1

# Preference Semantics for Hansson-type Dyadic Deontic Logic: A Survey of Results

Xavier Parent

Abstract. This chapter discusses the Hansson-type preference semantics for dyadic deontic logics. In that framework the conditional obligation operator is interpreted in terms of best antecedent-worlds. I survey results pertaining to the meta-theory of such logics, focusing on axiomatization issues. The goal is to provide a "roadmap" of the different systems that can be obtained, depending on the special properties envisaged for the betterness relation, and depending on how the notion of "best" is understood (optimality *vs.* maximality, stringent *vs.* liberal maximization). In addition, the systems' decidability and automated theorem-proving for them are discussed, and variant truth-conditions for the conditional obligation operator are reviewed.



This chapter is dedicated to the memory of Lennart Åqvist, who died on 7 March 2019 at the age of 86. This work was supported by WWTF MA16-028. I am indebted to Lou Goble for his careful reading of the chapter and for valuable comments. I wish to thank Christoph Benzmueller and Alexander Steen for feedback on the section devoted to automated theorem-proving, and Jose Carmo and Dov Gabbay for feedback on the section dealing with decidability. I also would like to extend my gratitude to Cleo Condoravdi for helpful discussions on Kratzer, and to Walter Bossert for useful discussions on rational choice theory and on Theorem 4.15. Last, I would like to thank those who have commented on this essay, or aspects of it, at various stages of its development. In particular I would like to mention Richard Booth, Joerg Hansen, Sven Ove Hansson, Jeff Horty, Paul McNamara, David Makinson and Leon van der Torre.

Handbook of Deontic Logic and Normative Systems, Volume 2

# 1 Introduction

Beginning with work by Danielsson [1968] and Hansson [1969], so-called Dyadic Deontic Logic (hereinafter referred to as "**DDL**") aims at providing a formal analysis of conditional obligation sentences within a preference-based semantics. The language of **DDL** employs a dyadic (or conditional) obligation operator $\bigcirc(-/-)$, where $\bigcirc(B/A)$ is read as "It is obligatory that $B$, given that $A$". This construct is interpreted using a preference relation, which orders all the possible worlds in terms of comparative goodness or betterness. In that framework $\bigcirc(B/A)$ is taken to hold, whenever all the best $A$-worlds are $B$-worlds.

**DDL** is a natural generalization of Monadic Deontic Logic (hereinafter referred to as "**MDL**"). The semantics of this one uses a binary classification of possible worlds into good/bad. For **DDL**, this binary classification is relaxed to allow for grades of ideality between these two extremes.[1] This leads to the use of a conditional obligation operator that is primitive rather than being defined in terms of the standard (monadic) obligation operator and some other familiar constructs like material implication or strict implication.

**DDL** uses the possible world semantics in novel ways with a view to solving issues related to two different kinds of deontic conditionals:

**Contrary-to-duty conditionals** Since the publication of Chisholm [1963], deontic logicians have struggled with what has become known as the "contrary-to-duty" (CTD) problem. It is the problem of giving a formal treatment to those obligations—called "contrary-to-duty" by Chisholm—which come into force when some other obligation is violated. **DDL** was initially developed in order to handle this first type of deontic conditional. According to Hansson and others, like van Fraassen [1972] and Lewis [1973; 1974], the problems raised by CTDs call for the use of an ordering on possible worlds, in terms of preference or relative goodness, and **MDL** fails in as much as its semantics does not allow for grades of ideality.

**Defeasible deontic conditionals** Independently of the above, the use of a preference relation has also been advocated in relation to the analysis of the notion of defeasible conditional obligation. In particular, Alchourrón [1993] argues that preferential models provide a better treatment of this notion than the usual Kripke-style models do. Indeed, a defeasible conditional obligation is one that leaves room for exceptions. Under a preference-based approach, we no longer have the deontic analogue of two laws, the failure of which constitutes the main formal feature expected of defeasible conditionals. One is "deontic" modus-ponens, also known as Factual Detachment (FD): $\bigcirc(B/A)$ and $A$ imply $\bigcirc B$. The other is

[1] A remark on my choice of name is in order. **MDL** is more commonly known as "Standard Deontic Logic" (SDL), and **DDL** as "Dyadic Standard Deontic Logic" (DSDL). Both names appear in Hansson's seminal paper. Throughout this chapter I will not use the label SDL, because it tends to carry the connotation that the framework in question is still a recognized "standard". As Hilpinen and McNamara [2013, p. 38] point out, to call SDL a standard is a misnomer. **MDL** refers to a family of systems, which were called **D**, **DS4**, **DM** and **DS5** by Hanson [1965]. (Other labels have been used in the literature.)

the law of Strengthening of the Antecedent (SA): $\bigcirc(B/A)$ entails $\bigcirc(B/A \wedge C)$.

There is an extensive literature on the treatment of these notions within a preference-based framework. Regarding contrary-to-duties, the reader may wish to consult [van Fraassen, 1972; Lewis, 1973; Tomberlin, 1981; Loewer and Belzer, 1983; Kratzer, 1991; Prakken and Sergot, 1997; van der Torre and Tan, 1999; Hilpinen and McNamara, 2013]. Concerning defeasible conditional obligations, the reader is referred to [Makinson, 1993; Boutilier, 1994; Alchourrón, 1995; Asher and Bonevac, 1997; van der Torre and Tan, 1997; Horty, 2014]. It is not the purpose of this chapter to evaluate such treatments, nor is it to discuss the relationship between dyadic deontic logic and frameworks developed in other closely related areas, like revealed preference theory (as introduced by the economist Samuelson), the logic of conditionals (as developed in the 1970's following Stalnaker and Lewis), or the theories of nonmonotonic inference operations (as constructed in the 1980's in the context of logics for artificial intelligence). All these frameworks share the idea of using a semantics based on a notion of minimality under a preference relation, or equivalently, a notion of maximality under its converse. For a good discussion of the interplay between these areas, the reader is referred to [Makinson, 1993].[2]

The aim of this chapter is to present a survey of recent results pertaining to the meta-theory of **DDL**. Since the publication of Hansson's seminal paper, substantial contributions have been made to enhance our understanding of the meta-theory of **DDL**, starting with work by Spohn [1975], and continuing with work by Åqvist [1987; 1993; 2002], Hansen [1999], Goble [2015; 2019] and myself [Parent, 2008; Parent, 2010; Parent, 2014; Parent, 2015]. However, there is still no systematic survey of the field. The present chapter aims at filling in this gap. The goal is to provide a "roadmap" of the different systems that can be obtained, based on two types of considerations or variations.

The first type of consideration is familiar from modal logic. Different systems can be obtained by varying the conditions on the preference relation. In general the imposition of a condition has the effect of validating a modal formula. In monadic modal logic, we have a clear picture of the different systems that can be obtained depending on the properties of the accessibility relation. In dyadic deontic logic, this picture is still missing. Results in the literature have so far mostly concerned classes of

[2] Makinson does not discuss the connection with rational choice theory. This one is examined by [Rott, 2001] among others.

structures with strong conditions on the betterness relation. One such condition is the property of transitivity, which has been called into question by moral philosophers and economists.[3] One would like to know what happens when such a condition is relaxed. What Lewis [1973] calls the limit assumption is another requirement that one would like to be able to drop. Roughly speaking, it says that a set of possible worlds should always have a best element. A number of deontic logicians objected to the limit assumption, Lewis [1973, p. 97-98] among them. It is not widely known what happens when these properties are not assumed.

This brings into the forefront so-called correspondence theory, devoted to the systematic study of relations between classes of frames and modal languages. Van Benthem [2001, §3.2] asks if, or to what extent, such a theory can be developed for conditional logic. Such a study falls outside the scope of the present chapter. But I hope the considerations it offers can be used as a stepping stone towards the development of such a theory.

The second type of consideration this chapter introduces concerns the notion of "best", in terms of which the truth conditions for $\bigcirc(-/-)$ are typically phrased. One can distinguish between two ways to understand the notion of a world being best: it can be either optimal or maximal. This distinction is well-known from rational choice theory where most authors follow Herzberger [1973] in using the terms "stringent" *vs.* "liberal" maximization for what (following Sen [1997]) I call optimality *vs.* maximality. For some item $x$ to qualify as an optimal element of $X$, it must be at least as good as every member of $X$. For $x$ to count as a maximal element, no other element in $X$ must be strictly better than it. Thus, while the optimal elements are all equally good, the maximal elements are either equally good or incomparable. Depending on what notion of "best" is used, one gets different truth conditions for $\bigcirc(-/-)$, but also different forms of the limit assumption.

I remark in passing that there is some variation in terminology. For instance, [Bossert and Suzumura, 2010] prefer the labels "maximal *vs.* greatest" element rationalizability. On the other hand, the choice to use "optimal" and "maximal" the way just described is not mine, but Sen's (see in particular [Sen, 1997, §5]). I have heard people swap the two terms, and take optimal as meaning "not-dominated", and maximal as meaning "dominates-all-others". (See, *e.g.*, the definition of optimal in [Horty, 2001, p. 72].) In the end, it does not matter which way we speak, so long as we understand and agree on what we mean and do not

[3] Cf. [Sen, 1971] and [Temkin, 1987].

allow the coexistence of two conflicting ways of speaking to engender confusion. In this chapter I will stick to Sen's terminology.

This investigation takes place in the conditional logic setting put forth by Åqvist in a series of publications [Åqvist, 1987; Åqvist, 1993; Åqvist, 2002] rather than in Hansson's original setting. That one is studied axiomatically by Spohn [1975] and Goble [2019]. Readers should be warned that there is far less standardization in preference semantics than in the usual Kripke-style semantics for deontic logic, and more room for variation. This is due to the fact that there are several factors that must be juggled all at once. Thus, even when sticking with Åqvist's approach, more semantical variations than the above two can be made. For instance, under the Åqvist account the ranking is not world-relative. However, as Makinson [1993] points out, one may want to allow for the ranking to vary across possible worlds. This extra choice (and some others) are studied axiomatically by Goble, who in his [2015] pursues a similar project. It falls outside the scope of the present paper to integrate his results. The present chapter is not, and does not pretend to be, a comprehensive survey of Hanssonian approaches to dyadic deontic logic, so much as a summary of certain results, mainly my own, that would help the reader understand some important aspects of the Hanssonian approach, but does not address the scope of that approach from either a formal or philosophical point of view.

As part of motivating the formal moves to be developed next, I briefly recall how the framework handles the standard CTD scenarios, like Chisholm's paradox.

**Example 1.1.** *[Chisholm's scenario] Consider the following set of sentences, where $h$ can be read as the fact that a certain man goes to the assistance of his neighbors and $t$ as the fact that he is telling them that he is coming:*

$$\Gamma = \{\bigcirc h, \bigcirc(t/h), \bigcirc(\neg t/\neg h), \neg h\}$$

*$\bigcirc h$ expresses what is usually called a primary obligation. $\bigcirc(\neg t/\neg h)$ is its associated CTD obligation, and $\bigcirc(t/h)$ is its associated ATD (according-to-duty) obligation. Figure 1 describes a typical preference model of $\Gamma$. Here the convention is that at each world $a \in W$, I list the propositional letters that $a$ satisfies, omitting those that it makes false. The best overall world is the one where both $h$ and $t$ hold, and the worst overall world is the one where $t$ holds but $h$ does not. In between one sees two worlds, one with $h$ but not $t$ and the other with neither $h$ nor $t$. All the formulas in $\Gamma$ are satisfied in $a_3$ and $a_4$. This shows that the set $\Gamma$ is consistent. The primary obligation holds, because the best overall*

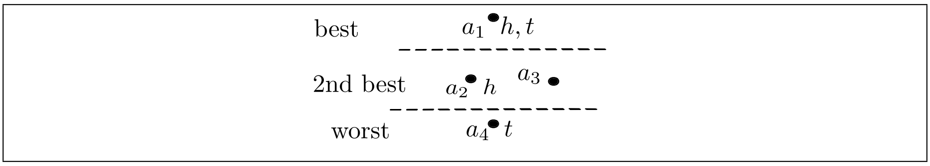



Figure 1: A typical model of Chisholm's scenario

*word satisfies $h$. The CTD obligation holds, because the best $\neg h$-world satisfies $\neg t$. The ATD obligation holds, because the best $h$-world satisfies $t$. It is worth mentioning that this approach to the CTD scenarios only works because neither (FD) nor (SA) are valid under this approach, as the model of Figure 1 demonstrates.*[4]

The layout of this chapter is as follows. In Section 2, the syntax and the semantics are described. In Section 3, the relevant proof systems are introduced. In Section 4, the determination results available at the time of writing this chapter are reviewed. In Section 5, the decidability of the theoremhood problem is established, and automated theorem-proving is discussed. In Section 6 variant truth-conditions are reviewed. Section 7 concludes. Supplementary material is gathered in three appendices. In particular, the proof of two new results is given.

# 2 Syntax and semantics

## 2.1 Syntax

**Definition 2.1.** *The language $\mathcal{L}$, or set of well-formed formulas (wffs), is generated from a set $\mathbb{P}$ of propositional atoms by the following BNF:*

$$A ::= p \in \mathbb{P} \mid \neg A \mid A \vee A \mid \Box A \mid \bigcirc(A/A)$$

$\neg A$ is read as "not-$A$", and $A \vee B$ as "$A$ or $B$". $\Box A$ is read as "$A$ is settled as true", and $\bigcirc(B/A)$ as "$B$ is obligatory, given $A$". $A$ is called the antecedent, and $B$ the consequent.

The following derived connectives are introduced. $P(B/A)$ ("$B$ is permitted, given $A$") is short for $\neg\bigcirc(\neg B/A)$, $\bigcirc A$ ("$A$ is unconditionally obligatory") and $PA$ ("$A$ is unconditionally permitted") are short for $\bigcirc(A/\top)$ and $P(A/\top)$, respectively. $\Diamond A$ is short for $\neg\Box\neg A$. Other Boolean connectives are defined as usual.

[4] (FD) yields $\bigcirc\neg t$. This "contradicts" the fact that the best overall world satisfies $t$, so that $\bigcirc t$ holds. (SA) warrants the move from $\bigcirc t$ to $\bigcirc(t/\neg h)$. This "contradicts" the third formula in $\Gamma$.

Åqvist's language goes beyond Hansson's by including alethic modalities, mixed formulas (in which deontic formulas are combined with Boolean ones) and iterated deontic modalities.

## 2.2 Semantics–basic setting

**Definition 2.2** (Preference model)**.** *A preference model is a structure*

$$M = (W, \succeq, v)$$

*in which*

*(i)* $W \neq \emptyset$ *(*$W$ *is a non-empty set of "possible worlds");*

*(ii)* $\succeq \subseteq W \times W$ *(intuitively,* $\succeq$ *is a betterness or comparative goodness relation; "*$a \succeq b$*" can be read as "world* $a$ *is at least as good as world* $b$*");*

*(iii)* $v : \mathbb{P} \to \mathcal{P}(W)$ *(*$v$ *is an assignment, which associates a set of possible worlds to each propositional atom* $p$*).*

$\succ$ denotes the strict relation induced by $\succeq$, defined as its "strengthened converse complement" and obtained by putting $a \succ b$ whenever $a \succeq b$ and $b \not\succeq a$. $a \succ b$ may be read as "$a$ is strictly better than $b$". Note that $\succ$ is by definition irreflexive (*i.e.*, for all $a$, $a \not\succ a$). Two worlds $a$ and $b$ are said to be equally good or indifferent, $a \equiv b$, whenever $a \succeq b$ and $b \succeq a$. They are said to be incomparable, $a||b$, whenever $a \not\succeq b$ and $b \not\succeq a$.[5]

**Definition 2.3** (Satisfaction relation)**.** *Given a model* $M = (W, \succeq, v)$ *and a world* $a \in W$*, the satisfaction relation* $M, a \models A$ *(read as "world* $a$ *satisfies* $A$ *in model* $M$*") is defined by induction on the structure of* $A$*. The clauses are as usual for the Boolean connectives and* $\Box$*:*

$$M, a \models p \textit{ iff (if and only if) } a \in v(p)$$
$$M, a \models \neg A \textit{ iff } M, a \not\models A$$
$$M, a \models A \vee B \textit{ iff } M, a \models A \textit{ or } M, a \models B$$
$$M, a \models \Box A \textit{ iff } \forall b \;\; M, b \models A$$

*The clause for the dyadic obligation operator is:*

$$M, a \models \bigcirc(B/A) \quad \textit{iff} \quad \mathrm{best}_{\succeq}(\|A\|^M) \subseteq \|B\|^M$$

[5] The betterness relation $\succeq$ may be defined in terms of some more basic ingredients in the semantics. (See, for instance, [Kratzer, 2012] and [Prakken and Sergot, 1997]). However, most articles in the field do not consider this course, and neither will I in this chapter. Kratzer's theory is discussed in more detail in the chapter in this volume "Deontic logic and natural language" by F. Cariani.

As usual $\|A\|^M$ denotes the truth-set of $A$ (*i.e.*, the set of worlds at which $A$ holds). The notation $\text{best}_{\succeq}(\|A\|^M)$ is a shorthand for the set of best (according to $\succeq$) worlds in which $A$ is true. Intuitively, $\bigcirc(B/A)$ holds at $a$ whenever all the best $A$-worlds are $B$-worlds. Note that, by definition of $P(–/–)$, $M, a \vDash P(B/A)$ iff $\text{best}_{\succeq}(\|A\|^M) \cap \|B\|^M \neq \emptyset$. Intuitively: $P(B/A)$ holds whenever at least one best $A$-world is a $B$-world. I will postpone the definition of $\text{best}_{\succeq}(\|A\|^M)$ until the next section. When the context allows, I will drop the symbol $M$ and just write $\|A\|$ and $a \models A$.

The notions of semantic consequence, validity and satisfiability are defined as usual.

## 2.3 Two notions of "best"

As mentioned in Section 1, there are two ways to formalize the notion of best antecedent-worlds: one may do it using the notion of optimality, or the notion of maximality.[6] They are not clearly distinguished in the deontic logic literature even though their differences can be significant. They may be defined thus:

$$\text{opt}_{\succeq}(\|A\|^M) = \{b \in \|A\|^M \mid \forall c \ \ (c \vDash A \to b \succeq c)\}$$
$$\max_{\succeq}(\|A\|^M) = \{b \in \|A\|^M \mid \forall c \ \ ((c \vDash A \ \& \ c \succeq b) \to b \succeq c)\}$$

Maximality can equivalently be defined in terms of $\succ$:

$$\max_{\succ}(\|A\|^M) = \{b \in \|A\|^M \mid \nexists c \ \ (c \vDash A \ \& \ c \succ b)\}$$

It is easy to see that $\text{opt}_{\succeq}(\|A\|^M) \subseteq \max_{\succeq}(\|A\|^M)$ although the converse inclusion may fail. Typically, it will fail if there are "gaps" in the ranking.

**Example 2.4** (Gaps)**.** *Define $M = (W, \succeq, v)$, with $W = \{a, b\}$, $v(p) = W$, and $\succeq = \{(a, a), (b, b)\}$. We have $a \| b$. $\max_{\succeq}(\|p\|^M) = \{a, b\}$ but $\text{opt}_{\succeq}(\|p\|^M) = \emptyset$.*

Totalness of $\succeq$ ("for all $a, b \in W, a \succeq b$ or $b \succeq a$") may be shown to be a sufficient condition for the two notions of "best" to coincide. We have already seen that $\text{opt}_{\succeq}(\|A\|^M) \subseteq \max_{\succeq}(\|A\|^M)$. Now,

**Observation 2.5.** $\max_{\succeq}(\|A\|^M) = \text{opt}_{\succeq}(\|A\|^M)$ *if $\succeq$ is total.*

*Proof.* The right-in-left inclusion holds by definition. The left-in-right inclusion calls upon totalness. To see why, assume $\succeq$ is total, and let

[6] As mentioned, I adopt this terminology from Sen [1997].

$a \in \max_{\succeq}(\|A\|^M)$. Consider $b \in \|A\|^M$. By totalness, $a \succeq b$ or $b \succeq a$. In the second case, $a \succeq b$, since $a \in \max_{\succeq}(\|A\|^M)$. Either way, $a \succeq b$, and so $a \in \mathrm{opt}_{\succeq}(\|A\|^M)$. □

Thus, one gets two different pairs of evaluation rules depending on which of the following two equations is adopted:

$$\mathrm{best}_{\succeq}(\|A\|^M) = \max_{\succeq}(\|A\|^M) \tag{max rule}$$

$$\mathrm{best}_{\succeq}(\|A\|^M) = \mathrm{opt}_{\succeq}(\|A\|^M) \tag{opt rule}$$

Both definitions can be found in the literature.[7] From now onward, I will refer to the first equation (*resp.* second equation) as the max rule (*resp.* opt rule). From Observation 2.6, it immediately follows that, in a given model $M$ with $\succeq$ total, the same deontic formulas are true at a given world whatever rule is used.

This chapter focuses on the above two definitions of "best". As a matter of fact, variant definitions have been proposed. The purpose of these variations is often to remedy the emptiness of the set of best worlds when the betterness relation admits cycles, like in Figure 2. Condorcet's well-known voting paradox [Sen, 1969] is often used to show the plausibility of this kind of situations.

