Contrary-to-duty reasoning, preference and violation

About an obligation

Yingqin Fu[1] Markus Davydov[2]

11.03.2021

Table of Contents

Table of Contents

Introduction

- What is contrary-to-duty obligation?

Introduction

What is contrary-to-duty obligation?

- A contrary-to-duty obligation is an obligation telling us what ought to be the case if something forbidden is true.

Examples

- If she is guilty, she should confess
- If he has hurt his friend, heshould apologise to her
- If you are not going to keep your promise to him, you ought to call him
- If the books are not returned by the due date, you must pay a fine

Table of Contents

Chisholm's paradox

... known as one of the earliest CTD paradox. Chisholm's paradox consists of the following four sentences:

Chisholm's paradox

1. It ought to be that a certain man go to the assistance of his neighbours.
2. It ought to be that if he does go, he tell them he is coming.
3. If he does not go then he ought not to tell them he is coming.
4. He does not go.

Chisholm's paradox

Chisholm's paradox is a contrary-to-duty paradox, since it contains both a primary obligation to go, and a secondary obligation not to tell if the agent does not go. Yet intuitively the natural-language expressions that make up the paradox are consistent and independent from each other: this is why it is called a paradox.

Table of Contents

Monadic deontic logic

- Traditional or ‘standard’ deontic logic, often referred to as SDL
- SDL regards as the most basic system of deontic logic
- But also have the most drawbacks

Monadic deontic logic

Quick overview about the Language, Semantics and Axioms

Language

- $\varphi := \perp |p| \neg\varphi | (\varphi \wedge \varphi) | \bigcirc \varphi | \Box\varphi$
- $\perp$:the empty symbol
- $\neg\ \varphi$: classical negation
- $\varphi \vee \phi$: classical disjunction
- $\bigcirc\varphi$: φ is obligatory
- $\neg \bigcirc \neg\varphi$: φ is forbidden
- $\Box\varphi$: φ is necessary

Quick overview about the Language, Semantics and Axioms

Semantics

- M, s $\vDash$ p iff s $\in \nu$ (p)
- M, s $\vDash \neg\varphi$ iff not M, s $\vDash \varphi$
- M, s $\vDash (\varphi \wedge \phi)$ iff M,s $\vDash \varphi$ and M,s $\vDash \phi$
- M, s $\vDash \bigcirc\varphi$ iff for all t,if Rst then M, t$\vDash \varphi$
- M, s $\vDash \Box\varphi$ iff for all t $\in$ W, M, t $\vDash \varphi$

Monadic deontic logic

Quick overview about the Language, Semantics and Axioms

Axioms

1. Taut: All tautologies are Well formed formulas of the language.
2. $AxiomK : O(p \to q) \to (O(p) \to O(q))$
3. $AxiomD : O(p) \to \neg O(\neg p)$
4. $ModusPonens : ((p \to q) \wedge p) \to q$
5. $Necessity : p \to O(p)$

The Limitations of SDL

Now lets focus on the Chisholms paradox again.

Chisholm's paradox

1. It ought to be that a certain man go to the assistance of his neighbours.
2. It ought to be that if he does go, he tell them he is coming.
3. If he does not go then he ought not to tell them he is coming.
4. He does not go.

Chisholm's paradox in SDL

1. $(1)O(h); (2)O(h \rightarrow t); (3)O(\neg h \rightarrow \neg t); (4)\neg h$
2. $(1)O(h); (2)O(h \rightarrow t); (3)\neg h \rightarrow O(\neg t); (4)\neg h$
3. $(1)O(h); (2)h \rightarrow O(t); (3)\neg h \rightarrow O(\neg t); (4)\neg h$
4. $(1)O(h); (2)h \rightarrow O(t)(3)O(\neg h \rightarrow \neg t)(4)\neg h$

The Limitations of SDL

Chisholm's paradox in SDL

1. $(1)O(h);(2)O(h \to t);(3)O(\neg h \to \neg t);(4)\neg h$
2. $(1)O(h);(2)O(h \to t);(3)\neg h \to O(\neg t);(4)\neg h$
3. $(1)O(h);(2)h \to O(t);(3)\neg h \to O(\neg t);(4)\neg h$
4. $(1)O(h);(2)h \to O(t)(3)O(\neg h \to \neg t)(4)\neg h$

1. From (4) and Axiom 5 ,we have $O(\neg h)$;From (1) and (2),we have O(t);From $O(\neg h)$ and (3),we have $O(\neg t)$. Obviously,O(t) $\wedge$ $O(\neg t)$ is paradox.
2. From (3) and (4) with Axiom MP,we have $O(\neg t)$;From (2) and Axiom K,we have $O(h) \to O(t)$;which with (1) and Axiom MP,we have O(t)
3. We know that,$\neg h \to (h \to O(t))$ is Taut,which with (4) can we have $h \to O(t)$.it is identical to (2).Hence,it's redundant.
4. It's redundant with the same reason of 3.

