I. Introduction

1. Standard Deontic Logic (SDL)

The logic of reasoning about obligation and permission.

\[ \text{Op} \equiv \text{It ought to be that } p \]
\[ \text{Pp} \equiv \text{It is permitted that } p \quad \text{Pp} =_{df} \neg \neg p \]

a) Syntax of SDL

The system D, but with L replaced by O and M by P

Rules

US, MP, N

In particular: \( O \) for any tautology

Axioms

PC

(KO) \( O(p \supset q) \supset (Op \supset Oq) \)

(DO) \( Op \supset Pp, \text{ OR } \neg(Op \land O\neg p) \)

Two results of particular significance:

CO \( (OA \land OB) \supset O(A \land B) \) (Closure under composition)

MO \( O(A \land B) \supset (OA \land OB) \)

These follow from KO.

b) Semantics of SDL

- Sound and complete with respect to serial frames
- Interpretation:

  \( wRw' \) if \( w' \) is an “ideal” world accessible from \( w \) (all obligations met)
  Seriality: there is always at least one such world accessible from any \( w \)

  \( V(Op, w) = 1 \text{ iff p is true at all accessible ideal worlds.} \)
2. Ought-to-do (prescriptive) statements and ought-to-be (evaluative) statements

The former concern actions, the latter states of affairs.

Not every ought-to-be is equivalent to an ought-to-do:

It **ought to be** that nobody is hungry, but it doesn’t follow that any individual **ought to do** something that puts an end to hunger.

3. Paradoxes and puzzles

Many stem from problems in applying ought-to-be to the ought-to-do:

1) Ross’s paradox.

M: mail the letter  
B: burn the letter  
\[ P(M) \Rightarrow P(M \lor B) \]

2) Defeasible obligations.

F: eat with your fingers  
E: eat asparagus  
\[ O(\neg F); E \supset O(F) \]

3) Good Samaritan.

B: Arthur bandages the man he will murder a week hence  
M: Arthur murders a man a week hence  
\[ O(B); B \supset M \Rightarrow OM \text{ (makes use of OK)} \]

4) Contrary-to-duty obligations and conditional obligations.

V: Arthur visits his mother  
C: Arthur calls to tell his mother he won’t visit  
\[ O(V); \neg V \supset OC; O(V \supset \neg C); \neg V \Rightarrow OC \land O(\neg C) \]

4. Why the continued use of SDL and the evaluative ought-to-be?

1) Technical convenience:
   - universality of \textit{O} (takes any sentence as its complement)  
   - normal modal system (can lift results from alethic modal logic)

2) Philosophical argument:
   The ought-to-do should be subsumed by the ought-to-be:  
   “Fred ought to visit his mother” = “It ought to be that Fred visits his mother.”
**Horty’s approach**

- A new semantics for the **ought to do** is developed by building on a semantics for **agency** (due to Belnap) which is itself built on a semantics for **tense logic** (due to Prior).

- Entirely semantic: no proof theory

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Concept</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Agency:</td>
</tr>
<tr>
<td></td>
<td>[α stit: A] ≡ “α sees to it that A”</td>
</tr>
<tr>
<td>3</td>
<td>Obligation: 1^{st} attempt</td>
</tr>
<tr>
<td></td>
<td>O[α stit: A]</td>
</tr>
<tr>
<td></td>
<td>Argument (“Gambling problem”): this semantics fails</td>
</tr>
<tr>
<td>4</td>
<td>Obligation: 2^{nd} attempt</td>
</tr>
<tr>
<td></td>
<td>☐[α stit: A]</td>
</tr>
<tr>
<td></td>
<td>Employs simple ideas from decision theory</td>
</tr>
<tr>
<td>5</td>
<td>Extension to conditional obligation</td>
</tr>
</tbody>
</table>
II. Indeterminism and agency

1. Grammar

Atomic formulas: A, B, C, …

Connectives:
\( \land \) (conjunction)
\( \lnot \) (negation)
\( \lor \) (disjunction)
\( \supset \) (material implication)
\( \equiv \) (biconditional)

Temporal Operators:
\( PA: \) It was the case that A (at some point)
\( FA: \) It will be the case that A (at some point)
\( A: \) It is settled true that A / historically necessary that A
\( \Diamond A: \) It is historically possible that A

Agents:
\( \alpha, \beta, \ldots \)

Stit operators:
\[ [\alpha \textit{ distit: } A]: \alpha \text{ (deliberatively) sees to it that } A \]
\[ [\alpha \textit{ cstit: } A]: \alpha \text{ sees to it that } A \text{ (in sense of Chellas)} \]