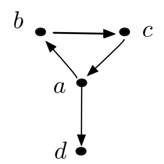



Figure 2: A top cycle. An arrow from $a$ to $b$ represents $a \geq b$. No arrow from $b$ to $a$ means $b \not\geq a$.

Hansson [2009] suggests maximizing with respect to the transitive closure $\succeq^\star$ rather than $\succeq$ itself.[8] Recall that $a \succeq^\star b$ iff there are $c_1, ..., c_n$

[7]For instance, Hansson [1969], Makinson [1993, §7.1], Schlechta [1995], Prakken and Sergot [1997], van der Torre and Tan [1997, p. 95], Horty [2001, p. 72] and Stolpe [2020] use the max rule. In contrast, Spohn [1975], Åqvist [1987; 2002], Fehige [1994, p. 43], Alchourrón [1995, p. 76], McNamara [1995], Hansen [2005, §6] work with the opt rule. Neither Goldman [1977], nor Jackson [1985], nor Hilpinen [2001, §8.5] specifies what notion of "best" is meant. (The last one uses "best" and "deontically optimal" interchangeably, but leaves optimality undefined.)

[8]There is room for variation here. Hansson considers four alternative constructions, and finally settles on that one.

such that $a \succeq c_1 \succeq ... \succeq c_n \succeq b$. I will call this variant the "quasi-maximality" (quasi-max, for short) rule:

$$\text{best}_{\succeq}(\|A\|^M) = \max_{\succeq^\star}(\|A\|^M) \qquad \text{(quasi-max rule)}$$

where

$$\max_{\succeq^\star}(\|A\|^M) = \{b \in \|A\|^M \mid \forall c \ \ ((c \vDash A \ \&\ c \succeq^\star b) \rightarrow b \succeq^\star c)\}$$

It is worth noticing that, if $\succeq$ is transitive, then $\succeq = \succeq^\star$, so that the quasi-max rule coincides with the original max rule:

**Observation 2.6.** $\max_{\succeq}(\|A\|^M) = \max_{\succeq^\star}(\|A\|^M)$ *if* $\succeq$ *is transitive.*

A thorough study of such alternative definitions must be postponed to another occasion. I will report a completeness result for the interpretation under the quasi-max rule in Section 4.3.

## 2.4 Properties of $\succeq$

The properties usually envisaged for $\succeq$ are reflexivity, transitivity, totalness, and the so-called limit assumption. The first three may be given the form:

- reflexivity: for all $a \in W, a \succeq a$;
- transitivity: for all $a, b, c \in W$, if $a \succeq b$ and $b \succeq c$, then $a \succeq c$;
- totalness: for all $a, b \in W, a \succeq b$ or $b \succeq a$.

The exact formulation of the limit assumption varies among authors. It can be given two basic forms:

Limitedness
If $\|A\| \neq \emptyset$ then $\text{best}_{\succeq}(\|A\|) \neq \emptyset$
Smoothness (or stopperedness)
If $a \vDash A$, then: either $a \in \text{best}_{\succeq}(\|A\|)$ or
$\exists b$ s.t. $b \succ a \ \&\ b \in \text{best}_{\succeq}(\|A\|)$

The name "limitedness" is from Åqvist [1987; 2002], "smoothness" from Kraus & al. [1990], and "stopperedness" from Makinson [1989]. Each of limitedness and smoothness may be specified further by identifying $\text{best}_{\succeq}(X)$ with either $\max_{\succeq}(X)$ or $\text{opt}_{\succeq}(X)$. A betterness relation $\succeq$ will be called "opt-limited" or "max-limited" depending on whether limitedness holds with respect to $\text{opt}_{\succeq}$ or $\max_{\succeq}$. Similarly, it will be called "opt-smooth" or "max-smooth" depending on whether smoothness holds with respect to $\mathrm{opt}_{\succeq}$ or $\max_{\succeq}$.[9]

This gives us four versions of the limit assumption. With the strong assumptions of transitivity and totalness, these different forms of the limit assumption coincide. However, with weaker constraints on $\succeq$, they may well diverge.

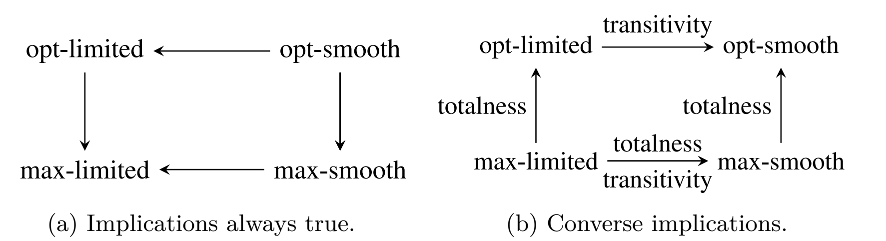
**Theorem 2.7.**
*(a) (i) opt-limitedness implies max-limitedness;*
*(ii) given totalness of $\succeq$, max-limitedness implies opt-limitedness;*
*(b) (i) opt-smoothness implies max-smoothness;*
*(ii) given totalness of $\succeq$, max-smoothness implies opt-smoothness.*

*Proof.* This follows at once from the definitions involved and Observation 2.6. □

**Theorem 2.8.**
*(a) (i) max-smoothness implies max-limitedness;*
*(ii) given transitivity and totalness of $\succeq$, max-limitedness implies max-smoothness;*
*(b) (i) opt-smoothness implies opt-limitedness;*
*(ii) given transitivity of $\succeq$, opt-limitedness implies opt-smoothness.*

*Proof.* See [Parent, 2014, Proposition 2]. □

Figure 3 represents the relationships just established in an Implication Diagram with the direction of the arrow representing that of implication. The implication relations shown in the picture on the left-hand




(a) Implications always true. (b) Converse implications.

Figure 3: Forms of the limit assumption, and their relationships.

[9] Hansson [1969] and Prakken and Sergot [1997] use max-limitedness, while Lewis [1974, p. 6], Spohn [1975], Åqvist [1987; 2002], Fehige [1994, p. 44], Alchourrón [1995, p. 84], McNamara [1995] and Hansen [2005, §6] use opt-limitedness, and Makinson [1993] and Schlechta [1995] max-smoothness. I am not aware of any authors who have considered opt-smoothness explicitly.

side hold without restriction. By contrast, those shown on the right-hand side hold under the hypothesis that $\succeq$ meets the property (or pair of properties) displayed as labeled.

In this chapter I only want to understand how the choice of a given version of the limit assumption affects the logic. The philosophical aspects of the limit assumption will not be discussed here—the reader should consult [Lewis, 1973; Fehige, 1994; McNamara, 1995; Hilpinen and McNamara, 2013]. Note that in linguistics the limit assumption has been given even more variant forms. (See, *e.g.*, the discussion in [Kaufmann, 2017].)

## 2.5 Where the opt rule *vs.* the max rule makes a difference

In this section, I give two examples of a valid formula for which the choice between the opt rule and the max rule makes a difference.

First, there is the example of the principle of rational monotony [Lehmann and Magidor, 1992], also called CV by Lewis [1973]. This is the principle

$$(P(B/A) \wedge \bigcirc(C/A)) \rightarrow \bigcirc(C/A \wedge B) \tag{RM}$$

(RM) expresses a restricted principle of strengthening of the antecedent: one can strengthen an antecedent when the added condition $B$ is permitted under the main condition $A$. Hence, doing the permitted has no effect on our other obligations.

**Observation 2.9.** *Under the opt rule, (RM) is valid if $\succeq$ is required to be transitive. Under the max rule, (RM) is valid if $\succeq$ is required to be both transitive and total.*

*Proof.* Assume that (i) $\text{opt}_{\succeq}(\|A\|) \subseteq \|C\|$, (ii) $\text{opt}_{\succeq}(\|A\|) \cap \|B\| \neq \emptyset$, and (iii) $\text{opt}_{\succeq}(\|A \wedge B\|) \not\subseteq \|C\|$. From (iii), there is some $a$ such that $a \in \text{opt}_{\succeq}(\|A \wedge B\|)$ and $a \not\models C$. From (i), $a \notin \text{opt}_{\succeq}(\|A\|)$, because $a \not\models C$. But $a \models A$. So there is some $b \models A$ with $a \not\succeq b$. From (ii), there is also some $c$ such that $c \in \text{opt}_{\succeq}(\|A\|)$ and $c \models B$. Since $c \models A \wedge B$, $a \succeq c$. Also, $c \succeq b$, since $c \in \text{opt}_{\succeq}(\|A\|)$. By transitivity, $a \succeq b$. Contradiction. Hence, under the opt rule, (RM) is valid if $\succeq$ is transitive.

For the max rule, it suffices to invoke the above along with Observation 2.6. □

While under the opt rule transitivity is sufficient for the validity of law (RM), by contrast under the max rule it is not sufficient.

**Observation 2.10.** *There is a preference model $M = (W, \succeq, v)$, in which $\succeq$ is transitive, such that (RM) fails in $M$ under the max rule.*

*Proof.* Put $M = (W, \succeq, v)$, with $W = \{a, b, c\}$, $\succeq = \{(a, b)\}$ and $v(p) = W$, $v(q) = \{b, c\}$ and $v(r) = \{a, c\}$. The model is depicted in Figure 4, where $\succeq$ is (vacuously) transitive. We have $\max_{\succeq}(\|p\|) = \{a, c\}$, $\max_{\succeq}(\|p \wedge q\|) = \{b, c\}$, $\|q\| = \{b, c\}$ and $\|r\| = \{a, c\}$. Under the max rule, (RM) fails, since $\bigcirc(r/p)$ and $P(q/p)$ hold while $\bigcirc(r/p \wedge q)$ does not (witness: $b$). □

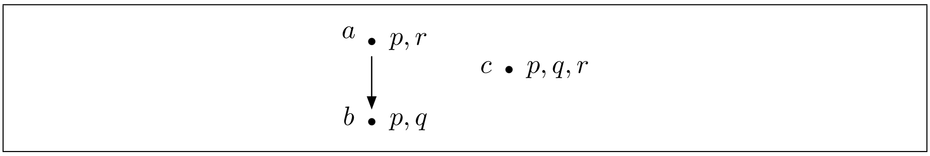



Figure 4: A countermodel to (RM)

What I say here about (RM) applies analogously to the following formula, named after Spohn [1975], who used it in his axiomatization of Hansson's system DSDL3:

$$(P(B/A) \wedge \bigcirc(B \rightarrow C/A)) \rightarrow \bigcirc(C/A \wedge B) \tag{Sp}$$

We will see that (Sp) and (RM) are equivalent.[10] Spohn [1975, p. 247] himself argues that the assumption of totalness is iddle. He can do so only because he uses the opt rule instead of the max rule.

Here is my second example of a validity for which the choice between the opt rule and the max rule makes a difference:

$$P(A/A \vee B) \wedge P(B/B \vee C) \rightarrow P(A/A \vee C) \tag{$\gg$-trans}$$

($\gg$-trans) expresses a principle of transitivity for a notion of weak preference over formulas given by $A \gg B =_{def} P(A/A \vee B)$.[11] This says that $A$ is ranked as at least as high as $B$ iff it is permitted that $A$ on the condition that either $A$ or $B$.

**Observation 2.11.** *Under the opt rule, ($\gg$-trans) is valid if $\succeq$ is required to be transitive. Under the max rule, ($\gg$-trans) is valid if $\succeq$ is required to be both transitive and total.*

[10] Cf. Theorem 3.3 in Section 3.1.
[11] Cf. [Lewis, 1973, p. 54].

*Proof.* Assume that (i) $\mathrm{opt}_{\succeq}(\|A \vee B\|) \cap \|A\| \neq \emptyset$, (ii) $\mathrm{opt}_{\succeq}(\|B \vee C\|) \cap \|B\| \neq \emptyset$, and (iii) $\mathrm{opt}_{\succeq}(\|A \vee C\|) \cap \|A\| = \emptyset$. From (i), there is some $a$ such that $a \in \mathrm{opt}_{\succeq}(\|A \vee B\|)$ and $a \models A$. From (ii), there is some $b$ such that $b \in \mathrm{opt}_{\succeq}(\|B \vee C\|)$ and $b \models B$. From (iii), $a \notin \mathrm{opt}_{\succeq}(\|A \vee C\|)$. Since $a \models A \vee C$, there is some $c$ such that $c \models A \vee C$ and $a \not\succeq c$. Since $a \in \mathrm{opt}_{\succeq}(\|A \vee B\|)$ and $a \not\succeq c$, $c \not\models A \vee B$, and so $c \not\models A$ and $c \not\models B$. Hence $c \models C$, and so $c \models B \vee C$. Thus, $b \succeq c$. By transitivity of $\succeq$, $a \not\succeq b$. On the other hand, since $a \in \mathrm{opt}_{\succeq}(\|A \vee B\|)$ and $b \models A \vee B$, $a \succeq b$. Contradiction.

For the max rule, it suffices to invoke the above along with Observation 2.6. □

While under the opt rule transitivity is sufficient for ($\gg$-trans), under the max rule it is not:

**Observation 2.12.** *There is a preference model $M = (W, \succeq, v)$, with $\succeq$ transitive, such that ($\gg$-trans) fails in $M$ under the max rule.*

*Proof.* Put $M = (W, \succeq, v)$, with $W = \{a, b, c\}$, $\succeq = \{(a, b)\}$ and $v(p) = \{b\}$, $v(q) = \{b, c\}$ and $v(r) = \{a\}$. This is shown in Figure 5. We have $\max_{\succeq}(\|p \vee q\|) = \{b, c\}$, $\max_{\succeq}(\|q \vee r\|) = \{a, c\}$, $\max_{\succeq}(\|p \vee r\|) = \{a\}$, $\|p\| = \{b\}$, $\|q\| = \{b, c\}$ and $\|r\| = \{a\}$. □

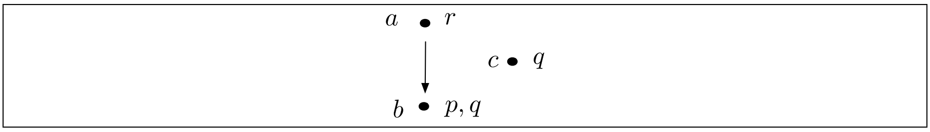



Figure 5: A countermodel to ($\gg$-trans)

(RM) and ($\gg$-trans) are two sample formulas for which the choice between the max rule and the opt rule makes a difference. To get the whole picture (or to get closer to it), we need to extend the scope of our study to examine not individual formulas (chosen randomly) but axiomatic systems. This will be done in Section 4.

## 2.6 Selection functions

This section provides some background information on so-called selection function semantics. It may seem a distraction from the focus on Hanssonian-type preference semantics. However, this material is needed for subsequent developments, especially in Section 5.1.

Stemming from Stalnaker [1968] and generalized by Chellas [1975], such a semantics was adapted to the present setting by Åqvist [2002].

I call these new structures "selection function models", to distinguish them from those described above. In models of this sort, the betterness relation $\succeq$ is replaced with a so-called selection function $\mathfrak{f}$ from formulas to subsets of $W$, such that, for all $A$ in $\mathcal{L}$, $\mathfrak{f}(A) \subseteq W$. Intuitively, $\mathfrak{f}(A)$ outputs all the best worlds satisfying $A$. The evaluation rule for the dyadic obligation operator is phrased thus:

$$M, a \vDash \bigcirc(B/A) \quad \text{iff} \quad \mathfrak{f}(A) \subseteq \|B\|^M$$

From this, on derives the following evaluation rule for permission:

$$M, a \vDash P(B/A) \quad \text{iff} \quad \mathfrak{f}(A) \cap \|B\|^M \neq \emptyset$$

The relevant constraints for $\mathfrak{f}$ are:

| | | |
|---|---|---|
| ($\mathfrak{f}$0) | If $\|A\|^M = \|B\|^M$ then $\mathfrak{f}(A) = \mathfrak{f}(B)$ | (Syntax-independence) |
| ($\mathfrak{f}$1) | $\mathfrak{f}(A) \subseteq \|A\|^M$ | (Inclusion) |
| ($\mathfrak{f}$2) | $\mathfrak{f}(A) \cap \|B\|^M \subseteq \mathfrak{f}(A \wedge B)$ | (Chernoff) |
| ($\mathfrak{f}$3) | If $\|A\|^M \neq \emptyset$ then $\mathfrak{f}(A) \neq \emptyset$ | (Consistency-preservation) |
| ($\mathfrak{f}$4) | If $\mathfrak{f}(A) \subseteq \|B\|^M$ then $\mathfrak{f}(A \wedge B) \subseteq \mathfrak{f}(A)$ | (Aizerman) |
| ($\mathfrak{f}$5) | If $\mathfrak{f}(A) \cap \|B\|^M \neq \emptyset$ then $\mathfrak{f}(A \wedge B) \subseteq \mathfrak{f}(A) \cap \|B\|^M$ | (Arrow) |

The reason why these conditions may be regarded as most central will become apparent in Section 3, when moving to the proof theory. Åqvist does not use ($\mathfrak{f}$4). It is weaker than ($\mathfrak{f}$5) in the following sense.

**Fact 2.13.** *Given* ($\mathfrak{f}$0) *and* ($\mathfrak{f}$3), ($\mathfrak{f}$5) *implies* ($\mathfrak{f}$4)*, but not vice versa (even in the presence of* ($\mathfrak{f}$1) *and* ($\mathfrak{f}$2)*).*

*Proof.* Let $\mathfrak{f}(A) \subseteq \|B\|$. Either (i) $\|A\| \neq \emptyset$ or (ii) $\|A\| = \emptyset$. In case (i), $\mathfrak{f}(A) \neq \emptyset$, by ($\mathfrak{f}$3). Thus, $\mathfrak{f}(A) \cap \|B\| \neq \emptyset$. ($\mathfrak{f}$5) then yields the desired result. In case (ii), $\|A\| = \|A \wedge B\|$. By ($\mathfrak{f}$0), $\mathfrak{f}(A) = \mathfrak{f}(A \wedge B)$. So $\mathfrak{f}(A \wedge B) \subseteq \mathfrak{f}(A)$ as required.

To show that the converse implication may fail even in the presence of ($\mathfrak{f}$0)-($\mathfrak{f}$3), let $M = (W, \mathfrak{f}, v)$ be such that $W = \{a, b, c\}$, $v(p) = \{a, b\}$ and $v(q) = W$ for all $q$ other than $p$, and

$$\mathfrak{f}(A) = \begin{cases} \{a, c\} & \text{if } \|A\| = W \\ \|A\| & \text{otherwise} \end{cases}$$

($\mathfrak{f}$0), ($\mathfrak{f}$1), ($\mathfrak{f}$2) and ($\mathfrak{f}$3) hold, and so does ($\mathfrak{f}$4). But ($\mathfrak{f}$5) fails:

$$\mathfrak{f}(q \wedge p) = \{a, b\} \nsubseteq \mathfrak{f}(q) \cap \|p\| = \{a\} \neq \emptyset$$

This concludes the proof. $\square$

The names used for the first four constraints are from Parent [2015]. All these constraints have known counterparts within the framework of rational choice theory (for an overview, see Moulin [1985]). ($\mathfrak{f}2$) is identical to so-called Chernoff's [1954] condition also known as Sen's condition $\alpha$. ($\mathfrak{f}4$) may be regarded as a reformulation of the condition called "Aizerman" in Moulin [1985] and in Lindström [1991]. Therefore it will henceforth be referred to as the Aizerman condition. Strictly speaking, this one is:

$$(\mathfrak{f}4^{\star}) \qquad \text{If } \mathfrak{f}(A) \subseteq \|B\| \subseteq \|A\| \text{ then } \mathfrak{f}(B) \subseteq \mathfrak{f}(A)$$

It is not difficult to see that, given ($\mathfrak{f}0$) and ($\mathfrak{f}1$), ($\mathfrak{f}4^{\star}$) and ($\mathfrak{f}4$) are equivalent.

**Fact 2.14.** *Given (*$\mathfrak{f}0$*) and (*$\mathfrak{f}1$*), (*$\mathfrak{f}4^{\star}$*) and (*$\mathfrak{f}4$*) are equivalent.*

*Proof.* I first verify that, given ($\mathfrak{f}1$), ($\mathfrak{f}4^{\star}$) implies ($\mathfrak{f}4$). Assume $\mathfrak{f}(A) \subseteq \|B\|$. By ($\mathfrak{f}1$), $\mathfrak{f}(A) \subseteq \|A\| \cap \|B\| = \|A \wedge B\| \subseteq \|A\|$. By ($\mathfrak{f}4^{\star}$), $\mathfrak{f}(A \wedge B) \subseteq \mathfrak{f}(A)$, as required. For the converse implication, let $\mathfrak{f}(A) \subseteq \|B\| \subseteq \|A\|$. On the one hand, by ($\mathfrak{f}0$) $\mathfrak{f}(A \wedge B) = \mathfrak{f}(B)$, since $\|A \wedge B\| = \|B\|$. On the other hand, a direct application of ($\mathfrak{f}4$) to $\mathfrak{f}(A) \subseteq \|B\|$ yields $\mathfrak{f}(A \wedge B) \subseteq \mathfrak{f}(A)$. Putting the two together, one gets $\mathfrak{f}(B) \subseteq \mathfrak{f}(A)$ as required. □

($\mathfrak{f}5$) may similarly be regarded as a reformulation of the condition which Hansson [1968] calls Arrow, and so I will henceforth refer to it as the Arrow condition. Strictly speaking, this one is:

$$(\mathfrak{f}5^{\star}) \qquad \text{If } \|A\| \subseteq \|B\| \text{ and } \mathfrak{f}(B) \cap \|A\| \neq \emptyset \text{ then } \mathfrak{f}(A) = \mathfrak{f}(B) \cap \|A\|$$

**Fact 2.15.** *Given (*$\mathfrak{f}0$*)-(*$\mathfrak{f}3$*), (*$\mathfrak{f}5^{\star}$*) and (*$\mathfrak{f}5$*) are equivalent.*

*Proof.* See Hansen [1998]. □

Not much more will be needed later about selection functions.

