Department of Philosophy - Utrecht University

Preference-based Semantics for Nonmonotonic Logics

θ

F. Voorbraak

π

Logic Group
Preprint Series
No. 96
August 1993

Department of Philosophy
Utrecht University
Heidelberglaan 8
3584 CS Utrecht
The Netherlands

ISBN 90-393-0458-0
ISSN 0929-0710
Dr. A. Visser, Editor

Preference-based Semantics for Nonmonotonic Logics*

Frans Voorbraak

Department of Philosophy
Utrecht University
P.O. Box 80.126
3508 TC Utrecht

Abstract

A variant is proposed of the preference-based semantics for nonmonotonic logics that was originally considered by Shoham (1987,1988). In this variant it is not assumed that preferences between standard models are aggregated into one preference order. This allows the capturing of *all* main nonmonotonic formalisms, including Default Logic of Reiter (1980). The preferential models introduced in this paper are motivated from an epistemic point of view, and are therefore called epistemic preference models. The consequence operations induced by epistemic preference models are characterized. Further, the view is defended that the rationality of cumulative monotonicity does not imply that nonmonotonic logics have to be cumulative, but only that a rational agent should not believe a set of default rules that induces a noncumulative consequence operation.

* This is an extended version of a paper that appears in the Proceedings of the 13th International Joint Conference on Artificial Inteligence (IJCAI-93). The investigations were supported by the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Researsch (NWO).

1 Introduction

Shoham (1987,1988) introduces preferential semantics as a possible unifying framework for all nonmonotonic logics. In that framework, a nonmonotonic logic is reduced to a standard logic plus a preference order on the models of that standard logic, and nonmonotonic entailment is considered to be preferential entailment, where Γ preferentially entails φ iff φ is true in every model M such that (1) Γ is true in M and (2) Γ has no model M' which is preferred to M.

In this paper we propose a variant of Shoham's framework which is epistemically motivated and uses (not necessarily transitive) preference *relations* between *sets* of models of a standard logic. From a technical point of view, considering *sets* of standard models is not an essential generalization, since sets of models of a standard logic **L** induce a partial variant of **L**, which is again standard. However, using arbitrary preference relations instead of preference orders *is* essential, since it allows more freedom to express preferences.

A consequence of the additional freedom is that in our approach Default Logic (**DL**) of Reiter (1980), which is one of the most important nonmonotonic formalisms, can be given a preference-based semantics, whereas in Shoham's original approach this turned out to be difficult, if not impossible. However, our relaxation of the constraints on preference-based semantics is not an *ad hoc* move solely motivated by the need to capture **DL**, but it follows from our interpretation of the preference relation.

We interpret the preferences between (sets of) models as epistemic preferences (of an ideally rational agent) between world descriptions. Since an agent usually has only partial information about the world, epistemic preferences are most naturally captured by preferences between *sets* of models, where the sets of models correspond to *partial* world descriptions. In our view, an agent has in general epistemic preferences of different kinds, which are not always easily combined into one preference order. For example, an agent may have preferences induced by factual information, induced by default information, or induced by lack of information:

EXAMPLE 1.1 Consider an agent with the following beliefs: (1) Typically, it does not rain in California, (2) It now rains in California. On account of belief (1), the agent prefers world descriptions in which it does not rain in California. However, this default preference is overridden by the preference induced by the factual belief (2) that it does rain. Since the agent does not have any (factual or default) beliefs about the weather conditions in Kansas, he prefers world descriptions which leave undecided the question whether it rains in Kansas.

In our analysis of the situation, an agent has a preference for the less specific or more ignorant world descriptions, unless there is some (default or factual) information to the contrary. This leads us to consider the lexicographic aggregation of strict partial orders, which itself is in general not a strict partial order.

DEFINITION 1.2 Let $>_1$ and $>_2$ be two asymmetric (preference)relations. The *lexicographic aggregation* of $>_1$ and $>_2$ is the relation $\lambda(>_1,>_2) =_{def} >_1 \cup (\sim_1 \cap >_2)$, where $x \sim_1 y \Leftrightarrow_{def}$ not $x >_1 y$ and not $y >_1 x$. In other words, $\langle x,y \rangle \in \lambda(>_1,>_2)$ iff y is 1-preferred to x, or x and y are 1-indifferent and y is 2-preferred to x.

It is easy to construct an example showing that the lexicograhic aggregation of two strict partial orders is itself not necessarily transitive. A few more properties of the lexicographic aggregation can be found in Voorbraak (1993).

Requiring the aggregated preferences of an agent to form a strict partial order should in our opinion be considered to be a rationality requirement on the (default and factual) beliefs of the agent and should not be interpreted as a requirement on nonmonotonic *logic*. Similar remarks hold for the smoothness condition of Kraus et al. (1990), the boundedness condition of Makinson (1989), etcetera. We return to this issue in some detail.

Since we use preference relations to denote preferences of an ideally rational agent between sets of models corresponding to partial world descriptions, it follows that the preference relations should not distinguish sets of models that correspond to the same world description, i.e., that validate the same set of formulas of the language under consideration. Instead of imposing this constraint on the preference relations, we will divide out the indistinguishability relation. In other words, if $\mathcal{L}_{\mathbf{L}}$ is the language of the standard logic $\mathbf{L}$ and $\mathrm{MOD}_{\mathbf{L}}$ is the class of $\mathbf{L}$-models, then the preference relation is defined on $\{\mathcal{M} \subseteq \mathrm{MOD}_{\mathbf{L}} \mid \text{for some } \Sigma \subseteq \mathcal{L}_{\mathbf{L}}, \mathcal{M} = \{M \mid M \models_{\mathbf{L}} \Sigma\}\}$. Hence each relevant set $\mathcal{M}$ of $\mathbf{L}$-models is characterized by some subset of $\mathcal{L}_{\mathbf{L}}$. Equivalently, one can define the preferences on $\{\Sigma \subseteq \mathcal{L}_{\mathbf{L}} \mid Cn_{\mathbf{L}}(\Sigma) = \Sigma\}$, the set of $\mathbf{L}$-theories.

