Credulous Reasoning About Defaults*

Yao-Hua Tan† Leendert W.N. van der Torre‡

May 7, 1997

Abstract

In this paper we propose a preference-based conditional logic for credulous reasoning about defaults. A conditional default "if β then by default α" is either formalized by the strong preference "$\beta \wedge \neg\alpha$ is not preferred to or equivalent to $\beta \wedge \alpha$", or by the weak preference "the preferred β is an α". We show that these two expressions, instances of what we call the *ordering* and *minimizing* usages of preference orderings, can be considered as duals of each other. Moreover, we give a formalization of ordering and minimizing in Boutilier's modal logic CT40 and we show how to combine them in a two-phase default logic.

1 Introduction

Conditional logic is a popular framework to formalize defeasible reasoning [Del88, AM91, Alc94, Bou94a, Mor95]. The conditional sentence "if β (the antecedent or condition) then by default α (the consequent or conclusion)" is represented in this approach by the formula $\beta > \alpha$, where '$>$' is the implication of conditional logic. The popularity of this framework is based on the lack of strengthening of the antecedent of the conditional implication, which is used to formalize the specificity principle. For example, from the 'birds fly' default $b > f$ the 'penguins fly' default $(p \wedge b) > f$ cannot be derived. However, the lack of strengthening of the antecedent also has drawbacks, in particular the so-called irrelevance problem. For example, from the 'birds fly' default $b > f$ the 'red birds fly' default $(r \wedge b) > f$ cannot be derived. In the framework, the set of defaults derivable in the conditional logic is called the 'conservative core' and conditional formulas are added to this core by Delgrande's irrelevance

*This research was partially supported by the ESPRIT III Basic Research Project No.6156 DRUMS II and the ESPRIT III Basic Research Working Group No.8319 MODELAGE.

†Erasmus University Research Institute for Decision and Information Systems (EURIDIS). Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. E-Mail: {ytan,ltorre}@@euridis.fbk.eur.nl. Tel: (+31)10-4082601. Fax: (+31)10-4526134. Http://www.euridis.fbk.eur.nl/Euridis/welcome.html.

principle [Del88] or mechanisms equivalent to system Z [Pea90] like rational closure [LM92], the minimum specificity principle [BDP92] or Boutilier's 'only knowing' construction [Bou92]. However, these solutions suffer from a specific instance of the irrelevance problem called the inheritance problem. For example, given that penguins are exceptional birds $b > \neg p$, from the 'birds have wings' default $b > w$ the 'penguins have wings' default $(b \wedge p) > w$ cannot be derived by system Z.

In this paper, we propose a conditional logic for credulous reasoning about defaults, i.e. a logic in which the two defaults $\top > p$ and $\top > \neg p$ can consistently exist together, where $\top$ stands for any tautology. Because it is a credulous logic, we can accept strengthening of the antecedent to solve the irrelevance and inheritance problems. Obviously, we loose the possibility to formalize specificity by accepting unrestricted strengthening of the antecedent. However, the following example shows another problem with unrestricted strengthening of the antecedent.

Example 1 (apples and pears problem) *Consider the defaults that (1) you normally buy apples or pears, and (2) you normally do not buy apples. It is counterintuitive to derive that if you buy apples, then you normally buy pears. Assume a conditional default logic that validates at least substitution of logical equivalents and the following Gentzen-style inference patterns* Strengthening of the Antecedent (SA)*,* Weakening of the Consequent (WC) *and* Conjunction (AND)*.*

$$\text{SA}: \frac{\beta_1 > \alpha}{(\beta_1 \wedge \beta_2) > \alpha}$$

$$\text{WC}: \frac{\beta > \alpha_1}{\beta > (\alpha_1 \vee \alpha_2)}$$

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

Furthermore, assume as premises the defaults $\top > (a \vee p)$ *and* $\top > \neg a$*, where* a *can be read as "buying apples" and* p *as "buying pears". The intuitive default* $\top > (\neg a \wedge p)$ *can be derived by* AND*. From the latter default, the default* $a > (\neg a \wedge p)$ *can be derived by* SA*. Unfortunately, from this default, the counterintuitive default* $a > p$ *can be derived by* WC*. This default is considered to be counterintuitive, because it is not grounded in the premises. If* a *is true, then the first premise is fulfilled and the second one is violated. This inference can be blocked by replacing unrestricted strengthening of the antecedent by the following version of restricted strengthening of the antecedent,*[1] *in which* $\Diamond$ *is a modal*

[1] This restriction in strengthening of the antecedent is uncontroversial, because the expression $p > q$ with $p \wedge q$ inconsistent is counterintuitive for logics of defeasible reasoning. However, in some other usages of conditionals and preference logics like deontic logic, such expressions do have an intuitive (and important) reading, see e.g. [TvdT96].

operator and $\Diamond\phi$ is true for all consistent propositional formulas ϕ.

