Sceptical Reasoning About Defaults*†

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

May 7, 1997

Abstract

In this paper we investigate preference-based logics for sceptical reasoning about defaults. In preference-based default logics a default is either formalized by a strong or by a weak preference, instances of what we call the *ordering* and *minimizing* usages of preference orderings. In a previous paper, we showed how ordering and minimizing can be formalized in Boutilier's modal logic CT40 and how they can be combined in a two-phase default logic. In this paper, we extend these results from the credulous case to the more complex sceptical case.

1 Introduction

The conditional sentence "if β then by default α" has (at least) two different interpretations, 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 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 investigate a logic for sceptical 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 shows that $\frac{\beta\wedge\gamma:\alpha}{\alpha}$ can be derived from $\frac{\beta:\alpha}{\alpha}$, which illustrates that type-1 defaults (like normal Reiter default rules) validate strengthening of the antecedent (SA). Similarly, it can be shown that $\frac{\beta:\alpha\vee\gamma}{\alpha\vee\gamma}$ *cannot* be derived from $\frac{\beta:\alpha}{\alpha}$, which illustrates that type-1 defaults do *not* validate weakening of the consequent (WC). In contrast, type-2 defaults do *not* validate SA (of b to $b \wedge p$ in the example above), but they do validate WC (because extensions are deductively closed). Notice that both types are duals as far as we consider the properties SA and WC. In [TvdT95a] (see also Section 2) it is shown that this duality is no coincidence, because combining SA and WC is problematic. Moreover, it is shown that they can be combined only by a technique, which might look odd at first sight, but which turns out to work well, namely to forbid application of SA after WC has been applied. This means that in derivations first SA 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, by combining two usages of a preference ordering in a preference-based semantics of a default logic. 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. The two phases are closely related to the two types of defaults, because the two types correspond to two different ways to evaluate formulas in a preference ordering. Type-1 defaults can be 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 can be 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. Summarizing, the two-phase approach combines SA and WC by combining type-1 and type-2 defaults, which semantically corresponds to combining ordering and minimizing. We combine type-1 and type-2 defaults by making type-1 defaults are strictly stronger than type-2 defaults, i.e. type-1 defaults *derive* corresponding type-2 defaults.[1]

*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.

†To submit to: Third Asian Logic Conference96, deadline 1 dec 95

‡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.

[1] In Reiter's default logic, this condition means that for every default, if the factual sentence is identical to the default rule's antecedent, then there is an extension which contains the default's consequent. This seems a reasonable constraint on default rules.

In this paper we extend the results of formalizing the two types of defaults and the two-phase default logic in a preference logic from the credulous case to the more complex sceptical case. In the sceptical case, a famous problem of preference logics is encountered, which is discussed in Example 1 below. 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, in some logics a value is 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. In such preference orderings there can be incomparable worlds. Incomparable worlds can be used to formalize 'multiple extensions' like the Nixon diamond in a consistent way. 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 $\neg p \wedge q$. In [TvdT95a], a strong preference for p means that "$\neg p$ is not preferred to or equivalent to p, regardless of other things".[2] This reading of preferences formalizes credulous reasoning about defaults, because the two expressions "by default p" and "by default $\neg p$" are consistent. The following example analyzes von Wright's problem from a proof theoretic point. It shows that the sceptical case is more complex than the credulous case, because SA interferes with scepticism.

Example 1 (SA+ D problem) *Assume a conditional default logic that validates at least substitution of logical equivalents and the following Gentzen-style inference pattern* Restricted Strengthening of the Antecedent (RSA). *We represent conditional defaults by $\beta > \alpha$, to be read as "if β (the antecedent) then by default α (the consequent)" and $\Diamond$ is a modal operator such that $\Diamond\phi$ is true for all consistent propositional formulas ϕ.*

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

Furthermore, assume the following D *axiom*.[3]

$$\text{D}'' : \Box\neg(\alpha_1 \wedge \alpha_2 \wedge \beta) \rightarrow \neg(\beta > \alpha_1 \wedge \beta > \alpha_2)$$

