

Avoiding pragmatic oddity: A bottom-up defeasible deontic logic

GUIDO GOVERNATORI, *Centre for Computational Law, Yong Pung How School of Law Singapore Management University, 55 Armenian Street, Singapore 179943.*
E-mail: guido@governatori.net

SILVANO COLOMBO TOSATTO, *Data61, CSIRO, 41 Boggo Road, Dutton Park, 4102 Queensland, Australia.*
E-mail: silvano.colombotosatto@data61.csiro.au

ANTONINO ROTOLO, *Alma Mater Research Centre for Human Artificial Intelligence, University of Bologna, Via Galliera 3, 40121 Bologna, Italy.*
E-mail: antonino.rotolo@unibo.it

Abstract

This paper presents an extension of Defeasible Deontic Logic to deal with the Pragmatic Oddity problem. The logic applies three general principles: (i) the Pragmatic Oddity problem must be solved within a general logical treatment of contrary-to-duty (CTD) reasoning; (ii) non-monotonic methods must be adopted to handle CTD reasoning; (iii) logical models of CTD reasoning must be computationally feasible and, if possible, efficient. The proposed extension of Defeasible Deontic Logic elaborates a preliminary version of the model proposed by Governatori and Rotolo [15]. The previous solution was based on particular characteristics of the (constructive, top-down) proof theory of the logic. However, that method introduces some degree of non-determinism. To avoid the problem, we provide a bottom-up characterization of the logic. The new characterization offers insights for the efficient implementation of the logic and allows us to establish the computational complexity of the problem.

1 Introduction

A key difference between norms and other constraints is that, typically, norms can be violated. Moreover, normative systems (especially the legal ones) contain provisions about norms that become effective when violations occur. Since the seminal work by Chisholm [3] the obligations in force triggered by violations have been dubbed contrary-to-duty (CTD) obligations. The treatment of CTDs has proven problematic for formal (logical) representations of normative systems. Accordingly, CTDs are the source of many paradoxes and problems and also the driver for criticizing Standard Deontic Logic (SDL) and for the development of many new deontic formalisms (see [2, 4]).

One well-known problem of CTDs is the so-called Pragmatic Oddity paradox, which was introduced by Prakken and Sergot [24] and is illustrated by the following example.

EXAMPLE 1 There should be no dog. $O\text{-}d$ (1)

If there is a dog, then there ought to be a warning sign. $d \rightarrow Os$ (2)

There is a dog. d (3)

In SDL, we have both $O\neg d$ and O_s . However, according to Prakken and Sergot,

‘Surely, it is strange to say that in all ideal worlds there is no dog and also a warning sign that there is no dog. [...] This oddity—we might call it a “pragmatic oddity”—seems to be absent from the natural language version, which means that the SDL representation is not fully adequate.’ [24, pp. 96, 95]

The oddity of Example 1, its counter-intuitiveness, seems to depend on the fact that the two obligations $O\neg d$ and O_s are in force at the same time: when you fail to have no dog, you are obliged to have no dog and obliged to hang a warning sign. The solutions proposed by Prakken and Sergot [24] consist of representing (2) as $d \Rightarrow O_{d_s}$, where \Rightarrow is a suitable conditional operator. The sentence O_{d_s} means that ‘there is a secondary obligation that s , presupposing the sub-ideal context d ’. The problem, for Prakken and Sergot, is avoided because O_{d_s} does not imply O_s : ‘primary and CTD obligations are obligations of a different kind: a CTD obligation pertains to, or presupposes, a certain context in which a primary obligation is already violated’ [24, p. 91].

Prakken and Sergot’s analysis is thus based on two basic principles:

PRINCIPLE 1.

The Pragmatic Oddity problem must be solved within a general logical treatment of CTD reasoning.

PRINCIPLE 2.

Primary obligations and CTD obligations are of a different kind.

In fact, most of the work on Pragmatic Oddity (in addition to [24], see, among others [2])¹ focuses on the issue of how to distinguish the mechanisms leading to the derivation of the two individual obligations and create different classes of obligations insofar as they express different ideality levels. One solution is to prevent the conjunction when the obligations are from different classes. Accordingly, if the problem is to avoid having a conjunctive obligation in force when the individual obligations are in force themselves, the simplest way is to have a deontic logic that does not support the aggregation axiom:

$$(Oa \wedge Ob) \rightarrow O(a \wedge b).$$

This solution, among other things, was discussed by [6]: adopting a non-normal deontic logic (i.e. weaker than **K**), each obligation semantically corresponds to a distinct norm that selects a set of ideal worlds (see [5]) and aggregation cannot be allowed.

However, as suggested by Horty [17] and also recalled by Goble [6], some restricted forms of agglomeration should be accepted: several examples seem to hold if CTDs are not considered.

Therefore, a more amenable option, as suggested by Parent and van der Torre [22, 23], is to admit aggregation for obligations that are independent of the violation of the other obligations. We agree with them. Indeed, in our view, what is odd is not that the two obligations are in force at the same time, but that if one admits forming a conjunctive obligation from the two individual obligations, then we get an obligation that is impossible to comply with.

¹ For excellent overviews of the literature, see [2, 7, 23].

Based on the intuition above, Governatori and Rotolo [15] proposed, in a preliminary work, an extension of Defeasible Deontic Logic [11] to handle Pragmatic Oddity. Their solution was based on the constructive proof theory of the logic with specific proof conditions: more specifically, to admit the derivation of $\mathbf{O}(a \wedge b)$ requires that $\mathbf{O}a$ and $\mathbf{O}b$ are already provable, and $\neg a$ does not appear in the derivation of $\mathbf{O}b$ (similarly for $\mathbf{O}a$ and $\neg b$).

The extension mentioned above of Defeasible Deontic Logic was based on Principle 1 by using the new non-classical operator \otimes : the reading of an expression like $a \otimes b \otimes c$ is that a is primarily obligatory, but if this obligation is violated, the secondary obligation is b , and, if the secondary (CTD) obligation b is violated as well, then c is obligatory (see [11]). This approach falls within a proof-theoretic line of inquiry about CTDs, which clearly distinguishes in the language and the logic structures representing norms from those representing obligations [19–21].

We also complied with another principle:

PRINCIPLE 3.

Non-monotonic methods must be adopted to handle CTD reasoning.

This principle was notably defended by Horty [17] and van der Torre and Tan [25], even though, according to Parent and van der Torre [23] it seemed not directly involved in the Pragmatic Oddity problem. Since we stick to Principle 1, we also adopt Principle 3. However, we will see in Section 4 that the idea of defeasibility plays a role in some scenarios of Pragmatic Oddity, too.

Finally, our concern was computational since CTDs are so pervasive in normative reasoning (e.g. in the law):

PRINCIPLE 4.

Logical models of CTD reasoning must be computationally feasible and, possibly, efficient.

The solution we proposed in [15] was based on particular characteristics of the (constructive, top-down) proof theory of the logic. However, that method introduces some degree of non-determinism, insofar as the solution requires the existence of a proof satisfying certain conditions, and alternative proofs are possible.

The contribution of this paper is to present a new logical framework for the Pragmatic Oddity problem, which

- revises the solution we advanced in [15],
- complies with the same general principles we set in [15],
- provides a bottom-up characterization of the logic that avoids the problem with the top-down solution,
- studies the complexity of the problem (the resulting logic is computationally feasible, i.e. polynomial in the size of the input theory); and offers insights for the efficient implementation of the logic.

The layout of the article is as follows. Section 2 offers a high-level introduction to Defeasible Deontic Logic, while Section 3 presents the technical details of the Defeasible Deontic Logic used in the paper, a logic equipped with the \otimes -operator to identify pragmatic oddity instances. Section 4 discusses some examples and scenarios of pragmatic oddity. Section 5 provides a bottom-up characterization of the logic. Section 6 studies the computational complexity of the problem of computing whether a conjunctive obligation is derivable from a given defeasible theory. The paper ends with some brief conclusions.

2 A gentle overview of Defeasible Deontic Logic

This section provides a gentle overview of Defeasible Deontic Logic and how to use it for normative reasoning (see also Section 3). For more detailed presentations of the logic and its uses to model different aspects of normative reasoning, we refer the readers to [9, 11, 14, 16].

Defeasible Deontic Logic [11] is a sceptical computationally oriented rule-based formalism designed for the representation of norms. The logic extends Defeasible Logic [1] with deontic operators to model obligations and (different types of) permissions and provides an integration with the logic of violation developed by Governatori and Rotolo [13]. The resulting formalism offers features for the natural and efficient representation of exceptions, constitutive and prescriptive rules and compensatory norms. The logic is based on a constructive proof theory that allows for full traceability of the conclusions and flexibility to handle and combine different facets of non-monotonic reasoning.

Knowledge in Defeasible Logic is structured in three components:

- A set of facts (corresponding to indisputable statements represented as literals, where a literal is either an atomic proposition or its negation).
- A set of rules. A rule establishes a connection between a set of premises and a conclusion. In particular, for reasoning with norms, it is reasonable to assume that a rule provides the formal representation of a norm (though, it is possible to have norms that are represented by a set of rules). Accordingly, the premises encode the conditions under which the norm is applicable, and the conclusion is the normative effect of the norm.
- A preference relation over the rules. The preference relation just gives the relative strength of rules. It is used in contexts where two rules with opposite conclusions fire simultaneously to determine that one rule overrides the other in that context.

The rules establish a relationship between a set of premises (the antecedent) of a rule and a conclusion. We can classify rules based on (1) the strength of the relationship and (2) the type of relationship, more precisely, the type (or mode) of conclusion or effect a rule produces. Accordingly, a rule is an expression

$$a_1, \dots, a_n \leftrightarrow \square c, \quad (4)$$

where a_1, \dots, a_n is the antecedent, c is the conclusion, \leftrightarrow indicates the strength and \square the mode. For the strength Defeasible Logic provides three kinds of rules: *strict rules* (represented by \rightarrow), *defeasible rules* (represented by \Rightarrow) and *defeaters* (represented by \rightsquigarrow). A strict rule is a rule in the classical sense; every time the antecedent holds, so does the conclusion. On the other hand, a defeasible rule can produce its effect (or conclusion) when it is applicable and when there are no (applicable) rules for the opposite or such rules are defeated (by stronger rules). Finally, defeaters are rules that do not directly produce a conclusion but prevent the opposite conclusion from holding.

For the type or mode of the conclusion, we distinguish between *constitutive rules* and *normative rules*. Constitutive rules are used to define terms as defined in the normative systems the rules are meant to formalize. Therefore, constitutive rules specify the institutional facts or statements that hold in a given situation. Thus, for example, the constitutive rule

$$person, age < 18y \Rightarrow minor$$

establishing the institutional fact that minors are persons whose age is less than 18 years (for the notation we drop the \square for constitutive rules). On the contrary, a normative rule determines the conditions under which the conclusion is in force as an obligation or permission (one of the two

modal operators of Defeasible Deontic Logic). Consider, for instance, the following two normative rules:

$$r_1: \text{vehicle, redLight} \Rightarrow_{\mathbf{O}} \text{stop} \quad (5)$$

$$r_2: \text{emergency, redLight} \Rightarrow_{\mathbf{P}} \neg \text{stop}. \quad (6)$$

The first, r_1 is a prescriptive rule (indicated by the obligation \mathbf{O} modality) prescribing the obligation to stop for vehicles approaching a set of red traffic lights. Thus, when the conditions set in the antecedent hold (*vehicle* and *redLight*), the rule allows us to conclude the obligation to stop $\mathbf{O}\text{stop}$ is in force. r_2 is a *permissive* rule derogating or establishing an exception to r_1 for emergency vehicles. When its antecedent holds, we can conclude that it is permitted not to stop ($\mathbf{P}\neg\text{stop}$). The two rules conflict with each other, and we can use the superiority relation to state that r_2 overrides r_1 , namely $r_2 > r_1$.

As we mentioned, a characteristic of normative reasoning is its ability to deal with violations and conditions triggered by them. To this end, Defeasible Deontic Logic extends the language with a compensation operator \otimes to form expressions like

$$c_1 \otimes c_2 \otimes \dots \otimes c_n$$

called *compensation chains*. Compensation chains are only allowed as the conclusion of prescriptive rules (and thus asserting that obligations are in force). Their meaning as proposed by Governatori and Rotolo [13] and further discussed by Governatori [8] is that $\mathbf{O}c_1$ is the primary obligation, and when violated (i.e. $\neg c_1$ holds), then $\mathbf{O}c_2$ is in force, and it compensates for the violation of the obligation of c_1 . Moreover, when $\mathbf{O}c_2$ is violated, then $\mathbf{O}c_3$ is in force, and so on until we reach the end of the chain when a violation of the last element is a non-compensable violation where the norm corresponding to the rule in which the chain appears is not complied with.

