sciendo

Commitment: Some Formal Interpretations

Daniel Rönnedal
Stockholm University

Disputatio Vol. 4, No. 33

November 2012

DOI: 10.2478/disp-2012-0010

ISSN: 0873-626X

Commitment: some formal interpretations

Daniel Rönnedal
Stockholm University

BIBLID [0873-626X (2012) 33; pp. 445-457]

Abstract
We often use sentences that seem conditional in nature when we reason about normative issues, e.g. 'If you have promised to do something, you should keep your promise' and 'If you have done something bad, you should apologize'. We seem to think that promise-making in some sense commits us to promise-keeping and that acting bad in some sense creates an obligation to apologize. It is, however, not obvious how we should symbolize such sentences in a formal language. The purpose of this essay is to investigate some different possible formalizations of different conditional obligation sentences. I consider seven different interpretations of the concept of commitment or conditional obligation and I say something about the logical properties of these different interpretations.

Keywords
Commitment, conditional obligation, deontic logic, dyadic deontic logic, conditional norms.

Introduction

Many interesting normative sentences and principles seem to be conditional in nature. Here are some examples: 'If you have promised to do something, you should keep your promise', 'If you have borrowed an item from someone, you should return it', 'If you have hurt someone, you should apologize', 'You ought not to lie if you want to maintain your reputation' and 'If you want to be treated with respect, you should treat others with respect'. Sentences such as these seem to express some kind (or kinds) of conditional obligation (or obligations) or commitment. Since we often use such sentences in moral and normative arguments, it is important to try to understand their

logical form and what follows from them. However, it is not at all obvious how they should be formalized logically and one can find many different suggestions in the literature (for an introduction to deontic logic, which is the branch of logic that investigates such sentences, and some relevant references, see e.g. Åqvist 2002). The purpose of this paper is to investigate seven different ways of rendering sentences like this. First I will describe a logical system, which includes seven different dyadic symbols that can be used to explicate the concept of commitment or conditional obligation. Then I will describe several theorems and rules of inference and see which of our notions satisfy these theorems and rules of inference in our system. I will not try to decide which interpretation is the most reasonable. I think that there could well be many different forms of conditional obligations, and different kinds of conditional obligations should perhaps be formalized differently.

The formal system G

The system ***G*** described here is essentially the same as the system G (also called **G**) described in e.g. Åqvist 2002. However, a different notation is used and several new definitions are added.

Syntax

Alphabet. (i) A denumerably infinite set Prop of proposition letters p, q, r, s, t, p_1, q_1, r_1, s_1, t_1, p_2, q_2, r_2, s_2, t_2… (ii) The primitive truth-functional connectives $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction), $\rightarrow$ (material implication) and $\leftrightarrow$ (material equivalence). (iii) $\top$ (verum), $\bot$ (falsum). (iv) The modal operator $\Box$ (necessity). (v) The dyadic deontic operator $\blacktriangleright$ (dyadic obligation). (vi) The brackets (,).

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

'$\Box A$' is read 'it is necessary that A', and '$(A \blacktriangleright B)$' is read '$A$ commits us to B', 'if A, then it is obligatory that B', 'it is obligatory that if A, then B' or 'it is obligatory that B given A'.

Capital letters A, B, C ... are used to represent arbitrary (not necessarily atomic) formulas of the object language. The upper case Greek letter Σ represents an arbitrary set of formulas. $\varnothing$ denotes the empty set. Outer brackets around sentences are usually dropped if the result is not ambiguous.

Definitions. (Di) $OA = \top \blacktriangleright A$. (Dii) $A \rhd B = \neg(A \blacktriangleright \neg B)$. (Diii) $PA = \neg O \neg A$. (Div) $A > B = O(A \rightarrow B)$. (Dv) $A \supset B = A \rightarrow OB$. (Dvi) $A \rightrightarrows B = PA \wedge (A \rightarrow OB)$. (Dvii) $A \rightarrowtail B = ((A \wedge PB) \rightarrow OB)$. (Dviii) $A \Rrightarrow B = \Box(A \rightarrow OB)$. (Dix) $A \Longrightarrow B = (A \rhd \top) \wedge (A \blacktriangleright B)$. (Dx) $FA = O \neg A$.

'OA' is read 'it is obligatory that A' or 'it ought to be the case that A', 'PA' is read 'it is permitted that A', 'FA' is read 'it is forbidden that A' and '$A \rhd B$' is read 'if A, then it is permitted that B', 'it is permitted that if A, then B' or 'it is permitted that B given A'.

$\blacktriangleright$ and the symbols introduced in definitions (Div) to (Dix) are called conditional obligation symbols.

