THREE CHARACTERIZABILITY PROBLEMS IN DEONTIC LOGIC

LENNART ÅQVIST

We consider an infinite hierarchy of systems of Alethic Modal Logic with so-called *Levels of Perfection,* and add to them suitable definitions of such interesting deontic categories as those of *supererogation, offence, conditional obligation* and *conditional permission.* We then state three problems concerning the proper characterization of the resulting logic(s) for our defined notions, and discuss two of these problems in some detail.

1. Introduction

Consider an infinite hierarchy Hm [m = 1, 2, …] of systems of Alethic Modal Logic with Levels of Perfection, as presented in Section 2 *infra.* The primitive logical vocabulary of those systems is extremely simple: it contains, in addition to the usual Boolean sentential connectives (including T and ⊥), just the alethic modal S5-operators

N (for universal necessity)

M (for universal possibility)

as well as a family

$$\{Qi\} \quad i = 1, 2, \ldots$$

of *systematic frame constants,* indexed by the set of positive integers and intended to represent various "levels of perfection" among the possible worlds in the models of the systems. The semantics and axiomatic proof theory of the logics Hm [m = 1, 2, …] are also extremely simple and straightforward. The purpose of the present paper is to raise (and to discuss to some extent) certain problems which arise from our adding to the systems Hm suitable *definitions* of interesting deontic categories, notably those of supererogation ("non-obligatory well-doing"), offence ("permissive ill-doing"), conditional obligation and conditional permission. Before stating these problems I must comment somewhat more in detail on the intended interpretation of the frame constants Qi [i = 1, 2, …]. In any logic Hm (with $m \geqslant 2$) we take the constant Qm to denote the class of *worst,* and indeed *definitely not acceptable* members of our set of possible worlds, whereas the

Nordic Journal of Philosophical Logic, Vol. 5, No. 2, pp. 65–82.

remaining constants Q1, …, Qm-1 are taken to represent classes of possible worlds that are all *acceptable* to a *sufficiently high,* albeit varying, degree: Q1 represents the *best*, or most acceptable ones, Q2 the second best, and so on down to Qm-1, which denotes the *worst among the acceptable* ones in the basic set of possible worlds considered in any Hm-model. Against this background we can state the following three worthwhile problems.

PROBLEM 1 (Supererogation and Offence; see Chisholm and Sosa 1966, McNamara 1994, 1996, Åqvist 1999).
Consider any system Hm with $m \geqslant 3$, and add to Hm the following series of definitions of m-1 monadic deontic O-operators (O for oughtness or obligation):

$$\text{Def O1. O1A} =_{df} N((Q1 \vee \ldots \vee Qm\text{-}2 \vee Qm\text{-}1) \supset A)$$

$$\text{Def O2. O2A} =_{df} N((Q1 \vee \ldots \vee Qm\text{-}2) \supset A)$$

$$\vdots$$

$$\text{DefOm-2. Om-2A} =_{df} N((Q1 \vee Q2) \supset A)$$

$$\text{DefOm-1. Om-1A} =_{df} N(Q1 \supset A)$$

Note here that the "ugly" constant Qm never appears as a disjunct in the antecedent of the implication inside N in these *definientia*. Then continue the list of definitions as follows (where F stands for wrongness or forbiddenness):

$$\text{DefFi. FiA} =_{df} \text{Oi}\neg A \ (i = 1, \ldots, m\text{-}1)$$

$$\text{DefSup. SupA} =_{df} \neg \text{O1A} \wedge \text{Om-1A}$$

$$\text{DefOff. OffA} =_{df} \neg \text{F1A} \wedge \text{Fm-1A}$$

Here, the operators Sup and Off are to be read, respectively, as "it is supererogatory that" and "it is offensive that".

Problem 1 can now be succinctly stated as follows. What is the logic of Oi, Fi ($i = 1, \ldots, m\text{-}1$), Sup and Off, which is generated from Hm by the above definitions? Is there a unique such logic so that we can meaningfully speak of *the* logic of those defined operators? Or are there several such logics possibly varying along with the systems Hm with $m = 3, 4, \ldots$, to which the above definitions were added?

PROBLEMS 2 and 3 (Conditional Obligation and Conditional Permission; see Hansson 1969, 1971, von Kutschera 1974, Åqvist 1997, 1997a).

We are now interested in adding to the systems Hm [$m = 1, 2, \ldots$] the following definitions of the dyadic deontic operators O (for conditional obligation) and P (for conditional permission):

Def O. $O(A/B) =_{df} [M(Q1 \wedge B) \supset N((Q1 \wedge B) \supset A)] \wedge$
$[(\neg M(Q1 \wedge B) \wedge M(Q2 \wedge B)) \supset N((Q2 \wedge B) \supset A)] \wedge \ldots \wedge$
$[(\neg M(Q1 \wedge B) \wedge \ldots \wedge \neg M(Qm\text{-}1 \wedge B) \wedge M(Qm \wedge B))$
$\supset N((Qm \wedge B) \supset A)]$.

Def P. $P(A/B) =_{df} M(Q1 \wedge B \wedge A) \vee (\neg M(Q1 \wedge B) \wedge M(Q2 \wedge B \wedge A))$
$\vee \ldots \vee (\neg M(Q1 \wedge B) \wedge \ldots \wedge \neg M(Qm\text{-}1 \wedge B) \wedge M(Qm \wedge B \wedge A))$.

