

# Contrary-to-Duty Paradoxes

BACHELOR-SEMINAR INFORMATIK (WS 20/21)

Yingqin Fu 702065 Markus Davydov 690672

December 2020

## Contents

<b>1</b>	<b>Chisholm's paradox</b>	<b>2</b>
<b>2</b>	<b>Monadic deontic logic</b>	<b>3</b>
2.1	language . . . . .	3
2.2	Semantics . . . . .	3
2.3	Limitations . . . . .	4
<b>3</b>	<b>Dyadic deontic logic</b>	<b>5</b>
3.1	Language . . . . .	5
3.2	Semantics . . . . .	5
3.3	Limitations . . . . .	6
<b>4</b>	<b>Input/Output logic</b>	<b>7</b>
4.1	Definition . . . . .	7
4.2	Contrary to duty reasoning . . . . .	9
4.2.1	Definition(Maxfamilies) . . . . .	9
4.2.2	Contrary to duty reasoning . . . . .	9
<b>5</b>	<b>Counterfactual Deontic Logic</b>	<b>10</b>
5.1	Definition . . . . .	10
5.2	Contrary to duty reasoning . . . . .	10
5.3	Limitations . . . . .	11
<b>6</b>	<b>Alternative approaches</b>	<b>11</b>
<b>7</b>	<b>Conclusion</b>	<b>12</b>
	<b>References</b>	<b>13</b>

## Abstract

We will summarize the first problem of Deontic Logic, the so-called Contrary-to-duty (CTD) in this paper. A contrary-to-duty obligation tells us, what we ought to do if something is forbidden. Which is very important in our moral and legal thinking. So we want to discuss the problem with different deontic logic systems, and try to find a way to explain it.

Keywords: contrary-to-duty Paradoxes (CTD), Chisholm's paradox, Standard deontic logic (SDL)

## Introduction

What is contrary-to-duty obligation? And why is it a Paradoxes? A contrary-to-duty obligation is an obligation telling us what ought to be the case if something forbidden is true. Here are some examples of sentences that often express such obligations: 'If she is guilty, she should confess', 'If he has hurt his friend, he should 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'. In this paper we will summarize the first part of the paper "Multiagent Deontic Logic and its Challenges from a Normative Systems Perspective" [1]

## 1 Chisholm's paradox

First of all, We are going to introduce first the so-called Chisholm's paradox [2], which is known as one of the earliest CTD paradox. Chisholm's paradox consists of the following four sentences:

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 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.

There are also various kinds of scenarios which are similar to Chisholm's scenario. For example, van Benthem, Grossi and Liu [3] give the following example, in the formulation proposed by Åqvist:

- (1) It ought to be that Smith refrains from robbing Jones.

- (2) Smith robs Jones.
- (3) If Smith robs Jones, he ought to be punished for robbery.
- (4) It ought to be that if Smith refrains from robbing Jones he is not punished for robbery.

Contrary-to-duty obligations turn up in discussions concerning guilt, blame, restoration, reparation, confession, punishment, repentance, retributive justice, compensation, apologies, damage control, etc. Therefore, they are important in our moral and legal thinking. The rationale of a contrary-to-duty obligation is the fact that most of us do neglect our primary duties from time to time and yet it is reasonable to believe that we should make the best of a bad situation, or at least that it matters what we do when this is the case. Everyone needs to find a way, that tells us what to do, when we don't finish the things we ought to do. In this paper, we will try to find an adequate formal representation of such obligations.

## 2 Monadic deontic logic

Traditional or 'standard' deontic logic, often referred to as SDL, was introduced by Von Wright[4]. SDL regards as the most basic system of deontic logic, but also maybe have the most drawback. In this section, we try to analyze the Chisholm's paradox with the basic SDL system.

### 2.1 language

$\varphi := \perp \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \bigcirc\varphi \mid \Box\varphi$

Given two arbitrary formulas  $\phi$  and  $\varphi$ , then

$\perp$ : the empty symbol

$\neg\varphi$ : classical negation

$\varphi \vee \phi$ : classical disjunction

$\bigcirc\varphi$ :  $\varphi$  is obligatory

$\neg\bigcirc\neg\varphi$ :  $\varphi$  is forbidden

$\Box\varphi$ :  $\varphi$  is necessary

### 2.2 Semantics

- let  $\nu$  be a valuation function on AP
- let  $M = (W, R)$  A kripke model with on worlds  $W$  related via  $R : W \times W$
- " $M, s \models \varphi$ " is read a state  $s$  satisfies  $\varphi$  in the Model  $M$ , we define the satisfaction relation  $M, s \models \varphi$  by induction on  $\varphi$  using the clauses:
  - $M, s \models p$  iff  $s \in \nu(p)$
  - $M, s \models \neg\varphi$  iff not  $M, s \models \varphi$

