Contextual Deontic Logic

Violation Contexts and Factual Defeasibility

Leendert W.N. van der Torre
IRIT, Paul Sabatier University
118 Route de Narbonne
31062 Toulouse
France

Yao-Hua Tan
EURIDIS, *Erasmus University Rotterdam*
P.O. Box 1738
3000 DR Rotterdam
The Netherlands

(Received ..... ; Accepted in final form .....)

Abstract. In this article we introduce Contextual Deontic Logic (CDL) to analyze the relation between deontic, contextual and defeasible reasoning. The optimal state, and therefore the set of active obligations, can change radically when the violation context changes. In such cases we say that the obligations only in force in the previous violation context are defeated; contextual deontic logic is therefore a defeasible deontic logic. This is expressed by the definition $O_\gamma(\alpha|\beta) =_{\text{def}} O(\alpha|\beta\backslash\neg\gamma)$: '$\alpha$ ought to be (done) if β is (done) in the context where γ is (done)' is defined as 'α ought to be (done) if β is (done) unless $\neg\gamma$ is (done).' The unless clause formalizes explicit exceptions and is analogous to the justification in Reiter's default rules. CDL is a *monotonic* defeasible deontic logic, because it has factual defeasibility but not overridden defeasibility.

1. Introduction

It is well-known in deontic logic that there are striking similarities between deontic reasoning and contextual reasoning. For example, in temporal deontic reasoning Thomason (1981) makes a distinction between the context of deliberation and the context of justification to distinguish a (deliberative) ought that implies 'practical-temporal can' from a (judgmental) ought that does not imply 'practical-temporal can.' The truth values of deontic sentences are not only time-dependent in the same, familiar way that the truth values of all tensed sentences are time-dependent, but they are also dependent of a set of choices or future options that varies as a function of time. The context of deliberation is the set of choices when you are looking for practical advice, whereas the context of justification is the set of choices for someone who is judging you. The following example illustrates the distinction between the two contexts.

Example 1. Yesterday (t_1) I had both the judgmental and deliberative obligation to go tonight (t_2) to my fathers' birthday in Rotterdam, so at t_1 it holds that $O_j f \wedge O_d f$. However, yesterday I took the airplane to Brazil, and today I therefore no longer have the deliberative obligation to go to my father's birthday. At t_2 it holds that $O_j f \wedge \neg O_d f \wedge \neg f$. Consequently, now I cannot be in time any more for my father's birthday and I have violated my judgmental ought, but not my deliberative ought.

Deontic logic is hampered by the so-called deontic paradoxes. The contrary-to-duty paradoxes like the notorious Chisholm paradox are the classic benchmark problems of deontic logics, which have initiated developments of monadic deontic logics (Chisholm, 1963; Forrester, 1984), dyadic deontic logics (Tomberlin, 1981) and temporal deontic logics (van Eck, 1982; van der Torre & Tan, 1998a). Contexts can be distinguished in several deontic paradoxes, for example in Castañeda's paradox of the second best plan (Castañeda, 1981; Feldmann, 1990; Yu, 1995).

Example 2 (Second best plan paradox). Suppose dr. Denton is going to give treatment to one of his patients. His best course of action would be to give an aspirin today and an aspirin tomorrow. His second best plan would be to give a Buffering today and another Buffering tomorrow. It would be very bad for him to give mixed medications. If dr. Denton does not give an aspirin today, then there are two contradictory obligations. *'The doctor ought to give an aspirin today and an aspirin tomorrow'* refers to the best plan context, and *'he ought not to give an aspirin tomorrow'* refers to the second best plan context.

In *a-temporal* deontic reasoning Prakken and Sergot (1996) distinguish between ideal and varying sub-ideal contexts to formalize contrary-to-duty reasoning. Obligations that refer to the ideal context are called ideal (or primary) obligations, and obligations that refer to the sub-ideal context are called contrary-to-duty (or secondary) obligations. Obviously there are relations between the distinction between justification and deliberation on the one hand, and the distinction between ideal and sub-ideal on the other hand. The essential contrast in both distinctions is a difference in what we take as being settled (Hansson, 1971; Thomason, 1981; Loewer & Belzer, 1983). The following example illustrates the distinction between the ideal and sub-ideal context by Prakken and Sergot's cottage housing regulations, in propositional modal logic alphabetic variants of the notorious Forrester and Chisholm paradoxes (Forrester, 1984; Chisholm, 1963).

Example 3 (Cottage housing regulations). First, consider the two obligations *'there ought to be no fence'* and *'if there is a fence, then it ought to be white.'* If there is a fence around the cottage, then there is an obligation *'there ought to be no fence'* referring to the ideal context there is no fence, and there is an obligation *'there ought to be a white fence'* referring to the sub-ideal context there is a fence. Removing the fence as well as painting it white are improvements of the present state of affairs, and therefore obligatory. Whether the fence ought to be removed or painted depends on the context referred to. The two derived obligations *'there ought to be no fence'* and *'there ought to be a white fence'* are contradictory. In a logic that does not distinguish between contexts, like Standard Deontic Logic, the two sentences of the regulations are inconsistent with the fact *'there is a fence.'*

Moreover, consider the three obligations *'there ought not to be a dog'*, *'if there is no dog, then there ought to be no sign'* and *'if there is a dog, then there ought to be a sign'*. If there is a dog, then the obligation *'there ought to be no sign'* is in force in the ideal context there is no dog nor a sign,[1] but it is not in force in the sub-ideal context there is a dog. Moreover, the two derived obligations *'there ought to be no sign'* (ideal context) and *'there ought to be a sign'* (sub-ideal context) are contradictory. Again, the three sentences of the regulations are inconsistent with the fact *'there is a dog'* in, for example, Standard Deontic Logic.

In this paper we introduce Contextual Deontic Logic (CDL). It does *not* formalize prima facie obligations, i.e. obligations *cannot* be overridden by other, stronger obligations, but it consistently formalizes contrary-to-duty reasoning as it occurs in the cottage housing regulations. The optimal state, and therefore the obligations, can change radically when the violation context changes. In such cases, we say that obligations only in force in the previous violation context are defeated; contextual deontic logic is therefore a defeasible deontic logic. We use two types of contextual obligations, in which we represent the context by a propositional sentence (as in (Prakken & Sergot, 1996)). We write $O_\gamma(\alpha|\beta)$ for 'α ought to be (done) if β is (done) in the context where γ is (done)' and we write $O(\alpha|\beta\backslash\neg\gamma)$ for 'α ought to be (done) if β is (done) unless $\neg\gamma$ is (done).' The unless clause can be compared to the justification in a Reiter default 'α is normally the case if β is the case unless $\neg\gamma$ is the case,' written as $\beta : \gamma/\alpha$ (Reiter, 1980). For example, *'birds fly unless they are penguins'* can be represented by $b : \neg p/f$, and *'penguins do not fly'* by $(b \wedge p) : \top/\neg f$. Hence, the unless clause formalizes a consistency check. The first type of contextual obligations are defined in terms of the second type of contextual obligations as follows.

$$O_\gamma(\alpha|\beta) =_{\text{def}} O(\alpha|\beta\backslash\neg\gamma)$$

Consequently, we disagree with Prakken and Sergot that "contrary-to-duty is not an instance of defeasible reasoning, and that methods of non-monotonic reasoning are inadequate since they are unable to distinguish between defeasibility and violation of primary obligations" (Prakken & Sergot, 1996).