Here, we write O(A/B) [P(A/B)] to render the ordinary language locution "if B, then it ought to be that A" ["if B, then it is permissible that A"]. We must now take seriously the potential ambiguity noted above of the question as to *"the"* logic of the dyadic operators O and P generated by these two definitions. Since the Hm form a whole hierarchy of logics, not just a single logic, I suggest that we distinguish between a *distributive* sense of the question and a *collective* one. And we give to *Problem 2* the following distributive import:

For each m = 1, 2, ...: which is the dyadic deontic O-P-fragment of Hm *including* the frame constants Qi that is generated by DefO and DefP *supra*? [*Problem 2*]

On the other hand, we take *Problem 3* to have the following collective import:

Which is the dyadic deontic O-P-fragment *without* the Qi and *common to all* the systems Hm [m = 1, 2, ...], that is generated by DefO and DefP? [*Problem 3*].

The distinction between these two problems, and their interrelationships, will become clearer as we go along.

In this paper we have to remain satisfied with having just stated *Problem 1*; there is no room here for dealing with it any further. Instead we are going to present some results that are relevant to *Problems 2* and *3* and their solutions. This decision gives rise to the following plan of the present paper. In Section 2 *infra* we give a quick description of the syntax, semantics and proof theory of the alethic modal logics Hm [m = 1, 2, ...], which provide the main motivation for raising our problems. In Section 3, then, we present an infinite hierarchy Gm [m = 1, 2, ...] of dyadic deontic logics, which are seen (Theorem 3 *infra*) to be deductively equivalent to the Hm on the basis of the two definitions DefO and DefP above; this affords a positive solution to *Problem 2*. Again, in Section 4 below, we describe a system G which is by far the most serious candidate for being the dyadic deontic O-P-fragment without the constants Qi and common to all the systems Hm, which is generated by Def O and Def P (see

Theorem 6 *infra*). Exactly in what sense that system G can be said to provide a positive solution to our *Problem 3* remains somewhat obscure, however. The matter will be discussed in some detail in the final Sections 5–6.

We hasten to remark that our logics Gm and G of conditional obligation and permission are all extensions of the system DSDL3, proposed by Bengt Hansson in his well-known pioneering paper (1969, 1971), as well as of the system D3 of von Kutschera (1974). As to the informal, philosophical motivation for Hansson's work in Dyadic Deontic Logic (and for that of David Lewis), see, e.g., the introduction to Åqvist 1997. As for current objections to the Hansson-style logic of conditional obligation that we endorse in this paper, we take them to apply only if that logic is considered in isolation, but not if combined with temporal logic, as seems to be suggested already by Hansson himself in the paper just mentioned. His suggestion has been more fully discussed in Spohn 1975 and Åqvist 1987: §8 and is further developed, e.g., in van Eck 1981, 1982 and Bailhache 1993, where a temporally dependent modality of *historical necessity* plays a crucial role. This observation should be carefully borne in mind when one tries to assess the adequacy of such logics as DSDL3, G and the Gm from an intuitive and applicational point of view.

2. The Alethic Modal Logics Hm [m = 1, 2, ...]: Syntax, Semantics and Proof Theory

The *language* of the systems Hm (with m any positive integer), has, in addition to an at most denumerable set Prop of propositional variables and the usual Boolean sentential connectives (including the constants *verum* and *falsum*, i.e. $\top$ and $\bot$), the following characteristic primitive *logical operators*:

N (for universal necessity)

M (for universal possibility)

as well as a family

$$\{Qi\} \quad i = 1, 2, \ldots$$

of *systematic frame constants*, indexed by the set of positive integers. The Qi are to represent different "levels of perfection" in the models of our systems, as explained above. The set Sent of well-formed sentences (formulas, wffs) is then defined in the obvious way—we think of the Qi as zero-place connectives on a par with $\top$ and $\bot$.

Let us next turn to the *semantics* for our alethic modal logics with frame constants. For any positive integer m, let a Hm-*structure* be any ordered quadruple

$$M = \langle W, V, \{opt\ i\} \quad i = 1, 2, \ldots; m \rangle$$

where:

(i) $W \neq \emptyset$ [W is a non-empty set of "possible worlds"].
(ii) V: Prop→pow(W) [V is a valuation function which to each propositional variable assigns a subset of W].
(iii) {opt i} i = 1, 2, … is an infinite sequence of subsets of W.
(iv) m is the positive integer under consideration.

