Remedial interchange, contrary-to-duty obligation and commutation

Xavier Parent

54 avenue de l'Elisa
83100 Toulon (France)
xp@up.univ-aix.fr

ABSTRACT. *This paper discusses the relation between deontic logic and the study of conversational interactions. Special attention is given to the notion of remedial interchange as analysed by sociologists and linguistic pragmaticians. This notion is close to the one of contrary-to-duty (reparational) obligation, which deontic logicians have been studying in its own right. The present article also investigates the question of whether some of the aspects of conversational interactions can fruitfully be described by using formal tools originally developed in the study of iterated belief change. I here adapt the latter tools to deontic logic, and attempt an account of remedial interchange (and, more generally, contrary-to-duty reasoning) in terms of commutation. This account brings the dynamics of obligations to the fore.*

KEYWORDS: *preference-based semantics for deontic logic, contrary-to-duty obligation, iterated revision, remedial interchange, conversational interaction, deontic tense logic.*

1. Introduction

The present article can be viewed as an attempt to explore the interface between argumentation theory and non-monotonic logics, a family of logics used to study reasoning about what is normally the case. In these formalisms, conclusions may be later withdrawn when additional information is obtained. Many researchers have proposed systems that perform such non-monotonic (or defeasible) inferences. The best known are probably: circumscription [MCC 80], default logic [REI 80], autoepistemic logic [MOO 85] and preferential models semantics [KRA 90, MAK 89]. This article focuses on applications of the latter kind of framework to the analysis of conditional obligation. Such applications were first envisaged by Danielsson [DAN 68] and Hansson [HAN 69], who suggested replacing the Kripke-type accessiblity relation by a preference relation measuring the comparative goodness of worlds. A number of researchers have followed this suggestion, providing a more comprehensive investigation of the treatment of conditional obligation within a preference-based approach ([CHE 74, LEW 74, JAC 85, MAK 93, ÅQV 93, PRA 97b, TOR 97, ÅQV 02]).

Over the past decades, non-monotonic logicians have been increasingly interested in the field of argumentation. Current research programmes in this area tend to fall into three main groups: those focusing on specific argumentative schemes, those dealing with the interface between semantics and pragmatics, and those developing a general theory of how arguments interact. The motivations of the last two groups are much less obvious than those of the first. Some comments are thus in order. First, the interest in pragmatic issues comes from the realization that some essential elements of the communicational content of an utterance, such as presupposition and conversational implicature, prove to be defeasible and subject to exceptions. This was already noticed by Grice [GRI 75], who viewed defeasibility as a defining property of so-called conversational implicatures. This observation has motivated a number of attempts to describe pragmatic inference as a non-monotonic one. Among the most representative works that have been done in this direction, one might mention the attempt made by Mercer [MER 88, MER 90] to apply default logic to the theory of presupposition, the theory of quantitative implicatures as put forth by Gazdar [GAZ 79a, GAZ 79b], or the attempt made by Perrault [PER 90], and Appelt and Konolige [APP 89] to import non-monotonic logic into speech-act theory.

Programmes whose ambition is to develop a general theory of how arguments interact mainly include systems for defeasible argumentation as developed by Pollock [POL 87], Prakken and Sartor [PRA 97a], Bench-Capon [BEN 03], and others. This third group of programmes can be viewed as an attempt to clarify and meet some of the criticisms made by Toulmin [TOU 58] against formal logic. The focus shifts here from pragmatic considerations to general questions of acceptability in the face of iterated defeat among arguments. The style of analysis that is usually adopted can be described as follows. Given a theory – i.e. a set of sentences representing the basic facts of the situation and a set of default rules, the entire set of arguments based on this theory is first determined. Next, the defeat relations among these arguments is computed. On the basis of this pattern of defeat relations, it is then possible to isolate a particular subset of arguments that are to be counted as justified (or, if you prefer, ultimately undefeated). There is a notion with which we must be very careful. It is the notion of reinstatement, which says that an argument should be counted as acceptable even if it is defeated, as long as its defeaters are themselves defeated. It is *via* this principle that the procedure isolates the set of justified arguments. As observed by Horty [HOR 01], the notion of reinstatement seems to be problematic. Intuitive counter-examples can be given to this idea, thereby setting a serious question mark against this group of theories. A more detailed discussion of this issue and related ones can be found in Prakken [PRA 02].

In this article my chief concern is to try to evaluate the extent to which deontic logic can be relevant to the study of argumentation. Special attention is given to the study of conversational interaction. In the best tradition of Goffman [GOF 71], who thought of remedial interchange as forming the nucleus of social activity, some writers tend to adopt a model of analysis in which reparational obligation plays a critical role [OWE 83, BRO 87, KER 94]. They often claim to be primarily concerned with the attempt to extract the formal pattern of conversational (face-to-face) interactions. Such a claim may, at first, be rather puzzling. Since the study of Chisholm [CHI 63], it has been increasingly clear that, as far as logic is concerned, the notion of a remedial interchange is not an easy one to grasp. The work of those interested in conversational interaction usually goes on as if the intricacies of contrary-to-duty norms had never been heard of.

This article also raises the question of whether deontic logic can benefit from concepts developed in the study of iterated belief change. My emphasis here is not on new formal results, but rather on introducing a way of thinking about contrary-to-duty contexts that is slightly different from more familiar ones. (For an overview, see [CAR 02].). The basic idea is to assume that, when they learn that an obligation has been transgressed, interactants minimally revise the ideality ordering (over possible worlds) to have the appropriate obligation deconditionalised (or detached). This process will be referred to as a *commutation*, because close examination reveals that, at the time of the violation, the set of 'second best' worlds commutes with the set of 'best' worlds. As will emerge in due course, this is very similar in structure to so-called natural revision, as developed by Boutilier [BOU 96] to give a semantic characterization of iterated belief change. All preference-based deontic logics known to me are static in the sense that there is no room left for updating the comparative goodness relation (in some way or other) so as to accomodate new information.[1] My contention here is that in the dynamics of conversational interaction there is room for such readjustments.

This paper is an attempt to substantiate these hypotheses further by elaborating the points outlined above. Section 2 presents Goffman's analysis of the notion of remedial interchange, as well as the difficulty I will focus on, which will take the form of a dilemma. Section 3 discusses one way of getting out of this dilemma, by taking time into account. This detour *via* so-called temporal deontic logic will serve as a means of introducing the basic idea underlying the account in terms of commutation, which will be the focus of attention in section 4. It will appear, then, that this second account provides a more satisfactory analysis of the notion of remedial interchange.

2. Remedial interchange

I begin by discussing the notion of remedial interchange as described by the sociologist E. Goffman in Ch. 3 of *Relations in public* (see [GOF 71]). My purpose in this paper is simply to try to connect two research traditions that have so far been developing on their own. No attempt will be made here to discuss the issue of whether Goffman's conception stands up to close scrutiny. On the other hand, presumably

1. One exception is the update semantics of van der Torre and Tan [TOR 99].

most deontic logicians will think that the notion of remedial interchange just provides one example of a contrary-to-duty (CTD) structure − one illustration among many others. I cannot here discuss in detail the many examples considered in the literature on deontic logic.

Goffman defends the view that interactions of agents can usefully be described as governed by norms. The norms Goffman is most concerned with are those to do with "territories of the self" [GOF 71, chap. 2] and the violations he is particularly interested in are those to do with "encroachments" on another's territory [GOF 71, p. 44]. The metaphor of territory is borrowed from the literature on ethology. This notion includes the following areas:

– Personal space.

– Use space: the territory immediately around or in front of an individual, his claim to which is in general respected because of instrumental needs. For instance, when I am close to a picture, others will make some effort to walk around my line of vision.

– The turn: the order in which someone receives an item with respect to someone else in the same situation. Examples abound: number-tickets; names on a receptionist's list.

– Possessional territory: any set of objects that can be identified with the self.

– Information preserve: the set of facts about himself to which an individual expects to control access while in the presence of others. These are traditionally treated under the heading of "privacy".

– Conversation preserve: the right of an individual to exert some control over who can summon him and when he can be summoned.

From the description of these types of territories, it emerges that the notion of territory is a normative one, in the sense that in general its occupier has the *right* to claim it as their own and the *right* to attempt to exclude others from it. When an individual encroaches on another's territory, a particular kind of activity occurs: remedial interchange. Its function is a negative one: it is a means of avoiding conflict, by

> "transforming what could be seen as offensive into what can be seen as acceptable". [GOF 71, p. 109]

Goffman sees remedial interchange as having three basic forms: accounts, apologies and requests. It is here unnecessary to examine these three kinds of communicative acts (and, more generally, the notion of remedial interchange) in the light of speech acts theory, since in this paper I aim simply to point out the relevance of deontic logic for those working on face-to-face interactions in the tradition of Goffman. One might mention Owen [OWE 83], Brown and Levinson [BRO 87] and Kerbrat-Orecchioni [KER 94]. (The list is necessarily selective.) They address the issue of language use and verbal interaction from the viewpoint of sociolinguistics and linguistic pragmatics. Their approach is inspired from the work of Goffman, in that apparently they all assume that remedial work has some role to play in any attempt to characterise interaction between agents. Such an approach has been criticized, on the ground that it exaggerates the aggressive character of face-to-face interaction, be it verbal or non-verbal. This is not germane for my present purpose.

At this point, it is interesting to consult the literature on deontic logic. In general these authors describe themselves as aiming to clarify the "logic" of conversational (face-to-face) interactions. For instance, Owen writes:

> "A broader aim in [my] work is to demonstrate that everyday conversation is not disordered, rambling, and 'casual' [...] but ordered, coherent and well-suited for the achieving of interactional goals". [OWE 83, p. 1]

Such a claim sets a problem, which is usually taken to be a hard one for deontic logicans. It is generally referred to as Chisholm's paradox,[2] and can be illustrated with a simple example:

EXAMPLE 1. — (Goffman [GOF 71, p. 140])

Deed:	A trips over B
A:	"Sorry"
B:	"S'okay"

The following sentences all appear to be true:

a) $\bigcirc\neg o$

b) $\bigcirc(r_1/o)$

c) $\Box(r_1 \rightarrow o)$

d) o

Here $\bigcirc(\psi/\phi)$ (resp. $\bigcirc\psi$) stands for the obligation of ψ given ϕ (resp. the obligation of ψ). The propositional letters o and r_1 can be read as (respectively) "offence" and "remedy". The $\Box$-operator expresses necessity of type S5. The sentence $\Box(r_1 \rightarrow o)$ (premiss c)) is read "r_1 presupposes o". The import of c) is essentially to rule out the case where the offender apologizes for an act that did not in fact occur. Intuitively, this amounts to restricting one's attention to the (not infrequently realised) case where the action observed is the same as the action performed. The kind of scenario Goffman is interested in is similar in pattern to the "gentle murderer" scenario (Forrester [FOR 84]) and to the "white fence" scenario (Prakken and Sergot [PRA 96]). The first one contains the following data: Smith ought not to kill his mother; if Smith kills his mother, he ought to kill her gently; gentle killing necessarily implies killing; Smith kills his mother. The second one contains the following data: there must be no fence; if there is a fence, it must be white; being a white fence necessarily implies being a fence; there is a fence.