# 3 Proof systems

This section presents the proof systems to be studied in this chapter.

## 3.1 Mixed alethic-deontic logics

I will primarily be concerned with four mixed alethic-deontic logics of increasing strength: **E**, **F**, **F**+(CM) and **G**. Systems **E**, **F** and **G** are from Åqvist [1987; 2002]. They correspond to his reconstruction of Hansson [1969]'s system DSDL1, DSDL2 and DSDL3, respectively. **F**+(CM) is from Parent [2014]. The list of all the relevant axioms is given below. For some of the axioms, I introduce special labels in order to facilitate reference to them later on.

The notions of theoremhood, deducibility and consistency (with respect to a given system) are defined as usual. I write $\vdash A$ if $A$ is provable, and $\Gamma \vdash A$ if $A$ is derivable from $\Gamma$, where $\Gamma$ is a set of wffs.

System **E** is defined by the following axioms and rules:

| | |
|---|---|
| Any axiomatization of classical propositional logic | (PL) |
| S5-schemata for $\Box$ | (S5) |
| $\bigcirc(B \to C/A) \to (\bigcirc(B/A) \to \bigcirc(C/A))$ | (COK) |
| $\bigcirc(B/A) \to \Box\bigcirc(B/A)$ | (Abs) |
| $\Box A \to \bigcirc(A/B)$ | (Nec) |
| $\Box(A \leftrightarrow B) \to (\bigcirc(C/A) \leftrightarrow \bigcirc(C/B))$ | (Ext) |
| $\bigcirc(A/A)$ | (Id) |
| $\bigcirc(C/A \wedge B) \to \bigcirc(B \to C/A)$ | (Sh) |
| If $\vdash A$ and $\vdash A \to B$ then $\vdash B$ | (MP) |
| If $\vdash A$ then $\vdash \Box A$ | (N) |

The abbreviations (PL), (S5), (MP) and (N) are self-explanatory. (COK) is the conditional analogue of the familiar distribution axiom K. (Abs) is the absoluteness axiom of [Lewis, 1973], and reflects the fact that the ranking is not world-relative. (Nec) is the deontic counterpart of the familiar necessitation rule. (Ext) permits the replacement of necessarily equivalent sentences in the antecedent of deontic conditionals. (Id) is the deontic analogue of the identity principle. (Sh) is named after Shoham [1988, p. 77], who seems to have been the first to discuss it. The question of whether (Id) is a reasonable law for deontic conditionals has been much debated. A defence of (Id) can be found in

Hansson [1969] and Prakken and Sergot [1997]–this line of defence is discussed in Parent [2012]. (For a different diagnosis, see also Spohn [1975], Makinson [1993], Alchourrón [1993] and Parent [2001].)

For future reference I introduce the following derived principles:

$$\text{If } \vdash A \leftrightarrow B \text{ then } \vdash \bigcirc(C/A) \leftrightarrow \bigcirc(C/B) \tag{LLE}$$

$$\text{If } \vdash B \rightarrow C \text{ then } \vdash \bigcirc(B/A) \rightarrow \bigcirc(C/A) \tag{RW}$$

$$\bigcirc(B/A) \wedge \bigcirc(C/A) \rightarrow (\bigcirc(B \wedge C/A) \tag{AND}$$

$$\bigcirc(C/A) \wedge \bigcirc(C/B) \rightarrow (\bigcirc(C/A \vee B) \tag{OR}$$

$$\bigcirc(C/A) \wedge \bigcirc(D/B) \rightarrow \bigcirc(C \vee D/A \vee B) \tag{OR$'$}$$

The labels (LLE) and (RW) are borrowed from the non-monotonic logic literature. (LLE) and (RW) are mnemonic for "Left Logical Equivalence" and "Right Weakening", respectively.

**Theorem 3.1.** *(LLE), (RW), (AND), (OR) and (OR′) are derivable in system* **E**.

*Proof.* The proofs of (LLE) and (RW) are straightforward, and left to the reader.

For (AND), assume $\bigcirc(B/A)$ and $\bigcirc(C/A)$. From the first, one gets $\bigcirc(C \rightarrow (B \wedge C)/A)$ by (RW). (COK) gives $\bigcirc(C/A) \rightarrow \bigcirc(B \wedge C/A)$. From this and the second hypothesis, one gets $\bigcirc(B \wedge C/A)$.

For (OR), assume $\bigcirc(C/A)$ and $\bigcirc(C/B)$. Using (Ext), one gets $\bigcirc(C/(A \vee B) \wedge A)$ and $\bigcirc(C/(A \vee B) \wedge B)$. By (Sh), $\bigcirc(A \rightarrow C/A \vee B)$ and $\bigcirc(B \rightarrow C/A \vee B)$. By (AND), $\bigcirc((A \rightarrow C) \wedge (B \rightarrow C)/A \vee B)$. By (RW), $\bigcirc((A \vee B) \rightarrow C/A \vee B)$. By (Id), $\bigcirc(A \vee B/A \vee B)$. By (COK), one then gets $\bigcirc(C/A \vee B)$.

(OR′) is easily derived using (OR) and (RW). □

Theorems 3.1 and 3.2 tell us that **E** is equivalently axiomatized by replacing, in **E**, (COK) and (Sh) with (RW), (AND) and (OR).

**Theorem 3.2.** *(COK) is derivable from (RW) and (AND). (Sh) is derivable from (RW), (Id), (OR) and (LLE).*

*Proof.* For (COK), assume $\bigcirc(B \rightarrow C/A)$ and $\bigcirc(B/A)$. By (AND), $\bigcirc((B \rightarrow C) \wedge B/A)$. By (RW), $\bigcirc(C/A)$.

For (Sh), suppose $\bigcirc(C/A \wedge B)$. By (RW), $\bigcirc(B \rightarrow C/A \wedge B)$. By (Id) and (RW), $\bigcirc(B \rightarrow C/A \wedge \neg B)$. By (OR) and (LLE), $\bigcirc(B \rightarrow C/A)$. □

The basis of **F** is that of **E** with the single extra axiom:

$$\Diamond A \rightarrow (\bigcirc(B/A) \rightarrow P(B/A)) \tag{D$^\star$}$$

(D$^\star$) is the conditional analogue of the familiar axiom D. Its import is simply that conflicts of obligations are ruled out, for possible antecedents.

**F**+(CM) and **G** are obtained by supplementing **F** with (CM) and (Sp), respectively:

$$(\bigcirc(B/A) \wedge \bigcirc(C/A)) \rightarrow \bigcirc(C/A \wedge B) \tag{CM}$$

$$(P(B/A) \wedge \bigcirc(B \rightarrow C/A)) \rightarrow \bigcirc(C/A \wedge B) \tag{Sp}$$

(CM) is the principle of cautious monotony from the non-monotonic logic literature.[12] It can be shown that (CM) and (D$^\star$) are independent of each other, given the other axioms of **F**. This is why their addition is considered separately of one another. In the presence of (CM), the following two principles are derivable:

$$\bigcirc(B/A) \wedge \bigcirc(A/B) \rightarrow (\bigcirc(C/A) \leftrightarrow \bigcirc(C/B)) \tag{CSO}$$

$$\bigcirc(A/A \vee B) \wedge \bigcirc(B/B \vee C) \rightarrow \bigcirc(A/A \vee C) \tag{$\geq$-trans}$$

(CSO) is familiar from the literature on conditional logic. It says that two “deontically” equivalent states of affairs trigger the same obligations. And ($\geq$-trans) expresses a principle of transitivity for a weak notion of preference defined by $A \geq B$ iff $\bigcirc(A/A \vee B)$.[13]

As mentioned, (Sp)—the distinctive axiom of **G**—is equivalent to the principle of rational monotony (RM):[14]

$$(P(B/A) \wedge \bigcirc(C/A)) \rightarrow \bigcirc(C/A \wedge B) \tag{RM}$$

**F**+(CM) is strictly included in **G**, because (CM) is derivable in **G**, but (Sp) is not derivable in **F**+(CM).

**Theorem 3.3.**
*(i) (CM) and (D$^\star$) are independent, given the other axioms of* **F***;*
*(ii) (CSO) is a theorem of* **F***+(CM);*
*(iii) ($\geq$-trans) is a theorem of* **F***+(CM);*
*(iv) (Sp) and (RM) are inter-derivable in* **E***;*
*(v) (CM) is a theorem of* **G***;*
*(vi) (Sp) is not a theorem of* **F***+(CM).*

*Proof.* The proof of (i) may be found in [Parent, 2014, Section 2.5].

[12] Cf. [Kraus *et al.*, 1990].
[13] Cf. [Kraus *et al.*, 1990, p. 194].
[14] Cf. Section 2.5.

For (ii), assume $\bigcirc(B/A)$, $\bigcirc(A/B)$ and $\bigcirc(C/A)$. From the first and third assumptions, $\bigcirc(C/A \wedge B)$, by (CM). This is equivalent to $\bigcirc(C/B \wedge A)$ by (Ext). Using (Sh), $\bigcirc(A \rightarrow C/B)$. From this together with the second assumption, one then gets $\bigcirc(C/B)$, by (RW). For the derivation of $\bigcirc(C/A)$ from $\bigcirc(C/B)$, the argument is similar. This establishes (CSO).

For (iii), assume $\bigcirc(A/A \vee B)$ and $\bigcirc(B/B \vee C)$. Using (OR$'$) and (Ext), $\bigcirc(A \vee B/A \vee B \vee C)$. By (Id) and (RW), $\bigcirc(A \vee B \vee C/A \vee B)$ is a theorem. Using (CSO), one immediately gets $\bigcirc(A/A \vee B \vee C)$. By (Id), $\bigcirc(C/C)$. By (OR$'$) and (Ext), one gets $\bigcirc(A \vee C/A \vee B \vee C)$. By (CM), $\bigcirc(A/(A \vee B \vee C) \wedge (A \vee C))$. By (Ext), $\bigcirc(A/A \vee C)$.

For (iv), suppose $P(B/A)$ and $\bigcirc(C/A)$. By (RW), $\bigcirc(B \rightarrow C/A)$, and so $\bigcirc(C/A \wedge B)$ by (Sp). Conversely, suppose $P(B/A)$ and $\bigcirc(B \rightarrow C/A)$. By (RM), $\bigcirc(B \rightarrow C/A \wedge B)$. Hence $\bigcirc(B \wedge (B \rightarrow C)/A)$ by (Sh). One then gets $\bigcirc(C/A)$ by (RW).

For (v). Suppose $\bigcirc(B/A)$ and $\bigcirc(C/A)$. Either $\Diamond A$ or $\neg\Diamond A$. In the first case, $P(B/A)$ by (D$^\star$), and so $\bigcirc(C/A \wedge B)$ by (RM). In the second case, $\Box(A \leftrightarrow (A \wedge B))$, and thus $\bigcirc(C/A \wedge B)$ by (Ext). Either way, $\bigcirc(C/A \wedge B)$.

The proof of (vi) is given in Appendix A, where I make use of an observation which will be available only later. $\square$

### 3.2 Pure deontic conditional logics

The above systems are mixed alethic-deontic logics. Goble [2015, p. 94] shows that each of **F**, **F**+(CM) and **G** has a "pure deontic conditional" counterpart. I borrow this term from Alchourrón [1995, p. 87], who uses the term "pure conditional axiomatisation" to refer to an axiomatisation in a language in which we only have the conditional (obligation) operator as a primitive connective added to those of classical propositional logic. This language still allows iterated modalities and mixed formulas, and thus is still distinct from the language of Hansson's systems.

The key point is that in systems **F**, **F**+(CM) and **G**, the alethic operators $\Box$ and $\Diamond$ become superfluous, because $\Box A$ and $\Diamond A$ turn out to be equivalent with $\bigcirc(\bot/\neg A)$ and $P(\top/A)$, respectively. (This is not the case in **E**, and this is why it is left out of the picture.) Thus, in the description of the three systems, one might eliminate all occurrences of $\Box$ and $\Diamond$ using these definitions, so that everything is written using the deontic modalities only. Drawing on this idea Goble defines three systems, called DDL-D-3, DDL-D-4 and DDL-D-5, using a language with no other primitive modality than $\bigcirc(–/–)$. (Nevertheless, to avoid cumbersome notation $\Box$ and $\Diamond$ are kept in the language as derived connectives.)[15] The distinctive axiom of DDL-D-4 is (CM), while that of DDL-D-5 is (RM). Roughly speaking, DDL-D-3 may be described as the system that results from **F** by leaving out ($D^\star$) (its pure conditional counterpart follows from the other axioms), and by replacing all occurrences of $\Box$ and $\Diamond$ by their definition throughout in (S5), (Abs), (Nec), (Ext) and (N). Goble's own axiomatic characterisation of DDL-D-3 is as follows:

$$
\begin{array}{lr}
\text{Any axiomatization of classical propositional logic} & \text{(PL)} \\
\Box A \to A \ (\textit{aka} \ \bigcirc(\bot/\neg A) \to A) & \text{(T)} \\
\text{If } \vdash A \leftrightarrow B \text{ then } \vdash \bigcirc(C/A) \leftrightarrow \bigcirc(C/B) & \text{(LLE)} \\
\text{If } \vdash B \to C \text{ then } \vdash \bigcirc(B/A) \to \bigcirc(C/A) & \text{(RW)} \\
\bigcirc(B/A) \wedge \bigcirc(C/A) \to (\bigcirc(B \wedge C/A) & \text{(AND)} \\
\bigcirc(B/A) \wedge \bigcirc(B/C) \to (\bigcirc(B/A \vee C) & \text{(OR)} \\
\bigcirc(A/A) & \text{(Id)} \\
\bigcirc(B/A) \to \bigcirc(\bigcirc(B/A)/C) & \text{(D}\bigcirc\text{4)} \\
P(B/A) \to \bigcirc(P(B/A)/C) & \text{(D}\bigcirc\text{5)} \\
\text{If } \vdash A \text{ and } \vdash A \to B \text{ then } \vdash B & \text{(MP)} \\
\text{If } \vdash A \text{ then } \vdash \Box A \ (\textit{aka} \ \bigcirc(\bot/\neg A)) & \text{(N}'\text{)}
\end{array}
$$

(D$\bigcirc$4) and (D$\bigcirc$5) are the dyadic generalization of the well-known principles (4) $\bigcirc A \to \bigcirc\bigcirc A$ and (5) $PA \to \bigcirc PA$. (T) and (N$'$) are self-explanatory.

Goble writes that "DDL-D-3 is equivalent to **F**, DDL-D-4 to **F**+(CM) and DDL-D-5 to **G**" [Goble, 2015, p. 102]. All the axioms and rules of each member of the DDL-D family are derivable in the corresponding mixed alethic-deontic logic. Hence the inclusions:

$$\text{DDL-D-3} \subseteq \mathbf{F} \qquad \text{DDL-D-4} \subseteq \mathbf{F}+\text{(CM)} \qquad \text{DDL-D-5} \subseteq \mathbf{G}$$

The converse inclusions also hold insofar as $\Box$ and $\Diamond$ are kept as derived connectives in the language of the pure deontic logics, and identified with those appearing in the language of the corresponding mixed alethic-deontic logics. The initial goal was to identify the pure conditional counterparts of Åqvist's systems. For the sake of consistency, one

[15] If in Goble's manuscript we look more closely at the two pairs of operators, we see a subtle difference in notation between them—Åqvist's operators are written as "□" and "◊", while Goble's operators are written as "□" and "◇". For simplicity's sake I will use the same notation for both pairs.

may prefer not to have $\Box$ and $\Diamond$ in the language of the pure deontic logics as a derived connective. In that case, the relationship between the two families of systems should be described differently. One suggestion is to say that each of **F**, **F**+(CM) and **G** can faithfully be embedded into their counterpart in the DDL-D family. That is: there is a translation $\star$ from the language of **F**, **F**+(CM) and **G** into the language of DDL-D-3, DDL-D-4 and DDL-D-5, such that $\star$ preserves both theoremhood and unprovability.

Figure 6 provides a map of the systems I have discussed. An arrow indicates (proper) containment in the sense that the system from which the arrow starts contains all the theorems of the system at which the arrow points, but not vice versa. The systems to the left of the dashed line are mixed alethic-deontic logics, while those to its right are pure deontic logics.

One can find more systems in the literature. In particular, there are also Hansson's DSDL1-3 as axiomatized by Goble [2019], or Lewis's system VN of [1973], which turns out to be equivalent with van Fraassen's system CD of [1972] and Goble's system SDDL of [2003]. However, none will be a part of the discussion. I mentioned that **F** and **G** were meant to be a reconstruction of Hansson's systems DSDL2 and DSDL3. Neither of **F** and **G** contains its DSDL counterpart. Both DSDL2 and DSDL3 have the rule "If $\nvdash \neg A$, then $\vdash P(\top/A)$", while neither of **F** and **G** does.

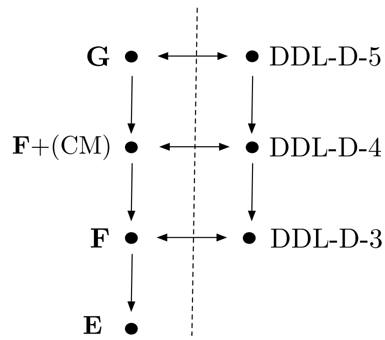



Figure 6: Systems

# 4 Determination results

This section gives a survey of the determination (*i.e.*, soundness and completeness) results available at the time of writing this chapter. Here I shall be primarily interested in the mixed systems put forth by Åqvist. Two determination results are new. Their proof may be found in the Appendices. To keep this chapter at a reasonable length, the proofs of the other results are omitted. Soundness and completeness are understood in their strong version: they conjointly establish a match between the deductibility and the semantic consequence relations, with no restriction on the cardinality of the premise set $\Gamma$. The statement of the theorem is written in the form "$\Gamma \vdash A$ iff $\Gamma \models A$".

## 4.1 Core results

A synopsis of the core determination results is given in Table 1.

| Properties of $\succeq$ | max | opt |
|---|---|---|
| binary relation | **E** | **E** |
| limitedness | **F** | **F** |
| smoothness | **F**+(CM) | **F**+(CM) |
| smoothness<br>transitivity | **F**+(CM) | **G** |

Table 1: Core results

This table must be read as follows. The leftmost column shows the constraints placed on $\succeq$. The top row covers the class of all preference models; one does not require any special properties of $\succeq$ apart from being a relation. The other two columns show the corresponding systems, the middle column for models applying the max rule, and the rightmost one for models applying the opt rule. It is understood that limitedness is defined for max in the max column, and for opt in the opt column.

Below I state formally the results reported in Table 1.

**Theorem 4.1.**

*(i) Under the opt rule (*resp.*, the max rule),* **E** *is sound and complete with respect to the class of all preference models;*

*(ii) Under the opt rule (*resp.*, the max rule),* **F** *is sound and complete with respect to the class of preference models in which* $\succeq$ *is opt-limited (resp. max-limited).*

*Proof.* See [Parent, 2015]. □

**Theorem 4.2.**

*(i) Under the opt rule (*resp.*, the max rule),* $\mathbf{F} + (CM)$ *is sound and complete with respect to the class of preference models in which* $\succeq$ *is opt-smooth (resp. max-smooth);*

*(ii) Under the max rule,* $\mathbf{F} + (CM)$ *is sound and complete with respect to the class of preference models in which* $\succeq$ *is max-smooth and transitive.*

*Proof.* For (i), see [Parent, 2014]. For (ii), see Appendix B. □

Theorem 4.2 (ii) tells us that, under the max rule, and given max-smoothness, the transitivity of $\succeq$ has no import. We will see that this also holds in the absence of max-smoothness. These results are in sharp contrast with those for the opt rule. For instance, in the presence of opt-smoothness, transitivity boosts the logic from **F**+(CM) to **G**.

**Theorem 4.3.** *Under the opt rule,* **G** *is sound and complete with respect to the class of preference models in which* $\succeq$ *is opt-smooth and transitive.*

*Proof.* See [Parent, 2014; Parent, 2008]. □

## 4.2 Adding reflexivity and totalness

Table 2 shows what happens when the constraints of reflexivity and of totalness are added. Reflexivity has no import. Totalness makes a difference only under the max rule in one case, when it is combined with transitivity and smoothness. Below I state formally the results shown in the table.