We can now tell what it means for any sentence A to be *true at* a point ("world") $x (\in W)$ *in* a Hm-structure M [in symbols: $M, x \models A$], starting out with obvious clauses like

$$M, x \models p \text{ iff } x \in V(p), \text{ for any p in the set Prop}$$

$$M, x \models T$$

$$\text{not: } M, x \models \bot$$

and so on for molecular sentences having Boolean connectives as their principal operator. We then handle sentences having the characteristic Hm-operators as their principal operator as follows:

$$M, x \models NA \text{ iff for each y in W: } M, y \models A$$

$$M, x \models MA \text{ iff for some y in W: } M, y \models A$$

$$M, x \models Qi \text{ iff } x \in opt\ i, \text{ for all positive integers i.}$$

We now focus our attention on a special kind of Hm-structures called "Hm-models". By a Hm-*model* we shall mean any Hm-structure M, where {opt i} and m satisfy the following additional condition: *Exactly* m *Non-Empty Levels of Perfection.* This condition requires the set {opt 1, opt 2, …, opt m} to be a *partition* of W in the sense that

(a) $opt\ i \cap opt\ j = \emptyset$, for all positive integers i, j with $1 \leqslant i \neq j \leqslant m$.
(b) $opt\ 1 \cup \ldots \cup opt\ m = W$.
(c) $opt\ i \neq \emptyset$, for each i with $1 \leqslant i \leqslant m$.
(d) $opt\ i = \emptyset$, for each i with $i > m$.

As usual, then, we say that a sentence A is Hm-*valid* iff $M, x \models A$ for all Hm-models M and all points x in W. And we say that a set Γ of sentences is Hm-*satisfiable* iff there exists a Hm-model M and a member x of W such that for all sentences A in Γ: $M, x \models A$.

It is now time to consider the *proof theory* of the systems Hm. Thus, for any positive integer m, the *axiomatic system* Hm is determined by the following rule of inference, rule of proof, and axiom schemata:

Rule of inference

$$\text{mp } (\textit{modus ponens}) \ \frac{A, A \supset B}{B}$$

Rule of proof

$$\text{nec (necessitation for N) } \frac{\vdash A}{\vdash NA}$$

Axiom schemata

A0 All tautologies over Sent
A1 S5-schemata for N, M (i.e. $MA \equiv \neg N \neg A$, $N(A \supset B) \supset (NA \supset NB)$, $NA \supset A$, $NA \supset NNA$, $MNA \supset A$)
A2 $Qi \supset \neg Qj$, for all positive integers i, j with $1 \leqslant i \neq j < \omega$
A3 $Q1 \vee \ldots \vee Qm$
A4 $MQ1 \wedge \ldots \wedge MQm$.

The above axiom schemata and rules then determine syntactic notions of Hm-*provability*, Hm-*deducibility*, and Hm-[*in*]*consistency* in the usual way. We now state the following result without proof:

Theorem 1 (Soundness and Completeness of the logics Hm).

Weak version: For any m = 1, 2, … and any sentence A: A is Hm-provable iff A is Hm-valid.

Strong version: For any m = 1, 2,… and any set Γ of sentences: Γ is Hm-consistent iff Γ is Hm-satisfiable.

(A sketch of proof is given in Åqvist 1997). □

3. Solution to *Problem 2*: The Dyadic Deontic Logics Gm [m = 1, 2, …]

The *language* of the systems Gm is like that of the Hm, except that their logical vocabulary has the dyadic deontic operators O and P among their *primitive* connectives. The definition of Sent is then the obvious one.

It is convenient to begin the presentation of these systems by outlining their *proof theory*. The rule of inference mp and the rule of proof nec (for N) are common to all the Gm [m = 1, 2, …]. Consider next the following list:

Axiom schemata

A0–A4 (as in the last section)
a1 $P(A/B) \equiv \neg O(\neg A/B)$
a2 $O(A \supset C/B) \supset (O(A/B) \supset O(C/B))$

a3 $O(A/B) \supset NO(A/B)$
a4 $NA \supset O(A/B)$

α0 $N(A \equiv B) \supset (O(C/A) \equiv O(C/B))$
α1 $O(A/A)$
α2 $O(C/A \wedge B) \supset O(B \supset C/A)$
α3 $MA \supset (O(B/A) \supset P(B/A))$
α4 $P(B/A) \supset (O(B \supset C/A) \supset O(C/A \wedge B))$

(α5 = A2)
α6 $P(Qi/B) \supset ((Q1 \vee \ldots \vee Qi\text{-}1) \supset \neg B)$, for all i with $1 < i \leqslant m$
α7 $Q1 \supset (O(A/B) \supset (B \supset A))$
α8 $(Qi \wedge O(A/B) \wedge B \wedge \neg A) \supset P(Q1 \vee \ldots \vee Qi\text{-}1/B)$, for all i with $1 < i \leqslant m$.

Then, the *axiomatic system* Gm [m = 1, 2, …] is determined by *all* these schemata (and the above rules).

And we define the notions of provability, deducibility, [in]consistency and maximal consistency for the axiomatic systems Gm in the usual way.

Turning next to the *semantics* for the logics Gm, we define, for any positive integer m, a Gm-*structure* as an ordered quintuple M = ⟨W, V, {opt i} i = 1, 2, …; m, best⟩ where the first four items are as in the definition of a Hm-structure, and where

(v) best: Sent→pow(W) [best is a function which to each sentence in the Gm-language assigns a subset of W, heuristically the set of *best* worlds in the extension (truth-set) of the sentence under consideration].