2. See Chisholm [CHI 63].

Now, consider a deontic logic that validates the following two principles, known as *Weakening the Consequent* (WC) and *Factual Detachment* (FD), respectively:

$$\big(\bigcirc(\psi/\phi) \wedge \Box(\psi \rightarrow \psi')\big) \rightarrow \bigcirc(\psi'/\phi) \qquad \text{(WC)}$$

$$\big(\bigcirc(\psi/\phi) \wedge \phi\big) \rightarrow \bigcirc\psi \qquad \text{(FD)}$$

As is well-known from the literature on the logic of norms, there is an undesirable interaction between these two rules. On the one hand, the obligation

$$\bigcirc r_1$$

can be derived from b) and d) by using (FD). On the other hand, c) entails $\Box(\neg o \rightarrow \neg r_1)$. From this together with a), it immediately follows by (WC) that

$$\bigcirc\neg r_1$$

Some may not be satisfied by this result, since intuitively a)-d) are perfectly consistent. In its traditional form, Chisholm's paradox highlights the problem of the relation between two detachment principles, the "Deontic Detachment" of $\bigcirc\neg\psi$ from $\Gamma_1 = \{\bigcirc\neg\phi, \bigcirc(\neg\psi/\neg\phi)\}$ and the "Factual Detachment" of $\bigcirc\psi$ from $\Gamma_2 = \{\bigcirc(\psi/\phi), \phi\}$.[3] Once again, the kind of contrary-to-duty (CTD) situation Goffman seems to have in mind is in fact close to the white fence/gentle murderer scenarios. These are obtained by substituting, in Chisholm's initial example, $\Box(\psi \rightarrow \phi)$ for $\bigcirc(\neg\psi/\neg\phi)$.

The problems raised by CTDs have many facets. In this article I choose to focus on one of them only, namely the problem caused by the interplay between the two rules FD and WC. One might refer to this as (to put it in Åqvist's terms) "the dilemma of Weakening the Consequent and Factual Detachment".[4] In section 4, I shall introduce and explore a way of getting out of this dilemma, by using tools originally developed in the context of the study of iterated belief change. The intuition that guides and motivates such an approach is better explained by analogy with earlier proposals based on temporal deontic logic. As we shall see in section 3, a number of contributors to the deontic literature suggested that the obligation to do $r1$ and the obligation not to do $r1$ both hold, but not at the same time point. In some ways, the

3. In standard deontic logic – a normal modal logic of type KD, a further complication arises from the fact that conditional norms can be represented in two ways, depending upon whether or not "$\bigcirc$" precedes "$\rightarrow$". In his 1963 paper, Chisholm proposed to represent $\bigcirc(\neg\psi/\neg\phi)$ and $\bigcirc(\psi/\phi)$ by $\bigcirc(\neg\phi \rightarrow \neg\psi)$ and $\phi \rightarrow \bigcirc\psi$, respectively. Both kinds of detachment are then validated, causing the above tension. If the formalization is modified in accordance with the above suggestion, the consistency of $\Gamma_1 \cup \Gamma_2$ is restored, but then another minimal requirement is not satisfied: the members of $\Gamma_1 \cup \Gamma_2$ must be analysed in such a way that none is a logical consequence of the remaining premisses. (See [ÅQV 67, CAR 02].)

4. See Åqvist [ÅQV 02]. This one uses "Deontic Detachment" where I use "Weakening the Consequent", for the reason just explained.

account in terms of iterated revision is an attempt to give a new guise to this idea,[5] which I will do by adding to the example the three further moves that, according to Goffman [GOF 71, p. 140-144], usually follow the remedy. These are:

– *relief*, by which the victim provides a sign that the remedy offered by the offender is sufficient;

– *appreciation*, by which the offender shows gratitude and thankfulness;

– *minimization*, by which the victim repeats in diminished form the relief he provided as the second move.

The full sequence might be illustrated by the following example:

EXAMPLE 2. — (Goffman [GOF 71, p. 143])

remedy	A:	Can I use your your phone to make a local call?
relief	B:	Sure, go ahead
appreciation	A:	That's very good of you
minimization	B:	It's okay

Here is an attempt to formalize the rules on which this four-move interchange is based. One might imagine that A makes his request at the same time that he lifts the receiver. In this case, A violates his primary obligation not to take another's property. This first obligation can be rendered as

$$\bigcirc \neg o$$

For the purposes of the present discussion, it is not necessary to specify what kind of territorial offence the letter o describes. At the same time that he violates his obligation not to do o, A tries to convert his offensive act into an acceptable one by providing a remedy of the "request" type. It does not seem unreasonable to assume that A provides such a remedy in accordance with a contrary-to-duty norm having the form (r_1 stands for the remedy move):

$$\bigcirc(r_1/o)$$

In the subsequent moves of the exchange (relief, appreciation and minimization), there is a shift in concern from the issue that a norm was violated to a focus on the way the participants handle their management of infractions. A similar formal treatment might be devised for these subsequent moves. For instance it does not seem unreasonable to believe that B answers "Sure, go ahead" in accordance with an according-to-duty (ATD) obligation having the form

$$\bigcirc(r_2/o \wedge r_1)$$

5. I say "in some ways", because this way of presenting the proposed account is not entirely correct, though convenient. (See my concluding remarks in subsection 4.4.)

where r_2 describes any kind of utterances that can be assimilated to a move of the type "relief". And so forth. Goffman [GOF 71, p. 141] suggests that one participant's making such-and-such move places the other under some obligation to make the next one. If this view is correct, then the use of a deontic operator to formalize the fully expanded sequence does make sense. Table 1 recapitulates the sentences that appear to be true before the interaction above has even occurred.

Table 1. *The remedial cycle (full sequence)*

	Normative premisses (α)	Integrity constraints (β)
(I)	$\bigcirc \neg o$	
(II)	$\bigcirc(r_1/o)$	$\Box(r_1 \rightarrow o)$
	$\bigcirc(r_2/o \wedge r_1)$	$\Box(r_2 \rightarrow (o \wedge r_1))$
(III)	$\bigcirc(a/o \wedge r_1 \wedge r_2)$	$\Box(a \rightarrow (o \wedge r_1 \wedge r_2))$
	$\bigcirc(m/o \wedge r_1 \wedge r_2 \wedge a)$	$\Box(m \rightarrow (o \wedge r_1 \wedge r_2 \wedge a))$

Row (II) lists the obligations on which the first round, remedy (r_1) and relief (r_2), is based. Row (III) mentions those producing the second round, appreciation (a) and minimization (m). A CTD obligation is attached to a primary obligation, and three ATD obligations are in turn associated with the CTD obligation. For lack of a better term, I use the label "integrity constraints" to cover the sentences mentioned in column (β). All these are of the form $\Box(\phi \rightarrow \psi)$. This is read "$\phi$ presupposes ψ". As already suggested, the import of these non-normative premisses is essentially to rule out some sequences of moves that seem to be unrealistic (notwithstanding the admitted fact that they are logically possible).

This completes the description of the fully expanded remedial cycle. In section 4 I will outline a possible analysis of this example, by using tools from iterated belief change theory. Since this account owes some inspiration to so-called temporal deontic approaches to the representation of CTD scenarios, it will be useful to look at them first. The next section is an attempt to discuss this kind of treatment.

3. Using correctly temporalized propositions

In the seventies and early eighties, various systems of temporal deontic logic were developed to give a consistent representation of CTD scenarios. In this section I will focus on the system DARB, first presented in Åqvist & Hoepelman [ÅQV 81] and further developed in Åqvist [ÅQV 91a]. It is beyond the scope of this paper to give a detailed account of this system. Rather, I will pick out those aspects of DARB that are relevant for the comparison with the account to be introduced in section 4.

The object language of DARB contains, in addition to the deontic operators alluded to above, the unary operator $\oplus$, to be read as "it will be the case at the next instant that". DARB has the following axiom schemata and inference rule for $\oplus$:

$$\oplus(\phi \rightarrow \psi) \rightarrow (\oplus\phi \rightarrow \oplus\psi) \qquad \text{(a1)}$$

$$\oplus\phi \leftrightarrow \neg \oplus \neg\phi \qquad \text{(a2)}$$

$$\text{If } \phi \text{ is a theorem then so is } \oplus\phi \qquad \text{(Nec)}$$

DARB also contains a modal operator of historical necessity $\Box$, to be read as "it is necessary on the basis of the past and the present that". $\Box$ is a modality of the S5 type, which also obeys the law: $\phi \rightarrow \Box\phi$ provided that ϕ is a propositional letter. An immediate consequence of the latter proviso is that the substitution rule fails in general. The reason why propositional letters are treated in this way has to do with the intended meaning of the $\Box$ operator. This one ranges over the set of histories that are still open, given the past and the present (which are fixed).

As far as the representation of CTD scenarios is concerned, the core of Åqvist's proposal is to distinguish between what is obligatory before the transgression and what is obligatory at the time of the violation. To understand how the solution works, it is not necessary to consider the three further moves that, according to Goffman, usually follow the remedy. For clarity's sake, in this section I will thus concentrate on the treatment given to the pair (offense, remedy). The proposed analysis can easily be extended to the full sequence.

Assuming that o and r_1 are executed in parallel, on might obtain a reasonable description of the situation before the transgression *via* the following quartet of DARB-formulae:

$$\bigcirc \oplus\neg o \qquad \text{(Ia)}$$

$$\Box(\oplus r_1 \rightarrow \oplus o) \qquad \text{(IIa)}$$

$$\oplus \bigcirc(r_1/o) \qquad \text{(IIIa)}$$

$$\oplus\, o \qquad \text{(IVa)}$$

It is easy to see that in DARB (IIa) entails $\Box(\oplus\neg o \rightarrow \oplus\neg r_1)$.[6] From this together with (Ia), it immediately follows by *Weakening the Consequent* (WC) that $\bigcirc \oplus \neg r_1$.

6. This might be proved by showing that (1), (2), (3) and (4) each are provable in DARB:

(1) $(\oplus r_1 \rightarrow \oplus o) \rightarrow (\neg \oplus o \rightarrow \neg \oplus r_1)$ contraposition

(2) $(\oplus r_1 \rightarrow \oplus o) \rightarrow (\oplus\neg o \rightarrow \oplus\neg r_1)$ $1, a2$

(3) $\Box\big((\oplus r_1 \rightarrow \oplus o) \rightarrow (\oplus\neg o \rightarrow \oplus\neg r_1)\big)$ $2,$ necessitation rule for $\Box$

(4) $\Box(\oplus r_1 \rightarrow \oplus o) \rightarrow \Box(\oplus\neg o \rightarrow \oplus\neg r_1)$ $3,$ rule RM for $\Box$ (see [CHE 80, p. 114])

This formula can be read as "it shall (now) be that it will not be the case at the next moment that r_1". On the other hand, from (IIIa) together with (IVa), it follows by *Factual Detachment* (FD) that $\oplus \bigcirc r_1$. This sentence can be read as "it will be the case at the next instant that it shall be the case that r_1". But care should be taken here. As Åqvist [ÅQV 91a, ÅQV 91b] emphasizes, in DARB the Factual Detachment principle does not hold unrestrictively. Using Åqvist's terminology, one might refer to the premiss $\bigcirc(\psi/\phi)$ as the normative major, and to the factual premiss ϕ as the boolean minor. The proviso placed on (FD) says that

> "the time-of-realization of the boolean minor [should] coincide[. . .] with or [be] earlier than [. . .] the time of being-in-force of the normative major". [ÅQV 91b, p. 137]

The meaning of such a proviso is likely to be opaque at first reading, and I shall briefly explain in a moment how I understand it. For present purposes, it is sufficient to observe that the above restriction is clearly satisfied in the case of:

(1) $(\bigcirc(r_1/o) \wedge o) \rightarrow \bigcirc r_1$

It immediately follows that (2) and (3) each are DARB-provable:

(2) $\oplus\big((\bigcirc(r_1/o) \wedge o) \rightarrow \bigcirc r_1\big)$ 1, necessitation rule for $\oplus$

(3) $(\oplus \bigcirc (r_1/o) \wedge \oplus o) \rightarrow \oplus \bigcirc r_1$ 2, rule RR for $\oplus$ (see [CHE 80, p. 114])

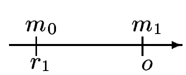
Thus, while *Weakening the Consequent* yields $\bigcirc \oplus \neg r_1$, *Factual Detachment* gives rise to $\oplus \bigcirc r_1$. Of the two competing obligations, none contradicts the other, since they do not hold at the same time point.

