Handout Lecture 1: “Standard” Deontic Logic

Xavier Parent and Leendert van der Torre
University of Luxembourg

October 18, 2016

1 Introduction

The topic of this handout is so-called “standard” deontic logic, SDL. The adjective “deontic” derives from the Greek word “deon,” meaning “duty,” “that which is binding.” The name SDL covers of family of systems based on an analogy between the deontic operator “It is obligatory that” and the alethic modal operator “It is necessary that”. The semantics of SDL is based on so-called relational semantics. It is a well-understood tool used to define the semantics of a modal logic.

⚠ The label “standard” is a misnomer. SDL has been a landmark system until the late 60s, when so-called Dyadic Standard Deontic Logic emerged as a new standard.

2 Language

Definition 1. *Let $\mathbb{P}$ be a set of atomic propositions. The language $\mathcal{L}$ of SDL is generated by the following BNF (Backus Normal Form):*

$$\phi ::= p \mid \neg\phi \mid (\phi \wedge \phi) \mid \bigcirc\phi$$

where $p \in \mathbb{P}$. The construct $\bigcirc\phi$ is read as “It is obligatory that ϕ.” Other connectives are introduced by the definitions:

disjunction	$\phi \vee \psi$	*is*	$\neg(\neg\phi \wedge \neg\psi)$
implication	$\phi \rightarrow \psi$	*is*	$(\neg\phi) \vee \psi$
equivalence	$\phi \leftrightarrow \psi$	*is*	$(\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi)$
verum	$\top$	*is*	$p \vee \neg p$ *(for some $p \in \mathbb{P}$)*
falsum	$\bot$	*is*	$\neg\top$
permission	$P\phi$	*is*	$\neg \bigcirc \neg\phi$
prohibition	$F\phi$	*is*	$\bigcirc\neg\phi$

The BNF appearing in Definition 1 says the same as Definition 2.

Definition 2. *The language of SDL is the smallest set $\mathcal{L}$ such that:*

1. $\mathbb{P} \subseteq \mathcal{L}$
2. *If $\phi \in \mathcal{L}$, then $\neg\phi \in \mathcal{L}$ and $\bigcirc\phi \in \mathcal{L}$*
3. *If $\phi \in \mathcal{L}$ and $\psi \in \mathcal{L}$, then $\phi \wedge \psi \in \mathcal{L}$*

We omit outermost parenthesis if doing so does not lead to confusion.

Remark 1. The definition allows for “iterated” modalities like $\bigcirc \bigcirc p$. “Mixed” formulas like $p \wedge \bigcirc q$ are also allowed.

Remark 2. $\mathcal{L}$ is the object-language. The Greek letters ϕ, ψ ... are not part of it; they belong to the meta-language. They are metavariables ranging over elements in $\mathcal{L}$, placeholders for object language formulas of a given category. The convention is that a metavariable is to be uniformly substituted with the same instance in all its appearances in a given more complex formula. This one is called a schema.

Exercise 1. Symbolize in SDL the following English statements:

1. You should pay tax;
2. Wash your dish or put it in the washing machine;
3. It ought to be the case that whatever ought to be the case be the case.

3 Relational semantics

This section describes the relational semantics for SDL.

Definition 3 (Relational model)**.** *A relational model M is a triplet (W, R, V) where:*

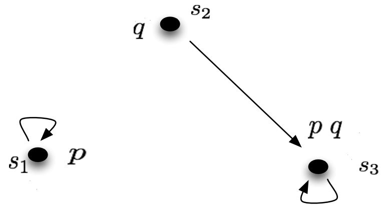
1. *W is a (non-empty) set of states (also called “possible worlds”) $s, t, \ldots$ W is called the universe of the model.*
2. *$R \subseteq W \times W$ is a binary relation over W. It is understood as a relation of deontic alternativeness: sRt (or, alternatively, $(s,t) \in R$) says that t is an ideal alternative to s, or that t is a “good” successor of s. It is “good” in the sense that all the obligations true in s are fulfilled in these. Furthermore, R is subject to the following constraint:*

$$(\forall s \in W)(\exists t \in W)(sRt) \qquad \text{(seriality)}$$

3. *$V : \mathbb{P} \mapsto 2^W$ is a valuation function assigning to each atom p a set $V(p) \subseteq W$ (intuitively the set of states at which p is true).*

We can represent a relational model using a graph. This is illustrated with Figure 1.