Semantics

Frames. An ordinary frame F is a relational structure $<W, \{R_A: A \in L\}>$, where W is a non-empty set of possible worlds and $\{R_A: A \in L\}$ is a set of dyadic relations on W, one for each sentence, A, in L, $R_A \subseteq W \times W$. A supplemented frame F_S is a relational structure $<W, \{R_A: A \in L\}, \geq>$, where W and $\{R_A: A \in L\}$ are exactly as in an ordinary frame and $\geq$ is a preference relation defined over the elements in W, i.e. $\geq \subseteq W \times W$. Intuitively $w_1 \geq w_2$ means that w_1 is at least as good as w_2. Let $w_1 > w_2$ if and only if (iff) $w_1 \geq w_2$ and not $w_2 \geq w_1$. Intuitively $w_1 > w_2$ means that w_1 is better than w_2. $||A||^M$ is the set of all worlds in the model M where the sentence A is true, the set of all A-worlds. $||A||^M = \{w \in W: M, w \Vdash A\}$. $M, w \Vdash A$ says that A is true at the world w in the model M. $\geq$ satisfies the following conditions: (i) $\forall x x \geq x$, (ii) if $||A||^M \neq \varnothing$ then $\{x \in ||A||^M: \forall y \in ||A||^M\ x \geq y\} \neq \varnothing$, (iii) $\forall x \forall y \forall z((x \geq y \wedge y \geq z) \rightarrow x \geq z)$, and (iv) $\forall x \forall y (x \geq y \vee y \geq x)$.

Models. An ordinary model M is a pair $<F, V>$, where F is an ordinary frame and V is a valuation function, which assigns a truth-value 1 (true) or 0 (false) to every proposition letter in Prop in each world $w \in W$. Intuitively $Vw(A) = 1$ says that the sentence A is true at the world w and $Vw(A) = 0$ that A is false at w. Every ordinary model satisfies the following conditions: for all A and for all B, (Mi)

$\forall x\forall y\forall z(xR_Ay \rightarrow zR_Ay)$, (Mii) if $||A||^M = ||B||^M$ then $R_A = R_B$, (Miii) $\forall x\forall y(xR_Ay \rightarrow M, y \Vdash A)$, (Miv) $\forall x\forall y((xR_Ay \wedge M, y \Vdash B) \rightarrow xR_{A\wedge B}y)$, (Mv) $\forall x((||A||^M \neq \varnothing) \rightarrow \exists yxR_Ay)$, and (Mvi) $\forall x\forall y\forall z((xR_Ay \wedge M, y \Vdash B) \rightarrow (xR_{A\wedge B}z \rightarrow (xR_Az \wedge M, z \Vdash B)))$. A supplemented model M_S is a pair $<F_S, V>$, where F_S is a supplemented frame and V is exactly as in an ordinary model. In a supplemented model the accessibility relations are defined in terms of the preference relation in the following way: For every A, xR_Ay iff $M, y \Vdash A \wedge \forall z(M, z \Vdash A \rightarrow y \geq z)$.

Truth conditions. The truth conditions for proposition letters, T, $\bot$ and sentences built by truth functional connectives are the usual ones. The truth conditions for the remaining sentences in L are given by the following clauses: (i) $M, w \Vdash \Box A$ iff for every $w \in W$: $M, w \Vdash A$, and (ii) $M, w \Vdash (A \blacktriangleright B)$ iff for all $w' \in W$ such that wR_Aw': $M, w' \Vdash B$. In supplemented models the truth conditions for $(A \blacktriangleright B)$ can be equivalently stated in the following way: $(A \blacktriangleright B)$ iff $\forall y((M, y \Vdash A \wedge \forall z(M, z \Vdash A \rightarrow y \geq z)) \rightarrow M, y \Vdash B)$. Informally this says that it is obligatory that B given A iff B is true in all the best A-worlds.

Validity, entailment etc. The concepts of validity, entailment etc. can be defined in the usual way. $M \Vdash A$ ($\boldsymbol{M} \Vdash A$) says that A is valid in the model M (the class of models $\boldsymbol{M}$) and $M, \Sigma \Vdash A$ ($\boldsymbol{M}, \Sigma \Vdash A$) says that Σ entails A in M ($\boldsymbol{M}$) or that A follows from Σ in M ($\boldsymbol{M}$).

