Forrester's paradox using typicality

Julian Chingoma and Thomas Meyer

CAIR and University of Cape Town, South Africa

Abstract. Deontic logic is a logic often used to formalise scenarios in the legal domain. Within the legal domain there are many exceptions and conflicting obligations. This motivates the enrichment of deontic logic with a notion of typicality which is based on defeasibility, with defeasibility allowing for reasoning about exceptions. Propositional Typicality Logic (PTL) is a logic that employs typicality. Deontic paradoxes are often used to examine logic systems as they provide undesirable results even if the scenarios seem intuitive. Forrester's paradox is one of the most famous of these paradoxes. This paper shows that PTL can be used to represent and reason with Forrester's paradox in such a way as to block undesirable conclusions without sacrificing desirable deontic properties.

1 Introduction

Logic has for a long time been used to formalise legal norms and study legal reasoning [5]. The difference between what is the case and what should be the case is fundamental to law and this naturally translates to deontic logic and its notions of obligation, permission and prohibition. This paper is part of a research study which focuses on introducing the notion of defeasibility into a deontic setting. Defeasibility allows for reasoning about exceptions in a domain, distinguishing between what is normally the case and what is actually the case. It is important to note there are already notions of defeasibility in the world of legal reasoning. The introduction of new information or a new regulation can cause laws to conflict and/or present exceptions which make existing laws inapplicable [5]. Therefore the combination of these notions is worth exploring, Typicality is based on defeasibility and is a notion used in Propositional Typicality Logic (PTL) where its extra expressivity makes it a more powerful version of defeasibility [1]. In deontic logic research, it is common for systems to be validated with the use of deontic paradoxes [18]. One of the more famous of these paradoxes is Forrester's paradox also known as the Gentle Murder paradox [15, 18]. The semantic connection between deontic logic and the logic systems of defeasibility and typicality will be discussed later in the paper. This paper will present the paradox and also examine the effectiveness of PTL when applied to the paradox.

Now we outline the structure of the paper. Firstly we present propositional logic as this is the logic that forms the foundation of the two logic systems will be working with. We then detail these two logic systems, deontic logic and propositional typicality logic. The deontic logic section will be where Forrester's paradox and its issues is detailed. Once these have been detailed we then look at the analysis of Forrester's paradox. Finally, we present the conclusions.

2 Propositional Logic

Propositional logic is a logic used to formalise statements that can either be true or false [7]. These statements are usually represented using propositional letters such as p,q and r. Given a set of propositional letters Φ, the language of propositional logic can be represented with the following constants and operators [7, 14, 15]. $\bot$ is a constant which represents a contradiction, $\neg$ and $\wedge$ are operators which represents negation and conjunction respectively. The generation of $\vee$, $\rightarrow$, $\leftrightarrow$ and $\top$ can be done in the usual way using the other parts of the language [14, 15]. Since the reasoning aspect is of interest to us it is important to mention the notion of entailment. Entailment refers to what conclusions logically follow from a set of premises and will be presented more formally in a later section [1].

3 KLM-style defeasible reasoning

We now briefly present a logic system which is a form of defeasible reasoning, which allows for conclusions to be retracted and therefore allows for dealing with exceptions. The logic often referred to as KLM approach is an enriched version of propositional logic with defeasible implications of the form $\alpha \mid\!\sim \beta$ [2, 3, 8]. Defeasible implications will then represent implications that we can reject in exceptional circumstances and are read as "α typically/usually implies β". Defeasible entailment, $\approx \alpha \mid\!\sim \beta$, means that all the minimal valuations that satisfy α also satisfy β [3, 8]. Minimal valuations will be detailed formally in the following section.

4 Deontic Logic

This section will formally present deontic logic and the specific logic system we will investigate. Deontic Logic is a field of logic which formalises normative concepts. These concepts include obligation ("what is an individual's duty", "what an individual ought to do"), permission ("what an individual may do") as well as other related concepts such as prohibition ("what an individual is forbidden from doing") [6, 13, 15]. The system we will be working with is the traditional Dyadic Standard Deontic Logic (DSDL) approach [12, 14, 15] although there are alternative approaches to deontic logic such as input/output logic [6, 9]. The reason we opted for the more traditional approach was that it has semantics based on valuations, similar to that of the other logic systems we deal with in the research study [12, 15].