Figure 1: A 3-state model



States are nodes, and the content of the deontic alternativeness relation is indicated by arrows from nodes to nodes. For instance, $s_1 \rightarrow s_2$ says that s_2 is a good successor of s_1. In addition, each state s is labelled with the atoms made true at s. Here $\mathbb{P} = \{p, q\}$, and $V(p) = \{s_1, s_3\}$ while $V(q) = \{s_2, s_3\}$.

The satisfaction relation determines the truth-value of sentences according to their form.

Definition 4 (Satisfaction)**.** *Given a relational model $M = (W, R, V)$ and a state $s \in W$, we define the satisfaction relation $M, s \vDash \phi$ (read as "state s satisfies ϕ in model M", or as " s makes ϕ true M") by induction on the structure of ϕ using the following clauses:*

1. *$M, s \vDash p$ iff $s \in V(p)$.*
2. *$M, s \vDash \neg\phi$ iff $M, s \nvDash \phi$.*
3. *$M, s \vDash \phi \wedge \psi$ iff $M, s \vDash \phi$ and $M, s \vDash \psi$.*
4. *$M, s \vDash \bigcirc\phi$ iff for all $t \in W$, if sRt then $M, t \vDash \phi$.*

Intuitively, $\bigcirc\phi$ is satisfied/true at s in model M just in case ϕ is satisfied/true at each of s'ideal alternatives.

We drop reference to M, and write $s \vDash \phi$, when it is clear what model is intended.

Example 1. Consider the model shown in Figure 1. We have

1. $s_1 \vDash p$ (since $s_1 \in V(p)$)
2. $s_1 \vDash \neg q$ (since $s_1 \nvDash q$)
3. $s_1 \vDash p \wedge \neg q$
4. $s_1 \vDash \bigcirc(p \wedge \neg q)$ (only s_1 is a good successor of s_1).

Exercise 2.

(1) Define the truth-conditions for $\vee$ and $\rightarrow$ in a relational model;

(2) Same question for P and F.

Validity, semantical consequence and satisfiability are always relative to a given class of models. In this subsection they are stated with respect to the class of relational models as defined *supra*.

Definition 5 (Validity)**.** *A formula ϕ is valid (notation: $\vDash \phi$) whenever, for all models $M = (W, R, V)$ and all $s \in W$, $s \vDash \phi$.*

Definition 6 (Semantical consequence)**.** *Given a set Γ of formulas, a formula ϕ is a semantical consequence of Γ (notation: $\Gamma \vDash \phi$) whenever, for all relational models $M = (W, R, V)$, and all $s \in W$, if $s \vDash \psi$ for all $\psi \in \Gamma$, then $s \vDash \phi$.*

We omit curly brackets from singletons for ease of reading.

Definition 7 (Satisfiability)**.** *A set Γ of formulas is satisfiable if and only if there is a relational model $M = (W, R, V)$, and some state $s \in W$ such that $s \vDash \phi$ for all $\phi \in \Gamma$.*

Remark 3. Satisfiability is the semantical counterpart of consistency.

Proposition 1. *We have*

$$\{\psi_1, ..., \psi_n\} \models \phi \text{ iff } \models (\psi_1 \wedge ... \wedge \psi_n) \rightarrow \phi \qquad \text{(Deduction theorem)}$$

$$\text{If } \Gamma \models \phi \text{ and } \Gamma \subseteq \Gamma', \text{ then } \Gamma' \models \phi \qquad \text{(Monotony)}$$

Proof. Left as an exercise. □

Exercise 3. Explain in what sense the notion of validity may be described as a limiting case of the notion of consequence.

Exercise 4. Give the proof of proposition 1.

4 Proof system

4.1 System D

A proof or a derivation in a given proof system, call it **X**, is a finite sequence of formulas each of which is an instance of an axiom schema or follows from earlier formulas in the sequence by a rule schema. If there is a derivation for ϕ in **X** we write $\vdash_{\mathbf{X}} \phi$, or, if the system **X** is clear from the context, we just write $\vdash \phi$. We then also say that ϕ is a *theorem* in **X**, or that **X** proves ϕ.

Definition 8. **D** *is the proof system consisting of the following axiom schemas and rule schemas.*

$\vdash \phi$, *where ϕ is a propositional tautology*	(PL)
$\vdash \bigcirc(\phi \rightarrow \psi) \rightarrow (\bigcirc\phi \rightarrow \bigcirc\psi)$	(K)
$\vdash \bigcirc\phi \rightarrow P\phi$	(D)
If $\vdash \phi$ *and* $\vdash \phi \rightarrow \psi$ *then* $\vdash \psi$	(MP)
If $\vdash \phi$ *then* $\vdash \bigcirc\phi$	(Nec)

Example 2. Below: a derivation of $\bigcirc(p \wedge q) \rightarrow \bigcirc p$.

