













Proofs are characterized by the following generic rules, where “ $A \quad \Delta$ ” abbreviates that  $A$  occurs in the proof on the condition  $\Delta$ :

PREM	If $A \in \Gamma$ :	$\frac{\dots \quad \dots}{A \quad \emptyset}$
		$A_1 \quad \Delta_1$
RU	If $A_1, \dots, A_n \vdash_{\text{LLL}} B$ :	$\frac{\vdots \quad \vdots}{A_n \quad \Delta_n}$
		$B \quad \Delta_1 \cup \dots \cup \Delta_n$
RC	If $A_1, \dots, A_n \vdash_{\text{LLL}} B \vee \text{Dab}(\Theta)$ :	$\frac{\vdots \quad \vdots}{A_n \quad \Delta_n}$
		$B \quad \Delta_1 \cup \dots \cup \Delta_n \cup \Theta$

Of course it is not enough to apply rules conditionally since there may be cases in which the condition turns out to be violated. Thus, we need (a) to define a criterion for what it means that a condition is violated and (b) introduce a mechanism to mark lines with violated conditions. The adaptive strategy and the corresponding marking conditions take care of this.

We shall need to consider *stages* of proofs, which are lists of lines obtained by applications of rules PREM, RU and RC (with the usual understanding that the justification of a line should only refer to lines preceding it in the list). The empty list will be considered as stage 0 of every proof. Where  $s$  is a stage,  $s'$  is an extension of  $s$  iff all lines that occur in  $s$  occur in the same order in  $s'$ . A (dynamic) proof is a chain of stages.

The following notions are useful for the definitions of the adaptive strategies.  $\text{Dab}(\Delta)$  is a *minimal Dab-formula* at a stage  $s$  of the proof iff it is the formula of a line with condition  $\emptyset$  and no  $\text{Dab}(\Delta')$  with  $\Delta' \subset \Delta$  is the formula of a line with condition  $\emptyset$ . Where  $\text{Dab}(\Delta_1), \dots, \text{Dab}(\Delta_n)$  are the minimal Dab-formulas at stage  $s$  with respect to a premise set  $\Gamma$ , we define the set of *unreliable formulas* at stage  $s$ ,  $U_s(\Gamma) = \Delta_1 \cup \dots \cup \Delta_n$ . The minimal Dab-formulas derivable with the lower limit logic we call *minimal Dab-consequences*. Where  $\text{Dab}(\Delta_1), \text{Dab}(\Delta_2), \dots$  are the minimal Dab-consequences with respect to a premise set  $\Gamma$ , we define the set of unreliable formulas,  $U(\Gamma) = \Delta_1 \cup \Delta_2 \cup \dots$ .

Let us have a look at the marking conditions of the two adaptive strategies used in this paper. Given a set of abnormalities  $\Omega$ , it is the job of the marking conditions to determine if lines are “in” or “out” of

the proof at a certain stage, i.e., to govern the internal dynamics of the proof procedure. For the *reliability strategy* lines are marked which have unreliable formulas in their condition.

DEFINITION 1 (Marking for Reliability). Line  $i$  is marked at stage  $s$  iff, where  $\Delta$  is its condition,  $\Delta \cap U_s(\Gamma) \neq \emptyset$ .

For the *minimal abnormality strategy* a few more notions need to be introduced. A *choice set* of  $\Sigma = \{\Delta_1, \Delta_2, \dots\}$  is a set that contains at least one element out of each member of  $\Sigma$ . A *minimal choice set* of  $\Sigma$  is a choice set of  $\Sigma$  of which no proper subset is a choice set of  $\Sigma$ .<sup>5</sup> Where  $\text{Dab}(\Delta_1), \dots, \text{Dab}(\Delta_n)$  are the minimal Dab-formulas that are derived on condition  $\emptyset$  from a premise set  $\Gamma$  at stage  $s$ ,  $\Phi_s(\Gamma)$  is the set of minimal choice sets of  $\{\Delta_1, \dots, \Delta_n\}$ .

DEFINITION 2 (Marking for Minimal Abnormality). Line  $i$  is marked at stage  $s$  iff, where  $A$  derived on the condition  $\Delta$  at line  $i$ ,

- (i) there is no  $\Delta' \in \Phi_s(\Gamma)$  such that  $\Delta' \cap \Delta = \emptyset$ , or
- (ii) for some  $\Delta' \in \Phi_s(\Gamma)$ , there is no line at which  $A$  is derived on a condition  $\Theta$  for which  $\Delta' \cap \Theta = \emptyset$ .

In order to highlight the difference between the two strategies consider the following example: suppose  $A$  has been derived at line  $l_1$  on condition  $\{B_1\}$  as well as at line  $l_2$  on condition  $\{B_2\}$ . Furthermore, suppose that at this stage of the proof we have one minimal Dab-formula, namely  $B_1 \vee B_2$ . Thus, at least one of the abnormalities,  $B_1$  or  $B_2$ , is valid. They are both considered as being unreliable. By the reliability strategy both lines,  $l_1$  and  $l_2$ , are marked. The minimal abnormality strategy, however, is more rigorous in interpreting the premises as normal as possible. In this case it ensures that one and only one of our abnormalities is interpreted as valid, so that it is not the case that both  $B_1$  and  $B_2$  are valid. Note that the minimal choice sets are at this stage  $\{B_1\}$  and  $\{B_2\}$ . Thus, since  $A$  has been derived on either condition and since at least one of the abnormalities is interpreted as false,  $A$  is considered as a consequence.

Note that a line might be marked at stage  $s$  of the proof, but become unmarked at a later stage  $s'$ . Indeed, even if  $\text{Dab}(\Delta)$  is a minimal Dab-formula at stage  $s$ , we might be able to derive  $\text{Dab}(\Delta')$  where  $\Delta' \subset \Delta$  at

---

<sup>5</sup>Let for instance  $\Sigma = \{\{1, 2\}, \{1, 3\}\}$ . Choice sets are  $\{1\}, \{1, 2\}, \{1, 3\}, \{2, 3\}$  and  $\{1, 2, 3\}$ . Minimal choice sets are  $\{1\}$  and  $\{2, 3\}$ .





stage  $s'$  and some  $A \in \Delta \setminus \Delta'$  might lose the status of being unreliable at this point. Consequently some formulas derived on conditions including these might get unmarked.

In order to define the consequence set of an adaptive logic we are interested in a stable criterion for derivability.

**DEFINITION 3.**  $A$  is *finally derived* from  $\Gamma$  on line  $i$  of a proof at stage  $s$  iff (i)  $A$  is the second element of line  $i$ , (ii) line  $i$  is not marked at stage  $s$  and (iii) every extension of the proof in which line  $i$  is marked may be further extended in such a way that line  $i$  is unmarked.

$\Gamma \vdash_{\mathbf{AL}} A$  ( $A$  is *finally AL-derivable* from  $\Gamma$ ) iff  $A$  is finally derived on a line of a proof from  $\Gamma$ .

Let us have a look at the semantics: here the strategy selects a set of **LLL**-models of a given premise set  $\Gamma$  in view of the abnormalities verified by the models. Given a logic  $\mathbf{L}$ , let  $\mathcal{M}_{\mathbf{L}}(\Gamma)$  be the set of all  $\mathbf{L}$ -models of  $\Gamma$ .

**DEFINITION 4.** An **LLL**-model  $M \in \mathcal{M}_{\mathbf{LLL}}(\Gamma)$  is *reliable* iff  $\text{Ab}(M) \subseteq U(\Gamma)$ , where  $\text{Ab}(M) =_{\text{df}} \{A \mid M \models A\} \cap \Omega$ . For the reliability strategy the selected set of **LLL**-models of  $\Gamma$  is the set of reliable **LLL**-models of  $\Gamma$ .

An **LLL**-model  $M \in \mathcal{M}_{\mathbf{LLL}}(\Gamma)$  is *minimally abnormal* iff there is no **LLL**-model  $M' \in \mathcal{M}_{\mathbf{LLL}}(\Gamma)$  such that  $\text{Ab}(M') \subset \text{Ab}(M)$ . For minimal abnormality strategy the selected set of **LLL**-models of  $\Gamma$  is the set of the minimal abnormal **LLL**-models of  $\Gamma$ .

$\Gamma \models_{\mathbf{AL}} A$  ( $A$  is an *AL-semantic consequence* of  $\Gamma$ ) iff  $A$  is verified by all members of the selected set of **LLL**-models of  $\Gamma$ .

The idea behind the minimal abnormality strategy is that only the models (of a given premise set) which validate a minimal set of abnormalities (that is, which are the “minimally abnormal” ones with respect to  $\subset$ ) are taken into account. Similarly the reliability strategy takes into account only models that validate no more than the unreliable formulas with respect to a given premise set.

Finally we want to state a representational result for adaptive logics in standard format (see [1]):

**THEOREM 1.**  $\Gamma \vdash_{\mathbf{AL}} A$  iff  $\Gamma \models_{\mathbf{AL}} A$ .



#### 4. Going adaptive – the monadic case

The authors in [16] propose adaptive logics on the basis of **DPM.1** (resp. **DPM.2'**).

**DEFINITION 5.** **ADPM. $\alpha$**  (where  $\alpha \in \{1, 2'\}$ ) is an adaptive logic in standard format defined by the following triple:

- lower limit logic: **DPM. $\alpha$**
- abnormalities:  $\Omega^c = \{OA \wedge O\neg A \mid A \in \mathcal{W}_1\}$
- strategy: minimal abnormality

The main motivation behind these logics stems from the following facts:

$$\text{If } A \vdash_{\mathbf{DPM}} B \text{ then } OA \vdash_{\mathbf{DPM}} OB \vee (OA \wedge O\neg A), \quad (1)$$

$$\{OA, OB\} \vdash_{\mathbf{DPM.2'}} O(A \wedge B) \vee ((OA \wedge O\neg A) \vee (OB \wedge O\neg B)), \quad (2)$$

where  $\mathbf{DPM} \in \{\mathbf{DPM.1}, \mathbf{DPM.2'}\}$ . This enables us to derive  $OB$  from  $OA$  on the condition  $\{OA \wedge O\neg A\}$  in (1). Furthermore in **ADPM.2'** aggregation may be applied to  $OA \wedge OB$  on the condition  $\{OA \wedge O\neg A, OB \wedge O\neg B\}$  due to (2). Thus, in the adaptive logics defined above the inheritance principle is applied conditionally. Furthermore, as demonstrated in (2), aggregation is applied conditionally in **ADPM.2'**.

The following results have been shown in [16].

**THEOREM 2.** *Adding (D) to **ADPM.1** (resp. **ADPM.2'**) as an axiom results in a logic equivalent to **SDL**. Furthermore, for all **SDL**-consistent premise sets **ADPM.1** (resp. **ADPM.2'**) is equivalent to **SDL**.*

Furthermore, none of Goble's explosion principles (**DEX-1**)–(**DEX-3**) is valid in these two logics (see [16]).

Additionally, the major benefit of these two logics is that they perform all the reasoning for the user. No permission statements have to be added as it was the case with **DPM.1** and **DPM.2'**. Note that in order to apply restricted inheritance to  $OA$  we also need  $PA$ . In cases in which  $PA$  is not derivable from the premises, the user of **DPM.1** and **DPM.2'** has to add  $PA$  manually to the premise set. The adaptive versions however apply inheritance conditionally to  $OA$  without user interference. Moreover, problematic lines leading to deontic conflicts get marked as a part of the proof dynamics. In this way deontic conflicts are located as part of the proof.