4.1 Language

Given a set of propositional letters Φ, the language of Dyadic Standard Deontic Logic (DSDL) can be represented with the following operator added to the propositional logic language [14, 15]: the $\bigcirc$-operator is added which represents obligation. This operator in DSDL better handles conditional obligations such as "if p is true then you must do q". Such statements can be represented using the "$|$" notation which is usually seen in conditional probability. The above example would be translated to $\bigcirc(q \mid p)$ in DSDL. Since many legal statements are of the conditional form, the conventional DSDL will be the logic used when we are dealing in the deontic environment instead of Standard Deontic Logic (SDL) which does not have the "$|$" mechanism for conditional obligations. The notion of permission is related to obligation by $\mathrm{P}p = \neg\bigcirc\neg p$ and that of prohibition being similarly related by $\mathrm{F}p = \bigcirc\neg p$. $\mathrm{P}p$ is to be read as "p is permitted" while $\mathrm{F}p$ can be read as "p is prohibited/forbidden" [14, 15]. Obligations without a conditional can be written in the conditional form in the following manner $\bigcirc p = \bigcirc(p|\top)$ [15].

4.2 Semantics

We can now formally define the preference-based semantics for DSDL as similarly presented by Parent et al. [14] and Pigozzi et al. [15]. We have preference models defined as $M = (V, \leq)$ where $V \subseteq W$, with W being a non-empty set of possible valuations. Possible valuations for a knowledge base containing the propositional letters p and q are $\{\{p, q\}, \{p, \neg q\}, \{\neg p, q\}, \{\neg p, \neg q\}\}$, where $\{p, \neg q\}$ is a valuation where p is true and q is false. Note that we will not allow for duplicate valuations. $\leq$ is not only a binary relation over V but a total preorder as it is reflexive, transitive and connected. The operator $\models$ represents the satisfaction of a formula. Given a model M and an $s \in V$ we can define the satisfaction of formulas in the language, $M, s \models p$ as follows [14]:

- $M, s \models p$ iff p is true in the valuation s
- $M, s \models \neg p$ iff not $M, s \models p$, as in p is false in s
- $M, s \models p \wedge q$ iff $M, s \models p$ and $M, s \models q$, as in p and q are both true in s
- $M, s \models \bigcirc(q \mid p)$ iff $\forall s'$, if $s' \in \{s \in \parallel p \parallel : s \leq t, \forall t \in \parallel p \parallel\}$, then $M, s' \models q$. Here $\parallel p \parallel = \{s \in W : M, s \models p\}$ and $s < s'$ means that $s \leq s'$ and $s' \not\leq s$. So $\bigcirc(q \mid p)$ means that given p being true, then only if the "minimal" or "most typical" valuations that satisfy p also satisfy q can we then can derive that q is obligatory

4.3 Properties

The following is an outline of some of the desirable deontic properties that commonly occur in deontic logic literature [4, 12, 15, 18]. These properties were chosen because they were presented as being important or at least relevant when assessing the usefulness of deontic logic systems. Thus they should be seen as properties that an ideal system of deontic logic would have. Note, this is not a full list of properties that can seem desirable for a deontic logic nor are they necessary properties for a reasonable deontic system. These are simply those needed for the analysis of Forrester's paradox in the paper..

Ought Implies Can $\neg \bigcirc (\alpha \wedge \neg\alpha)$

This property could also be represented as $\neg\bigcirc\bot$ as the conjunction of conflicting tasks, $\alpha \wedge \neg\alpha$, will be a logical contradiction and can therefore be represented by $\bot$. The property states that it is undesirable for contradictory tasks such as α and $\neg\alpha$ to be obligatory. Without *"ought implies can"*, the derivation of a contradiction, e.g $\bigcirc\bot$, would be acceptable and simply indicate that there has been a violation.

Factual Detachment If we have $\bigcirc(\beta \mid \alpha)$ and α then we can derive $\bigcirc\beta$

If we have an obligation to do a task β when α is satisfied, once we have that α has happened then it is intuitive that we are now obligated to do β.

Restricted Strengthening of the Antecedant If we have $\bigcirc(\beta \mid \alpha)$ then we can derive $\bigcirc(\beta \mid \gamma \wedge \alpha)$ if γ is true

Let's say we have the obligation to do β when α is true. It is intuitive that a more specific version of α being true would still make β obligatory. Note that this restricted version of the property requires $\alpha \wedge \gamma$ to be consistent. The property will also be referred to as RSA during this paper.

Conjunction If we have $\bigcirc(\beta \mid \alpha)$ and $\bigcirc(\gamma \mid \alpha)$ then we can derive $\bigcirc(\gamma \wedge \beta \mid \alpha)$

Let's say we have an obligation to do a task β when α is satisfied. And we also have an obligation to do γ when α is satisfied. By combining these two obligations, it is intuitive that we are now obligated to do both β and γ when we have α. We will be working with a restricted version of this property where we will require that $\beta \wedge \gamma$ be consistent.

