Vi skall inte göra några sådana antaganden i den här uppsatsen. Den allmänna normen att ingen får röka på offentliga platser kan t.ex. gälla vid en viss tidpunkt även om den inte gäller vid alla tidpunkter. På samma sätt kan man tänka sig att det är sant att ingen får röka på offentliga platser i någon möjlig värld, även om det inte är sant i alla möjliga världar. (Se avsnitt 2, för mer om detta.)

Vi skall inte heller anta att allmänna normer behöver vara *sanna*. Vi kan tala om *falska* allmänna normer. Vi kan t.ex. säga att följande sats *är* eller *uttrycker* en allmän norm: ”Ingen får borsta tänderna”, även om väl ingen någonsin har argumenterat för en sådan norm eller tror att den är sann. Vi är inte i den här uppsatsen intresserade av *sanningsvärdena* hos olika allmänna normer, utan av *den logiska formen hos* föreskrifter av detta slag.

Vi antar inte heller att en moralisk regel måste vara undantagslös. Det kan finnas allmänna regler som har undantag. Vi kan kalla en sådan princip för en *tumregel*. En tumregel är, enligt detta språkbruk, en allmän norm som kan ha undantag. Om denna regel har undantag, så innebär det att inte alla instanser av principen är sanna. Att det ändå handlar om en tum*regel* innebär att åtminstone många eller de flesta instanser är sanna. ”Ingen får ljuga” uttrycker kanske en sådan tumregel. Oavsett om det finns instanser av regeln att ingen får ljuga som är falska eller inte, så betraktar vi denna sats som en allmän norm i den här uppsatsen. Allmänna normer kan även explicit innehålla olika undantag eller villkor, som i följande fall: ”Ingen *som inte har körkort* får köra bil”.

Inte sällan backas allmänna normer upp av olika sanktioner. Man kan t.ex. få straff för olika brott mot lagen. Vi skall dock inte anta att brott mot allmänna normer nödvändigtvis måste ha några påföljder. Inte heller skall vi anta att allmänna normer i regel efterlevs. En allmän norm kan vara sann även om ingen någonsin följer den.

Vi är i den här uppsatsen framförallt intresserade av *den logiska formen hos* allmänna normer och *strukturen hos* normativa system. Moralfilosofer

[1] För mer information om olika betydelser av ”moraliska regler”, ”standarder” och ”principer”, se Goldman (2013).

talar ofta om moraliska regler och principer, men förvånansvärt lite har sagts om den här artikelns huvudtema i den filosofiska litteraturen. Följande uppsats är därför mer än väl motiverad.

Uppsatsen är indelad i fem avsnitt. I avsnitt 2 säger jag mer om den logiska formen hos allmänna och partikulära normer och jag visar hur det är möjligt att härleda enskilda normer ur generella normer. Avsnitt 3 handlar om hur det är möjligt att härleda allmänna normer ur allmänna normer, med eller utan information om olika nödvändiga samband. I avsnitt 4 visar jag hur man med hjälp av en eller flera allmänna normer kan bygga upp ett helt normativt system. Jag tar också upp frågan om hur man möjligtvis kan berättiga de första principerna eller de mest grundläggande normerna i sådana system. Avsnitt 5 innehåller en kort sammanfattning och slutsats.

2. Härledning av partikulära normer ur allmänna normer

Det är viktigt att klargöra skillnaden mellan allmänna och *partikulära*, *enskilda* eller *singulära* normer. Enskilda normer uttalar sig inte om *alla* entiteter av ett visst slag, utan handlar om *enskilda* individer. Följande satser är exempel på partikulära normer.

11. Lisa bör tala sanning.
12. Det är förbjudet att Johan misshandlar Erik.
13. Det är tillåtet att Sandra röstar i valet.

(11) handlar t.ex. om vad *Lisa* bör göra, inte om vad någon annan bör göra. Normer av detta slag kan generaliseras med avseende på flera olika faktorer: tid, plats, situation/möjlig värld, individ etc. Betrakta sats (11). Den allmänna norm som svarar mot denna enskilda norm uttrycks bäst av följande sats:

14. Alla bör tala sanning.

Här är ett par andra allmänna normer som påminner om (14).

15. Du bör tala sanning.
16. Alla lögner är fel.

Om vi antar att ”du” inte refererar till någon speciell person, så tycks (15) vara ekvivalent med (14). Om ”du” däremot refererar till en enskild individ, så är (15) en partikulär norm. Ibland tycks det som om satser av detta slag används för att uttrycka normer som är nödvändiga, omnitemporalt sanna och sanna överallt. ”Alla lögner är fel” antas då handla om absolut *alla* lögner, *överallt*, vid *alla* tidpunkter och i *alla* möjliga världar. På samma sätt kan ”Alla” i (14) antas handla om *alla* möjliga personer i *alla* möjliga situationer, vid *alla* tidpunkter och platser. Det är möjligt att tolka satserna på detta sätt, men vill man vara mer noggrann är det bättre att uttrycka sig mer precist. ”Alla” behöver inte förstås så generellt. Uttryckets räckvidd kan begränsas av kontexten.

Låt oss betrakta ett exempel som inte handlar om normer. Om någon tittar ut över en parkeringsplats där alla bilar är blåa och hävdar att alla bilar är blåa, så är det rimligt att tolka detta påstående på ett sådant sätt att det handlar om alla bilar som står på parkeringsplatsen vid just detta tillfälle. Påståendet är inte nödvändigtvis falskt om det finns bilar på någon annan plats som inte är blåa, eller om inte alla bilar på parkeringsplatsen är blåa vid någon annan tidpunkt, t.ex. om en vecka eller för en månad sedan, eller om det i någon annan möjlig situation (värld) är fallet att inte alla bilar på parkeringsplatsen är blåa. På liknande sätt kan det förhålla sig med normer.

Påståendet att alla bör tala sanning medför inte att det är *nödvändigt* att alla bör tala sanning, eller att det är sant *vid alla tidpunkter*, eller *på alla platser*. Det är möjligt att det är sant att alla bör tala sanning även om det inte är nödvändigt sant. Det är möjligt att det är sant att alla bör tala sanning även om det inte är sant vid alla tidpunkter. Det är möjligt att det är sant att alla bör tala sanning även om det inte är sant på alla platser. När jag talar om *allmänna* eller *generella* eller *universella* normer i den här uppsatsen menar jag normer som handlar om alla *personer*, *människor*, *individer*, eller *levande eller medvetna varelser*. Sådana normer behöver inte vara nödvändiga, sanna vid alla tidpunkter, platser osv., även om de *kan* vara det. Om man menar att en norm är nödvändig osv. bör man explicit uttrycka detta, t.ex. på följande sätt: ”Det är *nödvändigt* att alla bör tala sanning”.

Med hjälp av en kvantifierad deontisk logik kan vi visa att enskilda normer följer ur allmänna normer; (14) medför t.ex. (11), dvs. följande argument är giltigt.

Argument 1
Alla bör tala sanning.
Alltså bör Lisa tala sanning.

Vi kan symbolisera detta argument i en kvantifierad deontisk logik på följande sätt: $\forall xOTx : OTl$, där ”Tx” läses ”x talar sanning” och ”l” refererar till Lisa. För att visa att ett argument är giltigt kommer vi att använda s.k. semantiska tablåsystem. Den intuitiva tanken är att en slutledning är giltig om och endast om det är omöjligt att slutsatsen är falsk om alla premisser är sanna. För att visa att ett argument är giltigt antar vi därför att alla premisser är sanna och slutsatsen falsk. Om detta antagande leder till en motsägelse, kan vi sluta oss till att resonemanget är giltigt. För mer information om kvantifierad deontisk logik och de semantiska tablåsystem som används i den här uppsatsen, se Rönnedal (2012), (2012b), (2015). O, P och F är satsoperatorer som tar välformade formler som argument och ger välformade formler som värde. ”OA” läses ”Det är obligatoriskt att A” eller ”Det bör vara fallet att A”, ”PA” läses ”Det är tillåtet att A” eller ”Det får vara fallet att A” och ”FA” läses ”Det är förbjudet att A” eller ”Det är fel att A”. I den här uppsatsen bortser vi dock från att dessa system är inbäddade i en temporal dimension. Vi använder i regel possibilistiska kvantifikatorer om ingenting annat anges. Följande semantiska tablå bevisar att argument 1 är giltigt.

$$
\begin{array}{c}
\forall xOTx, 0 \\
\neg OTl, 0 \\
P\neg Tl, 0 \\
0s1 \\
\neg Tl, 1 \\
OTl, 0 \\
Tl, 1 \\
*
\end{array}
$$

I det här beviset har vi antagit att “Alla” i premissen varierar över alla objekt i vår domän. Om vi antar att ”Alla” varierar över alla människor, måste vi lägga till premissen att Lisa är en människa för att argumentet skall bli giltigt. Detta argument ser ut på följande sätt.

Argument 2
Alla människor bör tala sanning.
Lisa är en människa.
Alltså bör Lisa tala sanning.

Om det är uppenbart att Lisa är en människa och vi begränsar våra kvantifikatorer till människor, är det ofta onödigt att explicit omnämna denna premiss. Här följer ett annat exempel på ett giltigt argument som innehåller en allmän norm som uttalar sig om en begränsad mängd individer.

Argument 3
Alla svenska medborgare får rösta i valet.
Patrik är en svensk medborgare.
Alltså får Patrik rösta i valet.

Detta argument kan formaliseras på följande sätt: $\forall x(Sx \rightarrow PRx)$, Sp : PRp, där ”Sx” läses ”x är en svensk medborgare”, ”Rx” ”x röstar i valet”, och ”p” refererar till Patrik.

$$\forall x(Sx \rightarrow PRx), 0$$
$$Sp, 0$$
$$\neg PRp, 0$$
$$O\neg Rp, 0$$
$$Sp \rightarrow PRp$$
$$PRp, 0$$
$$0s1$$
$$Rp, 1$$
$$\neg Rp, 1$$
$$*$$

Bevisen av dessa argument är enkla och kräver egentligen inte att vi inför en speciell kvantifierade deontisk logik. De skulle kunna formaliseras i vanlig predikatlogik, om vi betraktar ”bör tala sanning” och ”får rösta i valet” som monadiska predikat. Sådana formaliseringar blottlägger dock inte premissernas logiska form lika fullständigt som symboliseringarna ovan och inte alla slutledningar är av detta relativt enkel slag.

Låt oss undersöka några fler argument med premisser och slutsats som har en något mer komplex form.

Argument 4
Om man är berusad, så är det förbjudet att man kör bil.
Gunnar är berusad.
Alltså är det inte tillåtet att Gunnar kör bil.

Argument 4 kan i en kvantifierad deontisk logik symboliseras på följande sätt: $\forall x(Bx \rightarrow FKx)$, Bg : $\neg$PKg, där ”Bx” läses ”x är berusad”, ”Kx” ”x kör bil" och "g" refererar till Gunnar. Hela resonemanget tolkas alltså på följande vis. Premiss 1: Det gäller för alla x att om x är berusad, så är det förbjudet att x kör bil. Premiss 2: Gunnar är berusad. Slutsats: Det är inte tillåtet att Gunnar kör bil. Argumentet är intuitivt giltigt. Följande semantiska tablå visar att slutsatsen följer ur premisserna.

$$
\begin{array}{c}
\forall x(Bx \to FKx), 0 \\
Bg, 0 \\
\neg\neg PKg, 0 \\
PKg, 0 \\
Bg \to FKg, 0 \\
FKg, 0 \\
O\neg Kg, 0 \\
0s1 \\
Kg, 1 \\
\neg Kg, 1 \\
*
\end{array}
$$

Argument 5 nedan är också giltigt och kan bevisas på liknande sätt. Här är en symbolisering av premisser och slutsats: $\forall x(PKx \to \neg Bx)$, Bg : FKg. Predikat och termer tolkas som ovan. Beviset lämnas till läsaren.

Argument 5
Det är tillåtet att man kör bil endast om man inte är berusad.
Gunnar är berusad.
Alltså är det förbjudet att Gunnar kör bil.

Låt oss avluta det här avsnittet med ytterligare ett exempel på ett giltigt argument med en partikulär norm som slutsats.

Argument 6
Ingen som saknar körkort får köra bil.
Alla som är under 18 år saknar körkort.
Karin är under 18 år.
Alltså är det förbjudet att Karin kör bil.

Argument 6 kan symboliseras på följande sätt: $\neg\exists x(Sx \wedge PKx)$, $\forall x(Ux \to Sx)$, Uk : FKk, där "Sx" läses "x saknar körkort", "Kx" "x kör bil", "Ux" "x är under 18 år” och ”k” refererar till Karin. Argument 6 är intuitivt giltigt och med hjälp av en kvantifierad deontisk logik kan vi bevisa detta.

$$\neg\exists x(Sx\wedge PKx),\ 0$$
$$\forall x(Ux\rightarrow Sx),\ 0$$
$$Uk,\ 0$$
$$\neg FKk,\ 0$$
$$\forall x\neg(Sx\wedge PKx),\ 0$$
$$PKk,\ 0$$
$$Uk\rightarrow Sk,\ 0$$
$$Sk,\ 0$$
$$\neg(Sk\wedge PKk),\ 0$$
$$\swarrow\quad\searrow$$

$\neg Sk,\ 0$	$\neg PKk,\ 0$
*	$O\neg Kk,\ 0$
	$0s1$
	$Kk,\ 1$
	$\neg Kk,\ 1$
	*

Vi har nu gått igenom ett antal exempel som visar hur man kan använda allmänna normer för att härleda enskilda normer. I flera fall krävs även ”faktiska” premisser för att slutsatsen skall följa. Argumenten ovan är intuitivt giltiga och kan enkelt bevisas med hjälp av en kvantifierad deontisk logik. Men allmänna normer kan också användas för att härleda (andra) allmänna normer. I nästa avsnitt skall vi se några exempel på detta.