In Kraus et al. (1990) a state is merely labelled and *not* identified with a set of worlds. We take a similar position in Voorbraak (1992), but there we describe the same (epistemic) states in different formal languages, whereas in this chapter we restrict ourselves to a single language, viz. $\mathcal{L}_{\mathbf{L}}$, which makes it reasonable to identify objects which cannot be distinguished in that language. In Kraus et al. (1990, p. 181) it is claimed that the additional freedom obtained by not identifying states with the same label is vital for their representation theorem of cumulative consequence relations. However, this is not true, since in the cumulative model constructed on p. 185 of Kraus et al. (1990) the states with the same label are in fact identified.

2 Epistemic preference models

Before we give the formal definition of our models for preferential entailment, we introduce some notation. Throughout this paper, **L** denotes some standard logic with language $\mathcal{L}_{\mathbf{L}}$, consequence relation $\models_{\mathbf{L}}$, and consequence operation $\mathrm{Cn}_{\mathbf{L}}$. $\mathrm{MOD}_{\mathbf{L}}$ is the set of models of **L** and $\models_{\mathbf{L},h}$ denotes the *hypervaluation consequence relation* defined as follows: for any set $\mathcal{M} \subseteq \mathrm{MOD}_{\mathbf{L}}$, $\mathcal{M} \models_{\mathbf{L},h} \varphi$ iff for all $M \in \mathcal{M}$, $M \models_{\mathbf{L}} \varphi$. (Hypervaluation consequence is introduced in Voorbraak (1991) and some of its properties are studied in Voorbraak (1993).)

DEFINITION 2.1 For any $\mathcal{M} \subseteq \mathrm{MOD}_{\mathbf{L}}$, $\mathrm{Th}_{\mathbf{L}}(\mathcal{M})$ denotes the ***L**-theory of* $\mathcal{M}$, i.e., the set $\{\varphi \in \mathcal{L}_{\mathbf{L}} \mid \mathcal{M} \models_{\mathbf{L},h} \varphi\}$. For any $\Sigma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Mod}_{\mathbf{L}}(\Sigma) =_{\mathrm{def}} \{M \in \mathrm{MOD}_{\mathbf{L}} \mid M \models_{\mathbf{L}} \Sigma\}$. For any $\mathcal{M} \subseteq \mathrm{MOD}_{\mathbf{L}}$, we define $\mathcal{M}^{+} =_{\mathrm{def}} \{M \in \mathrm{MOD}_{\mathbf{L}} \mid$ for every $\varphi \in \mathrm{Th}(\mathcal{M})$, $M \models_{\mathbf{L}} \varphi\}$. Further, the set of *world descriptions* of **L**, $\mathrm{WD}_{\mathbf{L}} =_{\mathrm{def}} \{\mathrm{Mod}_{\mathbf{L}}(\Sigma) \mid \Sigma \subseteq \mathcal{L}_{\mathbf{L}}\}$ The subscript **L** will often be omitted when confusion is unlikely to occur.

The following proposition lists some properties of the introduced notions. The (straightforward) proofs are omitted.

PROPOSITION 2.2 Let **L** be a standard logic, $\Sigma \subseteq \mathcal{L}_{\mathbf{L}}$, and $\mathcal{M} \subseteq \mathrm{MOD}_{\mathbf{L}}$.
(i) $\Sigma \subseteq \mathrm{Th}_{\mathbf{L}}(\mathcal{M})$ iff $\mathcal{M} \subseteq \mathrm{Mod}_{\mathbf{L}}(\Sigma)$*
(ii) $\mathrm{Th}_{\mathbf{L}}(\mathrm{Mod}_{\mathbf{L}}(\Sigma)) = \mathrm{Cn}_{\mathbf{L}}(\Sigma)$
(iii) $\mathrm{Mod}_{\mathbf{L}}(\mathrm{Th}_{\mathbf{L}}(\mathcal{M})) = \mathcal{M}^{+}$
(iv) If $\mathrm{Cn}_{\mathbf{L}}(\Sigma) = \mathrm{Th}_{\mathbf{L}}(\mathcal{M})$, then $\mathrm{Mod}_{\mathbf{L}}(\Sigma) = \mathcal{M}^{+}$
(v) For every $\mathcal{M} \in \mathrm{WD}_{\mathbf{L}}$, $\mathcal{M} = \mathcal{M}^{+}$.

The notion of $\mathcal{M}^{+}$ and its notation derive from Levesque (1990). He shows that $\mathcal{M}^{+}$ is a maximal set of **L**-models, in the sense that there is no proper superset of $\mathcal{M}^{+}$ which is equivalent to $\mathcal{M}^{+}$, and that $\mathcal{M}^{+}$ is the unique maximal set of **L**-models equivalent to $\mathcal{M}$.

We call our models for preferential entailment *epistemic* preference models, in order to distinguish them from several other models proposed in the literature, and in order to emphasize that we explicitly assumed the preference relation to express preferences between epistemic (belief) states.

---

* It follows that the pair $\langle \mathrm{Mod}_{\mathbf{L}}, \mathrm{Th}_{\mathbf{L}} \rangle$ is a Galois connection between the ordered sets $\mathcal{P}(\mathcal{L}_{\mathbf{L}})$ and $\mathcal{P}(\mathrm{MOD}_{\mathbf{L}})$.

DEFINITION 2.3 Let $\mathbf{L}$ be a standard logic. An *epistemic preference model* (for $\mathbf{L}$) is a triple $\langle \mathrm{WD}_{\mathbf{L}}, >, \models \rangle$, where $>$ is an irreflexive relation on the set of world descriptions $\mathrm{WD}_{\mathbf{L}}$, and $\models$ is the satisfaction relation $\models_{\mathbf{L},h}$ restricted to $\mathrm{WD}_{\mathbf{L}} \times \mathcal{L}_{\mathbf{L}}$.

DEFINITION 2.4 Let $\langle \mathrm{WD}_{\mathbf{L}}, >, \models \rangle$ be an epistemic preference model.

(i) $\mathcal{M} \in \mathrm{WD}_{\mathbf{L}}$ *preferentially satisfies* φ, notation $\mathcal{M} \models_{\mathbf{L},>} \varphi$, iff $\mathcal{M} \models_{\mathbf{L}} \varphi$ and for all $\mathcal{N} < \mathcal{M}$, $\mathcal{N} \not\models \varphi$.