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

The default $a > (\neg a \wedge p)$ cannot be derived from the default $\top > (\neg a \wedge p)$ by RSA*. Unfortunately, the counterintuitive $a > p$ can still be derived in another way. From the intuitive default $\top > (\neg a \wedge p)$ the intuitive $\top > p$ can be derived by* WC*. From this latter obligation, the counterintuitive $a > p$ can be derived by* RSA*. Both derivations are represented in Figure 1.*

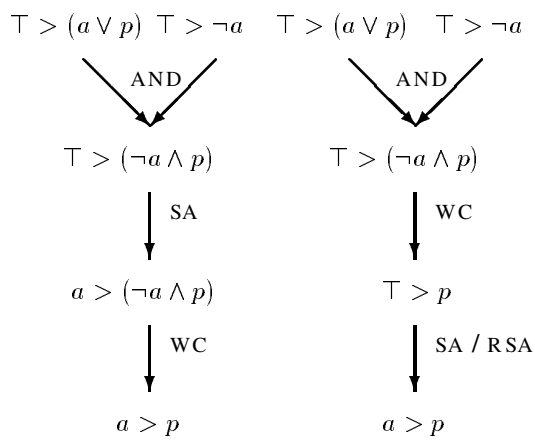

Figure 1: apples and pears problem

The apples and pears problem 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 RSA after WC has been applied. This means that in derivations first RSA has to be applied, and only afterwards WC may be applied. We call this the *two-phase approach* in default logic. 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.

In this paper we show that the two-phase approach can be obtained by combining two usages of a preference ordering in a preference-based semantics of a default logic. For the two usages we define two different types of default conditionals, which we call type-1 and type-2 defaults. The two types of defaults correspond to two different ways to evaluate formulas in a preference ordering. Type-1 defaults 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. Type-2 defaults 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 [Sho88, KLM90, Mak93, Bou94a]. In this paper, we consider credulous reasoning about defaults. In such reasoning, there can be several equivalence classes of preferred models. Hence, the definition of a type-2 default only considers truth in an equivalence class of preferred models instead of truth in all preferred models. In the two-phase approach, the first phase corresponds to ordering, and the second phase corresponds to minimizing. 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.

In this paper we formalize type-1 and type-2 defaults in Boutilier's modal preference logic CT40 [Bou92]. The logic CT40 and the minimizing approach are well-known, but the logic of ordering and the two-phase approach to defeasible reasoning introduced in this paper are new.

This paper is organized as follows. In Section 2 we give the preference logic in which we formalize type-1 and type-2 defaults as strong and weak preferences. We show that these preferences are instances of the ordering and minimizing usages of preference orderings. In Section 3 we show how ordering and minimizing can be combined in a two-phase default logic and how this solves the problems of Example 1. Finally, in Section 4 we briefly consider the sceptical case and in Section 5 we mention some related research.

2 A logic for reasoning about defaults

Preference-based default logics are default logics of which the semantics contains a preference ordering (usually on worlds of a Kripke style possible world model). This preference ordering reflects different degrees of 'normality': a world is preferred to another world if it is, in some sense, more normal than the other world. For example, a value can be associated with each world; in such cases, the ordering is connected (for all w_1 and w_2 we have $w_1 \leq w_2$ or $w_2 \leq w_1$). However, in general the preference ordering can be any partial pre-ordering. Hence, only reflexivity and transitivity are assumed. An expression "by default p" is expressed by a preference for p, which may mean that

1. "p is preferred to $\neg p$ regardless of other things", or that
2. "p is preferred to $\neg p$ other things being equal", or
3. some intermediate reading.

Many authors (for example [TP94, Bou94a]) take the second (*ceteris paribus*) reading, because the first reading does not allow for two or more unconditional preference statements to exist consistently together, as observed by von Wright in [vW63]. For example, the preferences for p and q will quickly run into conflict when considering the worlds $p \wedge \neg q$ and $\neg p \wedge q$.

2.1 Ordering

In this paper, a strong preference for p means that "$\neg p$ is not preferred to or equivalent to p, regardless of other things". Obviously, in a connected partial pre-ordering, the expression "$\neg p$ is not preferred to or equivalent to p" is equivalent to "p is preferred to $\neg p$". Hence, for connected orderings, which are quite popular in preferential semantics, our reading has the problem described by von Wright. However, we do not restrict ourselves to connected orderings, but we allow any partial pre-ordering. With such orderings, the preferences for p and q will not run into conflict when considering the worlds $p \wedge \neg q$ and $\neg p \wedge q$: these worlds are only incomparable. Notice that the whole ordering is taken into account when a default is evaluated. That is why we call it the *ordering* approach to default logic. Similarly, a conditional default "if q then by default p" is represented by "no $\neg p \wedge q$ is preferred to or equivalent to some $p \wedge q$, regardless of other things".

The preferences are formalized in Boutilier's logic CT40, for the details and completeness results of this logic see [Bou94b]. CT40 is a bimodal propositional logic of inaccessible worlds. Boutilier notes that many applications of preference logics (like the one we describe below) do not need the complexity of inaccessible worlds, but it makes the definitions easier and the semantics clearer.