- First, it follows directly from the definition above that contextual reasoning is defeasible, and given that "a contrary-to-duty obligation pertains to, or presupposes, a certain context in which a primary obligation is already violated," contrary-to-duty reasoning is defeasible too! Obligations lack unrestricted strengthening of the antecedent, the typical property of defeasible conditionals (Alchourrón, 1993), to block the combination of obligations from different contexts. Combining obligations of different contexts is counterintuitive, whether these obligations conflict (as in Example 3) or not (called pragmatic oddities in (Prakken & Sergot, 1996)[2]).

- Second, the diagnostic problem of *defeasible deontic logic* – i.e. to determine whether 'α ought to be (done) and $\neg\alpha$ is (done)' is a violation or an exception – can be solved by methods of non-monotonic reasoning. In the proof theory 'defeasibility' (in Prakken and Sergot's sense) and violation of primary obligations can be distinguished by distinguishing between different types of defeasibility (see below), and in the semantics they can be distinguished by multi-preference semantics for the strong overridden defeasibility of defeasible obligations 'normally, α ought to be (done)' and by strengths of obligations (priorities) for the weak overridden defeasibility of prima facie obligations (van der Torre & Tan, 1995)(1997a; 1998c).

We use CDL to analyze the relation between deontic, contextual and defeasible reasoning. For this analysis, we distinguish between three different types of defeasibility in defeasible reasoning: factual defeasibility (FD), conflict defeasibility (CD) and overridden defeasibility (OD). FD is the only one that does not depend on conflicts between formulas (when written as preferences, as explained later in this paper). The distinction between CD and OD is that in the last case the conflict is resolved, for example by the specificity principle or by priorities. All three types of defeasibility lead to restrictions on strengthening of the antecedent, but they can be discriminated with respect to non-monotonicity in (extensions of) dyadic logic without factual detachment. In such a logic OD leads to non-monotonicity whereas FD does not. Moreover, whether CD gives rise to non-monotonicity depends on the philosophical stance towards dilemmas.

In our view the relation between deontic, contextual and defeasible reasoning is as follows. Deontic contrary-to-duty reasoning is a kind of contextual reasoning, and is therefore defeasible. The pragmatic oddities show that non-monotonic 'consistency restoring' techniques are insufficient to formalize contrary-to-duty reasoning (van der Torre & Tan, 1995)(1997a), although the inconsistency of the two sets of obligations of the cottage housing regulations in Example 3 in SDL might suggest otherwise (Ryu & Lee, 1993; McCarty, 1994; Yu, 1995). Consequently, CDL should not be based on OD. It has to be based on factual defeasibility and it is therefore monotonic.[3]

This paper is organized as follows. In Section 2 we introduce contextual obligations $O_\gamma(\alpha|\beta)$ and $O(\alpha|\beta\backslash\gamma)$ in a preference-based logic and in Section 3 we axiomatize contextual deontic logic in a *modal* preference logic. In Example 4 we analyze contrary-to-duty reasoning in CDL and in Section 5 we discuss the relation between obligations and different types of defeasibility.

2. Preference-based contextual deontic reasoning

We restrict our analysis to *preference-based* contextual deontic logic. We propose the following simple formalization of contextual deontic reasoning, where $\alpha_1 \succ \alpha_2$ is read as 'α_1 is deontically preferred to α_2' or as 'α_1 is better than α_2.'

$$O_\gamma(\alpha|\beta) =_{\text{def}} O(\alpha|\beta\backslash\neg\gamma) =_{\text{def}} \alpha \wedge \beta \wedge \gamma \succ \neg\alpha \wedge \beta$$

This is a straightforward generalization of preference-based deontic logics defined by $O(\alpha|\beta) =_{\text{def}} \alpha \wedge \beta \succ \neg\alpha \wedge \beta$, see e.g. (van der Torre, 1997). In fact, preference-based dyadic obligations can be defined as a special case of preference-based contextual obligations.

$$O(\alpha|\beta) =_{\text{def}} O_\top(\alpha|\beta)$$

Despite their advantages, preference-based deontic logics also validate the theorem $O(\alpha \mid \beta) \leftrightarrow O(\alpha \wedge \beta \mid \beta)$, which is counterintuitive at first sight. For a discussion on this theorem, see e.g. (Hansson, 1971). Moreover, it is easily seen that preference-based contextual deontic logic has a similar theorem for the context γ, namely

$O_\gamma(\alpha \mid \beta) \leftrightarrow O_{\gamma\wedge\beta\wedge\alpha}(\alpha \mid \beta)$. These two theorems, characteristic for preference-based logics, are not further discussed in this paper, because we are interested in the relation between deontic, contextual and defeasible reasoning.

In this paper we only consider strong preferences, i.e. preferences that have the following property of restricted (left and right) strengthening, that formalizes that there is no overridden defeasibility in the logic. Strengthening is restricted by a 'consistency check' that formalizes that 'ought implies can,' in which $\overset{\leftrightarrow}{\Diamond}\alpha$ stand for 'α is possible' or 'α is propositionally consistent.'

$$(\alpha_1 \succ \beta_1) \wedge \overset{\leftrightarrow}{\Diamond}(\alpha_1 \wedge \alpha_2) \rightarrow (\alpha_1 \wedge \alpha_2 \succ \beta_1 \wedge \beta_2)$$

This restricted strengthening of the preferences, i.e. lack of overridden defeasibility, corresponds for the contextual obligations to restricted strengthening of the antecedent as well as restricted strengthening of the context.

$$\frac{O_\gamma(\alpha|\beta), \overset{\leftrightarrow}{\Diamond}(\alpha \wedge \beta \wedge \beta' \wedge \gamma)}{O_\gamma(\alpha|\beta \wedge \beta')} \qquad \frac{O(\alpha|\beta \backslash \neg\gamma), \overset{\leftrightarrow}{\Diamond}(\alpha \wedge \beta \wedge \beta' \wedge \gamma)}{O(\alpha|\beta \wedge \beta' \backslash \neg\gamma)}$$

$$\frac{O_\gamma(\alpha|\beta), \overset{\leftrightarrow}{\Diamond}(\alpha \wedge \beta \wedge \gamma \wedge \gamma')}{O_{\gamma\wedge\gamma'}(\alpha|\beta)} \qquad \frac{O(\alpha|\beta \backslash \neg\gamma), \overset{\leftrightarrow}{\Diamond}(\alpha \wedge \beta \wedge \gamma \wedge \gamma')}{O(\alpha|\beta \backslash \neg\gamma \vee \neg\gamma')}$$

A side-effect of the definition of the preference-based contextual obligation and restricted strengthening of the preferences is the following weakening of the consequent. It illustrates the interaction between antecedent and context, which is useful in the analysis of contrary-to-duty reasoning. For example, consider the second best plan paradox of Example 2. From the obligation *'give an aspirin today and give one tomorrow'* $O_\top(a_1 \wedge a_2|\top)$ we can derive the obligation *'give an aspirin tomorrow unless no-one is given today'* $O(a_2|\top \backslash \neg a_1)$, but not an obligation *'give an aspirin tomorrow'* simpliciter $O_\top(a_2|\top)$.

$$\frac{O_\gamma(\alpha \wedge \alpha'|\beta)}{O_{\gamma\wedge\alpha'}(\alpha|\beta)} \qquad \frac{O(\alpha \wedge \alpha'|\beta \backslash \neg\gamma)}{O(\alpha|\beta \backslash \neg\gamma \vee \neg\alpha')}$$

Whether we use $O_\gamma(\alpha \mid \beta)$ or $O(\alpha \mid \beta \backslash \gamma)$ is arbitrary, of course, but in many cases the the distinction between context and condition (antecedent) is unclear. The interpretation of the first type is therefore more difficult, and we advise to use the second type. In the following section we axiomatize a preference-based contextual deontic logic.