1. $\vdash ((p \wedge q) \rightarrow p)$ PL
2. $\vdash \bigcirc((p \wedge q) \rightarrow p)$ Nec, 1

3. $\vdash \bigcirc((p \wedge q) \to p) \to (\bigcirc(p \wedge q) \to \bigcirc p)$ K
4. $\vdash \bigcirc(p \wedge q) \to \bigcirc p$ MP, 2,3

Exercise 5. Show that
1. $\vdash Pp \to P(p \vee q)$
2. $\vdash (\bigcirc(p \wedge q) \wedge Pp) \to Pq$

Definition 9 (Syntactical consequence)**.** *Given a set Γ of formulas, we say that ϕ is a syntactical consequence of Γ (notation: $\Gamma \vdash \phi$) iff there are formulas $\psi_1, \ldots, \psi_n \in \Gamma$ such that $\vdash (\psi_1 \wedge \ldots \wedge \psi_n) \to \phi$. (In case where $n = 0$, this means that $\vdash \phi$.)*

Definition 10 (Consistency)**.** *A set Γ is consistent if $\Gamma \nvdash \bot$, and inconsistent otherwise.*

4.2 Soundness and completeness theorem

Theorem 1 (Soundness, weak version)**.** *If $\vdash \phi$ then $\models \phi$.*

Proof. It is enough to show that the axiom schemas are valid in the relevant class of models, and that the rules schema preserve validity in the relevant class of models.

We give the argument for (K). Let M and s be such that $s \models \bigcirc(\phi \to \psi)$ and $s \models \bigcirc\phi$. For the result that $s \models \bigcirc\psi$, we need to show that

For all t in M, if sRt, then $t \models \psi$.

So let t in M be such that sRt. By the two opening assumptions, $t \models \phi \to \psi$ and $t \models \phi$. By the evaluation rule for $\to$, $t \models \psi$ as required. □

Theorem 2 (Soundness, strong version)**.** *If $\Gamma \models \phi$, then $\Gamma \vdash \phi$.*

Proof. Let $\Gamma \vdash \phi$. By definition, there are $\psi_1, \ldots, \psi_n \in \Gamma$ such that $\vdash (\psi_1 \wedge \ldots \wedge \psi_n) \to \phi$. By weak soundness, $\models (\psi_1 \wedge \ldots \wedge \psi_n) \to \phi$. By the deduction theorem, $\{\psi_1, ..., \psi_n\} \models \phi$. By monotony, $\Gamma \models \phi$. □

Theorem 3 (Completeness, weak and strong version)**.** *i) if $\models \phi$ then $\vdash \phi$; ii) if $\Gamma \models \phi$ then $\Gamma \vdash \phi$.*

Proof. Weak completeness follows from strong completeness, which in turn can be established using the method of canonical models. For details, see Chellas [4]. □

Corollary 1 (Semantical compactness)**.** *Γ is satisfiable, if every finite subset $\Delta \subseteq \Gamma$ is satisfiable.*

Exercise 6.
(1) Complete the proof of soundness, theorem 1;
(2) Describe a model in which R is not serial, and in which the D axiom is not valid.

Exercise 7. To show that a set Γ of formulas is inconsistent, we usually turn to the semantics, and show that Γ is not satisfiable. Explain why unsatisfiability shows inconsistency. (Hint: use the soundness and completeness theorem.)

4.3 Decidability

A model is called finite, its universe W is finite.

Theorem 4 (Finite model property)**.** *If ϕ is satisfiable in a model, then ϕ is satisfiable in a finite model.*

Proof. Using the filtration method. Details can be found in Chellas [4]. □

Theorem 5 (Decidability)**.** *The theoremhood problem in* **D** *is decidable.*

Proof. This follows from the finite model property. □

5 Stronger systems

This section describes three extensions of **D** that have been considered in the literature: **DS4** (Deontic S4), **DS5** (Deontic S5) and **DM** (Deontic M).