Proof theory

Axioms. (A0) All truth-functional tautologies. (A1) $(C \blacktriangleright (A \rightarrow B)) \rightarrow ((C \blacktriangleright A) \rightarrow (C \blacktriangleright B))$. (A2) $(B \blacktriangleright A) \rightarrow \Box(B \blacktriangleright A)$. (A3) $\Box A \rightarrow (B \blacktriangleright A)$. (A4) An appropriate set of S5-schemata for $\Box$ (e.g. $\Box A \rightarrow A$, $A \rightarrow \Box\neg\Box\neg A$ and $\Box A \rightarrow \Box\Box A$). (A5) $\Box(A \leftrightarrow B) \rightarrow ((A \blacktriangleright C) \leftrightarrow (B \blacktriangleright C))$. (A6) $A \blacktriangleright A$. (A7) $((A \wedge B) \blacktriangleright C) \rightarrow (A \blacktriangleright (B \rightarrow C))$. (A8) $\neg\Box\neg A \rightarrow ((A \blacktriangleright B) \rightarrow \neg(A \blacktriangleright \neg B))$. (A9) $\neg(A \blacktriangleright \neg B) \rightarrow ((A \blacktriangleright (B \rightarrow C)) \rightarrow ((A \wedge B) \blacktriangleright C))$.

Rules of inference. (R1) (Modus Ponens (MP)) $A, A \rightarrow B/B$. (R2) ($\Box$-necessitation) $A/\Box A$. (R3) ($\blacktriangleright$-necessitation) $A/B \blacktriangleright A$ is derivable in $\boldsymbol{G}$.

Proofs, derivations, theorems etc. The concepts of proof, theorem, derivation etc. can now be defined in the usual way. $\vdash A$

says that A is a theorem in $\boldsymbol{G}$ and $\Sigma \vdash\!\!\!- A$ says that A is derivable from Σ in $\boldsymbol{G}$.

Soundness and completeness

We define the concepts of soundness and completeness in the usual way, i.e. $\boldsymbol{G}$ is (strongly) sound with respect to the class of models $\boldsymbol{M}$ iff $\Sigma \vdash\!\!\!- A$ entails $\boldsymbol{M}$, $\Sigma \Vdash\!\!\!- A$ and $\boldsymbol{G}$ is (strongly) complete with respect to the class of models $\boldsymbol{M}$ iff $\boldsymbol{M}$, $\Sigma \Vdash\!\!\!- A$ entails $\Sigma \vdash\!\!\!- A$.

Theorem 1. (i) $\boldsymbol{G}$ is sound and complete with respect to the class of all ordinary models. **(ii)** $\boldsymbol{G}$ is sound with respect to the class of all supplemented models.

Proof. See e.g. Åqvist 2002. A sketch of the proof is given in the appendix. ■

Theorems and rules of inference

Now we will consider some inference patterns and theorem schemas (see table 1). Later we will see which of these inference patterns and theorem schemas hold in $\boldsymbol{G}$ when the symbol © is replaced by our conditional obligation symbols throughout. We give every inference pattern and theorem schema a particular name. We can read 'A © B' as 'A commits us to B'.

Factual detachment, or Material detachment (FD)	$A, A \copyright B \vdash\!\!\!- OB$
Deontic detachment (DD)	$OA, A \copyright B \vdash\!\!\!- OB$
Strict detachment (SD)	$\Box A, A \copyright B \vdash\!\!\!- OB$
Contraposition (CP)	$A \copyright B \vdash\!\!\!- \neg B \copyright \neg A$
Strengthening of the antecedent, or Augmentation (SA)	$A \copyright C \vdash\!\!\!- (A \wedge B) \copyright C$
Weakening of the consequent (WC)	$A \copyright B \vdash\!\!\!- A \copyright (B \vee C)$
Transitivity (T)	$A \copyright B, B \copyright C \vdash\!\!\!- A \copyright C$
Idempotence (IP)	$\vdash\!\!\!- A \copyright A$
Trivial contrary-to-duty (TD)	$\vdash\!\!\!- FA \rightarrow (A \copyright B)$
Automatic commitment (AC)	$\vdash\!\!\!- OB \rightarrow (A \copyright B)$
Trivial commitment (TC)	$\vdash\!\!\!- \neg A \rightarrow (A \copyright B)$
Full commitment (FC)	$\vdash\!\!\!- A \copyright \mathrm{T}$
Vacuous commitment (VC)	$\vdash\!\!\!- \bot \copyright \mathrm{A}$