**Theorem 4.4.**

*(i) Under the opt rule (*resp.*, the max rule),* **E** *is sound and complete with respect to:*
  - *(a) the class of preference models in which* $\succeq$ *is reflexive;*
  - *(b) the class of preference models in which* $\succeq$ *is total.*

*(ii) Under the opt rule (*resp.*, the max rule),* **F** *is sound and complete with respect to:*
  - *(a) the class of preference models in which* $\succeq$ *is opt-limited (*resp.*, max-limited) and reflexive;*
  - *(b) the class of preference models in which* $\succeq$ *is opt-limited (*resp.*, max-limited) and total.*

*Proof.* See [Parent, 2015]. □

**Theorem 4.5.**

*(i) Under the opt rule (*resp.*, the max rule),* $\mathbf{F} + (CM)$ *is sound and complete with respect to:*

*(a) the class of preference models in which* $\succeq$ *is opt-smooth (*resp.*, max-smooth) and reflexive;*

*(b) the class of preference models in which* $\succeq$ *is opt-smooth (*resp.*, max-smooth) and total.*

*(ii) Under the max rule,* $\mathbf{F} + (CM)$ *is sound and complete with respect to the class of preference models in which* $\succeq$ *is max-smooth, transitive and reflexive.*

*Proof.* For (i), see [Parent, 2014]. For (ii), see Appendix B. □

| Properties of $\succeq$ | max | opt |
|---|---|---|
| reflexivity | **E** | **E** |
| totalness | **E** | **E** |
| limitedness<br>reflexivity | **F** | **F** |
| limitedness<br>totalness | **F** | **F** |
| smoothness<br>reflexivity | **F**+(CM) | **F**+(CM) |
| smoothness<br>totalness | **F**+(CM) | **F**+(CM) |
| smoothness<br>transitivity<br>reflexivity | **F**+(CM) | **G** |
| smoothness<br>transitivity<br>totalness | **G** | **G** |

Table 2: Adding reflexivity and totalness

**Theorem 4.6.**

*(i) Under the opt rule,* $\mathbf{G}$ *is sound and complete with respect to:*

*(a) the class of preference models in which* $\succeq$ *is opt-smooth, transitive and reflexive;*

(b) *the class of preference models in which $\succeq$ is opt-smooth, transitive and total.*

(ii) *Under the max rule,* **G** *is sound and complete with respect to the class of preference models in which $\succeq$ is max-smooth, transitive and total.*

*Proof.* See [Parent, 2014]. □

### 4.3 Transitivity without smoothness (max rule)

This section reports two determination results for the transitive (and not necessarily smooth) case.

**Theorem 4.7.** *Under the max rule,* **E** *is sound and complete with respect to:*

(i) *the class of preference models in which $\succeq$ is transitive;*

(ii) *the class of preference models in which $\succeq$ is transitive and reflexive.*

*Proof.* See Appendix C. □

**Theorem 4.8.** *Under the max rule,* **F** *is sound and complete with respect to:*

(i) *the class of preference models in which $\succeq$ is max-limited and transitive;*

(ii) *the class of preference models in which $\succeq$ is max-limited, transitive and reflexive.*

*Proof.* See Appendix C. □

I summarize these results in Table 3.

| Properties of $\succeq$ | max | opt |
|---|---|---|
| transitivity | **E** | ? |
| transitivity<br>reflexivity | **E** | ? |
| transitivity<br>limitedness | **F** | **G** |
| transitivity<br>limitedness<br>reflexivity | **F** | **G** |

Table 3: Non-smooth transitive betterness under the max rule

The middle column tells us that, under the max rule, transitivity alone has no import, be it combined or not with reflexivity, and be it combined or not with limitedness. This observation does not carry over to the opt rule. Transitivity combined with opt-limitedness boosts the logic from **F**+(CM) to **G**. (Given transitivity, opt-limitedness and opt-smoothness are equivalent.) On the other hand, consider ($\gg$-trans):

$$P(A/A \vee B) \wedge P(B/B \vee C) \rightarrow P(A/A \vee C) \tag{$\gg$-trans}$$

We know that under the opt-rule ($\gg$-trans) is valid if $\succeq$ is required to be transitive (cf. Observation 2.11). Thus, under the opt rule, the system obtained by supplementing **E** with (Sp) and ($\gg$-trans) is sound with respect to the class of preference models in which $\succeq$ is transitive and with respect to the class of those in which it is also reflexive. It is not known whether it is also complete with respect to these two classes of models.[16] This is indicated by a question mark in Table 3.

In Section 2.3, I mentioned the possibility of defining "best" in terms of maximization under the transitive closure $\succeq^{\star}$ of $\succeq$. I called this rule of interpretation the quasi-max rule. One has:

**Theorem 4.9.** *Under the quasi-max rule,* **E** *is sound and complete with respect to:*
*(i) the class of all preference models;*
*(ii) the class of preference models in which $\succeq$ is reflexive.*

*Proof.* This follows from Theorem 4.7 and Observation 2.6. □

## 4.4 Pure deontic conditional counterparts

Analogous results have been obtained by Goble for his pure deontic systems DDL-D-3, DDL-D-4 and DDL-D-5. Table 4 summarizes these results. As far as the contrast between maximality and optimality is concerned, the story seems to remain the same. I shall make two comments.

- First, there is no known determination result for (i) the class of all preference models (ii) the class of those in which $\succeq$ is required to be reflexive, and (iii) the class of those in which $\succeq$ is required to be total. Hence the presence of a question mark in the relevant cells.

[16] This is also pointed out by [Goble, 2015].

- Second, the axiomatic counterpart of the limitedness assumption changes. In Åqvist's systems, limitedness corresponds to (D$^\star$), whose pure deontic conditional counterpart is a theorem of DDL-D-2. The limitedness assumption validates the (T) axiom; this one takes over the role of (D$^\star$):

$$\Box A \to A \ (aka \bigcirc (\bot/\neg A) \to A) \tag{T}$$

| Properties of $\succeq$ | max | opt |
|---|---|---|
| Binary relation | ? | ? |
| reflexivity | ? | ? |
| totalness | ? | ? |
| limitedness | DDL-D-3 | DDL-D-3 |
| limitedness<br>reflexivity | DDL-D-3 | DDL-D-3 |
| limitedness<br>totalness | DDL-D-3 | DDL-D-3 |
| smoothness | DDL-D-4 | DDL-D-4 |
| smoothness<br>reflexivity | DDL-D-4 | DDL-D-4 |
| smoothness<br>totalness | DDL-D-4 | DDL-D-4 |
| smoothness<br>transitivity | DDL-D-4 | DDL-D-5 |
| smoothness<br>transitivity<br>reflexivity | DDL-D-4 | DDL-D-5 |
| smoothness<br>transitivity<br>totalness | DDL-D-5 | DDL-D-5 |

Table 4: Pure deontic conditional counterparts

## 4.5 Methods for proving completeness

Some remarks on the methods for proving the completeness part of the above determination results are in order. They will help the reader to get a feeling of what is involved.

### 4.5.1 Direct canonical model construction

All the proofs of completeness mentioned above are based on canonical models (see, for instance, [Chellas, 1980]). The proofs of completeness of **F**+(CM) and **G** in [Parent, 2008; Parent, 2014] use a direct canonical model construction. Adapting the canonical model technique to a preference-based setting is not as straightforward as might seem at first sight. Roughly speaking, the worlds in a canonical model are maximal consistent sets (MCSs) of sentences. The main difficulty is to define the comparative goodness relation in such a way that the semantic truth conditions for formulas starting with a deontic operator coincide with the set-membership relation between formulas and maximal consistent sets. In [Åqvist, 1987; Åqvist, 2002], the technique of so-called systematic frame constants is used to define the betterness relation part of the canonical model of **G**. Hansen [1999, p. 130] has shown that the method fails with respect to strong completeness.

For **F**+(CM) and **G**, one can think of suitable constructions. I start with **G**. The basic idea is to work with a point-generated canonical model. The set of all the MCSs is denoted by $\Omega$. Where $a$ is a MCS, $a^A$ denotes $\{B : \bigcirc(B/A) \in a\}$.

**Definition 4.10** (Canonical model, **G**). *Let $w$ be a fixed element of $\Omega$. The canonical model generated by $w$ is the structure $M^w = (W, \succeq, V)$ defined by*

*(i)* $W = \{a \in \Omega : \{A : \Box A \in w\} \subseteq a\}$

*(ii)* $a \succeq b$ *iff*

- *(a) there is no consistent wff $A$ such that $w^A \subseteq b$, or*
- *(b) there is some $A \in a \cap b$ such that $w^A \subseteq a$*

*(iii)* $v(p) = \{a \in W : p \in a\}$ *for all* $p \in \mathbb{P}$.

Condition (i) says that $W$ is the restriction of $\Omega$ to the set of MCSs containing all the wffs $A$ for which $\Box A$ is in the "generating" world $w$. This is needed to deal with the alethic modalities. The import of condition (ii) is that the best (according to $\succeq$) MCSs among those containing $A$ are precisely those containing all the wffs $B$ for which $\bigcirc(B/A)$ is in the "generating" world $w$.

The required construction for **F**+(CM) is more complex. The worlds in the universe of the canonical models are not just MCS's, but MCS's labeled with some suitable sentence. This is needed to rank them in terms of goodness. To be more precise, a world becomes a pair whose first element is a MCS $a$, and whose second element is some formula $A$

such that $w^A \subseteq a$, where $w$ is the MCS used to generate the canonical model. However, the method also demands that the selected MCS is part of the universe $W$ of the canonical model. Given a MCS $w$, there may not be any $A$ such that $w^A \subseteq w$. Due to this extra complication, one needs to distinguish between a principal case and a limiting case. I give the full details. For the sake of brevity, $A \geq B$ is used as a shorthand for $\bigcirc(A/A \vee B) \in w$, where $w$ is some MCS.

**Definition 4.11** (Canonical model, **F**+(CM), principal case)**.** *Let $w$ be a MCS such that $w^A \subseteq w$ for some $A$. The canonical model generated by $(w, A)$ is the structure $M^{(w,A)} = (W, \succeq, V)$ defined by*

*(i)* $W = \{(a, B) : a \in \Omega \;\&\; w^B \subseteq a\}$
*(ii)* $(a, B) \succeq (b, C)$ *iff: either* $C \not\geq B$ *or* $B \in b$
*(iii)* $v(p) = \{(a, B) \in W : p \in a\}$ *for all* $p \in \mathbb{P}$*.*

**Definition 4.12** (Canonical model, **F**+(CM), limiting case)**.** *Let $w$ be a MCS such that $w^A \subseteq w$ for no $A$. Take an arbitrarily chosen wff $A$. The canonical model generated by $(w, A)$ is the structure $M^{(w,A)} = (W, \succeq, V)$ defined by*

*(i)* $W = \widetilde{W} \cup \{(w, A)\}$*, where* $\widetilde{W} = \{(a, B) : a \in \Omega \;\&\; w^B \subseteq a\}$
*(ii)* $\succeq = \trianglerighteq \cup \{((w, A), (w, A))\} \cup \{(\alpha, (w, A)) : \alpha \in \widetilde{W}\}$ *where* $\trianglerighteq \subseteq \widetilde{W} \times \widetilde{W}$ *is defined as in Definition 4.11, putting* $(a, B) \trianglerighteq (b, C)$ *iff either* $C \not\geq B$ *or* $B \in b$
*(iii)* $v(p) = \{(a, B) \in W : p \in a\}$ *for all* $p \in \mathbb{P}$*.*

In [Parent, 2014] these two constructions are used to establish the completeness of **F**+(CM) with respect to the class of models in which $\succeq$ is opt-smooth (*resp.*, max-smooth), Theorem 4.2 (i), and with respect to the class of those in which $\succeq$ is also reflexive or total, Theorem 4.5 (i).

Under the max rule, **F**+(CM) is also sound and complete with respect to the class of models in which $\succeq$ is max-smooth and transitive, and with respect to the class of those in which $\succeq$ is also reflexive. This is Theorem 4.2 (ii) and Theorem 4.5 (ii). These results are new to the aforementioned paper. Their proof is given in Appendix B.

### 4.5.2 Completeness-via-selection-functions

Contrasting with this direct approach, the method used for **E** and **F** in [Parent, 2015] is indirect, and takes a detour through the alternative modeling in terms of selection functions described in Section 2.6. The proposed approach is related to the two-step methodology used by [Schlechta, 1997, chap. 2] when discussing representation problems for non-monotonic structures. There are two main steps.

The first step consists in showing soundness and completeness of the systems with respect to appropriate classes of selection function models. I state the needed results in the following theorem, which covers **F** +(CM) and **G** as well.

**Theorem 4.13.**

*(i)* **E** *is sound and complete with respect to the class of selection function models* $M = (W, \mathfrak{f}, v)$ *in which* $\mathfrak{f}$ *meets syntax-independence* ($\mathfrak{f}0$)*, inclusion* ($\mathfrak{f}1$) *and Chernoff* ($\mathfrak{f}2$)*;*

*(ii)* **F** *is sound and complete with respect to the class of selection function models* $M = (W, \mathfrak{f}, v)$ *in which* $\mathfrak{f}$ *meets in addition consistency-preservation* ($\mathfrak{f}3$)*;*

*(iii)* **F** *+(CM) is sound and complete with respect to the class of selection function models* $M = (W, \mathfrak{f}, v)$ *in which* $\mathfrak{f}$ *meets in addition Aizerman* ($\mathfrak{f}4$)*;*

*(iv)* **G** *is sound and complete with respect to the class of selection function models* $M = (W, \mathfrak{f}, v)$ *in which* $\mathfrak{f}$ *meets in addition Arrow* ($\mathfrak{f}5$)*.*

*Proof.* See [Åqvist, 2002, Theorem 77, p. 251]. Let $w$ be a fixed element of $\Omega$. Define the canonical model generated by $w$ as the model $M^w = (W, \mathfrak{f}, v)$ where

- $W = \{a \in \Omega : \{A : \Box A \in w\} \subseteq a\}$
- $\mathfrak{f}(A) = \{a \in \Omega : \{B : \bigcirc(B/A) \in w\} \subseteq a\}$
- $v(p) = \{a \in \Omega : p \in a\}$

Åqvist does not consider (CM). It is a straightforward matter to verify that, in the canonical model for **F**+(CM), $\mathfrak{f}$ meets Aizerman ($\mathfrak{f}4$). Details are omitted. $\square$

The second step consists in showing that the selection function semantics is equivalent with the preference-based semantics. One half of the equivalence is relatively easy to establish. This is Theorem 4.14 below. For the reason explained in Section 4, care should be taken with the Arrow condition. Under the max rule it calls for both transitivity and totalness of $\succeq$, while under the opt rule the constraint calls for transitivity only.

**Theorem 4.14.**

*(i) For every preference model* $M = (W, \succeq, v)$ *applying the opt rule, there is an equivalent selection function model* $M' = (W, \mathfrak{f}, v)$ *(with*

*$W$ and $v$ the same) in which $\mathfrak{f}$ meets syntax-independence* ($\mathfrak{f}0$), *inclusion* ($\mathfrak{f}1$) *and Chernoff* ($\mathfrak{f}2$). *If $\succeq$ is opt-limited, then $\mathfrak{f}$ meets consistency-preservation* ($\mathfrak{f}3$). *If $\succeq$ meets opt-smoothness, then $\mathfrak{f}$ meets Aizerman* ($\mathfrak{f}4$). *If $\succeq$ is transitive, then $\mathfrak{f}$ meets Arrow* ($\mathfrak{f}5$).

*(ii) For every preference model $M = (W, \succeq, v)$ applying the max rule, there is an equivalent selection function model $M' = (W, \mathfrak{f}, v)$ (with $W$ and $v$ the same) in which $\mathfrak{f}$ meets syntax-independence* ($\mathfrak{f}0$), *inclusion* ($\mathfrak{f}1$) *and Chernoff* ($\mathfrak{f}2$). *If $\succeq$ is max-limited, then $\mathfrak{f}$ meets consistency-preservation* ($\mathfrak{f}3$). *If $\succeq$ meets max-smoothness, then $\mathfrak{f}$ meets Aizerman* ($\mathfrak{f}4$). *If $\succeq$ is transitive and total, then $\mathfrak{f}$ meets Arrow* ($\mathfrak{f}5$).

*Proof.* For (i): starting with $M = (W, \succeq, v)$, define $M' = (W, \mathfrak{f}, v)$ by putting $\mathfrak{f}(A) = \mathrm{opt}_{\succeq}(\|A\|^M)$ for all wff $A$. For (ii): starting with $M = (W, \succeq, v)$, define $M' = (W, \mathfrak{f}, v)$ by putting $\mathfrak{f}(A) = \max_{\succeq}(\|A\|^M)$ for all wff $A$. $\square$

The hard part of the proof of equivalence is contained in the following theorem. This one extends a known result from rational choice theory (see, *e.g.*, [Herzberger, 1973]) to the case where the Arrow condition is no longer available.

**Theorem 4.15.** *For every selection function model $M = (W, \mathfrak{f}, v)$ in which $\mathfrak{f}$ meets syntax-independence, inclusion and Chernoff, there is a preference model $M' = (W', \succeq, v')$ such that, under the opt rule, $M'$ is equivalent to $M$. Furthermore, if $\mathfrak{f}$ meets consistency-preservation, then $\succeq$ is opt-limited.*

*Proof.* See [Parent, 2015, Theorem 3.5]. I recall the proposed construction. Let $M = (W, \mathfrak{f}, v)$. For all $a \in W$, define $\mathcal{Y}_a = \{\|C\|^M \subseteq W \mid a \in \|C\|^M - \mathfrak{f}(C)\}$. Let $\mathcal{Y}_a = \{X_i\}_{i \in I}$. Put $F_a := \prod_{i \in I} X_i$. Intuitively, $F_a$ is the (possibly infinite) cartesian product of all the sets in $\mathcal{Y}_a$. Formally, $F_a$ is the set of all the functions $g$ defined on the index set $I$ such that the value of the function $g$ at a particular index $i$ is an element of $X_i$:

$$\{g : I \to \bigcup_{i \in I} X_i \mid (\forall i \in I)(g(i) \in X_i)\}$$

The axiom of choice is assumed. Define $M' = (W', \succeq, v')$ as follows:

- $W' = \{\langle a, g\rangle \mid a, b \in W, g \in F_a\}$
- $\langle a, g\rangle \succeq \langle b, g'\rangle$ iff $b \notin \mathrm{Rng}(g))$
- $v'(p) = \{\langle a, g\rangle : a \in v(p)\}$

Rng($g$) denotes the range of $g$, *viz* $\{c \mid \langle i, c\rangle \in g \text{ for some } i \in I\}$. The verification that the construction above actually does the desired job proceeds via a series of lemmas, for which the reader is referred to the above paper. □

Combined with Theorem 4.13 (i) and (ii), Theorem 4.15 yields completeness of **E** with respect to the class of all preference models for the interpretation under the opt rule, and completeness of **F** with respect to the class of those where $\succeq$ is opt-limited under the same interpretation. These two core completeness results carry over to the class of models where $\succeq$ is also total or reflexive, and to the interpretation under the max rule. This is made possible by the following "bridge" result:

**Theorem 4.16.** *For every preference model $M = (W, \succeq, v)$ in which deontic formulas are interpreted under the opt-rule, there is an equivalent preference model $M = (W', \succeq', v')$ in which $\succeq'$ is total (and hence reflexive), and in which deontic formulas are interpreted under the max-rule (or, equivalently, the opt-rule). Furthermore, if $\succeq$ is opt-limited, then $\succeq'$ is max-limited.*

*Proof.* See [Parent, 2015, Theorem 3.3]. I recall the proposed construction. Starting with $M = (W, \succeq, v)$, one defines $M' = (W', \succeq', v')$ as follows:

- $W' = \{\langle a, n\rangle \mid a \in W, n \in \mathbb{N}\}$
- $\langle a, n\rangle \succeq' \langle b, m\rangle$ iff $a \succeq b$ or $n \geq m$
- $v'(p) = \{\langle a, n\rangle \mid a \in v(p)\}$

The universe in the output structure is the product set $W \times \mathbb{N}$. Thus, each world $a$ in $W$ has infinitely (albeit countably) many "duplicates" in $W'$. The order relation on the product set is the lexicographic ordering (or sort of). $\geq$ is total, and so is $\succeq'$. Equivalence between models follows from the fact that the set of optimal elements of $X \subseteq W$ in the input model "matches" the set of maximal elements of $X \times \mathbb{N}$ in the output model, in the sense that:

$$\mathrm{opt}_{\succeq}(X) \times \mathbb{N} = \mathrm{opt}_{\succeq'}(X \times \mathbb{N}) = \max_{\succeq'}(X \times \mathbb{N})$$

This suffices to establish the desired result. □

# 5 Decidability and automated theorem-proving

## 5.1 Decidability

The basic result we prove in this section is the decidability of the theoremhood problem "Is $A$ a theorem in such-and-such system?" This will be shown by establishing the so-called finite model property (FMP): any satisfiable formula is satisfiable in a finite model. To simplify matters, this property is shown to hold only with respect to models equipped with a selection function. Decidability of the theoremhood problem in **E**, **F**, **F**+(CM) and **G** follows in the usual way. (See [Chellas, 1980].) The FMP with respect to preference models is put aside.

The FMP with respect to selection function models may be established using the filtration method as adapted by Åqvist [1997; 2000] to a conditional logic setting. I will make a small change to one of his definitions in order to resolve a problem that was pointed out to me by Carmo [2009].