- $M, s \models (\varphi \wedge \phi)$  iff  $M, s \models \varphi$  and  $M, s \models \phi$
- $M, s \models \bigcirc\varphi$  iff for all  $t$ , if  $Rst$  then  $M, t \models \varphi$
- $M, s \models \Box\varphi$  iff for all  $t \in W$ ,  $M, t \models \varphi$

Axioms for Standard Deontic Logic:

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

### 2.3 Limitations

Now we focus on the Chisholm'paradox with the Monadic deontic logic. We can describe the four sentences of Chisholm'paradox with the Monadic Deontic logic ( $h$ := he goes to help his neighbours;  $t$ :=he tell them), we have four different ways to describe them:

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

Now we focus on the above-mentioned four SDL descriptions:

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

Hence, we can conclude, that 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. In a nutshell, 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? For example, one can add distinct modal operators for primary and

secondary obligations, where a secondary obligation is a kind of reparational obligation. So with the Chisholm's paradox, we regard 1 and 2 as primary obligation (he ought to go), 3 as secondary obligation (he ought to tell his neighbours). So the sentences are now:

- (1)  $\bigcirc_1 h$
- (2)  $\bigcirc_1 (h \rightarrow t)$
- (3)  $\neg h \rightarrow \bigcirc_2 (\neg t)$
- (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. It's not easy for us to describe the ideal world with this solution.

### 3 Dyadic deontic logic

Dyadic deontic logic (or Dyadic Standard Deontic Logic) is the 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  $p$ . The class of logics is called Dyadic 'Standard' Deontic Logic or DSDL. The notation is inspired by the representation of conditional probability. In this section, we will reasoning CTD with DDL.

#### 3.1 Language

$\varphi := \perp \mid p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid \bigcirc(\varphi|\varphi)$

Given two arbitrary formulas  $\phi$  and  $\varphi$ , then

$\perp$ : the empty symbol

$\neg\varphi$ : classical negation

$\varphi \vee \phi$ : classical disjunction

$\bigcirc\varphi$ :  $\varphi$  is obligatory

$\neg\bigcirc\neg\varphi$ :  $\varphi$  is forbidden

$\Box\varphi$ :  $\varphi$  is necessary

$\bigcirc(\varphi|\phi)$ : It ought to be  $\varphi$ , given  $\phi$

$P(\varphi|\phi)$ :  $\varphi$  is permitted, given  $\phi$ , as an abbreviation of  $\neg\bigcirc(\neg\varphi|\phi)$

$\Diamond(\varphi)$ : possibly  $\varphi$ , as an abbreviation of  $\neg\Box\neg\varphi$

#### 3.2 Semantics

A preference model  $M = (W, \geq, V)$  is a structure where:

- $W$  is a nonempty set of worlds.

- $\geq$  is a reflexive, transitive relation over  $W$  satisfying the following limitedness requirement: if  $\|\varphi\| \neq \emptyset$ , then  $x \in \|\varphi\| : (\forall y \in \|\varphi\|) x \geq y \neq \emptyset$ . Here  $\|\varphi\| = \{x \in W : M, x \models \varphi\}$ .
- $V$  is a standard propositional valuation such that for every propositional letter  $p$ ,  $V(p) \subseteq W$ .
  - $M, s \models p$  iff  $s \in V(p)$
  - $M, s \models \neg\varphi$  iff not  $M, s \models \varphi$
  - $M, s \models (\varphi \wedge \phi)$  iff  $M, s \models \varphi$  and  $M, s \models \phi$
  - $M, s \models \bigcirc(\phi \mid \varphi)$  iff  $\forall t ((M, t \models \varphi) \Rightarrow (M, t \models \phi))$
  - $M, s \models \Box\varphi$  iff for all  $t \in W$ ,  $M, t \models \varphi$

### 3.3 Limitations

The following example is a variant of the scenario originally phrased by Chisholm in 1963. There is widespread agreement in the literature that, from the intuitive point of view, this set of sentences is consistent, and its members are logically independent of each other.

- 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.

