Aggregative Deontic Detachment for Normative Reasoning

Xavier Parent and Leendert van der Torre
x.parent.xavier@gmail.com, leon.vandertorre@uni.lu

UNIVERSITÉ DU LUXEMBOURG
CSC COMPUTER SCIENCE AND COMMUNICATIONS RESEARCH UNIT

The problem

Deontic Detachment

$$\text{DD}\ \frac{\bigcirc(y \mid x) \qquad \bigcirc x}{\bigcirc y}$$

Counterexample

a. You ought to exercise hard
b. If you exercise hard, you ought to eat heartily
c. ?*You ought to eat heartily

Broome: "What, if you do not take exercise?" [1]

This counterexample (and others, cf. [3, 4]) suggests an alternative (call it **aggregative**) form of detachment:

$$\text{ADD}\ \frac{\bigcirc(y \mid x) \qquad \bigcirc x}{\bigcirc(x \wedge y)}$$

This form of detachment has been overlooked in the literature.

Question

- Is there any interesting system supporting ADD, but not DD?

Accepting ADD, but not DD, implies rejecting W (*Weakening*)

$$\text{W}\ \frac{\bigcirc(x \mid a) \qquad x \vdash y}{\bigcirc(y \mid a)} \qquad \text{ADD} + \text{W} \to \text{DD}$$

Tasks

- 2-step semantics
  - remove W from standard systems
  - add ADD
- Sound and complete axiomatization

Our approach

In our work, we use so-called **input/output (I/O) logic** [5, 6]. The meaning of deontic concepts in given in terms of a set of procedures yielding outputs for inputs.

In I/O logic, a conditional obligation is represented as a pair (a, x) of boolean formulae, where a and x are the body (antecedent) and the head (consequent), respectively.

A normative system N is a set of such pairs.

Below: our main construct

$$x \in O(N, a)$$

Intuitively: given input a (state of affairs), x (obligation) is in the output under norms N.

Equivalent notation: $(a, x) \in O(N)$.

References

[1] J. Broome, *Rationality Through Reasoning* (2013)

[2] H. Prakken and M. Sergot, Contrary-to-duty obligations, *Studia Logica* (1996)

[3] S. O. Hansson, Situationist deontic logic, *Journal of Philosophical Logic* (1997)

[4] D. Makinson, On a fundamental problem in deontic logic, P. McNamara & H. Prakken (eds), *Logic, Norms and Information Systems* (1999)

[5] D. Makinson and L. van der Torre, Input/output logics, *Journal of Philosophical Logic* (2000)

[6] L. van der Torre and X. Parent, Input/output logics, D. Gabbay & al. (eds), *Handbook of Deontic Logic and Normative Systems* (2013)

[7] X. Parent and L. van der Torre, 'Sing and dance!': Input/output logics with weakening, F. Cariani & al. (eds), *Deontic Logic and Normative Systems*, *ΔEON 2014 Proceedings.*

Standard I/O system

Definition 1 (Simple-minded, [5]). *$x \in out(N, a)$ iff $x \in Cn(N(Cn(a)))$, where $Cn(X) = \{y : X \vdash y\}$, and $N(X) = \{y : (b, y) \in N, b \in X\}$.*

Cf. Boghossian: modus-ponens is constitutive of the possession of the notion of conditional.

Removing W

$N[X] = \{x : x \dashv\vdash \wedge_{i=1}^{n} x_i\}$, where $N(X) = \{x_1, ..., x_n\}$. N is required to be finite.

Definition 2 (Semantics). *$x \in \mathcal{O}^\star(N, a)$ iff $\exists M \subseteq N$ s.t. $M(Cn(a)) \neq \emptyset$ and $x \in M[Cn(a)]$*

Define $\mathcal{O}^\star(N) = \{(a, x) : x \in \mathcal{O}^\star(N, a)\}$.

Definition 3 (Proof system). *$(a, x) \in \mathcal{D}^\star(N)$ iff there is a derivation of (a, x) from N using the rules {SI, EQ, AND}.*

$$\text{SI}\ \frac{(a, x) \qquad b \vdash a}{(b, x)} \qquad\Big|\qquad \text{EQ}\ \frac{(a, x) \qquad x \dashv\vdash y}{(a, y)} \qquad\Big|\qquad \text{AND}\ \frac{(a, x) \qquad (a, y)}{(a, x \wedge y)}$$

$\mathcal{D}^\star(N, a) = \{x : (a, x) \in \mathcal{D}^\star(N)\}$.

Theorem 1 (Soundness and completeness). $\mathcal{O}^\star(N, a) = \mathcal{D}^\star(N, a)$

Proof. See [7]. □

Adding ADD

Definition 4 (Semantics). *$x \in \mathcal{O}(N, a)$ iff $\exists M \subseteq N$ s.t. $M(Cn(a)) \neq \emptyset$ and $x \in M[B]$ for all B with $a \in B = Cn(B) \supseteq M[B]$. Such a M is called an a-witness for x.*

Definition 5 (Proof system). *$(a, x) \in \mathcal{D}(N)$ iff there is a derivation of (a, x) from N using the rules {SI, EQ, ACT}.*

$$\text{ACT}\ \frac{(a, x) \qquad (a \wedge x, y)}{(a, x \wedge y)}$$

ACT yields ADD as a special case (a is $\top$).

Theorem 2 (Soundness and completeness). $\mathcal{O}(N, a) = \mathcal{D}(N, a)$

Proof. See [7]. □

Properties

Property 1 (Bridge law). *$out'(N) = Cn(\mathcal{O}^\star(N))$, where out' is the standard "reusable" I/O operation [5]. (out' extends out to iterations of successive detachments.)*

Property 2 (Closure). *$\mathcal{O}^\star$ is a closure operator, viz*

$$(x, y) \in N \Rightarrow y \in \mathcal{O}^\star(N, x) \tag{1}$$

$$\mathcal{O}^\star(N) \subseteq \mathcal{O}^\star(N \cup M) \tag{2}$$

$$M \subseteq \mathcal{O}^\star(N) \Rightarrow \mathcal{O}^\star(N) = \mathcal{O}^\star(N \cup M) \tag{3}$$

(1), (2) and (3) express a principle of factual detachment, norm monotony, and norm induction, respectively.

Property 3 (Violation detection). $x \in \mathcal{O}^\star(N, a) \Rightarrow x \in \mathcal{O}^\star(N, a \wedge \neg x)$.

Intuitively: in a violation context, obligations do not 'drown'. (This is a departure from non-monotonic logics, which reject SI. Exceptions and violations should not be conflated.)

The way forward

Pragmatic oddity [6]

$$\text{ACT}\ \frac{\text{SI}\ \dfrac{(\top, \neg d)}{(d, \neg d)} \qquad \text{SI}\ \dfrac{(d, s)}{(d \wedge \neg d, s)}}{(d, \neg d \wedge s)}$$

in a cottage
d : there is a dog
s : there is a warning sign

Definition 6 (Backtesting). *$x \in \mathcal{O}'(N, a)$ iff: $a \vdash \wedge b$ with $x \in \mathcal{O}(N, \wedge b)$ and $\wedge b \cup \{x\} \nvdash \bot$.*

Intuitively: go back in time, and check if x was obligatory before the violation occurred.