*Example 1.* We take a look at what we shall dub the “life saver paradox”: Suppose the unfortunate situation that Anne and Bob are just about to drown. However, due to space and time related restrictions it is possible to save the life of only one of them. For each of them individually the obligation to save her or his life,  $OA$  and  $OB$ , is actual. The following proof is for **ADPM.1**:

1	$OA$	PREM	$\emptyset$
2	$OB$	PREM	$\emptyset$
3	$O(\neg(A \wedge B))$	PREM	$\emptyset$
4	$O(A \vee B)$	1; RC	$\{OA \wedge O\neg A\}$
5	$O(A \vee B)$	2; RC	$\{OB \wedge O\neg B\}$
6	$(OA \wedge O\neg A) \vee (OB \wedge O\neg B)$	1, 2, 3; RU	$\emptyset$

Note that it is possible to derive from lines 1, 2 and 3 line 6. Lines 4 and 5 contain conditional applications of the inheritance principle to  $OA$  resp.  $OB$ . Due to the fact that  $O(A \vee B)$  is derivable on both conditions,  $\{OA \wedge O\neg A\}$  and  $\{OB \wedge O\neg B\}$ , lines 4 and 5 are not marked. Clearly the obligation to save either Anne or Bob,  $O(A \vee B)$ , is a desired consequence of our logic.

Note that if we would define **ADPM.1** based on the reliability strategy, lines 4 and 5 would be marked. This is clearly undesired. Thus, the minimal abnormality strategy is the more intuitive choice.

## 5. The conditional case

Most of our moral or behavioral norms are in a conditional form. For instance “Being in an airplane, you ought to turn off your mobile phone.” One proposal to model this in deontic logics is to use  $P \supset OM$ . A disadvantage of this form is that it offers unrestricted “strengthening the antecedent” (SA): from  $P \supset OM$ ,  $(P \wedge B) \supset OM$  is derivable. In many cases this is as expected. But consider the following case:

1. In general, we’re supposed not to eat with fingers.
2. Eating asparagus, we’re allowed to eat with fingers.

Modeling (1) by  $\top \supset O\neg F$ ,  $A \supset O\neg F$  is derivable. This obviously is in conflict with (2)  $A \supset PF$  which is equivalent to  $A \supset \neg O\neg F$ . Similarly problematic is to use  $O(C \supset B)$  to represent the obligation  $B$  under condition  $C$ . In the example we can derive by the inheritance principle

the counter-intuitive  $O(A \supset \neg F)$  from  $O(\top \supset \neg F)$ . Also restricting the inheritance principle as proposed by Goble doesn't help, as nothing speaks against adding the harmless premise  $P(\top \supset \neg F)$ . However, this again enables the derivation of  $O(A \supset \neg F)$  by (RPM) from  $O(\top \supset \neg F)$ .

Monadic approaches were also the object of other kinds of criticism. The most prominent class of problems have to do with paradoxes such as the Chisholm Paradox (see [2]) or the Gentle Murderer Paradox (see [4]). It is commonly agreed that dyadic approaches are in general better suited to deal with these kind of problems.

Modeling conditional obligations with dyadic obligation operators also allows for a more subtle approach to (SA). We use  $O(A \mid B)$  in order to express that “under condition  $B$  it ought to be that  $A$ ”. As usual, the permission operator  $P(A \mid B)$  is defined as  $\neg O(\neg A \mid B)$ . We again use a propositional language, this time enriched by a dyadic obligation operator. Our well-formed formulas  $\mathcal{W}_2$  and the set  $\mathcal{W}_2^O$  are defined as follows:

$$\begin{aligned} \mathcal{W}_2 ::= & \perp \mid \top \mid \langle S \rangle \mid \langle \mathcal{W}_2 \rangle \wedge \langle \mathcal{W}_2 \rangle \mid \\ & \langle \mathcal{W}_2 \rangle \vee \langle \mathcal{W}_2 \rangle \mid \langle \mathcal{W}_2 \rangle \supset \langle \mathcal{W}_2 \rangle \mid \neg \langle \mathcal{W}_2 \rangle \mid O(\langle \mathcal{W}_2 \rangle \mid \langle \mathcal{W}_2 \rangle) \\ \mathcal{W}_2^O ::= & O(\langle \mathcal{W}_2 \rangle \mid \langle \mathcal{W}_2 \rangle) \mid \langle \mathcal{W}_2^O \rangle \wedge \langle \mathcal{W}_2^O \rangle \mid \langle \mathcal{W}_2^O \rangle \vee \langle \mathcal{W}_2^O \rangle \mid \\ & \langle \mathcal{W}_2^O \rangle \supset \langle \mathcal{W}_2^O \rangle \mid \neg \langle \mathcal{W}_2^O \rangle \end{aligned}$$

In this way we can formalize our example by: (A1)  $O(\neg F \mid \top)$  and (A2)  $P(F \mid A)$ . An unrestricted (SA)

$$\vdash O(B \mid A) \supset O(B \mid A \wedge C) \quad (\text{SA})$$

would lead to the counter-intuitive  $O(\neg F \mid A)$ .

As in the case of the restricted inheritance principle (RPM), a possible way of restricting (SA) is to require certain permission statements. Inspired by Kraus, Lehmann and Magidor's work (see [11, 12]) on non-monotonic consequence relations a candidate can be found in the principle of *Rational Monotonicity*

$$\vdash (O(B \mid A) \wedge P(C \mid A)) \supset O(B \mid A \wedge C) \quad (\text{RatMono})$$

Lou Goble proposed different axiomatizations of a standard dyadic deontic logic (which are equivalent to van Fraassen's **CD** in [17], David



Lewis's **VN** in [13] and his own **SDDL** in [7]). Let **RSDDL** be the logic consisting of all classical tautologies, (**RatMono**) and the following rules:

$$\text{If } \vdash A \equiv B \text{ then } \vdash \mathbf{O}(C \mid A) \equiv \mathbf{O}(C \mid B) \quad (\text{RCE})$$

$$\text{If } \vdash B \supset C \text{ then } \vdash \mathbf{O}(B \mid A) \supset \mathbf{O}(C \mid A) \quad (\text{RCM})$$

and axioms

$$\vdash \mathbf{O}(\top \mid \top) \quad (\text{CN})$$

$$\vdash \neg \mathbf{O}(\perp \mid A) \quad (\text{CP})$$

$$\vdash (\mathbf{O}(B \mid A) \wedge \mathbf{O}(C \mid A)) \supset \mathbf{O}(B \wedge C \mid A) \quad (\text{CAND})$$

$$\vdash \mathbf{O}(B \mid A) \supset \mathbf{O}(A \mid A) \quad (\text{QR})$$

$$\vdash \mathbf{O}(C \mid A \wedge B) \supset \mathbf{O}(B \supset C \mid A) \quad (\text{S})$$

But employing (**RatMono**) instead of (**SA**) is problematic as well. In our example it is in no way counter-intuitive to add the statement  $\mathbf{P}(A \mid \top)$ , “It is in general allowed to eat asparagus”, to the premise set. But now it is again possible to derive  $\mathbf{O}(\neg F \mid A)$ . A further restriction is needed. What about the following additional restriction?

$$\vdash (\mathbf{O}(B \mid A) \wedge \mathbf{P}(C \mid A) \wedge \mathbf{P}(B \mid A)) \supset \mathbf{O}(B \mid A \wedge C) \quad (\text{PRatMono})$$

However, this also is counter-intuitive. If just the harmless premise  $\mathbf{P}(\neg F \mid \top)$  is added, again  $\mathbf{O}(\neg F \mid A)$  is derivable.

### 5.1. Lou Goble's Proposal

Lou Goble proposes in [5, 8] the following *Weak Rational Monotonicity* principle:

$$\vdash (\mathbf{O}(B \mid A) \wedge \mathbf{P}(B \wedge C \mid A)) \supset \mathbf{O}(B \mid A \wedge C) \quad (\text{WRM})$$

On basis of his **DPM** systems he presents the following conditional logics:

**DEFINITION 6.** Enriching classical propositional logic by (**RCE**), (**CN**), (**CAND**) and the rules

$$\text{If } \vdash B \equiv C \text{ then } \vdash \mathbf{O}(B \mid A) \equiv \mathbf{O}(C \mid A) \quad (\text{CRE})$$

$$\text{If } \vdash B \supset C \text{ then } \vdash \mathbf{P}(B \mid A) \supset (\mathbf{O}(B \mid A) \supset \mathbf{O}(C \mid A)) \quad (\text{RCPM})$$

results in logic **CDPM.1**. The logic **CDPM.1c** is **CDPM.1** enriched by (**WRM**), (**QR**) and (**S**).



The idea behind the restricted inheritance principle (**RCPM**) is analogous to the monadic case: inheritance is only applied to non-conflicting obligations. That is to say, it is applied to  $O(A \mid B)$  only in case we also have  $P(A \mid B)$ .

DEFINITION 7. **CDPM.2'** is defined analogously to **CDPM.1**, just add (**CP**) and replace (**CAND**) by<sup>6</sup>

$$\vdash (O(A \mid C) \wedge O(B \mid C) \wedge P(A \mid C) \wedge P(B \mid C)) \supset O(A \wedge B \mid C) \quad (\text{CPAND}')$$

Logic **CDPM.2'c** is like **CDPM.1c** with the exception of (**CAND**) which is replaced by (**CPAND'**). Furthermore (**CP**) is added.

For the remainder of the paper, it is useful to introduce some writing conventions: We write  $\oplus(\{P_1, \dots, P_n\}, \mathbf{L})$  for the logic  $\mathbf{L}'$  that is defined as  $\mathbf{L}$  with the addition of principles  $P_1, \dots, P_n$  where each  $P_i \in \mathcal{P}$  and<sup>7</sup>

$$\mathcal{P} =_{\text{df}} \{\text{WRM}, \text{WRM}_*, \text{CAND}, \text{CPAND}', \text{QR}, \text{PS}, \text{S}, \text{CP}, \text{AWRM}_*, \text{CD}\}.$$

Define  $\mathcal{L} =_{\text{df}} \{\oplus(\Psi, \mathbf{CDPM}^-) \mid \Psi \in \wp(\mathcal{P})\}$  where  $\mathbf{CDPM}^-$  is defined by (**RCE**), (**CN**), (**CRE**) and (**RCPM**). Let  $\mathbf{L} = \oplus(\{P_1, \dots, P_n\}, \mathbf{CDPM}^-) \in \mathcal{L}$ . We write  $\ominus(\{P_i \mid i \in I\}, \mathbf{L})$  (where  $I \subseteq \{1, \dots, n\}$ ) for the logic  $\mathbf{L}' = \oplus(\{P_i \mid i \in \{1, \dots, n\} \setminus I\}, \mathbf{CDPM}^-)$ . We write  $\oplus_P \mathbf{L}$  (resp.  $\ominus_P \mathbf{L}$ ) instead of  $\oplus(\{P\}, \mathbf{L})$  (resp.  $\ominus(\{P\}, \mathbf{L})$ ). Further define for  $\mathbf{L} = \oplus(\{P_i \mid i \in I\}, \mathbf{CDPM}^-) \in \mathcal{L}$ ,  $\downarrow \mathbf{L}$  as the set of all sub-logics of  $\mathbf{L}$  in  $\{\ominus(\Psi, \mathbf{L}) \mid \Psi \in \wp(\{P_i \mid i \in I\})\}$ .