We try to reasoning the above-mentioned sentences with DSDL( $f$ :Jones eats fast food 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 detachment (FD) and deontic detachment (DD)

**Axiom FD**  $\bigcirc(m \mid f), f \Rightarrow \bigcirc m$

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

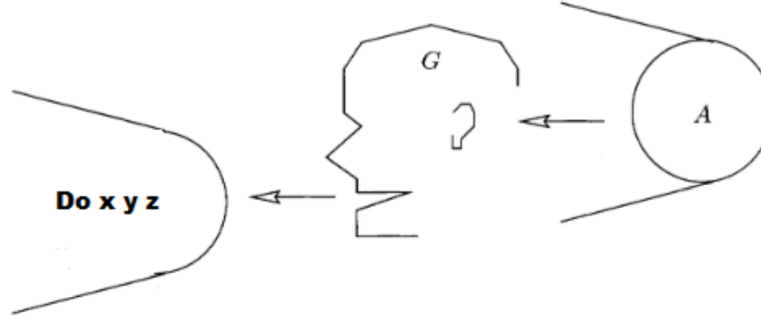
From A and B with (axiom DD), we have  $\bigcirc \neg m$ ; From C and D (axiom FD), we have  $\bigcirc m$ , as we simply derive a dilemma:  $\bigcirc \neg m \wedge \bigcirc m$ . Hence, there is also something wrong with the DSDL, when we try to reason the CTD paradox. The main drawback of DSDL is that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

## 4 Input/Output logic

Input/output logic [5], which is a non classical logic, takes its origin in the study of conditional norms. These may express desired features of a situation, obligations under some legal, moral or practical code, goals, contingency plans, advice, etc. Typically 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 – these locutions corresponding roughly to the kinds of norm mentioned. Input/output logic may be seen as an attempt to extract the essential mathematical structure behind the reconstruction of deontic logic.

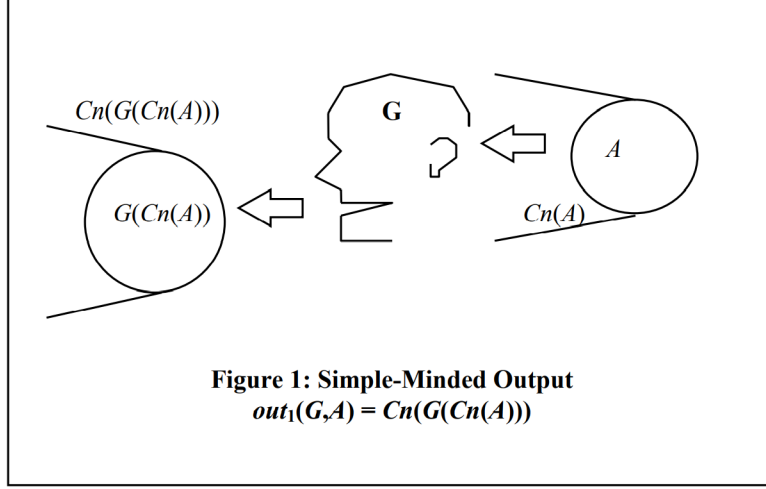
### 4.1 Definition

Technically, a normative code is seen as a set  $G$  of conditional norms, i.e. a set of such ordered pairs  $(a, x)$ . For each such pair, the body  $a$  is thought of as an input, representing some condition or situation, and the head  $x$  is thought of as an output, representing what the norm tells us to be desirable, obligatory or whatever in that situation.



$Cn(A)$  denotes the set of logical consequences of  $A$  in classical propositional logic. It returns the set of all provable propositional formulae provable assuming the fact in  $A$  (the set of answers of a set of inputs). For example,  $A = \{x, y\}$  then  $Cn(A) = \{x, y, x \vee y, x \wedge y \vee \dots\}$ . The simplest kind of input/output operation is depicted in Figure 1. A set  $A$  of

propositions serves as explicit input, which is prepared by being expanded to its classical closure  $Cn(A)$ . This is then passed into the ‘black box’ or ‘transformer’  $G$ , which delivers the corresponding immediate output  $G(Cn(A)) = \{x: \text{for some } a \in Cn(A), (a,x) \in G\}$ . Finally, this is expanded by classical closure again into the full output  $out_1(G,A) = Cn(G(Cn(A)))$ . We call this simple-minded output.



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

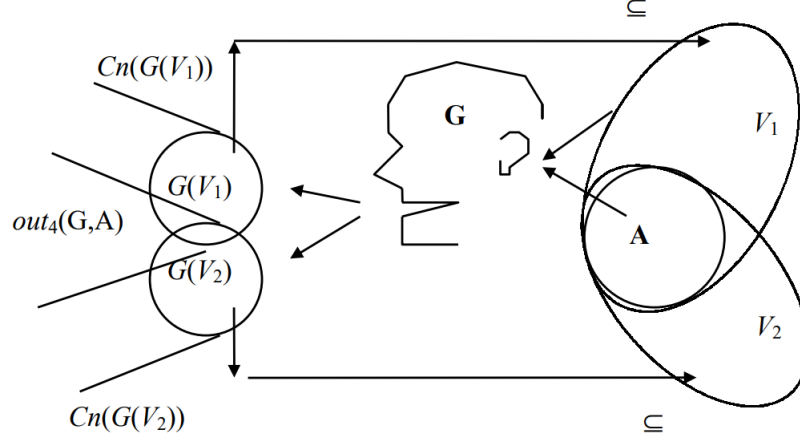
- Strengthening Input (SI): From  $(a,x)$  to  $(b,x)$  whenever  $a \in Cn(b)$
- Conjoining Output (AND): From  $(a,x), (a,y)$  to  $(a,x \wedge y)$
- Weakening Output (WO): From  $(a,x)$  to  $(a,y)$  whenever  $y \in Cn(x)$ .
- Disjoining input (OR): From  $(a,x), (b,x)$  to  $(a \vee b, x)$
- Cumulative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

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

- simple-minded alias  $out_1$  (as above)
- 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.





