Two-phase Exception Logic

Leendert W.N. van der Torre and **Yao-Hua Tan**

Abstract

In this paper we propose an exception logic – formalizing reasoning about exceptions. We use this logic to defend two claims. First, we argue that default logic – formalizing reasoning about default assumptions – is an extension of exception logic. A deconstruction argument shows that reasoning about exceptions is one of the first principles of reasoning about default assumptions. Second, we argue that two phases have to be distinguished in reasoning about exceptions, and therefore also in reasoning about default assumptions. We identify two causes of the distinction between two phases, the disjunction rule OR and right weakening RW. This sheds some new light on these 'standard' (according to the Kraus-Lehmann-Magidor paradigm) properties of default inference.

Introduction

In this paper we analyze the conditional logic approach to default logic, the logic that formalizes reasoning about default assumptions. Conditional logic is a popular framework to formalize defeasible reasoning, see e.g. (Delgrande 1988; Geffner & Pearl 1992; Goldszmidt, Morris, & Pearl 1993; Boutilier 1994; Veltman 1996). The conditional sentence "if β (the *antecedent* or condition) then by default α (the *consequent* or conclusion)" is represented in this framework by the formula $\beta > \alpha$, where '$>$' is some kind of implication of conditional logic. We distinguish between two approaches to formalize default reasoning in the conditional logic framework, which we call the strengthening and the weakening approach, respectively. The strengtening approach starts with a weak system (the so-called *core*) in which the conditionals do not have monotony, and adds defeasible reasoning schemes like System Z or Maximum Entropy to strengthen the system such that it has restricted monotony. The weakening approach starts with a strong system that has monotony, and adds constraints on monotony to incorporate defeasibility. In this paper we consider default logics of the weakening approach, which are extensions of exception logics – formalizing reasoning about exceptions. A typical reasoning structure is 'α is an exception, and if α then β is an exception,' which can *loosely* be read as 'by default $\neg\alpha$, and if α then by default $\neg\beta$,' or as 'normally $\neg\alpha$, and if α then normally $\neg\beta$.'

Consider the following example of reasoning about exceptions, which is typical for a class of examples that seems to have been ignored in default logic literature.

$\neg s > \neg g$	If the sun does not shine, then normally I do not sit in my garden,
$s > g$	If the sun shines, then normally I sit in my garden, and
$(s \wedge \neg g) > c$	If the sun shines and I do not sit in my garden, then normally I am behind my computer.

Surprisingly, some default logics do not deal satisfactorily with these sentences, as we illustrate below, because they focus on the normal cases neglecting the exceptional cases. For example, System Z (Pearl 1990) derives the counterintuitive default $((\neg s \wedge g) \vee (s \wedge \neg g \wedge \neg c)) > \neg s$. If only the most exceptional states are considered, then $\neg s$ is preferred because if s is false then only one rule is falsified (rank 1) and if s is true then two rules are falsified (rank 2). However, this violation-counting is highly counterintuitive, because the violation of the first default may be more exceptional than the violation of the latter two defaults. It has often been observed that System Z is too weak (Geffner & Pearl 1992; Goldszmidt, Morris, & Pearl 1993) in the sense that intuitive conclusions are not derived, but the example above illustrates that it is already too strong in the sense that counterintuitive conclusions are derived. Many systems are extensions of System Z, which means that they derive this counterintuitive conclusion too. It is not derived in the logic developed in this paper, which is a informal argument to use the weakening approach instead of the strengthening approach when formalizing default reasoning in the conditional logic framework. In the second part of this paper, we also give a more formal argument for the weakening approach based on a so-called deconstruction argument. But in the following sections we first discuss the main argument of this paper.

Two-phase logic

We argue that two phases are necessary to formalize reasoning about exceptions, and therefore also to formalize reasoning about default assumptions.[1] Otherwise counterintuitive conclusions follow. In the first phase exceptional circumstances are as important as the normal circumstances, whereas in the second phase distinctions between exceptional circumstances are ignored. The distinction between the two phases can be illustrated by the following metaphor. Phase-1 reasoning is what you do when you are told that α is an exception (i.e. that normally $\neg\alpha$), and phase-2 reasoning is what you do when you assume that the uncertain facts are unexceptional (i.e. that the actual world is the most normal world). These two are completely separated. You can know what is normally the case, without acting accordingly. For example, in decision theory reasoning about uncertainty (default reasoning) is combined with reasoning about utilities and a rational agent reasons about exceptional circumstances when such exceptional circumstances have very large risks. This explains that it may be rational to insure yourself against rare but extreme calamities. The following two examples illustrate two problems that occur if the two phases are not discriminated.

Example 1 (Disjunction rule) *Consider a conditional logic of exceptions, in which $\beta > \alpha$ can be read as 'if β is the case, then $\neg\alpha$ is an exception.' The formula $\beta > \alpha$ can* loosely *be read as 'if β is the case then by default α' in a credulous default logic (in a sense that is made precise later in this paper). Assume that the conditional logic validates at least substitution of logical equivalents and the following two inference patterns* Monotony MON *and the* Disjunction *rule* OR *(that enables reasoning by cases),*

$$\text{MON}: \frac{\beta_1 > \alpha}{(\beta_1 \wedge \beta_2) > \alpha} \qquad \text{OR}: \frac{\beta_1 > \alpha, \beta_2 > \alpha}{(\beta_1 \vee \beta_2) > \alpha}$$

and assume as premises $a > p$ and $\neg a > p$.[2] *The problem of this set is that we can derive the counterintuitive $(a \leftrightarrow p) > p$, as illustrated below.*

$$\cfrac{\cfrac{a > p \quad \neg a > p}{\top > p}\ \text{OR}}{(a \leftrightarrow p) > p}\ \text{MON}$$