3. Härledning av allmänna normer ur allmänna normer

I det här avsnittet visar vi hur man kan använda allmänna normer för att härleda (andra) allmänna normer. Vi börjar med ett par exempel som inte kräver några extra premisser. Sedan skall vi ta upp några argument som använder nödvändiga implikationer.

> Argument 7
> Om någon människa är oskyldig, så är det förbjudet att hon straffas.
> Det följer att ingen oskyldig människa får straffas.

Argument 7 är intuitivt giltigt. Både premissen och slutsatsen är generella normer, de uttalar sig om *alla* oskyldiga människor respektive *ingen* oskyldig människa. Argumentet kan symbolisera på följande sätt i en kvantifierad deontisk logik: $\forall x(Ox \to FSx) : \neg\exists x(Ox \wedge PSx)$, där ”Ox” läses ”x är en oskyldig människa” och ”Sx” läses ”x straffas”. Följande semantiska tablå bevisar att detta argument är giltigt.

$\forall x(Ox \to FSx), 0$
$\neg\neg\exists x(Ox \wedge PSx), 0$
$\exists x(Ox \wedge PSx), 0$
$Oc \wedge PSc, 0$
$Oc, 0$
$PSc, 0$
$Oc \to FSc, 0$
$FSc, 0$
$O\neg Sc, 0$
$0s1$
$Sc, 1$
$\neg Sc, 1$
*

Argument 8 nedan kan bevisas på liknande sätt. Här är en symbolisering: $\neg\exists x(Ox \wedge PSx) : \forall x(Ox \to O\neg Sx)$. Symbolerna tolkas som ovan.

Argument 8
Ingen oskyldig människa får straffas.
Det följer att om någon människa är oskyldig, så är det obligatoriskt att hon inte straffas.

$\neg\exists x(Ox \wedge PSx), 0$
$\neg\forall x(Ox \to O\neg Sx), 0$
$\forall x\neg(Ox \wedge PSx), 0$
$\exists x\neg(Ox \to O\neg Sx), 0$
$\neg(Oc \to O\neg Sc), 0$
$Oc, 0$
$\neg O\neg Sc, 0$
$P\neg\neg Sc, 0$
$\neg(Oc \wedge PSc), 0$
↙ ↘

$\neg Oc, 0$	$\neg PSc, 0$
*	$O\neg Sc, 0$
	$0s1$
	$\neg\neg Sc, 1$
	$\neg Sc, 1$
	*

Även argument 8 är intuitivt giltigt och innehåller en generell norm som premiss och en generell norm som slutsats. Det är ett positivt besked att vi kan använda en kvantifierad deontisk logik för att bevisa detta. Att argumenten ovan är giltiga är kanske inte så förvånande. Faktum är att premissen i argument 7 är logiskt ekvivalent med slutsatsen både i argument 7 och 8. Alla de allmänna normerna (17) – (19) nedan är alltså logiskt ekvivalenta.

17. Ingen oskyldig människa får straffas.
18. Om någon människa är oskyldig, så är det förbjudet att hon straffas.
19. Om någon människa är oskyldig, så är det obligatoriskt att hon inte straffas.

Vi skall nu undersöka hur man kan härleda allmänna normer från allmänna normer med hjälp av nödvändiga implikationer. I alla bevis nedan antar vi att vårt tablåsystem innehåller regeln T-MO (Rönnedal (2012)). Denna regel svarar mot det semantiska villkoret att den deontiska tillgänglighetsrelationen är inkluderad i den aletiska tillgänglighetsrelationen. I alla system som innehåller T-MO kan man bevisa att den s.k. mål-medel principen är giltig. Enligt denna princip följer det att varje nödvändig konsekvens av något som är obligatoriskt också är obligatorisk, eller – med andra ord – om det bör vara fallet att A och det är nödvändigt att om A så B, så bör det vara fallet att B. Detta är en intuitivt rimlig princip, som är mycket användbar.

Betrakta följande allmänna norm.

20. Ingen får mörda någon.

Denna norm, eller åtminstone någon norm som liknar denna norm väldigt mycket, tycks förekomma i alla kända normativa (juridiska och moraliska) system. Vi skall se hur man kan använda (20) för att härleda en mängd andra generella normer. Först skall vi emellertid säga något om den logiska formen hos denna föreskrift. Låt ”Mxy” vara ett två-ställigt predikat som läses ”x mördar y”. Då kan (20) symboliseras på följande sätt i en kvantifierad deontisk logik: $\neg\exists x\exists yPMxy$. Denna sats säger: ”Det är inte fallet att det finns ett x sådant att det finns ett y sådant att det är tillåtet att x mördar y”. $\neg\exists x\exists yPMxy$ är ekvivalent med (20b) $\neg\exists xP\exists yMxy$, (20c) $\neg P\exists x\exists yMxy$, (20d) $F\exists x\exists yMxy$, (20e) $\forall x\neg\exists yPMxy$, (20f) $\forall x\neg P\exists yMxy$, och (20g) $\forall xF\exists yMxy$. (20b) läses: ”Det är inte fallet att det finns ett x sådant att det är tillåtet att det finns ett y sådant att x mördar y”. (20c) läses: ”Det är inte tillåtet att det finns ett x sådant att det finns ett y sådant att x mördar y”. (20d) läses: ”Det är förbjudet att det finns ett x sådant att det finns ett y sådant att x mördar y”. (20e) läses: ”Det gäller för alla x att det inte är fallet att det finns ett y sådant att det är tillåtet att x mördar y”. (20f) läses: ”Det gäller för alla x att det inte är tillåtet att det finns ett y sådant att x mördar y”. Och (20g) läses: ”Det gäller för alla x att det är förbjudet att det finns ett y sådant att x mördar y”. Dessa ekvivalenser visar på ett tydligt sätt att (20) är en allmän norm. (20) handlar inte om enskilda personer, utan om *alla* individer. Däremot kan man givetvis använda (20) för att härleda en mängd enskilda normer, t.ex. att det är förbjudet att Björn mördar Harald, att det inte är tillåtet att Conny mördar Mark, och att det är förbjudet att Anna mördar sig själv (dvs. det följer att det är förbjudet att Anna begår självmord). (20) säger därför inte samma sak som följande norm.

21. Ingen får mörda någon annan (än sig själv).

(20) utesluter att det är tillåtet att Anna begår självmord, men det gör inte (21). (21) är förenlig med att det är tillåtet att Anna tar sitt liv, dvs. mördar sig själv. Det tycks finnas två möjliga läsningar av (21). Enligt den första medför (21) att det är tillåtet att begå självmord; enligt den andra medför (21) inte detta. Givet den första läsningen kan (21) symboliseras på följande sätt: $\neg\exists xP\exists y(\neg y = x \land Mxy) \land \forall xPMxx$; givet den andra på följande sätt: $\neg\exists xP\exists y(\neg y = x \land Mxy)$. Om (21) tolkas på det andra sättet, så följer (21) ur (20), men inte tvärtom. Dvs. om ingen får mörda någon, så följer det att ingen får mörda någon annan (än sig själv). När det kommer till kritan, kan man dock fråga sig om inte den andra tolkningen är bättre. I så fall kan man hävda att (21) *pragmatiskt implicerar* att det är tillåtet att begå självmord, men att (21) inte *medför* detta. Följande yttrande tycks inte vara inkonsistent: ”Ingen får mörda någon annan (än sig själv). Faktum är att ingen får mörda någon.” (21) *utesluter* då inte att det är tillåtet att ta sitt liv, men *medför* det *inte*. Oavsett hur det förhåller sig med detta, skall vi koncentrera oss på (20).

Betrakta följande argument.

Argument 9
Ingen får mörda någon.
Det är nödvändigt att om x dränker y, så mördar x y.
Alltså får ingen dränka någon.

Den andra premissen skall tolkas på ett sådant sätt att den gäller för alla x och y. Argument 9 är intuitivt giltigt. Det tycks vara omöjligt att premisserna är sanna och slutsatsen falsk. Hur skulle det kunna vara sant att ingen får mörda någon och att det är nödvändigt att om x dränker y, så mördar x y, samtidigt som det är falskt att ingen får dränka någon, dvs. samtidigt som det är sant att någon får dränka någon? Använder vi klassisk logik eller deontisk logik utan predikatlogik, tycks vi inte kunna symbolisera detta argument på ett sådant sätt att slutsatsen följer ur premisserna. Med hjälp av en kvantifierad deontisk logik kan vi dock bevisa att argument 9 är giltigt. Detta talar för behovet av en logik av den typ som används i den här uppsatsen. Argument 9 kan symboliseras på följande sätt i en kvantifierad deontisk logik: $\neg\exists x\exists y PMxy$, $\forall x\forall y\Box(Dxy \rightarrow Mxy)$: $\neg\exists x\exists y PDxy$, där ”Mxy” tolkas som ovan, ”Dxy” läses: ”x dränker y”, och ”□A” läses: ”Det är (historiskt) nödvändigt att A”. Följande semantiska tablå bevisar att argument 9 är giltigt.

$$
\begin{array}{c}
\neg\exists x\exists y PMxy, 0 \\
\forall x\forall y\Box(Dxy \rightarrow Mxy), 0 \\
\neg\neg\exists x\exists y PDxy, 0 \\
\exists x\exists y PDxy, 0 \\
\exists y PDcy, 0 \\
PDcd, 0 \\
\forall x\neg\exists y PMxy, 0 \\
\neg\exists y PMcy, 0 \\
\forall y\neg PMcy, 0 \\
\neg PMcd, 0 \\
O\neg Mcd, 0 \\
\forall y\Box(Dcy \rightarrow Mcy), 0 \\
\Box(Dcd \rightarrow Mcd), 0 \\
0s1 \\
Dcd, 1 \\
\neg Mcd, 1 \\
0r1 \\
Dcd \rightarrow Mcd, 1 \\
Mcd, 1 \\
*
\end{array}
$$

Vi har i symboliseringen av argument 9 använt *historisk* nödvändighet. Argumentet går även igenom om vi använder t.ex. s.k. *absolut* nödvändighet (Rönnedal (2012c)). Det är lätt att se att så är fallet. För om någonting är absolut nödvändigt, så är det historiskt nödvändigt. Premissen tycks vara rimlig, om vi talar om historisk nödvändighet. Men kanske är det t.o.m. absolut nödvändigt att om x dränker y, så mördar x y. Oavsett hur det förhåller sig med detta, så är vi här intresserade av argumentets *giltighet*, inte av deduktionens *hållbarhet*. Och ett argument kan, som bekant, vara *giltigt* även om premisserna inte är sanna.

Vi skall nu se hur den allmänna normen (20), Ingen får mörda någon, kan användas för att härleda en mängd andra allmänna normer. Alla nödvändiga implikationer (22) – (25) nedan tycks vara sanna.

22. Det är nödvändigt att om x skjuter ihjäl y, så mördar x y.
23. Det är nödvändigt att om x stryper y till döds, så mördar x y.
24. Det är nödvändigt att om x halshugger y, så mördar x y.
25. Det är nödvändigt att om x ger y en dödlig dos gift, så mördar x y.

Någon kanske vill invända att det inte är nödvändigt att x mördar y om x skjuter ihjäl y, eftersom det är möjligt att x skjuter ihjäl y av misstag. Låt oss därför anta att ”x skjuter ihjäl y” i detta sammanhang innebär att x skjuter ihjäl y med uppsåt att döda y; övriga implikationer ovan skall tolkas på samma sätt. Dessa nödvändiga samband tillsammans med den allmänna normen (20) kan t.ex. användas för att härleda de allmänna normerna (26) – (29) nedan.

26. Ingen får skjuta ihjäl någon.
27. Ingen får strypa någon till döds.
28. Ingen får halshugga någon.
29. Ingen får ge någon en dödlig dos gift.

(20) och (22) medför (26); (20) och (23) medför (27) osv. Bevisen ser likadana ut som beviset för argument 9 ovan. Faktum är att man med samma metod kan bevisa att alla sätt att mörda någon är förbjudna.

Vi skall gå igenom ytterligare några giltiga argument som innehåller allmänna normer. Betrakta argument 10.

Argument 10
Ingen får utsättas för tortyr.
Varje individ som utsätts för skendränkning utsätts nödvändigtvis för tortyr.
Alltså får ingen utsättas för skendränkning.

Argument 10 är intuitivt giltigt. Skendränkning betraktas ofta som en form av tortyr. Om detta är riktigt, är den andra premissen rimlig. Den första premissen är intuitivt mycket plausibel. FN:s allmänna förklaring om de mänskliga rättigheterna innehåller t.ex. ett förbud mot tortyr (se också Rönnedal (2014)). Följande semantiska tablå bevisar att argument 10 är giltigt. Slutledningen symboliseras på följande sätt: $\neg\exists xPTx$, $\forall x\Box(Dx \rightarrow Tx)$: $\neg\exists xPDx$, där ”Tx” läses ”x utsätts för tortyr” och ”Dx” läses ”x utsätts för skendränkning”.

$$\neg\exists xPTx, 0$$
$$\forall x\Box(Dx \rightarrow Tx), 0$$
$$\neg\neg\exists xPDx, 0$$
$$\forall x\neg PTx, 0$$
$$\exists xPDx, 0$$
$$PDc, 0$$
$$\neg PTc, 0$$
$$O\neg Tc, 0$$
$$\Box(Dc \rightarrow Tc), 0$$
$$0s1$$
$$Dc, 1$$
$$\neg Tc, 1$$
$$0r1$$
$$Dc \rightarrow Tc, 1$$
$$\swarrow \quad \searrow$$
$$\neg Dc, 1 \qquad\qquad Tc, 1$$
$$* \qquad\qquad\qquad *$$

Den allmänna normen att ingen får utsättas för tortyr kan på liknande sätt användas för att visa att ingen får utsättas för *någon* form av tortyr. Men den kan också användas för att härleda andra allmänna normer, t.ex. den allmänna normen att ingen får tortera någon.

Argument 11
Ingen får utsättas får tortyr.
Det gäller för alla x och y att det är nödvändigt att om x torterar y, så utsätts y för tortyr.
Det följer att ingen får tortera någon.

