Introduction

Topic of the lecture

Reasoning about norm violation

- Contrary-to-duties, CTDs

- Why do we need CTD for NMAS?
  - Norm enforcement
- Main issue in deontic logic
  - fish-hook
- Thorny issue too
- Deontic logic-based agent: a holy grail?

Layout

- The CTD problem
  - how to detach an obligation in a violation context
- Ways to handle it
  - Temporal DDL
  - 2-dimensional SDL
  - SDL with sub-ideality
  - Dynamic DDL
- Example from ACL
  - Dialogue rules as soft constraints

A bit of terminology

Primary obligation

- Unconditional or conditional

Contrary-to-duty (or secondary, CTD) obligation

- Condition: primary norm violated

According-to-duty (ATD) obligation

- Condition: primary norm fulfilled

Multiple CTD or ATD **levels**

A bit of terminology

Primary obligation

- Unconditional or conditional

Contrary-to-duty (or secondary, CTD) obligation

- Condition: primary norm violated

According-to-duty (ATD) obligation

- Condition: primary norm fulfilled

Multiple CTD or ATD **levels**

$$\bigcirc A$$

$$\bigcirc(B/\neg A)$$

$$\bigcirc(?/?)$$

A bit of terminology

Primary obligation

- Unconditional or conditional

Contrary-to-duty (or secondary, CTD) obligation

- Condition: primary norm violated

According-to-duty (ATD) obligation

- Condition: primary norm fulfilled

Multiple CTD or ATD **levels**

Hansen's infinitly bad set:

$$\{p_i : i \in N\} \cup \{\bigcirc\neg p_1\} \cup \{\bigcirc(\neg p_{i+1}/p_1 \wedge \ldots \wedge p_i) : i \in N\}$$

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

" How can we take seriously a conditional obligation if it cannot, by way of detachment, lead to an unconditional obligation?" (van Eck, 1982, p. 263).

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\,(FD)$$

“ How can we take seriously a conditional obligation if it cannot, by way of detachment, lead to an unconditional obligation?" (van Eck, 1982, p. 263).

Forrester

Weakening the Consequent

$$\frac{\bigcirc(B/A) \qquad \Box(B \to C)}{\bigcirc(C/A)}\ (WC)$$

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

$$\frac{\bigcirc\neg A \qquad \dfrac{\Box(B \to A)}{\Box(\neg A \to \neg B)}}{\bigcirc\neg B}\ (WC)$$

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

'Gentle murderer' scenario

$$\{\bigcirc\neg k, \Box(k \wedge g \to k), \bigcirc(k \wedge g/k), k\} \vdash \bot$$

Forrester

Weakening the Consequent

$$\frac{\bigcirc(B/A) \qquad \Box(B \rightarrow C)}{\bigcirc(C/A)} \; (WC)$$

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B} \; (FD)$$

$$\frac{\bigcirc\neg A \qquad \dfrac{\Box(B \rightarrow A)}{\Box(\neg A \rightarrow \neg B)}}{\bigcirc\neg B} \; (WC)$$

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B} \; (FD)$$

'Gentle murderer' scenario

$$\{\bigcirc\neg k, \Box(k \wedge g \rightarrow k), \bigcirc(k \wedge g/k), k\} \vdash \perp$$

Forrester

Weakening the Consequent

$$\frac{\bigcirc(B/A) \qquad \Box(B \rightarrow C)}{\bigcirc(C/A)}\ (WC)$$

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

$$\frac{\bigcirc\neg A \qquad \dfrac{\Box(B \rightarrow A)}{\Box(\neg A \rightarrow \neg B)}}{\bigcirc\neg B}\ (WC)$$

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

'Gentle murderer' scenario

$$\{\bigcirc\neg k, \Box(k \wedge g \rightarrow k), \bigcirc(k \wedge g/k), k\} \vdash \bot$$