The conditional $(a \leftrightarrow p) > p$ is considered to be counterintuitive, because it is not grounded in the premises. If $a \leftrightarrow p$ and a (the antecedent of the first premise) are true then p is trivially true, and if $a \leftrightarrow p$ and $\neg a$ (the antecedent of the second premise) are true then p is trivially false. With other words, if $a \leftrightarrow p$ then the first premise cannot be falsified and the second premise cannot be verified. Hence, the two premises do not ground the conclusion that for arbitrary $a \leftrightarrow p$ we have that $\neg p$ is an exception.

Example 2 (Right weakening) *Assume that the conditional logic validates at least substitution of logical equivalents,* MON *and the following two inference patterns* Right Weakening RW *and the* Conjunction *rule* AND*,*

$$\text{RW}: \frac{\beta > \alpha_1}{\beta > (\alpha_1 \vee \alpha_2)} \qquad \text{AND}: \frac{\beta > \alpha_1, \beta > \alpha_2}{\beta > (\alpha_1 \wedge \alpha_2)}$$

and as premises $\top > (a \vee p)$ and $\top > \neg a$. The following derivation shows that the counterintuitive $a > p$ can be derived.

$$\cfrac{\cfrac{\cfrac{\top > (a \vee p) \quad \top > \neg a}{\top > (\neg a \wedge p)}\ \text{AND}}{a > (\neg a \wedge p)}\ \text{MON}}{a > p}\ \text{RW}$$

The conditional $a > p$ is considered to be counterintuitive, because it is not grounded in the premises. If its antecedent a is true, then the first premise is verified and the second one is falsified. The derivation can be blocked by replacing MON *by the following version of* Restricted Monotony RMON*, in which $\Diamond$ is a modal connective and $\Diamond\phi$ is loosely read as 'ϕ is consistent.'*

$$\text{RMON}: \frac{\beta_1 > \alpha, \Diamond(\beta_1 \wedge \beta_2 \wedge \alpha)}{(\beta_1 \wedge \beta_2) > \alpha}$$

The derivation is blocked, because $a > (\neg a \wedge p)$ cannot be derived from $\top > (\neg a \wedge p)$ by RMON*. Unfortunately,*

[1] In fact, these two phases can easily be discriminated in most default logics. For example, in Reiter's default logic (Reiter 1980) there is a distinction between phase-1 default rules and phase-2 extensions, in conditional entailment (Geffner & Pearl 1992) there is a distinction between phase-1 default rules with abnormalities and phase-2 conditionally entailed conclusions, and in Veltman's logic (Veltman 1996) there is a distinction between phase-1 normally defaults and phase-2 presumably defaults. However, these approaches do not have a two-phase *proof theory* as developed in this paper.

[2] Following (Kraus, Lehmann, & Magidor 1990) we believe that giving a certain interpretation to propositional atoms in benchmark examples (like in this case a and p) does not make these examples more readable, but only causes a lot of confusion. However, if the reader insists on an interpretation, then she can read a as 'buying apples' and p as 'buying pears.'

the following derivation shows that the counterintuitive $a > p$ can still be derived in another way.

$$\cfrac{\cfrac{\cfrac{\top > (a \vee p) \quad \top > \neg a}{\top > (\neg a \wedge p)}\ \textsc{and}}{\top > p}\ \textsc{rw}}{a > p}\ \textsc{mon/rmon}$$

The problem in the Examples 1 and 2 above can be solved by a technique, which might look odd at first sight, but which turns out to work well, namely to forbid application of RMON after OR or RW has been applied. We call this the *two-phase approach*. Such a sequencing in derivations is rather unnatural and cumbersome from a proof-theoretic point of view. Surprisingly, the two-phase approach can be obtained very intuitively from a semantic point of view, by combining two usages of a preference ordering in a preference-based semantics. For the two usages we define two different types of conditionals, which we call phase-1 and phase-2 conditionals. Phase-1 conditionals are formalized by strong preferences and evaluated by what we call *Ordering*, a process in which the whole ordering is used to evaluate a formula. Phase-2 conditionals are formalized by weak preferences and evaluated by what we call *Minimizing*, in which the ordering is used to select the minimal elements that satisfy a formula. The minimizing approach is commonly taken in preferential semantics for non-monotonic logics, see for example (Shoham 1988; Kraus, Lehmann, & Magidor 1990; Boutilier 1994). In semantic terms the two-phase approach simply means that first a preference ordering has to be constructed by ordering worlds, and subsequently the constructed ordering can be used for minimization.

Exceptions

Credulous defaults are formalized in the normal modal logic S4. As is well-known, the normal modal system S4 contains the two axioms **T:** $\Box\alpha \rightarrow \alpha$ and **4:** $\Box\alpha \rightarrow \Box\Box\alpha$, and is characterized by reflexive transitive orderings (partial pre-orderings). We define phase-1 defaults as strong preferences (new) and phase-2 defaults as weak preferences (Boutilier 1994). We do not use the popular 'choice functions' or 'sphere semantics' for our conditionals, because they only give a bipartitioning in normal and exceptional elements, whereas we need varying degrees of exceptionality.[3] Intuitively, the formula $\Box p$ can be read as 'it is not more exceptional (at least as normal) that p.' We write $>_\exists$ for the phase-2 defaults, because they refer to *some* equivalence class of preferred worlds (see Proposition 1), in contrast to $>_\forall$ that is defined later in this paper and that refers to *all* equivalence classes of preferred worlds.

