Handout Lecture 2: Dyadic Standard Deontic Logic

Leendert van der Torre, Xavier Parent, Xin Sun
University of Luxembourg

1 Chisholm's paradox

The note of standard deontic logic (SDL) [11] discusses Chisholm's paradox:

(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.

The set of sentences $A4 - D4$ is a consistent representation in SDL (extended with an alethic modality), such that none of the sentences can be derived from the other ones.

A_4 $\bigcirc g$

B_4 $\Box(g \rightarrow \bigcirc t)$

C_4 $\Box(\neg g \rightarrow \bigcirc \neg t)$

D_4 $\neg g$

A drawback of the SDL representation $A_4 - D_4$ is that it does not represent that ideally, the man goes to the assistance and tells. Moreover, there does not seem to be a similar solution for the following variant of the scenario:

(AB) It ought to be that Jones goes to the assistance of his neighbours and he tells them he is coming.

(C) If Jones doesn't go to the assistance of his neighbours, then he ought not tell them he is coming.

(D) Jones does not go to their assistance.

Moreover, SDL only distinguishes between ideal and non-ideal worlds, whereas many ethical dilemmas are based on trade-offs between violations, or multiple levels of violation. The challenge is thus how to extend the semantics of SDL in this regard. For example, one can add distinct modal operators for primary and secondary obligations, where a secondary obligation is a kind of repair obligation. From $A_5 - D_5$ we can only derive $\bigcirc_1 t \wedge \bigcirc_2 \neg t$, but this is not a contradiction.

A_5 $\bigcirc_1 g$

B_5 $\bigcirc_1 (g \rightarrow t)$

C_5 $\neg g \rightarrow \bigcirc_2 \neg t$

D_5 $\neg g$

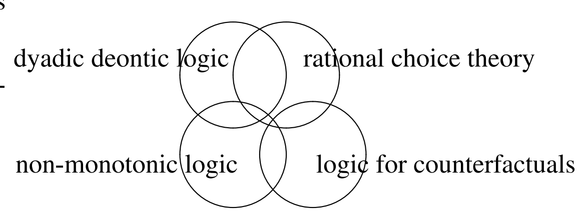
However, it may not always be easy to distinguish primary from secondary obligations, for example it may depend on the context whether an obligation is primary or secondary. Many people therefore put as an additional requirement for a solution of the paradox that **B** and **C** are represented in the same way. Finally, the distinction between $\bigcirc_1$ and $\bigcirc_2$ is insufficient for extensions of the paradox that seem to need also operators like $\bigcirc_3$, $\bigcirc_4$, etc, such as the following **E** and **F**.

(E) If Jones does not go to the assistance and he tell them he is coming, then he should apologize afterwards.

(F) If Jones does not go to the assistance and he tell them he is coming and he does not apologize afterwards, then ...

2 Dyadic standard deontic Logic

Dyadic standard deontic logic (DSDL) provides an alternative representation of Chisholm's paradox. The following figure visualizes a bird's eye view on DSDL and various related formalisms.



Inspired by rational choice theory in the sixties, preference-based semantics for DSDL became popular at the end of the sixties (by, for example, Danielsson [3], Hansson [5], van Fraassen [4], Spohn [13]). Various other formal approaches build on rational choice theory and dyadic deontic logic, such as:

- rational choice theory: Sen [12].

- logic for counterfactuals: Lewis [7].
- non-monotonic logic: KLM systems [6] .

For more information on the interplay between these areas, see Makinson's comparison [8].

2.1 Language

Definition 1. *Given a set Φ of propositional letters. The language of dyadic standard deontic logic $\mathfrak{L}_D$ is the smallest set such that:*

1. $\Phi \subseteq \mathfrak{L}_D$.
2. *if* $\phi \in \mathfrak{L}_D$*, then* $\neg\phi, \Box\phi \in \mathfrak{L}_D$.
3. *if* $\phi \in \mathfrak{L}_D$ *and* $\psi \in \mathfrak{L}_D$*, then* $\phi \wedge \psi \in \mathfrak{L}_D$.
4. *if* $\phi, \psi \in \mathfrak{L}_D$*, then* $\bigcirc(\phi/\psi) \in \mathfrak{L}_D$.