DS4 extends **D** with the axiom schema

$$\bigcirc \phi \to \bigcirc \bigcirc \phi \qquad (\bigcirc 4)$$

DS4 is sound and complete with respect to the class of relational models in which R is serial and transitive:

$$(\forall s)(\forall t)(\forall u)(sRt \;\&\; tRu \to sRu) \qquad \text{(transitivity)}$$

DS5 extends **DS4** with the axiom schema

$$\neg \bigcirc \phi \to \bigcirc \neg \bigcirc \phi \qquad (\bigcirc 5)$$

DS5 is sound and complete with respect to the class of relational models in which R is serial, transitive and euclidean:

$$(\forall s)(\forall t)(\forall u)(sRt \;\&\; sRu \to tRu) \qquad \text{(euclideanness)}$$

DM is obtained by replacing, in **DS5**, $\bigcirc$5 with the axiom schema

$$\bigcirc (\bigcirc\phi \to \phi) \qquad (\bigcirc \mathrm{M})$$

DM is sound and complete with respect to the class of relational models in which R is serial, transitive and secondary reflexive:

$$(\forall s)(\forall t)(sRt \to tRt) \qquad \text{(secondary reflexivity)}$$

Note that $\bigcirc$M is a theorem of **DS5**. Hence these systems form a series of systems of increasing strength:

$$\mathbf{D} \subset \mathbf{DS4} \subset \mathbf{DM} \subset \mathbf{DS5}$$

Exercise 8. Show that $\bigcirc$M is a theorem of **DS5**.

6 Chisholm's paradox

The original phrasing of the paradox by Chisholm [5] requires a formalisation of the following scenario in which the sentences are mutually consistent and logically independent.

(A) It ought to be that Jones goes to the assistance of his neighbours.
(B) It ought to be that if Jones goes to the assistance of his neighbours, then he tells them he is coming.
(C) If Jones doesn't go to the assistance of his neighbours, then he ought not to tell them he is coming.
(D) Jones does not go to their assistance.

First attempt is inconsistent.
(A_1) $\bigcirc g$
(B_1) $\bigcirc(g \to t)$
(C_1) $\neg g \to \bigcirc\neg t$
(D_1) $\neg g$

Second attempt does not meet the requirement of independency.
(A_2) $\bigcirc g$
(B_2) $\bigcirc(g \to t)$
(C_2) $\bigcirc(\neg g \to \neg t)$
(D_2) $\neg g$

Third attempt faces the same problem.
(A_3) $\bigcirc g$
(B_3) $g \to \bigcirc t$
(C_3) $\neg g \to \bigcirc\neg t$
(D_3) $\neg g$

A fourth attempt based on B_3 and C_2 is missing as we would derive nothing.

Proposition 2. *The set $\Gamma = \{A_1, B_1, C_1, D_1\}$ is not satisfiable, and thus it is inconsistent.*

Proof. Assume there exists a relational model $M = (W, R, V)$, and a world $s_1 \in W$, that satisfies all the formulas in Γ. Then $M, s_1 \vDash \bigcirc g$, $M, s_1 \vDash \bigcirc(g \to t)$, $M, s_1 \vDash \neg g \to \bigcirc\neg t$ and $M, s_1 \vDash \neg g$.

From $M, s_1 \vDash \neg g$ and $M, s_1 \vDash \neg g \to \bigcirc\neg t$ we deduce $M, s_1 \vDash \bigcirc\neg t$. By seriality we know there is $s_2 \in W$ such that $s_1 R s_2$. By the definition of $\bigcirc$ we know $M, s_2 \vDash \neg t$, $M, s_2 \vDash g$, $M, s_2 \vDash g \to t$, which is a contradiction.

The above shows that Γ is not satisfiable, definition 7. $\square$

Exercise 9. Show that the second and the third formalisation of Chisholm's paradox do not meet the requirement of logical independency.

7 The "Andersonian" reduction

The main purpose of this section is to describe an embedding (or translation) of SDL into alethic modal logic, and to prove its soundness and faithfulness. It is based on Anderson [1]'s suggestion that a statement like "it ought to be that ϕ" can be analysed as "if $\neg\phi$, then necessarily v", where v is a propositional constant in the language of the underlying logic expressing that there has been a violation of the norms, or that a bad state-of-affairs has occurred.

7.1 System $\mathbf{K}^v$

Definition 11. *The language of $\mathbf{K}^v$ is generated by the following BNF:*

$$\phi ::= p \mid v \mid \neg\phi \mid \phi \wedge \phi \mid \Box\phi$$

where $p \in \mathbb{P}$. v is a designated propositional constant read as "a violation has occurred." $\Box\phi$ is read as "necessarily, ϕ". We use the same derived propositional connectives as before. In addition, $\Diamond$ (for possibility) is introduced as the dual of $\Box$: $\Diamond\phi$ is a shorthand for $\neg\Box\neg\phi$.