Defeasible Logic is a constructive logic. Hence, the kernel of the logic is its proof theory, and for every conclusion we draw from a defeasible theory we can provide a proof for it, giving the steps used to reach the conclusion. At the same time, the derivation gives a (formal) explanation or justification of the conclusion. Furthermore, the logic distinguishes between *positive* and *negative* conclusion, the strength of a conclusion and its mode. This is achieved by labelling each step in a derivation with a proof tag. A derivation is a (finite) sequence of (tagged) formulas, each obtained from the previous ones using inference conditions. The inference conditions are formulated as proof conditions mandating the conditions that the previous steps in a derivation have to satisfy to append a new conclusion as the next step of a derivation. We adopt the following notation for proof tags: $+$ and $-$ indicate whether we have a positive or negative conclusion, Δ and ∂ denote, respectively, a definite or a defeasible conclusion, and they are subscripted by the modal (deontic) operator describing the mode of the conclusion. For example, the meaning of the tagged literal $-\Delta_{\mathbf{C}}p$ is that we definitely refute p as an institutional fact.² This means that we explored all possible ways to prove p using constitutive rules and facts, and we failed to derive it. On the other hand, $+\partial_{\mathbf{O}}\neg p$ means that we have a defeasible derivation for $\neg p$, where the rule used to conclude is a prescriptive rule. Finally, we say that $\square p$ is provable if we have a positive derivation for p with mode \square . Accordingly, $\mathbf{O}p$ holds if we derive $+\partial_{\mathbf{O}}p$ (or the stronger $+\Delta_{\mathbf{O}}p$).

Defeasible derivations have a three-phase argumentation-like structure. To show that $+\partial_{\square}p$ is provable at step n of a derivation we have to³:

²Similarly to the notation used for rules we drop the subscript for constitutive conclusions.

³Here we concentrate on proper defeasible derivations.

1. give an argument for p (where the last rule is a rule for \square);
2. consider all counterarguments for p ; and
3. rebut each counterargument by either:
 - (a) showing that the counterargument is not valid;
 - (b) providing a valid argument for p defeating the counterargument.

In this context, in the first phase, an argument is simply a strict or defeasible rule for the conclusion we want to prove, where all the elements are at least defeasibly provable. In the second phase, we consider all rules for the opposite or complement of the conclusion to be proved. Here, an argument (counterargument) is not valid if the argument is not supported. Here ‘supported’ means that all the elements of the body are at least defeasibly provable.

Finally, to defeasibly refute a literal, we have to show that either, the opposite is at least defeasibly provable, or an exhaustive search for a constructive proof for the literal fails (i.e. there are no rules for such a conclusion, or all rules are either ‘invalid’ arguments or they are not stronger than valid arguments for the opposite).

3 A Defeasible Deontic Logic for pragmatic oddity

In this section, we present a variant of Defeasible Deontic Logic designed to deal with the issue of Pragmatic Oddity. More specifically, we show how the proof theory can be used to propose a simple and (arguably) elegant treatment of the problem at hand.

We restrict ourselves to the fragment of Defeasible Deontic Logic that excludes permission and permissive rules since they do not affect the way we prevent Pragmatic Oddity from occurring: Definitions 12 and 13, the definitions that describe the mechanisms we adopt for a solution to Pragmatic Oddity are independent of any issue related to permission. In addition, for the sake of simplicity and to better focus on the non-monotonic aspects that the logic offers, we use only defeasible rules and defeaters. However, the definitions can be used directly in the full version of the logic. Accordingly, we consider a logic whose language is defined as follows.

DEFINITION 1

Let PROP be a set of propositional atoms and \mathbf{O} the modal operator for obligation.

- The set $\text{Lit} = \text{PROP} \cup \{\neg p \mid p \in \text{PROP}\}$ is the set of *literals*.
- The *complement* of a literal q is denoted by $\sim q$; if q is a positive literal p , then $\sim q$ is $\neg p$, and if q is a negative literal $\neg p$, then $\sim q$ is p .
- The set of *deontic literals* is $\text{DLit} = \{\mathbf{O}l, \neg\mathbf{O}l \mid l \in \text{Lit}\}$.
- If $c_1, \dots, c_n \in \text{Lit}$, then $\mathbf{O}(c_1 \wedge \dots \wedge c_n)$ is a *conjunctive obligation*.

In the rest of the paper, when relevant to the discussion, we will refer to elements of Lit as plain literals, and often we will use the unmodified term ‘literal’ to indicate either a plain literal or a deontic literal.

We formally introduce the compensation operator \otimes . This operator is used to build chains of compensation called \otimes -expressions. The formation rules for well-formed \otimes -expressions are:

1. every literal $l \in \text{Lit}$ is an \otimes -expression;
2. if $c_1, \dots, c_k \in \text{Lit}$, then $c_1 \otimes \dots \otimes c_k$ is an \otimes -expression;
3. nothing else is an \otimes -expression.

Given an \otimes -expression A , the *length* of A is the number of literals in it. Given an \otimes -expression $A \otimes b \otimes C$ (where A and C can be empty), the *index* of b is the length of $A \otimes b$. We also say that b appears at index n in $A \otimes b$ if the length of $A \otimes b$ is n .

DEFINITION 2

Let Lab be a set of arbitrary labels. Every rule is of the type

$$r: A(r) \leftrightarrow C(r),$$

where

1. $r \in \text{Lab}$ is the name of the rule;
2. $A(r) = \{a_1, \dots, a_n\}$, the *antecedent* (or *body*) of the rule, is the set of the premises of the rule (alternatively, it can be understood as the conjunction of all the elements in it). Each a_i is either a literal, a deontic literal or a conjunctive obligation;
3. $\leftrightarrow \in \{\Rightarrow, \Rightarrow_{\mathbf{O}}, \rightsquigarrow, \rightsquigarrow_{\mathbf{O}}\}$ denotes the type of the rule. If \leftrightarrow is \Rightarrow , the rule is a *defeasible rule*, while if \leftrightarrow is \rightsquigarrow , the rule is a *defeater*. Rules without the subscript \mathbf{O} are constitutive rules, while rules with such a subscript are prescriptive rules.
4. $C(r)$ is the *consequent* (or *head*) of the rule. It is a single literal for defeaters and constitutive rules, and an \otimes -expression for prescriptive defeasible rules.

Recall that prescriptive rules are used to derive obligations.

Given a set of rules R , we use the following abbreviations for specific subsets of rules:

- R_d denotes the set of defeasible rules in the set R ;
- $R[q, n]$ is the set of rules where q appears at index n in the consequent.⁴ The set of rules where q appears at any index n is denoted by $R[q]$;
- $R^{\mathbf{O}}$ denotes the set of prescriptive rules in R , i.e. the set of rules with \mathbf{O} as their subscript;
- R^C denotes the set of constitutive rules in R , i.e. $R \setminus R^{\mathbf{O}}$.

The above notations can be combined. Thus, for example, $R_d^{\mathbf{O}}[q, n]$ stands for the set of defeasible prescriptive rules such that q appears at index n in the consequent of the rule.

EXAMPLE 2

Let us consider the following set of rules R :

$$\begin{array}{lll} r_1: f_1 \Rightarrow_{\mathbf{O}} a \otimes b & r_2: f_2, g_2 \Rightarrow_{\mathbf{O}} b \otimes c & r_3: f_3 \Rightarrow \neg a \\ r_4: d \rightsquigarrow_{\mathbf{O}} \neg a & r_5: \neg \mathbf{O}a \Rightarrow \neg b & r_6: \mathbf{O}a, \mathbf{O}b \Rightarrow_{\mathbf{O}} \neg c \\ & r_7: f_7 \Rightarrow d. \end{array}$$

The set of prescriptive rules $R^{\mathbf{O}}$ is $\{r_1, r_2, r_4, r_6\}$; accordingly, the set of constitutive rules $R^C = \{r_3, r_5, r_7\}$. Moreover, the set of prescriptive defeasible rules $R_d^{\mathbf{O}} = \{r_1, r_2, r_6\}$. The set of rules for $\neg a$, $R[\neg a]$ is $\{r_3, r_4\}$; notice that this set contains a prescriptive and a constitutive rule; the corresponding set of defeasible constitutive rules $R_d^C[\neg a] = \{r_3\}$. When we consider the index where a literal appears we have the following sets: $R[b, 1] = \{r_2\}$, $R[b, 2] = \{r_1\}$ and $R[b] = \{r_1, r_2\}$.

⁴Strictly speaking, the notion of index is defined for \otimes -expressions and not for literals; however, according to the construction rules for \otimes -expressions, a plain literal is an \otimes -expression.

DEFINITION 3

A *Defeasible Theory* is a structure $D = (F, R, >)$ where F , the set of facts, is a set of (plain) literals, R is a set of rules and $>$, the superiority relation, is a binary relation over R .

A theory corresponds to a normative system, i.e. a set of norms, where every norm is modelled by some rules; accordingly, we do not admit deontic literals in the set of facts; obligations are determined by norms, and hence, in our framework by prescriptive rules. If both rules fire, the superiority relation is used for conflicting rules, i.e. rules whose conclusions are complementary literals. We do not restrict the superiority relation: it just determines the relative strength between two rules.

DEFINITION 4

A *proof* (or derivation) P in a defeasible theory D is a linear sequence $P(1) \dots P(z)$ satisfying the proof conditions given in Definitions 8–13, and each $P(i)$, $1 \leq i \leq z$, is a *tagged expression*, i.e. an expression of one of the forms: $+\partial q$, $-\partial q$, $+\partial_{\mathcal{O}}q$, $-\partial_{\mathcal{O}}q$, $+\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m$ and $-\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m$.

The tagged literal $+\partial q$ means that q is *defeasibly provable* as an institutional statement, or in other terms, that q holds in the normative system encoded by the theory. The tagged literal $-\partial q$ means that q is *defeasibly refuted* by the normative system. Similarly, the tagged literal $+\partial_{\mathcal{O}}q$ means that q is *defeasibly provable* in D as an obligation or that $\mathcal{O}p$ is defeasibly provable. In contrast, $-\partial_{\mathcal{O}}q$ means that q is *defeasibly refuted* as an obligation, thus $\mathcal{O}p$ cannot be proved. For $+\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m$ the meaning is that the conjunctive obligation $\mathcal{O}(c_1 \wedge \dots \wedge c_m)$ is defeasibly derivable; and that a conjunctive obligation $\mathcal{O}(c_1 \wedge \dots \wedge c_m)$ is defeasibly refuted corresponds to $-\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m$. The initial part of length i of a proof P is denoted by $P(1..i)$.

Defining when a rule is applicable or discarded is essential to characterize the notion of provability for constitutive rules and then for obligations. A rule is *applicable* for a literal q if q occurs in the head of the rule and all elements in the antecedent have been defeasibly proved (eventually with the appropriate modalities). On the other hand, a rule is *discarded* if at least one of the modal literals in the antecedent has not been proved. However, as literal q might not appear as the first element in an \otimes -expression in the head of the rule, some additional conditions on the consequent of rules must be satisfied. Accordingly, we first define the case for a constitutive rule (body-applicable) before moving to the condition for prescriptive rules with \otimes -expressions (Definition 6, where a literal is applicable if the previous element is provable as an obligation but violated, meaning that its complement is derivable).

DEFINITION 5

Given a proof P , a rule $r \in R$ is *body-applicable* at step $P(n+1)$ iff for all $a_i \in A(r)$:

1. if $a_i = \mathcal{O}l$ then $+\partial_{\mathcal{O}}l \in P(1..n)$;
2. if $a_i = \neg\mathcal{O}l$ then $-\partial_{\mathcal{O}}l \in P(1..n)$;
3. if $a_i = \mathcal{O}(c_1 \wedge \dots \wedge c_m)$ then $+\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m \in P(1..n)$;
4. if $a_i = l \in \text{Lit}$ then $+\partial l \in P(1..n)$.