Chisholm

Deontic detachment

$$\frac{\bigcirc A \qquad \bigcirc(B/A)}{\bigcirc B}\ (DD)$$

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

$$\frac{\bigcirc(\neg B/\neg A) \qquad \neg A}{\bigcirc \neg B}\ (FD)$$

Example

$$\{\bigcirc a, \bigcirc(t/a), \bigcirc(\neg t/\neg a), \neg a\} \vdash \bot$$

Chisholm

Deontic detachment

$$\frac{\bigcirc A \qquad \bigcirc(B/A)}{\bigcirc B}\ (DD)$$

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

$$\frac{\bigcirc(\neg B/\neg A) \qquad \neg A}{\bigcirc \neg B}\ (FD)$$

Example

$$\{\bigcirc a, \bigcirc(t/a), \bigcirc(\neg t/\neg a), \neg a\} \vdash \bot$$

Chisholm

Deontic detachment

$$\frac{\bigcirc A \qquad \bigcirc(B/A)}{\bigcirc B}\ (DD)$$

Factual Detachment

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD)$$

$$\frac{\bigcirc(\neg B/\neg A) \qquad \neg A}{\bigcirc \neg B}\ (FD)$$

Example

$$\{\bigcirc a, \bigcirc(t/a), \bigcirc(\neg t/\neg a), \neg a\} \vdash \bot$$

An inside: conditional obligation within SDL (1)

Chisholm set: $\{\bigcirc a, \bigcirc(t/a), \bigcirc(\neg t/\neg a), \neg a\}$

Below $\rightarrow$ is material implication, and O is like $\Box$

Two candidate definitions

Option 1	$\bigcirc(B/A) =_{def} O(A \rightarrow B)$
Option 2	$\bigcirc(B/A) =_{def} A \rightarrow OB$

Usual representation of Chisholm set:

$$\{\bigcirc a, O(a \rightarrow t), \neg a \rightarrow O\neg t, \neg a\} \vdash \bot$$

Why option 1 + option 2?

An inside: conditional obligation within SDL (2)

$$\{\bigcirc a, \bigcirc(t/a), \bigcirc(\neg t/\neg a), \neg a\}$$

option 1 + option 1:

$$\{\bigcirc a, O(a \rightarrow t), O(\neg a \rightarrow \neg t), \neg a\}$$

option 2 + option 2:

$$\{\bigcirc a, a \rightarrow Ot, \neg a \rightarrow O\neg t, \neg a\}$$

option 2 + option 1:

$$\{\bigcirc a, a \rightarrow Ot, O(\neg a \rightarrow \neg t), \neg a\}$$

Ways out

$$\frac{\bigcirc(B/A) \qquad A}{\bigcirc B}\ (FD) \qquad\qquad \frac{\bigcirc A \qquad \bigcirc(B/A)}{\bigcirc B}\ (DD)$$

$$\frac{\bigcirc(B/A) \qquad \Box(B \to C)}{\bigcirc(C/A)}\ (WC)$$

Keep (*FD*) - perhaps in a qualified form - but

Easy way: give up (*DD*)/(*WC*)

- So-called non-normal modal logic (Forrester)
- $\to$: relevant implication (Goble)

Hard way: make (*FD*) coexist with (*DD*)/(*WC*)

- This lecture: different ways to do it

Adding time

Basic idea

before the violation $\neq$ after the violation

Åqvist's system DARB: DDL supplemented with $\oplus$ ("next") and $\boxed{s}$ ("settledness")

Semantics of DARB (1)

Frame

- $(Tree, <)$: a tree-like structure, where
  - *Tree*: a set of moments, m_1, m_2, ...
  - $<$: temporal relation (says which moment is next to which)
    - h : a history (a maximal chain of moments)
    - H_m : the set of histories passing through moment m.
- $\geq$: a ranking of histories in terms of betterness

NB: Truth-value of formulae made relative to pairs m/h

$\mathfrak{M}, m/h \models A$: in model $\mathfrak{M}$, formula A is true at moment m in history h

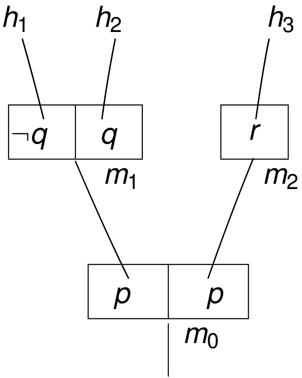
$[A]_m$: the set of histories making A true at m

Semantics of DARB (2)

Evaluation rules

$$\mathfrak{M}, m/h \models \oplus A \Leftrightarrow \mathfrak{M}, m+1/h \models A$$
$$\mathfrak{M}, m/h \models \boxed{\mathrm{s}} A \Leftrightarrow (\forall h' \in H_m)(\mathfrak{M}, m/h' \models A)$$
$$\mathfrak{M}, m/h \models \bigcirc(B/A) \Leftrightarrow \mathrm{best}_{\geq}([A]_m) \subseteq [B]_m.$$