To sum up, the basic idea that guides proposals based on temporal deontic logic such as DARB consists in post-dating the obligatoriness of the remedy. More specifically, it consists in assuming that the time-of-being-in-force of the reparational obligation coincides with the time-of-realization of the violation. At first sight this way of handling the paradox seems to be very natural. Interactants will not consider themselves as being unconditionnally obliged to make the remedy move unless the transgression has occurred. To say this amounts, more or less, to saying that the truth value of any sentence must be made relative to the available knowledge about what the future will be. Hence the idea of evaluating formulas on pairs m/h, where m is a moment and h a history passing through m. "ϕ is true at the pair m/h" can be understood to mean the same as "ϕ is true at m under the hypothesis that h is (other things being equal) the future of m". Now the special requirement placed on the Factual Detachment principle becomes clearer. The inference from $\{\bigcirc(\psi/\phi), \phi\}$ to $\bigcirc\psi$ fails when the time-of-realization (= m_1) of the boolean minor ϕ lies in the future of the time (= m_0) at which the normative major $\bigcirc(\psi/\phi)$ holds, because between m_0 and m_1 some "new" fact not historically or logically fixed at m_0 may occur and give rise to an overriding unconditional obligation.

The question of whether temporal deontic logic provides a satisfactory analysis of CTD scenarios has been much discussed within the literature. Among the issues that such a proposal has raised, the following two deserve special mention.

First of all, it is easy to see that (IVa) entails $\oplus \bigcirc o$, i.e., after o is done, the primary interdiction to do o is transformed into an obligation to do o. This might be verified as follows. As we saw previously, we have the law $\phi \rightarrow \Box\phi$ provided that ϕ is a propositional letter. On the other hand, we also have the law $\Box\phi \rightarrow \bigcirc\phi$.[7] Putting the two together, we get $\phi \rightarrow \bigcirc\phi$ whenever ϕ is a propositional letter, so that in particular $o \rightarrow \bigcirc o$. By the logic of $\oplus$, we finally get that $\oplus o \rightarrow \oplus \bigcirc o$ is DARB-provable. It is worth mentioning that a number of deontic logicians do not find this result unwelcome. For one might interpret it to mean "o, which is settled as true, is also *vacuously* obligatory". This is not to say that o should be reiterated, since $\oplus \bigcirc o$ is consistent with $\oplus \bigcirc \oplus\neg o$. Neither is it to say that (be it settled as true or not) o should not have been done. As observed by, e.g., Thomason [THO 81, p. 173], one must distinguish carefully between two uses of "ought", a use appropriate for deliberating and advising (deliberative "ought") and a use appropriate for passing judgement (judgemental "ought"). A framework such as DARB is intended to give an account of the first one only. For instance, the monadic "ought" is interpreted as follows. Some of the possible futures of h at m are marked as the most perfect ones; what holds in all of them are the obligations at m. Those histories that branched off in the past have become inaccessible, and are thus not considered. The dyadic obligation operator is interpreted in much the same way. In determining whether ψ is obligatory given ϕ, one look at whether ψ holds throughout all the best futures where ϕ holds. In both cases, the modality is forward-looking, rather than backward-looking.

Second, it is natural to ask what happens if o and r_1 are not done in parallel. First of all, there is the case where o precedes r_1 in time. As far as I can see, this scenario raises no problems. Next, there is the case where r_1 precedes o temporally. This leads to focus attention on this specific form of remedial work that Goffman calls "request". It consists of asking permission from a potentially offended person to engage in what could be considered as a violation of his rights. For example, requests to borrow goods convert the offensive taking of another's property into an acceptable act. Goffman observes that (as opposed to apologies) "requests [. . .] typically occur before the questionable event or, at the latest, during its initial phases" [GOF 71, p. 114]. This, because the request is orientated to the offence that *would* take place *if* the request was not made. It is unclear how (as it stands) the system DARB can handle this class of moves.[8] To see why, assume that the time-of-realization of r_1 and o is m_0 and m_1, respectively:



7. This is Åqvist and Hoepelman's Th 21, cf. [ÅQV 81, p. 213].
8. A similar point is made in [TOR 98, HAN 99], though the details are different. These authors consider Chisholm's paradox in its original version (see [CHI 63]), where the antecedent of the conditional obligations happens after the consequent. The former says that a certain man goes to the assistance of his neighbours, and the latter that he tells them he is coming.

The objective here is to obtain a description of the situation at the time point m_0. For this to be done, it is necessary to amend in one way or another the quartet (Ia)-(IVa). As it seems, neither (Ia) nor (IVa) needs to be altered. As for (IIa), it is enough to leave out the first $\oplus$. Hence:

$$\bigcirc \oplus \neg o \qquad \text{(Ib)}$$

$$\Box(r_1 \rightarrow \oplus o) \qquad \text{(IIb)}$$

$$\ldots \qquad \text{(IIIb)}$$

$$\oplus\, o \qquad \text{(IVb)}$$

It is easy to show that (IIb) entails $\Box(\oplus\neg o \rightarrow \neg r_1)$. So, from the first two premisses, the desired $\bigcirc\neg r_1$ can still validly be inferred, by using Weakening the Consequent. But it does not seem possible to reformulate the remaining premiss in such a way that some meaningful conclusion might be inferred from the couple (IIIb)-(IVb), by using the restricted form of Factual Detachment alluded to above. Consider the proviso this rule contains. It says that the time-of-realization of the violation must coincide with or be earlier than the time-of-being-in-force of the reparational obligation. Premiss (IVb) already says that the violation occurs at m_1. This means that the obligation must come into force at m_1, or perhaps later. Suppose it is m_1 (the argument would be similar if the norm came into force at a later time). Then the correct rendering is

$$\oplus \bigcirc(.../o) \qquad \text{(IIIb)}$$

It remains to specify the consequent in (IIIb). The trouble is that the time-of-realization of r_1 is supposed to be $m_0 < m_1$. One has no alternative but to use the past operator employed in DARB

$$\ominus = \text{ it was the case at the last instant that}$$

(IIIb) becomes

$$\oplus \bigcirc(\ominus r_1/o) \qquad \text{(IIIb*)}$$

(IIIb*) makes little sense, since it makes an act that belongs to the past obligatory. Can the past be undone?

The account in terms of iterated revision to which I now turn is an attempt to solve this second problem. However, the proposed construction will face a difficulty that is very similar to the first one. One possible way of refining the construction will be proposed, and briefly discussed.

4. An analysis in terms of commutation

The basic idea is to assume that a fact (in particular, a transgression) might affect the ideality ordering in the semantics. A similar intuition guides several constructions used in the study of iterated belief revision [SPO 88, DAR 94, WIL 94, LEH 95, BOU 96, FER 02, NAY 03]. For reasons to be made clear later, it is the notion of natural revision [BOU 96] that will be used here. As we shall see, its principal feature is that, when contrary-to-duty information is taken into account, the set of "second best" worlds commutes with the set of "best" worlds. I have said that this approach in terms of commutation owes some inspiration to the so-called preference-based temporal deontic solution. At the end of the present inquiry, it will appear that the two accounts are clearly distinct. Various other strategies for giving a consistent representation to CTD scenarios have been proposed in the literature on deontic logic. It falls outside the scope of this article to review them. (See, for example, [CAR 02] for an overview). It has often been suggested that CTD scenarios can be represented consistently if a distinction is made between different sorts of "oughts". Such a suggestion has taken various forms [ÅQV 67, JON 85, CAR 95, CAR 02], sometimes framed in preference-based terms [ART 96, TOR 99, CHO 01]. A thorough investigation of the relation between the latter approach and the one outlined in this paper remains to be done.

Note that, by switching to so-called iterated revision, I implicitly opt for a methodology that is very different from the one propounded by the authors of the AGM 1985 paradigm. The construction I will introduce in this section will cover the successive moves of the remedial cycle simultaneously. To put it another way, the construction will cover the "flat" or "first-degree" (non-iterated) case and the iterated case at once. This contrasts with the more analytical approach that is followed in AGM 1985. Here the exploration of the iterated case comes *after* the study of the flat case. It is beyond the scope of this article to discuss the merits or demerits of these two methods (for a more detailed discussion, see Makinson [MAK 03]). Noticeably, those sympathetic to the analytical method in general believe that, if one can deal with the simpler non-iterated case in a principled way (as needs to be done anyway), then it should be possible to apply the same procedures to the more complex iterated one. This is not always an easy task. A construction meant for the non-iterated case might not be suitable for the iterated one. Consider the case of the theory of Carmo and Jones [CAR 95]. They defend the view that CTD scenarios can be represented consistently if a distinction is made between ideal and actual obligations. In Carmo and Jones [CAR 02], it is shown that serious difficulties arise as soon as further levels of CTDs are introduced. The latter introduction will be assimilated below with the fact that iteration comes into the picture.

This section is organized as follows. The treatment presupposes some familiarity with the Hansson-Lewis semantics for deontic logic. First, in order to clear the ground, I recall the main ingredients of such a semantics (subsection 4.1). Second, I outline some of the ways the semantics could be refined if it is to deal with the analysis of a one-move interchange (subsection 4.2). Third, in order to complete the account, I consider the fully expanded remedial cycle (subsection 4.3). Fourth, I incorporate the construction into a branching-time logic, as an attempt to deal with the request-response more sensitively (subsection 4.4).

4.1. *Background*

Most of the definitions that I will use are fairly straightforward. The only new feature concerns the "dynamic" component that iterated revision theories aim to incorporate into preference-based semantics. However, to make the exposition clearer, I postpone discussion of this component until the next subsection.

By a model, I shall mean, as usual, any ordered triplet $\mathcal{M} = (W, \preceq, \iota)$ where

1) W is a set of possible worlds $w, w', \ldots$

2) $\preceq \subseteq W \times W$ is a pre-order, i.e. it is a reflexive and transitive relation on W; I take $w \preceq w'$ to mean that w is at least as preferred as w' (or w is at least as good as w')

3) ι is a function which assigns to each propositional letter p in Prop a subset of W.

Let $[\phi]_{\mathcal{M}}$ denote the truth-set of ϕ in $\mathcal{M}$, i.e. the set of worlds in $\mathcal{M}$ at which ϕ is true. Let $\min_{\mathcal{M}}(\phi)$ denote the set of $\preceq$-minima in $[\phi]_{\mathcal{M}}$, i.e.

$$\min_{\mathcal{M}}(\phi) \;=\; \{w \in [\phi]_{\mathcal{M}} \mid \forall w'(w' \in [\phi]_{\mathcal{M}} \;\rightarrow\; w \preceq w')\} \qquad \text{(Def-min)}$$

Intuitively $\min_{\mathcal{M}}(\phi)$ is understood to be the set of most perfect ϕ-worlds. For such a set not to be empty, two conditions must be fulfilled. One is that ϕ is satisfiable in $\mathcal{M}$, i.e. there is some world w in $\mathcal{M}$ at which ϕ is true. The other is that there are no infinite sequences of ever more perfect ϕ-worlds. The truth-conditions for the dyadic obligation operator are defined as follows:

$$\mathcal{M} \models \bigcirc(\psi/\phi) \Leftrightarrow \min_{\mathcal{M}}(\phi) \subseteq [\psi]_{\mathcal{M}}. \qquad \text{(Def}\bigcirc\text{)}$$

According to (Def$\bigcirc$), model $\mathcal{M}$ satisfies the obligation "ψ should be the case if ϕ is the case", $\mathcal{M} \models \bigcirc(\psi/\phi)$, whenever in the most perfect worlds, where ϕ holds, ψ holds too.[9] For the sake of simplicity, $\bigcirc\phi$ is taken to be equivalent to $\bigcirc(\phi/\top)$ – where $\top$ is a tautology.