Reduction of unconditional obligation to conditional obligation (RU)	$\vdash OA \leftrightarrow (\mathrm{T} \mathbin{©} A)$
Vacuous unconditional obligation (VUO)	$\neg A \mathbin{©} A \vdash OA$
Entailment of wide conditional obligation (EW)	$\vdash (A \mathbin{©} B) \rightarrow O(A \rightarrow B)$
Entailment of narrow conditional obligation (EN)	$\vdash (A \mathbin{©} B) \rightarrow (A \rightarrow OB)$
Entailment of strict narrow conditional obligation (ESN)	$\vdash (A \mathbin{©} B) \rightarrow \Box(A \rightarrow OB)$
Restricted material (factual) detachment (RFD)	$A, PA, A \mathbin{©} B \vdash OB$
Restricted strengthening of the antecedent, Restricted augmentation (RSA)	$P(A \wedge B), A \mathbin{©} C \vdash (A \wedge B) \mathbin{©} C$
Disjunction introduction (in the antecedent) (DIA)	$A \mathbin{©} C, B \mathbin{©} C \vdash (A \vee B) \mathbin{©} C$
Strong disjunction elimination (in the antecedent) (SDEA)	$(A \vee B) \mathbin{©} C \vdash (A \mathbin{©} C) \wedge (B \mathbin{©} C)$
Weak disjunction elimination (in the antecedent) (WDEA)	$(A \vee B) \mathbin{©} C \vdash (A \mathbin{©} C) \vee (B \mathbin{©} C)$
Conjunction introduction (in the consequent) (CIC)	$A \mathbin{©} B, A \mathbin{©} C \vdash A \mathbin{©} (B \wedge C)$
Conjunction elimination (in the consequent) (CEC)	$A \mathbin{©} (B \wedge C) \vdash (A \mathbin{©} B) \wedge (A \mathbin{©} C)$
Substitutivity in the antecedent (SUA)	If $\vdash A \leftrightarrow B$, then $\vdash (A \mathbin{©} C) \leftrightarrow (B \mathbin{©} C)$
Substitutivity in the consequent (SUC)	If $\vdash A \leftrightarrow B$, then $\vdash (C \mathbin{©} A) \leftrightarrow (C \mathbin{©} B)$

Table 1

Some of these inference patterns and theorem schemas seem intuitively plausible, e.g. (SUA) and (SUC), and they should probably hold for most kinds of conditional obligations; some are problematic, e.g. (TD), (AC) and (TC), and they should probably not hold for all kinds of conditional obligations. Contrary-to-duty obligations, obligations that tell us what ought to be the case if something forbidden is the case, for instance, should not satisfy (TD), which claims that we are committed to anything if something forbidden is the case. Some principles are controversial, e.g. (FD) and (DD). At least some able defenders think we should reject these inference patterns and at least some able defenders think that they should hold. I believe that the

best explanation of this fact is that there actually are several different kinds of conditional obligation. (FD) may hold for some of these and fail for others and (DD) may hold for some of these and fail for others.

Different interpretations of commitment

Now we can formulate seven different interpretations of the concept 'commitment'. I think that all of these interpretations have some plausibility and are all therefore worth considering. (i) $(A \copyright B) = (A \blacktriangleright B)$. (ii) $(A \copyright B) = (A > B)$. (iii) $(A \copyright B) = (A \supset B)$. (iv) $(A \copyright B) = (A \rightrightarrows B)$. (v) $(A \copyright B) = (A \rightarrowtail B)$. (vi) $(A \copyright B) = (A \Rightarrow B)$. (vii) $(A \copyright B) = (A \Rightarrow B)$.

Table 2 provides a summary of which theorems and rules of inference these notions satisfy (Y: Yes, N: No).

	$\blacktriangleright$	$>$	$\supset$	$\rightrightarrows$	$\rightarrowtail$	$\Rightarrow$	$\Rightarrow$
(FD)	N	N	Y	Y	N	Y	N
(DD)	Y	Y	N	N	N	Y	Y
(SD)	Y	Y	Y	Y	N	Y	Y
(CP)	N	Y	N	N	N	N	N
(SA)	N	Y	Y	N	Y	Y	N
(WC)	Y	Y	Y	Y	N	Y	Y
(T)	N	Y	N	N	N	N	N
(IP)	Y	Y	N	N	N	N	N
(TD)	N	Y	N	N	N	N	N
(AC)	N	Y	Y	N	Y	Y	N
(TC)	N	N	Y	N	Y	N	N
(FC)	Y	Y	Y	N	Y	Y	N
(VC)	Y	Y	Y	N	Y	Y	N
(RU)	Y	Y	Y	Y	N	Y	Y
(VUO)	Y	Y	N	N	N	N	Y
(EW)	Y	Y	N	N	N	Y	Y
(EN)	N	N	Y	Y	N	Y	N
(ESN)	N	N	N	N	N	Y	N
(RFD)	N	N	Y	Y	N	Y	N
(RSA)	N	Y	Y	Y	Y	Y	N
(DIA)	Y	Y	Y	Y	Y	Y	Y
(SDEA)	N	Y	Y	N	Y	Y	N
(WDEA)	Y	Y	Y	Y	Y	Y	Y
(CIC)	Y	Y	Y	Y	Y	Y	Y

(CEC)	Y	Y	Y	Y	N	Y	Y
(SUA)	Y	Y	Y	Y	Y	Y	Y
(SUC)	Y	Y	Y	Y	Y	Y	Y

Table 2

Theorem 2. All the results listed in table 2 hold in the system ***G***.