Example



$$h_1 = \{m_0, m_1\}$$
$$h_2 = \{m_0, m_1\}$$
$$h_3 = \{m_0, m_2\}$$

$$m_0/h_0 \models ?$$

Chisholm set

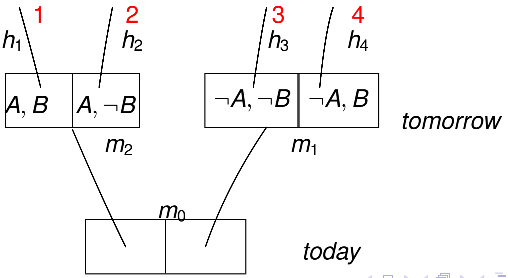
Case where A and B (will) occur simultaneously

Old representation	**New representation**
OA	$\bigcirc \oplus A$
$O(A \rightarrow B)$	$\bigcirc(\oplus B / \oplus A)$
$\neg A \rightarrow O\neg B$	$\oplus \bigcirc (\neg B / \neg A)$
$\neg A$	$\oplus \neg A$

Chisholm set

Case where A and B (will) occur simultaneously

Old representation	**New representation**
OA	$\bigcirc \oplus A$
$O(A \rightarrow B)$	$\bigcirc(\oplus B / \oplus A)$
$\neg A \rightarrow O\neg B$	$\oplus \bigcirc (\neg B / \neg A)$
$\neg A$	$\oplus \neg A$



Chisholm set

Case where A and B (will) occur simultaneously

Old representation	**New representation**
OA	$\bigcirc \oplus A$
$O(A \rightarrow B)$	$\bigcirc(\oplus B / \oplus A)$
$\neg A \rightarrow O\neg B$	$\oplus \bigcirc (\neg B / \neg A)$
$\neg A$	$\oplus \neg A$

Exercise

- Explain why the previous picture shows the new representation is consistent. (Use the evaluation rules for the connectives, and assume h_3 is the ‘actual’ history)

Chisholm set

Case where A and B (will) occur simultaneously

Old representation	**New representation**
OA	$\bigcirc \oplus A$
$O(A \rightarrow B)$	$\bigcirc(\oplus B / \oplus A)$
$\neg A \rightarrow O\neg B$	$\oplus \bigcirc (\neg B / \neg A)$
$\neg A$	$\oplus \neg A$

$$\frac{\bigcirc \oplus A \qquad \bigcirc(\oplus B / \oplus A)}{\bigcirc \oplus B}\,(DD)$$

$$\frac{\oplus \bigcirc (\neg B / \neg A) \qquad \oplus \neg A}{\oplus \bigcirc \neg B}\,(FD)$$

Chisholm set

Case where A and B (will) occur simultaneously

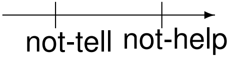
Old representation	**New representation**
OA	$\bigcirc \oplus A$
$O(A \rightarrow B)$	$\bigcirc(\oplus B / \oplus A)$
$\neg A \rightarrow O\neg B$	$\oplus \bigcirc (\neg B / \neg A)$
$\neg A$	$\oplus \neg A$

Remarks

- Forrester example: same idea
- But troubles with the case where B occurs before A

Consequent-before-the-antecedent

Chisholm words example



In DARB: qualified version of the factual detachment rule

$$\frac{\bigcirc_t(B/A) \qquad A_{t'}}{\bigcirc_t B}\ (t' \leq t)$$

time of B (within the scope of $\bigcirc$) (strictly) before t'.

Can the past be undone?

Exercise in class

Give a consistent representation of the Chisholm set for the case where B follows A.

Old representation	New representation
OA	$\bigcirc \oplus A$
$O(A \rightarrow B)$	$\bigcirc(\oplus \oplus B / \oplus A)$
$\neg A \rightarrow O\neg B$	$\oplus \bigcirc (\oplus\neg B/\neg A)$
$\neg A$	$\oplus\neg A$

Exercise in class

Give a consistent representation of the Chisholm set for the case where B follows A.

Old representation	**New representation**
OA	$\bigcirc \oplus A$
$O(A \rightarrow B)$	$\bigcirc(\oplus \oplus B / \oplus A)$
$\neg A \rightarrow O\neg B$	$\oplus \bigcirc (\oplus\neg B/\neg A)$
$\neg A$	$\oplus\neg A$

Basics

evaluation

reference

$$\langle t, s \rangle \Vdash A$$

Two senses of “ought"

- actual
- ideal or prima facie

Y modality: goes backwards in time

Material reading: *Synthese* paper on Moodle.