Finally, consider the set of defaults $\{\top > p_1, \top > p_2\}$, in which $\top$ stands for any tautology. From S the obligations $\neg(p_1 \wedge p_2) > p_1$ and $\neg(p_2 \wedge p_2) > p_2$ can be derived with RSA *(and $\Diamond(p_1 \wedge \neg p_2)$ and $\Diamond(\neg p_1 \wedge p_2)$). The two derived defaults are inconsistent with the* D'' *axiom when β is equivalent to $\neg(p_1 \wedge p_2)$. Obviously, S should be consistent; the derived inconsistency is what we call the* SA+ D *problem.*

There are two solutions for this problem, because the inconsistency is derived by only using RSA *and* D''*: weakening* D *or weakening* RSA. *An attempt of the first solution is to weaken the* D *axiom to the following axiom.*

$$\text{D}' : \Box\neg(\alpha_1 \wedge \alpha_2) \rightarrow \neg(\beta > \alpha_1 \wedge \beta > \alpha_2)$$

Unfortunately, S and D' *axiom are inconsistent with the following inference pattern* R, *and this inference pattern is accepted by most default logics.*

$$\text{R} : \frac{\beta > \alpha}{\beta > (\beta \wedge \alpha)}$$

From S we can derive $\neg(p_1 \wedge p_2) > (p_1 \wedge \neg(p_1 \wedge p_2))$ and $\neg(p_1 \wedge p_2) > (p_2 \wedge \neg(p_1 \wedge p_2))$ with R *and* RSA. *They are logically equivalent to $\neg(p_1 \wedge p_2) > (p_1 \wedge \neg p_2)$ and $\neg(p_1 \wedge p_2) > (p_2 \wedge \neg p_1)$, which are inconsistent with* D''' *when β is equivalent to $\neg(p_1 \wedge p_2)$. Obviously, if we weaken* D *to* D', *then we cannot accept the* R *rule. However, even without the* R *rule, the axiom* D' *is counterintuitive. For example, consider the set of defaults $S' = \{q_1 > p_1, q_2 > \neg p_1\}$. S' is intuitively consistent, but inconsistent with* D' *and* RSA. *The two defaults $(q_1 \wedge q_2) > p_1$ and $(q_1 \wedge q_2) > \neg p_1$ can be derived from S' by* RSA, *and they are inconsistent with* D'.

The second solution of the SA+ D *problem is to weaken* RSA. *This solution implies that the default $\neg(p_1 \wedge p_2) > p_1$ should not be derivable from the default $\top > p_1$ for the set of premises S.*[4] *An interesting issue is whether* RSA *should be weakenend such that the set of defaults $S'' = \{\top > p, q > \neg p\}$ becomes consistent. In this paper, we take the point of view that S'' should be inconsistent, and in Section 3 we show how inconsistency of S'' can be obtained in a two-phase default logic. Alternatively, it might be argued that S'' is consistent because the conflict should be resolved by a specificity argument, see Section 4). Notice that specificity is not part of Reiter's default logic, either.*

This paper is organized as follows. In Section 2, we give the preference logic CT40 in which we formalize credulous type-1 and type-2 defaults as strong and weak preferences, and we show how they can be combined in a two-phase default logic. In Section 3, we investigate the sceptical case. Finally, in Section 4, we mention the incorporation of specificity as further research.

[2]The condition is that no $\neg p$ world is strictly preferred over some p model, so the whole ordering is taken into account when a default is evaluated. That is why we call it the *ordering* usage of preference orderings. Similarly, the second reading is usually formalized by the *minimizing* usage of preference orderings, see for example [Sho88, KLM90, Mak93].

[3]The D axiom generalizes $\Box\neg(\alpha_1 \wedge \alpha_2) \rightarrow \neg(\top > \alpha_1 \wedge \top > \alpha_2)$ – which prohibits absolute defaults with inconsistent consequents –for arbitrary antecedents. Note that $\neg(\top > \alpha \wedge \top > \neg\alpha)$ – prohibiting contradictory absolute defaults – $\top > \alpha \wedge \top > (\neg\alpha \wedge \beta)$ consistent, because we did not assume weakening of the consequence.