None of the following “deontic explosion principles” are valid in **CDPM1.c** and **CDPM.2'c**:<sup>8</sup>

<sup>6</sup>Again we use a slight variation of Goble’s **CDPM.2** which employs  $\vdash (O(A \mid C) \wedge O(B \mid C) \wedge P(A \wedge B \mid C)) \supset O(A \wedge B \mid C)$  (**CPAND**) instead of our (**CPAND'**). Using **CDPM.2'c** instead of **CDPM.2c** as lower limit logic leads to technically more elegant adaptive logics. Furthermore, in contrast to **CDPM.2c**, **CDPM.2'c** fulfills criterion (**C\***), that is going to be introduced in a moment.

<sup>7</sup>Some of the principles in  $\mathcal{P}$  will be defined later on (namely **WRM\***, **PS**, **CD** and **AWRM\***).

<sup>8</sup>We slightly adjusted the criteria offered by Goble since his criteria were formulated in terms of theoremhood, while we focus on the consequences of premise sets. Models validating counter-instances of the criteria can be found in the proof of Theorem 13 in the Appendix.



If  $\not\vdash B$  then  $\{O(A | C), O(\neg A | C)\} \vdash O(B | C)$  (DEX-1)

$\{O(A | C), O(\neg A | C), P(B | C)\} \vdash O(B | C)$  (DEX-2)

$\{O(D | C), P(D | C), O(A | C), O(\neg A | C), P(B | C)\} \vdash O(B | C)$  (DEX-3)

Goble defined a dyadic variant of convention  $(\star)$  which is satisfied by both logics, **CDPM.1c** and **CDPM.2'c**:

$(C\star)$ : A dyadic deontic logic for dilemmas should be such that the result of adding  $(CD)$ , namely  $O(B | A) \supset \neg O(\neg B | A)$ , as an axiom is equivalent to **(R)SDDL**.

**THEOREM 3.** *Where  $\alpha \in \{1, 2'\}$ , **CDPM. $\alpha$ c** satisfies  $(C\star)$ .*

## 5.2. A critical analysis

**(A) CDPM.1c validates (PRatMono).** As has been demonstrated by Goble, one severe shortcoming of **CDPM.1c** is that **(PRatMono)** is derivable from it. As shown above, this leads to counter-intuitive behavior.

**(B) Explosive Behavior and other problems with (WRM).** We saw above that different restricted versions of **(SA)**, such as **(RatMono)** and **(PRatMono)**, are counter-intuitive as soon as we add further harmless premises to our asparagus example. Is the weak rational monotonicity principle proposed by Lou Goble robust to criticism of this kind? Goble demonstrates that it is not. It is not conflicting with our moral intuitions to add the premise  $P(\neg F \wedge A | \top)$ , “in general it is allowed not to eat with fingers and also to eat asparagus”, to the premise set consisting of **(A1)** and **(A2)**. But in this case, by applying **(WRM)** to  $O(\neg F | \top)$  and  $P(\neg F \wedge A | \top)$ , we arrive at  $O(\neg F | A)$ . This causes not just a deontic conflict, but a full-fledged explosion, as we also have  $P(F | A)$  which is equivalent to  $\neg O(\neg F | A)$ .

Note also that the following counter-intuitive statement is a consequence of **(WRM)**:  $\vdash (O(\neg F | \top) \wedge P(F | A)) \supset \neg P(\neg F \wedge A | \top)$ .

We add another problematic example: (1) in a hospital you ought not to smoke –  $O(\neg S | H)$ ; (2) If you're in a smoking room, you're allowed to smoke –  $P(S | R)$ ; (3) If you're in a hospital, you are allowed to be in a smokers room and not to smoke –  $P(R \wedge \neg S | H)$ ; (4) If you're in a smoking room (and) in a hospital, you're allowed to smoke –  $P(S | H \wedge R)$ .

Note that by **(WRM)** we can derive  $O(\neg S \mid H \wedge R)$  from (1) and (3). In face of (2) and (4) this is counter-intuitive and contradictory.

**Solving problems (A) and (B).** There is a way of tackling our problems in a non-adaptive way by further restricting **(WRM)**:

$$\vdash (O(B \mid A) \wedge P(B \wedge C \mid A) \wedge \neg P(\neg B \wedge A \mid C)) \supset O(B \mid A \wedge C) \quad (\text{WRM}_\star)$$

The idea is to strengthen the antecedent  $A$  of obligation  $B$  by  $C$  only if  $\neg B \wedge A$  is not allowed in the context described by  $C$ . Looking back at the asparagus example this obviously blocks the unwanted derivation. By **(WRM<sub>★</sub>)** we have  $\vdash (O(\neg F \mid \top) \wedge P(\neg F \wedge A \mid \top) \wedge \neg P(F \mid A)) \supset O(\neg F \mid A)$ . Note that  $P(F \mid A)$  is a premise and thus  $O(\neg F \mid A)$  is not derivable by **(WRM<sub>★</sub>)**. Also in the case of the second example we have no means to derive  $O(\neg S \mid H \wedge R)$ . We would need  $\neg P(S \wedge H \mid R)$ , which is obviously counter-intuitive.

As **(PRatMono)** was derivable in **CDPM.1c** while not being derivable in **CDPM.2'c**, this leads to an undesirable asymmetry. However, this asymmetry disappears in case of  $\oplus_{\text{WRM}_\star} \ominus_{\text{WRM}} \mathbf{CDPM.1c}$  and  $\oplus_{\text{WRM}_\star} \ominus_{\text{WRM}} \mathbf{CDPM.2'c}$ .

In both systems neither **(PRatMono)** nor **(WRM)** is derivable. This is clearly as desired. However, the price to pay for this is that neither  $\oplus_{\text{WRM}_\star} \ominus_{\text{WRM}} \mathbf{CDPM.1c}$  nor  $\oplus_{\text{WRM}_\star} \ominus_{\text{WRM}} \mathbf{CDPM.2'c}$  is equivalent to **(R)SDDL** if we add (CD) as an axiom.

Both systems do not validate **(DEX-1)**—**(DEX-3)** and are therefore sufficiently robust with respect to deontic conflicts.

Note that in **(R)SDDL** all instances of

$$\vdash (O(B \mid A) \wedge P(B \wedge C \mid A)) \supset O(A \supset B \mid C) \quad (\text{AWRM}_\star)$$

are valid. Where  $\alpha \in \{1, 2'\}$ , if all instances of (CD) are added to  $\oplus(\{\text{WRM}_\star, \text{AWRM}_\star\}, \ominus_{\text{WRM}} \mathbf{CDPM.}\alpha\text{c})$  then the resulting logic is equivalent to **(R)SDDL**.

**THEOREM 4.**  $\oplus(\{\text{WRM}_\star, \text{AWRM}_\star\}, \ominus_{\text{WRM}} \mathbf{CDPM.}\alpha\text{c})$  satisfies **(C★)**, where  $\alpha \in \{1, 2'\}$ .

One of the major problems with restricting **(WRM)** is analogous to the problem we already pointed out in connection with the restricted





inheritance principle (**RCPM**): in order to apply (**WRM<sub>\*</sub>**) the user has to manually add permission statements. However, on the one hand, if the relationships between various obligations are of a complicated nature, the manual addition of permission statements might lead to undesired results such as explosion. On the other hand we would like to delegate as much reasoning as possible from the user to the logic itself. We will therefore in Section 6.3 propose an adaptive logic which applies (**WRM**) “as much as possible” without the need of user interference.

**(C) A problem with aggregation and (S).** Consider the following strict version of the asparagus example:  $O(\neg F \mid \top)$  (1),  $O(F \mid A)$  (2),  $P(A \mid \top)$  (3), and  $P(\neg F \wedge \neg A \mid \top)$  (4). From (2) we get by (S),

$$O(A \supset F \mid \top) \quad (5)$$

Note that it is a consequence of (**RCPM**) that  $\vdash P(C \mid B) \supset (P(C' \mid B) \vee O(C' \mid B))$  where  $C \vdash C'$ . Therefore, by (3),  $P(A \vee F \mid \top) \vee O(A \vee F \mid \top)$ . Since (4) is equivalent to  $\neg O(A \vee F \mid \top)$  we arrive at

$$P(A \vee F \mid \top) \quad (6)$$

In case we apply (**CAND**) to (1) and (5) we get  $O(\neg F \wedge (A \supset F) \mid \top)$  and therefore, by (**CRE**),  $O(\neg F \wedge \neg A \mid \top)$ . But this is equivalent to  $\neg P(A \vee F \mid \top)$  — a contradiction with (6). Employing the weaker aggregation principle (**CPAND'**) doesn't help either. In this case we add the harmless statements  $P(\neg F \mid \top)$  and  $\neg O(A \wedge \neg F \mid \top)$  to the premises. By (**CPAND'**),  $O(\neg F \wedge \neg A \mid \top)$  is again derivable (as the reader can easily verify herself) — in contradiction with (6).

Thus, the example shows that (**CAND**) and (**CPAND'**) do not behave well together with (S). One possible solution is to use a restricted version of (S), namely<sup>9</sup>

$$\vdash (P(C \mid A) \wedge O(C \mid A \wedge B)) \supset O(B \supset C \mid A) \quad (\text{PS})$$

Note that in  $\oplus_{\text{PS}} \ominus_{\text{S}} \mathbf{CDPM.1c}$  and  $\oplus_{\text{PS}} \ominus_{\text{S}} \mathbf{CDPM.2'c}$  the derivation of (5) is blocked since we would need  $P(F \mid \top)$ , but we have  $O(\neg F \mid \top)$  which is equivalent to  $\neg P(F \mid \top)$ . It is also worth mentioning that, where

---

<sup>9</sup>(PS) was proposed by Goble in [5] in connection with another problem w.r.t. (**RatMono**).

$\alpha \in \{1, 2'\}$ ,  $\oplus(\{\text{CD}, \text{PS}\}, \ominus_{\text{S}}\mathbf{CDPM}.\alpha\mathbf{c})$  is equivalent to  $\oplus_{\text{CD}}\mathbf{CDPM}.\alpha\mathbf{c}$  and therefore to **(R)SDDL**.

**THEOREM 5.** *Where  $\alpha \in \{1, 2'\}$ ,  $\oplus_{\text{PS}}\ominus_{\text{S}}\mathbf{CDPM}.\alpha\mathbf{c}$  satisfies  $(\text{C}\star)$ .*

While we so far tackled the problem above by restricting **(S)**, it can be argued that the application of **(S)** leading to **(5)** is not the problem. One motivation would be to preserve one of the traditional monadic representations of condition obligations as much as possible. **(S)** can be seen in this sense as a bridging principle from the dyadic,  $\text{O}(F \mid A)$ , to the monadic representation,  $\text{O}(A \supset F)$ , of conditional obligations. It can also be argued that **(5)**, which is equivalent to  $\neg\text{P}(A \wedge \neg F \mid \top)$  should be derivable since it is intuitive. In view of **(2)**, the obligation to eat with fingers when being served asparagus, one should in general not be allowed to eat asparagus without using fingers. The punchline of this line of argument is that instead of blaming **(S)**, **(CAND)** (resp. **(CPAND')**) is considered to be too strong. It should not be applicable in order to derive  $\text{O}(\neg F \wedge \neg A \mid \top)$ . In the next section it will be demonstrated how this strategy can be followed by use of an adaptive logic which applies aggregation conditionally.