Argument 11 kan i en kvantifierad deontisk logik symboliseras på följande sätt: $\neg\exists xPTx$, $\forall x\forall y\Box(Txy \rightarrow Ty)$: $\neg\exists x\exists yPTxy$, där ”Txy” läses ”x torterar y” och ”Ty” läses ”y utsätts för tortyr”. Följande semantiska tablå bevisar att denna slutledning är giltig.

$$
\begin{array}{c}
\neg\exists xPTx, 0 \\
\forall x\forall y\Box(Txy \rightarrow Ty), 0 \\
\neg\neg\exists x\exists yPTxy, 0 \\
\exists x\exists yPTxy, 0 \\
\exists yPTcy, 0 \\
PTcd, 0 \\
\forall x\neg PTx, 0 \\
\neg PTd, 0 \\
O\neg Td, 0 \\
0s1 \\
Tcd, 1 \\
\neg Td, 1 \\
\forall y\Box(Tcy \rightarrow Ty), 0 \\
\Box(Tcd \rightarrow Td), 0 \\
0r1 \\
Tcd \rightarrow Td, 1 \\
\swarrow \quad \searrow \\
\begin{array}{cc}
\neg Tcd, 1 \qquad\qquad & Td, 1 \\
* \qquad\qquad & *
\end{array}
\end{array}
$$

Notera att härledda allmänna normer i sin tur kan användas för att härleda enskilda normer. Påståendet att ingen får utsättas för skendränkning medför t.ex. att Lena inte får utsättas för skendränkning, att det är förbjudet att Joakim utsätts för skendränkning och att det är obligatoriskt att Oskar inte utsätts för skendränkning osv.

Inte bara allmänna normer som uttalar sig om vad som är förbjudet eller otillåtet kan medföra andra allmänna normer, utan även generella föreskrifter som talar om vad som bör vara fallet. Här följer ett par intuitivt giltiga argument som innehåller den allmänna normen att alla bör vara ärliga.

Argument 12
Alla bör vara ärliga.
Alla som är ärliga håller nödvändigtvis sina löften.
Alltså bör alla hålla sina löften.

Argument 12 kan i en kvantifierad deontisk logik symboliseras på följande sätt: $\forall x O Äx$, $\forall x \Box(Äx \rightarrow Lx)$: $\forall x O Lx$, där ”Äx” läses ”x är ärlig” och ”Lx” läses ”x håller sina löften”. Följande semantiska tablå visar att slutsatsen följer ur premisserna.

$$
\begin{array}{c}
\forall x O Äx,\ 0 \\
\forall x \Box(Äx \rightarrow Lx),\ 0 \\
\neg \forall x O Lx,\ 0 \\
\exists x \neg O Lx,\ 0 \\
\neg O Lc,\ 0 \\
P \neg Lc,\ 0 \\
O Äc,\ 0 \\
\Box(Äc \rightarrow Lc),\ 0 \\
0s1 \\
\neg Lc,\ 1 \\
Äc,\ 1 \\
0r1 \\
Äc \rightarrow Lc,\ 1 \\
Lc,\ 1 \\
*
\end{array}
$$

Det tycks s.a.s. ligga i ärlighetens väsen att alla som är ärliga nödvändigtvis håller sina löften. Detta medför inte att det *alltid* är sant att alla ärliga personer håller sina löften eller att detta är *nödvändigt*. Men det är inte vad våra premisser säger. Premiss 2 säger bara att det gäller för alla x att det är nödvändigt att om x är ärlig så håller x sina löften. Detta kan vara sant vid en viss tidpunkt och falskt vid en annan tidpunkt eller sant i en möjlig värld och falskt i någon annan möjlig värld. Argumentets giltighet innebär bara att det är nödvändigt att om premisserna är sanna, så är också slutsatsen sann.

På samma sätt tycks det ligga i ärlighetens väsen att det är nödvändigt att om någon är ärlig, så ljuger hon inte. Ärliga människor talar sanning. Betrakta nu följande argument.

Argument 13
Alla bör vara ärliga.
Det är nödvändigt att om x är ärlig, så ljuger hon inte.
Det följer att ingen får ljuga.

Argument 13 kan i en kvantifierad deontisk logik symboliseras på följande sätt: ∀xOÄx, ∀x□(Äx → ¬Lx) : ¬∃xPLx, där ”Äx” tolkas som ovan och ”Lx” läses ”x ljuger”. Även detta argument är intuitivt giltigt. Här är ett semantiskt tablåbevis.

∀xOÄx, 0
∀x□(Äx → ¬Lx), 0
¬¬∃xPLx, 0
∃xPLx, 0
PLc, 0
OÄc, 0
□(Äc → ¬Lc), 0
0s1
Lc, 1
Äc, 1
0r1
Äc → ¬Lc, 1
¬Lc, 1
*

Vi har nu gått igenom ett antal exempel på hur både enskilda och allmänna normer kan härledas ur allmänna normer. Det borde vara tämligen uppenbart hur den grundläggande analysen kan utvidgas och tillämpas på andra generella normer. I nästa avsnitt skall vi se hur man kan konstruera hela normativa system med hjälp av en eller flera allmänna normer.

4. Normativa system

I det här avsnittet skall vi se hur man kan konstruera hela normativa system med hjälp av en eller flera allmänna normer. Först skall vi emellertid ta upp några epistemologiska eller kunskapsteoretiska frågor. Vi har visat hur man kan härleda både enskilda och allmänna normer ur allmänna normer. Det tycks därför som om vi skulle kunna använda generella normer för att berättiga våra partikulära normer eller enskilda moraliska omdömen, samt även åtminstone några icke-grundläggande normer. Betrakta t.ex. följande konversation.

Moralfilosofen.	Det är förbjudet att Mats dränker Stefan.
Skeptikern.	Varför det? (Hur vet du det?)

Moralfilosofen.	Därför att ingen får dränka någon. Och om ingen får dränka någon, så följer det att Mats inte får dränka Stefan.
Skeptikern.	Varför får ingen dränka någon? (Hur vet du att ingen får dränka någon?)
Moralfilosofen.	Därför att ingen får mörda någon och det är nödvändigt att om x dränker y, så mördar x y. Från dessa påståenden följer det att ingen får dränka någon.
Skeptikern.	Varför får ingen mörda någon? (Hur vet du att ingen får mörda någon?) …

Moralfilosofens svar på skeptikerns frågor är i varje fall relevant och berättigat. Det är nödvändigt att om ingen får dränka någon, så får Mats inte dränka Stefan. Vi har ovan bevisat att detta argument är giltigt. Så om premissen är sann, är också slutsatsen sann. Det är nödvändigt att om ingen får mörda någon och det är nödvändigt att om x dränker y så mördar x y, så får ingen dränka någon. Vi har ovan bevisat att detta argument är giltigt. Så om premisserna är sanna, är också slutsatsen sann.

Vi har dock inte sagt någonting om sanningsvärdena hos de olika premisserna och slutsatserna i våra argument. Det är lätt att se att om skeptikern fortsätter sina frågor, så kommer vi förr eller senare fram till någon allmän norm som inte kan härledas från någon annan mer grundläggande allmän norm på detta sätt. De enda alternativen tycks vara att vi åtminstone i princip kan fortsätta i oändlighet och berättiga generella principer med mer generella principer eller att det är tillåtet att vi rör oss i en cirkel och att en allmän norm kan berättiga sig själv (direkt eller indirekt via andra normer, värden och/eller faktapåståenden)[2]. Om vi bortser ifrån dessa alternativ, tycks det som om vi måste anta att det finns vissa allmänna normer som inte är berättigade genom att vara härledbara ur andra allmänna normer, om vi någonsin är berättigade att tro på några normer överhuvud taget. Men vad har vi för skäl att tro på en allmän norm som inte härleds ur en annan mer allmän norm? Finns det några allmänna normer av detta slag som är sanna? Och hur vet vi i så fall att de är sanna? Dessa frågor är epistemologiska och metaetiska snarare än logiska. Det tycks i regel vara fallet att vi *inte* kan bevisa sådana normer. Normen att ingen får mörda någon är t.ex. inte logiskt

[2] Varje sats medför sig själv. Men det är inte rimligt att anta att varje sats som medför sig själv är berättigad. För i så fall skulle alla satser överhuvudtaget vara berättigade, även t.ex. satser som är uppenbart falska och logiskt motsägelsefulla.

sann. Men hur kan vi då berättiga sådana allmänna normer? Är det någonsin förnuftigt att tro på en grundläggande generell moralisk princip?

Det finns inte utrymme att i den här uppsatsen diskutera dessa epistemologiska och metaetiska frågor i detalj. Men jag vill ändå nämna några möjliga svar.[3]

Enligt *nihilisten* och *errorteoretikern* är inga normativa satser sanna.[4] Om kunskap implicerar sanning, vilket många anser, kan vi då inte ha kunskap om några fundamentala allmänna normer.

Skeptikern förnekar i regel inte att vissa normativa satser kanske är sanna. Men hon ”hävdar” att vi inte kan ha någon moralisk kunskap eller att vi åtminstone *faktiskt* inte har någon sådan kunskap. Vi vet därför inte om det finns några grundläggande allmänna normer som är sanna.[5]

Den *fundamentistiska generalisten* menar att det finns åtminstone någon allmän norm som är fundamental och berättigad oberoende av sina logiska relationer till andra normer.[6] Vi kan t.ex. omedelbart inse att de fundamentala normerna är sanna med hjälp av en intuition (*intuitionisten*)[7], eller vårt förnuft (*rationalisten*)[8]. Eller också kan vi på ett eller annat sätt sluta oss till de allmänna normerna från partikulära normer som vi t.ex. kan ha kunskap om tack vare att vi har ett moralisk ”sinne” eller moraliska perceptioner (*moral sense-teoretikern* och *teoretikern som tror att det finns moralisk*

---

[3] För en allmän inledning till kunskapsteorin, se t.ex. Dancy (1985), Dancy och Sosa (red.). (1992), och Sosa och Kim (red.). (2000). För en introduktion till moralisk epistemologi, se Sayre-McCord (2013), Sinnott-Armstrong och Timmons (red.) (1996), och Zimmerman (2010). Se också Acton (1939), Brink (2014), Justin (2013), Machan (1982), McPherson (2013), och Mothersill (1959).

[4] Joyce (2001), Mackie (1977), Olson (2014).

[5] För en allmän introduktion till skepticismen, se t.ex. Hookway (1990). Mer information om den moraliska skepticismen hittar man bl.a. i Copp (1991), Sinnott-Armstrong (1996), (2006), (2008), Walker (1996), och Zimmerman (2013).

[6] Fundamentismen är en av de äldsta teorierna inom epistemologin. Ett antal historiskt inflytelserika moralfilosofer tycks ha varit fundamentister, t.ex. Sidgwick (1907) och Moore (1902). För en allmän inledning, se Hare (1996), Timmons (1987).

[7] En rad filosofer brukar klassificeras som intuitionister, t.ex. Samuel Clarke, John Balguy, Richard Price, William Whewell, Henry Sidgwick, G. E. Moore, W. D. Ross, C. D. Broad, H. A. Prichard, A. C. Ewing, Robert Audi (2004), Michael Huemer (2005), Derek Parfit, och Russ Shafer-Landau (2003) (se Russell (2013)). För mer information om etiska intuitioner och etisk intuitionism, se Audi (1993), (2004), Brody (1979), Bruce (2013), Dancy (2014), Huemer (2005), Ross (1930), Sandberg och Juth (2011), Shafer-Landau (2003), Shaw (1980), Singer (2005), Sinnott-Armstrong (2002), van Thiel och van Delden (2009), Väyrynen (2008).

[8] Bl.a. R. Cudworth, S. Clarke, J. Balguy. och I. Kant brukar klassificeras som rationalister (se Birondo (2013)). Se också Gill (2007), Kant (1785), och Peacocke (2004).

observation eller perception)[9], ett samvete, eller särskilda ”moraliska” erfarenheter, känslor, emotioner eller begär (t.ex. *vissa fenomenologer*)[10].

Induktivisten hävdar att vi kan härleda de grundläggande allmänna normerna med hjälp av observation, eller partikulära intuitioner eller känslor med hjälp av *enkel* induktion. Hur vet vi t.ex. att alla bör vara ärliga? Induktivisten kan använda ett induktivt argument av följande slag.

a bör vara ärlig
b bör vara ärlig,
c bör vara ärlig…
Alltså bör alla vara ärliga.

Enligt induktivisten har vi först kunskap om enskilda fall och sluter oss induktivt till den allmänna normen. Denna allmänna norm kan sedan användas för att härleda enskilda normer som kan testas mot våra moraliska observationer, intuitioner, känslor e.dyl.

Enligt *abduktivisten* eller anhängaren av den *hypotetisk-deduktiva metoden*, kan de grundläggande allmänna normerna berättigas abduktivt. Vi antar dem hypotetiskt som förklaring till våra moraliska observationer, intuitioner, känslor e.dyl. Här följer ett exempel på en abduktiv slutledning.

a bör vara ärlig
b bör vara ärlig,
c bör vara ärlig…
Den bästa förklaringen till att a bör vara ärlig, och att b bör vara ärlig, och att c bör vara ärlig… är att alla bör vara ärliga. (Om det vore sant att alla bör vara ärliga, så skulle det vara sant att a bör vara ärlig…)
Alltså bör alla vara ärliga.

Varken induktiva eller abduktiva argument är logiskt giltiga; det är möjligt att premisserna i ett induktivt eller abduktivt argument är sanna och slutsatsen falsk. Men premisserna antas likväl ge stöd åt slutsatsen. Om vi vet att premisserna är sanna, kan vi veta att slutsatsen är sann. Men vår kunskap är inte ofelbar.

---

[9] För mer information om moralisk observation, se t.ex. Cuneo (2003), McBrayer (2010a), (2010b), McGrath (2004), Prinz (2007), Savigny (1983). För några olika tolkningar av ”moralisk perception”, se t.ex. Blum (1991), (1994), Jacobson (2005), och Starkey (2006).