3. Axiomatization

Contextual obligations are formalized in Boutilier's modal preference logic CT4O, a bimodal propositional logic of inaccessible worlds. For the details and completeness proof of this logic see (Boutilier, 1994). The advantages of our formalization in a modal framework, where an operator is given by a definition in an underlying logic, is that we get an axiomatization for free! We do not have to look for a sound and complete set of inference rules and axiom schemata, because we simply take the axiomatization of the underlying logic together with the new definition. In other words, the problem of finding a sound and complete axiomatization is replaced by the problem of finding a definition of a contextual obligation in terms of a monadic modal preference logic. In the logic we abstract from actions, time and individuals.

Definition 1 (CT4O). The logic CT4O is a propositional bimodal system with the two normal modal connectives $\Box$ and $\overleftarrow{\Box}$. Dual 'possibility' connectives $\Diamond$ and $\overleftarrow{\Diamond}$ are defined as usual and two additional modal connectives $\overleftrightarrow{\Box}$ and $\overleftrightarrow{\Diamond}$ are defined as follows.

$$
\begin{array}{lll@{\qquad}lll}
\Diamond\alpha & =_{\mathrm{def}} & \neg\Box\neg\alpha & \overleftrightarrow{\Box}\alpha & =_{\mathrm{def}} & \Box\alpha \wedge \overleftarrow{\Box}\alpha \\
\overleftarrow{\Diamond}\alpha & =_{\mathrm{def}} & \neg\overleftarrow{\Box}\neg\alpha & \overleftrightarrow{\Diamond}\alpha & =_{\mathrm{def}} & \Diamond\alpha \vee \overleftarrow{\Diamond}\alpha
\end{array}
$$

CT4O is axiomatized by the following axioms and inference rules.

K	$\Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$	**Nes**	From α infer $\overleftrightarrow{\Box}\alpha$
K′	$\overleftarrow{\Box}(\alpha \to \beta) \to (\overleftarrow{\Box}\alpha \to \overleftarrow{\Box}\beta)$	**MP**	From $\alpha \to \beta$ and α infer β
T	$\Box\alpha \to \alpha$		
4	$\Box\alpha \to \Box\Box\alpha$		
H	$\overleftrightarrow{\Diamond}(\Box\alpha \wedge \overleftarrow{\Box}\beta) \to \overleftrightarrow{\Box}(\alpha \vee \beta)$		

Kripke models $M = \langle W, \leq, V\rangle$ for CT4O consist of W, a set of worlds, $\leq$, a binary transitive and reflexive accessibility relation, and V, a valuation of the propositional atoms in the worlds. The partial pre-ordering $\leq$ expresses preferences: $w_1 \leq w_2$ if and only if w_1 is as least as preferable as w_2. The modal connective $\Box$ refers to accessible worlds and the modal connective $\overleftarrow{\Box}$ to inaccessible worlds.

$M, w \models \Box\alpha$ iff $\forall w' \in W$ if $w' \leq w$, then $M, w' \models \alpha$
$M, w \models \overleftarrow{\Box}\alpha$ iff $\forall w' \in W$ if $w' \not\leq w$, then $M, w' \models \alpha$

Contextual obligations are defined in CT4O as follows. In this paper, we do not discuss the properties of $\succ$ but we focus on the properties of the contextual obligations.[4]

*Definition 2 (*CDL*).* The logic CDL is the logic CT4O extended with the following definitions of contextual obligations. The contextual obligation 'α should be the case if β is the case unless γ is the case', written as $O(\alpha|\beta\backslash\gamma)$, is defined as a strong preference of $\alpha\wedge\beta\wedge\neg\gamma$ over $\neg\alpha\wedge\beta$.

$$\begin{array}{lll}
\alpha_1 \succ \alpha_2 & =_{\mathrm{def}} & \overleftrightarrow{\Box}(\alpha_1 \rightarrow \Box\neg\alpha_2) \wedge \overleftrightarrow{\Diamond}\alpha_1 \\
O(\alpha|\beta\backslash\gamma) & =_{\mathrm{def}} & (\alpha\wedge\beta\wedge\neg\gamma) \succ (\neg\alpha\wedge\beta) \\
 & = & \overleftrightarrow{\Box}((\alpha\wedge\beta\wedge\neg\gamma) \rightarrow \Box(\beta\rightarrow\alpha)) \wedge \overleftrightarrow{\Diamond}(\alpha\wedge\beta\wedge\neg\gamma) \\
O_\gamma(\alpha|\beta) & =_{\mathrm{def}} & O(\alpha|\beta\backslash\neg\gamma)
\end{array}$$

From the definitions follows immediately the following satisfiability conditions for the two additional modal connectives: $M,w \models \overleftrightarrow{\Box}\alpha$ if and only if $\forall w' \in W\ M,w' \models \alpha$ and $M,w \models \overleftrightarrow{\Diamond}\alpha$ if and only if $\exists w' \in W\ M,w' \models \alpha$. As a consequence, the truth value of a contextual obligation does not depend on the world in which the obligation is evaluated. For a model $M = \langle W, \leq, V\rangle$ we have $M \models O(\alpha\,|\,\beta\backslash\gamma)$ (i.e. for all worlds $w \in W$ we have $M,w \models O(\alpha\,|\,\beta\backslash\gamma)$) if and only if there is a world $w \in W$ such that $M,w \models O(\alpha|\beta\backslash\gamma)$.

The following proposition shows the truth conditions of contextual obligations.

Proposition 1 (Contextual obligation). Let$M = \langle W, \leq, V\rangle$ be a CT4O model and let $|\alpha|$ be the set of worlds that satisfy α. For a world $w \in W$, we have $M,w \models O(\alpha|\beta\backslash\gamma)$ if and only if for all $w_1 \in |\alpha\wedge\beta\wedge\neg\gamma|$ and all $w_2 \in |\neg\alpha\wedge\beta|$ we have $w_2 \not\leq w_1$, and such a world w_1 exists.

Proof Follows directly from the definition of $\succ$.