## 6. Going adaptive – the conditional case

After having located various problems of **CDPM.1c** and **CDPM.2'c** we are now going to introduce an adaptive logic framework that can deal with these problems. One of our goals is to develop conflict-tolerant logics that are able to derive from the non-conflicting ‘parts’ of a given premise set as much as possible without the need of manually adding premises. Due to this, the criterion  $(\text{C}\star)$  is not adequate anymore since it measures the derivative power of a logic in view of adding all instances of **(CD)** to it. We alter it in the following way:

**(C‡):** For all premise sets for which **(R)SDDL** is non-explosive, a dyadic deontic logic for dilemmas should be equivalent to **(R)SDDL**.

Note that logics satisfying this criterion are in a sense stronger than logics only satisfying  $(\text{C}\star)$  since, in order to achieve equivalence with **(R)SDDL**, **(CD)** does not have to be added to the former ones.



### 6.1. Applying Inheritance conditionally

In order to apply the dyadic inheritance principle (**RCPM**) conditionally we can proceed in the same way as in the monadic case (see Section 4).

DEFINITION 8.  $\mathbf{A}_c(\mathbf{L})$  is defined by the triple

$$\langle \mathbf{L}, \Omega_d^c, \text{minimal abnormality} \rangle$$

where  $\Omega_d^c =_{\text{df}} \{O(A \mid B) \wedge O(\neg A \mid B) \mid A, B \in \mathcal{W}_2\}$  and  $\mathbf{L} \in \downarrow \mathbf{CDPM.1c} \cup \downarrow \mathbf{CDPM.2'c}$ .

*Example 2.* As we discussed the monadic variant already in Section 4, we are only going to take a look at a very simple example for  $\mathbf{A}_c(\mathbf{CDPM.2'c})$ . Let as usual  $F$  express ‘eating with fingers’,  $N$  ‘using a napkin’ and let  $B$  stand for ‘belching at the table’. We abbreviate  $O(C \mid D) \wedge O(\neg C \mid D)$  by  $!_c O(C \mid D)$ .

1	$O(\neg F \wedge \neg B \mid \top)$	PREM	$\emptyset$
2	$O(\neg B \mid \top)$	1; RC	$\{!_c O(\neg F \wedge \neg B \mid \top)\}$
3	$O(N \mid \top)$	PREM	$\emptyset$
4	$O(\neg F \wedge \neg B \wedge N \mid \top)$	1, 3; RC	$\{!_c O(\neg F \wedge \neg B \mid \top), !_c O(N \mid \top)\}$

In line 2 we have a conditional application of the inheritance principle, in line 4 one of the aggregation principle.

Similar as in the monadic case (see Theorem 2) we have:

THEOREM 6. Where  $\alpha \in \{1, 2'\}$ ,  $\mathbf{A}_c(\mathbf{CDPM.}\alpha\mathbf{c})$  satisfies  $(C\ddagger)$ .

### 6.2. Applying Aggregation conditionally

There are two ways of tackling the problem pointed out in Section 5.2 (C). One is to use a restricted version of (S), namely (PS). The other is to alter the behavior of aggregation. This can be done via an adaptive logic which enables conditional applications of (CAND).

As pointed out in Section 5.2 (C), neither (CAND) nor (CPAND') are acceptable in combination with (S). We will therefore, unlike the way we proceeded for **ADPM.2'c**, fully reject the aggregation principle from our lower limit logic. Instead we will apply the aggregation principle conditionally. We define the following abnormalities:  $\Omega_d^\wedge =_{\text{df}} \{O(A \mid C) \wedge O(B \mid C) \wedge \neg O(A \wedge B \mid C) \mid A, B, C \in \mathcal{W}_2\}$ . Let  $!_{\wedge} O(A \wedge B \mid C) =_{\text{df}}$



$O(A | C) \wedge O(B | C) \wedge \neg O(A \wedge B | C)$ . Propositional logic warrants the following:

$$\begin{aligned} O(A | C) \wedge O(B | C) &\vdash O(A \wedge B | C) \vee !_{\wedge} O(A \wedge B | C) \\ O(A | C) \wedge \neg O(A \wedge B | C) &\vdash \neg O(B | C) \vee !_{\wedge} O(A \wedge B | C) \\ \neg O(A \wedge B | C) &\vdash \neg O(A | C) \vee \neg O(B | C) \vee !_{\wedge} O(A \wedge B | C) \end{aligned}$$

Note that for  $\mathbf{L}$  in  $\downarrow \ominus_{\text{CPAND}} \mathbf{CDPM.2'c}$  validating **(CP)** we have  $\vdash_{\mathbf{L}} (O(A | B) \wedge O(\neg A | B)) \equiv (O(A | B) \wedge O(\neg A | B) \wedge \neg O(A \wedge \neg A | B))$

Therefore the abnormalities in  $\Omega_d^c$  can be considered as a sub-case of the abnormalities in  $\Omega_d^{\wedge}$  for the lower limit logic  $\mathbf{L}$ . Hence,

$$\begin{aligned} O(A | B) &\vdash_{\mathbf{L}} P(A | B) \vee !_{\wedge} O(A \wedge \neg A | B) \\ \text{If } A &\vdash_{\mathbf{L}} B \text{ then } O(A | C) &\vdash_{\mathbf{L}} O(B | C) \vee !_{\wedge} O(A \wedge \neg A | C) \end{aligned}$$

We define the following adaptive logics:

DEFINITION 9. Logic  $\mathbf{A}_{\wedge}^{\mathbf{m}/\mathbf{r}}(\mathbf{L})$  is defined by the triple

$$\langle \mathbf{L}, \Omega_d^{\wedge}, \text{minimal abnormality / reliability} \rangle$$

where lower limit logic  $\mathbf{L}$  belongs to

$$\downarrow \ominus_{\text{CAND}} \mathbf{CDPM.1c} \cup \{ \oplus_{\text{CD} \ominus_{\text{CAND}}} \mathbf{CDPM.1c} \} \cup \downarrow \ominus_{\text{CPAND}} \mathbf{CDPM.2'c} \cup \{ \oplus_{\text{CD} \ominus_{\text{CPAND}}} \mathbf{CDPM.2'c} \}.$$

Here are some conditional derivations for these adaptive logics:

$$\begin{aligned} O(A | B) &\vdash_{\mathbf{A}_{\wedge}^{\mathbf{x}}(\mathbf{L.2})} P(A | B) \text{ on the condition } \{ !_{\wedge} O(A \wedge \neg A | B) \} \\ \text{If } A &\vdash_{\mathbf{L.2}} B, \text{ then } O(A | C) &\vdash_{\mathbf{A}_{\wedge}^{\mathbf{x}}(\mathbf{L.2})} O(B | C) \text{ o.t.c. } \{ !_{\wedge} O(A \wedge \neg A | C) \} \\ O(A | C) &\wedge O(B | C) &\vdash_{\mathbf{A}_{\wedge}^{\mathbf{x}}(\mathbf{L})} O(A \wedge B | C) \text{ o.t.c. } \{ !_{\wedge} O(A \wedge B | C) \} \end{aligned}$$

where  $\mathbf{x} \in \{\mathbf{m}, \mathbf{r}\}$ ,  $\mathbf{L}$  and  $\mathbf{L.2}$  are any of the lower limit logics defined in Definition 9, and  $\mathbf{L.2}$  validates **(CP)**.

THEOREM 7. Where  $\mathbf{x} \in \{\mathbf{m}, \mathbf{r}\}$ , (i)  $\mathbf{A}_{\wedge}^{\mathbf{x}}(\oplus_{\text{CD} \ominus_{\text{CAND}}} \mathbf{CDPM.1c})$  and (ii)  $\mathbf{A}_{\wedge}^{\mathbf{x}}(\ominus_{\text{CPAND}} \mathbf{CDPM.2'c})$  satisfy  $(C_{\dagger}^{\dagger})$ .

*Example 3.* Let us have a look at the problematic example from Section 5.2 (C) again and let the lower limit be for instance  $\ominus_{\text{CAND}} \mathbf{CDPM.1c}$ .

1	$O(\neg F \mid \top)$	PREM	$\emptyset$
2	$O(F \mid A)$	PREM	$\emptyset$
3	$P(A \mid \top)$	PREM	$\emptyset$
4	$P(\neg F \wedge \neg A \mid \top)$	PREM	$\emptyset$
5	$O(A \supset F \mid \top)$	2; <b>S</b>	$\emptyset$
<sup>136</sup>	$O(\neg F \wedge (A \supset F) \mid \top)$	1, 5; RC	$\{\neg \wedge O(\neg F \wedge (A \supset F) \mid \top)\}$
<sup>137</sup>	$O(\neg F \wedge \neg A \mid \top)$	6; <b>CRE</b>	$\{\neg \wedge O(\neg F \wedge (A \supset F) \mid \top)\}$
8	$\neg O(F \vee A \mid \top)$	4; Def	$\emptyset$
9	$P(F \vee A \mid \top) \vee O(F \vee A \mid \top)$	3; <b>RCPM</b>	$\emptyset$
10	$P(F \vee A \mid \top)$	8, 9; DSyl	$\emptyset$
11	$\neg O(\neg F \wedge \neg A \mid \top)$	10; Def	$\emptyset$
12	$\neg O(\neg F \wedge (A \supset F) \mid \top)$	11; <b>CRE</b>	$\emptyset$
13	$\neg \wedge O(\neg F \wedge (A \supset F) \mid \top)$	1, 5, 12; Agg	$\emptyset$

As desired, the logic is able to block the derivation of  $O(\neg F \wedge \neg A \mid \top)$  at line 7.

### 6.3. Applying Weak Rationality conditionally

Furthermore, our logics can be enabled for conditional applications of (**WRM**). Compared with **CDPM.1c** and **CDPM.2'c** the advantage is that in case an application of (**WRM**) leads to unwanted results (e.g. conflicts or explosions) our logics block the application. Unlike the logic presented in Section 5.2 that makes use of a further restriction of (**WRM**), namely (**WRM<sub>\*</sub>**), there is no need to add auxiliary permission statements to the premise set for the adaptive logic that is presented in this section. This accords with our goal to reduce the reasoning and interference of the user as much as possible.

We define the set of abnormalities  $\Omega_d^r = \{O(A \mid B) \wedge P(A \wedge C \mid B) \wedge \neg O(A \mid B \wedge C) \mid A, B, C \in \mathcal{W}_2\}$ . By propositional logic we have

$$O(A \mid B) \wedge P(A \wedge C \mid B) \vdash O(A \mid B \wedge C) \vee (O(A \mid B) \wedge P(A \wedge C \mid B) \wedge \neg O(A \mid B \wedge C)).$$

This enables us to derive  $O(A \mid B \wedge C)$  from  $O(A \mid B)$  and  $P(A \wedge C \mid B)$  on the condition  $\{O(A \mid B) \wedge P(A \wedge C \mid B) \wedge \neg O(A \mid B \wedge C)\}$ .

DEFINITION 10. Logic  $\mathbf{A}_r^{\mathbf{m}/r}(\mathbf{L})$  is defined by the following triple:

$$\langle \mathbf{L}, \Omega_d^r, \text{minimal abnormality / reliability} \rangle$$

where  $\mathbf{L} \in \bigcup_{\alpha \in \{1, 2'\}} \downarrow \ominus_{\text{WRM}} \mathbf{CDPM}.\alpha \mathbf{c} \cup \{\oplus_{\text{CD}} \ominus_{\text{WRM}} \mathbf{CDPM}.\alpha \mathbf{c}\}$ .