As usual, a model $M$ is said to be finite whenever its universe $W$ is finite. $\Gamma$ denotes a non-empty and finite set of sentences closed under sub-formulas. § stands for a designated propositional atom in $\Gamma$. Put $\top = \S \to \S$ and $\bot = \neg\top$.

For any selection function model $M = (W, \mathfrak{f}, v)$, the equivalence relation $\sim_\Gamma$ on $W$ is defined by setting

$$a \sim_\Gamma b \text{ iff for every } A \text{ in } \Gamma : a \vDash A \text{ iff } b \vDash A$$

Given $a \in W$, $[a]$ will be the equivalence class of $a$ under $\sim_\Gamma$.

**Definition 5.1.** *Given some $\Gamma$, define the translation function $\tau$, taking every wff into a wff whose propositional atoms are all in $\Gamma$, as follows:*

$$\tau(p) = \begin{cases} p & \text{if } p \in \Gamma \\ \S & \text{if } p \notin \Gamma \end{cases}$$

$$\tau(\neg A) = \neg\tau(A) \qquad \tau(A \vee B) = \tau(A) \vee \tau(B)$$
$$\tau(\Box A) = \Box\tau(A) \qquad \tau(\bigcirc(B/A)) = \bigcirc(\tau(B)/\tau(A))$$

Since neither $\top$ nor $\bot$ are primitive symbols, and $\Gamma$ is non-empty, there is always one such propositional atom § in $\Gamma$.

**Lemma 5.2.** *Let $\Gamma$, $\tau$ and $M$ be as above. Then, for all wffs $A$ and all $a, b \in W$, if $a \sim_\Gamma b$, then $a \vDash \tau(A)$ iff $b \vDash \tau(A)$.*

*Proof.* By induction on $A$. If $A = p$, then either (i) $p \in \Gamma$ or (ii) $p \notin \Gamma$. In case (i), $\tau(p) = p$. In case (ii), $\tau(p) = \S$. In both cases, the claim holds, because $a \sim_\Gamma b$. If $A = B \vee C$ or $A = \neg B$, the result follows directly from the inductive hypothesis. If $A = \Box B$ or $A = \bigcirc(C/B)$, the result follows directly from the evaluation rules for $\Box$ and for $\bigcirc(-/-)$. $\Box$

**Definition 5.3** (Filtration)**.** *The filtration of* $M = (W, \mathfrak{f}, v)$ *through* $\Gamma$ *is the model* $M^\star = (W^\star, \mathfrak{f}^\star, v^\star)$ *where:*

*(i)* $W^\star = \{[a] : a \in W\}$
*(ii)* $\mathfrak{f}^\star(A) = \{[a] : \exists b \in [a] \ \&\ b \in \mathfrak{f}(\tau(A))\}$
*(iii)* $v^\star(p) = \{[a] : a \in v(\tau(p))\}$ *for all* $p \in \mathbb{P}$.

**Fact 5.4.** *(i) If* $a \in W$ *then* $[a] \in W^\star$*; (ii)* $W^\star \neq \emptyset$.

*Proof.* (i) follows from the reflexivity of $\sim_\Gamma$ and Definition 5.3 (i). (ii) follows from (i) and $W \neq \emptyset$. $\Box$

**Fact 5.5.** $W^\star$ *is finite.*

*Proof.* The cardinality of $W^\star$ is at most $2^n$, where $n$ is the cardinality of $\Gamma$. $\Box$

A comment on $\mathfrak{f}^\star$ in Definition 5.3 is in order. It is easy to see that $\mathfrak{f}^\star$ is well-defined, in the sense that its definition does not depend upon any particular class representative. That is,

**Fact 5.6.** *If* $a \sim_\Gamma b$, *then* $[a] \in \mathfrak{f}^\star(A) \leftrightarrow [b] \in \mathfrak{f}^\star(A)$.

*Proof.* Assume $a \sim_\Gamma b$ and $[a] \in \mathfrak{f}^\star(A)$. It follows that $c \in \mathfrak{f}(\tau(A))$ for some $c \in [a]$. Since $a \sim_\Gamma b$, $c \in [b]$ too, and thus $[b] \in \mathfrak{f}^\star(A)$ as required. For the other direction, the argument is similar. $\Box$

Åqvist [1997; 2000] uses the following simpler definition:

$$\mathfrak{f}^\star(A) = \{[a] : a \in \mathfrak{f}(\tau(A))\} \tag{1}$$

Carmo [2009] points out that, if $\mathfrak{f}^\star$ is defined as in (1), then Fact 5.6 may fail as shown in the following example.

**Example 5.7.** *Put* $M = (W, \mathfrak{f}, v)$ *with* $W = \{a, b\}$, $v(p) = W$, *and* $\mathfrak{f}$ *such that*

$$\mathfrak{f}(A) = \begin{cases} \{a\} & \textit{if } a \vDash A \\ \|A\|^M & \textit{otherwise} \end{cases}$$

$\mathfrak{f}$ *meets syntax-independence, inclusion, Chernoff, consistency-preservation, Aizerman and Arrow. Let* $M^\star = (W^\star, \mathfrak{f}^\star, v^\star)$ *be the filtration of* $M$ *through* $\Gamma = \{p\}$*. We have* $a \sim_{\{p\}} b$*. Assume* $\mathfrak{f}^\star$ *is defined as in (1). We then have* $[a] \in \mathfrak{f}^\star(p)$ *and* $[b] \notin \mathfrak{f}^\star(p)$*. For* $\mathfrak{f}(\tau(p)) = \mathfrak{f}(p) = \{a\}$*.*

Clause (ii) of Definition 5.3 does not face the above problem. It remains to verify that the entire proof still goes through.

**Theorem 5.8** (Filtration Theorem)**.** *Let* $\Gamma$*,* $\tau$*,* $M$ *and* $M^\star$ *be as above. Then,*

*(i) For all wffs* $A$*, if* $A \in \Gamma$*, then* $\tau(A) = A$*.*
*(ii) For all wffs* $A$ *and all* $a \in W$*:*

$$M^\star, [a] \vDash A \text{ iff } M, a \vDash \tau(A).$$

*(iii) For all wffs* $A$ *in* $\Gamma$ *and all* $a \in W$*:*

$$M^\star, [a] \vDash A \text{ iff } M, a \vDash A.$$

*Proof.* (i) and (ii) are established by induction on $A$, using the relevant definitions. Clause (iii) is an immediate consequence of (i) and (ii).

For (i), the fact that $\Gamma$ is closed under subformulas allows us to apply the inductive hypothesis.

I give the full details for (ii) only, focusing on the cases where $A = \Box B$ and $A = \bigcirc(C/B)$, and assuming for induction that the theorem holds for $B$ and $C$.

- $A = \Box B$. From left-to-right, assume $[a] \vDash \Box B$. By the truth-conditions for $\Box$, we get $[b] \vDash B$ for all $[b]$ in $W^\star$. By the inductive hypothesis, $b \vDash \tau(B)$ for all $b$ in $W$. Hence $a \vDash \Box\tau(B)$. By definition of $\tau$, $a \vDash \tau(\Box B)$ as required. For the converse direction, argue in reverse.
- $A = \bigcirc(C/B)$. From left-to-right, assume $[a] \vDash \bigcirc(C/B)$, so that $\mathfrak{f}^\star(B) \subseteq \|C\|^{M^\star}$. By definition of $\tau$, to show that $a \vDash \tau(\bigcirc(C/B))$ amounts to showing that $a \vDash \bigcirc(\tau(C)/\tau(B))$. Let $b \in \mathfrak{f}(\tau(B))$. Since $b \in [b]$, $[b] \in \mathfrak{f}^\star(B)$, by Definition 5.3 (ii). Thus, $[b] \vDash C$. By the inductive hypothesis, $b \vDash \tau(C)$, which suffices for $a \vDash \bigcirc(\tau(C)/\tau(B))$. For the other direction, assume $a \vDash \tau(\bigcirc(C/B))$. By definition of $\tau$, $a \vDash \bigcirc(\tau(C)/\tau(B))$. Hence $\mathfrak{f}(\tau(B)) \subseteq \|\tau(C)\|^M$. Let $[b] \in \mathfrak{f}^\star(B)$. By Definition 5.3 (ii), there is some $c \in [b]$ such that $c \in \mathfrak{f}(\tau(B))$. So, $c \vDash \tau(C)$. By Lemma 5.2, $b \vDash \tau(C)$. By the inductive hypothesis, $[b] \vDash C$, which suffices for $[a] \vDash \bigcirc(C/B)$. $\Box$

**Theorem 5.9.** *Let $M^\star = (W^\star, \mathfrak{f}^\star, v^\star)$ be the filtration of $M = (W, \mathfrak{f}, v)$ through $\Gamma$. If $\mathfrak{f}$ meets syntax-independence, inclusion, Chernoff, consistency-preservation, Aizerman or Arrow, then so does $\mathfrak{f}^\star$.*

*Proof.* This is a matter of running through all the conditions, and showing that they are met.

Syntax-independence. Let $\|A\|^{M^\star} = \|B\|^{M^\star}$. By Theorem 5.8 (ii), $\|\tau(A)\|^M = \|\tau(B)\|^M$. Let $[a] \in \mathfrak{f}^\star(A)$. By Definition 5.3 (ii), $b \in \mathfrak{f}(\tau(A))$ for some $b \in [a]$. Since $\mathfrak{f}$ satisfies syntax-independence, $b \in \mathfrak{f}(\tau(B))$, and hence $[a] \in \mathfrak{f}^\star(B)$. For the other direction, the argument is similar.

Inclusion. Suppose that $[a] \in \mathfrak{f}^\star(A)$. By Definition 5.3 (ii), $b \in \mathfrak{f}(\tau(A))$ for some $b \in [a]$. Since $\mathfrak{f}$ satisfies inclusion, $b \vDash \tau(A)$. By Lemma 5.2, $a \vDash \tau(A)$. By Theorem 5.8 (ii), $[a] \vDash A$.

Chernoff. Suppose that $[a] \in \mathfrak{f}^\star(A) \cap \|B\|^{M^\star}$. By Definition 5.3 (ii), $b \in \mathfrak{f}(\tau(A))$ for some $b \in [a]$. By Theorem 5.8 (ii), $a \vDash \tau(B)$. By Lemma 5.2, $b \vDash \tau(B)$. So, $b \in \mathfrak{f}(\tau(A)) \cap \|\tau(B)\|^M$. Since $\mathfrak{f}$ satisfies Chernoff, $b \in \mathfrak{f}(\tau(A) \wedge \tau(B))$. By definition of $\tau$, $b \in \mathfrak{f}(\tau(A \wedge B))$. By Definition 5.3 (ii), $[a] \in \mathfrak{f}^\star(A \wedge B)$, as required.

Consistency-preservation. Assume $\|A\|^{M^\star} \neq \emptyset$. Hence, there is some $[a] \in W^\star$ such that $[a] \vDash A$. By Theorem 5.8 (ii), $a \vDash \tau(A)$. Since $\mathfrak{f}$ satisfies consistency-preservation, there is $b \in W$ such that $b \in \mathfrak{f}(\tau(A))$. But $b \in [b]$. By Definition 5.3 (ii), $[b] \in \mathfrak{f}^\star(A)$. Hence, $\mathfrak{f}^\star(A) \neq \emptyset$, as required.

Aizerman. Suppose $\mathfrak{f}^\star(A) \subseteq \|B\|^{M^\star}$ and $[a] \in \mathfrak{f}^\star(A \wedge B)$. We need to show that $[a] \in \mathfrak{f}^\star(A)$. By Definition 5.3 (ii) there is some $b \in [a]$ with $b \in \mathfrak{f}(\tau(A \wedge B))$. We show $\mathfrak{f}(\tau(A)) \subseteq \|\tau(B)\|^M$. Let $c \in \mathfrak{f}(\tau(A))$. Since $c \in [c]$, $[c] \in \mathfrak{f}^\star(A)$, Definition 5.3 (ii). By the opening hypothesis, $[c] \vDash B$. By Theorem 5.8 (ii), $c \vDash \tau(B)$, as required. Since $\mathfrak{f}$ satisfies Aizerman, $\mathfrak{f}(\tau(A \wedge B)) \subseteq \mathfrak{f}(\tau(A))$, and thus $b \in \mathfrak{f}(\tau(A))$, which suffices for $[a] \in \mathfrak{f}^\star(A)$, Definition 5.3 (ii).

Arrow. Let $\mathfrak{f}^\star(A) \cap \|B\|^{M^\star} \neq \emptyset$. To show: $\mathfrak{f}^\star(A \wedge B) \subseteq f^\star(A) \cap \|B\|^{M^\star}$. Let $[a] \in \mathfrak{f}^\star(A \wedge B)$. By Definition 5.3 (ii), there is some $b \in [a]$ with $b \in \mathfrak{f}(\tau(A \wedge B))$. By the opening hypothesis, there is some $[c] \in \mathfrak{f}^\star(A)$ with $[c] \vDash B$. By Definition 5.3 (ii), there is some $d \in [c]$ such that $d \in \mathfrak{f}(\tau(A))$. By Theorem 5.8 (ii), $c \vDash \tau(B)$. By Lemma 5.2, $d \vDash \tau(B)$. Hence, $\mathfrak{f}(\tau(A)) \cap \|\tau(B)\|^M \neq \emptyset$. Since $\mathfrak{f}$ meets Arrow, $\mathfrak{f}(\tau(A) \wedge \tau(B)) \subseteq \mathfrak{f}(\tau(A)) \cap \|\tau(B)\|^M$. By definition of $\tau$, $\mathfrak{f}(\tau(A \wedge B)) \subseteq \mathfrak{f}(\tau(A)) \cap \|\tau(B)\|^M$. Hence, $b \in \mathfrak{f}(\tau(A))$ and $b \in \|\tau(B)\|^M$. From the former, $[a] \in \mathfrak{f}^\star(A)$, Definition 5.3 (ii). From the latter, $a \in \|\tau(B)\|^M$, by Lemma 5.2. It follows that

$[a] \vDash B$, Theorem 5.8 (ii). Thus, $\mathfrak{f}^\star(A \wedge B) \subseteq \mathfrak{f}^\star(A) \cap \|B\|^{M^\star}$, as required. □

**Theorem 5.10.** *The FMP holds with respect to the following classes of selection function models:*

*(i) the class of those in which $\mathfrak{f}$ meets syntax-independence, inclusion and Chernoff;*
*(ii) the class of those in which $\mathfrak{f}$ meets syntax-independence, inclusion, Chernoff, and consistency-preservation;*
*(iii) the class of those in which $\mathfrak{f}$ meets syntax-independence, inclusion Chernoff, consistency-preservation, and Aizerman;*
*(iv) the class of those in which $\mathfrak{f}$ meets syntax-independence, inclusion, Chernoff, consistency-preservation, and Arrow.*

*Proof.* For (i). Suppose $A$ is satisfiable in some selection function model $M = (W, \mathfrak{f}, v)$ in which $\mathfrak{f}$ meets syntax-independence, inclusion and Chernoff. Thus, there is a world $a \in W$ such that $M, a \vDash A$. Consider the filtration $M^\star = (W^\star, \mathfrak{f}^\star, v^\star)$ of $M$ through the set $\Gamma$ of all the subformulas of $A$. By Fact 5.4 (i), $[a] \in W^\star$. By Fact 5.5, $W^\star$ is finite. By Theorem 5.9, $\mathfrak{f}^\star$ meets syntax-independence, inclusion and Chernoff. Furthermore, $A \in \Gamma$. By Theorem 5.8 (iii), $M^\star, [a] \vDash A$. Thus, $A$ is satisfiable in a finite model of the appropriate kind.

For (ii)-(iv), the argument is similar. Details are omitted. □

Since **E**, **F**, **F**+(CM), and **G** are finitely axiomatized, one gets the following spin-off result:

**Corollary 5.11.** *The theoremhood problem ("Is A a theorem?") in* **E**, **F**, **F***+(CM) and* **G** *is decidable.*

*Proof.* The argument is standard (see, *e.g.*, [Chellas, 1980]). □

The FMP w.r.t. selection functions is enough to establish the decidability of the theoremhood problem. The question of whether the FMP also holds w.r.t. preference models has an interest in its own right. It is left as a topic for future research.

## 5.2 Automated theorem proving

This section describes work by [Benzmüller *et al.*, 2019] in automated theorem proving (ATP) for the family of logics discussed in this chapter. Readers who are not interested in automated reasoning can skip this section and go to Section 6.

A specific method called Shallow Semantical Embedding (SSE) is used. The key idea is to use classical higher-order logic (HOL), *i.e.*, Church's type theory [Benzmüller and Andrews, 2019], as a meta-logic in order to represent and model the syntactic and semantic elements of a specific source logic. One can then use off-the-shelf HOL theorem-provers like Isabelle/HOL [Nipkow *et al.*, 2002] or Leo-III [Steen and Benzmüller, 2018; Steen, 2018] for automation. The method was successfully applied to a wide range of non-classical and modal logics—for an overview, see [Benzmüller, 2019] and the references therein. The scope of application of the method has recently been extended to include various prominent deontic logics, including Åqvist's system **E**.[17] The authors focus on the case where deontic formulas are interpreted using the opt rule. It is a straightforward matter to extend the approach to the case where they are interpreted using the max rule, or even an evaluation rule other than one in terms of best, like one of those discussed in Section 6.

In this section I will only briefly describe this work, omitting most of the formal details and proofs, which can be found in the aforementioned paper. The method can be seen as a variant of the so-called standard translation from modal logic to first-order logic [Blackburn *et al.*, 2001]. Possible worlds and propositional letters become individuals and unary predicates, respectively. A distinguished binary predicate symbol $r$ is added to the language of HOL to represent the betterness relation. The modalities are handled by explicit quantification over the set of individuals. One "mimics" the evaluation rules used when evaluating the truth of formulas in a preference model. For example, $\Box A$ translates into:

$$\lambda x.\forall y.Ay$$

And $\bigcirc(B/A)$ translates into:

$$\lambda x\ (\forall y(Ay \land (\forall z(Az \to ryz)) \to By))$$

This translation holds for the interpretation under the opt rule.

On the HOL side, the following two primitive types are used: $i$ for individuals (or possible worlds); $o$ for the Boolean values. A variant of the standard semantics is used. It is called "generalized" or (after its inventor) "Henkin" semantics. This variant semantics leads to an axiomatizable version of higher-order logic, because the set of functions in a given model need not be complete. (See [Henkin, 1950].)

[17] This is part of a larger project, which aims at mechanizing and automating deontic reasoning. The study [Benzmüller *et al.*, 2020] gives an overview of the project, and documents further the other deontic frameworks covered so far by the SSE method. The Isabelle/HOL theory files are available at `www.logikey.org`.

When working out the formal details, there are three main steps to follow. The first step is to specify an embedding $\lceil\cdot\rceil$, which translates a formula $A$ of **E** into a term $\lceil A\rceil$ of HOL. As mentioned, the clauses of the definition of $\lceil\cdot\rceil$ mirror the evaluation rules used in the semantics of **E**. The second step is to establish that the embedding is sound and complete, that is faithful, in the sense that it preserves both the validity and invalidity of formulas. The establishment of such a result is the main criterion of success. This is Theorem 5.12 below. Intuitively it tells us that under the opt rule a formula $A$ in the language of **E** is valid in the class of all preference models (notation: $\models A$) if and only if the HOL formula $\forall x.\lceil A\rceil x$ is a tautology in every Henkin model (notation: $\models_{\text{HOL}} \forall x.\lceil A\rceil x$).

**Theorem 5.12** (Faithfulness of the embedding, [Benzmüller *et al.*, 2019])**.**

$$\models A \textit{ if and only if } \models_{\text{HOL}} \forall x.\lceil A\rceil x$$

*Proof.* This is [Benzmüller *et al.*, 2019, Theorem 2]. The crux of the argument consists in relating preference models with Henkin models in a truth-preserving way. □

The third and last step consists in encoding the embedding in a concrete theorem-prover like Isabelle/HOL [Nipkow *et al.*, 2002]. Figure 7 displays the encoding obtained for **E**. Some explanations are in order. On line 5, a designated constant "aw" for the actual world is introduced. On lines 28–31, this constant is used to distinguish between global validity (*i.e.*, truth in all worlds) and local validity (*i.e.*, truth at the actual world). On lines 19–26, the dyadic deontic operators are defined by introducing first the notion of optimal $A$-world.

Here is a (non-exhaustive) list of queries that can be run:

- *Satisfiability:* Is the (finite) set $\Gamma$ of formulas satisfiable?
- *Validity*: Is formula $A$ valid? Does inference rule $R$ preserve global validity?
- *Entailment:* Does $A$ follow from $\Gamma$ (with $\Gamma$ finite)?
- *Correspondance:* Is such-and-such property of the betterness relation sufficient to validate $A$? Is such-and-such property of the betterness relation necessary to validate $A$?