Definition 12. *An Anderson model is a tuple $M = (W, S, B, V)$ such that*

1. *W is a (non-empty) set of states*
2. *$S \subseteq W \times W$ is a binary relation over W. It is understood as a relation of alethic accessibility: sSt (or, alternatively, $(s,t) \in S$) says that t is accessible from s, or that t is a successor of s.*
3. *$B \subseteq W$ is a set of "bad" states. It is required that,*

$$(\forall s \in W)(\exists t \in W)(sSt \,\&\, t \notin B)$$

   Intuitively: for every state, there is an accessible one that is good.
4. *V is as before.*

Definition 13 (Satisfaction)**.** *Given an Anderson model $M = (W, S, B, V)$ and a state $s \in W$, we define the satisfaction relation $M, s \vDash \phi$ using the same clauses as before plus:*

1. *$M, s \vDash v$ iff $s \in B$.*
2. *$M, s \vDash \Box\phi$ iff, for all $t \in W$, if sSt then $M, t \vDash \phi$.*

Definition 14. *$\mathbf{K}^v$ is the proof system obtained by supplementing so-called system $\mathbf{K}$ (after Kripke) with the axiom schema (AV).*

$\vdash \phi$, *where ϕ is a propositional tautology*	(PL)
$\vdash \Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$	($\Box$-K)
If $\vdash \phi$ *and* $\vdash \phi \to \psi$ *then* $\vdash \psi$	(MP)
If $\vdash \phi$ *then* $\vdash \Box\phi$	($\Box$-Nec)
$\vdash \Diamond\neg v$	(AV)

Intuitively, (AV) says that the bad thing, v, is avoidable.

Theorem 6 (Soundness and completeness, strong version)**.** *$\mathbf{K}^v$ is strongly sound and strongly complete with respect to the class of Anderson models.*

Proof. Soundness is a matter of showing that the axiom schemas of $\mathbf{K}^v$ are valid in the class of Anderson models, and that the rule schemas preserve validity in the class of Anderson models. We just show (AV). Let M be an Anderson model, and s be a state in M. By condition 3 in Definition 12, there is some t such that sSt and $t \notin B$. By the truth-conditions for v, $t \not\models v$, and so $t \models \neg v$. By the truth-conditions for $\Diamond$, $s \models \Diamond\neg v$. Hence, $\models \Diamond\neg v$.

Completeness is shown using canonical models. $\square$

7.2 Embedding D into $\mathbf{K}^v$

Definition 15. *An embedding of* **D** *into* $\mathbf{K}^v$ *is a function* τ *mapping each formula* ϕ *in the language of* **D** *into a formula* $\tau(\phi)$ *in the language of* $\mathbf{K}^v$. τ *is defined inductively as follows:*

$$\tau(p) = p \tag{1}$$

$$\tau(\neg\phi) = \neg\tau(\phi) \tag{2}$$

$$\tau(\phi \wedge \psi) = \tau(\phi) \wedge \tau(\psi) \tag{3}$$

$$\tau(\bigcirc\phi) = \Box(\neg\tau(\phi) \to v) \tag{4}$$

Theorem 7 (Soundness of τ)**.** *For every formula* ϕ *in the language of* **D**, *if* $\vdash_{\mathbf{D}} \phi$, *then* $\vdash_{\mathbf{K}^v} \tau(\phi)$.

Proof. By induction on the length of a proof. $\square$

Lemma 1. *For every relational model* $M = (W, R, V)$, *there is an Anderson model* $M' = (W, S, B, V)$ *(with* W *and* V *the same) such that:*

$$\textit{For all states } s, M, s \models \phi \textit{ iff } M', s \models \tau(\phi) \tag{5}$$

Proof. Put $S = R$ and $B = \{s : s \in W \ \&\ \neg\exists t\, tRs\}$. It follows that

$$sRt \text{ iff: } sSt \ \&\ t \notin B \tag{6}$$

Because R is serial, S and B meet the property required of them in Definition 12. The proof of (5) is by induction on the structure of ϕ. $\square$

Theorem 8 (Faithfulness of τ)**.** *For every formula* ϕ *in the language of* **D**, *if* $\vdash_{\mathbf{K}^v} \tau(\phi)$, *then* $\vdash_{\mathbf{D}} \phi$.

Proof. We show the contrapositive. Suppose $\not\vdash_{\mathbf{D}} \phi$. By the completeness theorem for **D**, there exists a relational model $M = (W, R, V)$, and a state s in M, such that $M, s \not\models \phi$. By Lemma 1, there exists an Anderson model $M' = (W, S, B, V)$, with s in M', and such that $M', s \not\models \tau(\phi)$. By the soundness theorem for $\mathbf{K}^v$, $\not\vdash_{\mathbf{K}^v} \tau(\phi)$. $\square$

7.3 Chisholm's scenario revisited

Using Anderson's reduction, and allowing oneself some liberty, one may formalize Chisholm's scenario as follows:

(A_4) $\bigcirc g$
(B_4) $\Box(g \to \bigcirc t)$
(C_4) $\Box(\neg g \to \bigcirc\neg t)$
(D_4) $\neg g$

The formula (A_4) must be understood as $\Box(\neg g \to v)$. Same for the formulas prefixed with $\bigcirc$ appearing in (B_4) and (C_4).