**THEOREM 8.** *Where  $\alpha \in \{1, 2'\}$  and  $\mathbf{x} \in \{\mathbf{r}, \mathbf{m}\}$ ,  $\mathbf{A}_r^{\mathbf{x}}(\oplus_{\text{CD}} \ominus_{\text{WRM}} \mathbf{CDPM}.\alpha\mathbf{c})$  satisfies  $(\text{C}\dagger)$ .*

Let us take a look at a few examples.

*Example 4.* Let us return to the problematic asparagus example from Section 5.2 (B) with lower limit logic  $\ominus_{\text{WRM}} \mathbf{CDPM}.\mathbf{1c}$ . Let  $!_r \mathbf{O}_A^B(C) =_{\text{df}} \mathbf{O}(C \mid A) \wedge \mathbf{P}(C \wedge B \mid A) \wedge \neg \mathbf{O}(C \mid A \wedge B)$ .

1	$\mathbf{O}(\neg F \mid \top)$	PREM	$\emptyset$
2	$\mathbf{P}(F \mid A)$	PREM	$\emptyset$
3	$\mathbf{P}(\neg F \wedge A \mid \top)$	PREM	$\emptyset$
6 4	$\mathbf{O}(\neg F \mid A)$	1, 3; RC	$\{!_r \mathbf{O}_\top^A(\neg F)\}$
5	$\neg \mathbf{O}(\neg F \mid A)$	2; RU	$\emptyset$
6	$\mathbf{O}(\neg F \mid \top) \wedge \mathbf{P}(\neg F \wedge A \mid \top) \wedge \neg \mathbf{O}(\neg F \mid A)$	1, 3, 5; RU	$\emptyset$

This demonstrates that unwanted derivations are successfully blocked. In order to show that desired consequences are actually reached, the example needs to be extended. Let  $X$  stand for “being in a country  $C$ ” and we know that eating with fingers is strictly forbidden (no exceptions!) in  $C$ .

7	$\mathbf{O}(\neg F \mid X)$	PREM	$\emptyset$
8	$\mathbf{P}(\neg F \wedge A \mid X)$	PREM	$\emptyset$
9	$\mathbf{O}(\neg F \mid X \wedge A)$	7, 8; RC	$\{!_r \mathbf{O}_X^A(\neg F)\}$

It can easily be shown that line 9 will not be marked in any extension of the proof.

*Example 5.* Also with the other problematic example in Section 5.2 (B) the adaptive logics block the undesired instances of (SA).

1	$\mathbf{O}(\neg S \mid H)$	PREM	$\emptyset$
2	$\mathbf{P}(S \mid R)$	PREM	$\emptyset$
3	$\mathbf{P}(S \mid H \wedge R)$	PREM	$\emptyset$
4	$\mathbf{P}(\neg S \wedge R \mid H)$	PREM	$\emptyset$
8 5	$\mathbf{O}(\neg S \mid H \wedge R)$	1, 4; RC	$\{!_r \mathbf{O}_H^R(\neg S)\}$
8 6	$\mathbf{O}(H \supset \neg S \mid R)$	5; S	$\{!_r \mathbf{O}_H^R(\neg S)\}$
7	$\neg \mathbf{O}(\neg S \mid H \wedge R)$	3; Def	$\emptyset$
8	$!_r \mathbf{O}_H^R(\neg S)$	1, 4, 7; Agg	$\emptyset$



Note that with **CDPM.αc** ( $\alpha \in \{1, 2'\}$ ) the counter-intuitive  $\mathbf{O}(\neg S \mid H \wedge R)$  is derivable from  $\mathbf{O}(\neg S \mid H)$  and  $\mathbf{P}(\neg S \wedge R \mid H)$  by (**WRM**) causing an explosion due to  $\mathbf{P}(S \mid H \wedge R)$ .

#### 6.4. Combining the various systems

This section offers insight in the way the various systems introduced above can be combined, allowing for different degrees of adaptiveness. This is desired since each of our logics so far only treated particular problems of the ones presented in Section 5.2. By combining them we are able to get the most powerful and intuitive systems. The matter is illustrated by means of some interesting selected examples.

**Inheritance and Aggregation.** We have argued in Section 5.2 (C) that in order to keep (**S**) we have to apply the aggregation principle conditionally.

A first proposal for an adaptive approach to inheritance and aggregation, paradigmatically based on **CDPM.1**, is presented in the following definition:

**DEFINITION 11.** Let  $\mathbf{A}_{c,\wedge}^m(\mathbf{L})$  be the adaptive logic defined by the triple  $\langle \mathbf{L}, \Omega_d^c \cup \Omega_d^\wedge, \text{minimal abnormality} \rangle$ , where  $\mathbf{L} \in \downarrow \ominus_{\text{CAND}} \mathbf{CDPM.1c}$ .

However this approach leads to difficulties:

*Example 6.* We take a look at the life-saver-paradox with

$$\mathbf{A}_{c,\wedge}^m(\ominus_{\text{CAND}} \mathbf{CDPM.1c}).$$

1	$\mathbf{O}(A \mid \top)$	PREM	$\emptyset$
2	$\mathbf{O}(B \mid \top)$	PREM	$\emptyset$
3	$\mathbf{O}(\neg(A \wedge B) \mid \top)$	PREM	$\emptyset$
<sup>7</sup> 4	$\neg \mathbf{O}(A \wedge B \mid \top)$	3; RC	$\{\! _c \mathbf{O}(\neg(A \wedge B) \mid \top) \!\}$
<sup>7</sup> 5	$\! _{\wedge} \mathbf{O}(A \wedge B \mid \top)$	1, 2, 4; RU	$\{\! _c \mathbf{O}(\neg(A \wedge B) \mid \top) \!\}$
<sup>7</sup> 6	$\mathbf{O}(A \wedge B \mid \top)$	1, 2; RC	$\{\! _{\wedge} \mathbf{O}(A \wedge B \mid \top) \!\}$
7	$\! _c \mathbf{O}(\neg(A \wedge B) \mid \top) \vee \! _{\wedge} \mathbf{O}(A \wedge B \mid \top)$	5; RU	$\emptyset$

Some readers will be dissatisfied with the fact that line 4 is marked since it is quite intuitive that, in a situation in which we cannot save the life of both people, we should not be obliged to save both.

However, there is a solution. If we first conditionally apply inheritance and then conditionally apply aggregation, the problem is solved.

Let us take a look at the case in which for the second logic the minimal abnormality strategy is used:

**DEFINITION 12.** Logic  $\mathbf{A}_\lambda^m \circ \mathbf{A}_c(\mathbf{L})$  is defined by the consequence relation  $Cn_{\mathbf{A}_\lambda^m(\mathbf{L})}(Cn_{\mathbf{A}_c(\mathbf{L})}(\Gamma))$ , where  $\mathbf{L} \in \downarrow \ominus_{\text{CAND}} \mathbf{CDPM.1c} \cup \downarrow \ominus_{\text{CPAND}'} \mathbf{CDPM.2'c}$ .

As we are dealing with a combined adaptive logic, the marking conditions for the second logic have to be slightly altered.

**DEFINITION 13.**  $\Phi_s^\wedge(\Gamma)$  is the set of minimal choice sets over  $\{\Delta_i \mid i \in I\}$  where  $\{\text{Dab}(\Delta_i) \mid i \in I, \Delta_i \subset \Omega_d^\wedge\}$  is the set of all minimal Dab-formulas derived at stage  $s$  at unmarked lines on a condition  $\Delta' \subset \Omega_d^c$ .

Line  $i$  is marked at stage  $s$  iff, where  $A$  is derived on the condition  $\Delta$  at line  $i$ ,

- (i) there is no  $\varphi \in \Phi_s^\wedge(\Gamma)$  such that  $\varphi \cap \Delta = \emptyset$ , or
- (ii) for some  $\varphi \in \Phi_s^\wedge(\Gamma)$ , there is no line at which  $A$  is derived on a condition  $\Theta$  for which  $\varphi \cap \Theta = \emptyset$ .

It is easy to see that in our example line 6 gets marked due to line 5. However, line 4 and 5 remain unmarked and thus,  $\neg \text{O}(A \wedge B \mid \top)$  is derivable. Furthermore, it can be shown that the very intuitive  $\text{O}(A \vee B \mid \top)$  is finally derivable for both,  $\mathbf{A}_\lambda^m \circ \mathbf{A}_c(\ominus_{\text{CAND}} \mathbf{CDPM.1c})$  and  $\mathbf{A}_\lambda^m \circ \mathbf{A}_c(\ominus_{\text{CPAND}'} \mathbf{CDPM.2'c})$ .

**THEOREM 9.**  $\mathbf{A}_\lambda^x \circ \mathbf{A}_c(\ominus_{\text{CAND}} \mathbf{CDPM.1c})$  and  $\mathbf{A}_\lambda^x \circ \mathbf{A}_c(\ominus_{\text{CPAND}'} \mathbf{CDPM.2'c})$  satisfy  $(C\ddagger)$ , where  $\mathbf{x} \in \{\mathbf{r}, \mathbf{m}\}$ .<sup>10</sup>

**Inheritance and Weak Rationality.** A natural first suggestion for an adaptive logic account of both inheritance and weak rational monotonicity is given by the following logic:

**DEFINITION 14.**  $\mathbf{A}_{\mathbf{r}, \mathbf{c}}^m(\mathbf{L})$  is defined by the following triple

$$\langle \mathbf{L}, \Omega_d^c \cup \Omega_d^r, \text{minimal abnormality} \rangle,$$

where  $\mathbf{L} \in \bigcup_{\alpha \in \{1, 2'\}} \downarrow \ominus_{\text{WRM}} \mathbf{CDPM.}\alpha\mathbf{c}$ .

---

<sup>10</sup>We only gave marking conditions for the cases  $\mathbf{A}_\lambda^m \circ \mathbf{A}_c(\ominus_{\text{CAND}} \mathbf{CDPM.1c})$  and  $\mathbf{A}_\lambda^m \circ \mathbf{A}_c(\ominus_{\text{CPAND}'} \mathbf{CDPM.2'c})$ . However, the marking conditions for  $\Omega_d^\wedge$  in case of the reliability strategy are defined similarly to Definition 16.





However, the following example shows that this logic is suboptimal.

*Example 7.* Suppose the following obligations and permissions:

- If your friends Beth and Mike are around, you're supposed to serve coffee. –  $O(C \mid F)$
- If your friend Anna, who has a coffee allergy, is around, you're supposed not to serve coffee. –  $O(\neg C \mid A)$
- If your friends Beth and Mike are around, you're allowed to serve coffee and to have Anne around. –  $P(C \wedge A \mid F)$
- If your friend Anna is around, you're allowed to have Beth and Mike around and to not serve coffee. –  $P(\neg C \wedge F \mid A)$

The following proof is in  $\mathbf{A}_{r,c}^m(\ominus\text{WRM}\mathbf{CDPM.1c})$ .