(ii) φ is a *preferential consequence* of Γ, notation $\Gamma \models_{\mathbf{L},>} \varphi$, iff for all $\mathcal{M} \in \mathrm{WD}_{\mathbf{L}}$, $\mathcal{M} \models_{\mathbf{L},>} \Gamma$ implies $\mathcal{M} \models \varphi$.

We write $\mathrm{Cn}_{\mathbf{L},>}(\Gamma)$ for $\{\varphi \in \mathcal{L}_{\mathbf{L}} \mid \Gamma \models_{\mathbf{L},>} \varphi\}$ and we sometimes omit the subscript $\mathbf{L}$ if it is clear which standard logic is used.

Notice that an epistemic preference model is determined by the standard logic $\mathbf{L}$ and the preference relation $>$ on $\mathrm{WD}_{\mathbf{L}}$. Before we characterize the preferential consequence operation $\mathrm{Cn}_{>}$ of epistemic preference models we make some general comments on the different preferential semantics and we show that **DL** can be captured in epistemic preferential semantics.

Epistemic preference models have a status quite similar to the one of possible worlds models in modal logic, since they can be enriched by adding further conditions on $>$. An obvious condition is transitivity. To give another example, an epistemic preference model $\langle \mathrm{WD}_{\mathbf{L}}, >, \models \rangle$ is called *proper* iff set inclusion is contained in $\succcurlyeq \cup \preccurlyeq$, i.e., for all $\mathcal{M},\mathcal{N} \in \mathrm{WD}_{\mathbf{L}}$, $\mathcal{M} \subseteq \mathcal{N}$ implies $\mathcal{M} \succcurlyeq \mathcal{N}$ or $\mathcal{M} \preccurlyeq \mathcal{N}$. Proper epistemic preference models implement the intuition that a rational agent should not be indifferent with respect to epistemic states $\mathcal{M}$ and $\mathcal{N}$ if $\mathcal{M} \subseteq \mathcal{N}$.

For comparison, we also provide here the definitions of the models for preferential entailment that were proposed by Shoham (1987), Makinson (1989), and Kraus et al. (1990).

DEFINITION 2.5 Let $\mathbf{L}$ be a standard logic. A *Shoham model* (for $\mathbf{L}$) is a triple $\langle \mathrm{MOD}_{\mathbf{L}}, >, \models_{\mathbf{L}} \rangle$, where $>$ is a strict partial order on $\mathrm{MOD}_{\mathbf{L}}$ and $\models_{\mathbf{L}}$ is the satisfaction relation of $\mathbf{L}$.

DEFINITION 2.6 Let $\mathbf{L}$ be a standard logic. A *Makinson model* (for $\mathbf{L}$) is a triple $\langle \mathfrak{M}, >, \models \rangle$, where $\mathfrak{M}$ is an arbitrary set, $>$ is a binary relation on $\mathfrak{M}$, and $\models$ an arbitrary satisfaction relation $\subseteq \mathfrak{M} \times \mathcal{L}_{\mathbf{L}}$. A Makinson model will be called ***L**-faithful* iff for every $\mathcal{M} \in \mathfrak{M}$ the set $\{\varphi \in \mathcal{L}_{\mathbf{L}} \mid \mathcal{M} \models \varphi\}$ is closed under $\mathrm{Cn}_{\mathbf{L}}$.

DEFINITION 2.7 Let **L** be a standard logic. A *KLM model* (for **L**) is a quadruple $\langle S,\ell,>,\models\rangle$, where S is an arbitrary collection of states, $\ell : S \rightarrow \mathcal{P}(MOD_{\mathbf{L}})$ is a labelling function, $>$ is a binary relation on S, and $\models \subseteq S\times\mathcal{L}_{\mathbf{L}}$ is defined as follows: for all $s \in S$ and for all $\varphi \in \mathcal{L}_{\mathbf{L}}$, $s \models \varphi$ iff $\ell(s) \models_{\mathbf{L},h} \varphi$.

Preferential entailment on Shoham models, on Makinson models, and on KLM models is defined completely analogous to 2.4.

It is easy to see that any epistemic preference model and any Shoham model is a (**L**-faithful) Makinson model and is equivalent to a KLM model. The precise relation between Makinson models and KLM models is determined in Dix and Makinson (1992). From a technical point of view, Makinson models or KLM models are perhaps the obvious choice for basic preference models. However, since often epistemic intuitions are used to informally justify the preferential models, it is advisable to take these intuitions serious and consider preferences between "real" epistemic states, as is done in epistemic preference models. Proposition 2.8 below shows that (proper) epistemic preference models are at least as general as Shoham models.

PROPOSITION 2.8 For any Shoham model $\langle MOD_{\mathbf{L}},>,\models_{\mathbf{L}}\rangle$ there exists a proper epistemic preference model $\langle WD_{\mathbf{L}},>',\models\rangle$ such that $Cn_{\mathbf{L},>} = Cn_{\mathbf{L},>'}$.

Proof. Let $\langle MOD_{\mathbf{L}},>,\models_{\mathbf{L}}\rangle$ be a Shoham model. Define $>' \subseteq WD_{\mathbf{L}}\times WD_{\mathbf{L}}$ as $\lambda(>,\subset)$, where $>$ is the irreflexive version of $\geq =_{def} \{\langle\mathcal{M},Mod(Cn_{\mathbf{L},>}(Th(\mathcal{M})))\rangle \mid \mathcal{M} \in WD_{\mathbf{L}}\}$, and $\subset$ is proper set-inclusion. Then $\langle WD_{\mathbf{L}},>',\models\rangle$ is a proper epistemic preference model whose preferential consequence operator $Cn_{\mathbf{L},>'}$ coincides with $Cn_{\mathbf{L},>}$, since $\mathcal{M} \models_{>'} \Sigma$ iff $\mathcal{M} = Mod(Cn_{\mathbf{L},>}(Th(Mod(\Sigma))))$. (For any other $\mathcal{M}$ such that $\mathcal{M} \models \Sigma$, $\mathcal{M} >' Mod(\Sigma)$.) ■

It follows that all nonmonotonic logics treated in Shoham (1987,1988) can also be captured in our approach. This includes predicate circumscription of McCarthy (1980), the minimal knowledge logic of Halpern and Moses (1985), and some variants of **DL**, but *not* **DL** itself. Also, it is shown in Voorbraak (1993) that **NML** (nonmonotonic modal logic) of McDermott (1982) can be captured by means of epistemic preference models. **AEL** (autoepistemic logic) of Moore (1984,1985) is a special case of **NML**.