[10] Brentano (1889), Lemos (1989), Oddie (2005), Scheler (1913–1916). Se också Audi (1998).

Enligt den *intuitiva induktivisten*, kan vi få kunskap om de grundläggande allmänna normerna genom en *intuitiv induktion*. En *intuitiv* induktion är någonting annat än en *enkel* induktion. Enligt den intuitiva induktivisten reflekterar vi först över enskilda fall, t.ex. över att det är fel att a mördar b, att det är fel att c mördar d, att det är fel att e mördar f osv. När vi har gjort detta tillräckligt många gånger inser vi att dessa enskilda exempel endast är instanser av en mer generell regel, nämligen att ingen får mörda någon. Detta bör jämföras med hur vi kan få kunskap om logiska lagar. Vi kan t.ex. inse att den s.k. motsägelselagen (det är inte fallet att A och inte-A) är sann generellt genom att först reflektera över enskilda instanser. Detta innebär inte att motsägelselagen är en empirisk och kontingent lag. Motsägelselagen är inte berättigad genom *enkel* induktion, utan med hjälp av *intuitiv* induktion. På samma sätt förhåller det sig enligt den intuitiva induktivisten med vissa allmänna normer.

Enligt *konstruktivisten* eller *kontraktualisten* har vi kommit överens om att acceptera de grundläggande allmänna normerna eller skulle komma överens om att acceptera dem om vi vore fullständigt rationella eller skulle välja grundläggande moraliska principer bakom en slöja av okunnighet eller liknande.[11]

Enligt *reliabilisten* är de fundamentala allmänna normerna berättigade om de har sitt upphov i en tillförlitlig kunskapskälla (t.ex. förnuftet).[12]

Enligt *koherentisten* är de grundläggande generella normerna berättigade om de ingår i ett koherent, sammanhängande system, där de olika normerna ger stöd åt varandra. Enligt koherentisten är det hela systemet som bedöms. Om de grundläggande normerna medför andra allmänna normer och enskilda normer som stämmer bra överens med våra moraliska intuitioner och bildar ett koherent, konsistent, system, så är de berättigade.[13]

Vi skall inte här ta ställning till om något av dessa svar är tillfredsställande. Jag skall inte heller säga något om vilka grundläggande

[11] Milo (1993), Morris (1996). Se också Korsgaard (1995) och Rawls (1971). För en allmän introduktion till den moraliska konstruktivismen, se James (2013).

[12] Shafer-Landau (2003) inkorporerar t.ex. vissa reliabilistiska element i sin moraliska epistemologi.

[13] För mer information om koherensteorin i allmänhet, se t.ex. BonJour (1985), Ewing (1934), och Lehrer (1990). För mer specifik information om teorin inom den moraliska epistemologin, se t.ex. Brink (1989), kap. 5, DePaul (2013), Tersman (1993), och Sayre-McCord (1996). Begreppet reflektivt ekvilibrium är i sammanhanget relevant, se t.ex. Brandt (1990), Daniels (1979), DePaul (1986), (1987), (1988), (2013b), Ebertz (1993), Holmgren (1989), Kappel (2006), Rawls (1971), Schroeter (2004). Se också Tännsjö (1995).

allmänna normer som möjligtvis är berättigade. Låt oss istället undersöka hur man *i princip* kan använda allmänna normer för att konstruera ett helt normativt system. Ett sådant system kan vara *monistiskt* eller *pluralistiskt*. Om ett system innehåller exakt en grundläggande allmän norm, så är det ett *monistiskt* system. Om det innehåller flera grundläggande normer som inte kan härledas från varandra, så är det ett *pluralistiskt* system.

Här följer ett exempel på ett monistiskt system.

Monistiskt system (exempel)

GA1 Ingen får skada någon.

- HA1.1 Ingen får skada någon fysiskt.
  - HA1.1.1 Ingen får hugga någon med en kniv i magen.
    - HE1.1.1.1 Det är förbjudet att Conny hugger Johny med en kniv i magen.
  - HA1.1.2 Ingen får sparka någon i huvudet.
    - HE1.1.2.1 Det är inte tillåtet att Marika sparkar Mathilda i huvudet.
    - HE1.1.2.2 Det är förbjudet att Esbjörn sparkar Karl i huvudet.
- HA1.2 Ingen får skada någon psykiskt.
  - HA1.2.1 Ingen får sprida lögner om någon i sociala medier.
    - HE1.2.1.1 Elin får inte sprida lögner om Astrid i sociala medier.
    - HE1.1.1.2 Det är obligatoriskt att Magnus inte sprider lögner om Emma i sociala medier.
  - HA1.2.1 Ingen får skicka kränkande SMS till någon.
    - HE1.2.1.1 Mia får inte skicka kränkande SMS till Thomas.
- HA1.3 Ingen får stjäla någons egendom.
  - HA1.3.1 Ingen får stjäla någons bil.
    - HE1.3.1.1 Det är förbjudet att Tim stjäl någons bil.
- HA1.4 Ingen får förstöra någons egendom.
  - HA1.4.1 Ingen får bränna ner någons hus.
    - HE1.4.1.1 Det är inte tillåtet att Kim bränner ner någons hus.

Det här monistiska systemet består av *en* grundläggande allmän norm: ingen får skada någon. GA läses ”grundläggande allmän norm (regel)”, HA läses ”härledd allmän norm (regel)”, och HE läses ”härledd enskild norm”. Vi har här nämnt 4 härledda regler på nivå 1, 6 härledda regler på nivå 2, och 8 enskilda regler. ”HA1.1” är namnet på en allmän norm som är härledbar ur den allmänna normen GA1 tillsammans med relevanta nödvändiga implikationer, ”HA1.1.1” är namnet på en allmän norm som är härledbar ur den allmänna normen HA1.1 tillsammans med relevanta nödvändiga implikationer, ”HE1.1.1.1” är namnet på en enskild norm som är härledbar ur den allmänna normen HA1.1.1 osv. Notera att om HA1.1 är härledbar ur {GA1, P1}, och HA1.1.1 är härledbar ur {HA1.1, P2}, så är HA1.1.1 (direkt) härledbar ur {GA, P1, P2}. Härledda allmänna normer på samma nivå är i regel inte härledbara ur varandra. HA1.2 är t.ex. inte härledbar ur HA1.1, och HA1.1 är inte härledbar ur HA1.2. Systemet innehåller i princip även många andra härledbara allmänna och partikulära normer. Men det är inte möjligt, åtminstone inte praktiskt, att explicit ange alla dessa.

Låt oss nu ge ett exempel på ett möjligt pluralistiskt system.

Pluralistiskt system (exempel)

GA1 Alla bör vara ärliga.
- HA1.1 Alla bör tala sanning.
  - HE1.1.1 Albin bör tala sanning.
- HA1.2 Ingen får ljuga.
  - HE1.2.1 Cecilia får inte ljuga.
- HA1.3 Alla bör hålla sina löften.

GA2 Ingen får mörda någon.
- HA2.1 Ingen får skjuta ihjäl någon.
  - HE2.1.1 Det är förbjudet att du skjuter ihjäl din granne.
- HA2.2 Ingen får strypa någon till döds.
- HA2.3 Ingen får dränka någon.
  - HE2.3.1 Det är förbjudet att Mats dränker Stefan.
  - HE2.3.2 Det är inte tillåtet att Diana dränker sitt barn.
- HA2.4 Ingen får halshugga någon.
- HA2.5 Ingen får ge någon en dödlig dos gift.

GA3 Ingen får misshandla någon fysiskt.
- HA3.1 Ingen får sparka någon i huvudet.
- HA3.2 Ingen får slå någon på smalbenen med ett järnrör.

GA4 Ingen få misshandla någon psykiskt.
- HA4.1 Ingen får hota någon till livet.
- HA4.2 Ingen får mobba någon.
  - HE4.2.1 Det är inte rätt om Jenny mobbar Gunilla.
  - HE4.2.2 Det är fel om Erik mobbar Fredrik.
  - HE4.2.3 Du bör inte mobba din arbetskamrat.

GA5 Ingen får ha olämpliga sexuella förbindelser med någon.

HA5.1 Ingen får våldta någon.

HA5.2 Ingen får ha sex med någon minderårig.

Det här pluralistiska systemet innehåller 5 grundläggande allmänna normer: alla bör vara ärliga, ingen får mörda någon osv. ”GA”, ”HA” etc. tolkas som ovan. Vi har nämnt 14 härledda regler på nivå 1, och 8 härledda enskilda normer. Pluralister som utvecklar pluralistiska system antar i regel att de olika grundläggande allmänna normerna är oberoende, dvs. att de inte kan härledas från varandra. GA2 ovan tycks t.ex. inte följa ur GA1, och GA2 tycks inte medföra GA1. Om någon av de grundläggande reglerna kan härledas från någon annan grundläggande regel, eller om flera olika grundläggande regler kan härledas från en och samma regel, kan systemet förenklas. Möjligtvis skulle det pluralistiska systemet ovan kunna förenklas. Man skulle t.ex. kunna hävda att både GA3 och GA4 följer ur den allmänna normen: Ingen får misshandla någon. Monisten hoppas att alla allmänna normer på detta sätt skall kunna reduceras till en enda ”supernorm”. Huruvida detta är möjligt råder det delade meningar om. Klassiska utilitarister och kantianer har ofta varit monister, medan t.ex. intuitionister ofta har varit pluralister.

De normativa system som jag har nämnt ovan är knappast de bästa tänkbara och jag har inte sagt något specifikt om hur de grundläggande reglerna möjligtvis skulle kunna berättigas. Syftet med att ta upp dessa system är att visa hur man kan använda allmänna normer för att bygga upp hela normativa system. *Hur* ett sådant system bör se ut och exakt *vilka* allmänna normer det bör innehålla är frågor som jag inte skall behandla i den här artikeln.

5. Slutsats

Vi har i den här uppsatsen undersökt allmänna normer och strukturen hos normativa system. Allmänna normer är normer som uttalar sig om *alla* entiteter eller individer eller fenomen av ett visst slag. Jag nämnde att det tycks förekomma åtminstone två olika typer av föreskrifter av denna typ: normer där vi kvantifierar över *handlingar* eller *beteenden* och normer där vi kvantifierar över *personer*, *människor* eller *levande eller medvetna varelser*. Vi har i den här uppsatsen koncentrerat oss på den senare typen. Vi har sett hur allmänna normer kan användas för att härleda enskilda normer och andra universella regler och hur de kan nyttjas för att bygga upp hela normativa system. Det här talar för att allmänna normer kan användas för att berättiga partikulära normer och även vissa härledda generella principer. Jag har inte försökt besvara frågan *om* och *hur* de mest grundläggande allmänna normerna, de som inte härleds från andra universella regler, själva är berättigade, även om jag nämnde några möjliga alternativ. Vi har använt oss av en kvantifierad deontisk logik för att analysera de olika typerna av normer. Diskussionen pekar därför på nyttan av en kvantifierad deontisk logik av det slag som jag har försökt utveckla i bl.a. Rönnedal (2015).

Referenser

Acton, H. B. (1939). Moral Knowledge. *Analysis*. Vol. 7, No. 1, ss. 25–29.

Audi, R. (1993). Ethical Reflectionism. *The Monist*. Vol. 76, Nr. 3, Justification in Ethics, ss. 295–315.

Audi, R. (1996). Intuitionism, Pluralism, and the Foundations of Ethics. I Sinnott-Armstrong, W., Timmons, M. (red.). *Moral Knowledge*. Oxford: Oxford University Press, 1996, ss. 101–136.

Audi, R. (1998). The Axiology of Moral Experience. *The Journal of Ethics*. Vol. 2, Nr. 4, Intrinsic Value, ss. 355–375.

Birondo, N. (2013). Rationalism in Ethics. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 4329–4338.

Blum, L. (1991). Moral Perception and Particularity. *Ethics*, Vol. 101, Nr. 4, ss. 701–725.

Blum, L. (1994). *Moral Perception and Particularity*. New York: Cambridge University Press.

BonJour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge: Harvard University Press.

Brandt, R. B. (1990). The Science of Man and Wide Reflective Equilibrium. *Ethics*, Vol. 100, Nr. 2, ss. 259–278.

Brentano, F. (1889). *The Origin of Our Knowledge of Right and Wrong*. London: Routledge & Kegan Paul, 1969.

Brink, D. (1989). *Moral Realism and the Foundations of Ethics*. New York: Cambridge University Press.

Brink, D. (2014). Principles and Intuitions in Ethics: Historical and Contemporary Perspectives. *Ethics*, Vol. 124, No. 4, ss. 665–694.

Brody, B. A. (1979). Intuitions and Objective Moral Knowledge. *The Monist*, Vol. 62, Nr. 4, Objectivity in Knowledge and Valuation, ss. 446–456.

Copp, D. (1991). Moral Skepticism. *Philosophical Studies*, Vol. 62, No. 3, ss. 203–233.

Cuneo, T. (2003). Reidian Moral Perception. *Canadian Journal of Philosophy*, vol. 33, ss. 229–58.

Dancy, J. (1985). *Introduction to Contemporary Epistemology*. Oxford: Basil Blackwell.

Dancy, J. (2014). Intuition and Emotion. *Ethics*, Vol. 124, Nr. 4, ss. 787–812.

Dancy, J., Sosa, E. (red.). (1992). *A Companion to Epistemology*. Oxford: Blackwell.

Daniels, N. (1979). Wide Reflective Equilibrium and Theory Acceptance in Ethics. *Journal of Philosophy*, vol 76, ss. 256–282.

DePaul, M. R. (1986). Reflective Equilibrium and Foundationalism. *American Philosophical Quarterly*, Vol. 23, Nr. 1, ss. 59–69.

DePaul, M. R. (1987). Two Conceptions of Coherence Methods in Ethics. *Mind*, New Series, Vol. 96, Nr. 384, ss. 463–481.