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

After the simple introduction of Input/Output logic, we reconsider our topic, the Contrary to duty obligations.

## 4.2 Contrary to duty reasoning

Like every other approach to deontic logic, input/output logic must face the problem of accounting adequately for the behaviour of ‘contrary-to-duty’ norms. The problem may be stated thus: given a set of norms to be applied, how should we determine which obligations are operative in a situation that already violates some among them. It appears that input/output logic provides a convenient platform for dealing with this problem by imposing consistency constraints on the generation of output.

### 4.2.1 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  $out^\cap$ :

$$out^\cap(G, A) = \cap\{out(H, A) | H \in maxfamily(G, A, A)\}$$

### 4.2.2 Contrary to duty reasoning

Let  $G$  a “Chisholm norm set”;

$x$  means the norm that a man goes to the assistance 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

Hence, the Input/output logic is able to handle the CTD paradox, which is a proposed solution of it.

## 5 Counterfactual Deontic Logic

Counterfactual Deontic Logic[6], which is introduced by R nnedal in 2016. The Counterfactual Deontic Logic combine the Counterfactual Logic with Deontic Logic. It aims at the relationship between the obligation and time.

### 5.1 Definition

- $A \Box \rightarrow B$  is often read ‘If  $A$  were the case, then  $B$  would be the case’
- $R$  is a temporal operator; ‘ $Rt_1 A$ ’ says that it is realised at time  $t_1$  (it is true at  $t_1$ ) that  $A$

### 5.2 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).

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**  $Rt_1ORt_2k$

**CF2**  $Rt_1O(Rt_2k\Box \rightarrow Rt_3\neg a)$

**CF3**  $Rt_1(Rt_2\neg k\Box \rightarrow Rt_2ORt_3a)$

**CF4**  $Rt_1Rt_2\neg k$

We can see:

- 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 is no sentence in this set follows from the rest.

### 5.3 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

## 6 Alternative approaches

In this section, we summarize easily the other Analyse of other researchers.

Carmo and Jones [7] 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.

Another approach to Chisholm's paradox is to detach both obligations of the dilemma  $\neg m \wedge m$ , and represent them consistently using some kind of minimal deontic logic, for example using techniques from paraconsistent logic. From a practical reasoning point of view, a drawback of this approach is that a dilemma is not very useful as a moral cue for action. Moreover, intuitively it is not clear that the example presents a true dilemma. More about the dilemmas will be in

Section 9.

A recent representation of Chisholm's paradox is to replace deontic 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  $\neg m$ .

$\bigcirc(m|f), f \Rightarrow \bigcirc m$  FD

$\bigcirc(\neg m|\neg f), \bigcirc\neg f \Rightarrow \bigcirc(\neg m \wedge \neg f)$  ADD

The limitation of this approach is that we can no longer accept the principle of weakening (also known as inheritance).  $\bigcirc(\neg m \wedge \neg f|\top) \Rightarrow \bigcirc(\neg m|\top)$

Nowadays more and more logic researchers focus on the CTD Paradoxes, and build different deontic logic systems. When a new deontic logic is proposed, the traditional contrary-to-duty examples are always the first benchmark examples to be checked.

## 7 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 Paradoxes is that how can we find an adequate formula to deal with the relationship between **obligation** and **condition**. 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:

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

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:

- $\neg f \rightarrow \bigcirc\neg m$

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

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...

## References

- [1] G. Pigozzi and L. van der Torre. Multiagent deontic logic and its challenges from a normative systems perspective. page 2929–2993, 2017.
- [2] R.M. Chisholm. Contrary-to-duty imperatives and deontic logic. *analysis* 24:33–36. 1963.
- [3] D. Grossi J. van Benthem and F. Liu. Priority structures in deontic logic. 2014.
- [4] G.H. von Wright. Deontic logic. *mind*,60:1-15. 1951.
- [5] D. Makinson and L. van der Torre. Input-output logics. *journal of philosophical logic*. page 29(4):383–408, 2000.
- [6] Daniel Rønneidal. Contrary-to-duty paradoxes and counterfactual deontic logic. 2019.
- [7] J. Carmo and A.J.I. Jones. Deontic logic and contrary-to-duties. 2007.