The Limitations of SDL

We can conclude:

- Under the Monadic Deontic Logic(SDL) the CTD is inconsistent and redundant.
- Something must be wrong with our formalisation, with SDL or with our intuitions.

Basically this puzzle is the contrary-to-duty (obligation) paradox.

So how can we overcome the limitation of SDL?Can we extend the semantics of SDL?

The Limitations of SDL

Yes for example, one can add distinct modal operators for primary and secondary obligations, where a secondary obligation is a kind of reparational obligation.

So the sentences are now:

(1) $\bigcirc_1 h$

(2) $\bigcirc_1 (h \to t)$

(3) $\neg h \to \bigcirc_2(\neg t)$

(4) $\neg h$

From 1-4 we can derive only $\bigcirc_1 \neg t \wedge \bigcirc_2 t$,which is consistent.

Seems like we find a good way to solve the limitation of SDL.However,it may not always be easy to distinguish primary from secondary obligations, because it may depend on the context whether an obligation is primary or secondary.

Table of Contents

Dyadic deontic logic

- Also known as Dyadic Standart deontic logic
- Logic for reasoning with dyadic obligations ("it ought to be the case that ... if it is the case that ...").
- The language is extended with dyadic operators $\bigcirc(p|q)$, which is true iff the preferred q worlds satisfy p.

Dyadic deontic logic

Quick overview of the language and semantics

Language

- $\varphi := \perp |p| \neg\varphi | (\varphi \wedge \varphi) | \Box\varphi | \bigcirc (\varphi | \varphi)$
  Given two arbitrary formulas ϕ and φ
- ...
- $\bigcirc(\varphi \mid \phi)$: It ought to be φ,given ϕ
- $P(\varphi \mid \phi)$:φ is permitted,given ϕ,as an abbreviation of $\neg \bigcirc (\neg\varphi \mid \phi)$
- $\Diamond(\varphi)$:possibly φ,as an abbreviation of $\neg\Box\neg\varphi$

Semantic

- ...
- M, s
  $\vDash \bigcirc(\phi \mid \varphi) iff \forall t(((M, t \vDash \varphi)\forall u(M, u \vDash \varphi) \Rightarrow t \geq u) \Rightarrow M, t \vDash \phi)$
- M, s $\vDash \Box\varphi$ iff for all t $\in$ W, M, t $\vDash \varphi$

Limitations of DSDL

A variant of the Chilsom's paradox

- A It ought to be that Jones does not eat fast food for dinner.
- B It ought to be that if Jones does not eat fast food for dinner, then he does not go to McDonald's.
- C If Jones eats fast food for dinner, then he ought to go to McDonald's.
- D Jones eats fast food for dinner.

Limitations of DSDL

A variant of the Chilsom's paradox

- A It ought to be that Jones does not eat fast food for dinner.
- B It ought to be that if Jones does not eat fast food for dinner, then he does not go to McDonald's.
- C If Jones eats fast food for dinner, then he ought to go to McDonald's.
- D Jones eats fast food for dinner.

A variant of the Chilsom's paradox in DSDL

f:Jones eats fastfood for dinner; m:he goes to McDonald's

- A $\bigcirc\neg f$
- B $\bigcirc(\neg m \mid \neg f)$
- C $\bigcirc(m \mid f)$
- D f

Limitations of DSDL

A variant of the Chilsom's paradox in DSDL

f:Jones eats fastfood for dinner; m:he goes to McDonald's

- A $\bigcirc\neg f$
- B $\bigcirc(\neg m \mid \neg f)$
- C $\bigcirc(m \mid f)$
- D f

The dyadic representation A - D highlights the dilemma between factual detach-ment (FD) and deontic detachment (DD)

- Axiom FD $\bigcirc(m \mid f), f \Rightarrow \bigcirc m$
- Axiom DD $\bigcirc(\neg m \mid \neg f), \bigcirc\neg f \Rightarrow \bigcirc\neg m$

Limitations of DSDL

From A and B wit axiom DD ,we have $\bigcirc\neg m$;From C and D(axiom FD),we have $\bigcirc m$, as we simply derive a dilemma:$\bigcirc\neg m \wedge \bigcirc m$
Hence,there is also something wrong with the DSDL,when we try to reasoning CTD paradox.
The main drawback of DSDL is that in a monotonic setting, we cannot detach the obligation $\bigcirc m$ from the four sentences.