Weakening If we have $\bigcirc(\beta \wedge \gamma \mid \alpha)$ then we can derive $\bigcirc(\beta \mid \alpha)$

Let's say that we have the obligation to do both γ and β when α is true. It is intuitive that we can derive an obligation to do only one of γ or β when α is satisfied. So Weakening can be applied in this scenario since we know $\beta \wedge \gamma \rightarrow \beta$.

4.4 Forrester's paradox

Forrester's paradox is one of the most frequently occurring paradoxes in the deontic logic literature [12, 15, 18]. One of the reasons that this paradox was chosen is that it is similar in structure to many other deontic examples as it is a contrary-to-duty scenario [18]. For obligations $\bigcirc(\alpha_1 \mid \beta_1)$ and $\bigcirc(\alpha_2 \mid \beta_2)$ we say that the second obligation is a contrary-to-duty obligation of the first if its antecedant β_2 is contradictory to the consequent of the first α_1. Intuitively, this means an obligation that informs us what must be the case when something forbidden has been done [16]. Another reason we look at the paradox is that it provides difficulties that the straightforward examples would not, as it has been a challenge for deontic logic researchers [10, 12, 18].

This paradox comprises three statements. *"You must not kill anybody"*, *"If you kill someone then you must kill them gently"* and *"You killed someone"*. With this we also have the background knowledge that *"Killing gently implies killing"* [12, 15, 18]. We will now detail two undesirable derivations that occur through the different combinations of the deontic properties on this paradox's set of statements. Both are presented as they illustrate different issues with the paradox and the properties. In the following figures, derivations of obligations are shown using an arrow with a subscript containing the abbreviation of the property which was used for the derivation. $\bigcirc(\beta \mid \alpha) \rightarrow_W \bigcirc(\gamma \mid \alpha)$ would mean the Weakening property was used to go from $\bigcirc(\beta \mid \alpha)$ to $\bigcirc(\gamma \mid \alpha)$. Weakening would be the applicable here if we have $\beta \rightarrow \gamma$. For derivations that involve more than one obligation as the premise, these obligations are displayed between braces and separated by a comma. $\{\bigcirc(\gamma \mid \alpha), \bigcirc(\beta \mid \alpha)\} \rightarrow_{Conj} \bigcirc(\gamma \wedge \beta \mid \alpha)$ means that the Conjunction property was used on the obligations $\bigcirc(\gamma \mid \alpha)$ and $\bigcirc(\beta \mid \alpha)$ to derive $\bigcirc(\gamma \wedge \beta \mid \alpha)$. The paradox's statements can be represented by the following deontic knowledge base: $\{\bigcirc\neg k, \bigcirc(g \mid k), k\}$

RSA, Weakening and Conjunction

1.	$\bigcirc\neg k \rightarrow_W \bigcirc\neg g$
2.	$\bigcirc\neg g \rightarrow_{RSA} \bigcirc(\neg g \mid k)$
3.	$\{\bigcirc(\neg g \mid k), \bigcirc(g \mid k)\} \rightarrow_{Conj} \bigcirc(\neg g \wedge g \mid k)$

The background knowledge is represented by $g \rightarrow k$. When we apply Weakening to the first obligation *"You must not kill anybody"*, we can then derive *"You must not kill gently"* since we have that *"Killing gently implies killing"* and the contrapositive that *"Not killing implies not killing gently"*. This is an intuitive derivation since killing gently is still killing, which we want to be forbidden. Then using RSA and the fact that *"You killed someone"*, we can go from *"You must not kill gently"* to *"If you kill then you must not kill gently"*. This derivation is the issue with the paradox with an obligation becoming the premise from which its own contrary-to-duty obligation is derived which is counter-intuitive [12, 15]. Then using Conjunction we can derive a contradiction from the obligations *"If you kill then you must not kill gently"* and *"If you kill then you must kill gently"*. If we have the aforementioned *"ought implies can"* property then this would be undesirable [12, 15, 18]. Without it, we would be satisfies with the derivation of a violation but *"ought implies can"* states we don't want to settle for a contradiction but rather to act as best as possible in the case of a violation [18].

Factual Detachment and Conjunction

1.	$\{\bigcirc(g \mid k),\ k\} \rightarrow_{FD} \bigcirc g$
2.	$\{\bigcirc\neg k,\ \bigcirc g\} \rightarrow_{Conj} \bigcirc(\neg k \wedge g)$

The rule of Factual Detachment gives us *"You should kill gently"* from the fact *"You have killed"* and the obligation *"If you kill then you should kill gently"*.