4.2. *One-move interchange*

This subsection shows how to refine the Hansson-Lewis semantics so as to deal with the analysis of a one-move interchange.

The first step involves constructing a typical model, say $\mathcal{M}_1$, of the initial premisses set as given in table 1. Given the subset (β) of integrity constraints, the follow-

9. Throughout this paper I work with minimality under $\preceq$, as is the custom in non-monotonic logic. Deontic logicians usually work with maximality under the converse relation. This is more in line with the idea that, the more perfect a world is, the farther from actuality it is. From a mathematical point of view, the two idioms are equivalent.

ing universe W_1 is complete in the sense of giving an exhaustive characterisation of how the interaction might evolve:

$$w_1 : \neg o, \neg r_1, \neg r_2, \neg a, \neg m \qquad w_2 : o, r_1, r_2, a, m$$
$$w_3 : o, \neg r_1, \neg r_2, \neg a, \neg m \qquad w_4 : o, r_1, \neg r_2, \neg a, \neg m$$
$$w_5 : o, r_1, r_2, \neg a, \neg m \qquad w_6 : o, r_1, r_2, a, \neg m.$$

In w_1 the primary obligation is fulfilled. In w_2 this obligation is violated, but the other four obligations are fulfilled. In $\{w_3, \ldots, w_6\}$, the primary obligation is violated, and one of the other four obligations is also violated. One might adopt the hierarchical ordering as depicted in Figure 1 below. Here the convention is that all points on a

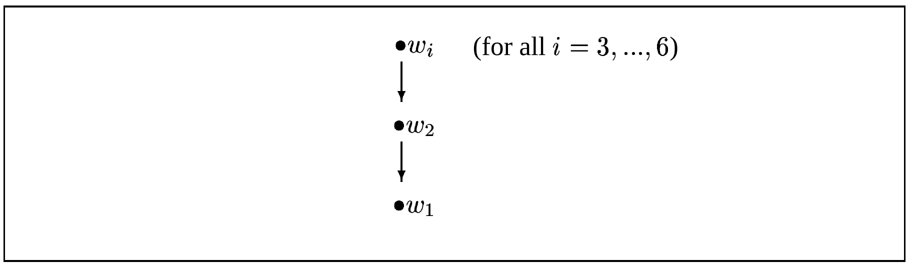


Figure 1. *Before the offence (model $\mathcal{M}_1$)*

lower level are more perfect than all those on a higher level. The arrow $w\bullet \leftarrow \bullet w'$ denotes the relation $w \prec w'$, where $\prec$ is for the strict partial order determined by $\preceq$, defined by putting $w \prec w'$ whenever $w \preceq w'$ and $w' \not\preceq w$. One might take $w \prec w'$ to mean that w is strictly better than w'. It can easily be verified that $\prec$ thus defined is irreflexive and transitive. $\mathcal{M}_1$ satisfies the primary obligation $\bigcirc\neg o$, since all the worlds gravitate towards w_1 (the best world). $\mathcal{M}_1$ also satisfies the remaining normative premisses, since w_2 (the 2nd best world) is better than w_3, w_4, w_5 and w_6. From a formal point of view, it is not necessary to refine further the ordering within the set $\{w_3, \ldots, w_6\}$. For simplicity's sake, it is assumed that (within this set) neither world is preferred over the other. Some may not be satisfied with such an assumption. In particular Prakken and Sergot (rightly) argue that "it is better to fulfil an obligation from a more ideal context and violate one from a less ideal context than the other way around" [PRA 97b, p. 152]. It would thus be more plausible to put $w_6 \prec w_5 \prec w_4 \prec w_3$. The point is that, as the authors observe, in itself the semantics of Hansson-Lewis does not require this, but must be augmented in some way or other. The strategy Prakken and Sergot opt for consists in using the lexicographic rule of Ryan [RYA 92]. The latter rule will not be used in what follows, as I do not wish to distract the reader from the main point of this paper.

Intuitively, model $\mathcal{M}_1$ depicts the pre-order *before the offence has occurred*. All the worlds gravitate towards w_1, where r_1 (remedy) is false. At this very moment, we thus have

- $\mathcal{M}_1 \models \bigcirc\neg r_1$

rather than

- $\mathcal{M}_1 \models \bigcirc r_1$

A similar remark applies to moves r_2 (relief), a (appreciation) and m (minimization). We have

- $\mathcal{M}_1 \models \bigcirc\neg r_2$
- $\mathcal{M}_1 \models \bigcirc\neg a$
- $\mathcal{M}_1 \models \bigcirc\neg m$

rather than

- $\mathcal{M}_1 \models \bigcirc r_2$
- $\mathcal{M}_1 \models \bigcirc a$
- $\mathcal{M}_1 \models \bigcirc m$

This is reminiscent of the step where in section 3 the norm $\bigcirc \oplus \neg r_1$ was inferred from (Ia) and (IIa).

It is now possible to bring the territorial offence into the picture. The idea is to assume that introducing a violation has a specific impact on the ideality ordering associated with the premisses set. The hypothesis here is twofold:

a) incorporating contrary-to-duty information, those taking part in the face-to-face interaction accommodate that ordering so as to have the appropriate unconditional obligation "detached";

b) in order to bring the remedial interchange to completion, they keep as much of the initial ordering as possible.

These assumptions can be expressed more rigorously as follows.

DEFINITION 3. — *Let $\mathcal{M} = \langle W, \preceq, \iota\rangle$ be the model reflecting the initial ordering. Given the contrary-to-duty context ϕ, the model obtained by revising $\mathcal{M}$ by ϕ is $\mathcal{M}^{\star}\phi = \langle W, \preceq', \iota\rangle$, with $\preceq'$ defined as follows:*

$(\mathbf{P}_1)$ *If $w_1 \in min_{\mathcal{M}}(\phi)$ then:*
(a) *$w_1 \preceq' w_2$ for all $w_2 \in W$ and*
(b) *$w_2 \preceq' w_1$ iff $w_2 \in min_{\mathcal{M}}(\phi)$;*

$(\mathbf{P}_2)$ *If $w_1, w_2 \notin min_{\mathcal{M}}(\phi)$ then: $w_1 \preceq' w_2$ iff $w_1 \preceq w_2$.*

Here I draw on the treatment by Boutilier [BOU 96] of iterated revision. His term for it is "natural revision". $(\mathbf{P}_1)$ aims to capture a). Roughly, $(\mathbf{P}_1)$ says that every minimal ϕ-world is shifted down to the bottom of the new ordering. $(\mathbf{P}_2)$ aims to capture b). Literally, $(\mathbf{P}_2)$ says that for w_1 and w_2 not in $\min_{\mathcal{M}}(\phi)$, then $\preceq'$ and $\preceq$ are the same. This has the effect of leaving $\preceq$ unaltered except as indispensably required by $(\mathbf{P}_1)$.

Intuitively, $(\mathbf{P}_1)$ means that $\min_{\mathcal{M}}(\phi)$ forms the set of "newly" best worlds. Expressed formally, this gloss amounts to saying that $\min_{\mathcal{M}}(\phi) = \min_{\mathcal{M}\star\phi}(W)$. The inclusion $\min_{\mathcal{M}}(\phi) \subseteq \min_{\mathcal{M}\star\phi}(W)$ follows from $(\mathbf{P}_1)$(a) alone. For the converse inclusion, we need $(\mathbf{P}_1)$(b) plus two further assumptions. One is that ϕ is satisfiable in $\mathcal{M}$. The other is that the ideality ordering in $\mathcal{M}$ satisfies a condition known as *stoppering* or *smoothness*. This says that whenever $w \in [\phi]_{\mathcal{M}}$, then either $w \in \min_{\mathcal{M}}(\phi)$ or there is a $w' \preceq w$ with $w' \in \min_{\mathcal{M}}(\phi)$.[10] The reason why these two extra assumptions are needed is better explained by a contrapositive argument. Assume $w \notin \min_{\mathcal{M}}(\phi)$. We wish to show that $w \notin \min_{\mathcal{M}\star\phi}(W)$. Since ϕ is satisfiable in $\mathcal{M}$, there is some $w' \in [\phi]_{\mathcal{M}}$. By smoothness, either $w' \in \min_{\mathcal{M}}(\phi)$ or there is a $w'' \preceq w'$ with $w'' \in \min_{\mathcal{M}}(\phi)$. By $(\mathbf{P}_1)$(b), we conclude that either $w \not\preceq' w'$ or $w \not\preceq' w''$. In both cases, (Def-min) guarantees that $w \notin \min_{\mathcal{M}\star\phi}(W)$ as desired.

I end this presentation with the following observation, which answers a question left open in [BOU 96]:

OBSERVATION 4. — The posterior ordering $\preceq'$ thus defined is reflexive and transitive. Moreover, if $\preceq$ satisfies the smoothness condition, then so does $\preceq'$.

PROOF. — That $\preceq'$ is reflexive follows immediately from the reflexivity of $\preceq$. So let us concentrate on transitivity. Assume that $w_1 \preceq' w_2$ and $w_2 \preceq' w_3$. We wish to show that $w_1 \preceq' w_3$. If $w_1 \in \min_{\mathcal{M}}(\phi)$, then a direct application of $(\mathbf{P}_1)$(a) yields the expected conclusion. So, assume that w_1 is not in $\min_{\mathcal{M}}(\phi)$. If w_2 was in $\min_{\mathcal{M}}(\phi)$, then by $(\mathbf{P}_1)$(b) we would get $w_1 \not\preceq' w_2$, contradicting the first assumption. Thus, w_2 is not in $\min_{\mathcal{M}}(\phi)$. Now, by $(\mathbf{P}_2)$ together with $w_1 \preceq' w_2$, we get $w_1 \preceq w_2$. Similarly, if w_3 was in $\min_{\mathcal{M}}(\phi)$, then by $(\mathbf{P}_1)$(b) we would have $w_2 \not\preceq' w_3$, contradicting the second assumption. Therefore, w_3 cannot be in $\min_{\mathcal{M}}(\phi)$. By $(\mathbf{P}_2)$ together with $w_2 \preceq' w_3$, we get $w_2 \preceq w_3$ and, as a result, $w_1 \preceq w_3$, since $\preceq$ is transitive. By $(\mathbf{P}_2)$, we finally get $w_1 \preceq' w_3$ as required.

As for the smoothness condition, one might argue contrapositively. Assume $\preceq'$ is not stoppered. We need to show that $\preceq$ is not stoppered either. Let w_1 and ψ be such that

I) $w_1 \in [\psi]_{\mathcal{M}\star\phi}$,

II) $w_1 \notin \min_{\mathcal{M}\star\phi}(\psi)$,

III) for all $w \preceq' w_1$ we have $w \notin \min_{\mathcal{M}\star\phi}(\psi)$.