We let $\bigcirc\phi := \bigcirc(\phi/\top)$*,* $P(\phi/\psi) := \neg\bigcirc(\neg\phi/\psi)$ *and* $\Diamond\phi := \neg\Box\neg\phi$*. Other boolean cases are defined as usual.*

$\mathfrak{L}_D$ is equivalently represented by the following BNF: for p range over Φ,

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

The intended reading of $\Box\phi$ is "ϕ is settled as true", $\bigcirc(\phi/\psi)$ is "ϕ is obligatory, given ψ". $P(\phi/\psi)$ is "ϕ is permitted, given ψ", and $\Diamond\phi$ is "possibly ϕ".

⚠ Iteration of deontic modality is allowed. For example $\bigcirc(B/\bigcirc(B/A) \wedge A)$ is a formula in $\mathfrak{L}_D$.

if one fails to pay his taxes, then one ought to be fined. ⊣

2.2 Semantics

In this section we introduce the preference-based semantics as defined by Åqvist [2].

Definition 2. *A preference model* $M = (W, \geq, V)$ *is a tuple where:*

- *W is a non-empty set of possible worlds.*
- *$\geq$ is a reflexive, transitive relation over W satisfying the following limitedness requirement: if $||\phi|| \neq \emptyset$ then $\{x \in ||\phi|| : (\forall y \in ||\phi||) x \geq y\} \neq \emptyset$. Here $||\phi|| = \{s \in W : M, s \vDash \phi\}$.* [1]

  $s > s'$ is short for $s \geq s'$ and $s' \not\geq s$.
- *V is a standard propositional valuation such that for every propositional letter p, $V(p) \subseteq W$.*

$s \geq s'$ is understood as s is at least as good as s'.

[1] Unlike Åqvist, we do not require $\geq$ to be connected. Parent [10] shows that connectedness does not modify the set of validities.

Definition 3 (Satisfaction)**.** *Given a preference model $M = (W, \geq, V)$ and a world $s \in W$, we define the satisfaction relation $M, s \vDash \phi$ by induction on the structure of ϕ using the following clauses:*

- *$M, s \vDash p$ iff $s \in V(p)$.*
- *$M, s \vDash \neg\phi$ iff not $M, s \vDash \phi$.*
- *$M, s \vDash \phi \wedge \psi$ iff $M, s \vDash \phi$ and $M, s \vDash \psi$.*
- *$M, s \vDash \Box\phi$ iff for all $s' \in W$, $M, s' \vDash \phi$.*
- *$M, s \vDash \bigcirc(\psi/\phi)$ iff for all s', if $s' \in Best(||\phi||)$, then $M, s' \vDash \psi$. Here for a set of worlds Q, $Best(Q) = \{s \in Q : s \geq s', \text{for all } s' \in Q\}$.*

$M, s \vDash \phi$ is read as "world s satisfies ϕ in model M".

Definition 4 (Validity)**.** *A formula ϕ is said to be valid, if for all preference model M and all state s in M, $M, s \vDash \phi$.*

Definition 5 (Consequence)**.** *Given a set Γ of formulas, we say that ϕ is a consequence of Γ (written as $\Gamma \vDash \phi$) iff for all model M, all state s in M, if $M, s \vDash \psi$ for all $\psi \in \Gamma$, then $M, s \vDash \phi$.*

The obligations involved in Chisholm's paradox can be represented by a preference ordering, for example:

$$g \wedge t > g \wedge \neg t > \neg g \wedge \neg t > \neg g \wedge t$$

Extensions like **E** and **F** can be incorporated by refining the ordering.