Definition 1 *Credulous phase-1 and phase-2 defaults 'if β then by default α,' written as $\beta > \alpha$ and $\beta >_\exists \alpha$ respectively, are defined in S4 as follows.*

$$\begin{array}{lll} \beta > \alpha & =_{\text{def}} & \Box((\beta \wedge \alpha) \rightarrow \Box(\beta \rightarrow \alpha)) \wedge \Diamond(\beta \wedge \alpha) \\ \beta >_\exists \alpha & =_{\text{def}} & \Diamond(\beta \wedge \Box(\beta \rightarrow \alpha)) \end{array}$$

Intuitively, a phase-1 default $\beta > \alpha$ expresses a strict preference of all $\beta \wedge \alpha$ over $\beta \wedge \neg\alpha$.[4] However, a preference of all $\beta \wedge \alpha$ worlds to every $\beta \wedge \neg\alpha$ world would be much too strong, because two independent defaults $\top > \alpha_1$ and $\top > \alpha_2$ would not have a model containing $\neg\alpha_1 \wedge \alpha_2$ and $\alpha_1 \wedge \neg\alpha_2$ worlds. The following proposition shows that this preference is represented by the negative condition that no $\beta \wedge \neg\alpha$ is preferred to a $\beta \wedge \alpha$.[5] The phase-2 default $\beta >_\exists \alpha$ is true iff α is true in an equivalence class of most preferred β worlds of the model, or it eventually becomes true in an infinite descending chain of β worlds (Boutilier 1994).

Proposition 1 *Let $M = \langle W, \leq, V\rangle$ be a Kripke model. We have $M, w \models \beta > \alpha$ iff for all $w_1, w_2 \in W$ such that $w_1 \leq w$, $w_2 \leq w$, $M, w_1 \models \beta \wedge \alpha$ and $M, w_2 \models \beta \wedge \neg\alpha$, it is true that $w_2 \not\leq w_1$, and there is such a world w_1. We have $M, w \models \beta >_\exists \alpha$ iff there is a world $w_1 \in W$ with $w_1 \leq w$ and $M, w_1 \models \beta \wedge \alpha$ such that for all worlds $w_2 \in W$ with $w_2 \leq w$ and $M, w_2 \models \beta \wedge \neg\alpha$, it is true that $w_2 \not\leq w_1$.*

Proof *Follows directly from the definition of $>$ and $>_\exists$ in Definition 1.*

In the introduction we mentioned the counterintuitive System Z derivation of $((\neg s \wedge g) \vee (s \wedge \neg g \wedge \neg c)) > \neg s$ from $\neg s > \neg g$, $s > g$ and $(s \wedge \neg g) > c$. This derivation is not made in the phase-1 and phase-2 logic, because

[3] Similar observations are made by Veltman (Veltman 1996) to solve the inheritance problem.

[4] The following definition shows how phase-1 defaults can be extended with explicit exceptions. Credulous phase-1 defaults with explicit exceptions 'if β then by default α unless γ' are defined by

$$\beta >_{\backslash\gamma} \alpha =_{\text{def}} \Box((\beta \wedge \alpha \wedge \neg\gamma) \rightarrow \Box(\beta \rightarrow \alpha)) \wedge \Diamond(\beta \wedge \alpha)$$

A typical example of such defaults are Reiter's default rules (Reiter 1980), given by $\frac{\beta : \neg\gamma}{\alpha}$, where $\neg\gamma$ is the so-called justification. Due to space limitations, we cannot make a more detailed comparison.

[5] In this paper we do not consider facts, we only consider the derivation of conditionals from conditionals. If we also consider facts, then we have to define the conditional in a bimodal logic by

$$\beta > \alpha =_{\text{def}} \Box_2((\beta \wedge \alpha) \rightarrow \Box_1(\beta \rightarrow \alpha)) \wedge \Diamond_2(\beta \wedge \alpha)$$

where the two modal operators are related by $\Box_2\alpha \rightarrow \Box_1\alpha$. Due to space limitations, we cannot discuss this kind of complications in this paper.

the previous Proposition 1 shows that there can be models of the latter three conditionals in which the $\neg s \wedge g$ worlds and $s \wedge \neg g \wedge \neg c$ worlds are equivalent. The following proposition gives several properties of the phase-1 and phase-2 defaults. In this paper we are interested in **RMon**, **OR** and **RW**.[6]

Proposition 2 *The logic S4 has the following theorems.*

RMon	$(\beta_1 > \alpha \wedge \Diamond(\beta_1 \wedge \beta_2 \wedge \alpha)) \to (\beta_1 \wedge \beta_2) > \alpha$
RAnd	$(\beta > \alpha_1 \wedge \beta > \alpha_2 \wedge \Diamond(\beta \wedge \alpha_1 \wedge \alpha_2)) \to \beta > (\alpha_1 \wedge \alpha_2)$
COR	$(\beta > \alpha_1 \wedge \beta > \alpha_2) \to \beta > (\alpha_1 \vee \alpha_2)$
Trans	$(\gamma > \beta \wedge \beta > \alpha \wedge \Diamond(\alpha \wedge \beta \wedge \gamma)) \to \gamma > (\alpha \wedge \beta)$
OR$_\exists$	$(\beta_1 >_\exists \alpha \wedge \beta_2 >_\exists \alpha) \to (\beta_1 \vee \beta_2) >_\exists \alpha$
RW$_\exists$	$\beta >_\exists \alpha_1 \to \beta >_\exists (\alpha_1 \vee \alpha_2)$