Proposition 3. *The set* $\Gamma = \{A_4, B_4, C_4, D_4\}$ *is satisfiable in an Anderson model, and thus it is consistent in* $\mathbf{K}^v$.

Proof. Put $M = (W, S, B, V)$ with $W = \{s_1, s_2, s_3\}$, s_1Ss_1, s_1Ss_2, s_2Ss_3, s_3Ss_3, $B = \{s_1\}$, $V(g) = \{s_2, s_3\}$, $V(t) = \{s_1, s_3\}$ and $V(v) = \{s_1\}$. This is illustrated with Figure 2, where the content of the alethic accessibility relation S is indicated by dashed arrows.

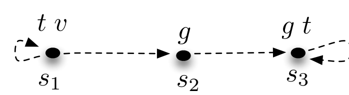


Figure 2: Chisholm's example–Anderson

The property required of S and B in Definition 12 is met. We have:

1. $s_1 \models \bigcirc g$
2. $s_1 \models \Box(g \to \bigcirc t)$
3. $s_1 \models \Box(\neg g \to \bigcirc\neg t)$
4. $s_1 \models \neg g$

$\square$

Exercise 10. Suppose Anderson's reduction schema is given the form of a pair of definitions (rather than that of an embedding function): $\bigcirc\phi =_{def} \Box(\neg\phi \to v)$; $P\phi =_{def} \Diamond(\phi \wedge \neg v)$. Show that all the axiom and rule schemas of **D** are derivable in $\mathbf{K}^v$.

Exercise 11. Give the relational model from which the Anderson model described in Figure 2 could have been derived. Hint: use backwards the construction given in the proof of Lemma 1.

Neighborhood semantics

There is a generalisation of the relational semantics, Chellas [4]'s minimal deontic logic, in terms of minimal or neighborhood models.

Definition 16 (Minimal model)**.** *A minimal model* $M = (W, N, V)$ *is a structure where* W *and* V *are as before, and* N*, called a "neighborhood" function, is a function assigning to each state* $s \in W$ *a set of subsets of W (i.e.* $N(s) \subseteq 2^W$ *for each* $s \in W$*).*

Definition 17 (Satisfaction)**.** *Given a minimal model $M = (W, N, V)$, and a world $s \in W$, we define the satisfaction relation $M, s \vDash \phi$ as before, except for deontic formulas, where:*

$$M, s \vDash \bigcirc\phi \textit{ iff } ||\phi|| \in N(s)$$

Here $||\phi||$ is the truth-set of ϕ, viz the set $\{t \in W : M, t \vDash \phi\}$.

Validity and consequence are defined as in the relational semantics.

The neighborhood approach allows for an extra degree of freedom here. The obligation operator as defined in Definition 17 is very weak. But extra constraints may be placed on $N(s)$ in order to make the operator validate more formulas, as one thinks fit. When sufficiently strong constraints are placed on $N(s)$, one obtains the relational semantics back.

For an illustration, consider Chellas' minimal deontic logic in [4], It has the law OD, but not the law OD*:

$$\neg \bigcirc \bot \qquad \text{(OD)}$$
$$\neg \bigcirc (\varphi \wedge \neg\varphi) \qquad \text{(OD}^\star\text{)}$$

OD expresses the seemingly uncontroversial principle "ought" implies "can". OD* rules out the possibility of conflicts between obligations, which seems counter-intuitive. In **D** and stronger systems, we have $\vdash (\neg \bigcirc \bot) \leftrightarrow (\neg \bigcirc (\varphi \wedge \neg\varphi))$. This makes it impossible to distinguish between OD* and OD, and to have one without the other.

Definition 18. *Chellas' minimal deontic logic consists of the following axiom schemata and rule schema:*

$$\textit{All tautologies of propositional logic} \qquad \text{(PL)}$$
$$\neg \bigcirc \bot \qquad \text{(OD)}$$
$$\textit{If } \vdash \phi \to \psi \textit{ then } \vdash \bigcirc\phi \to \bigcirc\psi \qquad \text{(ROM)}$$

The relevant constraints to be placed on N are:

$$\textit{If } U \subseteq V \textit{ and } U \in N(s) \textit{ then } V \in N(s) \qquad \text{(closure under superset)}$$
$$\emptyset \notin N(s) \qquad \text{(no-absurd)}$$

Exercise 12.
1. Show that ROM holds in view of the property of closure under superset.
2. Work out the evaluation rules for P and F in a minimal model.