A rule $r \in R$ is *body-discarded* at step $P(n+1)$ iff $\exists a_i \in A(r)$ such that

1. if $a_i = \mathcal{O}l$ then $-\partial_{\mathcal{O}}l \in P(1..n)$;
2. if $a_i = \neg\mathcal{O}l$ then $+\partial_{\mathcal{O}}l \in P(1..n)$;
3. if $a_i = \mathcal{O}(c_1 \wedge \dots \wedge c_m)$ then $-\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m \in P(1..n)$;
4. if $a_i = l \in \text{Lit}$ then $-\partial l \in P(1..n)$.

DEFINITION 6

Given a proof P , a rule $r \in R^O[q, j]$ such that $C(r) = c_1 \otimes \cdots \otimes c_m$ is *applicable* for literal q at index j at step $P(n+1)$ (or, simply, applicable for q), with $1 \leq j < m$, in the condition for $\pm\partial_O$ iff

1. r is body-applicable at step $P(n+1)$; and
2. for all $c_k \in C(r)$, $1 \leq k < j$, $+\partial_O c_k \in P(1..n)$ and $+\partial \sim c_k \in P(1..n)$.

Condition (1) represents the requirements on the antecedent stated in Definition 5; condition (2) on the head of the rule states that each element c_k before q has been derived as an obligation and a violation of such obligation has occurred.

DEFINITION 7

Given a proof P , a rule $r \in R^O[q, j]$ such that $C(r) = c_1 \otimes \cdots \otimes c_m$ is *discarded* for literal q at index j at step $P(n+1)$ (or, simply, discarded for q), with $1 \leq j \leq m$, in the condition for $\pm\partial_O$ iff

1. r is body-discarded at step $P(n+1)$; or
2. there exists $c_k \in C(r)$, $1 \leq k < l$, such that either $-\partial_O c_k \in P(1..n)$ or $-\partial \sim c_k \in P(1..n)$.

In this case, condition (2) ensures that an obligation before q in the chain is not in force or has already been fulfilled (thus, no reparation is required).

We now introduce the proof conditions for $\pm\partial$ and $\pm\partial_O$:

DEFINITION 8

The proof condition of *defeasible provability for an institutional statement* is

$+\partial$: If $P(n+1) = +\partial q$ then

- (1) $q \in F$ or
 - (2.1) $\sim q \notin F$ and
 - (2.2) $\exists r \in R_d[q]$ such that r is applicable for q , and
 - (2.3) $\forall s \in R[\sim q]$, either
 - (2.3.1) s is discarded for $\sim q$, or
 - (2.3.2) $\exists t \in R[q]$ such that t is applicable for q and $t > s$.

As usual, we use the strong negation to define the proof condition for $-\partial$

DEFINITION 9

The proof condition of *defeasible refutability for an institutional statement* is

$-\partial$: If $P(n+1) = -\partial q$ then

- (1) $q \notin F$ and
 - (2.1) $\sim q \in F$ or
 - (2.2) $\forall r \in R_d[q]$: either r is discarded for q , or
 - (2.2) $\exists s \in R[\sim q]$, such that
 - (2.3.1) s is applicable for $\sim q$, and
 - (2.3.2) $\forall t \in R[q]$ either t is discarded for q or $t \not> s$.

The proof conditions for $\pm\partial$ are the standard conditions in Defeasible Logic, see [1] for the full explanations.

DEFINITION 10

The proof condition of *defeasible provability for obligation* is

- $+\partial_{\mathcal{O}}$: If $P(n+1) = +\partial_{\mathcal{O}}q$ then
- (1) $\exists r \in R_q^{\mathcal{O}}[q, i]$ such that r is applicable for q , and
 - (2) $\forall s \in R^{\mathcal{O}}[\sim q, j]$, either
 - (2.1) s is discarded for $\sim q$, or
 - (2.2) $\exists t \in R^{\mathcal{O}}[q, k]$ such that t is applicable for q and $t > s$.

To show that q is defeasibly provable as an obligation, one must show that: the following two conditions must hold: (i) there must be a rule introducing the obligation for q , which can apply; (ii) every rule s for $\sim q$ is either discarded or defeated by a stronger rule for q . Observe that, since we do not admit deontic literals in F , we do not need the equivalent of conditions (1) and (2.1) for institutional statements to ensure that the logic is consistent.

The strong negation of Definition 10 gives the negative proof condition for obligation.

DEFINITION 11

The proof condition of *defeasible refutability for obligation* is

- $-\partial_{\mathcal{O}}$: If $P(n+1) = -\partial_{\mathcal{O}}q$ then
- (1) $\forall r \in R_q^{\mathcal{O}}[q, i]$ either r is discarded for q , or
 - (2) $\exists s \in R^{\mathcal{O}}[\sim q, j]$ such that
 - (2.1) s is applicable for $\sim q$, and
 - (2.2) $\forall t \in R^{\mathcal{O}}[q, k]$, either t is discarded for q or $t \not> s$.

Notice that, given the intended correspondence between $\mathcal{O}I$ and $+\partial_{\mathcal{O}}I$ (see Definition 5) we will refer to ‘the derivation of $\mathcal{O}I$ ’ when, strictly speaking, we should use ‘the derivation of $+\partial_{\mathcal{O}}I$ ’; similarly for when we say that $\mathcal{O}I$ has been refuted.

EXAMPLE 3

Let $D = (F, R, >)$ be a defeasible theory, where $F = \{f_1, f_2, g_2, f_3, f_7\}$, R is the set of rules given in Example 2, and $> = \{(r_6, r_2)\}$. D allows us to draw the following derivation:

- | | |
|---------------------------------|---|
| (1) $+\partial f_1$ | fact |
| ⋮ | |
| (5) $+\partial f_7$ | fact |
| (6) $+\partial d$ | from r_7 and $R^C[\sim f] = \emptyset$ |
| (7) $-\partial_{\mathcal{O}}a$ | from r_4 applicable and $r_1 \not> r_4$ |
| (8) $+\partial_{\mathcal{O}}b$ | from r_2 applicable and r_7 discarded |
| (9) $+\partial \neg b$ | from (7) r_5 applicable $-\partial_{\mathcal{O}}a \in P(1..8)$ |
| (10) $+\partial_{\mathcal{O}}c$ | from (8) and (9) r_2 applicable for c and (7) r_6 discarded |

Steps $P(1) \dots P(5)$: According to clause (1) of Definition 8 every fact is defeasible provable.

For $P(6)$ we have to satisfy the conditions given in Definition 8: We have that rule r_7 is applicable for d (clause 2.2), and $R[\neg d] = \emptyset$ satisfying clause (2.3) vacuously.

Step $P(7)$ follows from clauses (2.1) and (2.2) of Definition 11: Given that we have $+\partial d$ at step $P(6)$, rule r_4 is (body)-applicable, and the only (prescriptive) rule for a , rule r_1 is not stronger than r_4 .

The conclusion in step $P(8)$ is entailed by Definition 10: rule r_1 is body-applicable, but not applicable for b at index 2, since we have a at index 1, and $-\partial_{\mathcal{O}}a$ at $P(7)$; thus r_1 is discarded. However, r_2 is applicable for b at index 1 (the rule is clearly body-applicable given $A(r_2) \subset F$ and all facts are defeasibly provable). Thus, clause (1) holds. For clause (2), r_7 is (body)-discarded, $\mathcal{O}a \in A(r_7)$, we have $-\partial_{\mathcal{O}}a$ at $P(7)$, and there are no other rules in $R^{\mathcal{O}}[\neg b]$.

The justification for step $P(9)$ follows from $P(7)$ where we proved $-\partial_{\mathcal{O}}a$; thus r_5 is (body-)applicable (see item 2 of Definition 5); in addition $R^C[b] = \emptyset$.

Finally, for $P(10)$, as we have already argued r_2 is body-applicable, and we can use $P(8)$ and $P(9)$ to establish that the rule is applicable for c at index 2. In addition, $P(7)$ allows us to determine that r_6 is (body-)discarded since $\mathcal{O}a \in A(r_6)$, but the step proves $-\partial_{\mathcal{O}}a$ (item 2 of Definition 5, body-discarded part).

We are now ready to provide the proof condition under which a conjunctive obligation can be derived. The condition essentially combines two requirements. First, a conjunction holds only when all the conjuncts hold (individually). Second, the derivation of one of the individual obligations does not depend on the violation of the other conjunct. To achieve this, we determine the line of the proof when the obligation appears. Then we check that the negation of the other elements of the conjunction does not occur in the previous derivation steps.

DEFINITION 12

The proof condition of *defeasible provability for a conjunctive obligation* is

- If $P(n+1) = +\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m$, then
- $\forall c_i, 1 \leq i \leq m,$
 - (1) $+\partial_{\mathcal{O}}c_i \in P(1..n)$ and
 - (2) if $P(k) = +\partial_{\mathcal{O}}c_i, k \leq n$, then
 - $\forall c_j, 1 \leq j \leq m$ and $c_j \neq c_i, +\partial \sim c_j \notin P(1..k)$.

Again, the proof condition to refute a conjunctive obligation is obtained by strong negation from the condition to derive a conjunctive obligation defeasibly.

DEFINITION 13

The proof condition of *defeasible refutability for a conjunctive obligation* is

- If $P(n+1) = -\partial_{\mathcal{O}}c_1 \wedge \dots \wedge c_m$, then
- $\exists c_i, 1 \leq i \leq m$, such that either
 - (1) $-\partial_{\mathcal{O}}c_i \in P(1..n)$ or
 - (2) if $P(k) = +\partial_{\mathcal{O}}c_i, k \leq n$, then
 - $\exists c_j, 1 \leq j \leq m$ such that $c_j \neq c_i$ and $+\partial \sim c_j \in P(1..k)$.

In case of a binary conjunctive obligation, the positive proof condition boils down to

- $+\partial_{\mathcal{O}\wedge}$: If $P(n+1) = +\partial_{\mathcal{O}}p \wedge q$ then
- (1) $+\partial_{\mathcal{O}}p \in P(1..n)$ and
 - (2) $+\partial_{\mathcal{O}}q \in P(1..n)$ and

- (3) if $P(k) = +\partial_{\mathbf{O}p} (k \leq n)$, then $+\partial \sim q \notin P(1..k)$ and
- (4) if $P(k) = +\partial_{\mathbf{O}q} (k \leq n)$, then $+\partial \sim p \notin P(1..k)$.

Similarly, for the condition for $-\partial_{\mathbf{O}\wedge}$.

Consider a derivation where we have the following steps

$$\begin{array}{c} \vdots \\ P(x) + \partial_{\mathbf{O}a} \\ P(y) + \partial \neg a \\ P(w) + \partial_{\mathbf{O}b} \\ \vdots \end{array}$$

with $x < y < w < n$. Can we add $+\partial_{\mathbf{O}a \wedge b}$ at step $P(n + 1)$? Condition (1) of Definition 12 holds, but condition (2) does not, since we have $P(w) = +\partial_{\mathbf{O}b}$ and $P(y) = +\partial \neg a$, with $y < w$. On the contrary, if the derivation is

$$\begin{array}{c} \vdots \\ P(x) + \partial_{\mathbf{O}a} \\ P(y) + \partial_{\mathbf{O}b} \\ P(w) + \partial \neg a \\ \vdots \end{array}$$

both conditions hold and we can append $+\partial_{\mathbf{O}a \wedge b}$ to the derivation. In the second case, having $+\partial \neg a$ after the step where we concluded $+\partial_{\mathbf{O}b}$ ensures that the obligation of b does not depend on the violation of the obligation of a . Notice that in the first case, the order does not necessarily mean that $\mathbf{O}b$ depends on $\neg a$, but that the form of the derivation does not allow us to establish the independence of $\mathbf{O}b$ from $\neg a$.

Before proving some theoretical results about the logic, we give some examples to illustrate its behaviour.

4 Examples of Pragmatic Oddity scenarios

The scenarios in this section display some patterns of instances of Pragmatic Oddity and how they are dealt with based on the proof theory defined in the previous section. Moreover, as we will see, we use them to show a limitation of the proof theory: it introduces some non-determinism given that, in general, several derivations are possible and the order of the conclusion in a proof can affect what we can prove with specific orders.

In what follows, we use $\dots \Rightarrow c$ to refer to an applicable rule for c where we assume that the elements are not related (directly or indirectly) to the other literals used in the examples.