DePaul, M. R. (1988). The Problem of the Criterion and Coherence Methods in Ethics. *Canadian Journal of Philosophy*, Vol. 18, Nr. 1, ss. 67–86.

DePaul, M. R. (2013). Coherentism, Moral. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 866–875.

DePaul, M. R. (2013b). Reflective Equilibrium. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 4466–4475.

Ebertz, R. (1993). Is Reflective Equilibrium a Coherentist Model? *Canadian Journal of Philosophy*, vol 23, ss. 193–214.

Ewing, A. C. (1934). *Idealism A Critical Survey*. London: Methuen.

Gill, M. B. (2007). Moral Rationalism vs. Moral Sentimentalism: Is Morality More Like Math or Beauty? *Philosophy Compass,* vol. 2, ss. 16–30.

Goldman, A. H. Rules, Standards, and Principles. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 4676–4684.

Hare, R. M. (1996). Foundationalism and Coherentism in Ethics. I Sinnott-Armstrong, W., Timmons, M. (red.). *Moral Knowledge*. Oxford: Oxford University Press, 1996, ss. 190–214.

Holmgren, M. (1989). The Wide and Narrow of Reflective Equilibrium. *Canadian Journal of Philosophy*, Vol. 19, Nr. 1, ss. 43–60.

Jacobson, D. (2005). Seeing by Feeling: Virtues, Skills, and Moral Perception. *Ethical Theory and Moral Practice*, Vol. 8, No. 4, ss. 387–409.

James, A. (2013). Constructivism, Moral. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 1065–1079.

Joyce, R. (2001). *The Myth of Morality*. Cambridge: Cambridge University Press.

Justin P. M. (2013). A Posteriori Ethical Knowledge. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 1–5.

Kant, I. (1785). *Groundwork for the Metaphysics of Morals*. Edited and translated by Allen W. Wood. New Haven and London: Yale University Press, (2002).

Kappel, K. (2006). The Meta-Justification of Reflective Equilibrium. *Ethical Theory and Moral Practice*, Vol. 9, Nr. 2, ss. 131–147.

Korsgaard, C. (1996). *The Sources of Normativity*. Cambridge: Cambridge University Press.

Lehrer, K. (1990). *Theory of knowledge*. London: Routledge.

Lemos, N. (1989). Warrant, Emotion, and Value. *Philosophical Studies*, vol. 57, ss. 175–92.

Machan, T. R. (1982). Epistemology and Moral Knowledge. *The Review of Metaphysics*, Vol. 36, Nr. 1, ss. 23–49.

Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. London: Penguin.

McBrayer, J. (2010a). A Limited Defense of Moral Perception. *Philosophical Studies*, vol. 149, ss. 305–20.

McBrayer, J. (2010b). Moral Perception and the Causal Objection. *Ratio*, vol. 23, ss. 291–307.

McGrath, S. (2004). Moral Knowledge by Perception. *Philosophical Perspectives*, vol. 18, ss. 209–28.

McPherson, T. (2013). A Priori Ethical Knowledge. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 5–10.

Milo, R. D. (1993). Skepticism and Moral Justification. *The Monist*, Vol. 76, Nr. 3, Justification in Ethics, ss. 379–393.

Moore, G. E. (1902). *Principia Ethica*. Amherst, New York: Prometheus Books, 1988.

Morris, C. W. (1996). A Contractarian Account of Moral Justification. I Sinnott-Armstrong, W., Timmons, M. (red.). *Moral Knowledge*. Oxford: Oxford University Press, 1996, ss. 215–242.

Mothersill, M. (1959). Moral Knowledge. *The Journal of Philosophy*, Vol. 56, Nr. 19, ss. 755–763.

Oddie, G. (2005). *Value, Reality and Desire*. Oxford: Oxford University Press.

Olson, J. (2014). *Moral Error Theory: History, Critique, Defence*. Oxford: Oxford University Press.

Peacocke, C. (2004). Moral Rationalism. *Journal of Philosophy*, vol. 101, ss. 499–526.

Prinz, J. J. (2007). Can Moral Obligations Be Empirically Discovered? *Midwest Studies in Philosophy*, vol. 31, ss. 271–91.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.
Ross, W. D. (1930). *The Right and the Good*. Indianapolis: Hackett Publishing.
Russell, B. (2013). Intuitionism, Moral. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 2779–2788.
Rönnedal, D. (2012). Bimodal Logic. *Polish Journal of Philosophy*. Vol. VI, Nr. 2, ss. 71–93.
Rönnedal, D. (2012b). *Extensions of Deontic Logic: An Investigation into some Multi-Modal Systems*. Department of Philosophy, Stockholm University.
Rönnedal, D. (2012c). Temporal alethic-deontic logic and semantic tableaux. *Journal of Applied Logic*, 10, ss. 219–237.
Rönnedal, D. (2014). FN:s allmänna förklaring om de mänskliga rättigheterna och kvantifierad deontisk logik. *Tidskrift för Politisk Filosofi*. Årgång 18, Nr 2, ss. 22–34.
Rönnedal, D. (2015). Quantified Temporal Alethic Deontic Logic. *Logic and Logical Philosophy*. Vol 24, Nr 1, ss. 19–59.
Sandberg, J. och Juth, N. (2011). Ethics and Intuitions: A Reply to Singer. *The Journal of Ethics*, Vol. 15, Nr. 3, ss. 209–226.
Savigny, E. V. (1983). A Modest Concept of Moral Sense Perception. *Erkenntnis* (1975-), Vol. 19, Nr. 1/3, Methodology, Epistemology, and Philosophy of Science, ss. 331–344.
Sayre-McCord, G. (1996). Coherentist Epistemology and Moral Theory. I Sinnott-Armstrong, W., Timmons, M. (red.). *Moral Knowledge*. Oxford: Oxford University Press, 1996, ss. 137–189.
Sayre-McCord, G. (2013). Epistemology, Moral. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 1674–1688.
Singer, P. (2005). Ethics and Intuitions. *The Journal of Ethics*, vol. 9, ss. 331–52.
Scheler, M. (1913–1916). *Formalism in Ethics and Non-Formal Ethics of Values*. Northwestern University Press. Translated by Manfred S. F. and Roger L. F., 1973.
Schroeter, F. (2004). Reflective Equilibrium and Antitheory. *Noûs*, Vol. 38, Nr. 1, ss. 110–134.
Shafer-Landau, R. (2003). *Moral Realism. A Defence*. Oxford: Clarendon Press.

Shaw, W. H. (1980). Intuition and Moral Philosophy. *American Philosophical Quarterly*, Vol. 17, Nr. 2, ss. 127–134.

Sidgwick, H. (1907). *The Methods of Ethics*. Indianapolis: Hackett, 1981.

Sinnott-Armstrong, W. (1996). Moral Skepticism and Justification. I Sinnott-Armstrong, W., Timmons, M. (red.). *Moral Knowledge*. Oxford: Oxford University Press, 1996, ss. 3–48.

Sinnott-Armstrong, W. (2002). Moral Relativity and Intuitionism. *Noûs*, Vol. 36, Supplement: Philosophical Issues, 12, Realism and Relativism, ss. 305–328.

Sinnott-Armstrong, W. (2008). Moderate Classy Pyrrhonian Moral Scepticism. *The Philosophical Quarterly*, Vol. 58, Nr. 232, ss. 448–456.

Sosa, E. och Kim, J. (red.). (2000). *Epistemology An Anthology*. Blackwell.

Starkey, C. (2006). On the category of moral perception. *Social Theory and Practice*, 32(1), ss. 75–96.

Tersman, F. (1993). *Reflective Equilibrium: An Essay in Moral Epistemology*. Almqvist & Wiksell International.

van Thiel, G. J. M. W. och van Delden, J. J. M. (2009). The Justificatory Power of Moral Experience. *Journal of Medical Ethics*, Vol. 35, Nr. 4, ss. 234–237.

Timmons, M. (1987). Foundationalism and the Structure of Ethical Justification. *Ethics*, Vol. 97, Nr. 3, ss. 595–609.

Torbjörn T. (1995). In Defence of Theory in Ethics. *Canadian Journal of Philosophy*, Vol. 25, No. 4, ss. 571–593.

Walker, M. U. (1996). Feminist Skepticism, Authority, and Transparency. I Sinnott-Armstrong, W., Timmons, M. (red.). *Moral Knowledge*. Oxford: Oxford University Press, 1996, ss. 267–292.

Väyrynen, P. (2008). Some Good and Bad News for Ethical Intuitionism. *The Philosophical Quarterly*, Vol. 58, Nr. 232, ss. 489–511.

Zimmerman, A. (2010). *Moral Epistemology*. New York: Routledge.

Zimmerman, A. (2013). Skepticism, Moral. I Hugh LaFollette (red.). *The International Encyclopedia of Ethics*. ss. 4906–4916.

Daniel Rönnedal
Filosofiska institutionen
Stockholms universitet
daniel.ronnedal@philosophy.su.se

Quine and Plato's Beard Revisited

John F. Peterson

Abstract

To the extent that it allows individuating properties, Quine's answer to the puzzle of saying that Pegasus is not without assuming that Pegasus is is problematic. Alternatively, one might identify the referent of 'Pegasus' in 'Pegasus is not' with an unactualized possible. Yet, Quine's own objection that this compromises *reductio* proof seems to be decisive. So it seems that the best answer is Russell's. Unlike Quine's, it shuns individuating properties with all their attendant difficulties. Unlike Strawson's, it covers the *prima facie* truth of saying that Pegasus does not exist. And unlike Meinong's, it does both without recourse to non–existent particulars.

Dividing meaning and reference in singular terms is Quine's way of blocking commitment to an ontology containing Pegasus when we say that Pegasus is not.[1] For if a) 'Pegasus' is a name, b) 'Pegasus is not' is meaningful, c) the meaningfulness of 'Pegasus is not' requires the meaningfulness of 'Pegasus', and d) meaning and referent are identified in a name, then saying, meaningfully, that Pegasus is not implies that Pegasus is. But it implies this only if it is wrongly assumed in the first instance that the meaning of a singular term like 'Pegasus' is identified with the entity named by that term.[2] So marking off meaning and reference even in singular terms allows one to say with consistency that Pegasus is not.[3]

Quine's move translates singular terms like Pegasus into predicates. For it allows 'Pegasus is not' to be glossed as, say, 'It is not the case that there is an x such that x is a winged horse that opened the spring of Hippocrene and for

---

[1] Quine, W.V. (1961). *From a Logical Point of View*. Cambridge: Harvard, p. 9.

[2] Quine , W.V. (1961). *From a Logical Point of View*, p. 7.

[3] Meinong's way of avoiding inconsistency in saying "Pegasus is not" is to distinguish existent and subsistent objects. If 'Pegasus' names a subsistent and not an existent object, then one consistently says that Pegasus does not exist. For a defense of this distinction see Meinong, A. (1902). *Ueber Annahmen*. Leipzig: J.A. Barth, p. 74.

Filosofiska Notiser, Årgång 2, Nr. 3, December 2015, 99–106

all y if y is a winged horse that opened the spring of Hippocrene then y equals x.[4] But to this Russellian move Quine makes an addition which he illustrates in the case of Pegasus. For just in case Pegasus is so basic as to be insusceptible of analysis, Quine allows that ‘Pegasus is not’ be glossed as: ‘It is not the case that there is an x such that x is–pegasus (or pegasizes), where ‘is–pegasus’ or ‘pegasizes’ is a predicate. In any case, since under either Russell’s or Quine’s assay the alleged name ‘Pegasus’ is analyzed out without remainder, it is not implied that Pegasus is in saying that Pegasus is not. Thus what Quine calls the problem of Plato’s beard is solved.

Yet Quine’s nuance is problematic. To be true, ‘It is not the case that there is something that pegasizes’ must be meaningful. A condition of this is that the predicate ‘pegasizes’ is meaningful. Since it is like ‘is–green’ in being irreducible and unanalyzable, the property ‘pegasizes’ cannot be unpacked by using descriptive phrases. But unlike ‘is–green’, ‘pegasizes’ or ‘is–pegasis’ is not an object of acquaintance. So if it is neither analyzable nor an object of acquaintance how is ‘is–pegasus’ meaningful?

Besides, if ‘Pegasus’ is assimilated to a predicate then so too is any other singular term. And then the whole category of subject–predicate statements is swept away. That has the merit of economy. But for this logical elegance a price is paid in ontology. And that is the introduction of individuating properties. The latter go as far back as Scotus’ *haeccietas*. But the trouble with them is identifying the thing of which they are the property. Properties, individuating or otherwise, are the properties *of* something. But since all that is unique and individual is absorbed by them, there is nothing left for individuating properties to characterize but a Lockian I–know–not–what, a totally bare particular. So by allowing names to be replaced by individuating predicates, Quine invites something against which he himself recoils, i. e, bare *substrata*. An obvious answer to this is to identify an individual with a complex of properties one of which is individuating. Thus, being–Socrates is analytically predicated of a cluster of properties one of which is the individuating property of being–Socrates. And then bare particulars are avoided.

But this has troubles of its own. For one thing, it breeds circularity in the definition of an individual. For what under this assay is defined as being an individual is a bundle of properties one of which is individuating. For another, it fails to cover the unity of individuals like Socrates. The paradox is that when properties are distinguished from individuals they stand united in

[4] Quine, W.V. (1961). *From a Logical Point of View*, pp. 7–8.

individuals. For they are brought together by dint of inhering in the same subject. But when that distinction is dropped, so too is the unity. Since there is nothing to unite the properties, an individual like Socrates becomes a pile of predicates. Nor can it be said that the individuating property of being–Socrates unites the properties. As it is one of the properties to be united, it cannot be said to be what unites all the properties. Otherwise something is said to unite itself.

To avoid all this, one might try another tack. Under it, ‘Pegasus’ is a name just as it appears to be and not a disguised predicate. But what it names is the *idea* Pegasus. Then, one can say that Pegasus is not without assuming that Pegasus is. For since the ‘is not’ in that statement signifies real being and the referent of ‘Pegasus’ is mental being, then one consistently says that Pegasus is not.