The logic S4 does not *have the following theorems.*

OR	$(\beta_1 > \alpha \wedge \beta_2 > \alpha) \to (\beta_1 \vee \beta_2) > \alpha$
RW	$\beta > \alpha_1 \to \beta > (\alpha_1 \vee \alpha_2)$
NC	$\neg(\beta > \alpha \wedge \beta > \neg\alpha)$
Mon$_\exists$	$\beta_1 >_\exists \alpha \to (\beta_1 \wedge \beta_2) >_\exists \alpha$
And$_\exists$	$(\beta >_\exists \alpha_1 \wedge \beta >_\exists \alpha_2) \to \beta >_\exists (\alpha_1 \wedge \alpha_2)$
NC$_\exists$	$\neg(\beta >_\exists \alpha \wedge \beta >_\exists \neg\alpha)$

Proof *The (non)theorems can be verified by proving (un)satisfiability in the preference-based semantics.*

We now study the relation between the phase-1 and phase-2 defaults. First, Proposition 2 illustrates that they are duals of each other when we consider the properties monotony versus the disjunction rule and right weakening. The following proposition shows that phase-1 defaults are strictly stronger than phase-2 defaults, which is useful in the two-phase approach (which is explained later in this section).

Proposition 3 *The logic S4 has the theorem.*

Rel$_\exists$ $\quad \beta > \alpha \to \beta >_\exists \alpha$

Proof *The theorem can be verified by proving derivability in S4. It is equivalent to the formula*

$$(\Box((\beta \wedge \alpha) \to \Box(\beta \to \alpha)) \wedge \Diamond(\beta \wedge \alpha)) \to \Diamond(\beta \wedge \Box(\beta \to \alpha))$$

The following proposition gives another relation between phase-1 and phase-2 defaults. It shows that a phase-1 default is equivalent to a set of phase-2 defaults, when we impose a constraint on the models.

Proposition 4 *Let $M = \langle W, \leq, V \rangle$ be a Kripke model such that M does not contain duplicate worlds. This means that for all worlds $w_1, w_2 \in W$ such that $w_1 \neq w_2$, there is a propositional α such that $M, w_1 \models \alpha$ and $M, w_2 \not\models \alpha$.*[7] *We have $M, w \models \beta > \alpha$ iff for all β' such that $M, w \models \Box(\beta' \to \beta)$ and $M, w \models \Diamond(\beta' \wedge \alpha)$, we have $M, w \models \beta' >_\exists \alpha$.*

Proof $\Rightarrow$ *Follows directly from* **RMon** *and* **Rel**$_\exists$. $\Leftarrow$ *Every world is characterized by a unique propositional sentence. Let $\overline{w}$ denote this sentence that characterizes world w. Proof by contraposition. If $M, w \not\models \beta > \alpha$, then there is no w_1 such that $M, w_1 \models \alpha \wedge \beta$, or there are $w_1 \leq w, w_2 \leq w$ such that $M, w_1 \models \alpha \wedge \beta$, $M, w_2 \models \neg\alpha \wedge \beta$ and $w_2 \leq w_1$. The first case is trivial, so consider the latter case. Choose $\beta' = \overline{w_1} \vee \overline{w_2}$. The world w_2 is one of the preferred β' worlds, because there are no duplicate worlds. (If duplicate worlds are allowed, then there could be a β' world w_3 which is a duplicate of w_1, and which is strictly preferred to w_1 and w_2.) We have $M, w_2 \not\models \alpha$ and therefore $M, w \not\models \beta' >_\exists \alpha$.*

The latter result is rather surprising for the following reason. At first sight one would expect that the ordering defaults have at least the properties of the minimizing defaults, because the former are *defined* by a set of the latter. In particular, at first sight it seems that the ordering defaults have the disjunction rule and right weakening. A careful analysis of the definitions reveals that the argument is wrong due to the subtle consistency check part of **RMon**.[8]

We now proceed to explain the two-phase approach to defeasible reasoning. The two phases in a defeasible logic correspond to the two types of defaults $>$ and $>_\exists$. Semantically, the first phase corresponds to ordering ($>$) and the second phase to minimizing ($>_\exists$). From a proof theoretic point of view, the first phase corresponds to applying valid inferences of $>$ like RMON, RAND etc, and the second phase corresponds to applying valid inferences of $>_\exists$ like OR and RW. The two-phase reasoning is illustrated by the following two examples. They illustrate two causes of the distinction between the two phases: the disjunction rule and right weakening.

The first example illustrates that the two-phase approach solves the problem in Example 1. It also il-

[6] Another interesting property is default chaining. From 'birds have wings' $b > w$ and 'things with wings fly' $w > f$ we can derive $b > (w \wedge f)$ by **Trans** and $(b \wedge \neg f) > w$ by **RMon**.

[7] From a philosophical point of view, this means that the logical language is expressive enough to distinguish all worlds.