When the answer is "no" the model finder Nitpick [Blanchette and Nipkow, 2010] is able to give a counter-example. Similarly, when a formula (or a set of formulas) is satisfiable, Nitpick is able to give a model and a world satisfying the formula (or set of formulas) in question.

```
theory DDLE imports Main
begin
typedecl i (* type for possible worlds *)
type_synonym τ = "(i⇒bool)" (* type for propositions *)
consts aw::i (* actual world *)
consts r :: "i⇒τ"  (infixr "r" 70) (* comparative goodness relation *)
                                       (* Boolean connectives *)
definition ddetop    :: "τ" ("⊤")               where "⊤  ≡ λw. True"
definition ddebot    :: "τ" ("⊥")               where "⊥  ≡ λw. False"
definition ddeneg    :: "τ⇒τ" ("¬_"[52]53)      where "¬A ≡ λw. ¬A(w)"
definition ddeand    :: "τ⇒τ⇒τ" (infixr"∧"51) where "A∧B ≡ λw. A(w)∧B(w)"
definition ddeor     :: "τ⇒τ⇒τ" (infixr"∨"50) where "A∨B ≡ λw. A(w)∨B(w)"
definition ddeimp    :: "τ⇒τ⇒τ" (infixr"→"49) where "A→B ≡ λw. A(w)⟶B(w)"
definition ddeequivt :: "τ⇒τ⇒τ" (infixr"↔"48) where "A↔B ≡ λw. A(w)⟷B(w)"
                                      (* alethic operators *)
definition ddebox  :: "τ⇒τ" ("□") where "□ ≡ λA w.  ∀v. A(v)"
definition ddediomond  :: "τ⇒τ" ("◇") where "◇ ≡ λA w.  ∃v. A(v)"

definition ddeopt :: "τ⇒τ" ("opt<_>") (* deontic operators *)
  where "opt<A> ≡ (λv. ( (A)(v) ∧ (∀x. ((A)(x)  ⟶  v r x)) ) )"
abbreviation(input) msubset :: "τ⇒τ⇒bool" (infix "⊆" 53)
  where "A ⊆ B ≡ ∀x. A x ⟶ B x"
definition ddecond :: "τ⇒τ⇒τ" ("○<_|_>")
  where "○<B|A> ≡ λw. opt<A> ⊆ B"
definition ddeperm :: "τ⇒τ⇒τ" ("P<_|_>")
  where "P<B|A> ≡ ¬○<¬B|A>"

definition ddevalid :: "τ⇒bool" ("⌊_⌋"[8]109) (* global validity *)
  where "⌊p⌋ ≡ ∀w. p w"
definition ddeactual :: "τ⇒bool" ("⌊_⌋l"[7]105) (* local validity *)
  where "⌊p⌋l ≡ p(aw)"

end
```

Figure 7: Encoding of system **E** in Isabelle/HOL.

Theorem provers for KLM-style nonmonotonic and conditional logics have been developed, like, e.g., KLMLean 1.0 [Olivetti and Pozzato, 2005], KLM 2.0 [Giordano *et al.*, 2007] and Nescond [Olivetti and Pozzato, 2014]. It would be interesting to compare them with the one described here.

# 6 Alternative truth-conditions

Despite its length, the chapter does not purport to give an encyclopedic coverage of the field. In this section, I discuss two variant truth-conditions for the conditional obligation operator. As mentioned in the introduction section, more variations are possible. For details, the readers are referred to [Makinson, 1993; Goble, 2015] and references therein.

## 6.1 The Danielsson-van Fraassen-Lewis truth-conditions

These truth-conditions for deontic sentences are named by Åqvist [1987, p. 199] after their co-inventors: Danielsson [1968], van Fraassen [1972] and Lewis [1973]. One counts $\bigcirc(B/A)$ as true in a world $a$ whenever either there are no $A$-worlds, or there is some $A \wedge B$-world $b$ such that, as we go up in the ordering, the material implication $A \to B$ always holds. Hence, all worlds ranked as high as $b$ comply with the obligation in question. This evaluation rule is also used by van Kutschera [1974], Loewer and Belzer [1983] and Goble [2003], among others.

**Definition 6.1** ($\exists\forall$ rule)**.** *Given a preference model $M = (W, \succeq, V)$ and a world $a \in W$, we have*

$$M, a \vDash \bigcirc(B/A) \textit{ iff } \neg\exists b \ \ (b \models A) \textit{ or} \\ \exists b \ \ (b \models A \wedge B \ \ \& \ \ \forall c\,(c \succeq b \Rightarrow c \models A \to B)) \tag{$\exists\forall$}$$

I shall refer to the statement appearing at the right-hand-side of "iff" as the $\exists\forall$ rule. Lewis's preference for the $\exists\forall$ rule is based on his rejection of the limit assumption [1973, p. 98]. The $\exists\forall$ rule handles infinitely ascending chains better than the Hanssonian-type rule in terms of best worlds. Indeed when the $A$-worlds form an infinitely ascending chain (so that there is no best $A$-world) under the second rule the formula $\bigcirc(B/A)$ (where $B$ is an arbitrarily chosen formula) becomes (vacuously) true. Thus, when the limit assumption fails, everything is obligatory. With the $\exists\forall$ rule, this is not the case.[18]

Leaving the above issue aside, I now clarify how the $\exists\forall$ rule relates with the opt rule and the max rule.

**Theorem 6.2.**

*(i) The $\exists\forall$ rule implies the opt rule;*

*(ii) Given totalness of $\succeq$, the $\exists\forall$ rule implies the max rule.*

*Proof.* (ii) follows from (i). To show (i), suppose $\bigcirc(B/A)$ holds at world $a$ in virtue of the $\exists\forall$ rule. This means that either $\neg\exists b \ \ (b \models A)$ or $\exists b \ \ (b \models A \wedge B \ \ \& \ \ \forall c\,(c \succeq b \Rightarrow c \models A \to B))$. In the first case, we have $\text{opt}_{\succeq}(\|A\|) = \emptyset$, and so $\text{opt}_{\succeq}(\|A\|) \subseteq \|B\|$. In the second case, consider some $d \in \text{opt}_{\succeq}(\|A\|)$. We have $d \succeq b$ and $d \models A$. So $d \models B$, which suffices for $\text{opt}_{\succeq}(\|A\|) \subseteq \|B\|$ as required. $\square$

[18] Goble's own motivation for using the $\exists\forall$ rule is different. It is not directly related to the limit assumption but to the wish to accommodate conflicts between obligations (see *infra*).

**Theorem 6.3.**

*(i) Given transitivity and opt-limitedness of $\succeq$, the opt rule implies the $\exists\forall$ rule;*

*(ii) Given transitivity and max-limitedness of $\succeq$, the max rule implies the $\exists\forall$ rule.*

*Proof.* For (i), assume $\mathrm{opt}_{\succeq}(\|A\|) \subseteq \|B\|$. Either (a) $\mathrm{opt}_{\succeq}(\|A\|) = \emptyset$, or (b) $\mathrm{opt}_{\succeq}(\|A\|) \neq \emptyset$. In case (a), by opt-limitedness, $\|A\| = \emptyset$, and so the $\exists\forall$ rule is verified. In case (b), there is some $b$ such that $b \in \mathrm{opt}_{\succeq}(\|A\|)$. We have $b \models B$, by the opening assumption. Let $c$ be such that $c \succeq b$ and $c \models A$. Consider any $d$ such that $d \models A$. We have $b \succeq d$. By transitivity, we then get $c \succeq d$, so that $c \in \mathrm{opt}_{\succeq}(\|A\|)$, and hence $c \models B$, by the opening assumption. Thus, the $\exists\forall$ rule is verified too.

For (ii), the argument is similar. □

The question arises as to how to axiomatize the set of valid formulas for the interpretation under the $\exists\forall$ rule. This question was resolved very early by Lewis and van Fraassen in the case of total orders. Below I recast their result in terms of the systems studied in this chapter. As with Lewis's and van Fraassen's settings, the limit assumption has no impact.

**Theorem 6.4.** *Under the $\exists\forall$ rule,* **G** *is sound with respect to:*

*(i) the class of models in which $\succeq$ is transitive and total (and hence reflexive); and*

*(ii) the class of models in which $\succeq$ is transitive, total and opt/max-limited (or opt/max-smooth).*

*Proof.* In the presence of transitivity and totalness, opt-limitedness, max-limitedness, opt-smoothness and max-smoothness coincide. All that is required is to show that each axiom of **G** is valid in the class of models in which $\succeq$ is transitive and total, and that the inference rules of **G** preserve validity in this class of models. The argument is routine, and left to the reader. The arguments for (Abs), (Nec), (Ext), (Id) and (Sh) do not call for any of the properties of $\succeq$. (D$^\star$) calls for totalness. (Sp) calls for transitivity. (COK) and (CM) call for both totalness and transitivity. For the reader's convenience, I recap these points in the form of a table, Table 5. □

Completeness can be derived from the completeness of **G** under the interpretation applying the opt rule, with respect to the class of models in which $\succeq$ is transitive, total and opt-limited.

| Axiom of **G** | Property (or pair of properties) of $\succeq$ |
|---|---|
| (D$^\star$) | totalness |
| (Sp) | transitivity |
| (COK) | transitivity and totalness |
| (CM) | transitivity and totalness |

Table 5: Axioms and properties under the $\exists\forall$ rule

**Theorem 6.5.** *Under the $\exists\forall$ rule,* **G** *is complete with respect to:*
*(i) the class of models in which $\succeq$ is transitive and total; and*
*(ii) the class of models in which $\succeq$ is transitive, total and opt-limited (*resp. *max-limited, opt-smooth and max-smooth).*

*Proof.* Suppose that $\Gamma \nvdash_{\mathbf{G}} A$. By completeness under the opt rule with respect to the class of models in which $\succeq$ is transitive, total and opt-limited, $\Gamma \not\models A$ over that class of models. By Theorems 6.2 and 6.3, under the $\exists\forall$ rule $\Gamma \not\models A$ over the class of models in which $\succeq$ is transitive, total and opt-limited. Given transitivity and totalness, opt-limitedness, max-limitedness, opt-smoothness and max-smoothness coincide. This establishes (ii). Deleting a constraint on $\succeq$ does not increase the set of semantical consequences. This establishes (i). $\square$

Goble [2003] must be given credit for providing an axiomatization called **DP** in the case of partial orders. In the absence of totalness, (D$^\star$), which rules out the possibility of conflicting obligations, goes away. The choice of partial orders may thus be motivated by the need to accommodate conflicts between obligations, these being commonplace.[19] Note that (COK) and (CM) also go away while (Sp) remains. **DP** is a "pure" deontic logic: its language has no other primitive modal operator than $\bigcirc(-/-)$. Furthermore, its semantics uses a betterness relation relativized to worlds, and the truth-conditions make the obligation false when the antecedent is impossible. The proof of completeness for **DP** given by Goble takes a detour through an alternative semantics in terms of multiple preference models. The question as to whether the proof of com-

[19]Here lies Goble's reason for using the $\exists\forall$ rule. With the Hanssonian sort of interpretation, one needs to work with models without the limit assumption ; such models correspond to system **E**. However, **E** contains the following principle of "deontic explosion", $\bigcirc(B/A) \wedge \bigcirc(\neg B/A) \to \bigcirc(C/A)$, which says that if there is any instance of a deontic dilemma then everything is obligatory. (This is similar to the point made above in relation to the limit assumption, page 49). A survey of the state of the art regarding the treatment of conflicts between obligations may be found in [Goble, 2013].

pleteness for **DP** can be adapted to the present setting is left as a topic for future research. Furthermore, one would like to know what happens within this set-up when transitivity goes away. The question of how to axiomatize the corresponding logic is left as a topic for future research too.

### 6.2 The Burgess-Boutilier-Lamarre truth-conditions

The evaluation rule used by Burgess [1981], Boutilier [1994] and others has a "$\forall\exists\forall$" structure. This alternative evaluation rule has two technical attractions. First, as noted by Boutilier and independently by Lamarre [1991], it permits the reduction of the dyadic obligation operator to a monadic modal operator. Second, as mentioned by Lewis [1981, p. 230], it enables one to have a fairly strong dyadic deontic logic without committing to either a form of the limit assumption or totalness for $\succeq$. Makinson [1993, p. 346] gives a similar motivation. We see a similar rule in the Kratzer semantics for conditionals (see Kratzer [1991, Definition 13]) and in Veltman [1985]'s logic for conditionals.

**Definition 6.6** ($\forall\exists\forall$ rule)**.** *Given a preference model $M$, and some world $a$ in $M$, we have*

$$
\begin{aligned}
a \vDash \bigcirc(B/A) \textit{ iff } \forall b \textit{ if } b \models A \textit{ then}& \\
\exists c \textit{ s.t. } c \succeq b \;\&\; c \models A \;\&& \\
\forall d\, (d \succeq c \;\Rightarrow\; d \models A \to B)&
\end{aligned}
\tag{$\forall\exists\forall$}
$$

I will refer to the statement at the right-hand side of "iff" as the $\forall\exists\forall$ rule. I just described this rule as a way to avoid commitment to totalness for $\succeq$. This was Lewis's primary motivation. (See also [Kaufmann, 2017, §3].) It is worth mentioning that this benefit comes with a cost: (RM) goes away, while ($D^\star$) remains. The argument for ($D^\star$) is part of the proof of Theorem 6.10 below. I show the failure of (RM).

**Observation 6.7.** *There is a preference model $M = (W, \succeq, v)$, with $\succeq$ reflexive and transitive, in which (RM) fails under the $\forall\exists\forall$ rule.*

*Proof.* Put $M = (W, \succeq, v)$, with $W = \{a, b, c\}$, $\succeq$ the reflexive closure of $\{(b, a), (c, c)\}$ and $v(p) = W$, $v(q) = \{b, c\}$ and $v(r) = \{a, c\}$. This is shown in Figure 8, where reflexivity is left implicit. In this model, $\succeq$ is (vacuously) transitive. We have:

- $a \models \bigcirc(q/p)$
- $a \models \neg \bigcirc (\neg r/p)$ (witness: $c$)
- $a \not\models \bigcirc(q/p \wedge r)$ (witness: $a$) □

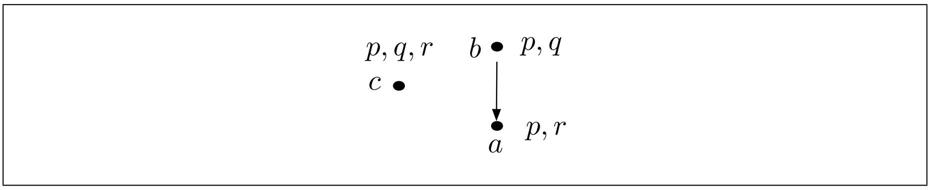



Figure 8: A countermodel to (RM) under the ∀∃∀ rule

Theorem 6.8 clarifies how the ∀∃∀ rule relates with the ∃∀ rule.

**Theorem 6.8.**
*(i) Given reflexivity of $\succeq$, the ∀∃∀ rule implies the ∃∀ rule;*
*(ii) Given both transitivity and totalness of $\succeq$, the ∃∀ rule implies the ∀∃∀ rule.*

*Proof.* For (i), suppose the ∀∃∀ rule holds, but not the ∃∀ rule. Hence, there is some $b_1$ such that $b_1 \models A$ and

$$\forall b \ (b \models A \wedge B \ \Rightarrow \ \exists c \, (c \succeq b \ \& \ c \models A \ \& \ c \not\models B)) \qquad (\alpha_1)$$

By the ∀∃∀ rule, there is some $c_1$ such that $c_1 \succeq b_1$, $c_1 \models A$ and

$$\forall d \, (d \succeq c_1 \ \Rightarrow \ d \models A \rightarrow B) \qquad (\alpha_2)$$

By reflexivity, $c_1 \succeq c_1$, and so $c_1 \models B$. By $(\alpha_1)$, there is some $d_1$ such that $d_1 \succeq c_1$, $d_1 \models A$ and $d_1 \not\models B$. This contradicts $(\alpha_2)$.

For (ii), suppose the ∃∀ rule holds, but not the ∀∃∀ rule. From the latter, there is some $b_1$ such that $b_1 \models A$ and

$$\forall c \, (c \succeq b_1 \,\&\, c \models A \ \Rightarrow \ \exists d \, (d \succeq c \,\&\, d \models A \,\&\, d \not\models B)) \qquad (\beta_1)$$

For the ∃∀ rule to hold, it must be the case that there is some $b_2$ such that $b_2 \models A \wedge B$ and

$$\forall c \, (c \succeq b_2 \ \Rightarrow \ c \models A \rightarrow B) \qquad (\beta_2)$$

By totalness, either (a) $b_1 \succeq b_2$ or (b) $b_2 \succeq b_1$. In case (a), $(\beta_2)$ yields $b_1 \models A \rightarrow B$. By reflexivity of $\succeq$, $b_1 \succeq b_1$. By $(\beta_1)$, there is some $d_1$ such that $d_1 \succeq b_1$, $d_1 \models A$ and $d_1 \not\models B$. By transitivity, $d_1 \succeq b_2$, and so by $(\beta_2)$, $d_1 \models A \rightarrow B$, a contradiction. In case (b), $(\beta_1)$ yields that there is some $d_1$ such that $d_1 \succeq b_2$ and $d_1 \models A$ and $d_1 \not\models B$, a result that immediately contradicts $(\beta_2)$. □

It is noteworthy that, in the presence of the limit assumption, the ∀∃∀ rule coincides with the max rule.

**Theorem 6.9.**
*(i) The* $\forall\exists\forall$ *rule implies the max rule;*
*(ii) Given reflexivity, transitivity and max-smoothness of* $\succeq$*, the max rule implies the* $\forall\exists\forall$ *rule.*

*Proof.* For (i), suppose the $\forall\exists\forall$ rule holds, and let $b \in \max_{\succeq}(\|A\|)$. Since $b \models A$, there is some $c$ such that $c \succeq b$, $c \models A$ and

$$\forall d\,(d \succeq c \;\Rightarrow\; d \models A \to B) \tag{$\gamma_1$}$$

Since $b \in \max_{\succeq}(\|A\|)$, $b \succeq c$. ($\gamma_1$) then yields $b \models B$, which suffices for $\max_{\succeq}(\|A\|) \subseteq \|B\|$.

For (ii), suppose the max rule holds, and let $b$ be such that $b \models A$. By max-smoothness either (a) $b \in \max_{\succeq}(\|A\|)$ or (b) there is $c$ such that $c \succ b$ and $c \in \max_{\succeq}(\|A\|)$. Suppose (a) applies. By reflexivity, $b \succeq b$. Also $b \models A$. Let $c$ be such that $c \succeq b$ and $c \models A$. Let $d$ be such that $d \succeq c$ and $d \models A$. By transitivity of $\succeq$, $d \succeq b$. By maximality of $b$, $b \succeq d$. By transitivity of $\succeq$ again, $c \succeq d$. Hence, $c \in \max_{\succeq}(\|A\|)$. It then follows that $c \models B$ as required. The argument for (b) is similar, working with $c$ instead of $b$. □

**Theorem 6.10.** *Under the* $\forall\exists\forall$ *rule,* **F**+*(CM) is sound with respect to the class of models in which* $\succeq$ *is reflexive and transitive.*

*Proof.* This is just a matter of verifying that the axioms of **F**+(CM) are valid. (Ext) and (Abs) hold independently of the reflexivity and transitivity of $\succeq$. (Nec), (Id) and (D$^\star$) each call for the reflexivity of $\succeq$. (CM) and (COK) call for transitivity of $\succeq$, while (Sh) calls for both transitivity and reflexivity. For the reader's convenience, I recap these points in the form of a table, Table 6. I give the argument for (D$^\star$)and (CM) only.

For (D$^\star$), suppose (i) $a \models \Diamond A$ and (ii) $a \models \bigcirc(B/A)$. To show: $a \models P(B/A)$, *i.e.*, $a \not\models \bigcirc(\neg B/A)$. From (i), there is some $b$ be such that $b \models A$. Let $c$ be such that $c \succeq b$ and $c \models A$. From (ii), there is some $d \succeq c$ such that $d \models A$ and

$$\forall e\,(e \succeq d \;\Rightarrow\; e \models A \to B) \tag{$\delta_1$}$$

By reflexivity, $d \succeq d$, and hence by ($\delta_1$) $d \models B$, *i.e.*, $d \not\models \neg B$. Hence, $a \not\models \bigcirc(\neg B/A)$ as required.

For (CM), suppose (i) $a \models \bigcirc(B/A)$ and (ii) $a \models \bigcirc(C/A)$. Let $b_1$ be such that $b_1 \models A \wedge B$. By (i), there is some $b_2 \succeq b_1$ such that $b_2 \models A$ and

$$\forall c\,(c \succeq b_2 \;\Rightarrow\; c \models A \to B) \tag{$\delta_2$}$$

By (ii), there is some $b_3 \succeq b_2$ such that $b_3 \models A$ and

$$\forall c\,(c \succeq b_3 \;\Rightarrow\; c \models A \to C) \tag{$\delta_3$}$$

By ($\delta_2$), $b_3 \models B$ and hence $b_3 \models A \wedge B$. By transitivity of $\succeq$, $b_3 \succeq b_1$. Let $d$ be such that $d \succeq b_3$ and $d \models A \wedge B$. Obviously, $d \models A$. By ($\delta_3$), $d \models C$, which suffices for $a \models \bigcirc(C/A \wedge B)$. □

| Axiom of **F**+(CM) | Property (or pair of properties) of $\succeq$ |
|---|---|
| (Nec) | reflexivity |
| (Id) | reflexivity |
| (D$^\star$) | reflexivity |
| (CM) | transitivity |
| (COK) | transitivity |
| (Sh) | reflexivity and transitivity |

Table 6: Axioms and properties under the $\forall\exists\forall$ rule

**Theorem 6.11.** *Under the $\forall\exists\forall$ rule,* **F**+*(CM) is complete with respect to the class of models in which $\succeq$ is reflexive and transitive.*

*Proof.* Suppose $\Gamma \nvdash_{\mathbf{F}+(CM)} A$. By Theorem 4.5 (ii), for the interpretation under the max rule we have that $\Gamma \not\models A$ over the class of models in which $\succeq$ is reflexive, transitive and max-smooth. By Theorem 6.9, the observation that $\Gamma \not\models A$ over the class of models in which $\succeq$ is reflexive, transitive and max-smooth carries over to the interpretation under the $\forall\exists\forall$ rule. That $\Gamma \not\models A$ continues to apply, *mutatis mutandis*, with respect to the class of models in which $\succeq$ is only reflexive and transitive. □

As with the $\exists\forall$ rule, the limit assumption has no impact.