Deontic operators:
\( OA: \) It ought to be that A (evaluative ought)
\( \Theta[\alpha \textit{ cstit: } A]: \alpha \text{ ought to see to it that } A \text{ (prescriptive ought)} \)
2. Branching time semantics

a) Basics

- Tree = set of moments \( m_1, m_2, \ldots \) (plays the role of \( W \))
- A relation \(<\) (or “earlier”’) on Tree:
  - transitive: if \( m_1 < m_2 \) and \( m_2 < m_3 \), then \( m_1 < m_3 \)
  - irreflexive: not \( m < m \)
  - tree property (no backward branching):
    if \( m_i < m_j \) and \( m_j < m_k \), then either \( m_i = m_j \) or \( m_i < m_j \) or \( m_j < m_i \).
- History \( h \) = maximal ordered set of moments
- \( H \) = set of all histories (also plays the role of \( W \))
- \( H_m \) = histories through \( m = \{ h / m \in h \} \)

**Both the moments and the histories function as worlds for evaluating the truth of modal formulas.

![Diagram](Fig 2.1)

Branching time:

1. Compatible with determinism: the actual history might be determined by laws of physics plus initial conditions.

2. Serious metaphysics or plausible meta-ethical presupposition for deliberating about obligations?
b) Basic Evaluation rules

As always, we want to assign values to all atomic statements A

**Initial proposal:** Assign V(p, m) for all m and define as usual for ¬A, A ∧ B, etc.

**Problem:** What is V(FA, m)? It depends which *history* you are talking about.

In Fig. 2.1: V(FA, m) is 1 on h₂, 0 on h₄ and h₅

**Modified proposal:** Assign V(p, m/h) for all *moment-history pairs* m/h.

**Notation 2:** Given a model <F, v>, write v(p) for the set of m-h pairs at which p is true. Can then say V(p, m/h) = 1 iff m/h ∈ v(p).

**Notation 3:** Instead of V, use turnstile notation. Write

\[ |A|_M^m = \text{set of histories } h \text{ through } m \text{ where } A \text{ comes out true on } M, \]

and

\[ M, m/h \ A \]

to indicate that A is true at m/h in model M. We identify the *proposition* \[|A|_M^m\] with the set of m/h pairs where A comes out true.

**Evaluation rules**

1. Atomic formulae:

\[ M, m/h \ A \iff m/h \in v(A) \text{ iff } h \in |A|_M^m \]

2. Basic operators

\[ M, m/h \ A \land B \iff M, m/h \ A \text{ and } M, m/h \ B \]

if h ∈ |A|_M^m and |B|_M^m

\[ M, m/h \ \neg A \iff M, m/h \ A \]

3. Temporal operators

\[ M, m/h \ PA \iff \text{there is an } m' \in h \text{ such that } m' < m \text{ and } M, m'/h \ A \]

\[ M, m/h \ FA \iff \text{there is an } m' \in h \text{ such that } m < m' \text{ and } M, m'/h \ A \]

\[ M, m/h \ A \iff \text{for all } h' \in H_m, \ M, m'/h' \ A \]

\[ M, m/h \ \lozenge A \iff \text{for some } h' \in H_m, \ M, m'/h' \ A \]

*Convention: Often omit M (“A is true at m/h”)

*Picture for A: A true at all histories through m.

**Remark:** Every pure P statement (about the past, with no embedded F operator) is settled-true or settled-false at m. (Follows from *no backward branching.*)
c) Agency

Agent: a set of individual agents \( \alpha, \beta, \ldots \) acting in time

Choices:
   i) Ignore intentionality, act descriptions, etc.
   ii) Ignore probability (truth of A must be guaranteed, not just highly probable)
   iii) Treat actions as momentary

\( \Rightarrow \) A choice by \( \alpha \) at \( m \) is simply constraining the histories that flow through \( m \)

- Histories through each moment are partitioned into “choice sets” for \( \alpha \) (for each \( \alpha \)).
  Each choice set is a cluster of histories between which \( \alpha \) cannot distinguish by a present choice. Two histories \( h \) and \( h' \) in the same choice set are said to be choice equivalent for \( \alpha \) at \( m \).

- We define a choice function: if \( m \in h \), then

  \[
  \text{Choice}_a^m(h) = \{ h' / m \in h' \text{ and } h, h' \text{ are choice equivalent for } \alpha \text{ at } m \}
  \]

- We also use \( \text{Choice}_a^m \) to denote the set of choice sets available to \( \alpha \) at \( m \).

Informally: choice sets are represented by dividing up a moment into ‘boxes’ of choice-equivalent histories. Typically, we suppress the reference to \( \alpha \) unless more than one agent is involved.

At \( m_1 \), there are three choice sets. \( h_1 \) and \( h_2 \) are choice-equivalent, as are \( h_4 \) – \( h_6 \).