1	$O(C \mid F)$	PREM	$\emptyset$
2	$O(\neg C \mid A)$	PREM	$\emptyset$
3	$P(C \wedge A \mid F)$	PREM	$\emptyset$
4	$P(\neg C \wedge F \mid A)$	PREM	$\emptyset$
<sup>9</sup> 5	$O(C \mid F \wedge A)$	1, 3; RC	$\{!_r O_F^A(C)\}$
<sup>9</sup> 6	$O(\neg C \mid F \wedge A)$	2, 4; RC	$\{!_r O_A^F(\neg C)\}$
7	$O(C \mid F \wedge A) \vee O(\neg C \mid F \wedge A)$	5; RU	$\{!_r O_F^A(C)\}$
8	$O(C \mid F \wedge A) \vee O(\neg C \mid F \wedge A)$	6; RU	$\{!_r O_A^F(\neg C)\}$
9	$!_c O(C \mid F \wedge A) \vee !_r O_F^A(C) \vee !_r O_A^F(\neg C)$	5, 6; RU	$\emptyset$

The minimal choice sets are  $\{\{!_c O(C \mid F \wedge A)\}, \{!_r O_F^A(C)\}, \{!_r O_A^F(\neg C)\}\}$ . Note that  $O(C \mid F \wedge A) \vee O(\neg C \mid F \wedge A)$  is derivable. This is undesired. One possible solution is to use the reliability strategy instead of the minimal abnormality strategy. In this case lines 5—8 are marked at line 9. Note however that we have already shown that the reliability strategy leads in the case of the life saver example to undesired consequences with respect to  $\Omega_d^c$ .<sup>11</sup>

A better solution to this problem is again presented in the form of a sequential adaptive logic:

**DEFINITION 15.**  $\mathbf{A}_r^r \circ \mathbf{A}_c(\mathbf{L})$  is the combined adaptive logic defined by the consequence relation  $Cn_{\mathbf{A}_r^r(\mathbf{L})}(Cn_{\mathbf{A}_c(\mathbf{L})}(\Gamma))$  where  $\mathbf{L} \in \bigcup_{\alpha \in \{1,2\}} \downarrow \ominus\text{WRM}\mathbf{CDPM.}\alpha\mathbf{c}$ .

The marking conditions for abnormalities in  $\Omega_d^r$  are defined in a similar way as Definition 13.

---

<sup>11</sup>We have shown this for the monadic case in Section 4. However, this can be easily generalized for the dyadic case.



DEFINITION 16. We define  $U_s^r(\Gamma) = \bigcup_I \Delta_i$  where  $\{\text{Dab}(\Delta_i) \mid i \in I, \Delta_i \subset \Omega_d^r\}$  is the set of all minimal Dab-formulas derived up to stage  $s$  at unmarked lines on a condition  $\Delta'$  such that  $\Delta' \cap \Omega_d^r = \emptyset$ .

Line  $i$  is marked at stage  $s$  iff, where  $\Delta$  is its condition,  $\Delta \cap U_s^r(\Gamma) \neq \emptyset$ .

Now the following line can be added to the proof of Example 7:

$$10 \quad !_r \text{O}_F^A(C) \vee !_r \text{O}_A^F(\neg C) \quad 9; \text{RC} \quad \{!_c \text{O}(C \mid F \wedge A)\}$$

Due to the reliability strategy we have to mark lines 5–8. Note that for minimal abnormality strategy lines 7 and 8 would not be marked at line 10. Thus, reliability strategy is preferable. Note furthermore that, although the reliability strategy is employed for  $\Omega_d^r$  and hence for the conditional applications of (WRM), we are free to use the minimal abnormality strategy for  $\Omega_d^c$  and hence for the conditional applications of (RCPM). Therefore this logic offers a satisfactory account for the life-saver example.

THEOREM 10. *Where  $\alpha \in \{1, 2'\}$ ,  $\mathbf{A}_r^\alpha \circ \mathbf{A}_c(\ominus_{\text{WRM}} \text{CDPM.}\alpha\mathbf{c})$  satisfies  $(\text{C}\ddagger)$ .*

**Inheritance, Aggregation and Weak Rationality.** Here an option is to use the sequential adaptive logic  $\mathbf{A}_r^\alpha \circ \mathbf{A}_\lambda^x \circ \mathbf{A}_c(\mathbf{L})$  defined by the consequence relation  $Cn_{\mathbf{A}_r^\alpha(\mathbf{L})}(Cn_{\mathbf{A}_\lambda^x(\mathbf{L})}(Cn_{\mathbf{A}_c(\mathbf{L})}(\Gamma)))$ , where  $\mathbf{x} \in \{\mathbf{r}, \mathbf{m}\}$  and  $\mathbf{L} \in \downarrow \ominus (\{\text{CAND}, \text{WRM}\}, \text{CDPM.1c}) \cup \downarrow \ominus (\{\text{CPAND}', \text{WRM}\}, \text{CDPM.2'c})$ . The marking conditions for  $\Omega_d^c$  are defined in Definition 2, the marking conditions for abnormalities in  $\Omega_d^\wedge$  are defined in Definition 13 and the ones for abnormalities in  $\Omega_d^r$  are defined in Definition 16.

THEOREM 11.  *$\mathbf{A}_r^\alpha \circ \mathbf{A}_\lambda^x \circ \mathbf{A}_c(\ominus(\{\text{CAND}, \text{WRM}\}, \text{CDPM.1c}))$  and  $\mathbf{A}_r^\alpha \circ \mathbf{A}_\lambda^x \circ \mathbf{A}_c(\ominus(\{\text{CPAND}', \text{WRM}\}, \text{CDPM.2'c}))$  satisfy  $(\text{C}\ddagger)$ , where  $\mathbf{x} \in \{\mathbf{r}, \mathbf{m}\}$ .*

It can be easily shown that for all the examples given in this section this logic has the desired consequences.

## 7. Conclusion

Lou Goble presented in [5, 8] dyadic deontic logics based on a restriction of the inheritance principle. These systems are able to deal with deontic



conflicts in the sense that they are not explosive when facing dilemmas and they block undesired derivations from these dilemmas. Furthermore, by including a restricted version of rational monotonicity, Goble is able to give an intuitive account of a restricted ‘strengthening the antecedent’ (SA) principle.

However, this paper locates various shortcomings of Goble’s logics. Most importantly the treatment of (SA) is counter-intuitive for some examples and can even lead to explosion. Furthermore, some of Goble’s principles do not behave well together.

An adaptive logic framework is developed which enables a satisfactory treatment of these problems. The price to pay for going adaptive is the lack of monotonicity. However, although Goble’s systems are monotonic logics, a defeasible approach is motivated by them due to the restrictiveness of some of their rules. Take for instance the restricted version of the inheritance principle. In order to apply the rule, in many cases the user needs to add permission statements to the premise sets. Especially for complicated premise sets this demands a great deal of reasoning by the user and can lead to explosive behavior in cases in which she is not able to foresee all consequences of the addition of certain premises. A defeasible approach is far more elegant. It not only avoids explosive behavior in such cases but also shifts the reasoning from the user to the logic. It is self-explanatory that one of the basic requirements for an adequate deontic logic is that it does as much reasoning with as little external interference as possible. The adaptive logics are well-suited for this task. By interpreting a premise set as “normally” as possible, the required additional permission statements are generated automatically as part of the proof dynamics. Furthermore these proof dynamics explicate our moral reasoning processes.

While this paper demonstrated that going adaptive improves **CDPM** in terms of elegance, strength and intuitiveness, in [15] the author shows that the adaptive logic approach on basis of **CDPM** is also able to give a solution to another deep problem of dyadic deontic logics, namely the lack of a proper treatment of (defeasible) detachment. The advantages of the logics presented in this paper and in [15] can easily be assembled by forming combined systems in the same manner as was done in Section 6.4.



## Appendix

### Semantics

The interested reader can find semantics for the monadic systems in [9] and [16]. Semantics for some of the dyadic **CDPM** systems, such as for instance **CDPM.1** and **CDPM.1c**, are introduced by Goble in [5]. The semantics for the variations introduced in this paper can be defined along the same lines.

One of the basic ideas for the neighborhood semantics is that propositions are interpreted in terms of sets of worlds. Moreover, each world has associated with it pairs of propositions, i.e., pairs of sets of worlds. The idea is that an obligation  $O(A \mid B)$  is true at a world  $w$ , in case  $\langle B, A \rangle$  is one of its associated pairs of propositions. Let us look the formal details.

Let a dyadic neighborhood frame  $F$  be a pair  $\langle W, \mathcal{O} \rangle$  where  $W$  is a set of worlds and  $\mathcal{O}$  assigns each world  $a \in W$  a set of ordered propositions, i.e.  $\mathcal{O}_a \subseteq \wp(W) \times \wp(W)$ . A model  $M$  on a frame  $F$  is a pair  $\langle F, v \rangle$  where  $v(p) \subseteq W$  for each propositional letter  $p$ . Where  $|A|_M =_{\text{df}} \{a \in W \mid M, a \models A\}$ , we define

$$\begin{aligned}
 M, a \models A &\text{ iff } a \in v(A), \text{ where } A \in \mathcal{S} \\
 M, a \models O(A \mid B) &\text{ iff } \langle |B|_M, |A|_M \rangle \in \mathcal{O}_a \\
 M, a \models \neg A &\text{ iff } M, a \not\models A \\
 M, a \models A \wedge B &\text{ iff } (M, a \models A \text{ and } M, a \models B) \\
 M, a \models A \vee B &\text{ iff } (M, a \models A \text{ or } M, a \models B) \\
 M, a \models A \supset B &\text{ iff } (M, a \models \neg A \text{ or } M, a \models B) \\
 M, a \models \top & \\
 M, a \not\models \top &
 \end{aligned}$$

For a model  $M = \langle W, \mathcal{O}, v \rangle$ ,  $M \models A$  iff  $M, a \models A$  for all  $a \in W$ . Where  $\Gamma \subseteq \mathcal{W}_2^O$  and  $M = \langle F, v \rangle$ , we say that  $M$  is an  $F$ -model of  $\Gamma$  iff  $M \models A$  for all  $A \in \Gamma$ . Moreover, for a frame  $F$ ,  $\Gamma \models_F A$  iff  $M \models A$  for all  $F$ -models of  $\Gamma$  and, for a class of frames  $\mathcal{F}$ ,  $\Gamma \models_{\mathcal{F}} A$  iff  $\Gamma \models_F A$  for all  $F \in \mathcal{F}$ .

Semantics for **CDPM.1c** are defined by means of the following frame conditions. Where  $F = \langle W, \mathcal{O} \rangle$  and  $\overline{X} =_{\text{df}} W \setminus X$ , we require for all  $a \in W$ :

$$\langle W, W \rangle \in \mathcal{O}_a \quad (\text{S-CN})$$



- If  $\langle X, Y \rangle \in \mathcal{O}_a$  and  $\langle X, Z \rangle \in \mathcal{O}_a$ , then  $\langle X, Y \cap Z \rangle \in \mathcal{O}_a$  (S-CAND)  
 If  $Y \subseteq Z$  and  $\langle X, Y \rangle \in \mathcal{O}_a$  and  $\langle X, \overline{Y} \rangle \notin \mathcal{O}_a$  then  $\langle X, Z \rangle \in \mathcal{O}_a$   
 (S-RCPM)  
 If  $\langle X, Y \rangle \in \mathcal{O}_a$ , for any  $Y \subseteq W$ , then  $\langle X, X \rangle \in \mathcal{O}_a$  (S-QR)  
 If  $\langle X \cap Y, Z \rangle \in \mathcal{O}_a$ , then  $\langle X, \overline{Y} \cup Z \rangle \in \mathcal{O}_a$  (S-S)  
 If  $\langle X, Y \rangle \in \mathcal{O}_a$  and  $\langle X, \overline{Y \cap Z} \rangle \notin \mathcal{O}_a$ , then  $\langle X \cap Z, Y \rangle \in \mathcal{O}_a$   
 (S-WRM)

Semantics for **CDPM.2'c** are defined by means of (S-CN), (S-RCPM), (S-QR), (S-S), (S-WRM) and the following frame conditions:

- If  $\langle X, Y \rangle \in \mathcal{O}_a$ ,  $\langle X, Z \rangle \in \mathcal{O}_a$ ,  $\langle X, \overline{Y} \rangle \notin \mathcal{O}_a$ ,  
 and  $\langle X, \overline{Z} \rangle \notin \mathcal{O}_a$ , then  $\langle X, Y \cap Z \rangle \in \mathcal{O}_a$  (S-CPAND)  
 $\langle X, \emptyset \rangle \notin \mathcal{O}_a$  (S-CP)

Moreover, the following frame conditions are useful to define the semantics for some of our logical variants:

- If  $\langle X, Y \rangle \in \mathcal{O}_a$ ,  $\langle X, \overline{Y \cap Z} \rangle \notin \mathcal{O}_a$  and  $\langle Z, Y \cup \overline{X} \rangle \in \mathcal{O}_a$ ,  
 then  $\langle X \cap Z, Y \rangle \in \mathcal{O}_a$  (S-WRM<sub>★</sub>)  
 If  $\langle X \cap Y, Z \rangle \in \mathcal{O}_a$ , and  $\langle X, \overline{Z} \rangle \notin \mathcal{O}_a$ , then  $\langle X, \overline{Y} \cup Z \rangle \in \mathcal{O}_a$  (S-PS)  
 If  $\langle Y, Z \rangle \in \mathcal{O}_a$ , and  $\langle Y, \overline{Z \cap X} \rangle \notin \mathcal{O}_a$ , then  $\langle X, \overline{Y} \cup Z \rangle \in \mathcal{O}_a$   
 (S-AWRM<sub>★</sub>)

**THEOREM 12.** *Where  $\Psi \in \wp(\mathcal{P})$ ,  $\mathbf{L} = \oplus(\Psi, \mathbf{CDPM}^-)$  is strongly complete and sound with respect to the class  $\mathcal{F}$  of frames that meet the conditions in  $\{\text{S-P} \mid \text{P} \in \Psi\}$ , (S-CN) and (S-RCPM), i.e., where  $\Gamma \subseteq \mathcal{W}_2^O$ ,  $\Gamma \models_{\mathcal{F}} A$  iff  $\Gamma \vdash_{\mathbf{L}} A$ .<sup>12</sup>*

In view of Theorem 1 this completeness result generalizes to all the adaptive logics in this paper.

---

<sup>12</sup>Goble offered a rather involved proof of weak completeness and soundness for **CDPM.1c** in [5]. The authors in [14] have proven strong soundness and completeness for all rank-1 modal logics (i.e., logics which are axiomatized by formulas containing exactly one level of modal operators) with respect to their canonical neighborhood semantics. Obviously our logics fall in this category.



### Proofs

PROOF OF THEOREM 3. Goble has proven the statement already for **CDPM.1c**. Note that from (CD) and (**CPAND'**), (**CAND**) is derivable. Since **CDPM.1c** together with (CD) validates (CN), **CDPM.2'c** together with (CD) is equivalent to **CDPM.1c** together with (CD). Thus, **CDPM.2'c** together with (CD) is equivalent to (**R**)**SDDL**.  $\dashv$

PROOF OF THEOREM 4.

Note that  $\oplus(\{\text{WRM}_\star, \text{AWRM}_\star\}, \ominus_{\text{WRM}} \mathbf{CDPM}.\alpha\mathbf{c})$  and **CDPM.αc** are equivalent. Due to (**WRM<sub>★</sub>**) and (**AWRM<sub>★</sub>**), (**WRM**) is valid in the former logic. Due to (**WRM**), (**WRM<sub>★</sub>**), and due to (**WRM**) and (S), (**AWRM<sub>★</sub>**) is valid in the latter logic. The rest is an immediate consequence of the fact that **CDPM.αc** satisfies (**C<sub>★</sub>**).  $\dashv$

LEMMA 1. Where  $\alpha \in \{1, 2'\}$ ,  $\oplus(\{\text{CD}, \text{PS}\}, \ominus_{\text{S}} \mathbf{CDPM}.\alpha\mathbf{c})$  is equivalent to  $\oplus_{\text{CD}} \mathbf{CDPM}.\alpha\mathbf{c}$ .

PROOF. Suppose  $\text{O}(C \mid A \wedge B)$  and  $\neg \text{O}(B \supset C \mid A)$ . By (**PS**)  $\neg \text{P}(C \mid A)$  which is equivalent to  $\text{O}(\neg C \mid A)$ . By  $\text{O}(C \mid A \wedge B)$  and (CD),  $\neg \text{O}(\neg C \mid A \wedge B)$ . This and  $\text{O}(\neg C \mid A)$  results by (**WRM**) in  $\neg \text{P}(B \wedge \neg C \mid A)$  which is equivalent to  $\text{O}(B \supset C \mid A)$ ,—a contradiction. The rest follows immediately.  $\dashv$

PROOF OF THEOREM 5. This is an immediate consequence of Lemma 1 and the fact that **CDPM.αc** satisfies (**C<sub>★</sub>**).  $\dashv$

PROOF OF THEOREM 6. Let  $\Gamma$  be a (**R**)**SDDL**-consistent premise set. As  $\oplus_{\text{CD}} \mathbf{CDPM}.\alpha\mathbf{c}$  is equivalent to (**R**)**SDDL**,  $\mathcal{M}_{\oplus_{\text{CD}} \mathbf{CDPM}.\alpha\mathbf{c}}(\Gamma) \neq \emptyset$ . By definition of  $\Omega_d^c$  these are the minimally abnormal models since for all these models  $M$ ,  $\text{Ab}(M) = \emptyset$  and moreover, for all  $M \in \mathcal{M}_{\mathbf{CDPM}.\alpha\mathbf{c}}(\Gamma) \setminus \mathcal{M}_{\oplus_{\text{CD}} \mathbf{CDPM}.\alpha\mathbf{c}}(\Gamma)$ ,  $\text{Ab}(M) \neq \emptyset$  since  $M$  validates a counter-instance of (CD),  $\text{O}(A \mid B) \wedge \text{O}(\neg A \mid B)$ . Hence **A<sub>c</sub>(CDPM.αc)** is equivalent to  $\oplus_{\text{CD}} \mathbf{CDPM}.\alpha\mathbf{c}$  for  $\Gamma$ . Therefore it is equivalent to (**R**)**SDDL** for all premise sets for which (**R**)**SDDL** is non-explosive.  $\dashv$

PROOF OF THEOREM 7. *Ad* (i): Let  $\Gamma$  be any (**R**)**SDDL**-consistent premise set. Then  $\mathcal{M}_{\oplus_{\text{CD}} \mathbf{CDPM}.\mathbf{1c}}(\Gamma) \neq \emptyset$  since  $\oplus_{\text{CD}} \mathbf{CDPM}.\mathbf{1c}$  is equivalent to (**R**)**SDDL**. Furthermore,

$$\mathcal{M}_{\oplus_{\text{CD}} \mathbf{CDPM}.\mathbf{1c}}(\Gamma) \subseteq \mathcal{M}_{\oplus_{\text{CD}} \ominus_{\text{CAND}} \mathbf{CDPM}.\mathbf{1c}}(\Gamma).$$



As models in  $\mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.1c}}(\Gamma)$  validate **(CAND)**, they do not validate any abnormalities in  $\Omega_d^\wedge$  and are therefore minimally abnormal. Let  $M \in \mathcal{M}_{\oplus_{\text{CD}}\ominus_{\text{CAND}}\text{CDPM.1c}}(\Gamma) \setminus \mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.1c}}(\Gamma)$ . Then there are  $\text{O}(A \mid C)$  and  $\text{O}(B \mid C)$  such that  $M \models \text{O}(A \mid C), \text{O}(B \mid C)$  and  $M \models \neg \text{O}(A \wedge B \mid C)$  since  $M$  does not validate all instances of **(CAND)**. But  $M$  cannot be minimally abnormal then. Hence  $\mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.1c}}(\Gamma)$  is the set of all minimally abnormal models.

*Ad* (ii)  $\mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.2'c}}(\Gamma)$  is a non-empty set of  $\ominus_{\text{CPAND}'}\text{CDPM.2c}$ -models that validate no abnormalities in  $\Omega_d^\wedge$  due to the fact that  $\oplus_{\text{CD}}\text{CDPM.2'c}$  is equivalent to **(R)SDDL** and hence validates **(CAND)**. Let  $M \in \mathcal{M}_{\ominus_{\text{CPAND}'}\text{CDPM.2'c}}(\Gamma) \setminus \mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.2'c}}(\Gamma)$ . In case  $M$  does validate a counter-instance of **(CPAND')** or **(CAND)**,  $M$  has a non-empty set of abnormalities and is thus not minimally abnormal resp. reliable. Assume that  $M$  validates all instances of **(CAND)**. Then there are  $\text{O}(A \mid B)$  and  $\text{O}(\neg A \mid B)$  such that  $M \models \text{O}(A \mid B), \text{O}(\neg A \mid B)$ . But then by **(CAND)**,  $M \models \text{O}(A \wedge \neg A \mid B)$ —a contradiction (due to **(CP)** and **(CRE)**). Hence  $\mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.2'c}}(\Gamma)$  is the set of all minimally abnormal models.  $\dashv$

**LEMMA 2.** *Where  $\alpha \in \{1, 2'\}$  and  $\mathbf{x} \in \{\mathbf{m}, \mathbf{r}\}$ , for all premise sets for which **CDPM.αc** (resp.  $\oplus_{\text{CD}}\text{CDPM.αc}$ ) is non-explosive,  $\mathbf{A}_F^{\mathbf{x}}(\ominus_{\text{WRM}}\text{CDPM.αc})$  is equivalent to **CDPM.αc** (resp.  $\mathbf{A}_F^{\mathbf{x}}(\oplus_{\text{CD}}\ominus_{\text{WRM}}\text{CDPM.αc})$  is equivalent to  $\oplus_{\text{CD}}\text{CDPM.αc}$ ).*

**PROOF.** It is immediately clear that all models in  $\mathcal{M}_{\text{CDPM.αc}}(\Gamma)$  are minimally abnormal  $\ominus_{\text{WRM}}\text{CDPM.αc}$ -models of  $\Gamma$  since they do not validate any abnormalities in  $\Omega_d^r$ . For all  $M \in \mathcal{M}_{\ominus_{\text{WRM}}\text{CDPM.αc}}(\Gamma) \setminus \mathcal{M}_{\text{CDPM.αc}}(\Gamma)$  there is a counter-instance for **(WRM)**. Therefore these  $M$  do validate abnormalities and are therefore not minimally abnormal. Hence  $\mathcal{M}_{\text{CDPM.αc}}(\Gamma)$  is the set of all minimally abnormal (and all reliable) models. The proof is similar for the other case.  $\dashv$

Theorem 8 is an immediate consequence.