Compensatory Obligations The first case we want to discuss is when the conjunctive obligation corresponding to the Pragmatic Oddity has as conjuncts an obligation and its compensation. This scenario is illustrated by the rule:

$$\dots \Rightarrow_{\mathbf{O}} a \otimes b.$$

In this case, when the rule is applicable, we derive $+\partial_{\mathcal{O}}a$. Also, if $+\partial\neg a$ holds (signalling that the obligation of a has been violated), the rule is applicable for b at index 2 (condition 2 of Definition 6), and we can derive $+\partial_{\mathcal{O}}b$ (corresponding to $\mathcal{O}b$). Thus, we have the two individual obligations $\mathcal{O}a$ and $\mathcal{O}b$, but we cannot derive the conjunctive obligation of a and b , since the proof condition that allows us to derive $+\partial_{\mathcal{O}}b$ explicitly requires that $+\partial\neg a$ has been already derived. Accordingly, it is impossible to have the obligation of b without the violation of the obligation of a . Hence, we conclude $-\partial_{\mathcal{O}}a \wedge b$.

Contrary-to-duty The second case is when we have a CTD. The following two rules provide the classical representation of a CTD:

$$\dots \Rightarrow_{\mathcal{O}} a \quad \neg a \Rightarrow_{\mathcal{O}} b$$

In this case, it is possible to have situations when the obligation of b is in force without violating the obligation of a , namely, when a is not obligatory. However, as soon as we have $\mathcal{O}a$, we need to derive $\neg a$ to trigger the derivation of $\mathcal{O}b$ (Definition 5). Similarly to the previous case, we have $+\partial_{\mathcal{O}}a$ and $+\partial_{\mathcal{O}}b$, but we cannot conclude $+\partial_{\mathcal{O}}a \wedge b$; instead $-\partial_{\mathcal{O}}a \wedge b$ holds.

Pragmatic Oddity via Intermediate Concepts The situations in the previous two cases can be easily detected by a simple inspection of the rules involved; nevertheless, there could be more complicated cases. Specifically, when the second conjunct does not immediately depend on the first conjunct, but it depends on a reasoning chain. The following three rules illustrate the simplest structure for this case:

$$\begin{aligned} \dots &\Rightarrow_{\mathcal{O}} a \\ \neg a &\Rightarrow b \\ b &\Rightarrow_{\mathcal{O}} c. \end{aligned}$$

Here to derive $\mathcal{O}c$, we need first to prove b . To prove b , we require that $\neg a$ has already been proved. Again, it is possible to conclude $\mathcal{O}a$ and $\mathcal{O}c$, but not $\mathcal{O}(a \wedge c)$.

Negative Support In the previous case, the support was through an intermediate concept. However, given the non-monotonic nature of Defeasible Deontic Logic, we can have cases where the support is not to derive the other obligation directly from the violation. The violation prevents the derivation of the prohibition (or the permission of the opposite) of the other conjunct. Consider the following set of rules⁵:

$$\begin{aligned} \dots &\Rightarrow_{\mathcal{O}} a \\ \dots &\Rightarrow_{\mathcal{O}} b \\ c &\Rightarrow_{\mathcal{O}} \neg b \\ \dots &\Rightarrow c \\ \neg a &\rightsquigarrow \neg c \end{aligned}$$

To derive $\mathcal{O}b$, we have to ensure that the rule for $\mathcal{O}\neg b$ is discarded. This means that c should be rejected (i.e. $-\partial c$). We have two options: the rule for c is discarded, or the rule for $\neg c$ is applicable.

⁵ It is worth noting that, in the theory below, the rules for $\neg b$ and $\neg c$ can be either defeasible rules or defeaters producing the same result as far as the derivation of $\mathcal{O}(a \wedge b)$ is concerned.

This latter implies that to prove $+∂_O b$ we have to prove first $+∂_O \neg a$. Thus, one of the two elements of the conjunctive obligation $O(a \wedge b)$ depends on the violation of the other.

Iterated Conjunctive Obligations The two previous examples show that the dependency of one of the conjuncts from the violation can be negative and indirect. Now, the logic allows for conjunctive obligations in the body of rules, so the intermediate concept could be a conjunctive obligation itself (and we have to use the mechanism to determine the independence iteratively). Consider the following theory:

$$\begin{aligned} & \dots \Rightarrow_O a \\ r_2: & \dots \Rightarrow_O b \\ O(a \wedge b) & \Rightarrow_O c \\ & \dots \Rightarrow_O d \\ O(c \wedge d) & \Rightarrow e. \end{aligned}$$

Here, to prove e , we have to determine if the conjunctive obligation $O(c \wedge d)$ holds. Accordingly, we have to show that Oc and Od are derivable (and neither depends on the violation of the other). For Oc , the problem reduces to determining whether the conjunctive obligation $O(a \wedge b)$ obtains or not, where we have to repeat the procedure for Oa and Ob . Given the theory above, there are no dependencies on violations so that we can conclude e . Suppose that we replace r_2 with

$$r'_2: \neg d \Rightarrow_O b.$$

In this situation, we are still able to derive the four individual obligations, and the conjunctive obligation $O(a \wedge b)$; however, we are no longer able to conclude $O(c \wedge d)$ because Oc depends (indirectly) on the violation of Od .

Multiple Conjuncts In the previous scenario, we consider only cases of binary conjunctions. In this example, and in the next one, we are going to examine the situation of pragmatic oddity with conjunctions involving more than two conjuncts. The first set of rules to analyse is

$$\begin{aligned} & \dots \Rightarrow_O a \\ & \dots \Rightarrow_O b \\ \neg a, \neg b & \Rightarrow_O c. \end{aligned}$$

Clearly, to derive Oc we need both $\neg a$ and $\neg b$; thus, we derive $-∂_O a \wedge b \wedge c$, $-∂_O a \wedge c$ and $-∂_O b \wedge c$. Finally, as far as conjunctive obligations are concerned we can conclude $+∂_O a \wedge b$, noticing that $O(a \wedge b)$ is not a Pragmatic Oddity instance.

Multiple Dependencies In contrast to the example we just examined where Oc depended on the conjunction of the two violations, what if it depends on them disjunctively? Thus, we have the following theory.

$$\begin{aligned} r_1: & \dots \Rightarrow_O a \\ r_2: & \dots \Rightarrow_O b \\ r_3: & \neg a \Rightarrow_O c \\ r_4: & \neg b \Rightarrow_O c. \end{aligned}$$

Let us consider the derivation below:

(1) $+∂¬a$	fact
(2) $+∂_O a$	from r_1
(3) $+∂_O b$	from r_2
(4) $+∂_O c$	from (1) and r_3
(5) $-∂_O a ∧ b ∧ c$	from (1)–(4), $¬a ∈ P(1..4)$
(6) $+∂¬b$	fact
(7) $+∂_O b ∧ c$	from (3) and (4), $¬b ∉ P(1..4)$
(8) $-∂_O a ∧ c$	from (1) and (4), $¬a ∈ P(1..4)$.

We can carry out a similar proof by swapping the positions of $¬a$ and $¬b$, using r_4 in step (4)—yielding $-∂_O a ∧ b ∧ c$, and $-∂_O b ∧ c$ —but proving $O(a ∧ c)$. Hence, we have a situation where it is impossible to prove $O(a ∧ b ∧ c)$, but we can prove both $O(a ∧ b)$ and $O(b ∧ c)$, though it is impossible to have both of them in a single derivation.

Pragmatic Un-pragmatic Oddity What about when there are multiple norms both prescribing the CTD obligation and at least one of the norms is not related to the violation of the primary norm?

$$r_1: \dots \Rightarrow_O a \otimes b$$

$$r_2: \dots \Rightarrow_O b$$

$$¬a$$

In this situation, you can have a derivation:

(1) $+∂¬a$	fact
(2) $+∂_O a$	from r_1
(2) $+∂_O b$	from r_1 and (1) and (2),

where the derivation of $O b$ ($+∂_O b$) depends on the violation of the primary obligation of r_1 . In this case, we cannot derive the conjunctive obligation of a and b . However, there is an alternative derivation, namely:

(1) $+∂_O a$	from r_1
(2) $+∂_O b$	from r_2
(3) $+∂¬a$	fact
(4) $+∂_O a ∧ b$	from (1) and (2).

The proof demonstrates the independence of $O b$ from $¬a$, given that the derivation of $¬a$ occurs in a line after the line where $+∂_O b$ is derived.

Iterated Un-pragmatic Pragmatic Oddity We have seen cases where multiple derivations are possible, leading to opposite results about the derivability of instances of conjunctive obligations (irrespective of whether they are pragmatic oddity instances). Furthermore, a conjunctive obligation can depend on a pragmatic oddity instance. For example, the following set of rules illustrates a

situation where we have a derivation refuting an instance of pragmatic oddity, and a second one where the same instance is derivable. In turn, this instance can make a conjunctive obligation derivable or not.

$$\begin{aligned}
 r_1: \dots \Rightarrow_{\mathbf{O}} a \otimes b \\
 r_2: \dots \Rightarrow_{\mathbf{O}} a \\
 r_3: \dots \Rightarrow_{\mathbf{O}} b \\
 r_4: \dots \Rightarrow \neg a \\
 r_5: \mathbf{O}(a \wedge b) \Rightarrow_{\mathbf{O}} \neg c \\
 r_6: \dots \Rightarrow_{\mathbf{O}} c \otimes d \\
 r_7: \dots \Rightarrow_{\mathbf{O}} d.
 \end{aligned}$$

The key point of this example is that we have a conjunctive obligation, $\mathbf{O}(a \wedge b)$, in the antecedent of a prescriptive rule, r_5 , and there is a second rule, r_6 , for the opposite of the conclusion of r_5 .

- | | |
|---|---|
| (1) $+\partial \neg a$ | from r_4 |
| (2) $+\partial_{\mathbf{O}} a$ | from r_2 or r_1 |
| (3) $+\partial_{\mathbf{O}} b$ | from r_1 or r_3 |
| (4) $-\partial_{\mathbf{O}} a \wedge b$ | from (1), $+\partial \neg a \in P(1..3)$ |
| (5) $-\partial_{\mathbf{O}} \neg c$ | from (4) and r_5 |
| (6) $+\partial_{\mathbf{O}} c$ | from r_6 |
| (7) $+\partial_{\mathbf{O}} d$ | from r_7 |
| (8) $+\partial_{\mathbf{O}} c \wedge d$ | from (6) and (7), $+\partial \neg c, +\partial \neg d \notin P(1..7)$. |

This proof blocks the derivation of $\mathbf{O}(a \wedge b)$, since $+\partial \neg a$ occurs in P before $+\partial_{\mathbf{O}} b$. Consequently, we can derive the conjunctive obligation $\mathbf{O}(c \wedge d)$, since rule r_5 is discarded. However, if we postpone the use of r_4 , namely, doing the proof with the sequence

- | | |
|---|---|
| (1) $+\partial_{\mathbf{O}} a$ | from r_1 or r_2 |
| (2) $+\partial_{\mathbf{O}} b$ | from r_3 |
| (3) $+\partial_{\mathbf{O}} a \wedge b$ | from (1) and (2), $+\partial \neg a, +\partial \neg b \notin P(1..2)$ |
| (4) $-\partial_{\mathbf{O}} c$ | from (3) and r_5 |
| (5) $-\partial_{\mathbf{O}} c \wedge d$ | from (4) |

we are allowed to derive $+\partial_{\mathbf{O}} a \wedge b$, making r_5 applicable, preventing the derivation of $\mathbf{O}c$. Suppose that r_5 , instead of being a prescriptive rule, is a constitutive rule, namely

$$r_5: \mathbf{O}(a \wedge b) \Rightarrow \neg c,$$

enabling us to prove or refute the violation of the first element of r_6 . Using the first derivation in step (5), we conclude $+\partial \neg c$. Now, we have two ways to derive $\mathbf{O}d$: using r_6 (leading to an instance

of pragmatic oddity), or using r_7 , and we can postpone the derivation of $\neg c$, allowing us to assert $O(c \wedge d)$.