The preference-based dyadic obligations $O(\alpha\,|\,\beta) =_{\mathrm{def}} O_\top(\alpha\,|\,\beta)$ are weak variants of the obligations defined in so-called Prohairetic Deontic Logic (PDL) in (van der Torre & Tan, 1997b). In PDL, obligations have an additional restriction and are defined by

$$O_{\mathrm{PDL}}(\alpha|\beta) =_{\mathrm{def}} (\alpha\wedge\beta \succ \neg\alpha\wedge\beta) \wedge I(\alpha|\beta)$$

where $I(\alpha|\beta)$ stands for 'the ideal (or preferred) β worlds are α worlds.' As a consequence of the additional condition, dilemmas are inconsistent in PDL. Nevertheless, weak variants of the properties of PDL can be given for CDL. We therefore immediately have that CDL has restricted strengthening of the antecedent, restricted deontic detachment, and a restricted conjunction rule. For the details see (van der Torre & Tan, 1997b; van der Torre & Tan, 1998b). The following proposition shows several properties of contextual obligations relevant for the analysis of contextual reasoning. In contrast to PDL, CDL also has a kind of weakening of the consequent.

Proposition 2 (Theorems of CDL*).* The logic CT4O validates the following theorems.

RSA: $(O(\alpha|\beta_1\backslash\gamma)\wedge \overleftrightarrow{\Diamond}(\alpha\wedge\beta_1\wedge\beta_2\wedge\neg\gamma))\rightarrow O(\alpha|\beta_1\wedge\beta_2\backslash\gamma)$

RSC: $(O(\alpha|\beta\backslash\gamma_1)\wedge \overleftrightarrow{\Diamond}(\alpha\wedge\beta\wedge\neg\gamma_1\wedge\neg\gamma_2))\rightarrow O(\alpha|\beta\backslash\gamma_1\vee\gamma_2)$

WC: $O(\alpha_1\wedge\alpha_2|\beta\backslash\gamma)\rightarrow O(\alpha_1|\beta\backslash\gamma\vee\neg\alpha_2)$

Proof The theorems can easily be proven in the preferential semantics. Consider **WC**. Assume $M\models O(\alpha_1\wedge\alpha_2|\beta\backslash\gamma)$. Let $W_1=|\alpha_1\wedge\alpha_2\wedge\beta\wedge\neg\gamma|$ and $W_2=|\neg(\alpha_1\wedge\alpha_2)\wedge\beta|$, and $w_2\not\leq w_1$ for $w_1\in W_1$ and $w_2\in W_2$. Moreover, let $W_1'=|\alpha_1\wedge\beta\wedge\neg(\gamma\vee\neg\alpha_2)|$ and $W_2'=|\neg\alpha_1\wedge\beta|$. We have $w_2\not\leq w_1$ for $w_1\in W_1'$ and $w_2\in W_2'$, because $W_1=W_1'$ and $W_2'\subseteq W_2$. Thus, $M\models O(\alpha_1|\beta\backslash\gamma\vee\neg\alpha_2)$. Verification of the other theorems is left to the reader.

To illustrate the properties of CDL, we compare it with Hansson's minimizing dyadic deontic logic. First we recall some well-known definitions and properties of this logic. In Hansson's classical preference semantics (Hansson, 1971), as studied by Lewis (Lewis, 1974), a dyadic obligation, which we denote by $O_{HL}(\alpha|\beta)$, is true in a model if and only if $I(\alpha|\beta)$ is true: 'the ideal (or preferred) β worlds satisfy α.' A weaker version of this definition for our models (which are not necessarily connected, i.e. we can have incomparable worlds) is that $O^w_{HL}(\alpha|\beta)$ is true in a model if and only if 'there is an *equivalence class* of minimal (or preferred) β worlds that satisfy α.' These obligations are not closed under the conjunction rule $Op\wedge Oq\rightarrow O(p\wedge q)$ and therefore allow for a consistent representation of dilemmas.

Definition 3 (Minimizing obligation). Let $M=\langle W,\leq,V\rangle$ be a Kripke model and $|\alpha|$ be the set of all worlds of W that satisfy α. The weak Hansson-Lewis obligation 'α should be the case if β is the case', written as $O^w_{HL}(\alpha|\beta)$, is defined as follows.

$$O^w_{HL}(\alpha|\beta) \quad =_{\text{def}} \quad \overleftrightarrow{\Diamond}(\beta\wedge\Box(\beta\rightarrow\alpha))$$

The model M satisfies the weak Hansson-Lewis obligation 'α should be the case if β is the case', written as $M\models O^w_{HL}(\alpha\,|\,\beta)$, if and only if there is a world $w_1 \in |\alpha \wedge \beta|$ such that for all $w_2 \in |\neg\alpha \wedge \beta|$ we have $w_2 \not\leq w_1$. The following proposition shows that the expression $O^w_{HL}(\alpha|\beta)$ corresponds to a weak Hansson-Lewis minimizing obligation. For simplicity, we assume that there are no infinite descending chains.

Proposition 3. Let $M = \langle W, \leq, V \rangle$ be a CT4O model, such that there are no infinite descending chains. As usual, we write $w_1 < w_2$ for $w_1 \leq w_2$ and not $w_2 \leq w_1$, and $w_1 \sim w_2$ for $w_1 \leq w_2$ and $w_2 \leq w_1$. A world w is a minimal β-world, written as $M, w \models_{\leq} \beta$, if and only if $M, w \models \beta$ and for all $w' < w$ holds $M, w' \not\models \beta$. A set of worlds is an equivalence class of minimal β-worlds, written as E_β, if and only if there is a w such that $M, w \models_{\leq} \beta$ and $E_\beta = \{w' \mid M, w' \models \beta \text{ and } w \sim w'\}$. We have $M \models O^w_{HL}(\alpha \mid \beta)$ if and only if there is an E_β such that $E_\beta \subseteq |\alpha|$.

Proof $\Leftarrow$ Follows directly from the definitions. Assume there is a w such that $M, w \models_{\leq} \beta$ and $E_\beta = \{w' \mid M, w' \models \beta \text{ and } w \sim w'\}$ and $E_\beta \subseteq |\alpha|$. For all $w_2 \in |\neg\alpha \wedge \beta|$ we have $w_2 \not\leq w$.

$\Rightarrow$ Assume that there is a world $w_1 \in |\alpha \wedge \beta|$ such that for all $w_2 \in |\neg\alpha \wedge \beta|$ we have $w_2 \not\leq w_1$. Let w be a minimal β-world such that $M, w \models_{\leq} \beta$ and $w \leq w_1$ (that exists because there are no infinite descending chains), and let $E_\beta = \{w' \mid M, w' \models \beta \text{ and } w \sim w'\}$.

Now we are ready to compare our contextual deontic logic with Hansson's dyadic deontic logic. The following proposition shows that under a certain condition, the contextual obligation $O(\alpha|\beta\backslash\gamma)$ is true in a model if and only if a set of the weak Hansson-Lewis minimizing obligations $O^w_{HL}(\alpha|\beta')$ is true in the model. The condition is that the formal language is expressive enough to distinguish between each two worlds.

Proposition 4. Let $M = \langle W, \leq, V \rangle$ be a CT4O model, that has no worlds that satisfy the same propositional sentences. Hence, we identify the set of worlds with a set of propositional interpretations, such that there are no duplicate worlds. We have $M \models O(\alpha \mid \beta \backslash \gamma)$ if and only if there are $\alpha \wedge \beta \wedge \neg\gamma$ worlds, and for all propositional β' such that $M \models \overset{\leftrightarrow}{\Box}(\beta' \to \beta)$ and $M \not\models \overset{\leftrightarrow}{\Box}(\beta' \to \alpha \to \gamma)$, we have $M \models O^w_{HL}(\alpha|\beta')$.