Definition 1 (Syntax of CT40) *The logic CT40 is a bimodal system with the two normal modal connectives* $\Box$ *and* $\overleftarrow{\Box}$*. The dual 'possibility' connectives are defined as usual:*

$\Diamond\alpha =_{def} \neg\Box\neg\alpha$ *and* $\overleftarrow{\Diamond}\alpha =_{def} \neg\overleftarrow{\Box}\neg\alpha$*.*

Moreover, the two following modal connectives are defined:

$\overleftrightarrow{\Box}\alpha =_{def} \Box\alpha \wedge \overleftarrow{\Box}\alpha$ *and* $\overleftrightarrow{\Diamond}\alpha =_{def} \Diamond\alpha \vee \overleftarrow{\Diamond}\alpha$*.*

CT40 is axiomatized by the following set of axioms and inference rules.

K $\Box(\alpha \rightarrow \beta) \rightarrow (\Box\alpha \rightarrow \Box\beta)$
K′ $\overleftarrow{\Box}(\alpha \rightarrow \beta) \rightarrow (\overleftarrow{\Box}\alpha \rightarrow \overleftarrow{\Box}\beta)$
T $\Box\alpha \rightarrow \alpha$
4 $\Box\alpha \rightarrow \Box\Box\alpha$
H $\overleftrightarrow{\Diamond}(\Box\alpha \wedge \overleftarrow{\Box}\beta) \rightarrow \overleftrightarrow{\Box}(\alpha \vee \beta)$

Nec *From* α *infer* $\overleftrightarrow{\Box}\alpha$
MP *From* $\alpha \rightarrow \beta$ *and* α *infer* β

(Semantics of CT40) *Kripke models* $M = \langle W, \leq, V\rangle$ *for CT40 consist of* W*, a set of worlds,* $\leq$*, a binary transitive and reflexive accessibility relation, and* V*, a valuation of the propositions in the worlds. The modal operator* $\Box$ *refers to accessible worlds and the modal operator* $\overleftrightarrow{\Box}$ *to inaccessible worlds.*

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

Given this modal preference logic, we define type-1 defaults as strong preferences.

Definition 2 *Type-1 defaults "if* β *then by default* α*", written as* $\beta > \alpha$*, are defined as follows.*

$$\beta > \alpha =_{def} \overleftrightarrow{\Box}((\beta \wedge \alpha) \rightarrow \Box(\beta \rightarrow \alpha))$$

Intuitively, a default $q > p$ expresses a strict preference of all $p \wedge q$ over $\neg p \wedge q$. The following proposition shows that this preference is represented by a negative condition: no $\neg p \wedge q$ is preferred to a $p \wedge q$,

Proposition 1 *Let* $M = \langle W, \leq, V\rangle$ *be a CT40 model.* $M, w \models \beta > \alpha$ *iff for all* $w_1, w_2 \in W$ *such that* $M, w_1 \models \beta \wedge \neg\alpha$ *and* $M, w_2 \models \beta \wedge \alpha$*, it is true that* $w_1 \not\leq w_2$*.*

Proof $\Rightarrow$ *Assume a model* $M = \langle W, \leq, V\rangle$ *with two worlds* $w_1, w_2 \in W$ *such that* $M, w_1 \models \beta \wedge \neg\alpha$*,* $M, w_2 \models \beta \wedge \alpha$ *and* $w_1 \leq w_2$*. We have* $M, w_2 \not\models (\beta \wedge \alpha) \rightarrow \Box(\beta \rightarrow \alpha))$*. Hence,* $M, w \not\models \beta > \alpha$ *for some world* $w \in W$*, because from the semantic definitions follows immediately that* $M, w \models \overleftrightarrow{\Box}\phi$ *iff for all worlds* $w' \in W$*,* $M, w' \models \phi$*.*

$\Leftarrow$ *Assume* $M, w \not\models \beta > \alpha$ *for some world* w*. Hence, there is a world* $w_2 \in W$ *such that* $M, w_2 \not\models (\beta \wedge \alpha) \rightarrow \Box(\beta \rightarrow \alpha))$*. It follows that* $M, w_2 \models \beta \wedge \alpha$ *and there is a world* $w_1 \in W$ *such that* $M, w_1 \models \beta \wedge \neg\alpha$ *and* $w_1 \leq w_2$*.*

Notice that the normality ordering is global (in the sense that the normality ordering is *not* relative to a world) and nested operators therefore do *not* have an intuitive reading, although they have a formal meaning in CT40. The following example illustrates the definition of type-1 defaults as strong preferences.

Example 2 *Let* $|\alpha|$ *denote a world that satisfies* α*. Assume models that consist of four worlds* $|p \wedge q|$*,* $|\neg p \wedge q|$*,* $|p \wedge \neg q|$ *and* $|\neg p \wedge \neg q|$*. A model satisfies the default* $\top > p$ *when neither* $|\neg p \wedge q|$ *nor* $|\neg p \wedge \neg q|$ *is preferred to* $|p \wedge q|$ *or*