Of course, proposition 2.8 does not provide a very satisfactory translation of Shoham models into epistemic preference models. One would like to have a more direct relation between $>'$ and $>$. However, instead of pursuing the matter of the relation between the two kinds of models, we concentrate on showing how **DL** can be captured by epistemic preference models.

3 Preference-based Semantics for Default Logic

We assume familiarity with Default Logic (**DL**). (See Reiter (1980) or almost any introduction to nonmonotonic reasoning.) For convenience, we repeat some conventions and notation.

The underlying standard logic **L** is assumed to be ordinary first-order logic. A *default rule* (or simply a *default*) is an expression of the form $\alpha : \beta_1,...,\beta_n / \omega$ $(n \geq 1)$, where α (the *prerequisite*), $\beta_1,...,\beta_n$ (the *justifications*), and ω (the *conclusion*) are first-order formulas. Without loss of generality, we assume these formulas to be closed. A *default theory* is a pair $\mathfrak{D} = \langle D,\Gamma\rangle$, where D is a set of defaults and $\Gamma \subseteq \mathcal{L}_L$. An extension of a default theory $\langle D,\Gamma\rangle$ is a set of formulas which is intended to represent a reasonable state of belief based on the defaults in D and on the propositions in Γ.

DEFINITION 3.1 An *extension* of the default theory $\langle D,\Gamma\rangle$ is a fixed point of the function $f : \mathcal{P}\mathcal{L} \to \mathcal{P}\mathcal{L}$ given by: $f(\Sigma)$ is the smallest set such that

D1 $\Gamma \subseteq f(\Sigma)$

D2 $Th(f(\Sigma)) = f(\Sigma)$

D3 If $\alpha : \beta_1,...,\beta_n / \omega \in D$, $\alpha \in f(\Sigma)$ and for all $i \in \{1,..,n\}$, $\neg\beta_i \notin \Sigma$, then $\omega \in f(\Sigma)$.

Our preferential semantics for **DL** is closely related to the semantics for **DL** given in Etherington (1988), which associates a preference relation to each default rule as follows:

DEFINITION 3.2 Let $\delta = \alpha : \beta_1,..,\beta_n / \omega$ be a default and let $\mathfrak{M}$ be a set of first-order models. The *preference relation* $>_\delta$ *corresponding to* δ over $\mathcal{P}\mathfrak{M}$ is defined as follows: $(\mathcal{M}, \mathcal{N} \in \mathcal{P}\mathfrak{M})$

$\mathcal{N} >_\delta \mathcal{M}$ iff Every model $M \in \mathcal{N}$ is a model of α,
for all $i \in \{1,..,n\}$, there exists a model $M_i \in \mathcal{N}$ of β_i,
and $\mathcal{M} = \mathcal{N} - \{M \in \mathcal{N} \mid M \not\models \omega\} \neq \mathcal{N}$.

Intuitively, $\mathcal{N} >_\delta \mathcal{M}$ means that on account of δ the (partial) world-description $\mathcal{M}$ is preferred to the (partial) world-description $\mathcal{N}$.* The preference relation corresponding to a finite set of defaults D is simply the transitive closure of the union of the preference relations corresponding to the elements of D. For the general case, we need a slightly more complicated definition. The definition given below is perhaps rather *ad hoc*, but it serves our purpose.

* We have changed some of Etherington's notation and terminology to obtain consistency with respect to the conventions used in preferential semantics.

DEFINITION 3.3 Let D be a set of defaults and let $\mathfrak{M}$ be a set of models.

(i) The preference relation corresponding to D, $>_D$, over $\mathcal{P}\mathfrak{M}$ is defined as follows: $\mathcal{N} >_D \mathcal{M}$ iff there exist $\delta_1, \delta_2, \ldots \in D$ and $\mathcal{N}_0, \mathcal{N}_1, \mathcal{N}_2, \ldots \subseteq \mathfrak{M}$ such that $\mathcal{N} = \mathcal{N}_0 >_{\delta_1} \mathcal{N}_1 >_{\delta_2} \mathcal{N}_2 >_{\delta_3} \ldots$, and $\mathcal{M} = \cap \mathcal{N}_i$.

(ii) For any default theory $\mathfrak{D} = \langle D, \Gamma \rangle$, $>_{\mathfrak{D}}$ denotes the relation $>_D$ defined over $\mathcal{P}(\mathrm{Mod}(\Gamma))$, where $\mathrm{Mod}(\Gamma)$ is the set of first-order models of Γ.

If $\mathfrak{D}$ is normal, then its extensions correspond to the minimal elements of $>_{\mathfrak{D}}$. This correspondence is essentially obtained in Łukaszewicz (1985). Etherington (1988) generalizes the result to arbitrary default theories as follows:

DEFINITION 3.4 Let $\mathfrak{D} = \langle D, \Gamma \rangle$ be a default theory. $\mathcal{M} \subseteq \mathrm{Mod}(\Gamma)$ is called $\mathfrak{D}$-*stable* iff there exists a $D' \subseteq D$ such that $\mathcal{M} \leq_{D'} \mathrm{Mod}(\Gamma)$ and $\mathcal{M}$ contains a model for every justification β of every default of D'.

PROPOSITION 3.5 (Etherington (1988)) Let $\mathfrak{D} = \langle D, \Gamma \rangle$ be a default theory.

(i) If E is an extension of $\mathfrak{D}$, then Mod(E) is $>_{\mathfrak{D}}$-minimal and $\mathfrak{D}$-stable.

(ii) If $\mathcal{M}$ is $\mathfrak{D}$-stable and $>_{\mathfrak{D}}$-minimal, then $\{\varphi \mid \forall M \in \mathcal{M} \;\; M \models \varphi\}$ is an extension of $\mathfrak{D}$.