Clearly, $w_1 \in [\psi]_{\mathcal{M}}$. We first argue that $w_1 \notin \min_{\mathcal{M}}(\psi)$. Since w_1 is a ψ-world, hypothesis (II) means that $w_1 \not\preceq' w_2$ for some ψ-world w_2. By $(\mathbf{P}_1)$(a), we conclude that $w_1 \notin \min_{\mathcal{M}}(\phi)$. Here we only need to show that $w_2 \notin \min_{\mathcal{M}}(\phi)$, since by $(\mathbf{P}_2)$ we immediately get $w_1 \not\preceq w_2$ and, thus, $w_1 \notin \min_{\mathcal{M}}(\psi)$ as expected. We argue *ad absurdum*. Let $w_2 \in \min_{\mathcal{M}}(\phi)$. By $(\mathbf{P}_1)$(a), we get $w_2 \preceq' w_1$. By hypothesis (III), $w_2 \notin \min_{\mathcal{M}\star\phi}(\psi)$. Given that w_2 is a ψ-world, this means that $w_2 \not\preceq' w_3$ for some

10. Such a condition essentially forbids infinite sequences of ever more perfect ϕ-worlds.

ψ-world w_3. Since $w_2 \in \min_{\mathcal{M}}(\phi)$, by $(\mathbf{P}_1)$(a) we also have $w_2 \preceq' w_3$. The expected contradiction has been reached.

Now, we need to show that for all $w \preceq w_1$ we have $w \notin \min_{\mathcal{M}}(\psi)$. Assume $w \preceq w_1$. If w is not a ψ-world, then the proof is completed. So let ψ be true in w. If $w \in \min_{\mathcal{M}}(\phi)$, then by $(\mathbf{P}_1)$(a) we get $w \preceq' w_1$. If $w \notin \min_{\mathcal{M}}(\phi)$, then by $(\mathbf{P}_2)$ we also get $w \preceq' w_1$, since we already knows that $w_1 \notin \min_{\mathcal{M}}(\phi)$. So, in both cases, $w \preceq' w_1$. From this together with hypothesis (III), we get $w \notin \min_{\mathcal{M}^\star\phi}(\psi)$. As w is a ψ-world, we conclude that there exists another ψ-world w_2 such that $w \not\preceq' w_2$. We are almost ready, if we can prove that $w \not\preceq w_2$. And, to prove this inequality, we only need to show that neither w nor w_2 is in $\min_{\mathcal{M}}(\phi)$. As for w, it suffices to invoke $(\mathbf{P}_1)$(a) and $w \not\preceq' w_2$. As for w_2, we argue as follows. Suppose $w_2 \in \min_{\mathcal{M}}(\phi)$. In the presence of $(\mathbf{P}_1)$(a), this assumption (which we want to reduce *ad absurdum*) implies that $w_2 \preceq' w'$ for all w' in W. In particular, $w_2 \preceq' w_1$. But, by hypothesis (III), we get $w_2 \not\preceq' w_3$ for some ψ-world w_3. This obviously contradicts the fact that $w_2 \preceq' w'$ for all w' in W.

This completes the proof. ■

It is now possible to apply definition 3 to the analysis of the notion of remedial interchange. To see how $(\mathbf{P}_1)$ and $(\mathbf{P}_2)$ work, and how they interact, let us apply them to the original model $\mathcal{M}_1$ (see figure 1). Consider the situation where the individual offends, viz. take ϕ to be o (offence). Figure 2 below depicts the resulting model $\mathcal{M}_1^\star o$. The hierarchical ordering is constructed as follows:

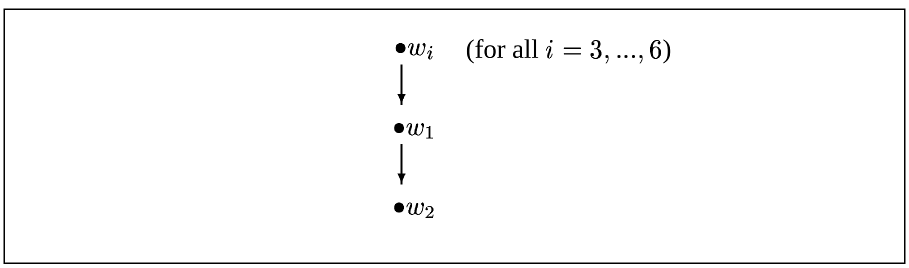


Figure 2. *At the time of offence ($\mathcal{M}_1^\star o$)*

• By $(\mathbf{P}_1)$, w_2 is strictly better than the other worlds, and must be shifted to the bottom of the new ordering. The details are as follows. On the one hand, by a) we get $w_2 \preceq' w$ for all $w \in W -$ since $w_2 \in \min_{\mathcal{M}_1}(o)$. On the other hand, by b) we get $w_i \not\preceq' w_2$ for $i \in \{1, 3, 4, 5, 6\} -$ since $w_i \notin \min_{\mathcal{M}_1}(o)$. Putting the two together, we thus have $w_2 \prec' wi$ for $i \in \{1, 3, 4, 5, 6\}$.

• It remains to re-order the members of $W - \{w_2\}$. None is in $\min_{\mathcal{M}_1}(o)$. So, by $(\mathbf{P}_2)$, inside that set the ideality ordering must be left unaltered. This means that within the subset $\{w_3, \ldots, w_6\}$ neither world is preferred over the other. This also implies that $w_1 \prec' w_i$ for all $i = 3, ..., 6$.

This completes the construction of the new ordering as described in figure 2. $\mathcal{M}_1^\star o$ depicts the ordering *at the time of the offence*. Of special interest is the fact that $\mathcal{M}_1^\star o$ is much like $\mathcal{M}_1$, except that the best world (viz. w_1) and the 2nd best one (viz. w_2) have permuted. For this reason, it is tempting to refer to this way of thinking about remedial interchange as the "commutation approach". It can easily be verified that, at this very moment of the face-to-face relation, we do not have

- $\mathcal{M}_1^\star o \models \bigcirc\neg r_1$
- $\mathcal{M}_1^\star o \models \bigcirc\neg r_2$
- $\mathcal{M}_1^\star o \models \bigcirc\neg a$
- $\mathcal{M}_1^\star o \models \bigcirc\neg m$

but rather

- $\mathcal{M}_1^\star o \models \bigcirc r_1$
- $\mathcal{M}_1^\star o \models \bigcirc r_2$
- $\mathcal{M}_1^\star o \models \bigcirc a$
- $\mathcal{M}_1^\star o \models \bigcirc m$

as one would expect. This is reminiscent of the step where in section 3 the sentence $\oplus \bigcirc r_1$ was inferred from (IIIa) and (IVa).

It is noteworthy that we no longer have

- $\mathcal{M}_1^\star o \models \bigcirc\neg o$,

but rather

- $\mathcal{M}_1^\star o \models \bigcirc o$.

Some may not be satisfied with this result. As noted earlier, something very similar happens in temporal deontic logic. As also observed, within this set-up the appearance of a paradox is considerably diminished: o is in fact settled as true; it is thus vacuously obligatory. The trouble with this line of defence is that it is difficult to interpret $\mathcal{M}_1^\star o$ as stating what is obligatory after that o has been settled as true. Indeed, in Figure 2 above, o is not true at each point in the structure. This means that, if (like in Prakken and Sergot's theory [PRA 97b]) settledness is encoded using a S5 $\Box$-modality, then strictly speaking fact o cannot be deemed a fixed, necessary, unalterable feature of the situation. How can the present account be brought more in line with our intuitions regarding settledness? One way to achieve this has been explored by van der Torre and Tan [TOR 99]. Their semantics is based upon the update semantics that Veltman [VEL 96] has developed for default reasoning. The revision procedure they use (their term for it is "zoom in") does not shift every minimal ϕ-world to the bottom of the ordering, but deletes all worlds in which ϕ is false.[11] This seems to reflect a rather common intuition. The worlds at which the input is false are eliminated, because they are too remote from the actual world, or outside any agent's control. To put

11. A similar device is employed for different purposes in Boutilier and Goldszmidt [BOU 95].

it in Hansson's terms, "an agent cannot 'undo' what he has actually done" [HAN 69, p. 142]. Such a revision policy, which serves as a device for limiting the worlds to be compared under the preference relation in the minimization process, can easily be adapted to the present framework. Let $\mathcal{M}/\phi$ denote the model obtained by eliminating all ϕ-worlds from $\mathcal{M} = (W, \preceq, \iota)$. Formally, $\mathcal{M}/\phi$ is just $(W', \preceq', \iota')$, where $W' = \{w \in W : \mathcal{M}, w \not\models \phi\}$, and $\preceq'$ and ι' are the restrictions of $\preceq$ and ι to W'.[12] I do not pursue this line here, because at first sight it is unclear how to deal with a sequence of inputs that are mutually inconsistent. For instance, if $\mathcal{M}_1$ is first revised by o, and then by $\neg o$, we are left with an empty universe W'', and deontic distinctions collapse: the set of best worlds is empty ; the model satisfies both $\bigcirc\neg r_1$ and $\bigcirc r_1$. As can easily be verified, if definition 3 is used instead, then the resulting model only satisfies the desired obligation $\bigcirc\neg r_1$.

It is natural to ask whether problems of that kind do not indicate the need for an alternative approach. In particular, some might be tempted not to revise the preference relation at all, but rather to distinguish between two distinct relations − a preference relation and an accessibility relation. The argument can be summarized as follows. As time passes the preference between possible worlds remain unchanged, but the accessibility of worlds differs. When o is followed by $\neg o$, it is not necessary to eliminate all possible worlds in the successive diminutions of the accessibility relation. For even if one possible world is no longer accessible, another one very much like it in many respects (but with a different time stamp) may still be accessible. Although I need to subject this point to further investigations, I believe that this way of thinking takes us back to the DARB treatment. This might be seen through a careful examination of how the truth-clauses work, when applied to CTD examples. For instance, in [ÅQV 91a] Åqvist discusses a version of the Chisholm paradox known as "the two medicines case". Commenting on the tree-like diagram showing the consistency in DARB of the premisses set under discussion, Åqvist draws the following conclusion: "Obviously, we may say that the second preference-ordering [i.e. as given at $m + 1$, when the transgression occurs] is the set-theoretical restriction of the first one [i.e. given at m] to those courses-of-events that remain accessible in the actual world as of $[m + 1]$" [ÅQV 91a, p. 130].

4.3. *The full sequence analysed*

Disregarding temporarily the aforementioned difficulty, I now apply the proposed construction to the fully expanded remedial sequence. The tables given below exhaust the moves that can in fact be made, and check whether the account gives the desired result for the matching obligations. For clarity's sake, I proceed step by step, starting with the first move in the sequence, up to the fourth one. I assume that interactants do not repeat identical moves.

12. This is Boutilier and Goldszmidt's definition 3.12, cf. [BOU 95, p. 288].

4.3.1. *One-move interchange*

Consider the following table:

Table 2. *One-move interchange*

Revision sequence	Induced ordering	Output
1. $\mathcal{M}_1^{\star} o$	$w_2 \prec w_1 \prec w_i$ (with $3 \leq i \leq 6$)	$\bigcirc r_1$
2. $\mathcal{M}_1^{\star} \neg o$	$w_1 \prec w_2 \prec w_i$ (with $3 \leq i \leq 6$)	$\bigcirc \neg r_1$

The first line recapitulates the analysis suggested in the previous section. This line says that, at the time of move o, then the best world w_1 commutes with the 2nd best world w_2, so that the obligation of r_1 comes into force. The second line expresses the proposition that, were move o not to be made, then the initial ordering $\mathcal{M}_1$ would be left untouched, and the obligation of $\neg r_1$ would still hold. This is intuitively plausible.

4.3.2. *Two-move interchange*

Table 3 below deals with a two-move interchange. The penultimate column shows how many times the worlds commute during the conversation. The first line considers the case where o is followed by r_1, and the second one the case where o is followed by $\neg r_1$. In both cases, the most recent ordering ($\mathcal{M}_1^{\star} o$) differs from the initial one $(\mathcal{M}_1)$. In the first case, the most recent ordering, with w_2 ranked below w_1, is left untouched, and so is the obligation of r_2. In the second case, we have that w_1 and w_2 commute again. This takes us back to $\mathcal{M}_1$ and, as a result, the obligation of $\neg r_2$ is reinstated.