Chisholm's quartet can be represented in DSDL as follows:

A_6 $\bigcirc g$

B_6 $\bigcirc(t|g)$

C_6 $\bigcirc(\neg t|\neg g)$

D_6 $\neg g$

The first example shows how to verify whether a deontic formula is satisfied in a model .

Example 2.1. Let $M = (W, \geq, V)$ be where $W = \{s_1, s_2\}$, $s_1 > s_2$ and $V(g) = \{s_1\}$. Show that $\bigcirc g$ is satisfied in M.

Solution. $M, s_1 \vDash \bigcirc g$ because $Best(||\top||) = \{s_1\}$ and $M, s_1 \vDash g$. ⊣

Exercise 2.1. Let $M = (W, \geq, V)$ such that $W = \{s_1, s_2, s_3\}$, $s_1 > s_2 > s_3$, $V(g) = \{s_2, s_3\}$ and $V(t) = \{s_2\}$. Show that $\bigcirc(t|g)$ is satisfiable in M.

The second example illustrates satisfiability of a set S of formulas. This means constructing a model, and a world in it, at which all the formulas in S are true.

Example 2.2. Show that $\{\bigcirc g, \bigcirc(t|g)\}$ is satisfiable.

Solution. Let $M = (W, \geq, V)$ such that $W = \{s_1, s_2\}$, $s_1 > s_2$, $V(g) = \{s_1\}$, $V(t) = \{s_1\}$. We have $M, s_1 \vDash \bigcirc g$ and $M, s_1 \vDash \bigcirc(t|g)$. ⊣

Exercise 2.2. Show that $\{\bigcirc g, \bigcirc(t|g), \bigcirc(\neg t|\neg g), \neg g\}$ is satisfiable.

The third example illustrates the notion of non-validity of a formula.

Example 2.3. Show that $\not\vDash \neg g \to \bigcirc(t|g)$.

Solution. We need to build a model such that for some world in the model, $\neg g$ is satisfied but $\bigcirc(t|g)$ is not. Let $M = (W, \geq, V)$ such that $W = \{s_1, s_2, s_3\}$, $s_1 > s_2$, $V(g) = \{s_1, s_2\}$, $V(t) = \{s_2\}$. It can be verified that $M, s_3 \vDash \neg g$ and $M, s_3 \not\vDash \bigcirc(t|g)$. ⊣

Exercise 2.3. Show that $\not\vDash \bigcirc g \to \bigcirc(\neg t|\neg g)$.

The fourth example illustrates the notion of validity of a formula.

Example 2.4. Show that $\vDash \bigcirc(t \vee \neg t|g)$.

Solution. Let M be an arbitrary model. We need to show that $Best(||g||) \subseteq ||t \vee \neg t||$. Let $s \in Best(||g||)$. By the evaluation rules for the propositional connectives, it follows at once that $M, s \vDash t \vee \neg t$, as required. ⊣

Exercise 2.4. Show that $\vDash \bigcirc(g|g)$.

Here is a list of valid and invalid formulas:

$$\vDash \Box(\phi \to \psi) \to (\Box\phi \to \Box\psi) \quad \text{(K)}$$

$$\vDash \Box\phi \to \Box\Box\phi \quad \text{(4)}$$

$$\vDash \neg\Box\phi \to \Box\neg\Box\phi \quad \text{(5)}$$

$$\vDash \bigcirc(\psi \to \chi/\phi) \to (\bigcirc(\psi/\phi) \to \bigcirc(\chi/\phi)) \quad \text{(COK)}$$

$$\vDash \bigcirc(\psi/\phi) \to \Box \bigcirc (\psi/\phi) \quad \text{(Abs)}$$

$$\vDash \Box\psi \to \bigcirc(\psi/\phi) \quad \text{(CON)}$$

$$\vDash \Box(\phi \leftrightarrow \psi) \to (\bigcirc(\chi/\phi) \leftrightarrow \bigcirc(\chi/\psi)) \quad \text{(Ext)}$$