Proof. Most of the parts of this theorem are straightforward, but I will go through some cases to illustrate the general proof method. ('PL' abbreviates 'by propositional logic'.)

Entailment of wide conditional obligation for $\blacktriangleright$: $\vdash (A \blacktriangleright B) \rightarrow O(A \rightarrow B)$.

1. $\Box(A \leftrightarrow (\top \wedge A))$ S5
2. $\Box(A \leftrightarrow (\top \wedge A)) \rightarrow ((A \blacktriangleright B) \leftrightarrow ((\top \wedge A) \blacktriangleright B))$ A5
3. $(A \blacktriangleright B) \leftrightarrow ((\top \wedge A) \blacktriangleright B)$ 1, 2, MP
4. $((\top \wedge A) \blacktriangleright B) \rightarrow (\top \blacktriangleright (A \rightarrow B))$ A7
5. $(A \blacktriangleright B) \rightarrow (\top \blacktriangleright (A \rightarrow B))$ 3, 4, PL
6. $(A \blacktriangleright B) \rightarrow O(A \rightarrow B)$ 5, (Di)

Transitivity for $>$: $A > B, B > C \vdash A > C$.

1. $A > B$ Assumption
2. $B > C$ Assumption
3. $\top \blacktriangleright (A \rightarrow B)$ 1, (Div), (Di)
4. $\top \blacktriangleright (B \rightarrow C)$ 2, (Div), (Di)
5. $\top \blacktriangleright ((A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C)))$ PL, R3
6. $5 \rightarrow ((\top \blacktriangleright (A \rightarrow B)) \rightarrow (\top \blacktriangleright ((B \rightarrow C) \rightarrow (A \rightarrow C))))$ A1
7. $(\top \blacktriangleright (A \rightarrow B)) \rightarrow (\top \blacktriangleright ((B \rightarrow C) \rightarrow (A \rightarrow C)))$ 5, 6, MP
8. $\top \blacktriangleright ((B \rightarrow C) \rightarrow (A \rightarrow C))$ 3, 7, MP
9. $8 \rightarrow ((\top \blacktriangleright (B \rightarrow C)) \rightarrow (\top \blacktriangleright (A \rightarrow C)))$ A1
10. $(\top \blacktriangleright (B \rightarrow C)) \rightarrow (\top \blacktriangleright (A \rightarrow C))$ 8, 9, MP
11. $\top \blacktriangleright (A \rightarrow C)$ 4, 10, MP
12. $A > C$ 11, (Div), (Di)

Disjunction introduction (in the antecedent) for $\Rightarrow$: $A \Rightarrow C, B \Rightarrow C \vdash (A \vee B) \Rightarrow C$.

1. $A \Rightarrow C$ Assumption
2. $B \Rightarrow C$ Assumption
3. $\Box(A \rightarrow OC)$ 1, (Dviii)
4. $\Box(B \rightarrow OC)$ 2, (Dviii)
5. $(\Box(A \rightarrow OC) \wedge \Box(B \rightarrow OC)) \rightarrow \Box((A \vee B) \rightarrow OC)$ S5
6. $\Box((A \vee B) \rightarrow OC)$ 3, 4, 5, PL
7. $(A \vee B) \Rightarrow C$ 6, (Dviii)

Conjunction elimination (in the consequent) for $\Rightarrow$: $A \Rightarrow (B \wedge C) \vdash\!\!\!- (A \Rightarrow B) \wedge (A \Rightarrow C)$.

1. $A \Rightarrow (B \wedge C)$ Assumption
2. $(A \rhd \mathrm{T}) \wedge (A \blacktriangleright (B \wedge C))$ 1, (Dix)
3. $A \blacktriangleright ((B \wedge C) \rightarrow B)$ PL, R3
4. $3 \rightarrow ((A \blacktriangleright (B \wedge C)) \rightarrow (A \blacktriangleright B))$ A1
5. $(A \blacktriangleright (B \wedge C)) \rightarrow (A \blacktriangleright B)$ 3, 4, MP
6. $(A \blacktriangleright B)$ 2, 5, PL
7. $(A \blacktriangleright C)$ Similarly
8. $(A \rhd \mathrm{T}) \wedge (A \blacktriangleright B)$ 2, 6, PL
9. $(A \rhd \mathrm{T}) \wedge (A \blacktriangleright C)$ 2, 7, PL
10. $A \Rightarrow B$ 8, (Dix)
11. $A \Rightarrow C$ 9, (Dix)
12. $(A \Rightarrow B) \wedge (A \Rightarrow C)$ 10, 11, PL

Since ***G*** is sound with respect to the class of all ordinary models and with respect to the class of all supplemented models, we can show that a sentence is not a theorem or that a sentence is not derivable from a set of sentences in ***G*** by producing either ordinary or supplemented countermodels. Here are some examples of such countermodels.