[8] For example, the default $\beta > \alpha_1$ seems equivalent to the set $\{\beta' >_\exists \alpha_1 \mid \Box(\beta' \to \beta)\}$. Phase-2 defaults have weakening, thus it implies $\{\beta' >_\exists (\alpha_1 \vee \alpha_2) \mid \Box(\beta' \to \beta)\}$. The latter set is equivalent to the default $\beta > (\alpha_1 \vee \alpha_2)$. Hence, it seems that the phase-1 default $\beta > \alpha_1$ implies the phase-1 default $\beta > (\alpha_1 \vee \alpha_2)$. However, the implication of $\beta > \alpha_1$ to the set $\{\beta' >_\exists \alpha_1 \mid \Box(\beta' \to \beta)\}$ is not valid due to the consistency check part of **RMon**.

lustrates that the non-validity of OR in the first phase can be used to analyze dominance arguments. A common sense dominance argument (1) divides possible outcomes into two or more exhaustive, exclusive cases, (2) points out that in each of these alternatives it is better to perform some action than not to perform it, and (3) concludes that this action is best unconditionally. Thomason and Horty (Thomason & Horty 1996) observe that, although such arguments are often used, and are convincing when they are used, they are invalid. In the second phase OR is accepted, because it is read as 'by default α given β and by default α given $\neg\beta$, then by default α without examining β.'

Example 3 (Disjunction rule, continued) *Let $S = \{a > p, \neg a > p\}$ be the S4 theory of Example 1. The solution of the problem in the two-phase default logic is that the application of* RMON *is blocked after* OR *has been applied. We have $S \not\models \top > p$ and $S \models \top >_\exists p$, $S \not\models (a \leftrightarrow p) > p$ and $S \not\models (a \leftrightarrow p) >_\exists p$. The derivable $\top >_\exists p$ expresses that p is true in the most normal world. It cannot be used to derive the counterintuitive $(a \leftrightarrow p) >_\exists p$, because $>_\exists$ does not have monotony. In fact, the model represented below shows that we can have the opposite $M \models (a \leftrightarrow p) > \neg p$. This figure should be read as follows. Every circle is a nonempty set of worlds, satisfying the propositions written in the circle. The arrows represent strict accessibility. The transitive closure is left implicit. Only worlds accessible from the actual world are represented.*

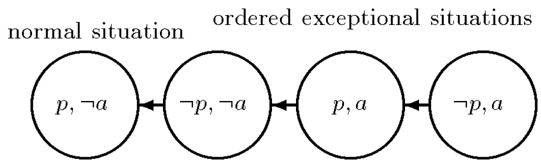

Figure 1: Model of $\{a > p, \neg a > p\}$.

The following example shows that the problem in Example 2 is also solved.

Example 4 (Weakening, continued)
Let $S = \{\top > (a \vee p), \top > \neg a\}$ be a S4 theory, where $\neg a$ does not entail $\neg p$. Reconsider the counterintuitive derivations in Example 2. The solution of the problem in the two-phase default logic is that the application of RMON *is blocked after* RW *has been applied. We have $S \models \Diamond(\neg a \wedge p)$, $S \models \top > (\neg a \wedge p)$ and $S \models \top >_\exists (\neg a \wedge p)$, $S \not\models \top > p$ and $S \models \top >_\exists p$. The crucial observation is that $a >_\exists p$ is not entailed by S, as represented by the blocked derivations below. A dashed line represents a blocked derivation step.*

$$\dfrac{\top > (a \vee p) \quad \top > \neg a}{\top > (\neg a \wedge p)}\ \text{RAND}$$
$$- - - - - -\ (\text{RW})$$
$$\dfrac{\dfrac{\top > p}{a > p}\ \text{RMON}}{a >_\exists p}\ \text{REL}$$

$$\dfrac{\dfrac{\dfrac{\top > (a \vee p) \quad \top > \neg a}{\top > (\neg a \wedge p)}\ \text{RAND}}{\top >_\exists (\neg a \wedge p)}\ \text{REL}}{\top >_\exists p}\ \text{RW}$$
$$- - - - - -\ (\text{RMON})$$
$$a >_\exists p$$

$$\dfrac{\top > (a \vee p) \quad \top > \neg a}{\top > (\neg a \wedge p)}\ \text{RAND}$$
$$- - - - - -\ (\text{RMON})$$
$$\dfrac{\dfrac{a > (\neg a \wedge p)}{a >_\exists (\neg a \wedge p)}\ \text{REL}}{a >_\exists p}\ \text{RW}$$

First of all, $a >_\exists p$ is not entailed by S via $\top > p$, because $\top > p$ is not entailed by S. Secondly, $a >_\exists p$ is not entailed by S via $\top >_\exists p$ either, because $>_\exists$ does not have monotony at all. Thirdly, it is not entailed by S via $a > (\neg a \wedge p)$, because $a > (\neg a \wedge p)$ is not entailed by $\top > (\neg a \wedge p)$ due to the restriction in RMON. *A typical model of S is represented in Figure 2 below.*

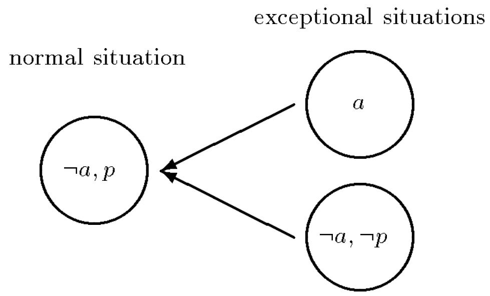

Figure 2: Model of $S = \{\top > (a \vee p), \top > \neg a\}$

The inference patterns OR and RW are standard properties of the Kraus-Lehmann-Magidor paradigm (Kraus, Lehmann, & Magidor 1990). However, the last two examples show that they are not as unproblematic as they seem at first sight, because the properties conflict with monotony. Now we have introduced our exception logic, we give an extensive argument in the following two sections to make the relation between exception logic and default logic clear. This is important, because the main use of exception logic is as a part of default logic.