Applying Conjunction to *"You should kill gently"* and the non-conditional obligation *"You ought to not kill someone"* gives us *"You should not kill and you should kill gently"* which is an undesirable derivation if one was to use the *"ought implies can"* principle [12, 15]. And as in the previous derivation, if we do not have *"ought implies can"* then the derivation is not problem.

5 Propositional Typicality Logic

5.1 Language

Given a set of propositional letters Φ, the language of the propositional typicality logic, denoted by $\mathcal{L}^{\bullet}$, can be represented with the following $\bullet$-operator added to the propositional logic language [1]: There is $\bullet\alpha$ with its intuition being that it represents the most typical situations where α holds. Note that this means that PTL is more expressive than KLM-style logic [1] from section 3 and the bullet operator can be applied to both the antecedant and consequent side of a conditional. The following example illustrates how it can be used. $\bullet\alpha \rightarrow \bullet\neg\beta$ stands for "the most typical situations where α holds, imply the most typical situations where β does not hold". Note, this is a similar reading to the semantics to that of DSDL conditionals as stated in section 4.2.

5.2 Semantics

For the semantics of PTL, it is done using ranked interpretations. With W being the set of possible valuations, ranked interpretations are pairs $< V, \leq >$, where $V \subseteq W$ and $\leq$ is a total preorder over V. Intuitively, the valuations pushed lower down the rankings are more typical than those that are higher [1]. And given a ranked interpretation R and a formula α, the set of valuations that satisfy α are represented as $[\![\alpha]\!]^R$ [1]. Satisfaction of a formula is done in the classical way, such as in section 4.2, with the omission of the $\bigcirc$-operator satisfaction and the addition of the following [1]: $v \models \bullet\alpha$ iff $v \models \alpha$ and there is not a $v' \leq v$ such that $v' \models \alpha$. So the valuations that satisfy $\bullet\alpha$ will be the minimal valuations that satisfy α. So $[\![\bullet\alpha]\!]^{\mathscr{R}} := min_{\leq}([\![\alpha]\!]^{\mathscr{R}})$ for a ranked interpretation $\mathscr{R}$.

Note that the typicality $\bullet$-operator can express any KLM-style conditionals. That is, for every ranked interpretation $\mathscr{R}$ and every α, $\beta \in \mathcal{L}$, $\mathscr{R} \vDash \alpha \mid\!\sim \beta$ if and only if $\mathscr{R} \vDash \bullet\alpha \rightarrow \beta$. There are $\mathcal{L}^{\bullet}$-sentences that cannot be expressed using KLM-style $\mid\!\sim$-statements on $\mathcal{L}$, so the converse does not hold.[1]. Now the method of entailment we will use, which is proposed by Booth et al. [1], is outlined.

5.3 LM-entailment

The first form of entailment to be looked at is one that produces a single ranked model that is constructed to be the LM-minimum model for the knowledge base. A sequence of ranked interpretations ($\mathscr{R}_0$,$\mathscr{R}_1$,$\mathscr{R}_2$,...) constructed during the algorithm will be used to construct $\mathscr{R}^*_{\mathcal{KB}}$, which will be used for entailment. The algorithm will make use of ranks in order to construct $\mathscr{R}^*_{\mathcal{KB}}$. The ranks represent a level in the ranked interpretation, where the rank of a valuation u is less than the rank of v if and only if $u < v$, as defined in section 5.2 [1].

The following is some of the notation used during the algorithm. In this algorithm, we say $\mathscr{R}^1_S$ is the ranked interpretation obtained when any valuation not in S, where $S \subseteq V^{\mathscr{R}}$, has its rank increased by 1. Similarly, $\mathscr{R}^\infty_S$ is the ranked interpretation obtained from $\mathscr{R}$ by setting the rank of all valuations not in S to ∞ [1]. These would represent those at the highest level of $\mathscr{R}^*_{\mathcal{KB}}$ and deemed to be atypical. Now to present the steps in the algorithm [1].

Step 1 Set the ranks of all valuations in the knowledge base to 0, define S_0 which is initially empty and have variable i equal to 1.
Step 2 Find the valuations which satisfy the knowledge base with respect to the current ranked interpretation $\mathscr{R}_0$ and put them into the set S_i.
Step 3 If S_i is equal to S_{i-1} then there hasn't been a change so set the rank of all the valuations that do not satisfy the knowledge base with respect to $\mathscr{R}_i$ to ∞ and return the interpretation that remains.
Step 4 Otherwise create a new ranked interpretation $\mathscr{R}_i$, by increasing the rank of every valuation not in S_i by 1.
Step 5 Find the valuations which satisfy the knowledge base with respect to the current ranked interpretation $\mathscr{R}_i$ and put them in the set S_{i+1} and finally, increment i.
Step 6 Go to Step 3.