**PROOF OF THEOREM 9.** Firstly note that for any **(R)SDDL**-consistent premise sets  $\Gamma$ ,  $\mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.1c}}(\Gamma)$  (resp.  $\mathcal{M}_{\oplus_{\text{CD}}\text{CDPM.2'c}}(\Gamma)$ ) is a non-empty sets of  $\ominus_{\text{CAND}}\text{CDPM.1c}$ -models (resp.  $\ominus_{\text{CPAND}}\text{CDPM.2'c}$ -models) that validate no abnormalities with respect to  $\Omega_d^c$  and  $\Omega_d^\wedge$ . Any model that validates a counter-instance of **(CD)** validates an abnormality w.r.t.  $\Omega_d^c$

and is therefore not minimally abnormal for  $\mathbf{A}_c(\ominus_{\text{CAND}}\mathbf{CDPM.1c})$  and  $\mathbf{A}_c(\ominus_{\text{CPAND}'}\mathbf{CDPM.2'c})$ . Furthermore, all models which validate all instances of (CD) validate no abnormalities w.r.t.  $\Omega_d^c$ . Hence,

$$\mathcal{M}_{\mathbf{A}_c(\ominus_{\text{CAND}}\mathbf{CDPM.1c})}(\Gamma) \subseteq \mathcal{M}_{\oplus_{\text{CD}}\ominus_{\text{CAND}}\mathbf{CDPM.1c}}(\Gamma)$$

and

$$\mathcal{M}_{\mathbf{A}_c(\ominus_{\text{CPAND}'}\mathbf{CDPM.2'c})}(\Gamma) \subseteq \mathcal{M}_{\oplus_{\text{CD}}\ominus_{\text{CPAND}'}\mathbf{CDPM.2'c}}(\Gamma).$$

Any model that validates counter-instances of (CAND) or (CPAND'), validates an abnormality w.r.t.  $\Omega_d^\wedge$ . Models which validate all instances of (CAND) do not validate any abnormalities w.r.t.  $\Omega_d^\wedge$  and are therefore minimally abnormal and reliable. Hence  $\mathcal{M}_{\oplus_{\text{CD}}\mathbf{CDPM.1c}}(\Gamma)$  is the set of all minimally abnormal models (resp. all reliable models) in

$$\mathcal{M}_{\mathbf{A}_\lambda^\times \circ \mathbf{A}_c(\ominus_{\text{CAND}}\mathbf{CDPM.1c})}(\Gamma).$$

Note further that all models in  $\mathcal{M}_{\oplus_{\text{CD}}\mathbf{CDPM.2'c}}(\Gamma)$  validate all instances of (CAND). Thus,  $\mathcal{M}_{\oplus_{\text{CD}}\mathbf{CDPM.2'c}}(\Gamma)$  is the set of all minimally abnormal (resp. all reliable) models in  $\mathcal{M}_{\mathbf{A}_\lambda^\times \circ \mathbf{A}_c(\ominus_{\text{CPAND}'}\mathbf{CDPM.2'c})}(\Gamma)$ . The rest follows due to the fact that  $\oplus_{\text{CD}}\mathbf{CDPM.}\alpha\mathbf{c}$  is equivalent to (R)SDDL.

–

The proofs of theorems 10 and 11 are very similar and are left to the reader.

**THEOREM 13.** *All adaptive logics defined in Section 6 with lower limit logics in  $\downarrow\mathbf{CDPM.1c} \cup \downarrow\mathbf{CDPM.2'c}$  falsify (DEX-1)–(DEX-3).*

**PROOF.** Let  $F = \langle W, \mathcal{O} \rangle$  where  $W = \wp(\mathcal{S})$ . We define  $W' = \{w \in W \mid p_1 \in w\}$  and  $\mathcal{O}_w$  is defined as follows for all  $w \in W$ :  $\mathcal{O}_w = \bigcup_{X \in \Psi} \{\langle X, Y \rangle \mid Y \supseteq X \cap W'\} \cup \{\langle W, \emptyset \rangle\}$  where  $\Psi = \{W'' \subseteq W \mid W'' \cap W' \neq \emptyset\}$ . It is easy to see that  $F$  satisfies the frame conditions for **CDPM.1c**.

The model validates all premise sets  $\Gamma \in \{\{\mathbf{O}(\top \mid \top), \mathbf{O}(\perp \mid \top)\}, \{\mathbf{O}(\top \mid \top), \mathbf{O}(\perp \mid \top), \mathbf{P}(p_2 \mid \top)\}, \{\mathbf{O}(\top \mid \top), \mathbf{O}(\perp \mid \top), \mathbf{O}(p_1 \mid \top), \mathbf{P}(p_1 \mid \top), \mathbf{P}(p_2 \mid \top)\}\}$  and falsifies the following instance of (DEX-3):  $M \models \mathbf{O}(\top \mid \top), \mathbf{O}(\perp \mid \top), \mathbf{O}(p_1 \mid \top), \mathbf{P}(p_1 \mid \top), \mathbf{P}(p_2 \mid \top)$  and  $M \not\models \mathbf{O}(p_2 \mid \top)$ . Thus, it does also not validate (DEX-1) and (DEX-2). Furthermore the model is minimally abnormal for  $\Gamma$  w.r.t.  $\Omega_d^c$ ,  $\Omega_d^r$  and  $\Omega_d^\wedge$  and any





lower limit  $\mathbf{L} \in \downarrow\mathbf{CDPM.1c}$ . For  $\Omega_d^c$  note that the only abnormalities are  $\{\mathbf{O}(A \mid \top) \wedge \mathbf{O}(\neg A \mid \top) \mid \vdash_{\mathbf{L}} \top \equiv A\}$ , but we also have  $\Gamma \vdash_{\mathbf{L}} \mathbf{O}(A \mid \top) \wedge \mathbf{O}(\neg A \mid \top)$  for all these abnormalities. For  $\Omega_d^\wedge$  and  $\Omega_d^r$  the set of abnormalities verified by  $M$  is empty. Therefore, given premises  $\Gamma$ , the model is selected in all adaptive logics based on lower limits in  $\downarrow\mathbf{CDPM.1c}$  that are defined in Section 6.

Let  $F = \langle W, \mathcal{O} \rangle$  where  $W = \wp(\mathcal{S})$ . We define  $W_a = \{w \in W \mid p_1 \notin w, p_2 \notin w\}$ ,  $W_b = \{w \in W \mid p_1 \notin w, p_2 \in w\}$ ,  $W_c = \{w \in W \mid p_1 \in w, p_2 \notin w\}$ ,  $W_d = \{w \in W \mid p_1 \in w, p_2 \in w\}$ , and  $\mathcal{O}_w$  for all  $w \in W$  as follows:  $\mathcal{O}_w = \bigcup_{X \in \Psi} \{\langle X, Y \rangle \mid Y \supseteq X \cap (W_c \cup W_d)\} \cup \{\langle W, W_a \cup W_b \rangle\}$  where  $\Psi = \{W'' \subseteq W \mid W'' \cap (W_c \cup W_d) \neq \emptyset\}$ . It is easy to see that  $F$  is a **CDPM.2'c**-frame.

The model validates all premise sets  $\Gamma \in \{\{\mathbf{O}(p_1 \mid \top), \mathbf{O}(\neg p_1 \mid \top)\}, \{\mathbf{O}(p_1 \mid \top), \mathbf{O}(\neg p_1 \mid \top), \mathbf{P}(p_2 \mid \top)\}, \{\mathbf{O}(p_1 \mid \top), \mathbf{O}(\neg p_1 \mid \top), \mathbf{O}(\top \mid \top), \mathbf{P}(\top \mid \top), \mathbf{P}(p_2 \mid \top)\}\}$  and falsifies the following instance of **(DEX-3)**:  $M \models \mathbf{O}(\top \mid \top), \mathbf{P}(\top \mid \top), \mathbf{O}(p_1 \mid \top), \mathbf{O}(\neg p_1 \mid \top), \mathbf{P}(p_2 \mid \top)$  and  $M \not\models \mathbf{O}(p_2 \mid \top)$ . Thus, it does also not validate **(DEX-1)** and **(DEX-2)**. Furthermore, the model is minimally abnormal for  $\Gamma$  w.r.t.  $\Omega_d^c$ ,  $\Omega_d^r$  and  $\Omega_d^\wedge$  and any lower limit  $\mathbf{L} \in \downarrow\mathbf{CDPM.2'c}$ . For  $\Omega_d^c$  note that the only abnormalities verified by  $M$  are  $\text{Ab}^c(M) = \{\mathbf{O}(A \mid \top) \wedge \mathbf{O}(\neg A \mid \top) \mid \vdash_{\mathbf{L}} A \equiv p_1\}$  and  $\Gamma \vdash_{\mathbf{L}} A$  for all  $A \in \text{Ab}^c(M)$ . For  $\Omega_d^r$  the set of abnormalities verified by  $M$  is empty. For  $\Omega_d^\wedge$  the set of abnormalities verified by  $M$  is  $\text{Ab}^\wedge(M) = \{\mathbf{O}(A \mid \top) \wedge \mathbf{O}(\neg A \mid \top) \wedge \neg \mathbf{O}(A \wedge \neg A \mid \top) \mid \vdash_{\mathbf{L}} A \equiv p_1\}$ . Note that  $\Gamma \vdash_{\mathbf{L}} A$  for all  $A \in \text{Ab}^\wedge(M)$ . Therefore, given premises  $\Gamma$ , the model is selected in all adaptive logics based on lower limits in  $\downarrow\mathbf{CDPM.2'c}$  that are defined in Section 6.  $\dashv$

## References

- [1] Diderik Batens, “A universal logic approach to adaptive logics”, *Logica Universalis* 1 (2007): 221–242.
- [2] Roderick M. Chisholm, “Contrary-to-duty imperatives and deontic logic”, *Analysis* 24 (1963): 33–36.
- [3] Newton Da Costa and Walter Carnielli, “On paraconsistent deontic logic”, *Philosophia* 16, 3 (1986): 293–305.
- [4] J. Forrester, “Gentle murder, or the adverbial samaritan”, *The Journal of Philosophy* 81 (1984): 193–197.

- [5] Lou Goble, “Dilemmas in deontic logic”. To appear.
- [6] Lou Goble, “Multiplex semantics for deontic logic”, *Logic* 5 (2000): 113–134.
- [7] Lou Goble, “Preference semantics for deontic logic, part I – simple models”, *Logique et Analyse* (2003): 383–418.
- [8] Lou Goble, “A proposal for dealing with deontic dilemmas”, pages 74–113 in: A. Lomuscio and D. Nute, editors, *DEON*, volume 3065 of *Lecture Notes in Computer Science*. Springer, 2004.
- [9] Lou Goble, “A logic for deontic dilemmas”, *Journal of Applied Logic* 3 (2005): 461–483.
- [10] John F. Horty, “Nonmonotonic foundations for deontic logic”, , pages 17–44 in: *Defeasible Deontic Logic*, 1997.
- [11] Sarit Kraus, Daniel J. Lehmann, and Menachem Magidor, “Nonmonotonic reasoning, preferential models and cumulative logics”, *Artificial Intelligence* 44 (1990): 167–207.
- [12] Daniel Lehmann and Menachem Magidor, “What does a conditional knowledge base entail?”, *Artificial Intelligence* 55, 1 (1992): 1–60.
- [13] David Lewis, *Counterfactuals*. Blackwell Publishers, December 2000.
- [14] Lutz Schröder and Dirk Pattinson, “Rank-1 modal logics are coalgebraic”, *Journal of Logic and Computation*, December 2008. Forthcoming. Online access <http://dx.doi.org/10.1093/logcom/exn096>.
- [15] Christian Straßer, “A deontic logic framework allowing for factual detachment”. Forthcoming.
- [16] Christian Straßer and Joke Meheus, “Avoiding deontic explosion by contextually restricting modal inheritance”. To appear.
- [17] Bas van Fraassen, “The logic of conditional obligation”, *Journal of Philosophical Logic* 1 (1972): 417–438.

CHRISTIAN STRASSER  
Centre for Logic and Philosophy of Science  
Ghent University (UGent)  
Ghent, Belgium  
[christian.strasser@UGent.be](mailto:christian.strasser@UGent.be)