Deconstruction

We use a deconstruction argument to argue that reasoning about exceptions is a first principle of reasoning about default assumptions. The crucial consequence is that reasoning about default assumptions like reasoning about exceptions contains two phases. A deconstruction argument assumes that to construct an inference relation it is useful to first decompose the logical entailment relation in small building blocks (the destruction) (Nienhuys-Cheng, van der Laag, & van der Torre 1993). In our deconstruction argument we argue that

1. reasoning about default assumptions consists of the three levels exception handling, conflict detection and conflict resolution,
2. reasoning about exceptions gives rise to factual defeasibility, in the sense that an exception logic only formalizes factual defeasibility, and
3. a logic of reasoning about exceptions can be extended with conflict defeasibility and overridden defeasibility to constitute a fully-fledged default logic.[9]

The argument is represented in Figure 3 below. Thus far, we introduced an exception logic that formalizes factual defeasibility, step 2. In the following section we discuss the destruction step of the deconstruction argument, and in the last section we discuss the third step, the construction step of the deconstruction argument.

Three levels of defeasibility

In this section we argue that reasoning about default assumptions consists of the three levels exception handling, conflict detection and conflict resolution. We illustrate the three levels of defeasible reasoning in the conditional logic framework, and we argue that each level has to deal with a different type of defeasibility. The major issue in the conditional logic framework is the derivation of conditionals from other conditionals. Defeasibility is represented by restrictions on the property *Monotony* (or *Strengthening of the Antecedent*), represented either by the inference pattern MON or by the formula **Mon**: $(\beta_1 > \alpha) \rightarrow ((\beta_1 \wedge \beta_2) > \alpha)$, regardless whether the logic is monotonic or non-monotonic. In fact, the level-1 logic proposed in this paper is monotonic and the level-2 logic is non-monotonic, whereas neither of them has **Mon** as a theorem. In the following three examples we show three reasons why unrestricted monotony cannot be accepted by default logics. The first example illustrates that MON has to be restricted, because it is incompatible with conflict resolution mechanisms like specificity. Specificity is often considered to be the most interesting of the conflict resolution principles, because it is founded in probability theory. Moreover, the main reason for the popularity of the conditional logic framework for defeasible reasoning is that conditional logic has a kind of 'built in' specificity, in the sense that MON is in general *not* valid.

[9]Usually, level-2 *conflict detection* makes a conflict inconsistent and level-3 *conflict resolution* uses a mechanism to restore consistency. For example, conditional entailment (Geffner & Pearl 1992) uses abnormalities at level-1, identifies conflicts at level-2 by making assumptions about the minimal worlds, and resolves conflicts at level-3 by introducing priorities (to deal with specificity). Most default logics attempt to solve all three levels at once, which makes the logics quite complicated and difficult to analyze. In particular, it is difficult to *classify* default logics, because they can differ at each level. For example, Veltman's normally-presumably logic (Veltman 1996) has the same first two phases as conditional entailment, but differs in phase-3 where a different construction is used to formalize specificity. Reiter's default logic, on the other hand, uses a different way of dealing with exceptions at level-1 (explicit exceptions), as noted in Footnote 4.

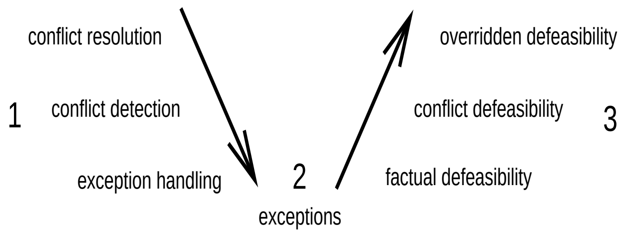

Figure 3: Deconstruction

Example 5 (Specificity) *Assume a conditional defeasible logic that validates at least substitution of logical equivalents, the inference patterns* OR, RW *and* AND *and the* Strong No-Conflict axiom **SNC**.

$$\textbf{SNC}: (\beta > \alpha_1 \wedge \beta > \alpha_2 \wedge \Diamond(\beta \wedge \alpha_1) \wedge \Diamond(\beta \wedge \alpha_2)) \rightarrow \Diamond(\beta \wedge \alpha_1 \wedge \alpha_2)$$

From the 'birds fly' default $b > f$ the 'penguins fly' default $(b \wedge p) > f$ should not be derivable, because 'penguins do not fly' $(b \wedge p) > \neg f$. We cannot accept MON, *because $(b \wedge p) > f$ and $(b \wedge p) > \neg f$ are inconsistent with* **SNC**. *However, from the 'birds fly' default $b > f$ the 'red birds fly' default $(r \wedge b) > f$ should be derivable and from the 'birds have wings' default $b > w$ the 'penguins have wings' default $(b \wedge p) > w$ should be derivable, although 'penguins are exceptional birds' $b > \neg p$. Hence, we have to weaken* MON *to the following inference pattern* Restricted Monotony $\text{RMON}_{\text{Spec}}$ *with a condition C_{Spec} that represents the specificity condition, but that allows for the latter two derivations.*

$$\text{RMON}_{\text{Spec}}: \frac{\beta_1 > \alpha, C_{\text{Spec}}}{(\beta_1 \wedge \beta_2) > \alpha}$$