Two constraints:
   1) Independence of agents
   2) No choice between undivided histories

A stit frame is a structure \(<\text{Tree}, <, \text{Agent}, \text{Choice}>\) satisfying the constraints.
d) Stit semantics

“α sees to it that A” means: the truth of A is guaranteed by a choice of α

i) \( \text{cstit} \)

\[
\begin{align*}
M, m/h & \quad [\alpha \text{ cstit: } A] \quad \text{iff for all } h' \in \text{Choice}_\alpha^m(h), \ M, m/h' \quad A \\
& \quad \text{iff Choice}_\alpha^m(h) \subseteq |A|_m^M
\end{align*}
\]

Informally: if α opts for the choice set that includes \( h \), then α guarantees the truth of A (since A is true no matter which history in the choice set we consider).

Picture (α closes the door):

\[
\begin{array}{ccccccc}
& D & D & D & \neg D & \neg D & \neg D \\
& h_1 & h_2 & h_3 & h_4 & h_5 & h_6
\end{array}
\]

[α cstit: D] at \( m/h_1 \) and \( m/h_2 \)
[α cstit: \neg D] at \( m/h_5 \) and \( m/h_6 \)
Neither at \( m/h_3 \) and \( m/h_4 \)

ii) \( \text{dstit} \)

\[
\begin{align*}
M, m/h & \quad [\alpha \text{ dstit: } A] \quad \text{iff} \\
& \quad (1) \text{ Positive condition: } \text{Choice}_\alpha^m(h) \subseteq |A|_m^M \\
& \quad (2) \text{ Negative condition: for some } h' \in H_m, \ M, m/h' \quad \neg A.
\end{align*}
\]

\( \text{dstit} = \text{cstit} + \text{Negative condition: } \text{A is capable of being false.} \)

\[
\begin{align*}
[\alpha \text{ dstit: } A] & \equiv [\alpha \text{ cstit: } A] \wedge \neg A \\
[\alpha \text{ cstit: } A] & \equiv [\alpha \text{ dstit: } A] \vee A
\end{align*}
\]

No difference between \( \text{dstit} \) and \( \text{cstit} \) in the picture. But they differ in two important cases:

i) If A is a logical or mathematical truth, \( [\alpha \text{ cstit: } A] \) but not \( [\alpha \text{ dstit: } A] \).

ii) More generally, if A happens to be settled true at \( m \), \( [\alpha \text{ cstit: } A] \) but not \( [\alpha \text{ dstit: } A] \).
e) Properties of the stit operators

i) cstit

\[
\begin{align*}
\text{RE. } & A \equiv B \rightarrow [\alpha \ cstit: A] \equiv [\alpha \ cstit: B] \\
\text{N } & [\alpha \ cstit: ] \\
\text{M } & [\alpha \ cstit: A \land B] \supset [\alpha \ cstit: A] \land [\alpha \ cstit: B] \\
\text{C } & [\alpha \ cstit: A] \land [\alpha \ cstit: B] \supset [\alpha \ cstit: A \land B]
\end{align*}
\]

In fact: if we just look at the histories through a moment as W, we have a normal modal system.
Actually, cstit is an S5 operator because in addition it satisfies T, 4 and B:

\[
\begin{align*}
\text{T } & [\alpha \ cstit: A] \supset A \\
\text{Lp } & \supset p \\
\text{4 } & [\alpha \ cstit: A] \supset [\alpha \ cstit: [\alpha \ cstit: A]] \\
\text{LLp } & \supset Lp \\
\text{B } & A \supset [\alpha \ cstit: \neg[\alpha \ cstit: \neg A]] \\
\text{LMp } & \supset Lp
\end{align*}
\]

RE, C, T are intuitively correct.
N is quite counter-intuitive.
M, 4 and B are problematic as well.

ii) dstit

\[
\begin{align*}
\text{RE, C, T and 4 hold.} \\
\text{N, M, and B fail: in fact, } \neg[\alpha \ dstit: ] \text{ is valid for any model.}
\end{align*}
\]

[\alpha \ dstit: A] respects the intuition that bringing about A involves the possibility that A could have been false.

Horty uses cstit because it is simpler.
3. Philosophical applications

a) Individual ability

How do we analyze the concept, “α is able to see to it that A”? 
Using cstit:
\[ \Diamond [\alpha \ cstit: \ A] \]

Picture: “α is able to see to it that A” holds at \( m/h \) if \([\alpha \ cstit: \ A]\) holds throughout some choice set at moment \( m \). It is not settled false that α does not see to it that A.

Kenny’s objection:

1. In most modal systems, possibility satisfies T and C:
   \[ T\Diamond: \ A \supset \Diamond A \]
   \[ C\Diamond: \ \Diamond (A \vee B) \supset \Diamond A \vee \Diamond B \]

2. Ability cannot be identified with any kind of possibility operator in such a system, because ability fails to satisfy both principles.
   Counterexample for T\( \Diamond \): \( A \equiv \alpha \) hits the bull’s eye
   (That A is true does not imply that α has the ability to see to it that A.)
   Counterexample for C\( \Diamond \): \( A \equiv \alpha \) hits the top half of the dartboard
   \( B \equiv \alpha \) hits the bottom half of the dartboard
   (α has the ability to hit the board, \( A \vee B \), but not to hit either half.)