8 Notes

Why should a given language be, of all the sets closed under the rules, the smallest one? This is because nothing else should be in it except what can be formed by repeated applications of the rules.

Axioms are particular formulas. It is common to specify them by giving axiom schemata. An axiom scheme is a pattern, and any formula matching that pattern is an axiom. When axiom schemata are used, a proof system will have a finite number of axiom schemata but infinitely many axioms. An alternative method is to specify a finite number of axioms and adopt substitution of formulas as an explicit rule. This tends to be more complicated, and we follow the axiom scheme approach.

The name **D** is given to the system described in subsection 4.1 in Hanson [7] and van Fraassen [9]. **D** is called **D*** in Chellas [4], **DL** in Åqvist [8], and **OK**$^+$ in Åqvist [3]. The names **DS4**, **DS5** and **DM** are given to the systems described in subsection 5 in H anson [7]. They are derived from the alethic modal logics **M**, **S4** and **S5**, to which they are similar.

What is a paradox? In the mathematical or strong sense, a paradox is a contradiction. Russell's paradox in set-theory is of this type. In the linguistic or weak sense, a paradox is something predicted by the theory, which a native speaker would not say. Example from propositional logic: the rule of $\vee$-introduction.

In deontic logic, we had paradoxes of both types. Chisholm's paradox is closer to the strong sense. The rule ($\Box$-Nec) is a paradox in the second sense.

Are paradoxes good or bad? As for paradoxes of the first type, usually people agree that they are bad, and that they should be resolved/avoided. Once spotted, they give rise to new developments, as is the case with the Chisholm example. As for paradoxes of the second type, some have argued that they can be explained away: the inference is only pragmatically odd in ways that are independently predictable by any adequate theory of the pragmatics of deontic language.

Anderson does not use the translation function τ, nor does he give a semantics to his system. He formulates his reduction as we do in Exercise 10. Anderson's own formulation is simpler, but it does not say anything about the status of the non-theorems of **D**: there is no guarantee that some will not be theorems of $\mathbf{K}^v$, and so there is guarantee that one will not get the laws of **D**, and some more. Needed is to take a detour through the embedding function τ. The fact that the embedding is faithful, theorem 8, tells us that the alethic counterpart of a non-theorem of **D** is also a non-theorem of the corresponding alethic modal logic.

In [2] , Anderson proposed that his reduction be worked out in the framework of the so-called logic **R** of relevant implication. This proposal is examined in Goble [6].

References

[1] A. R. Anderson. A reduction of deontic logic to alethic modal logic. *Mind*, 67(265):100–103, 1958.

[2] A. R. Anderson. Some nasty problems in the formal logic of ethics. *Nos*, 1(4):345–360, 1967.

[3] 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.

[4] B. F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, Cambridge, 1980.

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

[6] L. Goble. The Andersonian reduction and relevant deontic logic. In B. Byson and J. Woods, editors, *New Studies in Exact Philosophy: Logic, Mathematics and Science Proceedings of the 1999 Conference of the Society of Exact Philosophy*, pages 213–246. Hermes Science Publications, Paris, 1999.

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

[8] L. Åqvist. Good samaritans, contrary-to-duty imperatives, and epistemic obligations. *Nos*, 1(4):361–379, 1967.

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

Solutions

Exercise 1

Exercise 2.

(1) $s \vDash \phi \vee \psi$ iff $s \vDash \phi$ or $s \vDash \psi$.
$s \vDash \phi \to \psi$ iff $s \nvDash \phi$ or $s \vDash \psi$.

(2) $s \vDash P\phi$ iff there is some $t \in W$ such that sRt and $t \vDash \phi$.
$s \vDash F\phi$ iff for all $t \in W$, if sRt then $t \nvDash \phi$.

Exercise 3. Validity is a limiting case of consequence in the sense that a formula is valid iff it is the consequence of the empty set. Formally, $\vDash \phi$ iff $\emptyset \vDash \phi$.

Exercise 5. We give the derivations in a compressed form.