The two desired derivations of Example 5 are known as the irrelevance problem (red birds fly) and the inheritance problem (penguins have wings) of the conditional logic framework. Well-known solutions of the irrelevance problem are Delgrande's irrelevance principle (Delgrande 1988) and mechanisms equivalent to System Z (Pearl 1990) like rational closure, the minimum specificity principle of possibilistic logic and Boutilier's 'only knowing' construction (Boutilier 1994). Solutions of the inheritance problem are the maximum entropy approach (Geffner & Pearl 1992; Goldszmidt, Morris, & Pearl 1993) and Veltman's update semantics (Veltman 1996). We call these problems level-3 problems, because they are related to conflict resolution mechanisms like specificity.

Level-3 problems are not the only reason we cannot accept MON. From now on we accept MON to show two level-2 problems of MON. Note that the specificity condition C_{Spec} of $\text{RMON}_{\text{Spec}}$ is a kind of conflict resolution mechanism, because the two expressions $(b \wedge p) > f$ and $(b \wedge p) > \neg f$ derivable by MON conflict with axiom **SNC**. Hence, by accepting MON (at level 2!) we accept that $\{b > f, (b \wedge p) > \neg f\}$ represents an unresolved (inconsistent) conflict. The following example illustrates the second reason why MON cannot be accepted.

Example 6 (No-Conflict) *Assume the inference patterns* MON, OR, RW, AND, *the axiom* **SNC**, *and as premises the defaults* $\top > p_1$ *and* $\top > p_2$, *where* $\top$ *stands for any tautology. The intuitively consistent defaults can be strengthened to respectively the defaults* $\neg(p_1 \wedge p_2) > p_1$ *and* $\neg(p_1 \wedge p_2) > p_2$, *which conflict with* **SNC**. *As a solution,* **SNC** *can be replaced by the following* Weak No-Conflict axiom **WNC**.

$$\textbf{WNC} : (\beta > \alpha_1 \wedge \beta > \alpha_2 \wedge \Diamond(\beta \wedge \alpha_1) \wedge \Diamond(\beta \wedge \alpha_2)) \rightarrow \Diamond(\alpha_1 \wedge \alpha_2)$$

However, now consider the intuitively consistent 'quakers are pacifists' default $q > p$ *and the 'republicans are not pacifists' default* $r > \neg p$ *(known as the Nixon diamond). The defaults can be strengthened to* $(q \wedge r) > p$ *and* $(q \wedge r) > \neg p$, *which conflict with* **WNC**.

Example 6 shows that MON cannot be accepted (and thus has to be restricted) because of no-conflict axioms. However, the level-2 problems related to no-conflict axioms are again not the most fundamental problem of monotony. Example 7 shows that even in cases without specificity or no-conflict problems there are reasons not to accept MON in a defeasible logic.

Example 7 (Exceptions) *Reconsider Example 1 and 2 in which we assumed the inference patterns* MON, OR, RW, *and* AND. *We do not accept any no-conflict axiom like* **SNC** *or* **WNC**. *Nevertheless, we showed that we can still derive counterintuitive conclusions, namely* $(a \leftrightarrow p)) > p$ *from the two defaults* $a > p$ *and* $\neg a > p$, *and* $a > p$ *from the two defaults* $\top > (a \vee p)$ *and* $\top > \neg a$.

We call the derivations of $(a \leftrightarrow p) > p$ and $a > p$ level-1 problems. The level-2 problems only occur in sceptical inference relations. We therefore refer to sceptical and credulous defaults instead of level-2 and level-1 defaults.

The three different types of problems discussed in this section give rise to three different reasons to restrict the inference pattern monotony. That is, they give rise to three different types of defeasibility. We call these types respectively factual defeasibility (for level-1), conflict defeasibility (for level-2) and overridden defeasibility (for level-3), see Figure 3. Factual defeasibility is caused by facts, because exceptions are facts. The other types of defeasibility are caused by other conditionals, because there is a conflict. The distinction between conflict defeasibility and overridden defeasibility is that in the latter case the conflict between two conditionals has been resolved. In such a case, we say that the weakest conditional is overridden.

Conflict detection and resolution

For space reasons in this section we only *sketch* how the exception logic can be extended to a default logic, i.e. how no-conflict axioms and conflict resolution mechanisms can be added. Sceptical (i.e. level-2) defaults are defined in the modal preference logic (the sceptical phase-2 default is from (Lamarre 1991; Boutilier 1994)). The sceptical phase-2 default $\beta >_\forall \alpha$ is true in a model if α is true in all most preferred β worlds (and eventually become true in every infinite descending chain of β worlds), and the sceptical phase-1 default $\beta >_s \alpha$ is a combination of the credulous phase-1 default and the sceptical phase-2 default (trivially validating the theorem relating the two phases $\mathbf{Rel}_\forall$ $\beta >_s \alpha \rightarrow \beta >_\forall \alpha$).