[4]A simplistic solution to the SA+ D problem is not having any strengthening of the antecedent; however, then we do no longer have a type-1 default.

2 Credulous reasoning about defaults

In this section we formalize credulous reasoning about type-1 and type-2 defaults by strong and weak preferences. Moreover, we show how they can be combined in a two-phase default logic. The preferences are formalized in Boutilier's logic CT40, for the details of this logic see [Bou94b]. CT40 is a bimodal propositional logic of inaccessible worlds.

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

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

K	$\Box(\alpha \to \beta) \to (\Box\alpha \to \Box\beta)$
K′	$\overleftarrow{\Box}\,(\alpha \to \beta) \to (\overleftarrow{\Box}\,\alpha \to \overleftarrow{\Box}\,\beta)$
T	$\Box\alpha \to \alpha$
4	$\Box\alpha \to \Box\Box\alpha$
H	$\overleftrightarrow{\Diamond}\,(\Box\alpha \wedge \overleftarrow{\Box}\,\beta) \to \overleftrightarrow{\Box}\,(\alpha \vee \beta)$
Nec	*From* α *infer* $\overleftrightarrow{\Box}\,\alpha$
MP	*From* $\alpha \to \beta$ *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* $\overleftarrow{\Box}$ *to inaccessible worlds.*

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

Given this modal preference logic, we define type-1 defaults as strong preferences and type-2 defaults as weak preferences.[5]

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

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

$$\beta >_\exists \alpha =_{def} \overleftrightarrow{\Diamond}\,(\beta \wedge \Box(\beta \to \alpha))$$

Intuitively, a type-1 default $q > p$ expresses a strict preference of all $p \wedge q$ over $\neg p \wedge q$. This preference is represented by a negative condition: no $\neg p \wedge q$ is preferred to some $p \wedge q$. The type-2 default $q >_\exists p$ is true in a model if p is true in an equivalence class of most preferred q worlds of the model. Hence, the default $q >_\exists p$ refers to the preferred worlds where q is true, and $\top >_\exists p$ refers to the most preferred worlds. 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 and type-2 defaults as preferences.

[5] Boutilier defines weak preferences by the nearly equivalent $\beta >_\exists \alpha =_{def} \overleftrightarrow{\Box}\,\neg\beta \vee \overleftrightarrow{\Diamond}\,(\beta \wedge \Box(\beta \to \alpha))$.

Example 2 *Let* $|\alpha|$ *denote a world that satisfies* α*. Consider the Kripke model* M *that consist of four worlds ordered* $|p \wedge q| < |p \wedge \neg q| < |\neg p \wedge \neg q| < |\neg p \wedge q|$ *as represented in Figure 1. First,* M *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. Second,* M *satisfies the type-2 defaults* $\top >_\exists p$ *and* $\top >_\exists q$*. Since* $\top >_\exists q$ *is equivalent with* $\overleftrightarrow{\Diamond}\,\Box q$ *it is clear that* q *has to be true in some most preferred* $|\top|$ *world, and also that less preferred* $|\top|$ *worlds do not effect the truth of* $\overleftrightarrow{\Diamond}\,\Box q$*. Hence, in the evaluation of* $\top >_\exists q$ *only preferred elements are taken into account and not the whole ordering.*

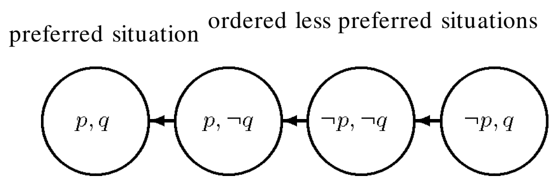

Figure 1: Preference relation with four worlds

The following proposition gives several properties of the type-1 and type-2 defaults, and illustrates that they are duals of each other when we consider strengthening of the antecedent and weakening of the consequent.