Example Now to present an example that illustrates the above steps. Let's take the knowledge base, $\{\bullet p \rightarrow \neg f, \bullet b \rightarrow f, p \rightarrow b\}$. The conditionals can be read as *"typical penguins do not fly"*, *"typical birds do fly"* and *"penguins are birds"*. Intuitively, the situations that are most reasonable given the information we have would be the situations where there are no penguins while the most typical birds do fly. Such a scenario would satisfy all the statements. It seems reasonable that the next best situation is when the most typical penguins don't fly while we can have that non-typical birds also don't fly. Then we can have that non-typical penguins do fly. The least desirable situations are when we have penguins that aren't birds at all as this violates a classical conditional, $p \rightarrow b$. Now to look if the reasoning matches our intuition.

Firstly we note that because of the last statement we can immediately discount the valuations $\{p, \neg b, f\}$ and $\{p, \neg b, \neg f\}$ as having infinite rank, and therefore on the highest level, as they will never satisfy the set of statements. So we begin by setting the rank of all the valuations to 0. The valuations that satisfy all the statements are $\{\neg p, b, f\}$, $\{\neg p, \neg b, f\}$ and $\{\neg p, \neg b, \neg f\}$. Therefore they become the first level of our model, $S_1 := [\![\mathcal{KB}]\!]^{\mathscr{R}_0} = \{\{\neg p, b, f\}, \{\neg p, \neg b, f\}, \{\neg p, \neg b, \neg f\}\}$. All the valuations not in S_1 obtain a rank of 1. The valuations that satisfy all the statements w.r.t. $\mathscr{R}_!$ are $\{p, b, \neg f\}$ and $\{\neg p, b, \neg f\}$. So we have $S_2 :=$

$[\![\mathcal{KB}]\!]^{\mathscr{R}_1} = \{\{p, b, \neg f\}, \{\neg p, b, \neg f\}\}$. The remaining valuation $\{p, b, f\}$ will be S_3 and $\{p, \neg b, f\}$ and $\{p, \neg b, \neg f\}$ will be S_4. As previously mentioned the valuations in S_4 will not satisfy the statements so S_4 will remain the same as S_5 and so on. The algorithm terminates at this stage. The ranked models for the Bird example generated during the execution of the LM-entailment algorithm are given in figure 1.

$\mathscr{R}_0$	0.	$\{\neg p, b, f\}, \{\neg p, \neg b, f\}, \{\neg p, \neg b, \neg f\}, \{p, b, \neg f\}$ $\{\neg p, b, \neg f\}, \{p, b, f\}, \{p, \neg b, f\}, \{p, \neg b, \neg f\}$

$\mathscr{R}_1$	1.	$\{p, b, \neg f\}, \{\neg p, b, \neg f\}, \{p, b, f\}, \{p, \neg b, f\}, \{p, \neg b, \neg f\}$
	0.	$\{\neg p, b, f\}, \{\neg p, \neg b, f\}\{\neg p, \neg b, \neg f\}$

$\mathscr{R}_2$	2.	$\{p, b, f\}, \{p, \neg b, f\}, \{p, \neg b, \neg f\}$
	1.	$\{p, b, \neg f\}, \{\neg p, b, \neg f\}, \{p, b, f\}, \{p, \neg b, f\}, \{p, \neg b, \neg f\}$
	0.	$\{\neg p, b, f\}, \{\neg p, \neg b, f\}\{\neg p, \neg b, \neg f\}$

$\mathscr{R}^*_{\mathcal{KB}}$	∞	$\{p, \neg b, f\}, \{p, \neg b, \neg f\}$
	2.	$\{p, b, f\}$
	1.	$\{p, b, \neg f\}, \{\neg p, b, \neg f\}$
	0.	$\{\neg p, b, f\}, \{\neg p, \neg b, f\}, \{\neg p, \neg b, \neg f\}$

Fig. 1. The ranked models for the Bird example generated during the execution of the LM-entailment algorithm. $\mathscr{R}^*_{\mathcal{KB}}$ is then the final model and gives us the entailment.

6 PTL Analysis

6.1 Representation

It is important to note that we restrict ourselves to the use of only a subset of PTL for this analysis. We will only allow PTL statements of the form, $\bullet\alpha \rightarrow \beta$ or $\bullet\alpha \rightarrow \bullet\beta$, where α and β could be any combination of the PTL language except $\bullet$-operator. The reason being that the examples we deal with can be represented reasonably with this limited language and this limiting also reduces the complexity of the analysis. Statements of the form $\alpha \rightarrow \bullet\beta$ do not have the intuitive reading we desire. Since we do not want the properties of β, whether they are the most typical or not, to apply to all α valuations. This is why we require that the antecedant have a bullet operator. We can represent the statements with the typicality bullet on antecedent side only where $\bullet\alpha \rightarrow \beta$ reads as *"the most typical α are β"*. As previously stated, bullets on the antecedant-side only make the conditionals equivalent to the KLM-style conditionals and this is examined in more detail in the greater research study. Thus we will use the alternative representation to examine typicality and its added expressive power. This is $\bullet\alpha \rightarrow \bullet\beta$ and can be read as *"the most typical α are the most typical β"*.