**Corollary 6.12.** *Under the $\forall\exists\forall$ rule,* **F**+*(CM) is sound and complete with respect to the class of models in which $\succeq$ is reflexive, transitive and max-smooth (*resp. *max-limited).*

*Proof.* Soundness follows from the fact that no axiom requires max-smoothness or max-limitedness. Completeness with respect to the class of models with max-smoothness has just been established as part of the proof of Theorem 6.11. Completeness with respect to the class of models with max-limitedness follows from this and Observation 2.8 (a) (i). □

It should be pointed out that Theorem 6.11 echoes the axiomatization result obtained by Goble [2014] for the Kratzer conditional.

I end this section by showing that the assumption of totalness boosts the logic from **F**+(CM) to **G**.

**Theorem 6.13.** *Under the* $\forall\exists\forall$ *rule,* **G** *is sound and complete with respect to:*

*(i) the class of models in which* $\succeq$ *is transitive and total (and hence reflexive); and*

*(ii) the class of models in which* $\succeq$ *is transitive, total and max-limited (*resp. *max-smooth, opt-limited and opt-smooth).*

*Proof.* For soundness, it suffices to verify that (Sp) holds is valid when $\succeq$ is required to be total. Consider a model $M$ and a world $a$ in $M$ such that (i) $a \models P(B/A)$, (ii) $a \models \bigcirc(B \to C/A)$ and (iii) $a \not\models \bigcirc(C/A \wedge B)$. From (iii), there is some $b_1$ such that $b_1 \models A \wedge B$ and

$$\forall c\ ((c \succeq b_1\ \&\ c \models A \wedge B) \Rightarrow \exists d\ (d \succeq c\ \&\ d \models A \wedge B\ \&\ d \not\models C)) \quad (\epsilon_1)$$

From (ii), there is some $b_2 \succeq b_1$ with $b_2 \models A$ and

$$\forall c\ (c \succeq b_2 \Rightarrow c \models A \to (B \to C)) \quad (\epsilon_2)$$

From (i), there is some $b_3$ such that $b_3 \models A$ and

$$\forall c\ ((c \succeq b_3\ \&\ c \models A) \Rightarrow \exists d\ (d \succeq c\ \&\ d \models A \wedge B)) \quad (\epsilon_3)$$

By totalness, either (1) $b_2 \succeq b_3$ or (2) $b_3 \succeq b_2$. We argue that, in both cases, there is some $b_4$ with $b_4 \succeq b_2$ and $b_4 \models A \wedge B$. In case (1), ($\epsilon_3$) immediately yields this result. In case (2), $b_3 \succeq b_3$ by reflexivity, and so ($\epsilon_3$) tells us that there is some $b_4$ with $b_4 \succeq b_3$ and $b_4 \models A \wedge B$. By transitivity, $b_4 \succeq b_2$. Thus, either way, there is some $b_4$ with $b_4 \succeq b_2$ and $b_4 \models A \wedge B$. By transitivity, $b_4 \succeq b_1$. ($\epsilon_1$) then yields that there is some $b_5$ with $b_5 \succeq b_4$ and $b_5 \models A \wedge B \wedge \neg C$. This contradicts ($\epsilon_2$), since $b_5 \succeq b_2$, by transitivity.

Completeness follows at once from Theorems 6.5 and 6.8. □

# 7 Conclusion

The chapter has provided a survey of results related to the meta-theory of dyadic deontic logics in Hansson's tradition, focusing on axiomatization issues. The goal was to provide a "roadmap" of the different systems that can be obtained, depending on the special properties envisaged for the betterness relation, and depending on whether "best" means "optimal" or "maximal". Four systems of increasing strength were discussed, and related to (combinations of) properties of the betterness relation. The most remarkable finding in this study is that the contrast between the two notions of "best" is not as significant as one may think,[20] because in an appreciable number of cases the determined logic remains the same no matter which definition is used. Another unexpected outcome is that an apparently strong condition like totalness (and also, sometimes, transitivity) is somewhat idle, because in quite a number of cases its imposition does not affect the logic.

At least two qualifications of these findings are worth noting. First, we have noticed an asymmetry between maximality and optimality in two cases, when transitivity interacts with totalness (and smoothness), and when transitivity is considered alone. The latter case is not fully understood yet because no completeness result for optimality has been reported. Second, the correlations between the properties of the betterness relation and the axioms are not the same when variant truth-conditions for the conditional are used in order to circumvent the limit assumption. Two such variant truth-conditions are the $\exists\forall$ rule and the $\forall\exists\forall$ rule. Under the former a completeness theorem is available for models with a transitive and total relation, and under the latter for models with a reflexive and transitive relation. But we still do not know the full picture. In particular it is not known what happens when transitivity goes away.

For the sake of exhaustiveness, decidability of the theoremhood problem and automated theorem-proving were also discussed. The decidability of the theoremhood problem in the four proof systems studied in this chapter was established, by taking a detour through a modeling in terms of a selection function. Reasoning tasks were automated via a faithful embedding into HOL. These topics have an interest in their own right. However no deeper insight on the above issues was gained. Looking at computational complexity is a natural next step.

# References

[Alchourrón, 1993] C. Alchourrón. Philosophical foundations of deontic logic and the logic of defeasible conditionals. In J.-J.Ch. Meyer and R.J. Wieringa,

[20]Bossert and Suzumara (personal communication) reached a similar conclusion within the framework of rational choice theory (cf. [Bossert and Suzumura, 2010, chapter 3].

editors, *Deontic Logic in Computer Science*, pages 43–84. John Wiley & Sons, Inc., New York, 1993.

[Alchourrón, 1995] C. Alchourrón. Defeasible logic: demarcation and affinities. In G. Crocco, L. Fariñas del Cerro, and A. Herzig, editors, *Conditionals: From Philosophy to Computer Science*, pages 67–102. Oxford University Press, Oxford, 1995.

[Åqvist, 1987] L. Åqvist. *An Introduction to Deontic logic and the Theory of Normative Systems*. Bibliopolis, Naples, 1987.

[Åqvist, 1993] L. Åqvist. A completeness theorem in deontic logic with systematic frame constants. *Logique & Analyse*, 36(141-142):177–192, 1993.

[Åqvist, 1997] L. Åqvist. On certain extensions of von Kutschera's preference-based dyadic deontic logic. In W. Lenzen, editor, *Das weite Spektrum der analytischen Philosophie: Festschrift fur Franz von Kutschera*, pages 8–23. Walter de Gruyter, Berlin, New York, 1997.

[Åqvist, 2000] L. Åqvist. Three characterizability problems in deontic logic. In R. Demolombe and R. Hilpinen, editors, *Proceedings of the* $5^{th}$ *International Workshop on Deontic Logic In Computer Science (ΔEON'00)*, pages 16–41. ONERA-DGA, 2000.

[Åqvist, 2002] L. Åqvist. Deontic logic. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 8, pages 147–264. Kluwer Academic Publishers, Dordrecht, Holland, 2nd edition, 2002. Originally published in [Gabbay and Guenthner, 1984, pp. 605–714].

[Asher and Bonevac, 1997] N. Asher and D. Bonevac. Common sense obligation. In Nute [1997], pages 159–203.

[Benzmüller and Andrews, 2019] C. Benzmüller and P. Andrews. Church's type theory. In E. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2019 edition, 2019. `https://plato.stanford.edu/entries/type-theory-church/`.

[Benzmüller *et al.*, 2019] C. Benzmüller, A. Farjami, and X. Parent. Åqvist's dyadic deontic logic E in HOL. *Journal of Applied Logics–IfCoLog*, 6:715–732, 2019.

[Benzmüller *et al.*, 2020] C. Benzmüller, X. Parent, and L. van der Torre. Designing normative theories of ethical reasoning: LogiKEy formal framework, methodology, and tool support. *Artificial Intelligence*, 287:103348, 2020.

[Benzmüller, 2019] C. Benzmüller. Universal (meta-)logical reasoning: Recent successes. *Science of Computer Programming*, 172:48–62, 2019.

[Blackburn *et al.*, 2001] P. Blackburn, M. de Rijke, and Y. de Venema. *Modal Logic*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 2001.

[Blanchette and Nipkow, 2010] J. C. Blanchette and T. Nipkow. Nitpick: A counterexample generator for higher-order logic based on a relational model finder. In M. Kaufmann and L. C. Paulson, editors, *Interactive Theorem Proving, First International Conference, ITP 2010, Edinburgh, UK, July 11-14, 2010. Proceedings*, volume 6172 of *Lecture Notes in Computer Science*, pages 131–146. Springer, 2010.

[Bossert and Suzumura, 2010] W. Bossert and K. Suzumura. *Consistency, Choice, and Rationality*. Harvard University Press, Cambridge, 2010.

[Boutilier, 1994] G. Boutilier. Toward a logic for qualitative decision theory. In J. Doyle, E. Sandewall, and P. Torasso, editors, *Proceedings of the 4th International Conference on Principles of Knowledge Representation and Reasoning (KR-94)*, pages 75–86. Morgan Kaufman, Bonn, 1994.

[Burgess, 1981] J. P. Burgess. Quick completeness proofs for some logics of conditionals. *Notre Dame J. Formal Logic*, 22(1):76–84, 1981.

[Carmo, 2009] J. Carmo. Private communication, 2009.

[Chellas, 1975] B. Chellas. Basic conditional logic. *Journal of Philosophical Logic*, 4(2):133–153, 1975.

[Chellas, 1980] B. Chellas. *Modal Logic*. Cambridge University Press, Cambridge, 1980.

[Chernoff, 1954] H. Chernoff. Rational selection of decision functions. *Econometrica*, 22(4):422–443, 1954.

[Chisholm, 1963] R. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24:33–36, 1963.

[Danielsson, 1968] S. Danielsson. *Preference and Obligation, Studies in the Logic of Ethics*. Filosofiska Färeningen, 1968.

[Fehige, 1994] C. Fehige. The limit assumption in deontic (and prohairetic) logic. In G. Meggle and U. Wessels, editors, *Analyomen 1*, pages 42–56. De Gruyter, Berlin, 1994.

[Gabbay and Guenthner, 1984] D. Gabbay and F. Guenthner, editors. *Handbook of Philosophical Logic*, volume II. Reidel, Dordrecht, Holland, 1st edition, 1984.

[Gabbay *et al.*, 2013] D. Gabbay, J. Horty, X. Parent, R. van der Meyden, and L. van der Torre, editors. *Handbook of Deontic Logic and Normative Systems*, volume 1. College Publications, London, 2013.

[Giordano *et al.*, 2007] L. Giordano, V. Gliozzi, and G. L. Pozzato. KLM-Lean 2.0: A theorem prover for KLM logics of nonmonotonic reasoning. In N. Olivetti, editor, *Automated Reasoning with Analytic Tableaux and Related Methods*, pages 238–244, Berlin, Heidelberg, 2007. Springer.

[Goble, 2003] L. Goble. Preference semantics for deontic logics. Part I: Simple models. *Logique & Analyse*, 46(183-184):383–418, 2003.

[Goble, 2013] L. Goble. Prima facie norms, normative conflicts and dilemmas. In Gabbay et al. [2013], pages 241–352.

[Goble, 2014] L. Goble. Further notes on Kratzer semantics for modality, with application to dyadic deontic logic, 2014. Unpublished.

[Goble, 2015] L. Goble. Models for dyadic deontic logics, 2015. Unpublished (version dated 8 October 2015).

[Goble, 2019] L. Goble. Axioms for Hansson's dyadic deontic logics. *Filosofiska Notiser*, 6(1):13–61, 2019.

[Goldman, 1977] H. Goldman. David Lewis's semantics for deontic logic. *Mind*, 86(342):242–248, 1977.

[Hansen, 1998] J. Hansen. Notes to my DEON'98 contribution, 1998. Available on-line at the address: `http://www.hh.shuttle.de/win/Joerg.Hansen/Deontic.html` (this document was initially distributed at the DEON conference held in Bologna in 1998, where a first version of [Hansen, 1999] was presented).

[Hansen, 1999] J. Hansen. On relations between Åqvist's deontic system G and van Eck's deontic temporal logic. In P. Mc Namara and H. Prakken, editors, *Norms, Logics and Information Systems*, Frontiers in Artificial Intelligence and Applications, pages 127–144. IOS Press, Amsterdam, 1999.

[Hansen, 2005] J. Hansen. Conflicting imperatives and dyadic deontic logic. *Journal of Applied Logic*, 3(3-4):484–511, 2005.

[Hanson, 1965] W. H. Hanson. Semantics for deontic logic. *Logique & Analyse*, 8:177–190, 1965.

[Hansson, 1969] B. Hansson. An analysis of some deontic logics. *Noûs*, 3(4):373–398, 1969. Reprinted in [Hilpinen, 1971, pp. 121-147].

[Hansson, 1968] B. Hansson. Choice structures and preference relations. *Synthese*, 18(4):443–458, 1968.

[Hansson, 2009] S.O. Hansson. Preference-based choice functions: a generalized approach. *Synthese*, 157, 2009.

[Henkin, 1950] L. Henkin. Completeness in the theory of types. *Journal of Symbolic Logic*, 15(2):81–91, 06 1950.

[Herzberger, 1973] H. Herzberger. Ordinal preference and rational choice. *Econometrica*, 41(2):187–237, 1973.

[Hilpinen and McNamara, 2013] R. Hilpinen and P. McNamara. Deontic logic: a historical survey and introduction. In Gabbay et al. [2013], pages 3–136.

[Hilpinen, 1971] R. Hilpinen, editor. *Deontic Logic: Introductory and Systematic Readings*. Reidel, Dordrecht, 1971.

[Hilpinen, 2001] R. Hilpinen. Deontic logic. In L. Goble, editor, *The Blackwell Guide to Philosophical Logic*, pages 159–182. Blackwell Publishers, Malden, 2001.

[Horty, 2001] J. Horty. *Agency and Deontic Logic*. Oxford University Press, New York, 2001.

[Horty, 2014] J. Horty. Deontic modals: Why abandon the classical semantics? *Pacific Philosophical Quarterly*, 95(4):424–460, 2014.

[Jackson, 1985] F. Jackson. On the semantics and logic of obligation. *Mind*, 94(374):177–195, 1985.

[Kaufmann, 2017] S. Kaufmann. The limit assumption. *Semantics and Pragmatics*, 10:1–29, 2017.

[Kratzer, 1991] A. Kratzer. Modality. In A. von Stechow and D. Wunderlich, editors, *Semantics: An International Handbook of Contemporary Research*, page 639–650. De Gruyter, Berlin, summer 2019 edition, 1991.

[Kratzer, 2012] A. Kratzer. The notional category of modality. In A. Kratzer, editor, *Modals and Conditionals*, pages 27–69. Oxford University Press, 2012.

[Kraus *et al.*, 1990] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44(1-2):167–207, 1990.

[Lamarre, 1991] P. Lamarre. S4 as the conditional logic of nonmonotonicity. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 357–367. Morgan Kaufmann, 1991.

[Lehmann and Magidor, 1992] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55(1):1–60, 1992.

[Lewis, 1973] D. Lewis. *Counterfactuals*. Blackwell, Oxford, 1973.

[Lewis, 1974] D. Lewis. Semantic analyses for dyadic deontic logic. In S. Stenlund, A.-M. Henschen-Dahlquist, L. Lindahl, L. Nordenfelt, and J. Odelstad, editors, *Logical Theory and Semantic Analysis*, volume 63 of *Synthese Library*, pages 1–14. Springer, Netherlands, 1974.

[Lewis, 1981] D. Lewis. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic*, 10:217–234, 1981.

[Lindström, 1991] S. Lindström. A semantical approach to nonmonotonic reasoning, 1991. Uppsala Prints and Preprints in Philosophy, Department of Philosophy, University of Uppsala.

[Loewer and Belzer, 1983] B. Loewer and M. Belzer. Dyadic deontic detachment. *Synthese*, 54:295–318, 1983.

[Makinson, 1989] D. Makinson. General theory of cumulative inference. In M. Reinfrank, J. de Kleer, M. Ginsberg, and E. Sandewall, editors, *Proceedings of the 2nd International Workshop on Non-monotonic Reasoning*, volume 346 of *Lecture Notes in Computer Science*, pages 1–18. Springer, New York, 1989.

[Makinson, 1993] D. Makinson. Five faces of minimality. *Studia Logica*, 52(3): 339–379, 1993.

[McNamara, 1995] P. McNamara. The confinement problem: How to terminate your mom with her trust. *Analysis*, 55(4):310–313, 1995.

[Moulin, 1985] H. Moulin. Choice functions over a finite set: A summary. *Social Choice and Welfare*, 2(2):147–160, 1985.

[Nipkow *et al.*, 2002] T. Nipkow, L.C. Paulson, and M. Wenzel. *Isabelle/HOL: A Proof Assistant for Higher-Order Logic*, volume 2283 of *LNCS*. Springer, Lecture Notes in Computer Science, 2002.

[Nute, 1997] D. Nute, editor. *Defeasible Deontic Logic*. Kluwer Academic Publishers, Dordrecht, 1997.

[Olivetti and Pozzato, 2005] N. Olivetti and G. Pozzato. KLMLean 1.0: a theorem prover for logics of default reasoning. In H. Schlingloff, editor, *Proceedings of the 4th International Workshop on Methods for Modalities (M4M-4)*, pages 235–245. 2005.

[Olivetti and Pozzato, 2014] N. Olivetti and G. Pozzato. NESCOND: an implementation of nested sequent calculi for conditional logics. In S. Demri, D. Kapur, and C. Weidenbach, editors, *Automated Reasoning - 7th International Joint Conference, IJCAR 2014*, volume 8562 of *Lecture Notes in Computer Science*, pages 511–518. Springer, 2014.

[Parent, 2001] X. Parent. Cumulativity, identity and time in deontic logic. *Fundam. Inform.*, 48(2-3):237–252, 2001.

[Parent, 2008] X. Parent. On the strong completeness of Åqvist's dyadic deontic logic G. In R. van der Meyden and L. van der Torre, editors, *Deontic Logic in Computer Science (DEON 2008)*, volume 5076 of *Lecture Notes in Artificial Intelligence*, pages 189–202. Springer, Berlin/Heidelberg, 2008.

[Parent, 2010] X. Parent. A complete axiom set for Hansson's deontic logic DSDL2. *Logic Journal of the IGPL*, 18(3):422–429, 2010.

[Parent, 2012] X. Parent. Why be afraid of identity? In A. Artikis, R. Craven, N. K. Cicekli, B. Sadighi, and K. Stathis, editors, *Logic Programs, Norms and Action–Essays in Honor of Marek J. Sergot on the Occasion of His 60th Birthday*, volume 7360 of *Lecture Notes in Artificial Intelligence*, pages 295–307, Heidelberg, 2012. Springer.

[Parent, 2014] X. Parent. Maximality vs. optimality in dyadic deontic logic. *Journal of Philosophical Logic*, 43(6):1101–1128, 2014.

[Parent, 2015] X. Parent. Completeness of Åqvist's systems E and F. *Review of Symbolic Logic*, 8(1):164–177, 2015.

[Prakken and Sergot, 1997] H. Prakken and M. Sergot. Dyadic deontic logic and contrary-to-duty obligations. In Nute [1997], pages 223–262.

[Rott, 2001] H. Rott. *Change, Choice and Inference.* Clarendon Press, Oxford, 2001.

[Schlechta, 1995] K. Schlechta. Preferential choice representation theorems for branching time structures. *Journal of Logic and Computation*, 5(6):783–800, 1995.

[Schlechta, 1997] K. Schlechta. *Nonmonotonic Logics.* Springer, 1997.

[Sen, 1969] A. Sen. Quasi-transitivity, rational choice and collective decisions. *The Review of Economic Studies*, 36(3):381–393, 1969.

[Sen, 1971] A. Sen. Choice functions and revealed preferences. *The Review of Economic Studies*, 38(3):307–317, 1971.

[Sen, 1997] A. Sen. Maximization and the act of choice. *Econometrica*, 65(4): 745–779, 1997.

[Shoham, 1988] Y. Shoham. *Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence.* MIT Press, Cambridge, MA, USA, 1988.