For M any Gm-structure and x ∈ W, we then define the locution "M, x⊧A" recursively just as in the case of Hm-structures, except that there are the following fresh clauses in the inductive step governing sentences having O and P as their principal operator:

M, x⊧O(A/B) iff for each y in best(B): M, y⊧A

M, x⊧P(A/B) iff for some y in best(B): M, y⊧A

Again, we focus our attention on a special kind of Gm-structures: by a Gm-*model* we understand any Gm-structure M, where {opt i} and m satisfy the condition *Exactly* m *Non-Empty Levels of Perfection*, clauses (a)–(d), and where the new function best satisfies:

γ0. x ∈ best(B) iff M, x⊧B and for each y in W: if M, y⊧B, then x *is-at-least-as-good-as* y.

Here, the weak preference relation *is-at-least-as-good-as* is to be understood as follows. First of all, by clauses (a) and (b) in the condition *Exactly* m *etc.*, we have that for each x in W there is *exactly one* positive integer i with $1 \leqslant i \leqslant m$ such that $x \in$ opt i. We then define a "ranking" function r from W into the closed interval [1, m] of integers by setting

$$r(x) = \textit{the } i, \text{ with } 1 \leqslant i \leqslant m, \text{ such that } x \in \text{opt } i.$$

Finally, we define *is-at-least-as-good-as* as the binary relation on W such that for all x, y in W:

$$x \textit{ is-at-least-as-good-as } y \text{ iff } r(x) \leqslant r(y).$$

Note the importance of the new condition $\gamma 0$: it captures well the intuitive meaning of our "choice" function <u>best.</u>

Given the notion of a Gm-model, those of Gm-*validity* and Gm-*satisfiability* are defined in the usual manner.

THEOREM 2 (Soundness and Completeness of the systems Gm [m = 1, 2, …]).
Weak version: For every sentence A: A is Gm-provable iff A is Gm-valid.
Strong version: For each set Γof sentences: Γ is Gm-consistent iff Γ is Gm-satisfiable.

Proof. See Åqvist 1997, Section 3. □

We can now deal successfully with our *Problem 2* posed in the Introduction to this paper. Its solution is embodied in the following result:

THEOREM 3 (Deductive Equivalence of Hm to Gm). Let Hm + DefO + DefP be the result of adding the definitions Def O and Def P *supra* to the alethic system Hm. Then, for all m = 1, 2, …, Hm + DefO + DefP is *deductively equivalent* to Gm in the sense that the following two conditions are satisfied:

(i) Hm + Def O + Def P contains Gm.
(ii) Each of Def O and Def P is provable in the form of an equivalence in Gm.

Proof. See Åqvist 1997, Section 4. One shows that (i) and (ii) both hold good. In the case of condition (i) one starts by observing that the rules mp and nec (for N) are common to Gm and Hm, and that the same thing holds for axiom schemata A0–A4. We then go on to prove in Hm every Gm-schema on the list a1–a4, $\alpha 0$–$\alpha 8$, using Def O and Def P in the obvious way. This task is tedious, but entirely routine. Similarly for condition (ii). □

We close this section by presenting an alternative, more "semantical" method of representing the dyadic deontic systems Gm in the alethic modal logics Hm: define recursively a certain *translation* ϕ from the set of Gm-sentences into the set of Hm-sentences by the stipulations:

$$\phi(p) = p, \text{ for each propositional variable p in Prop}$$
$$\phi(T) = T$$
$$\phi(\bot) = \bot$$
$$\phi(Qi) = Qi, \text{ for each positive integer i}$$
$$\phi(\neg A) = \neg \phi(A)$$
$$\phi(A \wedge B) = (\phi(A) \wedge \phi(B))$$

and similarly for Gm-sentences having $\vee$, $\supset$, $\equiv$ as their principal operator.

$$\phi(NA) = N\phi A$$
$$\phi(MA) = M\phi A$$

where we have written ϕA instead of $\phi(A)$ to the right. Finally, we have two characteristic clauses corresponding to to Def O and Def P:

$$\phi(O(A/B)) = [M(Q1 \wedge \phi B) \supset N((Q1 \wedge \phi B) \supset \phi A)]$$
$$\wedge [(\neg M(Q1 \wedge \phi B) \wedge M(Q2 \wedge \phi B)) \supset N((Q2 \wedge \phi B) \supset \phi A)]$$
$$\wedge \ldots \wedge [(\neg M(Q1 \wedge \phi B) \wedge \ldots \wedge \neg M(Qm\text{-}1 \wedge \phi B) \wedge M(Qm \wedge \phi B))$$
$$\supset N((Qm \wedge \phi B) \supset \phi A)]$$

Similarly for $\phi(P(A/B))$: write it out as an m-termed disjunction!

We then have the following result:

THEOREM 4 (Translation Theorem for the systems Gm). For each positive integer m, and for each Gm-sentence A:

$$Gm \vdash A \text{ iff } Hm \vdash \phi A.$$

(Here, "$\vdash$" is the usual sign for provability.)