1. $\vdash (\neg p \wedge \neg q) \to \neg p$ PL
2. $\vdash \bigcirc((\neg p \wedge \neg q) \to \neg p)$ Nec, 1
3. $\vdash \bigcirc(\neg p \wedge \neg q) \to \bigcirc \neg p$ K, MP, 2
4. $\vdash \neg \bigcirc \neg p \to \neg \bigcirc \neg (p \vee q)$ PL, 3
5. $\vdash Pp \to P(p \vee q)$ 4

1. $\vdash (p \wedge q) \to (\neg q \to \neg p)$ PL
2. $\vdash \bigcirc((p \wedge q) \to (\neg q \to \neg p))$ Nec,1
3. $\vdash \bigcirc(p \wedge q) \to \bigcirc(\neg q \to \neg p)$ K, MP, 2
4. $\vdash \bigcirc(\neg q \to \neg p) \to (\bigcirc \neg q \to \bigcirc \neg p)$ K
5. $\vdash \bigcirc(p \wedge q) \to (\bigcirc \neg q \to \bigcirc \neg p)$ PL, 3,4
6. $\vdash \bigcirc(p \wedge q) \wedge \neg \bigcirc \neg p) \to \neg \bigcirc \neg q)$ PL, 5
7. $\vdash (\bigcirc(p \wedge q) \wedge Pp) \to Pq$ 7

Exercise 6

(1) We show (Nec) and (D).

(Nec). Assume $\models \phi$. Let M, and s in M, be such that $s \models \phi$. Let t in M be such that sRt. From the opening assumption, $t \models \phi$. By the satisfaction condition for $\bigcirc$, $s \models \bigcirc\phi$. Hence, $\models \bigcirc\phi$.

(D). Let M, and s in M, be such that $s \models \bigcirc\phi$. By the condition of seriality, there is some t in M such that sRt. By the satisfaction condition for $\bigcirc$, $t \models \phi$. By the satisfaction condition for P, $s \models P\phi$. Hence, $s \models \bigcirc\phi \to P\phi$. This shows that $\models \bigcirc\phi \to P\phi$.

(2) $M = (W, R, V)$ with $W = \{s_1, s_2\}$, $R = \emptyset$, $V(p) = W$.

Exercise 7

Γ inconsistent $\Leftrightarrow \Gamma \vdash \bot$ (by def)
$\Leftrightarrow \Gamma \models \bot$ (by soundness and completeness)
$\Leftrightarrow \Gamma$ not satisfiable (by def)

Exercise 8

1. $\vdash p \to (\bigcirc p \to p)$ PL
2. $\vdash \bigcirc p \to \bigcirc(\bigcirc p \to p)$ Nec, K, 1
3. $\vdash \neg \bigcirc p \to (\bigcirc p \to p)$ PL
4. $\vdash \bigcirc \neg \bigcirc p \to \bigcirc(\bigcirc p \to p)$ Nec, K, 3
5. $\vdash \neg \bigcirc p \to \bigcirc \neg \bigcirc p$ $\bigcirc$5
6. $\vdash \neg \bigcirc p \to \bigcirc(\bigcirc p \to p)$ PL, 4,5
7. $\vdash \bigcirc(\bigcirc p \to p)$ PL, 2,6

Exercise 10

For K:

1. $\vdash (\neg(\phi \to \psi) \to v) \to ((\neg\phi \to v) \to (\neg\psi \to v))$ PL
2. $\vdash \Box(\neg(\phi \to \psi) \to v) \to (\Box(\neg\phi \to v) \to \Box(\neg\psi \to v))$
$\Box$-Nec, $\Box$-K, 1

For Nec:

1. $\vdash \phi$ Assumption
2. $\vdash \phi \to (\neg\phi \to v)$ PL
3. $\vdash \neg\phi \to v$ MP, 1,2
4. $\vdash \Box(\neg\phi \to v)$ $\Box$-Nec, 3

For D, we use the law $\vdash_{\mathbf{K}} (\Box\phi \wedge \Diamond\psi) \to \Diamond(\phi \wedge \psi)$, and we allow replacement of a formula by a logically equivalent one:

1. $\vdash (\Box(\neg v \to \phi) \wedge \Diamond\neg v) \to \Diamond((\neg v \to \phi) \wedge \neg v)$
2. $\vdash (\Box(\neg v \to \phi) \wedge \Diamond\neg v) \to \Diamond(\phi \wedge \neg v)$ Repl. , 1
3. $\vdash \Diamond\neg v \to (\Box(\neg v \to \phi) \to \Diamond(\phi \wedge \neg v))$ PL, 2
4. $\vdash \Diamond\neg v$ AV
5. $\vdash \Box(\neg v \to \phi) \to \Diamond(\phi \wedge \neg v)$ MP, 3,4

Exercise 11