Trivial commitment for $\blacktriangleright$ fails: it is not the case that $\vdash\!\!\!- \neg A \rightarrow (A \blacktriangleright B)$. Let $M = \langle W, \{R_A: A \in L\}, V\rangle$ be an ordinary model, where $W = \{w_1, w_2\}$ and $w_1 R_p w_2$. Furthermore, let M satisfy all conditions mentioned in the section on semantics. $Vw_1(p) = Vw_2(q) = 0$ and $Vw_2(p) = 1$. p is false at w_1 in M, so the antecedent in $\neg p \rightarrow (p \blacktriangleright q)$ is true at w_1 in M. $p \blacktriangleright q$ is true at w_1 in M iff q is true in all R_p-accessible worlds. But $w_1 R_p w_2$ and q is false in w_2. Hence, $p \blacktriangleright q$ is not true at w_1 in M. Accordingly, the consequent in $\neg p \rightarrow (p \blacktriangleright q)$ is false at w_1 in M. Thus $\neg p \rightarrow (p \blacktriangleright q)$ is false at w_1 in M. Hence, it is not the case that $M \Vdash\!\!\!- \neg p \rightarrow (p \blacktriangleright q)$ and so it is not the case that $M \Vdash\!\!\!- \neg A \rightarrow (A \blacktriangleright B)$. Let ***M*** be the class of all ordinary models. Then it is not the case that $\boldsymbol{M} \Vdash\!\!\!- \neg A \rightarrow (A \blacktriangleright B)$. Since ***G*** is sound with respect to the class of all ordinary models, it follows that it is not the case that $\vdash\!\!\!- \neg A \rightarrow (A \blacktriangleright B)$.

Trivial contrary-to-duty for $\Rightarrow$ fails: it is not the case that $\vdash\!\!\!- FA \rightarrow (A \Rightarrow B)$. $\vdash\!\!\!- FA \rightarrow (A \Rightarrow B)$ iff $\vdash\!\!\!- (\mathrm{T} \blacktriangleright \neg A) \rightarrow ((A \rhd \mathrm{T}) \wedge (A \blacktriangleright B))$ by (Di), (Dix) and (Dx). Let $Ms = \langle W, \{R_A: A \in L\}, \geq, V\rangle$ be a supplemented model, where $W = \{w_1, w_2\}$, $w_1 > w_2$, $Vw_1(p) = Vw_2(p) = Vw_1(q) = Vw_2(q) = 0$. Then $||p||^{Ms} = \varnothing$ and $||\mathrm{T}||^{Ms} = \{w_1, w_2\}$. The set of all the best p-worlds is empty since there are no p-worlds and the

set of all the best T-worlds is $\{w_1\}$. $\mathrm{T} \blacktriangleright \neg p$ is true at w_2 in *Ms* iff p is false in all the best T-worlds in *Ms*. Since p is false at w_1 in *Ms* p is false at all the best T-worlds in *Ms*. Hence, $\mathrm{T} \blacktriangleright \neg p$ is true at w_2 in *Ms*. $(p \rhd \mathrm{T}) \wedge (p \blacktriangleright q)$ is true at w_2 in *Ms* iff $p \rhd \mathrm{T}$ is true at w_2 in *Ms* and $p \blacktriangleright q$ is true at w_2 in *Ms*. $p \rhd \mathrm{T}$ is true at w_2 in *Ms* iff T is true in at least one of the best p-worlds. But there are no best p-worlds. Hence, $p \rhd \mathrm{T}$ is false at w_2 in *Ms*, and so, $(p \rhd \mathrm{T}) \wedge (p \blacktriangleright q)$ is false at w_2 in *Ms*. Accordingly, $(\mathrm{T} \blacktriangleright \neg p) \rightarrow ((p \rhd \mathrm{T}) \wedge (p \blacktriangleright q))$ is false at w_2 in *Ms*. Consequently, it is not the case that *Ms* $\Vdash (\mathrm{T} \blacktriangleright \neg p) \rightarrow ((p \rhd \mathrm{T}) \wedge (p \blacktriangleright q))$ and so it is not the case that *Ms* $\Vdash (\mathrm{T} \blacktriangleright \neg A) \rightarrow ((A \rhd \mathrm{T}) \wedge (A \blacktriangleright B))$. Let ***Ms*** be the class of all supplemented models. Then it is not the case that ***Ms*** $\Vdash (\mathrm{T} \blacktriangleright \neg A) \rightarrow ((A \rhd \mathrm{T}) \wedge (A \blacktriangleright B))$. Since ***G*** is sound with respect to the class of all supplemented models, it follows that it is not the case that $\vdash (\mathrm{T} \blacktriangleright \neg A) \rightarrow ((A \rhd \mathrm{T}) \wedge (A \blacktriangleright B))$. Consequently, it is not the case that $\vdash FA \rightarrow (A \Rightarrow B)$.