Since Forrester's paradox is a contrary-to-duty scenario, it would be reasonable to introduce a contrary-to-duty example and observe if the translation is sound. The obligations of the scenario are as follows: *"You should not be late for work"* and *"If you are late then you must apologise"*. These can be translated to $\{\bullet\top \rightarrow \bullet\neg l, \bullet l \rightarrow \bullet a\}$. The reading of the statements with bullets on both sides seems reasonable. This reading specifies that the most typical situations where one is late, $\bullet l$, must be the most typical situations where one apologises, $\bullet a$, as opposed to any general apologising scenario. So bullets on both sides seem to be reasonable for contrary-to-duty obligations and will be used to represent the paradox.

6.2 Properties

We check whether our restricted PTL satisfies the aforementioned desirable properties using LM-entailment. In other words, we check if the properties can be applied when we have obligations of the form similar to that of Forrester's paradox. We are not assessing whether these are general properties that are satisfied by PTL. Except for the "ought implies can" principle, we do the check for the different representations of obligations that we have, which are cases which involve non-conditional and conditional obligations. For each property, we present the knowledge bases and their corresponding LM-entailment models.

Ought Implies Can and Violations Let's say we have a knowledge base that contains the conditionals $\bullet\top \rightarrow \bullet\alpha$ and $\bullet\top \rightarrow \bullet\neg\alpha$. There will be no valuations that satisfy the knowledge base because of the conflicting conditionals, therefore we cannot reason with this knowledge base. This implies that we have the *"ought implies can"* property. Since having contradictory facts in the knowledge base stops us from using the LM-entailment reasoning, we will not have any facts in the knowledge base when using the LM-entailment algorithm. We will instead use facts after the LM-entailment algorithm constructs the ranked model. We will strip valuations from the model that contradict the facts we are presented with and then reason with the resultant model. This will give the best case scenario whenever an obligation has been violated.

Restricted Strengthening of the Antecedant We assume that we have $\bigcirc(\beta \mid \alpha)$ and then check if $\bigcirc(\beta \mid \gamma \wedge \alpha)$ can be derived.

1. We have $\{\bullet\top \rightarrow \bullet\beta\}$and ideally want to derive $\bullet\alpha \rightarrow \bullet\beta$ when α holds.

1	$\{\alpha,\neg\beta\}$,$\{\neg\alpha,\neg\beta\}$
0	$\{\alpha,\beta\}$,$\{\neg\alpha,\beta\}$

   In the case where α holds then it is clear that the most typical α valuation is also the most typical β valuation. This would be blocked if we had $\bullet\neg\alpha \rightarrow \bullet\beta$ or $\bullet\alpha \rightarrow \bullet\neg\beta$ in the knowledge base.
2. We have $\{\bullet\alpha \rightarrow \bullet\beta\}$ and ideally want to derive $\bullet(\alpha \wedge \gamma) \rightarrow \bullet\beta$ when γ holds.

1	$\{\alpha,\neg\beta, \gamma\}$,$\{\alpha,\neg\beta, \neg\gamma\}$
0	$\{\alpha,\beta, \gamma\}$,$\{\alpha,\beta, \neg\gamma\}$, $\{\neg\alpha,\beta, \gamma\}$, $\{\neg\alpha,\neg\beta, \gamma\}$, $\{\neg\alpha,\neg\beta, \neg\gamma\}$, $\{\neg\alpha,\beta, \neg\gamma\}$

In the case where γ holds then it is clear that the most typical $\alpha\wedge\gamma$ valuation is also the most typical β valuation. This would be blocked if we had $\bullet\neg(\alpha\wedge\gamma)\rightarrow\bullet\beta$ in the knowledge base.

Weakening We assume that we have $\bigcirc(\beta\wedge\gamma\mid\alpha)$ and then check if $\bigcirc(\beta\mid\alpha)$ can be derived.

1. We have $\{\bullet\top\rightarrow\bullet(\beta\wedge\gamma)\}$ and ideally want to derive $\bullet\top\rightarrow\bullet\beta$.