$|p \wedge \neg q|$. *For example, the Kripke model* $|p \wedge q| < |p \wedge \neg q| < |\neg p \wedge \neg q| < |\neg p \wedge q|$ *in Figure 2 is a model for* $\top > p$ *but not for* $\top > q$. *Note that* $\top > q$ *is not true, because* $|\neg p \wedge \neg q| < |\neg p \wedge q|$ *and* $|p \wedge \neg q| < |\neg p \wedge q|$. *This shows how in the ordering approach the whole ordering is taken into account in the evaluation of a formula, and not just the most preferred* $|p \wedge q|$ *worlds.*

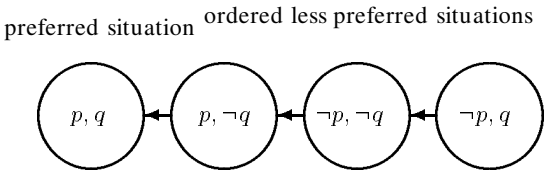

Figure 2: Preference relation with four worlds

In the beginning of this section, we mentioned von Wright's problem that absolute preferences for p and q are not mutually consistent, and the solution in our logic. This solution is illustrated in the following example.

Example 3 *Let* M *be a model of CT40 that satisfies the two defaults* $\top > p_1$ *and* $\top > p_2$ *and contains* $|p_1 \wedge \neg p_2|$ *and* $|\neg p_1 \wedge p_2|$ *worlds. The* $|p_1 \wedge \neg p_2|$ *and* $|\neg p_1 \wedge p_2|$ *worlds of* M *are incomparable, because the first is more normal than the latter with respect to default* $\top > p_1$, *but less normal with respect to default* $\top > p_2$. *An example of such a model* M *is represented in Figure 3.*

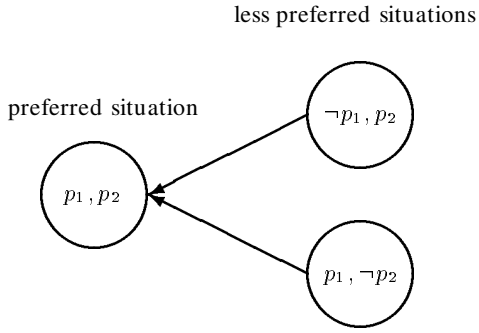

Figure 3: Preference relation with von Wright's problem

The following proposition gives several properties of the type-1 defaults.

Proposition 2 *The logic CT40 validates the following theorems of* Strengthening of the Antecedent (**SA**), Conjunction (**And**) *and* Disjunction (**Or**), *and a version of* Transitivity (**Trans**').

SA	$\beta_1 > \alpha \rightarrow (\beta_1 \wedge \beta_2) > \alpha$
And	$(\beta > \alpha_1 \wedge \beta > \alpha_2) \rightarrow \beta > (\alpha_1 \wedge \alpha_2)$
Or	$(\beta > \alpha_1 \wedge \beta > \alpha_2) \rightarrow \beta > (\alpha_1 \vee \alpha_2)$
Trans'	$(\gamma > \beta \wedge \beta > \alpha) \rightarrow \gamma > (\alpha \wedge \beta)$

The logic CT40 does not *validate the following theorems of* Weakening of the Consequent (**WC**)*,* Transitivity (**Trans**) *and the* sceptical axiom (**D**)*.*

WC	$\beta > \alpha_1 \rightarrow \beta > (\alpha_1 \vee \alpha_2)$
Trans	$(\gamma > \beta \wedge \beta > \alpha) \rightarrow \gamma > \alpha$
D	$\neg(\beta > \alpha \wedge \beta > \neg\alpha)$

Proof *The (non)theorems can easily be verified by proving (un)derivability in CT40.*

The validity of strengthening of the antecedent follows from the fact that a strong preference of p over $\neg p$ implies a preference of $p \wedge q$ over $\neg p \wedge q$. The most remarkable property of the logic is the invalidity of weakening of the consequent. Intuitively, the lack of weakening of $\top > p$ to $\top > (p \vee q)$ is the consequence of the fact that $\top > p$ expresses a preference of *all* p over $\neg p$, because from such a preference does not follow that $p \vee q$ is *always* preferred to $\neg p \wedge \neg q$. This is illustrated by the following example.

Example 4 *Reconsider the Kripke model in Figure 2. The model satisfies* $\top > p$ *but not* $\top > (p \vee q)$*. This illustrates that the type-1 defaults do not have strengthening of the antecedent.*

Transitivity, expressed by the theorem **Trans'**, is an unusual and remarkable property for default logics. The following example illustrates this property.

Example 5 *Let* S *be the set of conditional defaults* $\{\top > a, a > t\}$*, where* a *can be read as 'a certain man going to the assistance of his neighbors' and* t *as 'telling the neighbors that he will come'. Hence, the two defaults can be read as 'a certain man normally goes to the assistance of his neighbors', and 'if he goes, then he normally tells them he is coming'. The default* $\top > (a \wedge t)$ *can be derived from* S *with* **Trans'***, which expresses that normally, the man goes to the assistance of his neighbors* and *he tells them he is coming.*

2.2 Minimizing

Type-2 defaults are defined as weak preferences in the modal preference logic. In this paper, we consider credulous reasoning about defaults. In credulous reasoning, there can be several 'extensions' or equivalence classes of preferred models. Hence, the definition of a type-2 default only considers truth

in an equivalence class of preferred models instead of truth in all preferred models. For the details of this definition, see [Bou94b].[2]

Definition 3 *Type-2 defaults "if β then by default α", written as $\beta >_\exists \alpha$, are defined as follows.*

$$\beta >_\exists \alpha =_{def} \Diamond\!\!\!