Proof $\Rightarrow$ There exists a β' world that satisfies $\alpha \wedge \beta \wedge \neg\gamma$, because $M \not\models \overset{\leftrightarrow}{\Box}(\beta' \to \alpha \to \gamma)$. Each $\alpha \wedge \beta \wedge \neg\gamma$ world satisfies $\beta \wedge \Box(\beta \to \alpha)$, as a consequence of the definition of $O(al|\beta\backslash\gamma)$.

$\Leftarrow$ Every world is characterized by a unique propositional sentence. Let $\overline{w}$ denote the sentence that uniquely characterizes world w. Proof by contraposition. If $M \not\models O(\alpha|\beta\backslash\gamma)$, then either there are no $\alpha \wedge \beta \wedge \neg\gamma$ worlds (in which case the proof is trivial), or there are w_1, w_2 such that $M, w_1 \models \alpha \wedge \beta \wedge \neg\gamma$, $M, w_2 \models \neg\alpha \wedge \beta$ and $w_2 \leq w_1$. Choose $\beta' = \overline{w_1} \vee \overline{w_2}$. The world w_2 is an element of the preferred β' worlds, because there are no duplicate worlds. (If duplicate worlds are allowed, then there could be a β' world w_3 which is a duplicate of w_1, and which is strictly preferred to w_1 and w_2.) We have $M, w_2 \not\models \alpha$ and therefore $M \not\models O^w_{HL}(\alpha|\beta')$.

The examples in the following section illustrate the CDL analysis of the contrary-to-duty paradoxes.

4. Examples: the contrary-to-duty paradoxes

Preference-based deontic logics have proven useful to understand the nature of contrary-to-duty (CTD) and according-to-duty (ATD) reasoning. An obligation $O(\alpha|\beta)$ is a CTD (ATD) obligation of the *primary* obligation $O(\alpha_1|\beta_1)$ if and only if $\beta \wedge \alpha_1$ is inconsistent (β propositionally implies α_1), as represented in Figure 1.

$$\begin{tikzcd} O(\alpha_1|\beta_1) \arrow[d, leftrightarrow, "\text{inconsistent}"'] & O(\alpha_1|\beta_1) \\ O(\alpha|\beta) & O(\alpha|\beta) \arrow[u, "\text{implies}"] \end{tikzcd}$$

Figure 1. $O(\alpha|\beta)$ is a CTD / ATD obligation of $O(\alpha_1|\beta_1)$

Some typical examples are given below. The first two sets show how the preference-based representation reveals the structure of the CTDs and ATDs. In particular, it shows that they are *not* sets of conflicting preferences. Moreover, the third set TRANS combines the structural properties of CTD and ATD. It plays an important role in defeasible deontic logic, see Section 5.

$$\begin{array}{cc} CTD: & \\ O(a_1|\top), O(a_2|\neg a_1) & a_1 \succ \neg a_1, a_2 \wedge \neg a_1 \succ \neg a_2 \wedge \neg a_1 \\ ATD: & \\ O(a_1|a_1 \vee a_2), O(a_1 \vee a_2|\top) & a_1 \succ \neg a_1 \wedge a_2, a_1 \vee a_2 \succ \neg a_1 \wedge \neg a_2 \\ TRANS: & \\ O(a_1|a_1 \vee a_2), O(a_1 \vee a_2|\neg a_1) & a_1 \succ \neg a_1 \wedge a_2, \neg a_1 \wedge a_2 \succ \neg a_1 \wedge \neg a_2 \end{array}$$

The following two examples reconsider the contrary-to-duty examples discussed in the introduction.

Example 4 (Second best plan paradox, continued). Consider the obligations *'dr Denton ought to give an aspirin today and an aspirin tomorrow'* $O(a_1 \wedge a_2 \mid \top)$ and *'if he does not give an aspirin today, then he ought not to give one tomorrow'* $O(\neg a_2 \mid \neg a_1)$. The context is used to assure that counterintuitive derivations are blocked. If the context of $O(a_1 \wedge a_2 \mid \top)$ is 'always,' then the context of the derived obligation $O(a_2 \mid \top)$ is 'unless $\neg a_1$.' A simple consistency check ensures that $O(a_2 \mid \top)$ 'unless $\neg a_1$' cannot be used to derive $O(a_2 \mid \neg a_1)$, because the antecedent of the latter violates the context of the former. We can derive $O(a_2 \mid \top \setminus \neg a_1)$, but we cannot derive $O_\gamma(a_2 \mid \neg a_1)$ for any γ. The unless clause $\neg a_1$ prevents that the obligation from the ideal best plan context is strengthened to the second best plan context.

Example 5 (Cottage housing regulations, continued). Consider the obligations *'there ought to be no fence'* $O(\neg f \mid \top)$, *'if there is a fence, then it ought to be white'* $O(w \wedge f \mid f)$, *'there ought not to be a dog'* $O(\neg d \mid \top)$, *'if there is no dog, then there ought to be no sign'* $O(\neg s \mid \neg d)$ and *'if there is a dog, then there ought to be a sign'* $O(s \mid d)$. We cannot derive $O(\neg(f \wedge w) \mid f)$ from $O(\neg f \mid \top)$, as desired. Moreover, we can derive $O(\neg s \wedge \neg d \mid \top)$, but we cannot derive $O(\neg s \mid d)$ from $O(\neg d \mid \top)$ and $O(\neg s \mid \neg d)$, as desired. Finally, we can derive the obligation *'to have no sign'* in the ideal context there is no dog $O_{\neg d}(\neg s \mid \top)$. In other words, *'there should not be a sign, unless there is a dog'* $O(\neg s \mid \top \setminus d)$. Note that the latter two obligations are much stronger than the premise $O(\neg s \mid \neg d)$. On the other hand, we cannot derive an obligation *'there should not be a sign'* simpliciter $O_\top(\neg s \mid \top)$.

The cottage housing regulations in Example 5 illustrate that in CDL $O(\alpha_1 \mid \neg \alpha_2)$ cannot be derived from $O(\alpha_1 \wedge \alpha_2 \mid \top)$, because we cannot derive a CTD obligation from its primary obligation. We can derive $O_{\alpha_2}(\alpha_1 \mid \top)$ from $O_\top(\alpha_1 \wedge \alpha_2)$, but this derived obligation cannot be used to derive the counterintuitive $O(\alpha_1 \mid \neg \alpha_2)$ due to the context. The following example illustrates that the derivation of the obligation $O(\alpha_1 \mid \neg \alpha_2)$ from the obligation $O(\alpha_1 \wedge \alpha_2 \mid \top)$ is a fundamental problem underlying several contrary-to-duty paradoxes. Hence, the underlying problem of the contrary-to-duty paradoxes is that a contrary-to-duty obligation can be derived from its primary obligation.