Proof. Again, see Åqvist 1997, Section 4. The left-to-right direction is more or less immediate from the proof of the Deductive Equivalence result in Theorem 3 *supra*. The opposite direction is harder: we are to show that if Hm$\vdash\phi$A, then Gm $\vdash$A, or, contrapositively, that if A is *not* Gm-provable, then ϕA is *not* Hm-provable (either). To accomplish this task, we now make an interesting application of our (weak) completeness result (Theorem 2) for the Gm, as appears from the following "overall" argument:

1. A is not Gm-provable	hypothesis
2. A is not Gm-valid	from 1 by the completeness of Gm
3. Not: M, x⊧A, for some Gm-model M and some x in W	from 2 by the definition of Gm-validity

Consider that Gm-model M. We claim that the result M# of deleting the function best in M is a Hm-model (obviously!) satisfying what I call the

> *Crucial Lemma.* For each Gm-sentence A and each x in W:
> M, x⊧A iff M#, x⊧ϕA.

We then continue our overall argument as follows:

4. Not: M#, x⊧ϕA	from 3 by the Crucial Lemma
5. ϕA is not Hm-valid	from 4 by the definition of Hm-validity
6. ϕA is not Hm-provable	from 5 by the soundness of Hm

where 6 is our desired conclusion.

Hence, in order to complete the argument, one shows that the Hm-model M# satisfies the Crucial Lemma. This is done by induction on the length of A as in my paper cited above. □

4. Weak Solution to *Problem 3*: The Dyadic Deontic Logic G

The *language* of the system G (without numerical index, then) is like that of the Gm except for lacking the systematic frame constants in its primitive logical vocabulary. The definition of Sent is then straightforward in the case of G as well.

As to the *proof theory* for G, it is like that of the Gm in having the same rule of inference, mp, and the same rule of proof, nec (for N). But the axiom schemata of the system G are just A0, A1, a1–a4, and α0–α4; i.e. what remains after we have dropped every schema in the Gm containing occurrences of the frame constants Qi. The notions of provability, deducibility, [in]consistency and maximal consistency are then the obvious ones in the present case of the system G.

As to the *semantics* for G: a G-*structure* is any ordered triple

$$M = \langle W, V, \underline{\text{best}} \rangle$$

where, as usual, (i) $W \neq \varnothing$, (ii) V: Prop→pow(W), and (v) best: Sent→pow(W). The relevant clauses in the truth-definition for G-sentences have then all been stated. Again, by a G-*model* we mean any G-structure M, where the function best satisfies the following five conditions parallelling the axioms α0–α4 [see Åqvist 1987: Ch.VI, p. 166; for any G-sentence A, we let |A| be the *extension* (or "truth-set") in M of A, i.e. $|A| = \{x \in W: M, x \models A\}$]:

$\sigma 0$ $|A| = |B|$ only if $\underline{best}(A) = \underline{best}(B)$
$\sigma 1$ $\underline{best}(A) \subseteq |A|$
$\sigma 2$ $\underline{best}(A) \cap |B| \subseteq \underline{best}(A \wedge B)$
$\sigma 3$ $|A| \neq \varnothing$ only if $\underline{best}(A) \neq \varnothing$
$\sigma 4$ $\underline{best}(A) \cap |B| \neq \varnothing$ only if $\underline{best}(A \wedge B) \subseteq \underline{best}(A) \cap |B|$

for all G-sentences A, B. G-validity and G-satisfiability are as usual.

THEOREM 5 (Soundness and Completeness of the System G).
Weak version: For each A in Sent: A is G-provable iff A is G-valid.
Strong version: For each $\Gamma \subseteq$ Sent: Γ is G-consistent iff Γ is G-satisfiable.

Proof. Omitted and left as an exercise. *Hint:* Pp. 160–165 of Åqvist 1987: Ch.VI are helpful, although the setting is somewhat different from the present one. □

The following result then provides a weak, or partial, solution to our *Problem 3*.

THEOREM 6 (Relation of G to the systems Gm and Hm).

(i) For each G-sentence A: $G \vdash A$ only if for all $m = 1, 2, \ldots$, $Gm \vdash A$.
(ii) For each G-sentence A: $G \vdash A$ only if for all $m = 1, 2, \ldots$, $Hm \vdash \phi A$,

where the translation ϕ is defined as in the last section (except that we don't need the clause for the $\underline{Q}i$ any longer).

Proof. Here, clause (ii) readily follows from clause (i) by virtue of our Translation Theorem 4 *supra*. And, in turn, clause (i) is immediate by the fact that each system Gm is an extension of G, which means, proof-theoretically, that every G-proof is also a Gm-proof, for any positive integer m. Again, semantically, clause (i) follows from the fact that every Gm-model is also a G-model (*via* the soundness of G and the completeness of each Gm)—as the reader easily verifies. □

5. CAN THEOREM 6 BE IMPROVED?

As it stands, Theorem 6 is clearly somewhat inconclusive and unsatisfactory. Can we improve it by strengthening the "only if" to an "iff" (= "if and only if") in the clauses (i)—(ii)? In Section 3 of my earlier paper (Åqvist 1997a) I claim to have established a result to precisely that effect, i.e., to have proved in addition that if A is *not* G-provable,then there exists a positive integer m such that A is *not* Gm-provable (and ϕA is *not* Hm-provable). The gist of the argument was an attempt to construct a *finite* Gm-model (for some positive integer m) from any given *finite* G-model, where both models falsify a given G-sentence. However, a detailed check of my alleged proof in Åqvist 1997a reveals that I only manage to show that from any given finite G-model we can construct a finite Gm-*structure* satisfying clauses (a)–(d) in the condition *Exactly* m *Non-Empty Levels of Perfection*; in order to show that this Gm-structure is indeed a Gm-*model* we must prove in addition that it satisfies the condition γ0 (Section 3 *supra*) on the function best. And it so turns out that it is far from easy to prove this to be really the case.