The preferential semantics for **DL** will be given in two stages. Before the models corresponding to sets of defaults are given, we first define preference models for default theories. Notice that for any default theory $\mathfrak{D}$ it is trivial to give a preferential model $\langle \mathfrak{M}, \succ, \models \rangle$ such that the $\succ$-minimal elements correspond to the extensions of $\mathfrak{D}$. (Just define $\mathcal{M} \succ \mathcal{N}$ iff $\mathcal{N}$ is an extension and $\mathcal{M}$ is not.) Our preferential models are slightly less trivial, since they are intended to help us towards a preferential semantics for sets of defaults. The preference relation associated with a default theory $\mathfrak{D}$ is a variant of the relation $>_{\mathfrak{D}}$, defined in 3.3(ii). This variant is obtained by adding a check for $\mathfrak{D}$-stability to the relation $>_{\mathfrak{D}}$.

DEFINITION 3.6 Let $\mathfrak{D} = \langle D, \Gamma \rangle$ be a default theory. The *epistemic preference model associated with* $\mathfrak{D}$ is the model $\langle WD_L, \succ_{\mathfrak{D}}, \models \rangle$, where $\succ_{\mathfrak{D}} = \lambda(>_0, >_{\mathfrak{D}})$ and $>_0$ is defined as follows: $\mathcal{N} >_0 \mathcal{M}$ iff $\mathcal{M} = \varnothing \neq \mathcal{N}$ and there is no $>_{\mathfrak{D}}$-minimal and $\mathfrak{D}$-stable $\mathcal{N}' \leq_{\mathfrak{D}} \mathcal{N}$. To avoid stacked subscripts, we write $\models_{\mathfrak{D}}$ and $Cn_{\mathfrak{D}}$ instead of $\models_{\succ_{\mathfrak{D}}}$ and $Cn_{\succ_{\mathfrak{D}}}$.

PROPOSITION 3.7 Let $\mathfrak{D} = \langle D, \Gamma \rangle$ be a default theory and $\langle WD_L, \succ_{\mathfrak{D}}, \models \rangle$ its associated epistemic preference model.

(i) If E is an extension of $\mathfrak{D}$, then $\mathrm{Mod}(E) \models_{\mathfrak{D}} E$.

(ii) If $\mathcal{M} \neq \emptyset$ is $\succ_{\mathfrak{D}}$-minimal, then $\mathrm{Th}(\mathcal{M})$ is a consistent extension of $\mathfrak{D}$.
(iii) If $\emptyset$ is the only $\succ_{\mathfrak{D}}$-minimal element, then $\mathfrak{D}$ has no consistent extension.
(iv) $\mathrm{Cn}_{\mathfrak{D}}(\Gamma) = \bigcap\{E \mid E \text{ is an extension of } \mathfrak{D}\}$.
(v) For any $\Delta \subseteq \mathcal{L}_L$, $\mathrm{Cn}_{\mathfrak{D}}(\Delta) = \bigcap\{E \mid E \text{ is an extension of } \mathfrak{D} \text{ and } \Delta \subseteq E\}$.

Proof. (i) If E is an extension of $\mathfrak{D}$, then it follows from proposition 3.5 that Mod(E) is $>_{\mathfrak{D}}$-minimal, and thus also $\succ_{\mathfrak{D}}$-minimal. Since Mod(E) $\models$ E, we have Mod(E) $\models_{\mathfrak{D}}$ E.
(ii) If $\mathcal{M}$ is $\succ_{\mathfrak{D}}$-minimal and $\mathcal{M} \neq \emptyset$, then $\mathcal{M}$ is $>_{\mathfrak{D}}$-minimal and $\mathfrak{D}$-stable. Hence, by proposition 3.5, $\mathrm{Th}(\mathcal{M})$ is an extension of $\mathfrak{D}$. Since $\mathcal{M} \neq \emptyset$, this extension is consistent.
(iii) If E is a consistent extension of $\mathfrak{D}$, then Mod(E) $\neq \emptyset$ is $\succ_{\mathfrak{D}}$-minimal.
(iv) is a special case of (v), since for any extension E of $\mathfrak{D}$, $\Gamma \subseteq E$.
(v) Assume $\varphi \in \mathrm{Cn}_{\mathfrak{D}}(\Delta)$ and let E be an extension of $\mathfrak{D}$ such that $\Delta \subseteq E$. Then, as in (i), Mod(E) is $\succ_{\mathfrak{D}}$-minimal. Since Mod(E) $\models \Delta$, we have Mod(E) $\models_{\mathfrak{D}} \Delta$. Hence Mod(E) $\models \varphi$, and thus $\varphi \in$ E. On the other hand, assume $\varphi \notin \mathrm{Cn}_{\mathfrak{D}}(\Delta)$. Then for some $\mathcal{M}$, $\mathcal{M} \models_{\mathfrak{D}} \Delta$ and $\mathcal{M} \not\models \varphi$. By (ii), $\mathrm{Th}(\mathcal{M})$ is an extension of $\mathfrak{D}$. Since $\varphi \notin \mathrm{Th}(\mathcal{M})$ and $\Delta \subseteq \mathrm{Th}(\mathcal{M})$, $\varphi \notin \bigcap\{E \mid E \text{ is an extension of } \mathfrak{D} \text{ such that } \Delta \subseteq E\}$. ■

Notice that $\emptyset$ being the only preferred model corresponds to $\mathfrak{D}$ having an inconsistent extension or having no extension at all. The collapse of these cases can be defended by pointing out that both are boundary cases added for technical convenience, rather than representations of belief states of truly rational agents. (In both cases a rational agent would have to revise his belief state.)

An immediate corollary of proposition 3.7(v) is the monotonicity of $\mathrm{Cn}_{\mathfrak{D}}$. Hence as long as one keeps the default theory constant, the reasoning is monotonic. Default consequence is nonmonotonic because (the facts of) default theories are updated in the light of new information. To capture this in terms of preferential semantics, we need preference models associated with sets of defaults. The preference relation for such a model will be more or less a global version of the relation used for a default theory.