Example 6 (Contrary-to-duty paradoxes). Assume a dyadic deontic logic that validates at least substitution of logical equivalents and the following inference patterns *Restricted Strengthening of the Antecedent* (RSA), *Weakening of the Consequent* (WC), *Conjunction of the Antecedent* (AND) and a version of *Deontic Detachment* (DD), in which $\overset{\leftrightarrow}{\Diamond}$is a modal operator and $\overleftrightarrow{\Diamond}\phi$ is true for all consistent propositional formulas ϕ.

$$\text{RSA}: \frac{O(\alpha|\beta_1), \overleftrightarrow{\Diamond}(\alpha \wedge \beta_1 \wedge \beta_2)}{O(\alpha|\beta_1 \wedge \beta_2)} \qquad \text{WC}: \frac{O(\alpha_1|\beta)}{O(\alpha_1 \vee \alpha_2|\beta)}$$

$$\text{AND}: \frac{O(\alpha_1|\beta), O(\alpha_2|\beta)}{O(\alpha_1 \wedge \alpha_2|\beta)} \qquad \text{DD}: \frac{O(\alpha|\beta), O(\beta|\gamma)}{O(\alpha \wedge \beta|\gamma)}$$

Furthermore, consider the sets

$$S = \{O(\neg k|\top), O(g|k), \vdash g \rightarrow k, k\}$$
$$S' = \{O(a|\top), O(t|a), O(\neg t|\neg a), \neg a\}$$
$$S'' = \{O(\neg a|\top), O(a \vee p|\top), O(\neg p|a)\}$$
$$S''' = \{O(\neg r \wedge \neg g|\top), O(r|g)\}$$

S formalizes Forrester's paradox (Forrester, 1984) when k is read as 'Smith kills Jones' and g as 'he kills him gently,' S' formalizes Chisholm's paradox (Chisholm, 1963) when a is read as 'a certain man going to the assistance of his neighbors' and t as 'the man telling his neighbors that he will come,' S'' formalizes the apples-and-pears example (Tan & van der Torre, 1996) when a is read as 'buying apples' and p as 'buying pears,' and finally, S''' formalizes a part of the Reykjavik scenario (Belzer, 1986; van der Torre, 1994) when r is read as telling a secret to Reagan and g as telling it to Gorbatsjov. The last obligation of each premise set is a contrary-to-duty obligation of the first obligation of the set, because its antecedent is contradictory with the consequent of the latter. The paradoxical consequences of the sets of obligations are represented in Figure 2. The underlying problem of the counter-intuitive derivations is the derivation of the obligation $O(\alpha_1 \mid \neg\alpha_2)$ from $O(\alpha_1 \wedge \alpha_2 \mid \top)$ by WC and RSA: respectively the derivation of $O(\neg(g \wedge k)|k)$ from $O(\neg k|\top)$, $O(t|\neg a)$ from $O(a \wedge t|\top)$, $O(p|a)$ from $O(\neg a \wedge p|\top)$, and $O(\neg r|g)$ from $O(\neg r \wedge \neg g|\top)$. These derivations are blocked in CDL.

5. Defeasible deontic logic

A defeasible deontic logic is a deontic logic in which the obligations do not have unrestricted strengthening of the antecedent.[5] Defeasibility, i.e. lack of strengthening of the antecedent, has been used to analyze the following four phenomena. The role of defeasibility in deontic reasoning is complex, because different types of defeasibility have to be distinguished for the different phenomena, see (van der Torre & Tan, 1995; van der Torre & Tan, 1997a).

$$\dfrac{\dfrac{\dfrac{O(\neg k|\top)}{O(\neg g|\top)}\,\text{WC}}{O(\neg g|k)}\,\text{RSA} \qquad O(g|k)}{O(\neg g \wedge g|k)}\,\text{AND}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{O(t|a) \quad O(a|\top)}{O(a \wedge t|\top)}\,\text{DD}}{O(t|\top)}\,\text{WC}}{O(t|\neg a)}\,\text{RSA} \qquad O(\neg t|\neg a)}{O(t \wedge \neg t|\neg a)}\,\text{AND}$$

$$\dfrac{\dfrac{\dfrac{\dfrac{O(\neg a|\top) \quad O(a \vee p|\top)}{O(\neg a \wedge p|\top)}\,\text{AND}}{O(p|\top)}\,\text{WC}}{O(p|a)}\,\text{RSA} \qquad O(\neg p|a)}{O(p \wedge \neg p|a)}\,\text{AND}$$

$$\dfrac{\dfrac{\dfrac{O(\neg r \wedge \neg g|\top)}{O(\neg r|\top)}\,\text{WC}}{O(\neg r|g)}\,\text{RSA} \qquad O(r|g)}{O(r \wedge \neg r|g)}\,\text{AND}$$

Figure 2. Four contrary-to-duty paradoxes

Prima facie obligations: overridden defeasibility. The first and most obvious phenomena is prima facie reasoning, or, more generally, the reasoning about obligations that can be overridden by stronger obligations. For example, *'you should keep your promises,'* but *'one may break a promise in order to prevent a disaster'* (Ross, 1930), *'normally you should not eat with your fingers,'* but *'if you are served asparagus, then normally you should eat with your fingers'* (Horty, 1994), *'normally there should not be a fence,'* but *'if the house in next to a cliff then normally you should have a fence'* (Prakken & Sergot, 1996), etc.

Dilemmas: conflict defeasibility. Horty (1994; 1997) gives a normative interpretation of Reiter's default logic, which is used to analyze dilemmas. He shows formal links between Reiter's default logic and Van Fraassen's analysis of dilemmas (van Fraassen, 1973).

Unresolved conflicts: conflict defeasibility. Related to dilemmas is the issue of unresolved conflicts. For example, the two obligations of Von Wright's window problem (von Wright, 1963), *'the window ought to be closed if it rains'* $O(c\,|\,r)$ and *'it should be open if the sun shines'* $O(\neg c|s)$, conflict when it rains and the sun shines. We cannot derive $O(c\,|\,r \wedge s)$ nor $O(\neg c\,|\,r \wedge s)$. The simplest example consists of two unrelated obligations *'be polite'* Op and *'be honest'* Oh, that conflict if we only consider $p \leftrightarrow \neg h$ worlds. Consequently, in some logics we cannot derive $O(p|p \leftrightarrow \neg h)$.

Violability: factual defeasibility. In this paper we illustrated how factual defeasibility can be used to analyze CTD and ATD reasoning. CTD and ATD reasoning is context-dependent, and contextual reasoning is defeasible.

In this paper we analyzed the relation between contextual reasoning and deontic reasoning in which obligations cannot be overridden by other obligations. The relation between contextual reasoning and *defeasible* deontic reasoning in which obligations can be overridden is much more complex. It can be analyzed by making strengths and impacts of obligations explicit in the formal language. A well-known example of the context-dependence of obligations is related to the specificity principle. In most cases the strengths of obligations are unrestricted, but if an obligation is more specific than and conflicting with another obligation, then the strength of the former is higher than the strength of the latter. In general, strengths of obligations are restricted to resolve conflicts. Moreover, there are types of context-dependence that do not depend on the strengths of obligations or conflicts between obligations, as illustrated in the example below. In the framework of (van der Torre & Weydert, 1998) we therefore distinguish between the strengths of obligations and what we call their impact. The following example illustrates the two different types of context-dependence and the distinction between strength and impact.