But Quine himself notes the confusion in this escape.[5] Even granting this mental entity we call the idea of Pegasus, it is not *that* to which we refer when we deny that Pegasus is. So to avoid assuming that Pegasus is in saying that Pegasus is not, it will not do to say that ‘Pegasus’ in that statement names the idea Pegasus. That just misidentifies what is denied when it is denied that Pegasus is.

So what is the solution to Plato’s beard? How do you construe ‘Pegasus is not’ without either implying that Pegasus is, saying things like there’s not something that pegasizes, with all its attendant difficulties, or misidentifying what is denied when it is denied that Pegasus is?

A Strawsonian answer is that, if we only cease identifying meaning and referent in a name, dropping Russell’s logically proper names, we can construe ‘Pegasus’ as a non–naming name. And then, since it is not used to talk about anything, the sentence ‘Pegasus does not exist’ does not make an assertion in the first place and hence is neither true nor false.[6] But in that case the problem of implying that Pegasus is in saying that Pegasus is not fails to arise. For no assertion is in the first instance made. Thus, the supposed problem of non–being is dissolved.

But unlike either Meinong’s, Russell’s, or Quine’s answer, Strawson’s ploy fails to cover the *prima facie* truth of the utterance in question. Typically, when one says that Pegasus does not exist one does not use that sentence to illustrate a point in grammar, to write a line of poetry, to send a secret message or anything like that. To all appearances, one uses it

[5] Quine, W.V. (1961). *From a Logical Point of View*, p. 2.
[6] Strawson, P. F. (1950). On Referring. *Mind* 59, pp. 320–344.

straightforwardly to make a true assertion. And Strawson himself agrees that sentences that make assertions must be about something and hence be either true or false. So the better course of action is to save the appearance and then try and avoid commitment to Meinong's non–existent Pegasus.

For example, Quine, Russell, and even Meinong would remind Strawson that if someone said that Pegasus does not exist and asked you whether you thought that what he said was true or false, you would answer, "true." You would not answer, "neither." It seems, then, that 'Pegasus does not exist' can count as an assertion. But if so, then the conundrum of non–being is not dissolved after all and the problem of the referent of 'Pegasus' in 'Pegasus is not' remains.

Some might favor another answer, according to Quine.[7] It is to identify the referent of 'Pegasus' not with an idea in the sense of the mental Pegasus-idea which is something actual. For it is evidently not this mental Pegasus-idea that one denies when one denies that Pegasus is. Instead, this subtler answer identifies the referent of 'Pegasus' with an idea in the sense of a group of properties which has possible being only. It is an unactualized possible.

An unactualized possible is in the same sense of 'is' as what is defined is. That is a different sense of 'is' from that which is accorded to an actualized possible. Following tradition, one might say that one signifies essence and the other existence. For that reason it is neither inconsistent nor self–defeating to say that Pegasus is not. In saying this, one says only that the possible being that is named by the subject 'Pegasus' does not have actual being. True, one does assume here that Pegasus is in denying that Pegasus is. But since the 'is' is different each time, the statement is innocuous. It just repeats Aristotle's advice that being is said in many senses. Nor does this answer risk admitting contradictions as unactualized possibles just in case it is said, say, that the round square window is not. Since contradictory subject–terms like 'the round square window' are meaningless and genuine statements require meaningful terms, then the round square window is not assumed to be when it is said that it is not. For no genuine statement has in the first instance been made.

---

[7] Quine attributes this answer to a mind more subtle than one that would identify the referent of 'Pegasus' with the mental Pegasus-idea. He names this mind 'Wyman' but does not say either that Wyman represents a real respondent or that Wyman's answer has actually been given. See Quine, W. V. (1961). *From a Logical Point of View*, pp. 2–5.

But despite the *prima facie* appeal of this gambit, Quine, for one, rejects it. To work, it requires the doctrine of the meaninglessness of contradictions. Yet for two reasons Quine balks at that idea.[8] The first is that it threatens proof by *reductio*. In the latter, affirming the premises and denying the conclusion implies a contradiction. So if contradictions are meaningless, so too is *reductio* proof. Either, then, contradictions are not meaningless or proof by *reductio* is compromised. Second, if contradictions are meaningless, then deciding whether or not an expression is meaningful depends on knowing whether or not it is contradictory. But with Church Quine agrees that there is no generally applicable test of whether an expression is contradictory.[9] It follows that the contradictoriness of expressions is ultimately undecipherable. If you have no generally applicable test of contradictoriness and knowing whether or not expressions are meaningful hangs on that test, then you never know whether or not expressions are meaningful. But since that is unacceptable, says Quine, it follows that the assumption in question, i.e. the meaninglessness of contradictions, is false.

But if it is, concludes Quine, then no one can say that 'Pegasus' in 'Pegasus is not' names an unactualized possible. If it cannot be said that the phrase 'round square window' in 'The round square window is not' is meaningless because it is contradictory, then defenders of the solution that 'Pegasus' names an unactualized possible are forced after all to count entities like round square windows as unactualized possibles or as unactualized impossibles when it is said that the round square window is not. And that nullifies their solution to the problem of Plato's beard.

Some might object that neither one of Quine's objections to the doctrine of the meaninglessness of contradictoriness, and hence to saying that 'Pegasus' names an unactualized possible, is conclusive. Taking the objections in reverse order, even if Church is right that there is no generally applicable test of contradictoriness, that does not mean that you cannot tell whether or not expressions are meaningful when contradictions are meaningless. For suppose that a test of contradictoriness is lacking not because none can be found but because none are necessary. Then you can tell whether expressions are contradictory or not without a test. And then the doctrine of the meaninglessness of contradictions fails to imply that we cannot tell whether or not expressions are meaningful. But in that case Quine

---

[8] Quine, W.V. (1961). *From a Logical Point of View*, p. 5.

[9] Quine, W.V. (1961). *From a Logical Point of View*, p. 5. See also, Church, A. (1936). A Note on the *Entschedungsproblem*". *Journal of Symbolic Logic* 1, p. 40f,, p. 101f.

cannot use that supposed implication as grounds for denying that Pegasus names an unactualized possible.

But as a matter of fact, a case can be made for saying that contradictoriness *is* the sort of thing for which a generally applicable test is unnecessary. For suppose that contradictoriness is the sort of thing for which a test T is necessary. Then since any expression's being meaningful requires that it pass T, then that very test T, to be meaningful, must either pass itself or some higher–order test of contradictoriness, T1. But the first makes something the test of itself, from which Quine himself recoils on account of the theory of types. And the second invites an infinite regress of higher–order tests of contradictoriness.

Thus, it seems that defenders of the view that 'Pegasus' names an unactualized possible can answer Quine's second objection. The meaninglessness of contradictoriness, on which their view hangs, rules out knowing whether or not a string of symbols is meaningful only if it is conceded that a general test of contradictoriness is in the first instance required. But if only for the reasons just given, no such concession would be made by those who hold that 'Pegasus' in 'Pegasus is not' names an unactualized possible.

However, Quine's first objection to the doctrine of the meaninglessness of contradiction is more convincing. As against it, defenders of the doctrine might counter that his argument is question–begging. Quine rejects the doctrine because it rules out proof by *reductio*. He thus *uses reductio* proof to refute a view because it excludes *reductio* proof. By analogy, suppose I use an argument from analogy to refute some belief of yours because it undermines argument from analogy. You then have a right to demand that I show your belief wrong *independently* of using an argument from analogy. Otherwise you have the right to complain that I beg the question in favor of argument from analogy.

But this objection is captious. For it simply plays on the term '*reductio* proof.' When he rejects the meaninglessness of contradictions because it compromises *reductio* proof, Quine means by the latter the narrowly logical sense of '*reductio* proof'. Under it, you show that the joint assertion of an argument's premises and the denial of its conclusion as an added assumption yields a contradiction. From this you conclude that the conclusion of the argument must be true. But his own argument against the meaninglessness of contradictions is a *reductio* proof in the broader sense of the term. This consists in showing that P is false because it implies what is unacceptable, in this case, the elimination of *reductio* proof in the narrow sense. It follows that there is no circle and that the objection is answered.

Even so, defenders of the meaninglessness of contradictions might counter that Quine's first objection fails to recognize important meaninglessness. Wittgenstein, for example, recognized the importance of the mystical even though putting the mystical into words was nonsensical.[10] Similarly, some nonsense might be conducive to its own disclosure and hence be useful and to that extent important nonsense. Such is the case in *reductio* proof. The combination of the premises and the negation of the conclusion is useful nonsense because, in the context of the proof in which it figures, it is a necessary step in its own disclosure. For as the proof proceeds, the contradiction that was implicit in the foregoing combination is explicitly generated in the penultimate step of the proof. And then the validity of the argument in question is shown. Thus, practicing his own pragmatism might have restrained Quine from concluding that the meaninglessness of contradictoriness ruins *reductio* proof.

But to all of this Quine has a good answer. Even if with Wittgenstein we recognize important nonsense and even supposing that some of this is useful nonsense, it seems that Quine is right that nonsense of any sort has no place in logical proofs.

Let us take stock. Suppose that Quine is right that claiming that 'Pegasus' names an unactualized possible threatens *reductio* proof. Suppose too that as was suggested at the outset, his own solution *via* individuating properties either invites bare particulars or else both implies circularity in the definition of individuals and excludes the unity of individuals. Then what can be done? What is the solution to the problem of Plato's beard?

To close, it seems that the best answer is the one that is behind Quine's. By many it is regarded as one of the major achievements in philosophy in the Twentieth Century. Remarkably, it clings even closer to Ockam's Razor than does Quine. Moreover, it differs from Quine's only in avoiding individuating properties like "pegasizes." That is its merit. For then it entirely bypasses saying things like "There is not something that pegasizes". Not just that but it also sidesteps the dilemma that follows on the heels of those properties. I refer to Russell's answer.[11] Shunning individuating properties like "pegasizes", Russell is not then caught between admitting bare particulars

[10] Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. Translation by D. F. Pears and B. F. McGuinness. London: Routledge & Kegan Paul. 6.52–6.522, pp. 149–151.

[11] Russell, B. (1905). On Denoting. *Mind* 14, pp. 479–493.

and both causing circularity in the definition of individuals and excluding the unity of individuals. In ‘Pegasus is not’ ‘Pegasus’ is just translated as a definite description such as the winged horse that opened the spring of Hippocrene. And the resulting negative existential statement that preserves the truth of the statement is that it is not the case that there is an x such that x is a winged horse that opened the spring of Hippocrene, and for all y, if y is a winged horse that opened the spring of Hippocrene then y equals x. Unlike Strawson’s move, this allows the common sense statement, “‘Pegasus does not exist’ is true”. Unlike Quine’s, it does this without the onus of properties like pegasizes. And unlike Meinong’s, it covers that same truth without the extravagance of non–existent particulars. And so it is that Russell kills three birds with one philosophical stone.

References

Church, A. (1936). A Note on the *Entschedungsproblem*”. *Journal of Symbolic Logic* 1, p. 40f,, p. 101f.

Meinong, A. (1902). *Ueber Annahmen*. Leipzig: J.A. Barth.

Quine, W.V. (1961). *From a Logical Point of View*. Cambridge: Harvard.

Russell, B. (1905). On Denoting. *Mind* 14, pp. 479–493.

Strawson, P. F. (1950). On Referring. *Mind* 59, pp. 320–344.

Wittgenstein, L. (1961). *Tractatus Logico-Philosophicus*. Translation by D. F. Pears and B. F. McGuinness. London: Routledge & Kegan Paul.

J.F. Peterson
jfpeterson@uri.edu
University of Rhode Island

On the Signpost Principle of Alternate Possibilities: Why Contemporary Frankfurt-Style Cases are Irrelevant to the Free Will Debate

William Simkulet

Abstract
This article contends that recent attempts to construct Frankfurt-style cases (FSCs) are irrelevant to the debate over free will. The principle of alternate possibilities (PAP) states that moral responsibility requires indeterminism, or multiple possible futures. Frankfurt's original case purported to demonstrate PAP false by showing an agent can be blameworthy despite not having the ability to choose otherwise; however he admits the agent can come to that choice freely or by force, and thus has alternate possibilities. Neo-FSCs attempt to show that alternate possibilities are irrelevant to explaining an agent's moral responsibility, but a successful Neo-FSC would be consistent with the truth of PAP, and thus is silent on the big metaphysical issues at the center of the free will debate.

Introduction

Frankfurt-style cases (FSCs) are modeled after a case in Harry Frankfurt's "Alternate Possibilities and Moral Responsibility," where in an agent is purported to be uncontroversially morally responsible despite lacking the ability to do otherwise.[1] If FSCs are as advertised, they would be counter-examples to the principle of alternate possibilities (PAP), according to which one is morally responsible for something only if she could do otherwise. Much has been written about FSCs, but the general consensus is that they fail to be genuine counterexamples to PAP.[2] The reason FSCs have garnered such attention is that PAP is said to play a vital role in the debate over whether free will is consistent with determinism. Contemporary proponents of FSCs have largely abandoned the goal of constructing a counterexample to PAP, and instead aim to show merely that alternate possibilities don't play a role in determining an agent's degree of moral responsibility. This article

[1] Frankfurt 1969.
[2] See Fischer 2010 and Widerker and McKenna 2003/2006 for strong work on the topic.

Filosofiska Notiser, Årgång 2, Nr. 3, December 2015, 107–120

argues that this concession by proponents of FSCs dramatically undermines their relevance to the free will debate.

The main goal of this article is to show that while a successful traditional FSC would demonstrate the falsity of PAP, a successful Neo-FSC, sometimes called a "buffer zone" FSC[3], would be irrelevant to the truth or falsity of PAP. PAP, the principle Frankfurt claims "is false" (1969, 829), is most often interpreted as asserting that indeterminism is metaphysical prerequisite for true moral responsibility[4]. Frankfurt says of PAP that "Its exact meaning is a subject of controversy, particularly concerning whether someone who accepts it is thereby committed to believing that moral responsibility and determinism are incompatible." (1969, 829) While traditional FSCs are meant to be genuine counterexamples to PAP[5], Neo-FSCs attempt to show only that alternate possibilities are irrelevant to explaining an agent's moral responsibility for her free actions.[6] The best way to illustrate the difference between these two approaches is in terms of their implications for a specific interpretation of PAP, known in this paper as the signpost interpretation of the principle of alternate possibilities (SPAP):

> SPAP - A necessary, but not sufficient, condition for agent *A*'s being morally responsible for something *s* is that *A* could have done otherwise.