Proposition 1 *The logic CT40 validates the following theorems.*

SA	$\beta_1 > \alpha \to (\beta_1 \wedge \beta_2) > \alpha$
And	$\beta > \alpha_1 \wedge \beta > \alpha_2 \to \beta > (\alpha_1 \wedge \alpha_2)$
Or	$\beta > \alpha_1 \wedge \beta > \alpha_2 \to \beta > (\alpha_1 \vee \alpha_2)$
Trans′	$\gamma > \beta \wedge \beta > \alpha \to \gamma > (\alpha \wedge \beta)$
WC$_\exists$	$\beta >_\exists \alpha_1 \to \beta >_\exists (\alpha_1 \vee \alpha_2)$
	$\beta >_\exists \alpha \to \overleftrightarrow{\Diamond}\,(\beta \wedge \alpha)$
	$\neg(\alpha >_\exists \bot)$
	$\overleftrightarrow{\Diamond}\,\alpha \to \alpha >_\exists \alpha$

The logic CT40 does not *validate the following theorems.*

WC	$\beta > \alpha_1 \to \beta > (\alpha_1 \vee \alpha_2)$
Trans	$\gamma > \beta \wedge \beta > \alpha \to \gamma > \alpha$
D	$\neg(\beta > \alpha \wedge \beta > \neg\alpha)$
SA$_\exists$	$\beta_1 >_\exists \alpha \to (\beta_1 \wedge \beta_2) >_\exists \alpha$
AND$_\exists$	$\beta >_\exists \alpha_1 \wedge \beta >_\exists \alpha_2 \to \beta >_\exists (\alpha_1 \wedge \alpha_2)$
DD$_\exists$	$\gamma >_\exists \beta \wedge \beta >_\exists \alpha \to \gamma >_\exists \alpha$
D$_\exists$	$\neg(\beta >_\exists \alpha \wedge \beta >_\exists \neg\alpha)$

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

The idea of combining ordering and minimizing is to combine formulas with $>$ and $>_\exists$ operators, where we demand that ordering is strictly stronger than minimizing. However, the combination $>$ and $>_\exists$ is not satisfactory in the logic CT40, because we cannot derive $\beta >_\exists \alpha$ from $\beta > \alpha$. For example, 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 $\beta \wedge \alpha$ is possible.

Definition 3 *Consistent type-2 defaults "if β then by default α", written as $\beta >^c \alpha$, are defined as follows.*

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

The version of ordering introduced in the previous definition is strictly stronger than minimizing, as is shown in the following proposition.

Proposition 2 *The logic CT40 validates the following theorem.*

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

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

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

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

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 4 *The logic CT40* is CT40 extended with the following axiom scheme:*

LP $\quad \overleftrightarrow{\Diamond} \phi$ for all satisfiable propositional ϕ

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

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 2, 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 2 we also obtain $(\beta \wedge \gamma) >_\exists \alpha$. If $\beta >_\exists \alpha$ but not $\beta >^c \alpha$, then there is no way we can apply RSA to this formula. Being a minimizing formula $\beta >_\exists \alpha$ lacks RSA. Hence, once it 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 can combine strengthening of the antecedent and weakening of the consequent.

Example 3 (SA+ WC problem) *Consider the defaults that (1) you normally either buy apples or you buy pears, and (2) you normally do not buy apples. Then, given that you already buy apples, it is counterintuitive to derive that 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). *We represent conditional defaults by $\beta > \alpha$, where $>$ stands for any conditional connective (not necessarily the one we defined in Boutilier's logic CT40).*

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

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

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

Furthermore, assume as premises the set of defaults $S = \{\top > (a \vee p), \top > \neg a$, where a can be read as "buying apples" and p as "buying pears". The intuitive default $\top > (\neg a \wedge p)$ can be derived from S by AND. *From this 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. Notice this inference is validated by $>^c$ in Boutilier's logic CT40.*

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

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 depicted in Figure 2.*

The solution of the problem in the two-phase default logic is to block the second derivation by disallowing the application of RSA *after* WC. *Let $S' = \{\top >^c (a \vee p), \top >^c \neg a\}$ be a CT40* theory, where $\neg a$ does not entail $\neg p$. We have $S' \models \overleftrightarrow{\Diamond} (\neg a \wedge p)$, $S' \models \top >^c (\neg a \wedge p)$ 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.