Mix and Match In all the previous cases the focus was on conjunctions where one of the conjuncts somehow depended on one of the other conjuncts. In other terms, the conjunction contains a primary obligation and a secondary obligation (an obligation in force after the violation of another obligation). Consider the rules

$$r_1: \dots \Rightarrow_O a \otimes b$$

$$r_2: \dots \Rightarrow_O c \otimes d$$

$$r_3: \dots \Rightarrow \neg a$$

$$r_4: \dots \Rightarrow \neg c.$$

Here, from r_1 , we obtain $Oa (+\partial_O a)$; similarly from r_2 we get Oc . r_3 and r_4 allow us to derive the literals corresponding to the violations of the two obligations, namely $+\partial \neg a$ and $+\partial \neg c$. Now, r_1 and r_2 applicable for their element at index 2. Hence, we conclude $Ob (+\partial_O b)$ and $Od (+\partial_O d)$; can we derive their conjunctive obligation? The answer is positive. Ob does not depend on the violation of Od (there is no way to derive $\neg d$ from the rules above) and the other way around. $O(b \wedge d)$ is a conjunctive obligation of two secondary obligations, what about other conjunctions, e.g. $O(a \wedge d)$ and $O(c \wedge b)$? Again the answer is positive: there is no need to derive $+\partial \neg a$ for the derivation of $+\partial_O d$; the argument for the second is the same.

5 A bottom-up characterization

The examples in the previous section illustrate cases where multiple derivations are possible and whether a conjunctive obligation is derivable or not depends on the specific derivation. Furthermore, the non-monotonicity of the logic presents other complications. Whether some conclusions are derivable depends on other elements being derivable, and these depend on specific derivations. Hence, we need to devise a mechanism that does not rely on a particular order in which a derivation sequence is laid out. The idea of the proof conditions for conjunctive obligations is to see that in the derivation of an obligation, the derivation of the violation of the other conjunct does not appear. Alternatively, we can say that an obligation is independent of the violation if we can push down in the proof the derivation of the violation. If the derivation is independent, then the rules to derive the violation do not contribute to the derivation of the obligation. Consequently, we could remove such rules without affecting the derivability of the obligation. Given that the derivations of the obligations of the conjuncts in a conjunctive obligation must be independent of the derivation of violations of those obligations and that when they are independent, we can run the derivations in parallel (using separate subsets of the rules), then we can inquire whether it is possible to carry out these derivations in a single construction. The answer is positive, and we can adapt the bottom-up construction of Maher and Governatori [18]. The idea of the bottom-up construction is that instead of working in a goal-directed fashion, we work in stages. For each stage, we determine all the conclusions that can be ‘derived’ at that stage, assuming that whatever was in a previous stage is already provable. Accordingly, we start from the empty set, and in the first stage, we determine what is provable from the empty set; then, for stage $n + 1$, we see what is provable from stage n . Before defining the extension of a defeasible theory, we need to provide a mechanism that guarantees that the derivation of an obligation does not depend on the violation of another obligation. To this end, we introduce a

construction, called reduction, that removes all rules for a particular element from a theory. In what follows, we are going to use the reduct to remove all rules for the violation of an obligation, and we are going to examine whether the other obligation is still derivable from the reduced theory. If it does, then the obligation does not depend on the violation.

DEFINITION 14

Given a defeasible theory $D = (F, R, >)$ and a set of (plain) literals $L = \{l_1, l_2, \dots\}$, the *reduct of D based on L* , noted as $red(D, L)$ is the defeasible theory $D' = (F', R', >)$ satisfying the following conditions:

1. $F' = F \setminus L$;
2. $R' = R \setminus \bigcup_{l \in L} R[l]$;
3. $>' = > \setminus \{(r, s) : r \notin R' \vee s \notin R'\}$.

The idea of the transformation is to create a theory similar to the original theory, as we said, without the literals in L . The condition on F is obvious. The second condition ensures that for each literal $l \in L$ the rules that can derive the literal are removed. Then the literal is no longer derivable since the resulting theory does not contain rules for the literal anymore. Given that $R'[l] = \emptyset$, the following result is immediate.

Observation 1.

Given a Defeasible Theory D , and a set L of literals, $-\partial l$ is derivable in $red(D, L)$ for $l \in L$.

It is worth noting that we do not have to remove rules where the literals in L appear in the antecedent of the rule. Such rules are immediately discarded. Similarly, for prescriptive rules where the complement of the removed literals appears in the head of the rules. Such rules are no longer applicable for elements appearing after one of the removed literals. Thus, if you have a rule with the \otimes -chain $c_1 \otimes \dots \otimes c_{n-1} \otimes \neg l \otimes c_{n+1} \dots$, the rule is in $R^O[c, m]$, but it is not applicable for any $m \geq n + 1$. Remember that to derive $+\partial_O c_{n+1}$ we have to prove both $+\partial_O \neg l$ and $+\partial l$.

EXAMPLE 4

Consider a theory D whose set of rules R consists of the rules presented in the Iterated Unpragmatic Pragmatic Oddity scenario described in the previous section including the additional rule $r'_5: \mathbf{O}(a \wedge b) \Rightarrow \neg c$. The reduct of D based on $L = \{-a\}$, $red(D, \{-a\})$ has the following set of rules $\{r_1, r_2, r_3, r_5, r'_5, r_6, r_7\}$. For $red(D, \{\neg g\})$, $R' = \{r_1, r_2, r_3, r_4, r_5, r_6, r_7\}$. Finally, for the reduct based on $L = \{-a, \neg c\}$, the resulting set of rules is $\{r_1, r_2, r_3, r_5, r_6, r_7\}$. Given that we removed rules for the elements in L , those literals cannot be derived positively; indeed, we derive $-\partial \neg a$ in $red(D, \{-a\})$, $-\partial \neg c$ in $red(D, \{\neg c\})$, and both of them in $red(D, \{-a, \neg c\})$;

We can now specify when a (deontic) literal is independent of a set of plain literals in Defeasible Deontic Logic

DEFINITION 15

Given a defeasible theory D , a set L of plain literals and a literal m , m is *independent from L* iff m is defeasibly provable in D and in $red(D, L)$.

We can now show that condition (2) in the proof conditions for a conjunctive obligation ensures the independence of the obligations from the violations. However, before proving this result, we have to recall a general property about Defeasible (Deontic) Logic: first of all, a defeasible theory

is consistent if F does not contain a literal l and its complement $\neg l$. Second, given a logical formula expressing a proof condition of the strong negation of the formula/conditions is obtained by replacing every occurrence of a positive proof tag with the corresponding negative proof tag, replacing conjunctions with disjunctions, disjunctions with conjunctions, existential with universal and universal with existential. Is it immediate to observe that each negative proof condition given in Section 2 is the strong negation of the corresponding positive one (and the other way around).

If corresponding proof conditions are defined using the principle of strong negation outlined above, then, given a derivation, it is impossible to have that the literal (conjunctive obligation) is both derivable and refutable in the same derivation.

PROPOSITION 1

[12] Given a consistent defeasible theory D , a derivation P , a literal l and proof tag $\# \in \{\partial, \partial_{\mathcal{O}}\}$, it is not possible that $+\#l, -\#l \in P$.

PROOF. The proof is an extension of the proof given in [12]. [12] proves that if the proof conditions for a pair of proof tags $+\#$ and $-\#$ are defined as the strong negation of each other, there is no theory D such that $+\#l$ and $-\#l$ both hold. The proof conditions for literals are the same as those in [11] and the result applies to them; and if the property holds for theories, it holds for individual proofs as well. The proof conditions for conjunctive obligations extend the constraints in [12] since they require that some elements are not in a derivation (clause (2)). Let us consider the case of the proof conditions for conjunctive obligations (Definitions 12 and 13). Suppose we have a derivation where we have both $+\partial_{\mathcal{O}}c_1, \dots, c_n$ at step n and $-\partial_{\mathcal{O}}c_1, \dots, c_n$ at step m . Let us assume that $m < n$ (the case $n < m$ is analogous). Thus, by clause (1) of Definition 12 we have that for all c_i , $+\partial_{\mathcal{O}}c_i \in P(1..n)$ (if an element is in $P(1..m)$ and $m < n$, the element is also in $P(1..n)$); if clause (1) of Definition 13 holds, then there is a c_i such that $-\partial_{\mathcal{O}}c_i \in P(1..n)$, contradicting the results for the other proof conditions. Thus clause (2) of Definition 12 must hold. Let k be the step where we derived $+\partial_{\mathcal{O}}c_i$. Now, $k \leq m < n$, and we have $+\partial \sim c_j \in P(1..k)$, while for clause (2) of Definition 12, it should be $+\partial \sim c_j \notin P(1..k)$, thus even in this case we obtain a contradiction. Accordingly, in any case, we have a contradiction; thus, it is impossible to derive and refute a conjunctive obligation in the same derivation. \square

Armed with this result, we can prove the result linking independence and the proof conditions for conjunctive obligations.

PROPOSITION 2

Given a consistent defeasible theory D , a deontic literal m and a set L of plain literals, m is independent from L iff there is a derivation P in D such that

1. $P(n) = +\partial_{\mathcal{O}}m$ and
2. $\forall l \in L, +\partial l \notin P(1..n)$.

PROOF. By the definition of independence (Definition 15), there is a derivation P' in $red(D, L)$ for $+\partial_{\mathcal{O}}m$. By construction of $red(D, L)$, for $l \in L$ $R[l] = \emptyset$; thus, we can add $-\partial l$ to any derivation in $red(D, L)$; by Proposition 1 $+\partial l$ is not derivable, and no derivation in $red(D, L)$ can contain it. What we have to do, is to show how to ensure that the derivation in $red(D, L)$ guarantees that there is a derivation in D . More specifically, we are going to give a (constructive) procedure to transform P' into a proof in D . The procedure removes all steps that do not contribute to the derivation of m . Let us assume that $P'(k) = +\partial_{\mathcal{O}}m$. We now consider the conclusion in $P'(k-1)$. If it is a justification for step $P'(k)$, we keep it; otherwise, we delete it. A conclusion (tagged literal or tagged conjunction) is a justification for a step of a derivation $P(n)$ if it makes discarded a rule attacking the conclusion in

$P(n)$ (see Definition 7), or it contributes to making applicable a rule for the conclusion in $P(n)$ (see Definition 6). We repeat step by step backward the procedure for all step $P'(k - z)$ to determine if they justify the step $P'(w)$, $k - z < w \leq k$ that have not been deleted in the previous iterations of the procedure. It is clear that the resulting sequence is still a proof for $+\partial_{\mathcal{O}m}$ in $red(D, L)$, and thus does not contain any step of the form $+\partial l$ for $l \in L$. Finally, since the rules used to check if the remaining steps are a subset of the rules in D , and again, if it does not involve any step with $+\partial l$, then the proof is also a proof for $+\partial_{\mathcal{O}m}$ in D .

For the other direction, a proof in D for $+\partial_{\mathcal{O}m}$ that does not contain any step of the form $+\partial l$ is trivially a proof in D . The rules missing in $red(D, L)$ are the rules in $R[l]$; these rules would be used to justify steps of the form $+\partial l$ (which are not present in P anyway); thus, the proof P is also a proof in $red(D, L)$. \square

EXAMPLE 5

Consider again a theory D containing the rules for the Iterated Unpragmatic Pragmatic Oddity scenario. Notice that the second derivation provided in that section is effectively a derivation in $red(D, \{-a\})$, and we can use it to show that we can derive $+\partial_{\mathcal{O}b}$ in $red(D, \{-a\})$. At the same time, the derivation is also a derivation in D , and thus it shows the independence of $\mathcal{O}a$ from $\neg a$. Hence, we can conclude $\mathcal{O}(a \wedge b)$.

Suppose we replace r_5 with its constitutive version, i.e. $r'_5: \mathcal{O}(A \wedge b) \Rightarrow \neg c$. Since we no longer have a prescriptive rule for $\mathcal{O}\neg c$, rule r_6 is unopposed and we can derive $+\partial_{\mathcal{O}c}$ from it. Also, $+\partial\neg c$ and $+\partial_{\mathcal{O}d}$ are derivable, and we can ask if the instance of the pragmatic oddity $\mathcal{O}(c \wedge d)$ is derivable. To this end, we compute the reduct $red(D, \{-c\})$. In this theory we do not have r'_5 , and it is easy to check that we can derive both $\mathcal{O}c$ and $\mathcal{O}d$. Accordingly, $\mathcal{O}d$ does not depend on $\neg c$, and we can conclude $+\partial_{\mathcal{O}c} \wedge d$.