Upshot: in spite of Åqvist (1997a, Section 3), the question whether Theorem 6 can be strengthened in the indicated way remains an open problem. *Either* the converses of its clauses (i)–(ii) just do not hold, *or else*, if they hold, we still have not fully proved those converse results. The difficulty in validating condition γ0 seems to be bound up with insufficient expressive power of the object-language of the system G: by means of the dyadic deontic operators O and P we are able to say of any set best(B) that it is *included in* a set |A|, and that it has *a non-empty intersection with* such a set; but we seem unable to *tell exactly what* the set best(B) *is*, using the expressive resources of the object-language of G. However, interestingly, this fact gives us a clue to what is needed in order for the type of argument presented in Åqvist 1997a, Section 3 to work. We need to study, I suggest, certain logics having a *primitive* one-place operator "b" in their object-language that *perfectly matches* the function best in their semantics.

We then close this paper by presenting some logics of that kind and by proving a few relevant results on them (Theorems 7 and 8 *infra*).

6. Two Alethic Modal Logics with a Primitive b-Operator: The Systems HBm [m = 1, 2, …] and HB

The *language* of the infinite hierarchy of axiomatic systems HBm is like that of the Hm, except their *primitive* logical vocabulary contains the new one-place operator "b". The rules mp and nec (for N) are common to the Hm [m = 1, 2, …]. Consider next the following list:

Axiom schemata for HBm

A0–A4 as in Section 2 *supra*

β0 $N(A \equiv B) \supset N(bA \equiv bB)$
β1 $bA \supset A$
β2 $(bA \wedge B) \supset b(A \wedge B)$
β3 $MA \supset MbA$
β4 $M(bA \wedge B) \supset N(b(A \wedge B) \supset (bA \wedge B))$

For each positive integer m, the *axiomatic system* HBm is determined by *all* these schemata (and the above rules). The usual proof-theoretical notions are defined for the systems HBm in the obvious way.

As to the *semantics* for the HBm, HBm-*structures* are ordered quintuples

$$M = \langle W, V, \{\text{opt i}\} \quad i = 1, 2, \ldots; m, \underline{\text{best}} \rangle$$

which are like Gm-structures except that best is now a function from the set of HBm-sentences into the power-set of W. In the truth-definition we have the following fresh condition for sentences of the form bA:

$$M, x \models bA \text{ iff } x \in \underline{\text{best}}(A)$$

HBm-*models* are then any HBm-structures M, where {opt i} and m satisfy *Exactly* m *Non-Empty Levels of Perfection,* clauses (a)–(d), and where the new function best satisfies γ0 (Section 3 *supra*). Please bear in mind that, as compared with the Gm-semantics above, best now operates on a new set of sentences, viz. that of HBm-sentences.

Soundness and completeness of the alethic modal logics (with primitive b) HBm [m = 1, 2, …] are then more or less immediate.

We next consider the system HB (without numerical indices, then). Its *language* is like that of the HBm except for lacking the constants Qi in its primitive logical vocabulary. The *proof theory* for HB is like that of the HBm in having the same rules, mp and nec (for N). But the axiom schemata of HB are just A0, A1 and β0–β4; i.e. what remains after we have dropped every schema in the HBm containing occurrences of the frame constants Qi. As to the *semantics* for HB, we consider HB-*structures* in the sense of triples

$$M = \langle W, V, \underline{\text{best}} \rangle$$

with W, V as usual, and with best a function from the set of HB-sentences into the power-set of W. We adopt the same truth-condition for sentences of the form bA as in the case of the systems HBm [m = 1, 2, …]. Again, HB-*models* are then any HB-structures satisfying the familiar set of conditions σ0–σ4 (parallelling the axioms β0–β4); see Section 4 *supra.* Please keep in mind that, as compared to the semantics for G, the function best now operates on a new set of sentences, viz. that of HB-sentences.

Soundness and completeness of the system HB are then more or less immediate.

Theorem 7 (On the Relation of HBm, HB to Gm, G).

Let the following definitions of the dyadic deontic operators O and P be added to the alethic logics HBm and HB:

Df.O $\quad O(A/B) =_{df} N(bB \supset A)$

Df.P $\quad P(A/B) =_{df} M(bB \wedge A)$

Then:

(i) For each m = 1, 2, ...: HBm + Df.O + Df.P contains Gm.
(ii) HB + Df.O + Df.P contains G.

Furthermore, let the following definition of the operator b be added to the dyadic deontic logics Gm:

Df.b $bA =_{df} A \wedge Q1$, if $m = 1$

$$bA =_{df} (A \wedge Q1) \vee (A \wedge Q2 \wedge N(A \supset \neg Q1) \vee \ldots \vee$$
$$(A \wedge Qm \wedge N(A \supset \neg(Q1 \vee \ldots \vee Qm\text{-}1))), \text{ if } m > 1$$

Then:

(iii) For each m = 1, 2, ...: Gm + Df.b contains HBm.
(iv) Each of Df.O and Df.P is provable in the form of an equivalence in Gm + Df.b.
(v) Df.b is provable in the form of an equivalence in HBm + Df.O + Df.P.