Strengthening of the antecedent, or Augmentation for $\blacktriangleright$ fails: it is not the case that $B \blacktriangleright A \vdash (B \wedge C) \blacktriangleright A$. Consider the following supplemented model $Ms = <W, \{R_A: A \in L\}, \geq, V>$, where $W = \{w_1, w_2\}$, $w_1 > w_2$, $Vw_1(p) = Vw_1(q) = Vw_2(q) = Vw_2(r) = 1$ and $Vw_1(r) = Vw_2(p) = 0$. Then $\|q\|^{Ms} = \{w_1, w_2\}$ and $\|q \wedge r\|^{Ms} = \{w_2\}$. The set of all the best q-worlds is $\{w_1\}$ and the set of all the best $(q \wedge r)$-worlds is $\{w_2\}$. $q \blacktriangleright p$ is true at w_2 in *Ms* iff p is true at all the best q-worlds in *Ms*. Since p is true at w_1 in *Ms* p is true at all the best q-worlds in *Ms*. Hence, $q \blacktriangleright p$ is true at w_2 in *Ms*. $(q \wedge r) \blacktriangleright p$ is true at w_2 in *Ms* iff p is true in all the best $(q \wedge r)$-worlds in *Ms*. Since p is false at w_2 in *Ms* p is not true in all the best $(q \wedge r)$-worlds in *Ms*. Hence, $(q \wedge r) \blacktriangleright p$ is false at w_2 in *Ms*. Consequently, it is not the case that *Ms* $q \blacktriangleright p \Vdash (q \wedge r) \blacktriangleright p$ and so it is not the case that *Ms* $B \blacktriangleright A \Vdash (B \wedge C) \blacktriangleright A$. Let ***Ms*** be the class of all supplemented models. Then it is not the case that ***Ms*** $B \blacktriangleright A \Vdash (B \wedge C) \blacktriangleright A$. Since ***G*** is sound with respect to the class of all supplemented models, it follows that it is not the case that $B \blacktriangleright A \vdash (B \wedge C) \blacktriangleright A$. ■

Appendix

The proofs in this appendix are essentially due to Lennart Åqvist, see Åqvist 1987 and Åqvist 2002.

Soundness with respect to ordinary models. ***G*** is (strongly) sound with respect to the class of all ordinary models.

Proof. To prove that ***G*** is sound with respect to the class of all ordinary models we must show that every axiom in ***G*** is valid and that every rule in ***G*** is validity preserving (in the class of all ordinary models). This is straightforward. So, we only consider two examples.

(A5) is valid. Suppose that $\Box(A \leftrightarrow B) \rightarrow ((A \blacktriangleright C) \leftrightarrow (B \blacktriangleright C))$ is false at some world w in a model M. Then $||A||^M = ||B||^M$ and either (i) $A \blacktriangleright C$ is true but $B \blacktriangleright C$ is false at w in M, or (ii) $B \blacktriangleright C$ is true but $A \blacktriangleright C$ is false at w in M. Assume (i). Then there is a world w' in M such that wR_Bw' and C is false at w' in M. Since M satisfies (Mii) and $||A||^M = ||B||^M$, $R_A = R_B$. Hence, there is a world w' in M such that wR_Aw'. But C is true at every world in M that is A-accessible from w, since $A \blacktriangleright C$ is true at w in M. Consequently, C is true at w' in M. But this is absurd. In a similar way it can be shown that (ii) leads to a contradiction.

(A9) is valid. Assume that $\neg(A \blacktriangleright \neg B) \rightarrow ((A \blacktriangleright (B \rightarrow C)) \rightarrow ((A \wedge B) \blacktriangleright C))$ is false at some world w in a model M. Then $\neg(A \blacktriangleright \neg B)$ is true at w in M and $(A \blacktriangleright (B \rightarrow C)) \rightarrow ((A \wedge B) \blacktriangleright C)$ is false at w in M. Hence, $A \blacktriangleright (B \rightarrow C)$ is true and $(A \wedge B) \blacktriangleright C$ is false at w in M. Since $\neg(A \blacktriangleright \neg B)$ is true at w in M there is a world w' in M that is A-accessible from w such that B is true at w' in M and since $(A \wedge B) \blacktriangleright C$ is false at w in M there is a world w'' in M $(A \wedge B)$-accessible from w such that C is false at w'' in M. Consequently, wR_Aw'' and B is true at w'' in M. For M satisfies (Mvi). Accordingly, $B \rightarrow C$ is true at w'' in M. It follows that C is true at w'' in M. But this is absurd. ■

Soundness with respect to supplemented models. ***G*** is (strongly) sound with respect to the class of all supplemented models.