Example 7. Consider the two sets $S_1 = \{O(\neg s|\top), O(s|i)\}$ and $S_2 = \{O(p|p \vee c), O(c|c \vee h)\}$, with background knowledge $\top \leftrightarrow (p \vee c \vee h)$, $\neg(p \wedge c)$, $\neg(p \wedge h)$, $\neg(c \wedge h)$ (the three variables p, c and h are mutually exclusive and exhaustive), and with for each set strengths s_1 and s_2, and impacts α_1 and α_2. The first type of context-dependence (set S_1) is a consequence of resolving conflicts and the second type of context-dependence (set S_2) follows from the relation between obligations and preferences; it is *not* a consequence of resolving conflicts. As represented in Figure 3, the representation of obligations suggests that S_2 is a conflict, because the consequents of the obligations are contradictory. However, this is not a sufficient reason for two obligations to conflict. In contrast, the set S_2 is conflict free, because the antecedents of both obligations are only true if c is the case, and in that case the first obligation can no longer be fulfilled. This is especially clear when the obligations are rewritten as preferences. An obligation $O(a\,|\,b)$ is formalized by a preference of the fulfilled obligation $a \wedge b$ over the violated obligation $\neg a \wedge b$, which can be written as $O(a\,|\,b) =_{\text{def}} a \wedge b \succ \neg a \wedge b$. Hence, the set S_2 can also be written as $\{p \succ c, c \succ h\}$. The representation of preferences reveals that S_2 is conflict-free, because $p \wedge h$ is inconsistent. There are not two states such that the first is preferred to the second one by the first preference, and vice versa by the second preference. The reason that the impact of $O(c\,|\,c \vee h)$ is higher than the impact of $O(p|p \vee c)$ also follows from the formalization of the obligations as preferences. The context-dependent meaning of the obligations is such that it reflects the preferences.

$$\begin{tikzcd} O(\neg s|\top) \arrow[d, leftrightarrow, shift right, "\text{inconsistent}"'] & O(p|p \vee c) \arrow[d, leftrightarrow, "\text{inconsistent}"] & p \succ c \arrow[d, leftrightarrow, "\text{inconsistent}"] \\ O(s|i) \arrow[u, shift right, "\text{implies}"'] & O(c|c \vee h) & c \succ h \end{tikzcd}$$

Figure 3. Context-dependent obligations: specificity and conflict-free

6. Related research

There are two types of dyadic deontic logics, dependent on how the antecedent is interpreted. The first type, as advocated by Chellas (Chellas, 1974; Alchourrón, 1993), defines a dyadic obligation in terms of a monadic obligation by $O(\alpha|\beta) =_{\text{def}} \beta > O\alpha$, where '$>$' is a strict implication. These dyadic deontic logics have strengthening of the antecedent, but they cannot represent the contrary-to-duty paradoxes in a consistent way. Dyadic deontic logics of the second type, as introduced by Hansson (Hansson, 1971) and further investigated by Lewis (Lewis, 1974), do not have strengthening of the antecedent and they therefore can represent the paradoxes in a consistent way.

In (Tan & van der Torre, 1996) we proposed the two-phase deontic logic 2DL to combine restricted strengthening of the antecedent (RSA) and weakening of the consequent (WC). In 2DL, the first phase validates RSA but not WC, and the second phase validates WC but not RSA. 2DL can be used to analyze the contrary-to-duty paradoxes, because the obligation $O(\alpha_1|\neg\alpha_2)$ cannot be derived from $O(\alpha_1 \wedge \alpha_2|\top)$ by respectively WC and RSA. Proof-theoretically, such sequencing of derivations is rather complicated, but semantically, the two-phase approach simply means that first an ordering has to be constructed before it can be used for minimization.

Prakken and Sergot formalize the context as a propositional sentence, and write contextual obligations 'α ought to be the case in context β' as $O_\beta\alpha$. In a second paper, they extend the classical dyadic deontic logics of Hansson and Lewis (Hansson, 1971; Lewis, 1974). They look for a consistency check to block strengthening of the context (Prakken & Sergot, 1996) or strengthening of the antecedent (Prakken & Sergot, 1997). Moreover, in (Prakken & Sergot, 1997) they show it is impossible to construct such a consistency check in the classical Hansson-Lewis semantics.

Makinson (Makinson, 1998) proposes a labeled deontic logic in which the inference relation is relative to a set of explicit promulgations following (Alchourrón & Bulygin, 1981). He uses an hypothesis that ensures that obligations are not strengthened to another context. However, this hypothesis is not represented in the formal language.

The contextual obligations $O_\gamma(\alpha \mid \beta)$ as defined in this paper are closely related to labeled obligations $O(\alpha|\beta)_L$, to be read as 'α ought to be (done) if β is (done) against the background of L' (van der Torre & Tan, 1995; van der Torre & Tan, 1997a; van der Torre, 1998). Labeled deontic logic LDL is a labeled deductive system as introduced by Gabbay in (Gabbay, 1996). It is a powerful technique to define complex inductive definitions with several arguments. The label formalizes the context in which an obligation is derived. The advantage of CDL over LDL is that it has a semantics, and that its language contains negations and disjunctions of obligations, as well as facts (and thus violations).

7. Conclusions

Contextual deontic logic CDL is a defeasible deontic logic that has factual defeasibility but not overridden defeasibility. This limited expressive power is useful to analyze contrary-to-duty reasoning without confusing factual defeasibility with overridden defeasibility, i.e. without confusing violations with exceptions.

Notes

[1] Contexts are not only useful to analyze contrary-to-duty reasoning, but they can also be used to analyze according-to-duty reasoning related to so-called deontic detachment, also called deontic transitivity. Deontic detachment is notorious for its role in the Chisholm paradox, where it is combined with contrary-to-duty reasoning.

However, as Prakken and Sergot (1996) observe, the problem of the Chisholm paradox and deontic detachment is different from the problem of formalizing contrary-to-duty reasoning.

[2] It is counterintuitive to derive the obligation *'you should keep a promise and apologize for not keeping it'* from the two obligations *'you should keep a promise'* and *'if you do not keep a promise, then you should apologize for it'* and the fact *'you break a promise.'* The two obligations refer respectively to the ideal context you keep your promise and to the sub-ideal context you do not keep your promise, and therefore they should not be combined.

[3] At first sight, the monotonicity property of CDL may seem surprising. However, this property is a consequence of the restricted language, and the lack of factual detachment. If a notion of factual detachment would be added, then the logic would become non-monotonic. Hence, we do use non-monotonic techniques! This surprising combination of defeasibility and monotonicity can be further explained by comparing contextual deontic logic with Reiter defaults, which are also monotonic. Non-monotonicity only arises in Reiter's default logic when default rules are applied, which is the default logic analogue of factual detachment, to construct so-called Reiter extensions.

[4] The preference relation $\succ$ is quite weak. For example, it is not anti-symmetric (we cannot derive $\neg(\alpha_2 \succ \alpha_1)$ from $\alpha_1 \succ \alpha_2$) and it is not transitive (we cannot derive $\alpha_1 \succ \alpha_3$ from $\alpha_1 \succ \alpha_2$ and $\alpha_2 \succ \alpha_3$). The lack of these properties is the result of the fact that we do not have connected orderings. Moreover, this a-connectedness is crucial for our preference-based deontic logics, see (Tan & van der Torre, 1996; van der Torre & Tan, 1997b).