1	$\{\beta,\neg\gamma\}$, $\{\neg\beta,\gamma\}$, $\{\neg\beta,\neg\gamma\}$
0	$\{\beta,\gamma\}$

   It is clear that the most typical valuation is the most typical β valuation which allows for the derivation of $\bullet\top\rightarrow\bullet\beta$. The most typical valuation in the model is also the most typical γ valuation thus the derivation of $\bullet\top\rightarrow\bullet\gamma$ also holds.
2. We have $\{\bullet\alpha\rightarrow\bullet(\beta\wedge\gamma)\}$ and ideally want to derive $\bullet\alpha\rightarrow\bullet\beta$.

1	$\{\alpha,\neg\beta,\gamma\}$,$\{\alpha,\neg\beta,\neg\gamma\}$, $\{\alpha,\beta,\neg\gamma\}$
0	$\{\alpha,\beta,\gamma\}$, $\{\neg\alpha,\beta,\gamma\}$, $\{\neg\alpha,\neg\beta,\gamma\}$, $\{\neg\alpha,\neg\beta,\neg\gamma\}$, $\{\neg\alpha,\beta,\neg\gamma\}$

   It is clear that the most typical α valuation, which is $\{\alpha,\beta,\gamma\}$, is also the most typical β valuation as well as the most typical γ valuation. This means that both $\bullet\alpha\rightarrow\beta$ and $\bullet\alpha\rightarrow\bullet\gamma$ also holds.

Factual Detachment We assume that we have $\bigcirc(\beta\mid\alpha)$ and α, and then check the if $\bigcirc\beta$ can be derived when using LM-entailment. There is only one case to look at as the non-conditional obligation check is trivial.

1. We have $\{\bullet\alpha\rightarrow\bullet\beta\}$ and ideally want to derive $\bullet\top\rightarrow\bullet\beta$ in when α holds.

1	$\{\alpha,\neg\beta\}$
0	$\{\alpha,\beta\}$, $\{\neg\alpha,\beta\}$, $\{\neg\alpha,\neg\beta\}$

   When α is true then the most typical valuation is $\{\alpha,\beta\}$ therefore the derivation holds.

Conjunction We assume that we have $\bigcirc(\beta\mid\alpha)$ and $\bigcirc(\gamma\mid\alpha)$, and then check if $\bigcirc(\beta\wedge\gamma\mid\alpha)$ can be derived. The cases with non-conditional obligations aren't checked since they will equivalent to Deontic Detachment.

1. We have $\{\bullet\top\rightarrow\bullet\beta,\bullet\top\rightarrow\bullet\gamma\}$ and ideally want to derive $\bullet\top\rightarrow\bullet(\beta\wedge\gamma)$.

1	$\{\beta,\neg\gamma\}$,$\{\neg\beta,\gamma\}$,$\{\neg\beta,\neg\gamma\}$
0	$\{\beta,\gamma\}$

   It is clear that we get $\bullet\top\rightarrow\bullet(\beta\wedge\gamma)$ as the best valuation is $\{\beta,\gamma\}$.
2. We have $\{\bullet\alpha\rightarrow\bullet\beta,\bullet\alpha\rightarrow\bullet\gamma\}$ and ideally want to derive $\bullet\alpha\rightarrow\bullet(\beta\wedge\gamma)$.

1	$\{\alpha,\neg\beta,\gamma\}$,$\{\alpha,\neg\beta,\neg\gamma\}$, $\{\alpha,\beta,\neg\gamma\}$
0	$\{\alpha,\beta,\gamma\}$, $\{\neg\alpha,\beta,\gamma\}$, $\{\neg\alpha,\neg\beta,\gamma\}$, $\{\neg\alpha,\neg\beta,\neg\gamma\}$, $\{\neg\alpha,\beta,\neg\gamma\}$

   'There is only one best α valuation and it is $\{\alpha,\beta,\gamma\}$ therefore we can derive $\bullet\alpha\rightarrow\bullet(\beta\wedge\gamma)$.

6.3 LM-entailment

We present the paradox once again and then translate it into a PTL version. The LM-entailment model is then presented and afterwards we show that the undesirable derivations, from section 4.4, can no longer be derived. This is despite the satisfaction of all the properties. The paradox's statements are translated into the following PTL knowledge base, $\{\bullet\top \rightarrow \bullet\neg k, \bullet k \rightarrow \bullet g\}$. With this knowledge base comes the background knowledge $g \rightarrow k$ and the fact k. The background knowledge means that the valuation $\{g, \neg k\}$ must be omitted from the model.