$$\vDash \bigcirc(\phi/\phi) \quad \text{(Id)}$$

$$\vDash \bigcirc(\chi/(\phi \wedge \psi)) \to \bigcirc((\psi \to \chi)/\phi) \quad \text{(C)}$$

$$\vDash \Diamond\phi \to (\bigcirc(\psi/\phi) \to P(\psi/\phi)) \quad \text{(D*)}$$

$$\vDash (P(\psi/\phi) \wedge \bigcirc((\psi \to \chi)/\phi)) \to \bigcirc(\chi/(\phi \wedge \psi)) \quad \text{(S)}$$

$$\not\vDash \bigcirc(q/p) \to \bigcirc(q/p \wedge r) \quad \text{(SI)}$$

$$\not\vDash (\bigcirc(q/p) \wedge \bigcirc(r/q)) \to \bigcirc(r/p) \quad \text{(T)}$$

$$\not\vDash \bigcirc(q/p) \wedge p \to \bigcirc q \quad \text{(factual detachment)}$$

Example 2.5. Strengthening of the Antecedent is invalid:

$$\not\vDash \bigcirc(q/p) \to \bigcirc(q/p \wedge r)$$

Intuitively, this means obligations are ***defeasible*** (cf. non-monotonic logic).

We construct the following model: $M = (W, \geq, V)$ such that $W = \{w_1, w_2, w_3\}$, $w_1 > w_2 > w_3$, $V(p) = W$, $V(q) = \{w_1, w_3\}$, $V(r) = \{w_2, w_3\}$. It can be verified that $M, w_1 \not\vDash \bigcirc(q/p) \to \bigcirc(q/p \wedge r)$. ⊣

Exercise 2.5. Show that

$$\not\vDash (\bigcirc(\neg t/\neg g) \wedge \bigcirc(g/\neg t)) \to \bigcirc(g/\neg g)$$

Exercise 2.6. Show that

- $\not\vDash \bigcirc(q/p) \wedge p \to \bigcirc q$
- $\vDash \bigcirc(q/p) \wedge \Box p \to \bigcirc q$

2.3 Proof system

In the 20th century, mathematical logic was developed in an attempt to answer a question concerning "truth" and "proof" in mathematics.

⋆ Can we give a rigorous definition of "proof" and use it to "prove" every "true mathematical statement"?

To make sense of this questions, careful concept formation is needed. What do we mean by:

- "mathematical statement"?
- "true mathematical statement"?
- "proof"?

In the deontic setting, those questions turn to

- What is "deontic statement"?
- What is "true deontic statement"?
- What is "proof"?
- Can we "prove" every "true deontic statement"?

In DSDL, the answers to those questions are the following:

- A deontic statements is a formula $\phi \in \mathfrak{L}_D$.
- True deontic statements are **valid** formulas.
- A proof is a **derivation** to be defined in Definition 6.
- We can prove every true deontic statement. This is Theorem 1 in the next subsection.

Definition 6. *Let* $\mathbf{X}$ *be a given proof system.* $\mathbf{X}$ *is characterized by axioms* $Ax_1, Ax_2, ..., Ax_n$ *and rules* $Ru_1, Ru_2, ...Ru_k$*, where each rule* $Ru_j (j \leq k)$ *is of the form "From* $\phi_1, \ldots \phi_{j_{ar}}$ *infer* ϕ_j*". We call* j_{ar} *the arity of the rule. Then, a derivation of* ϕ *in* $\mathbf{X}$ *is a finite sequence* $\phi_1, \ldots \phi_m$ *of formulas such that:*

1. $\phi_m = \phi$;
2. *every* ϕ_i *in the sequence is*
   (a) *either an* ***instance*** *of one of the axioms* $Ax_1, Ax_2, ...Ax_n$
   (b) *or else the result of the application of one of the rules* $Ru_j (j \leq k)$ *to* j_{ar} *formulas in the sequence that appear before* ϕ_i.

If there is a derivation for ϕ *in* $\mathbf{X}$ *we write* $\vdash_{\mathbf{X}} \phi$*, and we say that* ϕ *is a theorem in* $\mathbf{X}$.