Notice that we can combine the two independence results, and we are still able to conclude $+\partial_{\mathcal{O}d}$ when we remove the rules for $\neg a$ and $\neg c$. Finally, strictly speaking, to prove the two conjunctive obligations $\mathcal{O}(a \wedge b)$ and $\mathcal{O}(c \wedge d)$, we have to show that $\mathcal{O}a$ and $\mathcal{O}c$ are independent of, respectively, $\neg b$ and $\neg d$. However, since there are no constitutive rules for $\neg b$ and $\neg d$, the theory D itself is its reduct, i.e. $red(D, \{-b\}) = red(D, \{-d\}) = red(D, \{-b, -d\}) = D$.

We are now ready to introduce the notion of extension. Typically in Defeasible Logic, the extension of a theory is the set of all the literals that can be derived from the theory. However, in Defeasible Deontic Logic, the order of the elements in a derivation does not matter. On the contrary, as we have seen, the order matters if we want to capture the pragmatic oddity phenomenon properly, and different derivations are possible. In the definition below, we continue to speak of derivable/refutable literals/conjunctions. The fixed point construction will give the precise notion of derivation/refutation in Definition 21. Formally, the extension is a 6-tuple of sets, where every set contains the derivable/rejected literals/conjunctive obligations.

DEFINITION 16

Given a defeasible theory D the *extension* ($E(D)$) of the theory is the tuple:

$$E(D) = \langle \partial^+(D), \partial^-(D), \partial^{+\mathcal{O}}(D), \partial^{-\mathcal{O}}(D), \partial^{+\wedge}(D), \partial^{-\wedge}(D) \rangle,$$

where

- $\partial^+(D)$ is the set of literals appearing in D that are defeasible provable as institutional statements;

- $\partial^-(D)$ is the set of literals appearing in D that are defeasible refutable as institutional statements;
- $\partial^{+O}(D)$ is the set of literals appearing in D that are defeasible provable as obligations;
- $\partial^{-O}(D)$ is the set of literals appearing in D that are defeasible refutable as obligations;
- $\partial^{+\wedge}(D)$ is the set of conjunctive obligations defeasibly provable in D whose conjuncts are literals appearing in D ;
- $\partial^{-\wedge}(D)$ is the set of conjunctive obligations defeasibly refutable in D whose conjuncts are literals appearing in D .

Before we move to the procedure to construct the extension of a given defeasible theory, we need some auxiliary definitions (these definitions are the counterpart of Definitions 5–7 for extensions instead of derivations). First of all, we introduce some notation to identify the types of literal occurring in the body of rules.

DEFINITION 17

Given a rule r , we identify the following sets of literals (and conjunctions of literals).

- $r^\# = \{l \in \text{Lit} : l \in A(r)\}$;
- $r^O = \{l \in \text{Lit} : \mathbf{O}l \in A(r)\}$;
- $r^P = \{l \in \text{Lit} : \neg \mathbf{O}l \in A(r)\}$;
- $r^\wedge = \{c = l_1 \wedge \dots \wedge l_n : \mathbf{O}(c) \in A(r)\}$.

EXAMPLE 6

Consider the rule

$$r: a, \neg b, \mathbf{O}\neg c, \neg \mathbf{O}a, \mathbf{O}(c \wedge \neg d) \Rightarrow_{\mathbf{O}} e \otimes f. \quad (7)$$

Here, $r^\# = \{a, \neg b\}$, $r^O = \{\neg c\}$, $r^P = \{a\}$ and $r^\wedge = \{c \wedge \neg d\}$.

Given a conjunction $c_1 \wedge \dots \wedge c_m$, we use C to denote the set of complements of the literals in the conjunction, namely: $C = \{\sim c_1, \dots, \sim c_m\}$.

The next three definitions just mimic the definitions of when rules are applicable or rejected; instead of applying to steps of a derivation, they apply to elements of an extension. In the construction we are going to use to compute the extension of a theory, the reference is to the previous stage of the construction of the extension.

DEFINITION 18

A rule $r \in R[q, j]$ is *body-applicable* in an extension $E(D)$ iff

1. $r^\# \subseteq \partial^+(D)$ and
2. $r^O \subseteq \partial^{+O}(D)$ and
3. $r^P \subseteq \partial^{-O}(D)$ and
4. $r^\wedge \subseteq \partial^{+\wedge}(D)$.

A rule $r \in R[q, j]$ is *body-discarded* in an extension $E(D)$ iff

1. $r^\# \cap \partial^-(D) \neq \emptyset$ or
2. $r^O \cap \partial^{-O}(D) \neq \emptyset$ or
3. $r^P \cap \partial^{+O}(D) \neq \emptyset$ or
4. $r^\wedge \cap \partial^{-\wedge}(D) \neq \emptyset$.

DEFINITION 19

A rule $r \in R^O[q, j]$ such that $C(r) = c_1 \otimes \dots \otimes c_n$ is *applicable in an extension $E(D)$* for literal q at index j , with $1 \leq j < n$, in the condition for $\partial^{\pm O}$ iff

1. r is body-applicable in $E(D)$; and
2. for all $c_k \in C(r)$, $1 \leq k < j$, $c_k \in \partial^{+O}(D)$ and $\sim c_k \in \partial^+(D)$.

DEFINITION 20

A rule $r \in R[q, j]$ such that $C(r) = c_1 \otimes \dots \otimes c_n$ is *discarded in an extension $E(D)$* for literal q at index j , with $1 \leq j \leq n$ in the condition for $\partial^{\pm O}$ iff

1. r is body-discarded in $E(D)$; or
2. there exists $c_k \in C(r)$, $1 \leq k < l$, such that either $c_k \in \partial^{-O}(D)$ or $\sim c_k \in \partial^-(D)$.

According to Definition 5 a rule r is (body-)applicable if all the elements in the antecedent of the rule $A(r)$ have been proved in previous steps of the derivation. Similarly, r is (body-)discarded if there is an element in the antecedent that has been refuted. The idea behind the construction of the extension of a theory is to start from the set of facts and derive all conclusions (positive and negative) that can be obtained directly from the facts. Then, the procedure works as follows: at every iteration, we compute all the conclusions that follow directly from the elements calculated in the previous extension. A key aspect is determining what rules are applicable or discarded at a particular iteration.

EXAMPLE 7

When we consider again the rule r in (7), then r is applicable in an extension $E(D)$, if $\{a, \neg b\} \subseteq \partial^+(D)$, $\{\neg c\} \subseteq \partial^O(D)$, $\{a\} \subseteq \partial^P(D)$ and $\{c \wedge \neg d\} \subseteq \partial^\wedge(D)$. In addition, to check if it is applicable for f at index 2, we have to see if $e \in \partial^O(D)$ and $\neg e \in \partial^-(D)$. The rule is discarded if one of the given sets has a non-empty intersection with the corresponding negative sub-part of the extension, indicating, in this case, that one of the elements has been refuted.

We are now ready to give the definition providing the procedure to compute the extension of a defeasible theory.

DEFINITION 21 (Extension construction).

Given a defeasible theory $D = (F, R, >)$, the extension of D is built by the following construction:

$$\begin{aligned} E_{n+1}(D) &= (\partial_{n+1}^+(D), \partial_{n+1}^-(D), \partial_{n+1}^{+O}(D), \partial_{n+1}^{-O}(D), \partial_{n+1}^{+\wedge}(D), \partial_{n+1}^{-\wedge}(D)) \\ &= (\mathcal{T}(\partial_n^+(D)), \mathcal{T}(\partial_n^-(D)), \mathcal{T}(\partial_n^{+O}(D)), \\ &\quad \mathcal{T}(\partial_n^{-O}(D)), \mathcal{T}(\partial_n^{+\wedge}(D)), \mathcal{T}(\partial_n^{-\wedge}(D))), \end{aligned}$$

where

$$E_0(D) = (F, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset)$$

and

$$\begin{aligned}
\mathcal{T}(\partial_n^+(D)) &= \partial_n^+ \cup \{q: \sim q \notin F \text{ and} \\
&\quad \exists r \in R_d[q] r \text{ is applicable in } E_n(D) \text{ and} \\
&\quad \forall s \in R[\sim q] s \text{ is either discarded in } E_n(D) \text{ or} \\
&\quad \exists t \in R[q] t \text{ is applicable in } E_n(D) \text{ and } t > s\} \\
\mathcal{T}(\partial_n^-(D)) &= \partial_n^- \cup \{q: \sim q \in F \text{ or} \\
&\quad \forall r \in R_d[q] \text{ either } r \text{ is discarded in } E_n(D) \text{ or} \\
&\quad \exists s \in R[\sim q] s \text{ is applicable in } E_n(D) \text{ and} \\
&\quad \forall t \in R[q] \text{ either } t \text{ is discarded in } E_n(D) \text{ or } t \not> s\} \\
\mathcal{T}(\partial_n^{+O}(D)) &= \partial_n^{+O} \cup \{q: \exists r \in R^O[q, j] r \text{ is applicable in } E_n(D) \text{ and} \\
&\quad \forall s \in R^O[\sim q, k] s \text{ is either discarded in } E_n(D) \text{ or} \\
&\quad \exists t \in R^O[q, m] t \text{ is applicable in } E_n(D) \text{ and } t > s\} \\
\mathcal{T}(\partial_n^{-O}(D)) &= \partial_n^{-O} \cup \{q: \forall r \in R^O[q, j] \text{ either } r \text{ is discarded in } E_n(D) \text{ or} \\
&\quad \exists s \in R^O[\sim q, k] s \text{ is applicable in } E_n(D) \text{ and} \\
&\quad \forall t \in R^O[q, m] \text{ either } t \text{ is discarded in } E_n(D) \text{ or } t \not> s\} \\
\mathcal{T}(\partial_n^{+\wedge}(D)) &= \partial_n^{+\wedge} \cup \{c_1 \wedge \dots \wedge c_m: \forall c_i, c_i \in \partial_n^{+O} \text{ and} \\
&\quad c_i \in \partial_n^{+O}(\text{red}(D, C \setminus \{\sim c_i\}))\} \\
\mathcal{T}(\partial_n^{-\wedge}(D)) &= \partial_n^{-\wedge} \cup \{c_1 \wedge \dots \wedge c_m: \exists c_i, c_i \in \partial_n^{-O}(D) \text{ or} \\
&\quad c_i \notin \partial_n^{+O}(\text{red}(D, C \setminus \{\sim c_i\}))\}.
\end{aligned}$$

In the construction above, the first four sets replicate the proof conditions for the corresponding proof tags where we proceed in terms of stages instead of derivation steps. For conjunctive obligations, we first determine if the individual obligations are derivable at the current stage. At the same time, for each individual obligation, we check if it is in the extension of the reduct of the theory obtained by removing the literals corresponding to the violations of the other obligations in the conjunctive obligation. If it is, then the individual obligation is independent of the violation of the other obligations. Notice that for this last step, we are not looking if they are in the extension in a particular stage but in the extension at the end of the construction for the extension of the reduct. Also, further reducts (for other conjunctions) may be computed in the computation for a reduct. However, given that a reduct is a subset of a given theory, the process is guaranteed to terminate (provided that the initial theory has finitely many rules).

The set of extensions forms a complete lattice under the pointwise containment ordering, with E_0 as its least element. The least upper bound operation is the pointwise union. It is easy to see that \mathcal{T} is monotonic, and the Kleene sequence from E_0 is increasing. Thus the limit

$$\mathcal{L} = \langle \partial_L^+(D), \partial_L^-(D), \partial_L^{+O}(D), \partial_L^{-O}(D), \partial_L^{+\wedge}(D), \partial_L^{-\wedge}(D) \rangle,$$

of all finite elements in the sequence exists, and it has a least fixpoint

$$\mathcal{F} = \langle \partial_F^+(D), \partial_F^-(D), \partial_F^{+O}(D), \partial_F^{-O}(D), \partial_F^{+\wedge}(D), \partial_F^{-\wedge}(D) \rangle.$$

When D is a finite propositional defeasible deontic theory $\mathcal{F} = \mathcal{L}$. Accordingly, we take \mathcal{F} as the extension of D , $E(D) = \mathcal{F}$. Furthermore, \mathcal{F} being the least upper bound is unique and captures the

conditions that determine whether a conjunctive obligation is independent of the violations of its conjuncts.