Proof. Exercise. Obviously, clause (i) together with clauses (iii)–(v) amount to HBm + Df.O + Df.P being *deductively equivalent to* Gm + Df.b in the straightforward sense. □

THEOREM 8 (On the Relation of HB to the Logics HBm [m = 1, 2, ...]). For each HB-sentence A:

$$HB \vdash A \text{ [A is provable in HB] iff for each } m = 1, 2, \ldots: HBm \vdash A$$
$$\text{[A is HBm-provable].}$$

Proof. Clearly, there are no occurrences of frame constants in A, since A is a HB-sentence. Now, the left-to-right direction here is of course trivial, since each HBm is an extension of HB. The opposite direction is much harder, however, as can be seen from its contraposed version: if A is *not* HB-provable, then there exists a positive integer m such that A is *not* HBm-provable (either). We would like to establish this result by the following type of "overall" argument:

1. A is not HB-provable	assumption
2. A is not HB-valid	from 1 by the (weak) completeness of HB
3. Not: M, x$\models$A, for some HB-model M = ⟨W, V, <u>best</u>⟩ and some x in W	from 2 by the definition of HB-validity

Let $M^* = \langle W^*, V^*, \underline{best}^* \rangle$ be the *filtration of* M *through* the set of subsentences of A (in the sense of Åqvist (1997a, Section 3), *mutatis mutandis*), and let [x] be the equivalence class of x under a certain equivalence relation on W (see again Åqvist 1997a, Section 3). We then obtain:

4. Not: $M^*, [x] \models A$	from 3 by the Filtration Lemma for HB (Åqvist 1997a, Section 3, *mutatis mutandis*)

We now observe that the filtration M* is necessarily a *finite* HB-model, so that there can be at most a *finite* number of levels of perfection compatible with and definable on M*. Again, this means that we can construct, for some positive integer m, a HBm-model

$$M^*+ = \langle W^*, V^*, \{\text{opt } i\}\ i = 1, 2, \ldots; m, \underline{best}^* \rangle$$

with the property that

5. Not: $M^*+, [x] \models A$	from 4 by the fact that the new items {opt i} and m do not affect the truth-value of the HB-sentence A

and then argue:

6. A is not HBm-valid	from 5 by the definition of HBm-validity
7. A is not HBm-provable	from 6 by the soundness of each system HBm

where 7 is our desired conclusion. The proof of the present Theorem 8 will be finished, when all gaps in the above pattern of argument have been filled in. Clearly, Step 5 remains the crucial one, so we now turn to the problem of constructing the HBm-model M* + from the filtration M* that figures in Step 4.

First of all, we observe that the frame constants Qi [i = 1, 2, ...] are *definable in* HB by means of the following recursion:

$$Qi = bT, \text{ if } i = 1$$

$$Qi = b\neg(Q1 \vee \ldots \vee Qi\text{-}1), \text{ if } i > 1$$

The justification of Step 5 in our overall argument *supra* can then be seen to follow from the following Lemmata, the proofs of which will not be given here—they are available from the author of this paper.

Lemma 1 on HB. Let $M = \langle W, V, \underline{best} \rangle$ be any HB-model. Assume that there are *at least* m (> 0) HB-sentences MQ1, MQ2, ..., MQm such that for some x in W:

$$M, x \models MQ1 \wedge \ldots \wedge MQm$$

where the Qi [i = 1, …, m] are recursively defined as above. Then the cardinality of W is greater than or equal to m [in symbols: card(W) ⩾ m].

Lemma 2 on HB. Let M = ⟨W, V, best⟩ be a finite HB-model, i.e. with card(W) = k, for some positive integer k. Then:

(i) There are *at most* k HB-sentences MQ1, …, MQk such that for some x in W:

$$M, x \models MQ1 \wedge \ldots \wedge MQk.$$

(ii) For some integer m with 1 ⩽ m ⩽ k, there are *exactly* m HB-sentences MQ1, …, MQm such that for some x in W: $M, x \models MQ1 \wedge \ldots \wedge MQm \wedge \neg MQm+1$.

Lemma 3 on HB. The following theorem schemata are provable (and valid) in HB:

S4-principle for the operator b: $bA \supset bbA$

Syntactic version of *Arrow's Axiom*: $(N(A \supset B) \wedge M(A \wedge bB)) \supset N(bA \equiv (A \wedge bB))$

T1 $M(A \wedge Q1) \supset (bA \equiv (A \wedge Q1))$
T2 $M(A \wedge Q2) \wedge \neg M(A \wedge Q1) \supset (bA \equiv (A \wedge Q2))$
⋮
Tm $M(A \wedge Qm) \wedge \neg M(A \wedge (Q1 \vee \ldots \vee Qm\text{-}1)) \supset (bA \equiv (A \wedge Qm))$.