Example 7 *Reconsider the Kripke model of Figure 2. The model satisfies* $\top >_\exists q$ *but not* $\neg p >_\exists q$*. For example,* $\neg p >_\exists q$ *is false because the preferred* $|\neg p|$ *worlds are the* $|\neg p \wedge \neg q|$ *worlds. Hence,* $>_\exists$ *does not have strengthening of the antecedent.*

3 Two-phase approach to default logic

3.1 Combining ordering and minimizing

The idea of combining ordering and minimizing is to combine formulas with $>$ and $>_\exists$ operators, where ordering should be strictly stronger than minimizing. However, this combination is not satisfactory in the logic CT40, because we cannot derive $\beta >_\exists \alpha$ from $\beta > \alpha$. The following proposition gives the relation between the two operators.

Proposition 4 *The logic CT40 validates the following theorem.*

$$\beta > \alpha \wedge \overset{\leftrightarrow}{\Diamond} (\beta \wedge \alpha) \rightarrow \beta >_\exists \alpha$$

Proof *The theorem can easily be verified by proving derivability in CT40. It is equivalent to the following formula:*

$$(\overset{\leftrightarrow}{\Box} ((\beta \wedge \alpha) \rightarrow \Box(\beta \rightarrow \alpha)) \wedge \overset{\leftrightarrow}{\Diamond} (\beta \wedge \alpha)) \rightarrow \overset{\leftrightarrow}{\Diamond} (\beta \wedge \Box(\beta \rightarrow \alpha))$$

Hence, when $\overset{\leftrightarrow}{\Diamond} (\alpha \wedge \beta)$ is false, then ordering is not stronger than minimizing. For example, it can easily be verified that the logic CT40 validates the theorem $\alpha > \bot$ but it does not validate $\alpha >_\exists \bot$. In the following definition, $\beta >^c \alpha$ has an additional condition which works like a 'consistency check' to test whether $\alpha \wedge \beta$ is possible.

Definition 4 *Consistent type-1 defaults "if* β *then by default* α*", written as* $\beta >^c \alpha$*, are defined as follows.*

$$\beta >^c \alpha =_{def} \beta > \alpha \wedge \overset{\leftrightarrow}{\Diamond} (\beta \wedge \alpha)$$

This new type of ordering is strictly stronger than minimizing, as is shown in the following proposition.

Proposition 5 *The logic CT40 validates the following theorem.*

$$\beta >^c \alpha \rightarrow \beta >_\exists \alpha$$

Proof *Follows directly from Definition 4 and Proposition 4.*

The type-1 defaults $\beta >^c \alpha$ validate weaker versions of the theorems of Proposition 2, like for example the following Restricted Strengthening of the Antecedent (**RSA**) and Restricted Conjunction (**RAnd**).

RSA $\quad \beta_1 >^c \alpha \wedge \overset{\infty}{\Diamond} (\beta_1 \wedge \beta_2 \wedge \alpha) \rightarrow (\beta_1 \wedge \beta_2) >^c \alpha$

RAnd $\quad \beta >^c \alpha_1 \wedge \beta >^c \alpha_2 \wedge \overset{\infty}{\Diamond} (\beta \wedge \alpha_1 \wedge \alpha_2) \rightarrow \beta >^c (\alpha_1 \wedge \alpha_2)$

We already saw restricted strengthening of the antecedent in Example 1 in the introduction. We elaborate on Example 1 in Section 3.2. To strengthen the theorems above, we consider only models in which all propositionally satisfiable formulas ϕ are true in *some* world. This can be 'axiomatized' with Boutilier's axiom scheme **LP**, see [Lev90, Bou94b] for a discussion. The axiom scheme **LP** states that every formula ϕ without any occurrences of modal operators, which is propositionally satisfiable, is true in some world.

Definition 5 *The logic CT40* is CT40 extended with the following axiom scheme:*
LP : $\overset{\infty}{\Diamond} \phi$ for all satisfiable propositional ϕ

We write $\models$ *for logical entailment in CT40*.*

The following example illustrates the logic CT40* and the idea of combining ordering and minimizing.