We can revisit some of the scenarios presented in Section 4 using the bottom-up construction.

EXAMPLE 8

Let us consider again the theory $D = (F, R, \emptyset)$ we used to illustrate the Iterated Pragmatic Non Pragmatic Oddity scenario, where $F = \{f_1, f_2, f_3, f_4, f_6, f_7\}$, and R consists of the following rules:

$$\begin{aligned} r_1: f_1 \Rightarrow_{\circ} a \otimes b & & r_2: f_2 \Rightarrow_{\circ} a & & r_3: f_3 \Rightarrow_{\circ} b & & r_4: f_4 \Rightarrow \neg a \\ r_5: \mathbf{O}(a \wedge b) \Rightarrow_{\circ} \neg c & & r_6: f_6 \Rightarrow_{\circ} c \otimes d & & r_7: f_7 \Rightarrow_{\circ} d. \end{aligned}$$

According to Definition 16

$$E_0(D) = (\{f_1, f_2, f_3, f_4, f_6, f_7\}, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset).$$

We can now compute $E_1(D)$. All rules but r_5 are applicable since their antecedent is a subset of $\partial_0^+(D)$. Moreover, for r_1, r_2, r_3, r_4 and r_7 , there are no rules for the complement of their conclusion, thus, vacuously, the condition that all rules for the opposite are either defeated or discarded is satisfied. Hence we add $\neg a$ to $\partial_1^+(D)$, and $\partial^{+O} = \{a, b, d\}$. Given that there are no constitutive rules for $a, b, \neg b, c, \neg c, d, \neg d$, these literals are all in $\partial_1^-(D)$. For the same reason, $\partial^{-O} = \{\neg a, \neg b, \neg d\}$. Notice that, even if we have an applicable prescriptive rule for c (r_6), there is a prescriptive rule for $\neg c$ (r_5), but we are not able to assess, yet, whether it is applicable or discarded. We are not in the position to populate $\partial_1^{+\wedge}$ since ∂_0^{+O} . For ∂_1^{-O} , we can compute the reduct for all individual literals and determine what literals are not in $\partial^{+O}(\text{red}(D, \{l\}))$. In the theory, the only constitutive rule is r_4 , and we can repeat the argument for $\partial_1^{-O}(D)$. Accordingly, $\partial_1^{-\wedge}(D)$ contains all conjunctions where at least one element belongs to $\partial^{-O}(D)$, e.g. $\neg a \wedge b, a \wedge \neg b \wedge c$ and so on.

We proceed to the computation of E_2 , specifically $\partial_2^{+\wedge}(D)$. We have $a, b, d \in \partial_1^{+O}(D)$. Thus, we have to consider what combinations result in conjunctive obligations that are not pragmatic oddity instances. To this end, we compute $\text{red}(D, \{\neg a\})$, $\text{red}(D, \{\neg b\})$ and $\text{red}(D, \{\neg c\})$. Since there are no constitutive rules for $\neg b$ and $\neg c$, $\text{red}(D, \{\neg b\}) = \text{red}(D, \{\neg c\}) = D$, and we have seen that $a, b, d \in \partial^{+O}(D)$. Removing $\neg a$ results in $\neg a \in \partial^-(\text{red}(D, \{\neg a\}))$ making r_1 not applicable for b at index 2. However, we still have rule r_3 to include b in $\partial^{+O}(\text{red}(D, \{\neg a\}))$. Hence, $a \wedge b \in \partial_2^{+\wedge}(D)$, and so are $a \wedge d, b \wedge d$ and $a \wedge b \wedge d$.

For $E_3(D)$, we have two applicable prescriptive rules for complementary literals: r_5 for $\neg c$ and r_6 for c . However, we do not have instances of the superiority relation for them. Thus, c and $\neg c$ are not provable as obligations, and we include them in $\partial_3^{-O}(D)$. This, in turn, allows us to establish that $c \wedge d \in \partial_4^{-\wedge}(D)$. After this step, we no longer add elements to the extension, meaning that we have reached the fixed point.

EXAMPLE 9

Let us turn our attention to the theory D for the multiple dependencies scenarios, where the rules are

$$\begin{aligned} r_1: \Rightarrow_{\circ} a & & r_2: \Rightarrow_{\circ} b \\ r_3: \neg a \Rightarrow_{\circ} c & & r_4: \neg b \Rightarrow_{\circ} c \end{aligned}$$

where $F = \{\neg a, \neg b\}$. It is easy to verify that $a, b, c \in \partial_1^{+O}$. Let us consider the reducts for $\{\neg a\}$, $\{\neg b\}$ and $\{\neg a, \neg b\}$. For the first $F = \{\neg b\}$; therefore, $\neg a \in \partial^-(\text{red}(D, \{\neg a\}))$ and r_3 is discarded.

However, we can still use r_4 to conclude Ob ($b \in \partial^{+O}(red(D, \{-a\}))$). Accordingly, $a \wedge c \in \partial_2^{+\wedge}(D)$. We can repeat a similar argument for $red(D, \{-b\})$ to determine that $b \wedge c \in \partial_2^{+\wedge}(D)$. Finally, for $a \wedge b \wedge c$, we notice that when we remove both $\neg a$ and $\neg b$ from the set of facts in the computation of $red(D, \{-a, \neg b\})$, rules r_3 and r_4 are both discarded, and there are no remaining prescriptive rules for c ; ergo, $c \in \partial^{-O}(red(D, \{-a, \neg b\}))$, which implies $a \wedge b \wedge c \in \partial^{-\wedge}(D)$.

DEFINITION 22

An extension

$$\langle \partial^+(D), \partial^-(D), \partial^{+O}(D), \partial^{-O}(D), \partial^{+\wedge}(D), \partial^{-\wedge}(D) \rangle$$

is *coherent* if $\partial^+ \cap \partial^- = \emptyset$, $\partial^{+O} \cap \partial^{-O} = \emptyset$ and $\partial^{+\wedge} \cap \partial^{-\wedge} = \emptyset$.

An extension is *consistent* if for every set ∂^* , it is not the case that p and $\sim p$ are both in ∂^* .

Intuitively, coherence says that no literal is simultaneously provable and unprovable. Consistency says that a literal and its negation are not both defeasibly provable.

PROPOSITION 3

Given a theory D , $E(D)$ is coherent. If F does not contain a pair of complementary literals, and the transitive closure of $>$ is acyclic, then $E(D)$ is consistent.

PROOF. Notice that the conditions to establish that a literal/conjunction is a member of one of the positive sets of an extension at a given stage are de facto the strong negation of the condition to add the literal to the corresponding negative set. We have to replace $a(r) \subseteq \partial^+$ for $\forall a \in A(r)$, $+\partial a \in P(1..n)$, and $a(r) \cap \partial^-$ for $\exists a \in A(r)$, $-\partial a \in P(1..n)$. Hence, we can use the results of [12], see also Proposition 1. Here we show the key cases for coherence. For the cases of consistency, see the proof in [12].

We prove the proposition for coherence by induction on the extension's construction stage. The inductive base, the case for $E_0(D)$, is trivial by the definition of E_0 .

For the inductive base, let us assume that coherence holds up to the n -th extension, $E_n(D)$. By the monotonicity of the construction, if a rule is applicable at a step $m < n$, then the rule remains applicable at step n (similarly for discarded). For ∂^+ and ∂^- , the argument is as follows: for a literal l to be in ∂^{+n+1} , there must be a rule r that is applicable at $E_n(D)$: by the inductive hypothesis, and Definitions 18, 19 and 20 no rule is at the same time applicable and discarded for one and the same literal at the same time. This means that, for the condition for ∂^- , there is a rule s that is applicable in $E_n(D)$, but then there is a rule t applicable for l at $E_n(D)$ and $t > c$, but for ∂^- t should either be discarded or not stronger than s . Contradiction. Thus ∂_{n+1}^+ and ∂_{n+1}^- are disjoint.

For $\partial^{+O} \cap \partial^{-O} = \emptyset$, we remark in addition to what we have just proved, we have to consider conditions 2 of Definitions 19 and 20 to realize by the inductive hypothesis that no rule can satisfy the conditions in the two definitions.

Finally, for $\partial^{+\wedge} \cap \partial^{-\wedge} = \emptyset$, by the inductive hypothesis $\partial_n^{+O} \cap \partial_n^{-O} = \emptyset$, in addition, the extension of any theory is unique (being the least upper bound of a finite monotonically increasing construction), and the reducts we consider are subsets of the given theory (thus, the coherence property holds for them as well). \square

An inconsistency is possible only when the theory we started with was inconsistent (either because the facts are inconsistent or because the superiority relation induces a cycle in the superiority relations, meaning that a rule is at the same time stronger and weaker than another rule). Accordingly, defeasible inference for defeasible deontic logic for pragmatic oddity does not

introduce inconsistency. A logic is coherent (consistent) if the meaning of each theory of the logic, when expressed as an extension, is coherent (consistent).

6 Complexity

In this section, we study the computational complexity of the problem of computing whether a conjunctive obligation is derivable from a given defeasible theory. To this end, we adapt the algorithm proposed in [11] to compute the extension of a defeasible theory, where the computation of the extension is linear in the size of the theory. The algorithm is based on a series of transformations that reduce the complexity of the theory by either removing elements from rules when some elements are provable, or removing rules when they become discarded (and so no longer able to produce positive conclusions).

The paper aims to determine when conjunctive obligations are either provable or discarded. Accordingly, we have to extend the definition to account for conjunctive obligations. However, if we want to maintain a feasible computational complexity, we have to limit the conjunctions we consider: given a set of n literals, the set of all possible non-logically equivalent conjunctions that the n literals can form contains 2^n conjunctions; hence, we cannot compute in polynomial time for such a set if any element is derivable or refuted by the theory. However, we are going to show that for each individual conjunction, we can compute in polynomial-time whether it is derivable or refuted.

DEFINITION 23

Given a defeasible theory D , the *conjunctive extension* of the theory is the tuple:

$$\langle \partial^+(D), \partial^-(D), \partial_{\mathbf{O}}^+(D), \partial_{\mathbf{O}}^-(D), \partial_{\wedge}^+(D), \partial_{\wedge}^-(D) \rangle,$$

where $\partial^+(D)$, $\partial^-(D)$, $\partial_{\mathbf{O}}^+(D)$ and $\partial_{\mathbf{O}}^-(D)$ are as in Definition 16 and

- $\partial_{\wedge}^+(D)$ is the set of conjunctive obligations appearing in D (i.e. $c = \mathbf{O}(c_1 \wedge \dots \wedge c_n)$ and $\exists r \in R$ such that $c \in A(r)$) that are defeasibly provable in D ;
- $\partial_{\wedge}^-(D)$ is the set of conjunctive obligations appearing in D that are defeasibly refutable in D .

The algorithm to determine the conjunctive extension of a theory is based on the following data structure (for the full details, we refer the reader to [11]). We create a list of the atoms appearing in the theory. Every entry in the list of atoms has an array associated to it. The array has ten cells, where every cell contains pointers to rules depending on whether and how the atom appears in the rule. The first cell is where the atom appears in the head of a constitutive rule, the second where the negation of the atom appears in the head of a constitutive rule, the third where the atom appears in the head of a prescriptive rule, the fourth where the negation of atom appears in the head of a prescriptive rule, the fifth where the atom appears in the body of a rule, the sixth where the negation of the atom appears in the body of a rule, the seventh where the atom appears as an obligation in the body of a rule, the eighth where the negation of the atom appears as an obligation in the body of a rule, the ninth where the atom appears as a negative obligation in the body of a rule and the tenth where the negation of the atom appears as a negative obligation in the body of a rule. In addition, we maintain a list of conjunctive obligations occurring in the theory, and for every conjunction, we associate it to the rules where it appears in the body.