Line 3 considers the case where $\neg o$ is followed by r_1, and line 4 the case where $\neg o$ is followed by $\neg r_1$. In both cases, the most recent ordering (viz. $\mathcal{M}_1^{\star} \neg o$) is the same as the initial one $(\mathcal{M}_1)$. In the first case, we have that w_1 and w_2 permute, so that the obligation of r_2 comes into force. In the second case, the ordering is left untouched, and so is the obligation of $\neg r_2$.

Table 3. *Two-move interchange*

Revision sequence	Induced ordering	Perm	Output
1. $(\mathcal{M}_1^{\star} o)^{\star} r_1$	$w_2 \prec w_1 \prec w_i$ (with $3 \leq i \leq 6$)	1	$\bigcirc r_2$
2. $(\mathcal{M}_1^{\star} o)^{\star} \neg r_1$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg r_2$
3. $(\mathcal{M}_1^{\star} \neg o)^{\star} r_1$	$w_2 \prec w_1 \prec w_i$	1	$\bigcirc r_2$
4. $(\mathcal{M}_1^{\star} \neg o)^{\star} \neg r_1$	$w_1 \prec w_2 \prec w_i$	0	$\bigcirc \neg r_2$

One might summarize table 3 as follows. If move r_1 is made (line 1 and 3), then the obligation of r_2 comes into force $-$ whatever happened before r_1. Were move r_1 not to be made (line 2 and 4), then the obligation of $\neg r_2$ would still hold $-$ and this, whatever happened before $\neg r_1$. This is intuitively plausible.

4.3.3. *Three-move interchange*

Table 4 below deals with a three-move interchange. This table can be summarized as follows. Once move $r2$ has been made (lines 1, 3, 5 and 7), then the obligation of a comes into force $-$ whatever happened before r_2. Were move r_2 not to be made (lines 2, 4, 6 and 8), then the obligation of $\neg a$ would be true $-$ whatever happened before $\neg r_2$. Again, this is intuitively plausible.

Table 4. *Three-move interchange*

Revision sequence	Induced ordering	Perm	Output
1. $((\mathcal{M}_1^\star o)^\star r_1)^\star r_2$	$w_2 \prec w_1 \prec w_i$ (with $3 \leq i \leq 6$)	1	$\bigcirc a$
2. $((\mathcal{M}_1^\star o)^\star r_1)^\star \neg r_2$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg a$
3. $((\mathcal{M}_1^\star o)^\star \neg r_1)^\star r_2$	$w_2 \prec w_1 \prec w_i$	3	$\bigcirc a$
4. $((\mathcal{M}_1^\star o)^\star \neg r_1)^\star \neg r_2$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg a$
5. $((\mathcal{M}_1^\star \neg o)^\star r_1)^\star r_2$	$w_2 \prec w_1 \prec w_i$ (with $3 \leq i \leq 6$)	1	$\bigcirc a$
6. $((\mathcal{M}_1^\star \neg o)^\star r_1)^\star \neg r_2$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg a$
7. $((\mathcal{M}_1^\star \neg o)^\star \neg r_1)^\star r_2$	$w_2 \prec w_1 \prec w_i$	1	$\bigcirc a$
8. $((\mathcal{M}_1^\star \neg o)^\star \neg r_1)^\star \neg r_2$	$w_1 \prec w_2 \prec w_i$	0	$\bigcirc \neg a$

4.3.4. *Four-move interchange*

As shown in table 5 below, the account deals with the fourth move in much the same way as with the first three. Roughly table 5 says that, once move a has been made (lines marked with an odd number), then the obligation of m holds, and that if move a is not made (lines marked with an even number), then the obligation of $\neg m$ is true. Once again, this is intuitively plausible.

This ends the verification. It might be interesting to explore ways in which the concept of commutation can be used to give a general framework for iterated revision. A closer look at the above tables reveals that, once the first move has been made, the best world and the second best one commute each time the “polarity” of the input sentence changes. The question naturally arises whether it is possible to develop a model of iterated revision based on this phenomenon. This is part of my current research.

4.4. *The request-response within a tense logic*

I now indicate how to incorporate the construction outlined in the previous subsections into a branching-time framework, so as to deal with the request-response more sensitively. I say “more sensitively”, because there is something puzzling in the temporal interpretation (or gloss) given to the treatment in terms of commutation. If we assume that interactants cannot detach the unconditional obligation of providing a remedy (and, as a result, do what they believe to be the second best) unless the offence has been made, then it is unclear how the solution could work for those scenarios

Table 5. *Four-move interchange*

Revision sequence	Induced ordering	Perm	Output
1. $(((\mathcal{M}_1^\star o)^\star r_1)^\star r_2)^\star a$	$w_2 \prec w_1 \prec w_i$ $(3 \leq i \leq 6)$	1	$\bigcirc m$
2. $(((\mathcal{M}_1^\star o)^\star r_1)^\star r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$ $(3 \leq i \leq 6)$	2	$\bigcirc \neg m$
3. $(((\mathcal{M}_1^\star o)^\star r_1)^\star \neg r_2)^\star a$	$w_2 \prec w_1 \prec w_i$	3	$\bigcirc m$
4. $(((\mathcal{M}_1^\star o)^\star r_1)^\star \neg r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg m$
5. $(((\mathcal{M}_1^\star o)^\star \neg r_1)^\star r_2)^\star a$	$w_2 \prec w_1 \prec w_i$	3	$\bigcirc m$
6. $(((\mathcal{M}_1^\star o)^\star \neg r_1)^\star r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$	4	$\bigcirc \neg m$
7. $(((\mathcal{M}_1^\star o)^\star \neg r_1)^\star \neg r_2)^\star a$	$w_2 \prec w_1 \prec w_i$	3	$\bigcirc m$
8. $(((\mathcal{M}_1^\star o)^\star \neg r_1)^\star \neg r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg m$
9. $(((\mathcal{M}_1^\star \neg o)^\star r_1)^\star r_2)^\star a$	$w_2 \prec w_1 \prec w_i$ $(3 \leq i \leq 6)$	1	$\bigcirc m$
10. $(((\mathcal{M}_1^\star \neg o)^\star r_1)^\star r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$ $(3 \leq i \leq 6)$	2	$\bigcirc \neg m$
11. $(((\mathcal{M}_1^\star \neg o)^\star r_1)^\star \neg r_2)^\star a$	$w_2 \prec w_1 \prec w_i$	3	$\bigcirc m$
12. $(((\mathcal{M}_1^\star \neg o)^\star r_1)^\star \neg r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg m$
13. $(((\mathcal{M}_1^\star \neg o)^\star \neg r_1)^\star r_2)^\star a$	$w_2 \prec w_1 \prec w_i$	1	$\bigcirc m$
14. $(((\mathcal{M}_1^\star \neg o)^\star \neg r_1)^\star r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$	2	$\bigcirc \neg m$
15. $(((\mathcal{M}_1^\star \neg o)^\star \neg r_1)^\star \neg r_2)^\star a$	$w_2 \prec w_1 \prec w_i$	1	$\bigcirc m$
16. $(((\mathcal{M}_1^\star \neg o)^\star \neg r_1)^\star \neg r_2)^\star \neg a$	$w_1 \prec w_2 \prec w_i$	0	$\bigcirc \neg m$

where the remedy precedes the offence in time (as is typically the case for remedies of the "request" type). This subsection is an attempt to clarify this point.

Let us assume that the object language contains, in addition to the dyadic deontic operator $\bigcirc(/)$, a monadic operator $\oplus$ read as "it will be the case at the next moment that", and a modal operator of historical necessity $\Box$ to be read as "it is necessary on the basis of the past that". Let us define a deontic branching temporal frame as a structure $\mathfrak{M} = ((Tree, <), \preceq, \iota)$ where:

1) $(Tree, <)$ is a tree-like structure. $Tree$ is a non-empty set of times or moments $\{m_1, m_2, ...\}$. $<$ is the temporal relation (time is assumed to be discrete). A history h is defined as a maximal chain of moments. $\mathbf{H}$ denotes the set of all histories, and H_m the subset of those passing through moment m.

2) $\preceq \subseteq \mathbf{H} \times \mathbf{H}$ is a pre-order, i.e. it is a reflexive and transitive relation on W. Intuitively, $h \preceq h'$ means that h is at least as good as h'.

3) ι is a function that assigns to each propositional letter p the set of m/h pairs at which intuitively p is thought of as true.

The truth-conditions have the form:

$$\mathfrak{M}, m/h \models \phi : \text{in model } \mathfrak{M}, \text{ formula } \phi \text{ is true at moment } m \text{ in history } h.$$

Let $[\phi]_m$ be a shorthand for the set of histories making ϕ true at m, and $\min_{\mathfrak{M}}([\phi]_m)$ denote the $\preceq$-minima in $[\phi]_m$, i.e.

$$\min_{\mathfrak{M}}([\phi]_m) \;=\; \{h \in [\phi]_m \mid \forall h'(h' \in [\phi]_m \;\rightarrow\; h \preceq h')\}.$$

Define:

$$\mathfrak{M}, m/h \models \oplus\phi \Leftrightarrow \mathfrak{M}, m+1/h \models \phi$$
$$\mathfrak{M}, m/h \models \Box\phi \Leftrightarrow (\forall h' \in H_m)(\mathfrak{M}, m/h' \models \phi)$$
$$\mathfrak{M}, m/h \models \bigcirc(\psi/\phi) \Leftrightarrow \min_{\mathfrak{M}}([\phi]_m) \subseteq [\psi]_m.$$

It is now possible to analyse the case where the remedy precedes the offence. In the analysis I shall focus on the pair (offence, request) and ignore subsequent moves. The following triple of formulae (where r_1 denotes the request-response) gives a reasonable description of the situation when (so to say) the offence has not been anticipated yet:

Table 6. *The request-response*

	Normative premisses (α)	Integrity constraint (β)
(I)	$\bigcirc \oplus \neg o$	
(II)	$\bigcirc(r_1/\oplus o)$	$\Box(r_1 \rightarrow \oplus o)$

A typical model $\mathfrak{M}_1$ of these premisses is represented in Figure 3 below. This figure can be read as follows. The upward direction represents the forward direction of time. $\hbar$ denotes the actual evolution of the interaction (evidently this is an arbitrary choice). Our hierarchical ordering is (the transitive closure of) $h_2 \prec h_1 \prec \hbar$. Let us consider the situation from the point of view of $\hbar$ at m_0. We have that:

- $\mathfrak{M}_1, m_0/\hbar \models \bigcirc \oplus \neg o$, since $\min_{\mathfrak{M}_1}(\{h_1, h_2, \hbar\}) = \{h_2\} \subseteq [\oplus\neg o]_{m_0}$
- $\mathfrak{M}_1, m_0/\hbar \models \bigcirc(r_1/\oplus o)$, since $\min_{\mathfrak{M}_1}(\{h_1, \hbar\}) = \{h_1\} \subseteq [r_1]_{m_0}$.

At this very moment, we have that

- $\mathfrak{M}_1, m_0/\hbar \models \bigcirc\neg r_1$,

and not

- $\mathfrak{M}_1, m_0/\hbar \models \bigcirc r_1$.