Example 8 *Let S be the set of defaults $\{\top >^c (\neg r \wedge \neg g), g >^c r, r >^c g\}$. The intended model is given in Figure 4. We have $S \not\models \top >^c \neg r$, $S \not\models \top >^c \neg g$, $S \models \top >_\exists \neg r$ and $S \models \top >_\exists \neg g$.*

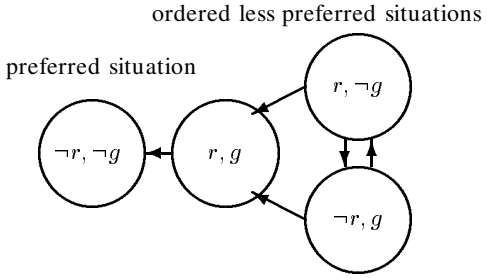

Figure 4: Preference relation

3.2 The two phases in a default logic

The two phases in a default logic correspond to the two different kinds of defaults $>^c$ and $>_\exists$. Semantically, the first phase corresponds to ordering ($>^c$)

and the second phase to minimizing ($>_\exists$). From a proof theoretic point of view, the first phase corresponds to applying valid inferences of $>^c$ like RSA, RAND etc, and the second phase corresponds to applying valid inferences of $>_\exists$ like WC. The basic technique of default logic as a two-phase logic is that a conclusion of the form $\beta >_\exists \alpha$ can be derived either with or without $\beta >^c \alpha$. In the first case $\beta >_\exists \alpha$ can be derived via $\beta >^c \alpha$ with Proposition 5, which says that the latter formula implies the first one. If so, we say that $\beta >_\exists \alpha$ is derived in the first phase. In the second case we say that $\beta >_\exists \alpha$ is second phase derived. The important difference is that in the first phase we can apply RSA to $\beta >_\exists \alpha$, because of the simultaneous occurrence of $\beta >^c \alpha$. We apply RSA to $\beta >^c \alpha$ to obtain, for example, $(\beta \wedge \gamma) >^c \alpha$, and then due to Proposition 5 we also obtain $(\beta \wedge \gamma) >_\exists \alpha$. If $\beta >^c \alpha$ does not occur simultaneously with $\beta >_\exists \alpha$, then there is no way we can apply RSA to this formula. Being a minimizing formula it lacks RSA. Hence, once $\beta >_\exists \alpha$ has been derived in the second phase, we loose RSA permanently for subsequent derivations of this formula. Analogously, we can say that $\beta >_\exists \alpha$ is first phase or second phase entailed by a set of premises, depending on whether S does or does not entail $\beta >^c \alpha$.

The following example shows that the two-phase approach solves the apples and pears problem of Example 1 in the introduction.

Example 9 (apples and pears problem, continued) *Let* $S = \{\top >^c (a \vee p), \top >^c \neg a\}$, *where* $\neg a$ *does not entail the negation of* p. *We have* $S \models^{\Diamond\!\!\!\!-}_{\Diamond} (\neg a \wedge p)$, $S \models \top >^c (\neg a \wedge p)$ *and* $S \models \top >_\exists (\neg a \wedge p)$, $S \not\models \top >^c p$ *and* $S \models \top >_\exists p$. *The crucial observation is that* $a >_\exists p$ *is not entailed by* S. *First of all,* $a >_\exists p$ *is not first phase entailed by* S *via* $\top >_\exists p$, *because* $\top >_\exists p$ *is not first phase entailed by* S. *Secondly,* $a >_\exists p$ *is not second phase entailed by* S *via* $\top >_\exists p$ *either, because in second phase entailment* $>_\exists$ *does not have strengthening of the antecedent at all. Thirdly, it is not second phase entailed by* S *via a first phase derivation of* $a >_\exists (\neg a \wedge p)$, *because* $a >^c (\neg a \wedge p)$ *is not entailed by* $\top >^c (\neg a \wedge p)$ *due to the restriction in* RSA.

4 Further research

Preference semantics are used in several different areas. In fact, the notions 'ordering' and 'minimizing' were introduced in a *deontic logic* based on preferences [TvdT96]. Although there are various subtle differences between these areas, see [Mak93] for a survey, we think that (at least) the following two discussions within deontic logic are relevant for default logic. First of all, the discussion about the validity of the deontic **D** axiom (see e.g. [TvdT96]) is relevant for sceptical reasoning about defaults. Secondly, defeasible deontic logic [vdTT95] and its multi-preference semantics [TvdT95] can be relevant to formalize specificity.

We can consider the two-phase approach with sceptical reasoning about defaults. The obvious candidate for the sceptical type-2 default is "a conditional α is true in *all* most preferred $|\beta|$ worlds", which we write as $\beta >_\forall \alpha$. The following definition of this default in CT40* is from [Bou94b].

Definition 6 *Type-2 sceptical defaults "if β then by default α", written as $\beta >_\forall \alpha$, are defined as follows.*

$$\beta >_\forall \alpha =_{def} \stackrel{\leftrightarrow}{\Box} (\beta \rightarrow \Diamond(\beta \wedge \Box(\beta \rightarrow \alpha)))$$

However, to combine $>_\forall$ with $>^c$, the latter has to be strictly stronger than the other. The following example illustrates that this type-2 default cannot satisfactorily be combined with the type-1 default $\beta >^c \alpha$, because the ordering of worlds can be too weak.