Exercise in class

Read sec. 2.2 of the paper. And

- Identify a condition on R to validate ax 3.5
- Use a direct argument to show the validity of ax 3.2
- Use a reductio ad absurdum to show the validity of ax 3.1, 3.3 and 3.4.

Chisholm example

$ObA =_{df} YO(r \rightarrow A) \qquad Ob(B/A) =_{df} YO(A \rightarrow ObB)$

Premisses are evaluated at $< t, t >$
t: violation context / ‘ideal’ u immediately before t

$$\frac{\begin{array}{c} < t, t > \models Ob(B/A) \\ < t, t > \models ObA \end{array}}{< u, t > \models r \rightarrow ObB}$$

$\Downarrow$

ideal (prima facie) “ought"

$< \boxed{u}, t >$

$$\frac{\begin{array}{c} < t, t > \models Ob(\neg B/\neg A) \\ < t, t > \models \neg A \end{array}}{< t, t > \models Ob\neg B}$$

$\Downarrow$

actual “ought"

$< \boxed{t}, t >$

Chisholm example

$$ObA =_{df} YO(r \rightarrow A) \qquad Ob(B/A) =_{df} YO(A \rightarrow ObB)$$

Premisses are evaluated at $< t, t >$
t: violation context / 'ideal' u immediately before t

$$\frac{\begin{array}{c} < t, t > \models Ob(B/A) \\ < t, t > \models ObA \end{array}}{< u, t > \models r \rightarrow ObB} \qquad \frac{\begin{array}{c} < t, t > \models Ob(\neg B/\neg A) \\ < t, t > \models \neg A \end{array}}{< t, t > \models Ob\neg B}$$

$$\Downarrow \qquad\qquad \Downarrow$$

ideal (prima facie) "ought" — actual "ought"

$< \boxed{u}, t >$ — $< \boxed{t}, t >$

Chisholm example

$$ObA =_{df} YO(r \to A) \qquad Ob(B/A) =_{df} YO(A \to ObB)$$

Premisses are evaluated at $< t, t >$
t: violation context / 'ideal' u immediately before t

$$\frac{\begin{array}{c}< t, t > \models Ob(B/A) \\ < t, t > \models ObA\end{array}}{< u, t > \models r \to ObB} \qquad \frac{\begin{array}{c}< t, t > \models Ob(\neg B/\neg A) \\ < t, t > \models \neg A\end{array}}{< t, t > \models Ob\neg B}$$

$\Downarrow$ ideal (prima facie) "ought"

$< u, t >$

$\Downarrow$ actual "ought"

$< t, t >$

Chisholm example

$ObA =_{df} YO(r \rightarrow A)$ $\qquad Ob(B/A) =_{df} YO(A \rightarrow ObB)$

Premisses are evaluated at $< t, t >$

t: violation context / 'ideal' u immediately before t

$$\frac{\begin{array}{c}< t, t > \models Ob(B/A) \\ < t, t > \models ObA\end{array}}{< u, t > \models r \rightarrow ObB} \qquad \frac{\begin{array}{c}< t, t > \models Ob(\neg B/\neg A) \\ < t, t > \models \neg A\end{array}}{< t, t > \models Ob\neg B}$$

$\Downarrow$ ideal (prima facie) "ought" $< \boxed{u}, t >$

$\Downarrow$ actual "ought" $< \boxed{t}, t >$

Exercises in class

- Show that the two inferences bottom of the previous slide are sound
- Show that $< t, t > \models (ObA) \wedge \neg A$ is inconsistent, if Ob is defined by $ObA =_{df} YOA$

Exercises in class

- Show that the two inferences bottom of the previous slide are sound
- Show that $< t, t > \models (ObA) \wedge \neg A$ is inconsistent, if Ob is defined by $ObA =_{df} YOA$

SDL with sub-ideality

Jones and Porn

Two relations

- R (ideality) and R' (sub-ideality)
- for all x, xRx or $xR'x$

Two pairs of modalities

- (O, P) and (O', P')

SDL with sub-ideality (2)

Jones and Porn

- **Ought** $A = OA \wedge P' \neg A$
- $\mathrm{Nec}\ A = OA \wedge O' A$

New Chisholm set
Ought A
$\mathrm{Nec}(A \rightarrow \textbf{Ought}\ B)$
$\mathrm{Nec}(\neg A \rightarrow \textbf{Ought}\ \neg B)$
$\neg A$

Exercise:

- Show the new Chisholm set is consitent