This article is divided into three sections. In the first, I discuss the virtues of traditional FSCs as purported counterexamples to PAP, but demonstrate why these cases fail.[7] In the second section, I show that Neo-FSCs are

---

[3] Franklin 2009.

[4] By "true moral responsibility" here I mean to capture, roughly, what Galen Strawson discusses in "The Impossibility of Moral Responsibility." See Strawson 1994/2002. For the purposes of this paper, "moral responsibility" is to be understood as "true moral responsibility."

[5] See Fischer 1992; Mele, Robb 1998.

[6] See Hunt 200, 2005; Pereboom 2001, 2005, 2008.

[7] Oddly, this approach is largely indifferent to Frankfurt's original goal of undermining PAP by undermining its appeal. Initially Frankfurt argued that PAP was appealing because of its relationship to a different commonsense moral principle, the coercion principle, which is sometimes said to leave an agent no alternative to doing as their coercer desires. (1969) Frankfurt's initial versions of the case were meant to provide a counterexample to coercion principle; his case, he says, called attention to an important distinction, "that making an action unavoidable *is not the same thing* as bringing it about that the action is performed." (2003/2006, 340) He says "Appreciating this distinction tends to liberate us from the natural but nonetheless erroneous supposition that it is proper to regard people as morally responsible for what they have

concerned with demonstrating the falsity of a robust principle of alternate possibilities (RPAP):

> RPAP - A necessary and sufficient condition for agent *A*'s being morally responsible for something *s* is that *A* had robust alternate possibilities to *s*, where a robust alternate possibility is relevant to explaining *A*'s degree of moral responsibility for *s*.

I argue Neo-FSCs fail to demonstrate the falsity of RPAP, and instead illustrate its truth, that an agent's genuine alternate possibilities play a vital role in determining her degree of moral responsibility in both traditional and Neo-FSC. In the third section, I argue that a hypothetically successful Neo-FSC that demonstrates the falsity of RPAP would fail to demonstrate the falsity of SPAP. Furthermore, such a case is consistent with SPAP and would be insufficient to undermine our commonsense commitment to SPAP. While a successful Neo-FSC would break significant ground in the study of how one's degree of moral responsibility is determined, it would be silent about the metaphysical prerequisites of moral responsibility that are at the heart of the debate between compatibilists and incompatibilists about free will.

I. On Frankfurt Style Cases

Prior to Frankfurt's attack, there was little doubt about our commitment to PAP. Frankfurt says of the principle, "Practically no one... seems inclined to deny or even to question that the principle of alternate possibilities (construed in some way or other) is true." (1969, 829) Frankfurt sought to convince the reader of the principle's falsity by tying it to another supposedly commonsense moral principle, the coercion principle. On Frankfurt's interpretation, both principles offered sufficient conditions to absolve an agent of moral responsibility; in PAP's case if the agent lacked the ability to do otherwise; in the coercion principle's case if the agent was coerced. Frankfurt believed that our commitment to PAP was contingent on the truth of the coercion principle, and by demonstrating the falsity of the coercion

---

done only if they could have done otherwise." The problem is that this distinction is largely irrelevant, as the supposition in question just is the supposition that moral responsibility and determinism are incompatible. In all cases where an agent is determined to act, the thing that makes it unavoidable just is that which causally determined the agent to act; the intuition in question just is the intuition that it would be inappropriate to blame someone in such a case. See Ginet 1996.

principle he would undermine the appeal of PAP.[8] He argued that when an agent acted indifferently to a coercive threat, and freely chose to do the same actions he might otherwise have been coerced into, that agent would be uncontroversially morally responsible for her actions despite the coercive threat.[9] After this, Frankfurt shifts to presenting a case meant to be a direct counterexample to PAP. Here is an updated version of this case:

> Neuroscientist Black wants Jones to kill Smith at time *t*. Black is fairly certain that Jones will do this freely, however he doesn't want to take any chances. Black secretly installs a device in Jones's brain that is designed to causally determine him to choose to kill Smith at time *t*. Unbeknownst to Black the device will remain dormant if Jones were to freely choose to kill Smith at time *t*; however if Jones would have freely chose not to kill Smith, the device would activate causing him to kill Smith.[10] As it so happens, Jones freely chooses to kill Smith and the device remains dormant.

The appeal of Frankfurt's case is that it attempts to sidestep the more thorny metaethical and metaphysical issues that have become the calling card of the free will debate, and is designed to be an open-ended counterexample to any reasonable interpretation of PAP, regardless of what kind of alternate possibilities one interprets PAP as requiring. Frankfurt stipulates that Jones's choice is free, allowing the reader to fill in whatever metaphysical

---

[8] Frankfurt contends our intuitions about PAP are related to our commitment to the truth of the coercion principle but offers no explanation for this being the case. Although it sometimes makes sense to say that coerced agents can't do otherwise; Frankfurt contends that this isn't "strictly speaking" true. (1969, 834) Rather, when faced with some threats, one shouldn't do otherwise, and would be blameworthy if they tried. To act to avoid the bad consequences of a sufficiently horrible threat, then, is not responsibility absolving; rather it is *prima facie* praiseworthy. Not only is the coercion principle not implied by PAP, it doesn't even have the same kind of implications as PAP. The coercion principle is about how we ought to calculate one's degree of moral responsibility; where as a thief might be *prima facie* blameworthy, where we to learn the thief acted under a coercive threat to save her children's life it would become clear that she acted in a praiseworthy manner. In contrast, according to PAP, if we were to learn that a thief was actually a complicated robot wholly causally determined by its programming to steal, PAP offers a quick explanation why that thing is not morally responsible like a person would be; because it had no say, no alternatives.

[9] Frankfurt worries that it doesn't make sense to say that such an agent is coerced; however he expands upon this view later to great effect. See Frankfurt 1973.

[10] This case is partially based one found in Alfred Mele and David Robb's 1998 article "Rescuing Frankfurt-Style Cases."

prerequisites they believe are necessary for moral responsibility. For incompatibilist readers, this means Jones inhabits a world where there are multiple possible futures at any given time. Jones is said to be both uncontroversially morally responsible for his free choice to kill Smith, and to be unable to do otherwise. Frankfurt says "Of course it is in a way up to him whether he acts on his own or as a result of Black's intervention. That depends upon what action he himself is inclined to perform." (1969, 836)

The problem with Frankfurt's original case, and traditional attempts to build upon it, is that either it is unable to effectively cut off alternate possibilities, or it is such that the agent is not uncontroversially morally responsible. This argument against FSCs is known as the Kane-Widerker objection, sometimes called the "dilemma defense."[11] In order for Black's device to interfere only when Jones would choose otherwise, critics contend that it must pick up upon some prior state of affairs that is causally related to Jones's choice, such that either his choice is wholly causally determined, and as such he isn't morally responsible for his action according to the incompatibilist, or there are some situations where in the device will fail to prevent Jones from choosing otherwise. Alfred Mele and David Robb propose that the device could be triggered by some step within Jones's natural decision making process, and thus only activate if he would choose otherwise. While there is some concern that such a device wouldn't be consistent with the kind of libertarian agency incompatibilists believe is necessary for moral responsibility; the question is largely beside the point as Frankfurt, Mele, and Robb seem to agree that Jones can either act freely, or be forced to act, and that these are distinct possibilities.

Compatibilist John Martin Fischer has argued that FSCs can still be persuasive even if one needs to assume determinism in order to ensure that Jones cannot do otherwise and that it doesn't even matter if Jones is uncontroversially morally responsible.[12] The problem with this approach is that it ignores and abandons the open-endedness and persuasiveness of Frankfurt's original case. Fischer contends "... it is not alleged by the Frankfurt-style compatibilist that the strategy is *knockdown* or *decisive*." (2007, 470) However, this seems to miss the point of FSCs; if Jones's action is wholly causally determined by events that occurred long before he was born, Black's machinations are irrelevant to the explanation of why Jones

---

[11] See Kane 1985, 1996; Widerker 1995; Ginet 1996; Wyma 1997; Goetz 2005; Simkulet 2012, 2014a.

[12] See Fischer 2000, 2007, 2010.

can't do otherwise. Fischer-style interpretations of FSCs might convince *compatibilists* of the falsity of PAP, but compatibilists already widely reject PAP.

Frankfurt's own account is inconsistent with Fischer's; Frankfurt volunteers that Jones has alternate possibilities, and that he can act either virtuously or viciously (1969, 826; 2003/2006 , 343); he stipulates that there are two possible futures open to Jones; Jones can either freely choose to kill Smith, or be forced by Black to choose to kill Smith. The former is vicious and blameworthy; the latter is virtuous, perhaps even praiseworthy![13] On Frankfurt's view it seems Jones would be praiseworthy for trying to make his choice in such a manner that would trigger Black's device; but because the device ultimately determines Jones's choice, it would be inappropriate to hold him morally accountable for the outcome of that deliberation.

Frankfurt's case is said to be a counterexample to SPAP because Black's device prevents Jones from choosing anything but to kill Smith, and intuitively he's morally responsible for that choice because he freely chose to do so. The problem is that Jones is only morally responsible for his choice if he freely chooses it; had he been forced to make the choice by Black's device, although his deliberation would have the same outcome, he wouldn't be morally responsible for its outcome. Jones has alternatives; he can freely choose to kill Smith, or freely choose to act in a way that, *sans* Black's device, would result in him choosing something else, but that thanks to Black's device instead results in him being (unfreely) caused to choose to kill Smith.

SPAP is agnostic on the role that one's alternate possibilities play in determining how morally responsible one is; it merely states that alternate possibilities are a necessary condition for moral responsibility; because Jones can act in either of two ways, Frankfurt's case fails to cut off alternate possibilities of the kind relevant to this principle, and thus fails to constitute a counterexample to the principle.

II. On Neo-Frankfurt-Style Cases

Recent attempts to construct FSCs have moved away from trying to construct scenarios in which an agent completely lacks alternate possibility. Instead

[13] Michael Otsuka similarly argues that the morally relevant kind of alternate possibilities in question are the alternate manner in which Jones can act; either freely (viciously) or be forced to act by Black (in such a way that makes it absurd to hold him accountable for his action). See Otsuka 1998.

these Neo-FSCs are meant to restrict an agent's alternate possibilities to possibilities that don't seem to be morally relevant. David Hunt and Derk Pereboom have each constructed cases of this kind, designed to show that the alternate possibilities open to the agents in question are insufficient to explain their moral responsibility.[14] Here are concise versions of these cases:

> *Hunt's Revenge case*
> Smith humiliates Jones, which causally determines Jones to have a strong desire to kill Smith. If Jones does not rid himself of this desire soon, it will causally determine him to choose to try to kill Smith. Jones knows acting on this desire would be wrong, and knows that to rid himself of this desire he must first *consider* his alternatives, then he must choose not to kill Smith. Black is monitoring Smith's thoughts, and if Jones considers not killing Smith, Black will intervene and force Jones to choose to kill Smith. As it so happens, Jones never considers his alternatives, and his desire to kill Smith causally determines that he chooses to kill Smith, and he does. (Hunt, 2005)
>
> *Pereboom's Tax Evasion case:*
> Joe believes that he can get away with cheating on his taxes, but that doing so would be wrong. His strong desire to advance his self interest will causally determine him to cheat on his taxes unless he chooses otherwise. However, he cannot choose otherwise *on a whim*; he knows a prerequisite for him to choose otherwise is for him to freely raise his moral attentiveness level through the use of his libertarian free will. If he raises it enough, he will be able to use his libertarian free will to reconsider; however he might then freely choose act either in his self interest, or act as he believes is moral. Unbeknownst to Joe, Black has implanted a device in his brain that is triggered by him reaching the appropriate level of moral attentiveness. When triggered, the device robs him of his libertarian free will and causally determines him to cheat on his taxes. As it so happens, Joe never uses his libertarian free will to raise his moral attentiveness level "and he chooses to evade taxes while the device remains idle." (Pereboom, 2008)

[14] See Hunt 2005, Pereboom 2001, 2005, 2008.

Jones and Joe are supposed to be uncontroversially morally responsible for their actions, and both have the ability to do otherwise - Jones can freely consider not killing Smith (which would prompt Black to intervene and rob him of his free will), while Joe can freely raise his moral attentiveness level (which would trigger Black's device and rob him of his free will). The primary difference between these cases and Frankfurt's original is that in the original, Frankfurt admits Jones can act virtuously - he can act in such a way that would counterfactually lead to him choosing not to kill Smith - and in doing so be *prima facie* praiseworthy; however in Hunt and Pereboom's cases, the best Jones can do is to freely *consider* not killing Smith, while the best Joe can do is raise his attentiveness level to *consider* not cheating on his taxes. Because after this consideration, Jones and Joe could go on to freely act immorally, they do not consider the alternate possibility to consider doing otherwise as a robust alternate possibility, where an alternate possibility is robust if and only if an agent knew she would be differently morally responsible if she chose that action.