Example 10 *Consider the default $\top >^c p$. All models that satisfy $|p| \not\leq |\neg p|$ are models of $\top >^c p$. Hence, $|p|$ worlds and $|\neg p|$ worlds are either incomparable, or $|p|$ worlds are strictly preferred to $|\neg p|$ worlds. Let M be a model in which all $|p|$ and $|\neg p|$ worlds are incomparable. M satisfies $\top >^c p$, but it does not satisfy $\top >_\forall p$. Hence, $\top >_\forall p$ is not entailed by $\top >^c p$. For minimization, we only want the models in which $|p|$ worlds are strictly preferred to $|\neg p|$.*

A solution of the previous problem is to define a preference ordering on models, which prefers models which are maximally connected with respect to the partial pre-ordering $\leq$, i.e. with the most binary relations of $\leq$. The preferred models of this ordering are the only models which are used for minimization.[3]

Definition 7 *Let $M_1 = \langle W_1, R_1, \leq_1, V_1 \rangle$ and $M_2 = \langle W_2, R_2, \leq_2, V_2 \rangle$ be two CT40* 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 $M_2 \not\sqsubseteq_{\tau^{-1}} M_1$.

The ordering on models ($\sqsubseteq$) should not be confused with the ordering on worlds ($\leq$). The ordering on models is a technical trick to ensure that the worlds within a model are maximally connected, whereas the ordering on worlds expresses the normality ordering. Given the preference ordering on models, we can define a notion of preferential entailment, see [Sho88, KLM90].

[3] In [Bou92], maximally connected models related to system Z are models in which *only* the premises are known, and are formalized with Levesque's 'only knowing' (alias 'all-I-know') operator [Lev90]. However, system Z defines a unique preferred model, whereas in our case there are many distinct preferred models. Hence, we cannot simply copy this 'only knowing' concept.

Definition 8 *Let $M = \langle W, R, \leq, V \rangle$ be a model and S be a set of sentences. A world $w \in W$ of M 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 following example illustrates the notion of preferential entailment.

Example 11 *Reconsider the set of defaults $S = \{\top >^c (\neg r \wedge \neg g), g >^c r, r >^c g\}$ of Example 8. Without the preference ordering on models, the $|\neg r \wedge \neg g|$ and $|r \wedge g|$ worlds could be incomparable. Such a model M would still satisfy $M \models \top >_{\exists} \neg r$ but it would not satisfy $M \models \top >_{\forall} \neg r$.*

Preferential entailment is a typical mechanism from non-monotonic reasoning. The combination of ordering and minimizing is non-monotonic, as the following example illustrates.

Example 12 *Let $S = \{\top >^c p\}$ and $S' = \{\top >^c p, \top >^c \neg p\}$. We have $S \models_{\sqsubset} \top >_{\forall} p$ and $S' \not\models_{\sqsubset} \top >_{\forall} p$. Hence, by addition of a formula we loose conclusions.*

Unfortunately, the previous example also shows that ordering is not stronger than minimizing, a condition for the two-phase approach. For this reason, we define a new type-1 default $>_D$ for the sceptical case, which trivially satisfies a counterpart of Proposition 5.

Definition 9 *Type-1 sceptical defaults "if β then by default α", written as $\beta >_D \alpha$, are defined as follows.*

$$\beta >_D \alpha =_{def} \beta > \alpha \wedge \beta >_{\forall} \alpha$$

The two-phase approach with $>_D$ and $>_{\forall}$ works exactly like the two-phase approach with $>^c$ and $>_{\exists}$. The following example illustrates several properties of this new type-1 default.

Example 13 *The set of defaults $\{\top >_D p, \top >_D q\}$ is consistent, but the set of defaults $\{\top >_D p, \top >_D \neg p\}$ is inconsistent. Moreover, the set of defaults $\{\top >_D p, q >_D \neg p\}$ is inconsistent too.*

The latter example shows that the specificity principle is not formalized in the proposed two-phase defeault logic. An interesting approach to this problem is the multi preference framework proposed in [TvdT95].

5 Related research

An example of a default logic with ordering and minimizing is the dynamic interpretation of defaults in Veltman's preference-based default logic [Vel91].

For example, in his logic the formula $\top[\textit{normally}\, p] \models \textit{presumably}\, (p \vee q)$ means that after ordering all worlds by preferring p to $\neg p$, $p \vee q$ is true in the preferred worlds. A distinction between the logic proposed in this paper and Veltman's logic is that in the latter, the order of the defaults influences the set of derivable formulas. Moreover, in Veltman's logic defaults are not conditionals and ordering and minimizing are not modeled in the same object language.

Two phases can be traced in several popular non-monotonic logics like Reiter's default logic, autoepistemic logic and circumscription, which we illustrate by Reiter's default logic [Rei80]. Consider default theories consisting of a set of normal default rules $\frac{\beta:\alpha}{\alpha}$, which express that α is part of an extension (a deductively closed set of formulas) if β is part of the extension and $\neg\alpha$ is not, and a factual sentence (for simplicity we assume that the facts can be represented by a single formula). For example, the 'birds fly' default rule $\frac{b:f}{f}$ expresses that f is part of an extension if b is part of the extension and $\neg f$ is not; hence, birds are assumed to fly unless there is knowledge of the contrary. Type-1 and type-2 defaults can be identified in Reiter's default logic as follows.

1. For type-1 defaults, assume a fixed factual sentence. The 'birds fly' default $\frac{b:f}{f}$ is stronger than the 'red birds fly' default $\frac{b\wedge r:f}{f}$ in the sense that if a default theory contains the first default, then the second one can be added to the default theory without changing its set of extensions.