Table of Contents

Input/Output logic

- Input/Output logic is a non classical logic
- they may be expressed in terms like: In such-and-such a situation, so-and-so should be the case, or . . . should be brought about, or . . . should be worked towards, or . . . should be followed.
- Input/output logic may be seen as an attempt to extract the essential mathematical structure behind the reconstruction of deontic logic.

Input/Output logic

Definition

1. Cn(A) denotes the set of logical consequences of A in classical propositionallogic.It returns the set of all provable propositional formulae provable assumingthe fact in A(the set of answers of a set of inputs).
2. G(A) is the set of answers of a set of inputs.

Examples

A= $\{x, y\}$ then $Cn(A) = \{x, y, x \vee y, x \wedge y \vee \ldots\}$

Examples

$G1 = \{(a1, x1),(a2, x2)\}$ and $A1\{a1, z\}$ and $A2 = \{a1, a2, x2\}$
$G1(A1) = x1$
$G1(A2) = x1, x2$

Rules

What's more,there are 5 rules in I/O logic:

- Strengthening Input (SI): From (a,x) to (b,x) whenever a $\in$ Cn(b)
- Conjoining Output (AND): From (a,x), (a,y) to (a,x$\wedge$y)
- Weakening Output (WO): From (a,x) to (a,y) whenever y $\in$ Cn(x).
- Disjoining input (OR): From (a,x), (b,x) to (a$\vee$b,x)
- Cumulative transitivity (CT): From (a,x), (a$\wedge$x,y) to (a,y).

Systems of input/output

There are also four very natural systems of input/output, which are labelled as follows:

- simple-minded alias $out_1(G,A) = Cn(G(Cn(A)))$
- basic (simple-minded plus input disjunction: out_2)
- reusable (simple-minded plus reusability: out_3)
- reusable basic alias out_4

For example, out_4 can be given by Figure 2.

systems of input/output

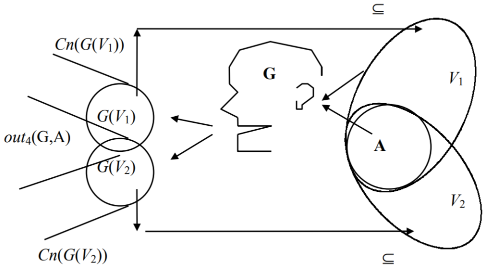

Figure 2: Reusable Basic Output:
$out_4(G,A) = \cap\{Cn(G(V)): A \subseteq V \supseteq G(V), V\ complete\}$

Contrary to duty reasoning

Definition(Maxfamilies)

Let G be a set of conditional norms and A and C two sets of propositional formulas.
Then maxfamily(G, A, C) is the set of maximal subsets H $\subseteq$ G such that out(H, A) $\cup$ C is consistent.(which means,out(H,A)) is consistent with C.)

For a possible solution to Chisholm's paradox, consider the following output operation $out^{\cap}$:
$out^{\cap}(G, A) = \cap\{out(H, A)|H \in maxfamily(G, A, A)\}$

Contrary to duty reasoning

Let G a "Chisholm norm set";
x means the norm that a man goes to the assistence of his neighbors;
z means the norm that he tells them he is coming

- G=$\{(\top, x), (x, z), (\neg x, \neg z)\}$, $(\top, x)$ means the norm that the man must go to the assistance of his neighbors;(x, z)means the norm that it ought to be that if he goes he ought to tell them he is coming;$(\neg x, \neg z)$ means that the norm that if he does not go he ought not to tell them he is coming.
- $A = \{\neg x\}$
- We have H $\in out_1(G, A) = \{x, \neg z\}$,which is inconsistent with A because of $\top$
- We have H $=$ maxfamily(G,A,A)=$\{G \setminus \{(\top, x)\}\} = \{\{(x, z), (\neg x, \neg z)\}\}$
- out(H,A)=$\neg z$,which is consistent with A
- $out^{\cap}(G, A) = Cn(\neg z)$,which means the form that the man must not tell his neighbors he is coming.