DEFINITION 3.8 Let D be a set of defaults. The *epistemic preference model associated with* D is the model $\langle \mathrm{WD}_L, \succ_D, \models \rangle$, where $\succ_D = \lambda(\rhd_D, \subset)$ and $\rhd_D$ is defined as follows: $\mathcal{N} \rhd_D \mathcal{M}$ iff $\mathcal{N} >_{\langle D, \mathrm{Th}(\mathcal{N})\rangle} \mathcal{M}$. To avoid stacked subscripts, we write $\models_D$ and Cn_D instead of $\models_{\rhd_D}$ and $\mathrm{Cn}_{\rhd_D}$.

PROPOSITION 3.9 Let D be a set of defaults and $\langle \mathrm{WD}_L, \succ_D, \models \rangle$ its associated epistemic preference model. Then $\mathrm{Cn}_D(\Gamma) = \bigcap\{E \mid E \text{ is an extension of } \langle D, \Gamma \rangle\}$.

Proof. Assume $\varphi \in Cn_D(\Gamma)$, or in other words, for any $\mathcal{M} \in WD_L$, $\mathcal{M} \models_D \Gamma$ implies $\mathcal{M} \models \varphi$. Let E be an extension of $\langle D,\Gamma\rangle$. Then $Mod(E) \models \Gamma$. Now suppose that for some $\mathcal{M} \in WD_L$, $Mod(E) >_D \mathcal{M}$ and $\mathcal{M} \models \Gamma$. Then either (1) $Mod(E) >_{\langle D,Th(Mod(E))\rangle} \mathcal{M}$, or (2) $Mod(E) \subset \mathcal{M}$ and $Mod(E) \not<_{\langle D,Th(\mathcal{M})\rangle} \mathcal{M}$. From (1) it follows that $Mod(E) >_{\langle D,E\rangle} \mathcal{M}$, which contradicts the fact that E is an extension of $\langle D,E\rangle$. (2) also leads to a contradiction, since any sequence of defaults $\delta_1,\delta_2,\ldots$ of D which can be used to go from $Mod(\Gamma)$ to $Mod(E)$, can also be used to go from $\mathcal{M}$ to $Mod(E)$. ($\mathcal{M} \models \Gamma$ guarantees the validity of the necessary prerequisites, and $Mod(E) \subset \mathcal{M}$ insures that the relevant justifications remain consistent.) But then $Mod(E) <_{\langle D,Th(\mathcal{M})\rangle} \mathcal{M}$. Hence $Mod(E) \models_D \Gamma$, and thus $Mod(E) \models \varphi$. Therefore, for any extension E of $\langle D,\Gamma\rangle$, $\varphi \in E$.
To show the other direction, assume $\varphi \in \bigcap\{E \mid E \text{ is an extension of } \langle D,\Gamma\rangle\}$. Suppose $\mathcal{M} \models_D \Gamma$. Then $\mathcal{M} \models \Gamma$ and $\neg\exists\, \mathcal{N} <_D \mathcal{M}$, $\mathcal{N} \models \Gamma$. It follows that either (1) $\mathcal{M} = Mod(\Gamma)$, or (2) $Mod(\Gamma) >_D \mathcal{M}$. In either case, $\mathcal{M} = \varnothing$ or $\mathcal{M}$ is a consistent extension of $\langle D,\Gamma\rangle$. Hence $\mathcal{M} \models_D \varphi$, and $\varphi \in Cn_D(\Gamma)$. ■

In contrast to $>_{\mathfrak{D}}$ of 3.7, $>_D$ of 3.9 is in general not a strict partial order. In fact, it can be shown that some sets of defaults cannot be captured by transitive epistemic preference models. However, in Voorbraak (1991) it is proved that transitive preference models can capture **DL**, provided the notion of preferred model is strengthened as follows:

$$\begin{aligned} \mathcal{M} \models_{L,>} \Gamma \text{ iff } & \mathcal{M} \models_L \Gamma, \\ & \text{for all } \mathcal{N} < \mathcal{M},\ \mathcal{N} \not\models_L \Gamma, \text{ and} \\ & \text{for all } \mathcal{N}' \ (\mathcal{M} \subset \mathcal{N}' \subseteq Mod(\Gamma) \text{ implies } \mathcal{M} < \mathcal{N}'). \end{aligned}$$

A similar result has been obtained independently by Lin and Shoham (1992). As far as we know, our preferential semantics for **DL** is the first one that uses the original definition of preferred model (at the cost of allowing intransitive preference relations).

4 Consequence operations of epistemic preference models

In this section we address the problem of characterizing the consequence operations of epistemic preference models. Notice that we have defined the intersection of extensions to represent the consequences of a default theory, even though this intersection is not necessarily itself an extension. Indeed, we prefer the sceptical interpretation of nonmonotonic logics. However, as a compromise to the more brave or credulous authors, we introduce extension operations next to the nonmonotonic consequence operations. Let **L** denote an arbitrary standard logic in the language $\mathcal{L}_{\mathbf{L}}$.

DEFINITION 4.1 An *extension operation* Ext for **L** is a function $\mathcal{P}(\mathcal{L}_{\mathbf{L}}) \to \mathcal{P}(\mathcal{P}(\mathcal{L}_{\mathbf{L}}))$ satisfying the following three conditions:

(i) If $\Delta \in \mathrm{Ext}(\Gamma)$, then $\Gamma \subseteq \Delta$ *(inclusion)*

(ii) If $\Gamma' \in \mathrm{Ext}(\Gamma)$ and $\Gamma \subseteq \Delta \subseteq \Gamma'$, then $\cap\mathrm{Ext}(\Delta) \subseteq \Gamma'$ *(cumulative transitivity)*

(iii) $\mathrm{Ext}(\Gamma) = \mathrm{Ext}(\mathrm{Cn}_{\mathbf{L}}(\Gamma))$ *(**L**-invariance)*

DEFINITION 4.2 Cn is called a *nonmonotonic consequence operation for **L*** iff for some extension operation Ext for **L**, for all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Cn}(\Gamma) = \cap\mathrm{Ext}(\Gamma)$.