2. For type-2 defaults, assume a fixed set of normal Reiter defaults. If the factual sentence of a default theory is b and one of the extensions contains f, then this does *not* imply that the default theory with facts $b \wedge p$ has an extension that also contains f (e.g., with default rules $\frac{\top:\neg p}{\neg p}$ and $\frac{b\wedge\neg p:f}{f}$).

In this paper we proposed a logic for credulous reasoning about both types of defaults. In the vocabulary of a logic for reasoning about type-1 defaults, the 'red birds fly' default is *derivable* from the 'birds fly' default. The example illustrates that $\frac{\beta\wedge\gamma:\alpha}{\alpha}$ can be derived from $\frac{\beta:\alpha}{\alpha}$, which means that normal Reiter defaults validate strengthening of the antecedent. Similarly, it can be shown that $\frac{\beta:\alpha\vee\gamma}{\alpha\vee\gamma}$ *cannot* be derived from $\frac{\beta:\alpha}{\alpha}$, which means that normal Reiter defaults do *not* validate weakening of the consequent. In contrast, type-2 defaults do *not* validate strengthening of the antecedent (of b to $b \wedge p$ in the example above), but they do validate weakening of the consequent (because extensions are deductively closed). In this paper, we combined type-1 and type-2 defaults by the additional condition that type-1 defaults are strictly stronger than type-2 defaults. For Reiter's default logic, this additional condition means (in the credulous case) that for every default theory with default rule $\frac{\beta:\alpha}{\alpha}$ and factual sentence β, there is an extension that contains α. Hence,

the reasonable additional condition means that for every default rule $\frac{\beta:\alpha}{\alpha}$, $\beta \wedge \alpha$ is consistent.

Acknowledgement

Thanks to Patrick van der Laag for several discussions on the issues raised in this paper.

References

[Alc94] C. E. Alchourrón. Philosophical foundations of deontic logic and the logic of defeasible conditionals. In *Deontic Logic in Computer Science: Normative System Specification*, pages 43–84. John Wiley & Sons, 1994.

[AM91] N. Asher and M. Morreau. Common sense entailment: a modal theory of nonmonotonic reasoning. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence (IJCAI'91)*, Palo Alto, 1991. Morgan Kaufman.

[BDP92] S. Benferhat, D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning (KR'92)*, pages 673–684, Cambridge, MA, 1992.

[Bou92] C. Boutilier. Conditional logics for defeault reasoning and belief revision (phd thesis). Technical Report 92-1, Department of Computer Science, University of British Colombia, 1992.

[Bou94a] C. Boutilier. Conditional logics of normality: a modal approach. *Artificial Intelligence*, 68:87–154, 1994.

[Bou94b] C. Boutilier. Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence*, 68, 1994.

[Del88] J.P. Delgrande. An approach to default reasoning based on a first-order conditional logic: revised report. *Artificial Intelligence*, 36, 1988.

[KLM90] S. Kraus, D. Lehmann, and M. Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207, 1990.

[Lev90] H. Levesque. All I know: a study in autoepistemic logic. *Artificial Intelligence*, 42:263–309, 1990.

[LM92] D. Lehmann and M. Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55:1–60, 1992.

[Mak93] D. Makinson. Five faces of minimality. *Studia Logica*, 52:339–379, 1993.

[Mor95] M. Morreau. Allowed consequence. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI'95)*. Morgan Kaufman, 1995.

[Pea90] J. Pearl. System Z: a natural ordering of defaults with tractable applications to default reasoning. In *Proceedings of Theoretical Aspects of Reasoning about Knowledge (TARK)*, San Mateo, 1990. Morgan Kaufmann.

[Rei80] R. Reiter. A logic for default reasoning. *Artificial Intelligence*, 13:81–132, 1980.

[Sho88] Y. Shoham. *Reasoning About Change*. MIT Press, 1988.

[TP94] S.-W. Tan and J. Pearl. Specification and evaluation of preferences under uncertainty. In *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning (KR'94)*, pages 530–539, 1994.

[TvdT95] Y.-H. Tan and L.W.N. van der Torre. Why defeasible deontic logic needs a multi preference semantics. In *Proceedings of the ECSQARU'95. Lecture Notes in Artificial Intelligence 946*. Springer Verlag, 1995.

[TvdT96] Y.-H. Tan and L.W.N. van der Torre. How to combine ordering and minimizing in a deontic logic based on preferences. In *Proceedings of the Third Workshop on Deontic Logic in Computer Science (Δeon'96)*. Springer Verlag, 1996.

[vdTT95] L.W.N. van der Torre and Y.-H. Tan. Cancelling and overshadowing: two types of defeasibility in defeasible deontic logic. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI'95)*. Morgan Kaufman, 1995.

[Vel91] F. Veltman. Defaults in update semantics. Technical Report ITLI Prepublication Series LP-91-02, Department of Philosophy, University of Amsterdam, 1991. To appear in the *Journal of Philosophical Logic*.

[vW63] G.H. von Wright. *The logic of preference*. Edinburgh, 1963.