[5] To analyze the relation between obligations and defeasibility, we can either look at non-monotonicity in a monadic deontic logic, or at lack of strengthening of the antecedent in (an extension of) dyadic deontic logic. In this paper we followed the latter approach, based on Alchourrón's definition of a defeasible conditional. A defeasible conditional is a conditional that does not have strengthening of the antecedent.

References

Alchourrón, C., and Bulygin. 1981. The expressive conception of norms. In Hilpinen, R., ed., *New Studies in Deontic Logic: Norms, Actions and the Foundations of Ethics*. D. Reidel. 95–124.

Alchourrón, C. 1993. Philosophical foundations of deontic logic and the logic of defeasible conditionals. In Meyer, J.-J., and Wieringa, R., eds., *Deontic Logic in Computer Science: Normative System Specification*. John Wiley & Sons. 43–84.

Belzer, M. 1986. A logic of deliberation. In *Proceedings of the Fifth National Conference on Artificial Intelligence (AAAI'86)*, 38–43.

Boutilier, C. 1994. Conditional logics of normality: a modal approach. *Artificial Intelligence* 68:87–154.

Castañeda, H. 1981. The paradoxes of deontic logic: the simplest solution to all of them in one fell swoop. In Hilpinen, R., ed., *New Studies in Deontic Logic: Norms, Actions and the Foundations of Ethics*. D.Reidel Publishing company. 37–85.

Chellas, B. 1974. Conditional obligation. In Stunland, S., ed., *Logical Theory and Semantical Analysis*. Dordrecht, Holland: D. Reidel Publishing Company. 23–33.

Chisholm, R. 1963. Contrary-to-duty imperatives and deontic logic. *Analysis* 24:33–36.

Feldmann, F. 1990. A simpler solution to the paradoxes of deontic logic. In Tomberlin, J., ed., *Philosophical perspectives 4: Action theory and Philosophy of Mind.* Atascadero: Ridgview.

Forrester, J. 1984. Gentle murder, or the adverbial Samaritan. *Journal of Philosophy* 81:193–197.

Gabbay, D. 1996. *Labelled Deductive Systems*, volume 1. Oxford University Press.

Hansson, B. 1971. An analysis of some deontic logics. In Hilpinen, R., ed., *Deontic Logic: Introductory and Systematic Readings.* Dordrecht, Holland: D. Reidel Publishing Company. 121–147.

Horty, J. 1994. Moral dilemmas and nonmonotonic logic. *Journal of Philosophical Logic* 23:35–65.

Horty, J. 1997. Nonmonotonic foundations for deontic logic. In Nute, D., ed., *Defeasible Deontic Logic.* Kluwer. 17–44.

Lewis, D. 1974. Semantic analysis for dyadic deontic logic. In Stunland, S., ed., *Logical Theory and Semantical Analysis.* Dordrecht, Holland: D. Reidel Publishing Company. 1–14.

Loewer, B., and Belzer, M. 1983. Dyadic deontic detachment. *Synthese* 54:295–318.

Makinson, D. 1998. On a fundamental problem of deontic logic. In McNamara, P., and Prakken, H., eds., *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science.* IOS Press.

McCarty, L. 1994. Defeasible deontic reasoning. *Fundamenta Informaticae* 21:125–148.

Prakken, H., and Sergot, M. 1996. Contrary-to-duty obligations. *Studia Logica* 57:91–115.

Prakken, H., and Sergot, M. 1997. Dyadic deontic logic and contrary-to-duty obligations. In Nute, D., ed., *Defeasible Deontic Logic.* Kluwer. 223–262.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence* 13:81–132.

Ross, D. 1930. *The Right and the Good.* Oxford University Press.

Ryu, Y., and Lee, R. 1993. Defeasible deontic reasoning: A logic programming model. In Meyer, J.-J., and Wieringa, R., eds., *Deontic Logic in Computer Science: Normative System Specification.* John Wiley & Sons. 225–241.

Tan, Y.-H., and van der Torre, L. 1996. How to combine ordering and minimizing in a deontic logic based on preferences. In *Deontic Logic, Agency and Normative Systems. Proceedings of the ∆eon'96. Workshops in Computing*, 216–232. Springer Verlag.

Thomason, R. 1981. Deontic logic as founded on tense logic. In Hilpinen, R., ed., *New Studies in Deontic Logic: Norms, Actions and the Foundations of Ethics.* D. Reidel. 165–176.

Tomberlin, J. 1981. Contrary-to-duty imperatives and conditional obligation. *Noûs* 16:357–375.

van der Torre, L., and Tan, Y. 1995. Cancelling and overshadowing: two types of defeasibility in defeasible deontic logic. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence (IJCAI'95)*, 1525–1532. Morgan Kaufman.

van der Torre, L., and Tan, Y. 1997a. The many faces of defeasibility in defeasible deontic logic. In Nute, D., ed., *Defeasible Deontic Logic.* Kluwer. 79–121.

van der Torre, L., and Tan, Y. 1997b. Prohairetic deontic logic and qualitative decision theory. In *Proceedings of the AAAI Spring Symposium on Qualitative Approaches to Deliberation and Reasoning.*

van der Torre, L., and Tan, Y. 1998a. The temporal analysis of Chisholm's paradox. In *Proceedings of 15th National Conference on Artificial Intelligence (AAAI'98).*

van der Torre, L., and Tan, Y. 1998b. An update semantics for deontic reasoning. In McNamara, P., and Prakken, H., eds., *Norms, Logics and Information Systems. New Studies on Deontic Logic and Computer Science.* IOS Press.

van der Torre, L., and Tan, Y. 1998c. An update semantics for prima facie obligations. In *Proceedings of the Thirteenth European Conference on Artificial Intelligence (ECAI'98)*, 38–42.

van der Torre, L., and Weydert, E. 1998. Goals, desires, utilities and preferences. In *Proceedings of the ECAI'98 Workshop Decision Theory meets Artificial Intelligence.*

van der Torre, L. 1994. Violated obligations in a defeasible deontic logic. In *Proceedings of the Eleventh European Conference on Artificial Intelligence (ECAI'94)*, 371–375. John Wiley & Sons.

van der Torre, L. 1997. *Reasoning about Obligations: Defeasibility in Preference-based Deontic Logic.* Ph.D. Dissertation, Erasmus University Rotterdam.

van der Torre, L. 1998. Labeled logics of conditional goals. In *Proceedings of the Thirteenth European Conference on Artificial Intelligence (ECAI'98)*, 368–369.

van Eck, J. 1982. A system of temporally relative modal and deontic predicate logic and its philosophical application. *Logique et Analyse* 100:249–381.

van Fraassen, B. 1973. Values and the heart command. *Journal of Philosophy* 70:5–19.

von Wright, G. 1963. *The logic of preference.* Edinburgh University Press.

Yu, X. 1995. *Deontic Logic with Defeasible Detachment.* Ph.D. Dissertation, University of Georgia.