Notice the explicit reference to the underlying standard logic in definition 4.1. For nonmonotonic logics formulated without an underlying standard logic, one can assume **L** to be trivial in the sense that for all Γ, $\mathrm{Cn}_{\mathbf{L}}(\Gamma) = \Gamma$. In that case, **L**-invariance trivially holds. We do not claim that **L**-invariance is reasonable for all nonmonotonic logics, but the condition seems inevitable for nonmonotonic reasoning of ideally rational agents, since they are able to draw all standard consequences of the premises.

Extension operations will be called *equivalent* iff they induce the same consequence operation. It is easy to see that for every extension operation Ext there exists an equivalent Ext' such that for all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Ext}'(\Gamma) \neq \varnothing$. (Just replace all values $\varnothing$ with $\{\mathcal{L}_{\mathbf{L}}\}$.) We will not distinguish between having no extensions and having only the inconsistent extension (since in both situations the belief state is inconsistent), and we simply assume from now on that only non-empty sets are in the range of extension operations.

In the following proposition some properties of nonmonotonic consequence operations for **L** are listed.

PROPOSITION 4.3 Let Cn be a nonmonotonic consequence operation for **L**.

(i) For all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\Gamma \subseteq \mathrm{Cn}(\Gamma)$ *(inclusion)*

(ii) For all $\Gamma,\Delta \subseteq \mathcal{L}_{\mathbf{L}}$, if $\Gamma \subseteq \Delta \subseteq \mathrm{Cn}(\Gamma)$, then $\mathrm{Cn}(\Delta) \subseteq \mathrm{Cn}(\Gamma)$ *(cumulative transitivity)*

(iii) For all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Cn}(\Gamma) = \mathrm{Cn}(\mathrm{Cn}_{\mathbf{L}}(\Gamma))$ *(**L**-invariance)*

(iv) For all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Cn}(\Gamma) = \mathrm{Cn}(\mathrm{Cn}(\Gamma))$ *(idempotency)*

(v) For all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Cn}(\Gamma) = \mathrm{Cn}_{\mathbf{L}}(\mathrm{Cn}(\Gamma))$

(vi) For all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Cn}_{\mathbf{L}}(\Gamma) \subseteq \mathrm{Cn}(\Gamma)$

Proof. (i), (ii), and (iii) follow immediately from their corresponding conditions on Ext mentioned in definition 4.1. Idempotency (iv) follows from inclusion and cumulative transitivity. (Apply cumulative transitivity for $\Delta = \mathrm{Cn}(\Gamma)$.) Property (v) follows from inclusion for $\mathrm{Cn}_{\mathbf{L}}$ and the fact that $\mathrm{Cn}_{\mathbf{L}}(\mathrm{Cn}(\Gamma)) \subseteq_{(i)} \mathrm{Cn}(\mathrm{Cn}_{\mathbf{L}}(\mathrm{Cn}(\Gamma))) =_{(iii)} \mathrm{Cn}(\mathrm{Cn}(\Gamma)) =_{(iv)} \mathrm{Cn}(\Gamma)$. To show (vi), notice that $\Gamma \subseteq \mathrm{Cn}(\Gamma)$ implies $\mathrm{Cn}_{\mathbf{L}}(\Gamma) \subseteq \mathrm{Cn}_{\mathbf{L}}(\mathrm{Cn}(\Gamma))$, and, by (v), $\mathrm{Cn}(\Gamma) = \mathrm{Cn}_{\mathbf{L}}(\mathrm{Cn}(\Gamma))$. ■

Nonmonotonic consequence operations for **L** are in fact characterized by the first three properties of 4.3.

PROPOSITION 4.4 Cn is a nonmonotonic consequence operation for **L** iff Cn satisfies inclusion, cumulative transitivity, and **L**-invariance.

Proof. One half of the proposition already follows from 4.3. To show the other half, let Cn satisfy inclusion, cumulative transitivity, and **L**-invariance. Define an operation $\mathrm{Ext} : \mathcal{P}(\mathcal{L}_{\mathbf{L}}) \to \mathcal{P}(\mathcal{P}(\mathcal{L}_{\mathbf{L}}))$ as follows: for every $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Ext}(\Gamma) = \{\mathrm{Cn}(\Gamma)\}$. Then it is easy to verify that Ext is an extension operator and that for every $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathrm{Cn}(\Gamma) = \bigcap \mathrm{Ext}(\Gamma)$. ■

Thus far, we only considered properties of extension operations that correspond to properties of consequence operations. However, in general, extension operations give a more detailed description of a logic than consequence operations, since different extension operations may correspond to the same consequence operation. Let us mention some possible properties of extension operations which have no counterparts formulated in terms of consequence operations.

If $\Gamma,\Delta \in \mathrm{Ext}(\Sigma)$ and $\Gamma \subseteq \Delta$, then $\Gamma = \Delta$. (*no nesting*)
If $\Gamma,\Delta \in \mathrm{Ext}(\Sigma)$, then $\Gamma \cap \Delta \in \mathrm{Ext}(\Sigma)$. (*intersection*)

The first property could perhaps be motivated by the intuition that a rational agent should decide between two (compatible) states of belief which mainly differ in their degree of ignorance. The intersection property expresses the non-existence of the multiple-extension problem for the logic at hand. For each nonmonotonic consequence operation Cn, the extension operation Ext, defined by $\mathrm{Ext}(\Gamma) =_{\mathrm{def}} \{\mathrm{Cn}(\Gamma)\}$, satisfies both no nesting and intersection, and the thus defined extension operation is the only one which satisfies both conditions, since they together imply $\mathrm{Ext}(\Sigma) = \{\bigcap \mathrm{Ext}(\Sigma)\}$.

Extension operations are particularly suited for the abstract study of credulous (or brave) nonmonotonic inference, but since we prefer the sceptical interpretation of nonmonotonic inference, we do not pursue the independent study of extension operations any further, and return to consequence operations.

It turns out that nonmonotonic consequence operations for **L** are exactly the consequence operations of epistemic preference models for **L**.

PROPOSITION 4.5 Cn is a nonmonotonic consequence operation for **L** iff Cn = $\mathrm{Cn}_{\succ}$, for some epistemic preference model $\langle \mathrm{WD}_{\mathbf{L}}, \succ, \models \rangle$.