2	$\{\neg g, k\}$
1	$\{g, k\}$
0	$\{\neg g, \neg k\}$

Fig. 2. LM-entailment model for Forrester's paradox

RSA, Weakening and Conjunction Now using Weakening we can go from $\bullet\top \rightarrow \bullet\neg k$ to $\bullet\top \rightarrow \bullet\neg g$ as the model shows that the most typical valuations are $\neg g$ valuations. This is equivalent to the derivation of *"You must not kill gently"* from *"You must not kill anybody"* in section 4.4. But unlike in section 4.4, one cannot derive $\bullet k \rightarrow \bullet\neg g$ using RSA. The model blocks this derivation since the best k valuations are g valuations in this model.

Factual Detachment and Conjunction The issue presented in section 4.4 is blocked because once we assume the fact k in the model, the $\neg k$ valuations are removed as seen in the following model. The model shows that the derivation of $\bullet\top \rightarrow \bullet\neg k$, which means *"You must not kill anybody"*, is not possible, and thus when k is assumed the derivation of $\bullet\top \rightarrow \bullet(\neg k \wedge g)$ is blocked in the model.

7 Conclusion

The focus on this paper was to explore the extent that PTL can be used to deal with Forrester's paradox. After detailing the paradox and its issues, we presented PTL and the LM-entailment algorithm. Section 6.1 then presents how we represent the paradox using PTL. We then see that there is a way to use LM-entailment to solves the issues with Forrester's paradox. Section 6.2 shows that PTL satisfies the deontic properties we desire in our restricted environment. These are the properties which are the source of the paradox's issues. Section 6.3 then shows that the undesirable derivations from section 4.4 are avoided by the models produced by LM-entailment algorithm. This shows the potential that PTL possesses when applied in a deontic setting and this potential is explored further in the ongoing research study. This approach differs from other Forrester's paradox solutions such as those by Sinnott-Armstrong [17] and Meyer [11] in that it avoids the need for the expansion of the representative language using actions and/or logic quantifiers. Now the question to be asked is if PTL can be used on a variety of other examples to similar effectiveness.

References

1. Booth, R., Casini, G., Meyer, T., Varzinczak, I.: On the entailment problem for a logic of typicality. Proceedings of the Twenty-Fourth International Joint Conference on Artificial Intelligence (2015)
2. Casini, G., Straccia, U.: Lexicographic closure for defeasible description logics. In: Proc. of Australasian Ontology Workshop. vol. 969, pp. 28–39 (2012)
3. Giordano, L., Gliozzi, V., Olivetti, N., Pozzato, G.L.: Semantic characterization of rational closure: From propositional logic to description logics. Artificial Intelligence **226**, 1–33 (2015)
4. Goble, L.: Prima Facie Norms, Normative Conflicts, and Dilemnas, chap. 4, pp. 241–352. College Publications (2013)
5. Grossi, D., Rotolo, A.: Logic in the law: A concise overview (2012), `https://pdfs.semanticscholar.org/e1d1/5fe4cf3c29cb5417b6c89368f6c790fe3c60.pdf`
6. Hilpinen, R., McNamara, P.: Deontic Logic: A historical Survey and Introduction (2013)
7. Klement, K.: Propositional logic, `https://www.iep.utm.edu/prop-log/#SH3a`
8. Kraus, S., Lehmann, D., Magidor, M.: Nonmonotonic reasoning, preferential models and cumulative logics. Artificial Intelligence **44**, 167–207 (1990)
9. Makinson, D., van der Torre, L.: Input/output logics. Journal of Philosophical Logic **29**, 383–408 (2000)
10. McNamara, P.: Deontic logic (February 2006)
11. Meyer, J.J.: A simple solution to the "deepest" paradox inn deontic logic. Logique et Analyse **30**, 81–90 (1987)
12. Parent, X., van der Torre, L.: Detachment in normative systems: Examples, inference patterns, properties. IFCoLog Journal of Logic and its Applications **4**(9) (2017)
13. Parent, X., van der Torre, L., Sun, X.: Handout lecture 1:standard deontic logic, unpublished notes
14. Parent, X., van der Torre, L., Sun, X.: Handout lecture 2: Dyadic standard deontic logic
15. Pigozzi, G., van der Torre, L.: Multiagent deontic logic and its challenges from a normative systems perspective. IFCoLog Journal of Logic and its Applications **4**(9), 2929–2994 (2017)
16. Rönnedal, D.: Contrary-to-duty paradoxes and counterfactual deontic logic, `https://link.springer.com/content/pdf/10.1007%2Fs11406-018-0036-0.pdf`
17. Sinnott-Armstrong, W.: A solution to forrester's paradox of gentle murder. The journal of Philosophy **82.3**, 162–168 (1985)
18. van der Torre, L.: Reasoning About Obligations, Defeasibility In Preference-based Deontic Logic. Ph.D. thesis, Erasmus University Rotterdam (1997)