Solution:
When ability is \( \Diamond [\alpha \ cstit: \ A] \), both principles are not valid.
b) Refraining

Concept: \( \alpha \) refrains from an action.

Examples: \( \alpha \) refrains from smoking

i) “Could have done otherwise”

With \( cstit \), no connection between doing and refraining.
With \( dstit \), the counter guarantees that, at the very least, \( \alpha \) could have made a choice that did not guarantee the truth of \( A \).

So \( dstit \) is a more natural concept to use for discussion of refraining.

Analysis of “doing implies could have done otherwise”:

- \([\alpha \ dstit: A] \supset \diamond[\alpha \ dstit: \neg A] \). FAILS.
- \([\alpha \ dstit: A] \supset \diamond[\alpha \ dstit: \neg[\alpha \ dstit: A]]\). VALID. (ACR)

ii) Refraining

*Def.* \( \alpha \) refrains from seeing to it that \( A \equiv [\alpha \ dstit: \neg[\alpha \ dstit: A]] \)

*Proposition:* Equivalent to \( \neg[\alpha \ dstit: A] \land \diamond[\alpha \ dstit: A] \)

The above becomes: “doing implies could have refrained”. We also get:

\([\alpha \ dstit: \neg[\alpha \ dstit: A]] \supset \diamond[\alpha \ dstit: A] \) (RCA)
\( \diamond[\alpha \ dstit: \neg[\alpha \ dstit: A]] \equiv \diamond[\alpha \ dstit: A] \) (CACR)

*Objection:* The Deer Hunter

Crossbow: 90\% chance of killing (K)
Rifle: 100\% chance of killing (K)

Shooting the crossbow counts as *refraining* from killing.
(Weakness due to decision to ignore probabilities.)

iii) Refref

\([\alpha \ dstit: \neg[\alpha \ dstit: [\alpha \ dstit: \neg[\alpha \ dstit: A]]]] \equiv [\alpha \ dstit: A] \) is VALID in all stit models.

“Refraining from refraining is equivalent to doing”.  

11
4. Group agency and ability

a) Two agents

- $\alpha$: left or right; $B$: up or down
- Neither $\alpha$ nor $\beta$ can see to it that $A$
- But $\{\alpha, \beta\}$ can see to it that $A$

Informal Idea:

- $h$ and $h'$ are choice-equivalent for $\{\alpha, \beta\}$ at $m$ if they belong to the same intersection of choice sets for $\alpha$ and $\beta$
- $M, m/h \quad \{\{\alpha, \beta\} \text{ cstit: } A\} \iff$ for all choice-equivalent $h'$, $M, m/h'$ $A$

b) Multiple agents

Independence of agents: each agent has no effect on simultaneous choices by other agents.

More formally: for each moment $m$ and each function $s$ that assigns a choice set $s(\alpha)$ to each agent $\alpha$, the intersection of the sets $s(\alpha), s(\beta), \ldots$ always contains at least one history. ("Something happens")

Still more formally:

$Select_m$ is the set of functions $s$ from Agent into subsets of $H_m$ that satisfies $s(\alpha) \in \text{Choice}_{\alpha}^m$.

Independence of agents: For any $m$ and any $s$ in $Select_m$, we have

$$\bigcap_{\alpha \in \text{Agent}} s(\alpha) \neq \Phi$$
If $\Gamma$ is a set of agents, then

$$Choice_{\Gamma}^m = \{ \bigcap_{s \in Select_m} s(\alpha) / s \in Select_m \}$$

is the set of joint choices available to $\Gamma$.

c) Group agency:

- $M, m/h [\Gamma_{cstit} : A]$ iff for all $h'$ in $Choice_{\Gamma}(h)$, $M, m/h'$ $A$
- $M, m/h [\Gamma_{dstit} : A]$ iff for all $h'$ in $Choice_{\Gamma}(h)$, $M, m/h'$ $A$, and in addition, there is some $h''$ with $M, m/h'' \neg A$.

Both cstit and dstit have the same properties for groups as for individual agents.

We can formalize group ability as either $\diamond [\Gamma_{cstit} : A]$ or $\diamond [\Gamma_{dstit} : A]$.

d) Frankfurt-style counterexamples

Story: Adam either pulls the trigger freely, or evil Dr. Blair pushes a button that causes him to pull the trigger involuntarily. In cases where he pulls the trigger freely, he acts even though there is no “ability to do otherwise”.

A problem for dstit?

Solution using joint agency:

By independence of agents, there is still a counter.