Proof. This result follows from the soundness proof above and the fact that any supplemented model satisfies (Mi) to (Mvi). ■

Completeness with respect to ordinary models. ***G*** is (strongly) complete with respect to the class of all ordinary models.

Proof. To prove that ***G*** is complete with respect to the class of all ordinary models we use the standard canonical model technique. Most of this is straightforward. But note how we define the canonical model. Let x be any set of sentences. Then we say that x is ***G***-saturated iff x is ***G***-consistent and for each sentence A, either $A \in x$ or $\neg A \in x$. Let R be a relation on the set of all ***G***-saturated sets defined as follows: xRy iff for each A in L, if $\Box A \in x$, then $A \in y$. By virtue of the S5-schemata in ***G***, the relation R is an equivalence relation on the set of all ***G***-saturated sets, which partitions that set into disjoint equivalence classes. For any ***G***-saturated set x, let $[x]_R = \{y: xRy\} =$ the R-equivalence class of x. For any ***G***-consistent set of sentences S,

we shall say that the canonical ***G***-model generated by S is the structure: $M = <W, \{R_A: A \in L\}, V>$, where: (i) $W = [S+]_R = \{y: y$ is ***G***-saturated and, for each A in L, if $\Box A \in S+$, then $A \in y\}$ ($S+$ is a ***G***-saturated extension of S). (ii) the accessibility relations are defined in the following way: xR_Ay iff $\forall C(A \blacktriangleright C \in x \rightarrow C \in y)$. (iii) V is as usual. Now the different steps in the completeness proof are more or less standard. The interesting part is to verify that the canonical model is of the right kind. We go through two examples to illustrate the method. In these proofs we use the lemma on saturated sets and the coincidence lemma. Let x be any ***G***-saturated set of sentences. Then according to the lemma on saturated sets $A \rightarrow B \in x$ iff, if $A \in x$ then $B \in x$. Let $M = <W, \{R_A: A \in L\}, V>$ be a canonical model as defined above and let $x \in W$. Then according to the coincidence lemma $M, x \Vdash A$ iff $A \in x$.

The canonical model M is a ***G***-model that satisfies (Miii). Suppose (1) xR_Ay. Then (2) $\forall C(A \blacktriangleright C \in x \rightarrow C \in y)$ [from 1]. (3) $A \blacktriangleright A \in x$ [from axiom A6]. Hence, (4) $A \in y$ [from 2 and 3]. So, (5) $M, y \Vdash A$ [from 4]. It follows that (6) $xR_Ay \rightarrow M, y \Vdash A$ [from 1 to 5].

The canonical model M is a ***G***-model that satisfies (Miv). Assume (1) xR_Ay and (2) $M, y \Vdash B$. Then (3) $\forall C(A \blacktriangleright C \in x \rightarrow C \in y)$ [from 1]. Hence, (4) $B \in y$ [from 2]. Suppose (5) $(A \wedge B) \blacktriangleright C \in x$. Then (6) $A \blacktriangleright (B \rightarrow C) \in x$ [from 5 and axiom A7]. Accordingly, (7) $B \rightarrow C \in y$ [from 3 and 6]. So, (8) $C \in y$ [from 4 and 7]. It follows that (9) $(A \wedge B) \blacktriangleright C \in x \rightarrow C \in y$ [from 5 to 8]. Consequently, (10) $xR_{A \wedge B}y$ [from 9]. In conclusion, (11) $xR_Ay \rightarrow (M, y \Vdash B \rightarrow xR_{A \wedge B}y)$ [from 1 to 10]. ■

Completeness with respect to supplemented models. In Parent 2008 Xavier Parent claims that Åqvist's system G is (strongly) complete with respect to the class of all supplemented models. I must confess that I have been unable to verify that this is true, because I have not studied the proof closely enough. If it is correct, it is good news. For, as far as I am aware, no one else has proved this. However, for our purposes in this essay it doesn't matter whether ***G*** is complete with respect to the class of all supplemented models or not.[1]

Daniel Rönnedal

[1] I would like to thank the anonymous referees, Lennart Åqvist and Paul Needham for valuable comments on an earlier version of this paper.

Department of Philosophy
Stockholm University
106 91 Stockholm, Sweden
daniel.ronnedal@philosophy.su.se

References

Parent, Xavier. 2008. On the Strong Completeness of Åqvist's Dyadic Deontic Logic G. In *Deontic Logic in Computer Science: 9th International Conference, DEON 2008*, ed. by R. van der Meyden and L. W. N. van der Torre. Berlin/Heidelberg: Springer-Verlag, 189-202.

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

Åqvist, Lennart. 2002. Deontic Logic. In *Handbook of Philosophical Logic*, *Vol. 8*, ed. by D. Gabbay and F. Guenthner. Kluwer Academic, 147-264.