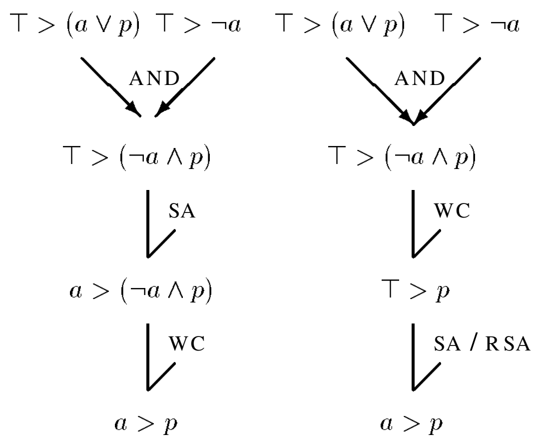

Figure 2: SA+ WC problem

3 Sceptical reasoning about defaults

In this section, we consider the sceptical case of the two-phase approach. 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 5 *Type-2 sceptical defaults "if β then by default α", written as $\beta >_\forall \alpha$, are defined as follows.*

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

However, this type-2 default cannot satisfactorily be combined with the type-1 default $\beta >^c \alpha$. The following example illustrates that the ordering of worlds can be too weak.

Example 4 *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 $|\neg p|$ worlds.*

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.[6]

Definition 6 *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].

Definition 7 *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 5 *Let the set of defaults $S = \{\top >^c (\neg r \wedge \neg g), g >^c r, r >^c g\}$ be a CT40* theory. The intended model is given in Figure 3. We have $S \not\models \top >^c \neg r$, $S \not\models \top >^c \neg g$, $S \models \top >_\forall \neg r$ and $S \models \top >_\forall \neg g$. 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 $M \models \top >_\forall \neg r$.*

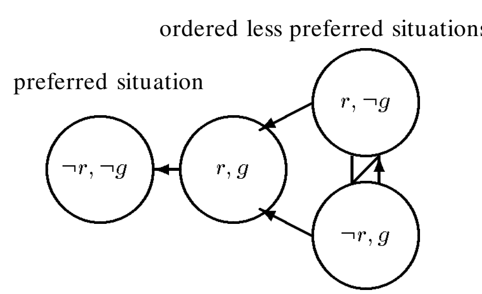

Figure 3: Preference relation

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 6 *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.*

[6] 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.

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.

Definition 8 *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 sceptical type-1 default trivially satisfies the following proposition, the counterpart of Proposition 2.

Proposition 3 *The logic CT40 validates the following theorem.*

$$\beta >_D \alpha \rightarrow \beta >_\forall \alpha$$

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 7 (SA+ D problem, continued) *Reconsider the sets of defaults discussed in Example 1: $S = \{\top >_D p_1, \top >_D p_2\}$, $S' = \{q_1 >_D p_1, q_2 >_D \neg p_1\}$ and $S'' = \{\top >_D p, q >_D \neg p\}$. S and S' are consistent, but S'' is inconsistent.*

Remark Another perspective on the sceptical two-phase default logic is that premises and conclusions are of type-2, and that the thus defined 'conservative core' is strengthened by preferential entailment (similar to system Z) and the additional premises $\beta >_D \alpha$. The additional premises result in a rule-counting mechanism.

4 Further research

A well-known disadvantage of normal Reiter default rules is that specificity cannot be modeled. The same problem occurs in credulous type-1 defaults, as expressed by strengthening of the antecedent. and sceptical type-1 defaults, as expressed by the inconsistency of S'' in Example 7. An interesting solution is the multi preference framework like in [TvdT95b].

References

[Bou92] C. Boutilier. PhD thesis, 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.

[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.

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

[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.

[TvdT95a] Y.-H. Tan and L.W.N. van der Torre. Credulous reasoning about defeaults. Technical report, EURIDIS, 1995.

[TvdT95b] 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.

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