Let us now imagine that interactants take the offending act into account. As before, it is assumed that this new information affects the hierarchical ordering of histories. Hence the following notation

$$\mathfrak{M}^{\circledast^m}\phi = ((Tree, <), \preceq', \iota)$$

to describe the model obtained once $\mathfrak{M}$ has been revised (at m) with ϕ. The formal definition of this notion parallels the one employed in a timeless framework:

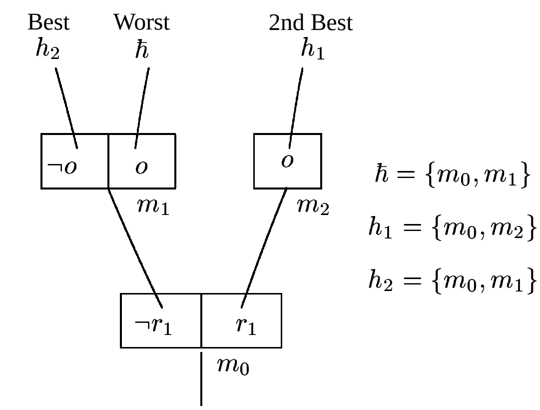


Figure 3. *Before the offence has been taken into account (model $\mathfrak{M}_1$)*

DEFINITION 5. — *Let $\mathfrak{M} = ((Tree, <), \preceq, \iota)$ be the model reflecting the initial ordering. Given the contrary-to-duty context ϕ, the model obtained by revising (at m) $\mathfrak{M}$ by ϕ is $\mathfrak{M}^{\circledast^m}\phi = ((Tree, <), \preceq', \iota)$, with $\preceq'$ defined as follows:*

$(\mathbf{P}_1)$ *If $h \in min_{\mathfrak{M}}([\phi]_m)$ then:*
(a) $h \preceq' h'$ *for all $h' \in H_m$ and*
(b) $h' \preceq' h$ *iff $h' \in min_{\mathfrak{M}}([\phi]_m)$;*

$(\mathbf{P}_2)$ *If $h, h' \notin min_{\mathfrak{M}}([\phi]_m)$ then: $h \preceq' h'$ iff $h \preceq h'$.*

Take ϕ to be $\oplus o$, and m to be m_0. We have

$$\min_{\mathfrak{M}_1}([\oplus o]_{m_0} = \{h_1, \hbar\}) = \{h_1\}.$$

From this it follows that $h_1 \prec' \hbar$ and $h_1 \prec' h_2$ [by $\mathbf{P}_1$]. It also follows that $h_2 \prec' \hbar$ [by $\mathbf{P}_2$]. As illustrated in figure 4 below, this means that the best history permutes with the 2nd best one.

Of special interest is the fact that we now have

- $\mathfrak{M}_1^{\circledast^{m_0}} \oplus o, m_0/\hbar \models \bigcirc r_1$

rather than

- $\mathfrak{M}_1^{\circledast^{m_0}} \oplus o, m_0/\hbar \models \bigcirc \neg r_1$.

As we have seen in section 3, proposals based on temporal deontic logic such as DARB can deal with the case where the remedy and the violation refer to the same time point, but not with the case where the remedy precedes the transgression temporally. The present framework has no difficulty coping with the latter scenario. The troublesome

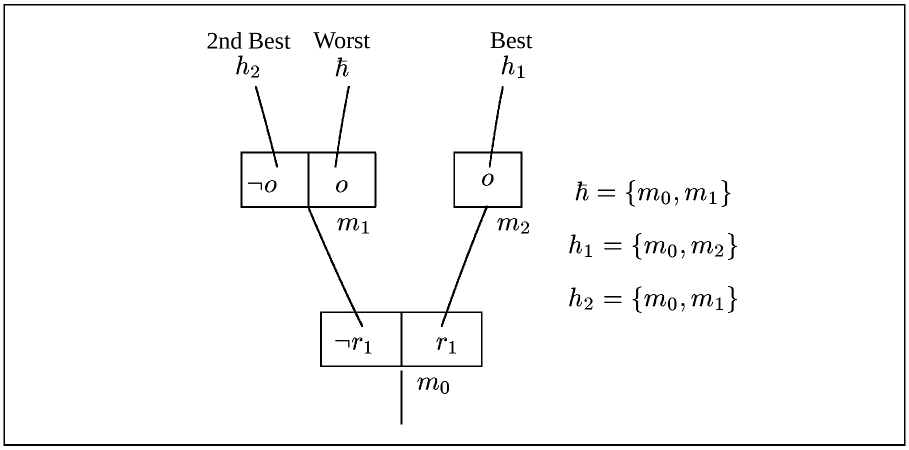


Figure 4. *Once the offence has been anticipated ($\mathfrak{M}_1^{\circledast^{m_0}} \oplus o$)*

premiss (IIIb) is rendered as $\bigcirc(r_1/ \oplus o)$. Thus, the distinction between $\bigcirc\oplus$ and $\oplus\bigcirc$, which is central to Åqvist's proposal, is not used anymore. The primary norm and the CTD norm both embed the "next" operator within the scope of "ought", and both structures make the norms true at one and the same instant, i.e. $m_0/\hbar$. In this connection, it is interesting to note that the non-sequential scenario can be handled without introducing any consideration relative to the passage of time, as illustrated in figure 5. Model $\mathfrak{M}_2$ makes the premisses $\bigcirc\neg o$, $\bigcirc(r_1/o)$ and $\Box(r_1 \to o)$ true at

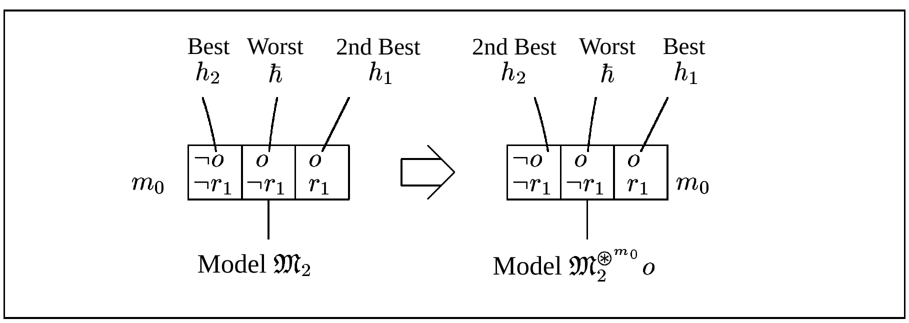


Figure 5. *Non-sequential scenario (No time-stretch)*

moment m_0 in history $\hbar$. And, as a result, the model makes the conclusion $\bigcirc\neg r_1$ true at $m_0/\hbar$. In contrast, the model obtained by revising $\mathfrak{M}_2$ (at m_0) by o makes $\bigcirc r_1$ true at $m_0/\hbar$.

Now that the construction has been incorporated within a branching-time framework, it appears that the parallel initially drawn (for heuristic purposes) between temporal deontic approaches and the account in terms of commutation breaks down. It remains to give a better intuitive explanation to the latter. I shall explore this issue in future research.

5. Concluding remarks

In this paper, I outlined a preference-based account in terms of commutation, which brings the dynamics of obligations to the fore. I also illustrated how the proposed construction may serve as a tool for analysing one kind of linguistic exchange: the remedial interchange.

Of course, the account outlined in this paper can only be suggestive of how future work should proceed. First of all, we have seen that there is a problem with the flat (non-iterated) part of the framework: primary norms do not "survive" an update by a contrary-to-duty context. Can such an anomaly be avoided? Second, it would be interesting to know what the construction has to say about the many other CTD structures that are discussed in deontic literature [PRA 96, PRA 97b, MAK 01, CAR 02]. In particular, it is natural to ask what happens if a "second-level" of CTDs is introduced. Third, it might be thought that any account of interactive communication must be linked, in some way or other, to speech-act theory. These are only hints, to which a later essay will be devoted.

It is clear that, of these three issues, the first one should be our prime target. But a further complication arises from the fact that the commutation approach can also deliver the right conclusion. Hence, such an approach should not be excluded entirely, as it may contain elements necessary to the solution of some other benchmark problems discussed in deontic literature. For instance, example 6 below shows that, due precisely to the choosen revision policy, the account does not generate what has become known as the "pragmatic oddity". The intended point in so-called pragmatic oddity is that the *Factual Detachment* rule is problematic *per se*. This can be illustrated in examples of much the same vein as those considered so far. Consider:

EXAMPLE 6. — (*Broken promise*)

a) $\bigcirc k$: You should keep your promise.

b) $\bigcirc(a/\neg k)$: If you have not kept your promise, you should apologize.

c) $\neg k$: You have not kept your promise.

Suppose $\bigcirc a$ can be inferred from the last two premisses, by using *Factual Detachment*. In this case, $\bigcirc k$ and $\bigcirc a$ both hold. As Prakken and Sergot remark, "it is a bit odd to say that in all [best] versions of this world you keep your promise and you apologize for not keeping it" [PRA 96, p. 95]. The following self-explanatory diagrams make it clear why the present account does not generate the pragmatic oddity here described in its simplest form. Given a typical model $\mathcal{M}_2$ of $\{\bigcirc k, \bigcirc(a/\neg k)\}$, we have $\mathcal{M}_2^{\star}\neg k \models \bigcirc a$ and $\mathcal{M}_2^{\star}\neg k \not\models \bigcirc k$.

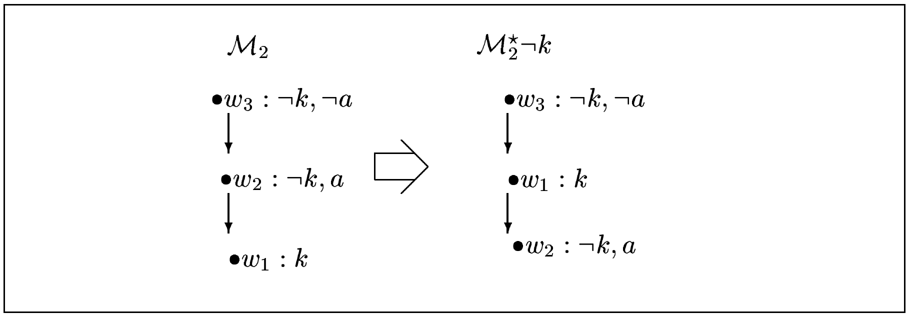


Figure 6. *Broken promise*

Acknowledgements

I am grateful to Paul Gochet, Andreas Herzig, Andrew J.I. Jones, Pierre Livet, David Makinson, Marek Sergot and Audun Stolpe for specific suggestions, doubts and critical remarks. They are, of course, not responsible for the views advocated in the present paper. I also thanks an anonymous referee for helpful comments. A preliminary version of the paper was presented at *MFI-03* (Formal Models for Interaction) held in Lille, France, 20-22 May 2003. I am extremely grateful to the audience of this workshop. I am also indebted to Clare Sibley, who checked my English.

6. References

[APP 89] APPELT D., KONOLIGE K., "A Non-Monotonic Logic for Reasoning about Speech Acts and Belief Revision", REINFRANK M., DE KLEER J., GINSBERG M. L., SANDEWALL E., Eds., *Non-Monotonic Reasoning, 2nd International Workshop*, vol. 346 of *Lecture Notes in Computer Science*, Springer-Verlag, 1989, p. 164–175.

[ÅQV 67] ÅQVIST L., "Good Samaritans, Contrary-To-Duty Imperatives, and Epistemic Obligations", *Noûs*, vol. 1, 1967, p. 361–379.

[ÅQV 81] ÅQVIST L., HOEPELMAN J., "Some Theorems about a 'Tree' System of Deontic Tense Logic", p. 187–221, 1981.

[ÅQV 91a] ÅQVIST L., "Deontic Tense Logic: Restricted Equivalence of Certain Forms of Conditional Obligation and a Solution to Chisholm's Paradox", SCHURZ G., DORN G., Eds., *Advances in Scientific Philosophy: Essays in Honour of Paul Weingartner*, p. 127–141, GA: Rodopi, 1991.