Lemma 4 on HB. The above series T1–Tm of "conditional definitions" of bA is equivalent in HB to a certain "categorical definition" of bA, given the assumption $MQ1 \wedge \ldots \wedge MQm \wedge \neg MQm+1$ (= Hyp), viz. as follows:

$$HB \vdash Hyp \supset (bA \equiv ((A \wedge Q1) \vee (A \wedge Q2 \wedge N(A \supset \neg Q1)) \vee \ldots \vee (A \wedge Qm \wedge N(A \supset \neg(Q1 \vee \ldots \vee Qm\text{-}1))))).$$

Let us now see how the above Lemmata on HB yield our desired justification of Step 5 in our overall argument.

Consider the filtration M* figuring in Step 4. M* is a finite HB-model so that card(W*) = k, for some positive integer k. By clause (ii) in Lemma 2 (which follows from, *inter alia,* clause (i) in that Lemma together with Lemma 1) we obtain that, for some integer m with 1 ⩽ m ⩽ k, there are *exactly* m HB-sentences MQ1, …, MQm such that for some x in W*: $M^*, x \models MQ1 \wedge \ldots \wedge MQm \wedge \neg MQm+1$ (where the Qi are defined as above). In other words we have, for that m, M*, x⊨Hyp. Then we define the following series of subsets of W*:

$$\text{opt } 1 = |Q1|$$
$$\text{opt } 2 = |Q2|$$
$$\vdots$$
$$\text{opt } m = |Qm|$$

and

$$\text{opt } i = \emptyset, \text{ for all positive integers i with } i > m.$$

Using this definition, we then define the ranking function r and the relation *is-at-least-as-good-as* on W* as in Section 3 *supra*. Finally, we take our desired M* + to be the structure

$$M^* + = \langle W^*, V^*, \{\text{opt i}\}\ i = 1, 2, \ldots; m, \underline{\text{best}}^* \rangle$$

where, as compared with the filtration M*, the only new items are the family {opt i} and the positive integer m. In order to show that M* +, as just constructed, is indeed a HBm-model we must then prove (i) that the sets opt 1, …, opt m satisfy clauses (a)–(d) in the condition *Exactly* m *Non-Empty Levels of Perfection* (which task is easily accomplished in HB and left to the reader), and (ii) that the function best* satisfies our condition γ0 already in the finite HB-model (filtration) M*, and hence in M* + as well. Lemma 3 and Lemma 4 are useful in establishing point (ii); the detailed validation of γ0 is available from the author. This completes the justification of Step 5 in our overall argument *supra* as well as the proof of Theorem 8 (notably its "if"-direction). □

References

Åqvist, L. 1987. *Introduction to Deontic Logic and the Theory of Normative Systems.* Bibliopolis, Napoli.

Åqvist, L. 1997. Systematic frame constants in defeasible deontic logic: a new form of Andersonian reduction. In D. Nute (ed.), *Defeasible Deontic Logic,* pp. 59–77. Kluwer, Dordrecht/Boston/London.

Åqvist, L. 1997a. On certain extensions of von Kutschera's preference-based dyadic deontic logic. In W. Lenzen (ed.), *Das weite Spektrum der analytischen Philosophie (Festschrift* für Franz von Kutschera), pp. 8–23. de Gruyter, Berlin/New York.

Åqvist, L. 1999. Supererogation and offence in deontic logic: an analysis within systems of alethic modal logic with levels of perfection. In R. Sliwinski (ed.), *Philosophical Crumbs* (*Festschrift* to Ann-Mari Henschen-

Dahlquist), pp. 261–276. Department of Philosophy, Uppsala University (Uppsala Philosophical Studies 49), Uppsala.

Bailhache, P. 1993. The deontic branching time: two related conceptions. *Logique et Analyse*, vol. 36, pp. 159–175.

Chisholm, R. M. and Sosa, E. 1966. Intrinsic preferability and the problem of supererogation. *Synthese*, vol. 16, pp. 321–331.

Eck, J. A. Van 1981, 1982. *A System of Temporally Relative Modal and Deontic Predicate Logic and its Philosophical Applications*. Department of Philosophy, University of Groningen 1981. Also in *Logique et Analyse*, vol. 25 (1982), pp. 249–290 and 339–381.

Hansson, B. 1969, 1971. An analysis of some deontic logics. *Nous*, vol. 3 (1969), pp. 373–398. Reprinted in R. Hilpinen (ed.), *Deontic Logic: Introductory and Systematic Readings*, pp. 121–147. Reidel, Dordrecht, 1971.

Kutschera, F. Von 1974. Normative Präferenzen und bedingte Gebote. In H. Lenk (ed.), *Normenlogik*, pp. 137–165. Verlag Dokumentation, Pullach bei Muenchen.

McNamara P. 1994, 1996. Doing well enough: toward a logic for common sense morality. In A. J. I. Jones and M. Sergot (eds.), *ΔEON '94: Second International Workshop on Deontic Logic in Computer Science* (pre-proceedings), pp. 165–197. Tano, Oslo. Also in *Studia Logica*, vol. 57 (1996), pp. 167–192.

Spohn, W. 1975. An analysis of Hansson's dyadic deontic logic. *Journal of Philosophical Logic*, vol. 4, pp. 237–252.

Department of Law,
Uppsala University,
Uppsala, Sweden
E-mail: Lennart.Aqvist@jur.uu.se