Definition 2 *Sceptical phase-1 and phase-2 defaults 'if* β *then by default* α*,' written as* $\beta >_s \alpha$ *and* $\beta >_\forall \alpha$ *respectively, are defined in S4 as follows.*

$$\begin{aligned} \beta >_\forall \alpha &=_{\text{def}} \Box(\beta \rightarrow \Diamond(\beta \wedge \Box(\beta \rightarrow \alpha))) \\ \beta >_s \alpha &=_{\text{def}} \beta > \alpha \wedge \beta >_\forall \alpha \end{aligned}$$

We define a preference ordering on models which prefers models which are maximally connected with respect to the partial pre-ordering $\leq$. Given the preference ordering on models, we can define a notion of preferential entailment, see (Shoham 1988; Kraus, Lehmann, & Magidor 1990). The preferred models of the ordering are the only models which are used for minimization. They are the maximal ignorant models with respect to the *credulous* obligations, i.e. the definition is based on the maximum entropy principle (Goldszmidt, Morris, & Pearl 1993).

Definition 3 *Let $M_1 = \langle W_1, \leq_1, V_1 \rangle$ and $M_2 = \langle W_2, \leq_2, V_2 \rangle$ be two S4 models. M_1 is preferred to M_2 for mapping τ, written as $M_1 \sqsubseteq_\tau M_2$, iff (1) τ is a one-to-one mapping of the worlds of W_2 to the worlds of W_1 such that the worlds satisfy the same propositions, and (2) if $w_1 \leq_2 w_2$ for $w_1, w_2 \in W_2$ then $\tau(w_1) \leq_1 \tau(w_2)$. We write $M_1 \sqsubset_\tau M_2$ iff $M_1 \sqsubseteq_\tau M_2$ and $M_2 \not\sqsubseteq_{\tau^{-1}} M_1$. A world $w \in W$ preferentially satisfies S, written as $M, w \models_\sqsubset S$, iff $M, w \models S$ and there is not a model M' and a mapping τ such that $M', \tau(w) \models S$ and $M' \sqsubset_\tau M$ (M is a preferred model of S). S preferentially entails ϕ, written as $S \models_\sqsubset \phi$, iff for all M and w, if $M, w \models_\sqsubset S$ then $M, w \models \phi$.*

The logic of $>_s$ and $>_\forall$ is not a level-3 default logic, because the 'penguin' theory $S = \{b >_s f, (b \wedge p) >_s \neg f\}$ is inconsistent (as can easily be shown). The inconsistency of S shows that specificity is not yet incorporated in the two-phase approach. A simple solution to incorporate specificity is to build a conflict resolution mechanism on top of the two-phase logic. For example, we can introduce a prioritization on the phase-1 defaults (Geffner & Pearl 1992). However, we agree with (Veltman 1996) that questions of priority should be decided at the level of semantics. For example, the fact that $b >_s f$ can be overridden by $(b \wedge p) >_s \neg f$ is enforced by what these rules mean. It is not something to be stipulated over and above the semantics – as most theories would have it – but something to be explained by it. One way to incorporate specificity is to reformulate Veltman's 'dynamic' defaults (Veltman 1996) in our 'static' framework. This can be accomplished by replacing the single preference ordering $\leq$ by a *set* of preference orderings called a frame, and by defining when defaults are applicable (which explains why a rule is sometimes overruled by other rules).

Conclusions

In this paper different usages of preference orderings for defeasible conditional logics are discussed. The different usages, so-called minimizing and ordering, are represented by different modal operators. It is shown that for the adequate representation of some examples a combination of these operators is needed. Each operator validates different inference rules. Hence, the combination of different modal operators imposes restrictions on the proof theory of the logic. The restriction discussed in this paper is that a proof rule can be blocked in a derivation due to the fact that another proof rule has been used earlier in the derivation. We call this the two-phase approach in the proof theory.

The logics discussed in this paper are closely related to the dynamic interpretation of defaults in the preference-based default logic of Veltman (Veltman 1996). Veltman also uses ordering and minimizing defaults, but he uses a different syntax (for example, he does not define conditional minimizing defaults). $\mathbf{0}$[*normally* p] $\vdash$ *presumably* $(p \vee q)$ means that after ordering all worlds by preferring p to $\neg p$, $p \vee q$ is true in the preferred worlds. A detailed comparison is subject of further research.

References

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

Delgrande, J. 1988. An approach to default reasoning based on a first-order conditional logic: revised report. *Artificial Intelligence* 36.

Geffner, H., and Pearl, J. 1992. Conditional entailment: bridging two approaches to default reasoning. *Artificial Intelligence* 53:209–244.

Goldszmidt, M.; Morris, P.; and Pearl, J. 1993. A maximum entropy approach to nonmonotonic reasoning. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15:220–232.

Kraus, S.; Lehmann, D.; and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence* 44:167–207.

Lamarre, P. 1991. S4 as the conditional logic of nonmonotonicity. In *Proceedings of the KR'91*, 357–367.

Nienhuys-Cheng, S.-H.; van der Laag, P.; and van der Torre, L. 1993. Constructing refinement operators by deconstructing logical implication. In Torasso, P., ed., *Advances in Artificial Intelligence: Proceedings of the AI*IA'93*, 178–189. Springer Verlag, LNAI 728.

Pearl, J. 1990. System Z: a natural ordering of defaults with tractable applications to default reasoning. In *Proceedings of the TARK'90.* San Mateo: Morgan Kaufmann.

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

Shoham, Y. 1988. *Reasoning About Change.* MIT Press.

Thomason, R., and Horty, R. 1996. Nondeterministic action and dominance: foundations for planning and qualitative decision. In *Proceedings of the TARK'96*, 229–250. Morgan Kaufmann.

Veltman, F. 1996. Defaults in update semantics. *Journal of Philosophical Logic* 25:221–261.