The algorithm works as follows: at every round, we scan the list of atoms. For every atom (excluding the entries for the conjunctions), we look if the atom appears in the head of some rules. If it does not appear in any of the cells for the heads, we can set the corresponding literals as refuted;

and we can remove rules from corresponding cells. So, for example, given an atom p , if there are no prescriptive rules for $\neg p$; then we can conclude that the theory proves $-\partial_{\mathbf{O}}\neg p$; accordingly, all rules where $\neg\mathbf{O}\neg p$ occurs in the body are (body)-discarded, and we can remove them from the data structure. Similarly, if there are no constitutive rules for $\neg p$, then we can prove $-\partial\neg p$, and then (i) all the rules where it appears in the body are body-discarded, but also (ii) for each rule r in whose head p appears as an obligation, no elements following p in r can any longer be derived using r , and such elements are removed from the appropriate cells. If an atom appears in the head of a rule, we determine (i) if the body of the rule is empty and (ii) for prescriptive rules, if the atom is the first element of the head. If this is the case, then the rule is applicable, and we check if there are rules for the negation. If there are no rules for the negation, or the rules are weaker than applicable rules, then the atom/literal is provable with the suitable proof tag. Then we remove the atom/literal from the appropriate rules. We repeat the above steps until we can no longer obtain new conclusions. When we are not able to derive new conclusions, we turn our attention to the list of the conjunctive obligations, where we invoke the following (sub)algorithm for every conjunction $c = (c_1 \wedge \dots \wedge c_n)$ in the list (where $C = \{\sim c_i, 1 \leq i \leq n\}$)

Algorithm 1 Evaluate Conjunctive Obligation $c = c_1 \wedge \dots \wedge c_n$

```

1: for  $i \in 1..n$  do
2:   if  $c_i \in \partial_{\mathbf{O}}^-(D)$  then
3:      $c \in \partial_{\wedge}^-(D)$  remove all rules  $r$  where  $c \in A(R)$ 
4:   Exit
5:   end if
6:   if  $c_i \in \partial_{\mathbf{O}}^+(D)$  then
7:     if  $\forall c_j \neg c_i, \sim c_j \in \partial^+(D)$  then
8:       if  $c_i \in +\partial_{\mathbf{O}}^+(\text{red}(D, C \setminus \{\sim c_i\}))$  then
9:          $i := i + 1$ 
10:      else  $c \in \partial_{\wedge}^-(D)$  remove all rules  $r$  where  $c \in A(R)$ 
11:      Exit
12:      end if
13:      if  $\exists c_j \neq c_i, \sim c_i \in \partial^-(D)$  then
14:         $i := i + 1$ 
15:      end if
16:      end if
17:    end if
18:    Exit
19:  end for
20:  $c \in \partial_{\wedge}^+(D)$ , remove  $c$  from all rules  $r$  where  $c \in A(r)$ 

```

For every conjunction, the algorithm iterates over the conjuncts. The conjunction is not provable if a conjunct is not provable as an obligation (lines 2–4). If the conjunct is provable as an obligation, it checks whether the violations of the other obligations are provable; if so, it has to check whether the obligation of the conjunct is independent of the violations. To determine this, we can repeat the whole algorithm with the sub-theory obtained by the transformation $\text{red}(D, C \setminus \{c_i\})$. If it is independent, we continue with the next element of the conjunction; otherwise, the conjunction is not derivable. Similarly, if some of the violations are not derivable, we continue with the iteration. The conjunction is provable when the iteration is successful for all the conjunction elements.

At the end of the sub-routine, we return to the main algorithm; if there are changes in the rules, we repeat the process; otherwise, the process terminates.

PROPOSITION 4

The algorithm to compute the conjunctive extension of a theory D computes the extension $E(D)$ when the language is restricted to the conjunctive obligations that occur in D .

PROOF. The algorithm consists of two parts. The first part is the algorithm presented in [11] to compute the extension of a Defeasible Deontic Logic. The proof conditions presented in this paper are restrictions of those in [11], and they are equivalent as far as the language in this paper is concerned. The language (and algorithm) in [11] does not allow for conjunctive obligations. Thus, we can consider each conjunctive obligation with a new literal. [11] proves that their algorithm is sound and complete to compute the extension (corresponding to $\langle \partial^+(D), \partial^-(D), \partial_0^+(D), \partial_0^-(D) \rangle$). The second part of the computation presented in this paper is Algorithm 1, that effectively acts as an external oracle to determine whether the conjunctive obligations (the new literal) hold or not (based on the reduct construction). If a conjunctive obligation holds, then it can be added to $\partial_\wedge^+(D)$, otherwise to $\partial_\wedge^-(D)$, and we can resolve the corresponding new literal. Thus, the correctness depends on the correctness of Algorithm 1 against the construction in Definition 21. The explanation of the algorithm above shows that the steps in the algorithm correspond to the steps to compute $\mathcal{T}(\partial_n^+(D))$ and $\mathcal{T}(\partial_n^-(D))$. \square

Concerning the computational complexity, [11] proves that the complexity of computing the extension of a defeasible theory without conjunctive obligations is linear in the size of the theory, where the size of the theory is determined by the number of symbols in the theory, and hence if n and r stand for, respectively, the number of atoms and the number of rules in the theory, the complexity is in $O(n * r)$. For the complexity of computing the conjunctive extension of a defeasible theory, we have to take into account the complexity of the Evaluate Conjunctive Obligation algorithm and the number of times we have to compute it. This can be determined as follows: let m be the number of conjunctive obligations in the theory, and k the number of conjuncts in the longest conjunctive obligation. For each of them, we have to compute the extension of $red(D, C)$, thus we have to perform $O(m * k * O(n * r))$ computations on top of the calculation of the extension (i.e. $O((m + n) * r)$).

PROPOSITION 5

The conjunctive extension of a theory can be computed in polynomial time.

Notice that the algorithm Evaluate Conjunctive Obligation can be used to evaluate any conjunctive obligation, not only the conjunctive obligations occurring in a theory. All we have to do is to compute the conjunctive extension of the theory and then evaluate the single conjunctive obligation, and as we have just seen, this can be calculated in polynomial time.

7 Summary and discussion

We have proposed an extension of Defeasible Deontic Logic that prevents the so-called Pragmatic Oddity paradox from occurring. The mechanism we used to achieve this result was to provide a schema that allows us to give a guard to the derivation of conjunctive obligations ensuring that each individual obligation does not depend on the violation of the other obligation. The proof theory of defeasible logic gives the mechanism; in addition, we presented a bottom-up characterization of the logic that avoids the problem of non-deterministically selected derivations. Furthermore, the bottom-up construction is the foundation of the algorithm presented in [11] to compute the extension of a

defeasible deontic theory (without conjunctive obligations) in linear time. This allows us to give a polynomial upper bound to the problem of computing the extension of a defeasible theory with pragmatic oddity (limiting to the conjunctive obligations appearing explicitly in the theory). First, we treat the conjunctive obligations in a theory as new literals, and then for each of them, we spin out the computation of the extensions for the relevant reducts. While the upper bound complexity of the logic is polynomial and hence feasible, the algorithm we just outlined is not optimal. Most practical real-life examples are likely to involve only a few conjunctive obligations, and ones with few conjuncts, so modest inefficiency of the algorithm for implementation is often not a serious practical problem. Nonetheless, it is desirable, as a next step, to devise an optimal algorithm to implement these novel proof conditions and the bottom-up procedure.

Acknowledgements

Preliminary versions of the paper proposing the idea of the logic were presented at Jurix 2019 [15] and DEON 2020/2021 [10]. We thank the anonymous reviewers for their valuable comments on earlier versions of the paper.

References

- [1] G. Antoniou, D. Billington, G. Governatori and M. J. Maher. Representation results for defeasible logic. *ACM Transactions on Computational Logic*, **2**, 255–287, 2001.
- [2] J. Carmo and A. J. I. Jones. Deontic logic and contrary-to-duties. In *Handbook of Philosophical Logic*, pp. 265–343. Springer, 2002.
- [3] R. M. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, **24**, 33–36, 1963.
- [4] D. Gabbay, J. Horty, X. Parent, R. van der Meyden and L. van der Torre, eds. *Handbook of Deontic Logic and Normative Systems*, vol. 1. College Publications, 2013.
- [5] L. Goble. Multiplex semantics for deontic logic. *Nordic Journal of Philosophical Logic*, **5**, 113–134, 2001.
- [6] L. Goble. A logic for deontic dilemmas. *Journal of Applied Logic*, **3**, 461–483, 2005.
- [7] L. Goble. Prima facie norms, normative conflicts and dilemmas. In *Handbook of Deontic Logic and Normative Systems*, D. Gabbay, J. Horty, X. Parent, R. van der Meyden and L. van der Torre, eds, vol. 1. College Publications, 2013.
- [8] G. Governatori. Thou shalt is not you will. In *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Law*, K. Atkinson., ed., pp. 63–68. ACM, New York, 2015.
- [9] G. Governatori. Practical normative reasoning with defeasible deontic logic. In *Reasoning Web 2018*, C. d’Amato and M. Theobald, eds, LNCS, vol. 11078, pp. 1–25. Springer International Publishing, Cham, 2018.
- [10] G. Governatori, S. Colombo Tosatto and A. Rotolo. A defeasible deontic logic for pragmatic oddity. In *Deontic Logic and Normative Systems: 15th International Conference (DEON2020/2021)*, F. Liu, A. Marra, P. Portner and F. Van De Putte, eds, pp. 239–254. College Publications, London, 2021.
- [11] G. Governatori, F. Olivieri, A. Rotolo and S. Scannapieco. Computing strong and weak permissions in defeasible logic. *Journal of Philosophical Logic*, **42**, 799–829, 2013.
- [12] G. Governatori, V. Padmanabhan, A. Rotolo and A. Sattar. A defeasible logic for modelling policy-based intentions and motivational attitudes. *Logic Journal of the IGPL*, **17**, 227–265, 2009.

- [13] G. Governatori and A. Rotolo. Logic of violations: a Gentzen system for reasoning with contrary-to-duty obligations. *Australasian Journal of Logic*, **4**, 193–215, 2006.
- [14] G. Governatori and A. Rotolo. Logics for legal dynamics. In *Logic in the Theory and Practice of Lawmaking, number 2 in Legisprudence Library*, A. Michał and K. Pleszka, eds, vol. 2, pp. 323–356. Springer, 2015.
- [15] G. Governatori and A. Rotolo. A computational model for pragmatic oddity. In *JURIX 2019: The 32nd International Conference on Legal Knowledge and Information Systems*, Araszkievicz M. and V. Rodríguez-Doncel, eds, vol. 332 of *Frontiers in Artificial Intelligence and Applications*, pp. 187–192, Amsterdam, 2019. IOS Press.
- [16] G. Governatori, A. Rotolo and G. Sartor. Logic and the law: philosophical foundations, deontics, and defeasible reasoning. In *Handbook of Deontic Logic and Normative Reasoning*, D. M. Gabbay, J. Horty, X. Parent, R. van der Meyden and L. van der Torre, eds, vol. 2, chapter 9, pp. 655–760. College Publications, London, 2021.
- [17] J. F. Horty. Deontic logic as founded on nonmonotonic logic. *Annals of Mathematics and Artificial Intelligence*, **9**, 69–91, 1993.
- [18] M. J. Maher and G. Governatori. A semantic decomposition of defeasible logic. In *Proc. American National Conference on Artificial Intelligence (AAAI-99)*, J. Hendler and D. Subramanian, eds, pp. 299–305. AAAI Press, Menlo Park, CA, 1999.
- [19] D. Makinson. On a fundamental problem of deontic logic. In *Norms, Logics and Information Systems. New Studies in Deontic Logic and Computer Science*, P. McNamara and H. Prakken, eds, pp. 29–54. IOS Press, Amsterdam, 1999.
- [20] D. Makinson and L. van der Torre. Input–output logics. *Journal of Philosophical Logic*, **29**, 383–408, 2000.
- [21] D. Makinson and L. van der Torre. Constraints for input/output logics. *Journal of Philosophical Logic*, **30**, 155–185, 2001.
- [22] X. Parent and L. van der Torre. Sing and dance! In *Deontic Logic and Normative Systems*, F. Cariani, D. Grossi, J. Meheus and X. Parent, eds, pp. 149–165. Springer International Publishing, Cham, 2014.
- [23] X. Parent and L. van der Torre. The pragmatic oddity in norm-based deontic logics. In *Proceedings of the 16th edition of the International Conference on Artificial Intelligence and Law*, pp. 169–178. ACM, 2017.
- [24] H. Prakken and M. J. Sergot. Contrary-to-duty obligations. *Studia Logica*, **57**, 91–115, 1996.
- [25] L. W. N. van der Torre and Y.-H. Tan. *The Many Faces of Defeasibility in Defeasible Deontic Logic*, pp. 79–121. Springer, Netherlands, Dordrecht, 1997.

Received 14 December 2021