Table of Contents

Counterfactual Deontic Logic

Counterfactual Deontic Logic,which is introduced by Rönnedal in 2016.The Counterfactual Deontic Logic combine the Counterfactual Logic with Deontic Logic.It aims at the relationship between the obligation and time.

Definition

- $A \Box \rightarrow B$ is often read ‘If A were the case, then B would be the case’
- R is a temporal operator; ‘Rt_1A’ says that it is realised at time t1 (it is true at t1) that A

Contrary to duty reasoning

Consider the following sentences about Contrary to duty paradox

1. (On Monday it is true that) You ought to keep your promise (and see your friend on Friday).
2. (On Monday it is true that) It ought to be that if you keep your promise, you do not apologise (when you meet your friend on Saturday).
3. (On Monday it is true that) If you do not keep your promise (i.e. if you do not see your friend on Friday and help her out), you ought to apologise (when you meet her on Saturday).
4. (On Monday it is true that) You do not keep your promise (on Friday).

Contrary to duty reasoning

With the help of Counterfactual Deontic Logic,the four sentences can now be symbolised in the following way:(R is a temporal operator;t1,t2,t3 are the different time;k is keep promise;and a is apologise)

CF1 $Rt_1 ORt_2 k$

CF2 $Rt_1 O(Rt_2 k\Box \rightarrow Rt_3 \neg a)$

CF3 $Rt_1 (Rt_2 \neg k\Box \rightarrow Rt_2 ORt_3 a)$

CF4 $Rt_1 Rt_2 \neg k$

Contrary to duty reasoning

- CF-CTD is consistent,It does not seem to be possible to deduce any contradiction from the four sentences. Hence, we want our symbolisation of this set to be consistent. In this respect, CF-CTD is an intuitively plausible formalisation of the four sentences.
- CF-CTD is non-redundant.There ist no sentence in this set follows from the rest.

Limitations

The Counterfactual Deontic Logic seems to be a proposed solution of CTD-paradox(at least in our example).However,it also has limitations.

- the Counterfactual Solution cannot Handle Timeless (or Parallel) Contrary-to-Duty Paradoxes;
- the Counterfactual Solution cannot Solve Beforehand Contrary-to-Duty Paradoxes

Table of Contents

Alternative approaches

approach 1

Carmo and Jones suggest that the representation of the facts is challenging, instead of the representation of the norms. In their approach, depending on the formalisation of the facts various obligations can be detached.

approach 2

A recent representation of Chisholm'paradox is to replace deonitc detachment by so-called aggregative deontic detachment(ADD),and to derive from A-D the obligation $(\neg f \wedge \neg m)$and m, but not ¬m.
$\bigcirc(m|f), f \Rightarrow \bigcirc m$ FD
$\bigcirc(\neg m|\neg f), \bigcirc\neg f \Rightarrow \bigcirc(\neg m \wedge \neg f)$ ADD
The limitation of this approaches is that we can no longer accept the principle of weakening(also known as inheritance).
$\bigcirc(\neg m \wedge \neg f|\top) \Rightarrow \bigcirc(\neg m|\top)$

Table of Contents

Conclusion

Accordingly, CTDs are the source for many paradoxes and the driver for the development of many formalisms and deontic logics.In our opinion,the key point of CTD Paradoxs is that how can we find a adequate formel to deal with the relationship between **obligation** and **condition**.

Example

Consider the sentence from section 3:

B It ought to be that if Jones does not eat fast food for dinner, then he does not go to McDonald's.

which can be formed as:

- $\bigcirc(\neg f \to \neg m)$

However, in our opinion,"if Jones does not eat fast food for dinner "is a **condition** (instead of obligation) of "he ought to not go to McDonald's".We cannot lump these two concepts together. So the better form should be:

- $\neg f \to \bigcirc \neg m$

However,this is also redundant due to $f \vDash \neg f \to \bigcirc \neg m$

Conclusion

We always find a limitation,when we try to form CTD with deontic logic,especially SDL.More and more approaches are proposed.Many researchers also build the binary deontic logic to solve them......When a new deontic logic is proposed, the traditional contrary-to-duty examples are always the first benchmark examples to be checked,which is always the biggest challenge of DL System.What's more,it is also a challenge for our own moral life and law system.
Perhaps there is not an eternal solution to solve the paradoxes entirely.But with the research on all the challenges for multiagent deontic logic,the law system will be more and more perfect,our obligation and violation will be more and more clearly...