For a set of formulas Γ*, we say* ϕ *is derivable from* Γ *in system* $\mathbf{X}$ *(notation:* $\Gamma \vdash_{\mathbf{X}} \phi$*) if there are formulas* $\psi_1, \ldots, \psi_n \in \Gamma$ *such that* $\vdash_{\mathbf{X}} (\psi_1 \wedge \ldots \wedge \psi_n) \rightarrow \phi$.

When it is clear what system is intended, we drop the subscript $\mathbf{X}$, and just write $\vdash \phi$ or $\Gamma \vdash \phi$.

Definition 7. *Åqvist [2]'s system* $\mathbf{G}$ *consists of the following axioms and rules (labels are from [10]):*

Axioms schemes:

$$\text{All tautologies of propositional logic} \quad \text{(PL)}$$
$$\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi) \quad \text{(K)}$$
$$\Box\phi \rightarrow \Box\Box\phi \quad \text{(4)}$$
$$\neg\Box\phi \rightarrow \Box\neg\Box\phi \quad \text{(5)}$$
$$\bigcirc(\psi \rightarrow \chi/\phi) \rightarrow (\bigcirc(\psi/\phi) \rightarrow \bigcirc(\chi/\phi)) \quad \text{(COK)}$$
$$\bigcirc(\psi/\phi) \rightarrow \Box\bigcirc(\psi/\phi) \quad \text{(Abs)}$$
$$\Box\psi \rightarrow \bigcirc(\psi/\phi) \quad \text{(Nec)}$$
$$\Box(\phi \leftrightarrow \psi) \rightarrow (\bigcirc(\chi/\phi) \leftrightarrow \bigcirc(\chi/\psi)) \quad \text{(Ext)}$$
$$\bigcirc(\phi/\phi) \quad \text{(Id)}$$
$$\bigcirc(\chi/(\phi \wedge \psi)) \rightarrow \bigcirc((\psi \rightarrow \chi)/\phi) \quad \text{(Sh)}$$
$$\Diamond\phi \rightarrow (\bigcirc(\psi/\phi) \rightarrow P(\psi/\phi)) \quad \text{(D}^\star\text{)}$$
$$(P(\psi/\phi) \wedge \bigcirc((\psi \rightarrow \chi)/\phi)) \rightarrow \bigcirc(\chi/(\phi \wedge \psi)) \quad \text{(Sp)}$$

Rules:

$$\text{If } \vdash \phi \text{ and } \vdash \phi \rightarrow \psi \text{ then } \vdash \psi \quad \text{(MP)}$$
$$\text{If } \vdash \phi \text{ then } \vdash \Box\phi \quad \text{(N)}$$

From now onwards for the sake of conciseness brackets are omited when no confusion can arise.

Remark 2.1. The axioms and the rule governing $\Box$ are those of the modal system known as S5.

Remark 2.2. (Sp) is the distinctive axiom of $\mathbf{G}$. It appears in Spohn's own axiomatization of Hansson's system DSDL3 (see [13]). (Sp) is equivalent to the so-called principle of rational monotony used in non-monotonic logic [6]:

$$P(\psi/\phi) \wedge \bigcirc(\chi/\phi) \rightarrow \bigcirc(\chi/\phi \wedge \psi) \quad \text{(RM)}$$

Example 2.6. In system $\mathbf{G}$, we have:

$$\{P(r/p), \bigcirc(q/p)\} \vdash \bigcirc(q/p \wedge r) \quad (1)$$
$$\{P(r/p), \bigcirc(q/p)\} \vdash \bigcirc(r \rightarrow q/p) \quad (2)$$

Below: the required derivations:
For 1:

1. $\vdash q \rightarrow (r \rightarrow q)$ PL
2. $\vdash \Box(q \rightarrow (r \rightarrow q))$ N, 1
3. $\vdash \Box(q \rightarrow (r \rightarrow q)) \rightarrow \bigcirc(q \rightarrow (r \rightarrow q)/p)$ Nec
4. $\vdash \bigcirc(q \rightarrow (r \rightarrow q)/p)$ MP, 2,3
5. $\vdash \bigcirc(q \rightarrow (r \rightarrow q)/p) \rightarrow (\bigcirc(q/p) \rightarrow \bigcirc(r \rightarrow q/p))$ COK
6. $\vdash \bigcirc(q/p) \rightarrow \bigcirc(r \rightarrow q/p)$ MP 4,5
7. $\vdash (P(r/p) \wedge \bigcirc(r \rightarrow q/p)) \rightarrow \bigcirc(q/p \wedge r)$ Sp
8. $\vdash \bigcirc(r \rightarrow q/p) \rightarrow (P(r/p) \rightarrow \bigcirc(q/p \wedge r))$ PL, 7
9. $\vdash \bigcirc(r/p) \rightarrow (P(r/p) \rightarrow \bigcirc(q/p \wedge r))$ PL, 8, 6
10. $\vdash (P(r/p) \wedge \bigcirc(r/p)) \rightarrow \bigcirc(q/p \wedge r)$ PL, 9

For 2:

1. $\vdash P(r/p) \wedge \bigcirc(q/p) \rightarrow \bigcirc(q/p \wedge r)$ RM
2. $\vdash \bigcirc(q/p \wedge r) \rightarrow \bigcirc(r \rightarrow q/p)$ Sh
3. $\vdash P(r/p) \wedge \bigcirc(q/p) \rightarrow \bigcirc(r \rightarrow q/p)$ PL 1,2

$\dashv$

Exercise 2.7. Show that

1. $\vdash \bigcirc(t/g) \rightarrow \bigcirc(g \rightarrow t)$
2. $\vdash P(p \wedge q) \rightarrow P(p/q)$

Exercise 2.8. Show that, for axiom Sp to be valid, $\geq$ must be transitive. Show that, for $\text{D}^\star$ to be valid, $\geq$ must be limited.

2.4 Completeness and decidability

Theorem 1 (Completeness [9])**.** $\Gamma \vdash \phi$ *iff* $\Gamma \models \phi$

Theorem 2 (Decidability [1])**.** *The derivability problem in* $\mathbf{G}$ *is decidable. That is, the question as to whether or not* $\Gamma \vdash \phi$ *can be answered in finitely many steps.*

References

[1] L. Åqvist. Three characterizability problems in deontic logic. In R. Demolombe and R. Hilpinen,

editors, *Proceedings of the 5th International Workshop on Deontic Logic in Computer Science (DEON 2000), Toulouse, France, January 20-22, 2000*, pages 15–41. ONERA, Informal Proceedings, 2000.

[2] 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.

[3] S. Danielsson. *Preference and Obligation: Studies in the Logic of Ethics*. Filosofiska foreningen, Uppsala, 1969.

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

[5] B. Hansson. An analysis of some deontic logics. In Hilpinen R. and Follesdal D., editors, *Deontic Logic: Introductory and Systematic Readings*, volume 33 of *Synthese Library*, pages 121–147. Springer Netherlands, 1971.

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

[7] D. Lewis. *Counterfactuals*. Blackwells, 1973.

[8] D. Makinson. Five faces of minimality. *Studia Logica*, 52(3):339–380, 1993.

[9] X. Parent. On the strong completeness of Åqvist's dyadic deontic logic G. In R. van der Meyden and L. van der Torre, editors, *Deontic Logic in Computer Science, 9th International Conference, DEON 2008, Luxembourg, Luxembourg, July 15-18, 2008. Proceedings*, volume 5076 of *Lecture Notes in Computer Science*, pages 189–202. Springer, 2008.

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

[11] X. Parent, L. van der Torre, and X. Sun. Handout lecture 1: Standard deontic logic. Unpublished notes.

[12] A. Sen. Choice functions and revealed preference. *Review of Economic Studies*, 38:307–317, 1971.

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