[Spohn, 1975] W. Spohn. An analysis of Hansson's dyadic deontic logic. *Journal of Philosophical Logic*, 4(2):237–252, 1975.

[Stalnaker, 1968] R. Stalnaker. A theory of conditionals. In N. Rescher, editor, *Studies in Logical Theory*, pages 98–112. Blackwell, Oxford, 1968.

[Steen and Benzmüller, 2018] A. Steen and C. Benzmüller. The higher-order prover Leo-III. In D. Galmiche, S. Schulz, and R. Sebastiani, editors, *Automated Reasoning–9th International Joint Conference, IJCAR 2018, Proceedings*, pages 108–116. Springer, 2018.

[Steen, 2018] A. Steen. *Extensional Paramodulation for Higher-Order Logic and its Effective Implementation Leo-III*, volume 345 of *DISKI*. Akademische Verlagsgesellschaft AKA GmbH, Berlin, 9 2018. Dissertation, Freie Universität Berlin, Germany.

[Stolpe, 2020] A. Stolpe. Unsettling preferential semantics. *Journal of Philosophical Logic*, 49:371–399, 2020.

[Temkin, 1987] L. S. Temkin. Intransitivity and the mere addition paradox. *Philosophy and Public Affairs*, 16(2):138–187, 1987.

[Tomberlin, 1981] J.E. Tomberlin. Contrary-to-duty imperatives and conditional obligation. *Noûs*, pages 357–375, 1981.

[van Benthem, 2001] J. van Benthem. Correspondence theory. In D. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 3, pages 325–408. Kluwer, Dordrecht, 2nd edition, 2001. Originally published in [Gabbay and Guenthner, 1984, pp. 167–248].

[van der Torre and Tan, 1997] L. van der Torre and Y.-H. Tan. The many faces of defeasibility in defeasible deontic logic. In Nute [1997], pages 79–121.

[van der Torre and Tan, 1999] L. van der Torre and Y. H. Tan. Contrary-to-duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence*, 27(1-4):49–78, 1999.

[van Fraassen, 1972] B. C. van Fraassen. The logic of conditional obligation. *Journal of Philosophical Logic*, 1(3/4):417–438, 1972.

[van Kutschera, 1974] F. van Kutschera. Normative Präferenzen und bedingte Gebote. In H. Lenk, editor, *Normenlogik*, pages 137–165. Verlag Dokumentation, Pullach bei Munchen, 1974.

[Veltman, 1985] F. Veltman. *Logics for Conditionals.* PhD thesis, University of Amsterdam, 1985.

# Appendix A: Proof of Thm 3.3 (vi)

It is enough to describe a selection function model $M = (W, \mathfrak{f}, v)$ in which $\mathfrak{f}$ meets syntax-independence ($\mathfrak{f}0$), inclusion ($\mathfrak{f}1$), Chernoff ($\mathfrak{f}2$), consistency-preservation ($\mathfrak{f}3$) and Aizerman ($\mathfrak{f}4$), and in which (Sp) is falsified. The claim that (Sp) is not derivable in **F**+(CM) follows at once from Theorem 4.13 (iii). The same holds for (RM).

Our counter-model for (Sp) is similar to the model used in the proof of Fact 2.13. Define $M = (W, \mathfrak{f}, v)$ as follows: $W = \{a, b, c\}$; $\mathfrak{f}$ is defined by

$$\mathfrak{f}(A) = \begin{cases} \{a, c\} & \text{if } \|A\| = W \\ \|A\| & \text{otherwise;} \end{cases}$$

$v(p) = W$, $v(q) = \{b, c\}$, $v(r) = \{a, c\}$ and $v(s) = \emptyset$ for all the other propositional atoms $s$. ($\mathfrak{f}$0), ($\mathfrak{f}$1), ($\mathfrak{f}$2), ($\mathfrak{f}$3) and ($\mathfrak{f}$4) hold. But (Sp) is falsified at, e.g., world $a$:

- $\mathfrak{f}(p) = \{a, c\} \cap \|q\| = \{b, c\} \neq \emptyset \Rightarrow a \models P(q/p)$
- $\mathfrak{f}(p) = \{a, c\} \subseteq \|q \to r\| = \{a, c\} \Rightarrow a \models \bigcirc(q \to r/p)$
- $\mathfrak{f}(p \wedge q) = \{b, c\} \not\subseteq \|r\| = \{a, c\} \Rightarrow a \not\models \bigcirc(r/p \wedge q)$

# Appendix B: Proof of Thms 4.2 (ii) and 4.5 (ii)

For the reader's convenience, I restate the theorems to be proven:

*Theorem 4.2 (ii).* Under the max rule, **F**+(CM) is sound and complete with respect to the class of preference models in which $\succeq$ is max-smooth and transitive.

*Theorem 4.5 (ii).* Under the max rule, **F**+(CM) is sound and complete with respect to the class of preference models in which $\succeq$ is max-smooth, transitive, and reflexive.

Soundness is straightforward. Completeness for models in which $\succeq$ is max-smooth and transitive follows from completeness for models in which $\succeq$ is max-smooth, transitive and reflexive. Therefore, I will focus on the latter. I find it more convenient to use an indirect approach, and show how the result can be obtained from the completeness theorem for a betterness relation max-smooth and reflexive, Theorem 4.5 (i), page 32. The detailed proof of the latter result may be found in [Parent, 2014]. The betterness relation in the canonical model as defined there does not satisfy the property of transitivity. Nevertheless, the desired result follows, because one can transform the model into one in which $\succeq$ is transitive in a truth-preserving way.[21]

Call $\succeq$ virtually connected whenever $a \succeq b$ implies $a \succeq c$ or $c \succeq b$. Given reflexivity, virtual connectivity implies totalness, but not the other way around. In [Parent, 2014] it is argued that on the canonical model

[21] A direct proof is also possible. We need only change the definition of $\succeq$ in the canonical model, and adapt the initial proof accordingly. The definition used by Goble for his systems DDL-4 [Goble, 2015, p. 176 *et seq.*] and DDL-c [Goble, 2019] achieves the result we want. The definition puts $(a, B) \succeq (b, C)$ whenever $(a, B) = (b, C)$ or ($B \geq C$ and $C \notin a$). For simplicity's sake, I choose the indirect method.

of **F**+(CM) as defined there (cf. Definitions 4.11 and 4.12, page 37 *supra*) the betterness relation $\succeq$ is total (hence reflexive) and opt-smooth (*resp.*, max-smooth). The first step is to realize that $\succeq$ is also virtually connected, because the relation $\geq$ (in terms of which $\succeq$ is defined) is transitive. Recall that $A \geq B$ is a shorthand for $\bigcirc(A/A \vee B)$, and that (in the principal case) $(a, B) \succeq (b, C)$ iff: either $C \not\geq B$ or $B \in b$.

The following fact from [Parent, 2014] will also be helpful:

**Fact B.1.** *If $A \geq B \geq C$, $w^A \subseteq a$, and $C \in a$, then $w^B \subseteq a$.*

*Proof.* This is [Parent, 2014, Lemma 2 (iii)]. □

Now for the main observation:

**Fact B.2.** *In the canonical model $M^{(w,A)}$ of **F**+(CM) (as defined in Definitions 4.11 and 4.12, on page 37), $\succeq$ is virtually connected.*

*Proof.* Let $(a, B)$, $(b, C)$ and $(c, D)$ be such that $(a, B) \not\succeq (c, D)$ and $(c, D) \not\succeq (b, C)$.

Case 1: $w^A \subseteq w$ for some $A$. In that case, the canonical model generated by $(w, A)$ is as in Definition 4.11. So $C \geq D$, $D \geq B$ and $D \notin b$. From the first two, $C \geq B$, by law ($\geq$-trans) in Theorem 3.3. By construction, $w^C \subseteq b$. By (Id), $D \in w^D$ and so $w^D \not\subseteq b$. By Fact B.1, $B \notin b$. By Definition 4.11 (ii), $(a, B) \not\succeq (b, C)$ as required.

Case 2: $w^A \subseteq w$ for no $A$. In that case, the canonical model generated by $(w, A)$ is as in Definition 4.12. When it is supposed that $(c, D) \not\succeq (b, C)$, that entails that $(b, C) \in \widetilde{W}$, by definition of $\succeq$. Either (i) $(a, B) := (w, A)$ or (ii) $(a, B) \in \widetilde{W}$. In case (i), $(a, B) \not\succeq (b, C)$ as required. In case (ii), the hypothesis $(a, B) \not\succeq (c, D)$ entails that $(c, D) \in \widetilde{W}$, and the claim follows for the same reason as in case 1. □

The second step is to realize that in the presence of reflexivity virtual connectivity and transitivity do not make much difference as long as we are only interested in maximal elements. To be more precise, a reflexive and virtually connected relation can be transformed into a reflexive and transitive (albeit not necessarily total) one in a truth-preserving way with respect to the max rule. (It does not matter which rule is applied in the input model, since its betterness relation is total.)

**Theorem B.3.** *For every preference model $M = (W, \succeq, v)$ in which $\succeq$ is reflexive and virtually connected, there is a preference model $M' = (W, \succeq', v)$ (with $W$ and $v$ the same) in which $\succeq'$ is reflexive and transitive, such that $M$ and $M'$ are equivalent under the max rule. Furthermore, if $\succeq$ is max-smooth, then $\succeq'$ is max-smooth.*

*Proof.* Starting with $M = (W, \succeq, v)$, define $M' = (W, \succeq', v)$ by putting $a \succeq' b$ whenever $a = b$ or $b \not\succeq a$.

Reflexivity of $\succeq'$ is immediate. Transitivity of $\succeq'$ follows from virtual connectivity of $\succeq$. Let $a \succeq' b$ and $b \succeq' c$. If one of $a = b$, $b = c$ and $a = c$ is the case, then we are done. So assume $a \neq b$, $b \neq c$ and $a \neq c$ . Then $a \succeq' b$ and $b \succeq' c$ mean that $b \not\succeq a$ and $c \not\succeq b$. By virtual connectivity, $c \not\succeq a$, and so $a \succeq' c$ as required.

To show equivalence, it is enough to note that:

**Lemma B.4.** $\succ = \succ'$.

*Proof of Lemma B.4.* The argument for the $\subseteq$-direction appeals to the reflexivity of $\succeq$. Let $a \succ b$. Hence $a \succeq b$ but $b \not\succeq a$. The latter implies $a \succeq' b$, but also that $a \neq b$ (since $\succeq$ is reflexive). On the other hand, $a \succeq b$ and $a \neq b$ in turn imply $b \not\succeq' a$. Hence $a \succ' b$ as required.

For the $\supseteq$-direction, let $a \succ' b$. Hence $a \succeq' b$ but $b \not\succeq' a$. The latter means that $a \neq b$ and $a \succeq b$. For $a \succeq' b$ to hold, it must be the case that $b \not\succeq a$, which suffices for $a \succ b$. □

With Lemma B.4 in hand, the argument is straightforward since we have that, under the inductive hypothesis,

$$\max_{\succeq}(\|B\|^M) = \max_{\succeq'}(\|B\|^{M'}) \tag{2}$$

It is also straightforward to show that max-smoothness of $\succeq$ implies max-smoothness of $\succeq'$. Details are omitted. □

From this, Theorem 4.5 (ii) follows quickly. Suppose $\Gamma \nvdash_{\mathbf{F}+(\mathrm{CM})} A$. A similar argument as in the proof of Theorem 5 of [Parent, 2014] yields that the universe of the canonical model $M$ of **F**+(CM) contains a point $a$ such that under the max rule $a$ verifies all of $\Gamma$ and falsifies $A$. On that model $\succeq$ is reflexive, max-smooth and virtually connected, Fact B.2. By Theorem B.3, $M$ can be transformed into a model $M'$ whose relation $\succeq'$ is reflexive, transitive and max-smooth. The two models share the same universe, so $a$ is in $M'$. Under the max rule $a$ verifies all of $\Gamma$ and falsifies $A$, since the two models are equivalent. Thus, it is not the case that under the max rule $\Gamma \models A$ over the class of models in which the betterness relation is reflexive, transitive and max-smooth.

## Appendix C: Proof of Thms 4.7 and 4.8

For the reader's convenience, I restate the theorems to be proven:

> *Theorem 4.7.* Under the max rule, **E** is sound and complete with respect to (i) the class of models in which $\succeq$ is transitive, and (ii) the class of models in which $\succeq$ is transitive and reflexive.
>
> *Theorem 4.8.* Under the max rule, **F** is sound and complete with respect to (i) the class of models in which $\succeq$ is transitive and max-limited, and (i) the class of models in which $\succeq$ is transitive, max-limited and reflexive.

Soundness is straightforward. For the completeness half, it suffices to invoke the following theorem.[22]

**Theorem C.1** (Goble [2015; 2019]). *For every model $M = (W, \succeq, v)$, there is a model $M' = (W', \succeq', v')$ in which $\succeq'$ is reflexive and transitive, and such that under the max rule $M$ and $M'$ are equivalent. Furthermore, if $\succeq$ is max-limited, then $\succeq'$ is also max-limited.*

*Proof.* Let $M = (W, \succeq, v)$. Define $M' = (W', \succeq', v')$ as follows:

- $W' = \{\langle a, b, n\rangle \mid a, b \in W, n \in \omega\}$
- $\langle a, b, n\rangle \succeq' \langle c, d, m\rangle$ iff (1) $\langle a, b, n\rangle = \langle c, d, m\rangle$ or

$$(2)\begin{cases} \text{(a)}\ \ b = d \ \&\ n \geq m \\ \text{and} \\ \text{(b}_1\text{)}\ \ c \neq d \ \&\ a = c \text{ or } \text{(b}_2\text{)}\ \ c = d \ \&\ a \succ c \end{cases}$$

- $v'(p) = \{\langle a, b, n\rangle \mid a \in v(p)\}$

The following applies.

**Fact C.2.** *$W' \neq \emptyset$.*

*Proof.* This follows from the fact that $W \neq \emptyset$. □

**Fact C.3.** *$\succeq'$ is reflexive.*

*Proof.* This follows at once from clause (1) of the definition of $\succeq'$. □

**Fact C.4.** *$\succeq'$ is transitive.*

[22] [Goble, 2019, p. 44] describes the theorem as a modification and generalization of a theorem due to myself, planned for inclusion in the current chapter. At the time Goble wrote his paper, such an inclusion was indeed planned. But Goble's result leaves out certain non-essential details, and for this reason I have decided to include it instead.

*Proof.* Assume $\langle a, b, n\rangle \succeq' \langle c, d, m\rangle$ and $\langle c, d, m\rangle \succeq' \langle e, f, l\rangle$.

In case one of these holds by clause (1) of the definition of $\succeq'$, then we are done. So suppose both hold by clause (2). By (2.$a$), we have $b = d$ and $d = e$, from which $b = e$ follows. We also have $n \geq m$ and $m \geq l$. By transitivity of $\geq$, one gets $n \geq l$.

Note that $\langle a, b, n\rangle \succeq' \langle c, d, m\rangle$ and $\langle c, d, m\rangle \succeq' \langle e, f, l\rangle$ cannot hold in virtue of (2.$b_2$) and (2.$b_1$), respectively. The first implies $c = d$, while the second implies $e \neq f$ and $c = e$. One then gets $e = c = d = f$, a contradiction. Similarly, $\langle a, b, n\rangle \succeq' \langle c, d, m\rangle$ and $\langle c, d, m\rangle \succeq' \langle e, f, l\rangle$ cannot both hold in virtue of (2.$b_2$). For in this case, $c \succ e$ and $e = f = d$ would imply $c \succ d$, and so $c \succ c$, given that $c = d$. This contradicts the irreflexivity of $\succ$. I consider the remaining cases in turn.

Suppose $\langle a, b, n\rangle \succeq' \langle c, d, m\rangle$ and $\langle c, d, m\rangle \succeq' \langle e, f, l\rangle$ both hold in virtue of (2.$b_1$). In that case, $c \neq d$, $a = c$, $e \neq f$ and $c = e$. From $a = c$ and $c = e$, one gets $a = e$, and so we are done.

Suppose $\langle a, b, n\rangle \succeq' \langle c, d, m\rangle$ holds in virtue of (2.$b_1$) and $\langle c, d, m\rangle \succeq' \langle e, f, l\rangle$ holds in virtue of (2.$b_2$). In that case, $c \neq d$, $a = c$, $e = f$ and $c \succ e$. One gets $a \succ e$, and so we are done.[23] □

**Lemma C.5.** *Under the max rule, $M'$ is equivalent to $M$. That is, for all $a, b \in W$, and all $n \in \omega$, $a \models A \Leftrightarrow \langle a, b, n\rangle \models A$.*

*Proof.* By induction on $A$. I only handle the case where $A = \bigcirc(C/B)$. For the left-to-right direction, it will help to note that, under the inductive hypothesis,

**Sub-lemma C.6.** *If $\langle c, d, m\rangle \in \max_{\succeq'}(\|B\|^{M'})$, then $c = d$.*

*Proof of Sub-lemma C.6.* Assume that $\langle c, d, m\rangle \in \max_{\succeq'}(\|B\|^{M'})$ and that $c \neq d$. We have $\langle c, d, m\rangle \models B$. Also $\langle c, d, m+1\rangle \in W'$. By the inductive hypothesis, $\langle c, d, m+1\rangle \models B$. Since $c \neq d$, we have

$$\langle c, d, m+1\rangle \succeq' \langle c, d, m\rangle$$

But $m + 1 > m$, and so

$$\langle c, d, m\rangle \not\succeq' \langle c, d, m+1\rangle$$

Thus, $\langle c, d, m\rangle \notin \max_{\succeq'}(\|B\|^{M'})$, contrary to assumption, and one must conclude that $c = d$, after all. □

[23]Fact C.4 is Lemma 31 in [Goble, 2019, p. 33]. I have modified the part of the argument dealing with the case where the two opening suppositions hold in virtue of (2.$b_2$). In the paper the case is described as a possible one. But it is not, because the second supposition would hold only if (in the author's notation) $c < e$; this is a contradiction since $c = e$.

One can now turn to the proof of equivalence, starting with the right-to-left direction.

($\Leftarrow$) Assume $\langle a,b,n\rangle \models \bigcirc(C/B)$. Let $c \in \max_{\succeq}(\|B\|^M)$. We have $c \models B$. By construction $\langle c,c,n\rangle \in W'$. Assume for a reductio that $\langle c,c,n\rangle \notin \max_{\succeq'}(\|B\|^{M'})$. By the inductive hypothesis, $\langle c,c,n\rangle \models B$. So there is some $\langle d,e,m\rangle \in \|B\|^{M'}$ such that

$$\langle d,e,m\rangle \succeq' \langle c,c,n\rangle \tag{$\alpha$}$$

$$\langle c,c,n\rangle \not\succeq' \langle d,e,m\rangle \tag{$\beta$}$$

By ($\beta$), $\langle c,c,n\rangle \neq \langle d,e,m\rangle$. Thus, ($\alpha$) holds because condition (2.a) of the definition of $\succeq'$ is met along with one of (2.b$_1$) and (2.b$_2$). Since $c = c$, (2.b$_2$) applies, viz. $d \succ c$. By the inductive hypothesis, $d \models B$. But, then, $c \notin \max_{\succeq}(\|B\|^M)$. So one must conclude that $\langle c,c,n\rangle \in \max_{\succeq'}(\|B\|^{M'})$. But one then gets $\langle c,c,n\rangle \models C$ from the opening assumption. By the inductive hypothesis, we get $c \models C$, which suffices for $a \models \bigcirc(C/B)$.

($\Rightarrow$) Assume $a \models \bigcirc(C/B)$. Let $\langle c,d,m\rangle \in \max_{\succeq'}(\|B\|^{M'})$. By the inductive hypothesis, $c \models B$. By Sub-lemma C.6, $c = d$, *viz.* $\langle c,d,m\rangle$ is $\langle c,c,m\rangle$. Assume for a reductio that $c \notin \max_{\succeq}(\|B\|^M)$. There is some $d$ such that $d \models B$ and $d \succ c$. But $\langle d,c,m+1\rangle \in W'$. By the inductive hypothesis, $\langle d,c,m+1\rangle \models B$. By the definition of $\succeq'$, one gets $\langle d,c,m+1\rangle \succ' \langle c,c,m\rangle$, a contradiction. So one must conclude that $c \in \max_{\succeq}(\|B\|^M)$. From the opening assumption, $c \models C$, and so $\langle c,d,m\rangle \models C$ by the inductive hypothesis. This establishes the desired claim $\langle a,b,n\rangle \models \bigcirc(C/B)$. $\square$

It remains to verify that, if $\succeq$ is max-limited, then $\succeq'$ is max-limited. Assume that there exists some $\langle a,b,n\rangle \in W'$ such that $\langle a,b,n\rangle \models A$. By Lemma C.5, $a \models A$. Since $\succeq$ is max-limited, there is some $c$ with $c \in \max_{\succeq}(\|B\|^M)$. Re-running the same argument as that for the right-to-left-direction of Lemma C.5, one gets $\langle c,c,n\rangle \in \max_{\succeq'}(\|A\|^{M'})$, and thus $\succeq'$ is max-limited.

**Xavier Parent**
Institute of Logic and Computation, Technical University of Vienna, Vienna, Austria.
Email: xavier@logic.at