Hunt and Pereboom contend that Jones and Joe, respectively, are morally culpable for their actions, despite lacking robust alternate possibilities and if this is the case RPAP is false. The problem for this account is that both Jones and Joe *know* that considering alternatives and raising one's moral attentiveness level, respectively, are necessary, but not sufficient, conditions for freely choosing to do otherwise.[15] While it is true that the needlessly complicated (and evolutionarily deficient) buffer-step that Hunt and Pereboom incorporate into their agents' decision making process might lead to either a morally praiseworthy or blameworthy decision by that agent, both agents know that this step is a necessary prerequisite to doing what they believe is right.[16]

If one has a moral obligation to do something *y*, and *x* is a necessary, but not sufficient, step for obtaining *y*, one has a moral obligation to do *x*. As such Jones and Joe are each morally obligated to take this step; Jones has a moral obligation to freely consider his alternatives, and Joe has a moral obligation to raise his moral attentiveness level. Although Black's

[15] In Hunt and Pereboom's original cases they fail to specify whether Jones or Joe have sufficient working knowledge of their bizarre mental faculties to *know* the steps they need to take to act otherwise. This article stipulates that they have this knowledge, because if they were ignorant of such things it would be absurd to expect them to do otherwise as they would have no reason to do so, and thus absurd to hold them morally accountable for failing to take the steps to change their minds.

[16] See Simkulet 2014a.

monitoring makes it impossible that Jones can ever freely choose not to kill Smith, and Black's device makes it impossible that Joe can ever freely choose not to cheat on his taxes; both Jones and Joe know that they have a strong moral obligation to do otherwise, and fail to do so. Implicit in the idea of a moral obligation is that if one fails, one is differently morally responsible than if one succeeds. As such, the buffer-step that Hunt and Pereboom construct constitutes a robust alternate possibility - if Jones and Joe take this step intending it to be the first step in avoiding a blameworthy action (killing Smith, cheating) in favor of a praiseworthy action (not killing Smith, not cheating), they are inherently praiseworthy for doing so.

I've argued that Jones and Joe have a strong moral obligation to engage in the steps they believe are necessary prerequisites for freely choosing to do what each thinks is right - not killing Smith, and not cheating on taxes, respectively - and that if they do these prerequisites for these reasons, they are morally praiseworthy for doing so. However, it is possible that Jones and Joe engage in these steps for other reasons. Suppose that Jones knows that unless he considers his alternatives, he will be causally determined to kill Smith, but that Jones wants Smith to experience worse, say by letting Smith live and systematically killing everyone and everything Smith loves. By stipulation, Jones knows the only way he can choose to do such a thing is to freely consider his alternatives. If, intent upon getting his revenge, Jones considers his alternatives (with the hope of choosing this long drawn out torture over a swift death), Black will intervene and rob Jones of his free will and causally determine that he kills Smith then and there. If this were the case, it doesn't make sense to say that Jones is morally blameworthy for Smith's death... but he is blameworthy for something. He is blameworthy for freely acting in such a way that he believes is a necessary prerequisite for his murdering Smith's friends and loved ones to bring about that very consequence.

Although critics of Neo-FSCs could focus on demonstrating the *prima facie* praiseworthy possibilities of Jones or Joe acting otherwise, it is clear that whether they would be praiseworthy or blameworthy for their alternate possible actions depends upon the intentions they take those actions with.[17]

[17] It strikes me as possible that a well-intentioned Jones might still freely choose to either kill Smith, or get revenge on Smith by killing Smith's friends and family; however such a Jones would still be *prima facie* praiseworthy for acting in a manner such that he believed was necessary for him to do what is right. He is, however, blameworthy for his latter, wrong, free choice.

Rather than demonstrate the falsity of RPAP, it seems as though Neo-FSCs actually demonstrate its truth; an agent's beliefs about the moral character of their possible intentional actions play a vital role in determining her degree of moral responsibility. If Jones were to consider his alternatives hoping to do worse than kill Smith, he would be blameworthy for doing so (and blameless for being forced to kill Smith by Black), while if Joe were to raise his moral attentiveness level in hopes of doing the right thing, he would be praise-worthy for doing so (and as Pereboom contends Joe's choice to cheat on his taxes is the result of his being determined to act in his self interest, Joe would be blameless for being caused by his beliefs about his self interest to cheat on his taxes as a result).

III. Why Neo-Frankfurt-Style Cases are Irrelevant

In the previous two sections I've argued that traditional FSCs fail to demonstrate the falsity of SPAP, and Neo-FSCs fail to demonstrate the falsity of RPAP. The goal of this section is to compare the implications of a hypothetically successful traditional FSC with those of a hypothetically successful Neo-FSC.

For our purposes a successful traditional FSC is stipulated to have the following characteristics: The agent, Jones*, is actually and uncontroversially truly morally blameworthy for his free choice, *s*, and he is blameworthy to non-zero degree *d* for *s*. Jones*'s moral responsibility for *s* is not derivative of some prior act *r* that preceded it. Jones* could not do otherwise; by this I mean that from the moment that immediately preceded Jones*'s deliberative process in which he chose *s*, there was only one possible way in which he could choose *s* and no possible way that he could choose other than *s*. There was no way such that he could choose *s* and be morally blameworthy in any other degree than *d*.

A successful traditional FSC would be a decisive counterexample to both SPAP and RPAP, directly demonstrating the falsity of the principles. Jones* would be morally responsible despite lacking the ability to do otherwise, but even if he had the ability to do otherwise it would be irrelevant to explaining how morally responsible he is. This case does not show that moral responsibility is impossible within an indeterministic universe, but it would demonstrate the truth of compatibilism.[18]

[18] Note that a successful Fischer-style interpretation of a FSC would prove the truth of compatibilism. It strikes me that such a successful case would also undermine the majority of

For our purposes, a successful Neo-FSC is stipulated to have the following characteristics: The agent, Joe*, is actually and uncontroversially truly morally responsible for something, his free choice *c*, and he is blameworthy to degree *b* for *c*. Joe* had alternate possibilities, but these alternate possibilities are not relevant to explaining his moral responsibility for *c* to degree *b*. By this I mean that Joe* has a set of alternate possibilities *a*, composed of one or more alternate possibilities, and that had he had acted on any of the alternate possibilities within set *a*, he would still have been blameworthy to degree *b*.

A successful Neo-FSC demonstrates the falsity of RPAP; Joe* would be morally responsible despite lacking robust alternate possibilities because the alternate possibilities within set *a* would be irrelevant to explaining Joe*'s blame. However, at least as formulated above, a successful Neo-FSC is not a counterexample to SPAP, Joe* is stipulated to have alternate possibilities. If this were the case, then it is still possible that there is some connection between moral responsibility and indeterministic metaphysics, such that compatibilism is false.

Assuming the existence of a successful Neo-FSC, critics of SPAP have two options to attack SPAP: First, they might revise the successful Neo-FSC such that Joe* would be blameworthy to degree *b* for choice *c* regardless the size of the set of alternate possibilities he has, such that if Joe* had no alternate possibilities (an empty set *e*), Joe* would be blameworthy to degree *b* for choice *c* in the same exact way as if he had alternate possibilities. If successful, all this approach does is turn the Neo-FSC into a traditional FSC, which by assumption would be a counterexample to SPAP.

The second way in which a critic might argue against SPAP, given a successful Neo-FSC, is to argue that our belief in SPAP is contingent on our belief in RPAP. This style of argument is similar to Frankfurt's original assertion that our commitment to the truth of PAP is based, at least in part, on our commitment to the coercion principle. I don't see how this argument could gain any traction among proponents of SPAP. The appeal of SPAP is that alternate possibilities serve as signposts that indicate an agent might have the kind of control over their actions necessary for moral responsibility. In contrast, the truth of RPAP might actually be a problem for incompatibilists; the thought that such alternate possibilities might play a role in determining an agent's moral responsibility is actually quite troubling, raising the specter

our beliefs about moral responsibility and undermine the idea that commonsense moral intuitions could be a reliable guide to either metaphysical or moral truth.

of moral luck.[19] Analytically, the truth of RPAP necessitates the truth of SPAP; but proponents of SPAP leery of the problem of moral luck are far more likely to reject RPAP than attribute their belief in SPAP to it.[20]

Not only would a successful Neo-FSC fail prove the falsity of SPAP, but there is no reason to think that a Neo-FSC's proving RPAP false would give us any reason to rethink our commonsense commitment to SPAP. To my knowledge, proponents of Neo-FSCs don't argue RPAP is connected to our beliefs about any other principles relevant to the free will debate; and thus I have to conclude that a hypothetically successful Neo-FSCs would be silent on the issues relevant to the debate over free will; in contrast a successful FSC would have substantial implications for the debate, demonstrating the truth of compatibilism. This is not to say that a successful Neo-FSC would be irrelevant; quite the contrary. A successful Neo-FSC would seem to show that some actions have fixed moral value; for example maybe killing Smith always wrong to the exact same degree and always wrong in the same manner. On this view, killing Smith for embarrassing you at a party might be as wrong and equally as bad as killing Smith because it is the only way to stop him from murdering thousands of helpless newborn infants.

If that sounds wrong, it's probably because it is. I've argued there are no successful traditional FSCs and that there are no successful Neo-FSCs. Rather than demonstrate the falsity of RPAP, I've argued (unsuccessful) Neo-FSCs demonstrate its truth; killing Smith out of revenge is *prima facie* blameworthy; however killing Smith because it is the only way to prevent him from murdering infants is *prima facie* praiseworthy. The reason why killing Smith in the former case is morally abhorrent is because there are better alternatives; the reason it is morally necessary in the second is because there are no better alternatives. However all of this is quite beside the point; what this article set out to show was that while a successful traditional FSC would have substantive implications for the free will debate, a successful Neo-FSC would do nothing of the sort; I believe it has succeeded in this goal.

---

[19] For more on moral luck see Nagel 1976, Zimmerman 2002, 2006, Simkulet 2014b.

[20] Moral luck would occur if and only if something outside of an agent's control would determine their moral responsibility; however as demonstrated in Hunt and Pereboom's cases, the existence of alternate possibilities don't determine the moral responsibility of either Jones or Joe; however the agent's beliefs about their alternate possibilities plays a vital role in explaining why they are morally culpable to the degree in which they are, much as Frankfurt argues that an agent's beliefs and intentions surrounding a coercive threat determines how morally culpable they are for acting in accordance with that threat, because of the threat, or indifferent to the threat.

References

Fischer, John Martin. (1982). Responsibility and Control. *Journal of Philosophy* 79, pp. 24–40.

Fischer, John Martin. (2000). As Go the Frankfurt Examples, so Goes Deontic Morality. *The Journal of Ethics* 4, pp. 361–363.

Fischer, John Martin. (2007). The Importance of Frankfurt-Style Argument. *The Philosophical Quarterly* 57, pp. 464–471.

Fischer, John Martin. (2010). The Frankfurt Cases: The Moral of the Stories. *Philosophical Review* 119, pp. 315–336.

Frankfurt, Harry G. (1969). Alternate Possibilities and Moral Responsibility", *The Journal of Philosophy,* 66, pp. 829–839.

Frankfurt, Harry G.. (1973). Coercion and Moral Responsibility. *Essays on Freedom of Action*, ed. Ted Honderich, London: Routledge & Kegan Paul, pp. 41–42.

Frankfurt, Harry G. (2003/2006). Some Thoughts Concerning PAP. *Moral Responsibility and Alternative Possibilities Essays on the Importance of Alternative Possibilities*, ed. Widerker, David and McKenna, Michael, Aldershot, England, Ashgate Publishing Limited. pp. 339–345.

Franklin, Christopher. (2009). Neo-Frankfurtians and Buffer Cases: the New Challenge to the Principle of Alternative Possibilities. *Philosophical Studies* 152, pp. 189–207.

Ginet, Carl. (1996). In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing. *Philosophical Perspectives* 10, pp. 403–417.

Goetz, Stewart. (2005). Frankfurt-Style Counterexamples and Begging the Question. *Midwest Studies in Philosophy* 29, pp. 83–105.

Hunt, David. (2000). Moral Responsibility and Unavoidable Action. *Philosophical Studies* 97, pp. 195–227.

Hunt, David. (2005). Moral responsibility and buffered alternatives. *Midwest Studies in Philosophy* 29, pp. 126–145.

Kane, Robert. (1985). *Free Will and Values*, Albany: State University of New York Press.

Kane, Robert. (1996). *The Significance of Free Will*, New York: Oxford University Press.

Mele, Alfred R.; Robb, David. (1998). Rescuing Frankfurt-Style Cases. *The Philosophical Review* 107, pp. 97–112.

Nagel, Thomas. (1976). Moral Luck. *Proceedings of the Aristotelian Society Supplementary Volumes* 50, pp. 137–151.

Otsuka, Michael. (1998). Incompatibilism and the Avoidability of Blame. *Ethics* 108, pp. 685–701.

Pereboom, Derk. (2001). *Living Without Free Will.* Cambridge: Cambridge University Press.

Pereboom, Derk. (2005). Defending Hard Incompatibilism. *Midwest Studies* 29, pp. 228–247.

Pereboom, Derk. (2008). Defending Hard Determinism Again. *Essays on Free Will and Moral Responsibility*, ed. Nick Trakakis and Daniel Cohen, Newcastle-upon-Tyne: Cambridge Scholars Press, pp. 1–33.

Simkulet, William. (2012). On Moral Enhancement. *American Journal of Bioethics Neuroscience* 3, pp. 17–18.

Simkulet, William. (2014a). On Robust Alternate Possibilities and the Tax Evasion Case. *Southwest Philosophy Review* 31(1), pp. 101–107.

Simkulet, William. (2014b). Lucky Assassins: On Luck and Moral Responsibility. *Lyceum* 13(1), pp. 58–93.

Strawson, Galen. (1994/2002). The Impossibility of Moral Responsibility. *Philosophical Studies* 75, pp. 5–24. reprinted in *Ethical Theory Classic and Contemporary Readings* by Louis P. Pojman.

Wyma, Keithm. (1997). Moral Responsibility and Leeway for Action. *American Philosophical Quarterly* 34, pp. 57–70.

Widerker, David. (1995). Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities. *Philosophical Review* 104, pp. 247–361.

Zimmerman, Michael J. (2002). Taking Luck Seriously. *The Journal of Philosophy* 99, pp. 553–576.

Zimmerman, Michael J. (2006). Moral Luck: A Partial Roadmap. *Canadian Journal of Philosophy* 36, pp. 585–608.

William Simkulet
Cleveland State University
simkuletwm@yahoo.com