[ÅQV 91b] ÅQVIST L., "Review Essay: Doing the Best We Can", *Philosophy and Phenomenological Research*, vol. 51, 1991, p. 215–225.

[ÅQV 93] ÅQVIST L., "A Completeness Theorem in Deontic Logic with Systematic Frame Constants", *Logique et Analyse*, vol. 141-142, 1993, p. 177–192.

[ÅQV 02] ÅQVIST L., "Deontic Logic", GABBAY D., GUENTHNER F., Eds., *Handbook of Philosophical Logic*, vol. 8, p. 147–264, Kluwer Academic Publishers, Dordrecht, Holland, 2002.

[ART 96] ARTOSI A., GOVERNATORI G., SARTOR G., "Towards a Computational Treatment of Deontic Defeasibility", BROWN M., CARMO J., Eds., *Deontic Logic, Agency and Normative Systems*, p. 27–46, Springer-Verlag, London, 1996.

[BEN 03] BENCH-CAPON T., "Persuasion in Practical Argument Using Value Based Argumentation Framework", *Journal of Logic and Computation*, vol. 13, 2003, p. 429–448.

[BOU 95] BOUTILIER C., GOLDSZMIDT M., "On the Revision of Conditional Belief Sets", CROCCO G., FARIÑAS DEL CERRO L., HERZIG A., Eds., *Conditionals: from Philosophy to Computer Science*, p. 267–300, Clarendon Press, Oxford, 1995.

[BOU 96] BOUTILIER G., "Iterated Revision and Minimal Change of Conditional Beliefs", *Journal of Philosophical Logic*, vol. 25, 1996, p. 263–305.

[BRO 87] BROWN P., LEVINSON S. C., *Politeness. Some Universals in Language Usage*, Cambridge University Press, Cambridge, 1987.

[CAR 95] CARMO J., JONES A. J. I., "Deontic Logic and Different Levels of Ideality", Research report num. RRDMIST 1/95, 1995, Department of Mathematics, Technological University of Lisbon.

[CAR 02] CARMO J., JONES A. J. I., "Deontic Logic and Contrary-To-Duties", GABBAY D., GUENTHNER F., Eds., *Handbook of Philosophical Logic*, vol. 8, p. 265–343, Kluwer Academic Publishers, Dordrecht, Holland, 2002.

[CHE 74] CHELLAS B. F., "Conditional Obligation", STENLUND S., Ed., *Logical Theory and Semantic Analysis*, p. 23–33, D. Reidel, Dordrecht-Holland, 1974.

[CHE 80] CHELLAS B. F., *Modal Logic: an Introduction*, Cambridge University Press, Cambridge, 1980.

[CHI 63] CHISHOLM R. M., "Contrary-To-Duty Imperatives and Deontic Logic", *Analysis*, vol. 24, 1963, p. 33–36.

[CHO 01] CHOLVY L., GARION C., "An Attempt to Adapt a Logic of Conditional Preferences for Reasoning with Contrary-To-Duties", *Fundamenta Informaticae*, vol. 48, 2001, p. 183–204.

[DAN 68] DANIELSSON S., *Preference and Obligation. Studies in the Logic of Ethics*, Filosofiska föreningen, Uppsala, 1968.

[DAR 94] DARWICHE A., PEARL J., "On the Logic of iterated Belief Revision", FAGIN R., Ed., *Proceedings of the Fifth Conference on Theoretical Aspects of Reasoning about Knowledge*, San Mateo, California, 1994, Morgan Kaufmann Publishers, p. 5–23.

[FER 02] FERMÉ E., ROTT H., "Revision by Comparison", Manuscript, December 2002, University of Regensburg.

[FOR 84] FORRESTER J. W., "Gentle Murder, or the Adverbial Samaritan", *Journal of Philosophy*, vol. 81, 1984, p. 193–197.

[GAZ 79a] GAZDAR G., *Pragmatics: Implicature, Presupposition and Logical Form*, Academic Press, New York, 1979.

[GAZ 79b] GAZDAR G., "A Solution to the Projection Problem", OH C.-K., DINNEEN D., Eds., *Syntax and Semantics 11: Presuppositions*, p. 57–89, Academic Press, 1979.

[GOF 71] GOFFMAN E., *Relations in Public: Microstudies of the Public Order*, Basic Books, New York, 1971, [Page refs. to this ed.].

[GRI 75] GRICE H. P., "Logic and Conversation", COLE P., MORGAN J. L., Eds., *Syntax and Semantics 3: Speech Acts*, p. 41–58, Academic Press, New York, 1975.

[HAN 69] HANSSON B., "An Analysis of Some Deontic Logics", *Noûs*, vol. 3, 1969, p. 373–398, Reprinted in [**?**, p. 121–147].

[HAN 99] HANSEN J., "On Relations between Åqvist's Deontic System G and Van Eck's Deontic Temporal Logic", MCNAMARA P., PRAKKEN H., Eds., *Norms, Logics and Information Systems*, p. 127–144, IOS Press, Amsterdam, 1999.

[HOR 01] HORTY J., "Argument Construction and Reinstatement in Logics for Defeasible Reasoning", *Artificial Intelligence and Law*, vol. 9, 2001, p. 1–28.

[JAC 85] JACKSON F., "On the Semantics and Logic of Obligation", *Mind*, vol. 94, 1985, p. 177–196.

[JON 85] JONES A. J. I., PÖRN I., "Ideality, Sub-Ideality and deontic logic", *Synthese*, vol. 65, 1985, p. 275–290.

[KER 94] KERBRAT-ORECCHIONI C., *Les Intéractions Verbales*, vol. III, Colin, Paris, 1994.

[KRA 90] KRAUS S., LEHMANN D., MAGIDOR M., "Nonmonotonic Reasoning, Preferential Models and Cumulative Logics", *Artificial Intelligence*, vol. 44, 1990, p. 167–207.

[LEH 95] LEHMANN D., "Belief Revision, Revised", *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, Montreal, Canada, 1995, p. 1534–1541.

[LEW 74] LEWIS D. K., "Semantical Analysis for Dyadic Deontic Logic", STENLUND S., Ed., *Logical Theory and Semantic Analysis*, p. 1–14, D. Reidel, Dordrecht, 1974.

[MAK 89] MAKINSON D., "General Theory of Cumulative Inference", REINFRANK M., DE KLEER J., GINSBERG M. L., SANDEWALL E., Eds., *Non-Monotonic Reasoning, 2nd International Workshop*, vol. 346 of *Lecture Notes in Computer Science*, p. 1–18, Springer-Verlag, 1989.

[MAK 93] MAKINSON D., "Five Faces of Minimality", *Studia Logica*, vol. 52, 1993, p. 339–379.

[MAK 01] MAKINSON D., VAN DER TORRE L., "Constraints for Input/Output Logics", *Journal of Philosophical Logic*, vol. 30, 2001, p. 155–185.

[MAK 03] MAKINSON D., "Ways of Doing Logic: What was Different about AGM 1985", *Journal of Logic and Computation*, vol. 13, 2003, p. 3–13.

[MCC 80] MCCARTHY J., "Circumscription - A Form of Non-Monotonic Reasoning", *Artificial Intelligence*, vol. 13, 1980, p. 27–39.

[MER 88] MERCER R. E., "Using Default Logic to Derive Natural Language Presuppositions", GOEBEL P., Ed., *Proceedings of the Seventh Biennal Conference of the Canadian Society for Computational Studies of Intelligence*, Morgan Kaufmann, 1988, p. 14–21.

[MER 90] MERCER R. E., "Deriving Natural Language Presuppositions from Complex Conditionals", *Proceedings of the Eighth Biennal Conference of the Canadian Society for Computational Studies of Intelligence*, Ottawa, May 1990, p. 114–120.

[MOO 85] MOORE R. C., "Semantical Considerations on Non-Monotonic Logic", *Artificial Intelligence*, vol. 25, 1985, p. 75–94.

[NAY 03] NAYAK A. C., PAGNUCCO M., PEPPAS P., "Dynamic Belief Revision Operators", *Artificial Intelligence*, vol. 146, 2003, p. 193–228.

[OWE 83] OWEN O., *Apologies and Remedial Interchanges. A Study of Language Use in Social Interaction*, Mouton Publishers, Berlin, 1983.

[PER 90] PERRAULT C. R., "An Application of Default Logic to Speech Act Theory", COHEN P. R., MORGAN J., POLLACK M. E., Eds., *Intentions in Communication*, p. 161–185, MIT Press, Cambridge, 1990.

[POL 87] POLLOCK J. L., "Defeasible Reasoning", *Cognitive Science*, vol. 11, 1987, p. 481–518.

[PRA 96] PRAKKEN H., SERGOT M. J., "Contrary-To-Duty Obligations", *Studia Logica*, vol. 57, 1996, p. 91–115.

[PRA 97a] PRAKKEN H., SARTOR G., "Argument-Based Extended Logic Programming with Defeasible Priorities", *Journal of Applied Non-Classical Logics*, vol. 7, 1997, p. 25–75.

[PRA 97b] PRAKKEN H., SERGOT M. J., "Dyadic Deontic Logic and Contrary-To-Duty Obligations", NUTE D., Ed., *Defeasible Deontic Logic*, p. 223–262, Kluwer Academic Publishers, Dordrecht, 1997.

[PRA 02] PRAKKEN H., "Intuitions and the Modelling of Defeasible Reasoning: some Case Studies", *Proceedings of the Ninth International Workshop on Non-Monotonic Reasoning*, Toulouse, 2002, p. 91–99.

[REI 80] REITER R., "A Logic for Default Reasoning", *Artificial Intelligence*, vol. 13, 1980, p. 81–132.

[RYA 92] RYAN M., "Representing Defaults as Sentences with Reduced Priority", NEBEL B., RICH C., SWARTOUT W., Eds., *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*, Cambridge, 1992, Morgan Kaufmann Publishers, p. 649–660.

[SPO 88] SPOHN W., "Ordinal Conditional Functions: A Dynamic Theory of Epistemic States", HARPER W. L., SKYRMS B., Eds., *Causation in Decision, Belief Change and Statistics*, vol. 2, p. 105–134, Kluwer Academic Publishers, Dordrecht, 1988.

[THO 81] THOMASON R. H., "Deontic Logic as Founded on Tense Logic", p. 165–176, 1981.

[TOR 97] VAN DER TORRE L., TAN Y.-H., "The Many Faces of Defeasibility", NUTE D., Ed., *Defeasible Deontic Logic*, p. 79–121, Kluwer Academic Publishers, Dordrecht, 1997.

[TOR 98] VAN DER TORRE L., TAN Y.-H., "The Temporal Analysis of Chisholm's Paradox", *Proceedings of the Fifteenth National Conference on Artificial Intelligence and Tenth Innovative Applications of Artificial Intelligence Conference, AAAI 98, IAAI 98*, Madison, Wisconsin, USA, July 26-30, 1998, AAAI Press / The MIT Press,, p. 650–655.

[TOR 99] VAN DER TORRE L., TAN Y.-H., "An Update Semantics for Deontic Reasoning", MCNAMARA P., PRAKKEN H., Eds., *Norms, Logics and Information Systems*, p. 73–90, IOS Press, 1999.

[TOU 58] TOULMIN S. E., *The Uses of Argument*, Cambridge University Press, Cambridge, 1958.

[VEL 96] VELTMAN F., "Defaults in Update Semantics", *Journal of Philosophical Logic*, vol. 25, 1996, p. 221–261.

[WIL 94] WILLIAMS M.-A., "Transmutations of Knowledge Systems", DOYLE J., SANDEWALL E., TORASSO P., Eds., *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning (KR'94)*, Bonn, Germany, May 24-27, 1994, Morgan Kaufmann, p. 619–629.