Proof. Let $\langle \mathrm{WD}_{\mathbf{L}}, \succ, \models \rangle$ be an epistemic preference model. Since epistemic preference models are Makinson models, it follows from observation 5 of Makinson (1989) that $\mathrm{Cn}_{\succ}$ satisfies inclusion and cumulative transitivity. Since for any $\mathcal{M} \in \mathrm{WD}_{\mathbf{L}}$, $\mathcal{M} \models \Sigma$ iff $\mathcal{M} \models \mathrm{Cn}_{\mathbf{L}}(\Sigma)$, $\mathrm{Cn}_{\succ}$ is **L**-faithful. To show the other direction, assume Cn is a nonmonotonic consequence operation for **L**. Define $\succ = \lambda(>,\subset)$, where $>$ is the asymmetric version of $\geq = \{\langle \mathcal{N}, \mathcal{M} \rangle \mid$ for some $\Sigma \subseteq \mathcal{L}_{\mathbf{L}}$, $\mathcal{N} \subseteq \mathrm{Mod}(\Sigma)$ and $\mathcal{M} = \mathrm{Mod}(\mathrm{Cn}(\Sigma))\}$. Then Cn = $\mathrm{Cn}_{\succ}$, since $\mathcal{M} \models_{\succ} \Sigma$ iff for some $\Delta \subseteq \Sigma$, $\mathcal{M} = \mathrm{Mod}(\mathrm{Cn}(\Delta)) \subseteq \mathrm{Mod}(\mathrm{Cn}(\Sigma))$. ■

Proposition 4.5 is also true if epistemic preference models are replaced by **L**-faithful Makinson models.

5 Cumulativity and rationality

Gabbay (1985) introduces the notion of weakly monotonic, or cumulative, consequence, which is characterized in Makinson (1989) and Kraus et al. (1990) in terms of preferential models. We repeat Makinson's characterization of cumulative consequence operations in terms of Makinson models.

DEFINITION 5.1 A *cumulative* consequence operation is a consequence operation which satisfies inclusion ($\Gamma \subseteq \mathrm{Cn}(\Gamma)$), idempotency ($\mathrm{Cn}(\Gamma) = \mathrm{Cn}(\mathrm{Cn}(\Gamma))$), and cumulative monotonicity ($\Gamma \subseteq \Delta \subseteq \mathrm{Cn}(\Gamma) \Rightarrow \mathrm{Cn}(\Gamma) \subseteq \mathrm{Cn}(\Delta)$).

DEFINITION 5.2 A Makinson model $\langle \mathfrak{M}, \succ, \models \rangle$ is called *stoppered* iff for all $\Gamma \subseteq \mathcal{L}_{\mathbf{L}}$ and for all $\mathcal{M} \in \mathfrak{M}$, $\mathcal{M} \models \Gamma$ implies $\mathcal{M} \models_{\succ} \Gamma$ or $\exists \mathcal{N} \prec \mathcal{M}\ \mathcal{N} \models_{\succ} \Gamma$.

PROPOSITION 5.3 (Makinson (1989)) Cn is a cumulative consequence operation iff Cn = $\mathrm{Cn}_{\succ}$, for some stoppered Makinson model $\langle \mathfrak{M}, \models, > \rangle$.

Below we show that this result readily generalizes to a characterization of cumulative consequence relations for **L** in terms of **L**-faithful Makinson models (see proposition 5.6) or in terms of epistemic preference models (see proposition 5.8).

DEFINITION 5.4 An extension operation Ext for **L** is called *cumulative* iff it satisfies the following condition of *cumulative monotonicity**: If $\Gamma' \in \text{Ext}(\Gamma)$, $\Delta' \in \text{Ext}(\Delta)$ and $\Gamma \subseteq \Delta \subseteq \Gamma'$, then $\Delta' \in \text{Ext}(\Gamma)$. In other words, if $\Gamma' \in \text{Ext}(\Gamma)$ and $\Gamma \subseteq \Delta \subseteq \Gamma'$, then $\text{Ext}(\Delta) \subseteq \text{Ext}(\Gamma)$. A *cumulative consequence operation for* ***L*** is a nonmonotonic consequence operation induced by a cumulative extension operation for **L**.

PROPOSITION 5.5 Cumulative consequence operations for **L** are characterized by inclusion, cumulative transitivity, **L**-invariance, and the property of cumulative monotonicity for consequence operations ($\Gamma \subseteq \Delta \subseteq \text{Cn}(\Gamma) \Rightarrow \text{Cn}(\Gamma) \subseteq \text{Cn}(\Delta)$.)

Proof. Let Ext be a cumulative extension operator for **L** and let $\text{Cn}(\Gamma) = \bigcap\text{Ext}(\Gamma)$. Then, by proposition 4.3, the operator Cn satisfies inclusion, cumulative transitivity, and **L**-invariance. To show that Cn satisfies cumulative monotonicity, assume $\Gamma \subseteq \Delta \subseteq \text{Cn}(\Gamma)$. Then $\Gamma \subseteq \Delta \subseteq \bigcap\text{Ext}(\Gamma)$. Hence for some $\Gamma' \in \text{Ext}(\Gamma)$, $\Gamma \subseteq \Delta \subseteq \Gamma'$. Therefore, $\text{Ext}(\Delta) \subseteq \text{Ext}(\Gamma)$, and $\text{Cn}(\Gamma) = \bigcap\text{Ext}(\Gamma) \subseteq \bigcap\text{Ext}(\Delta) = \text{Cn}(\Delta)$. To show the other direction, let $\text{Cn} : \mathcal{P}(\mathcal{L}_\mathbf{L}) \rightarrow \mathcal{P}(\mathcal{L}_\mathbf{L})$, and define $\text{Ext}(\Gamma) = \{\text{Cn}(\Gamma)\}$. In the previous section we noted that inclusion, cumulative transitivity, and **L**-invariance for Cn imply these properties for Ext. So we only have to show that additionally assuming cumulative monotonicity for Cn, implies this property for Ext. Let $\Gamma' \in \text{Ext}(\Gamma)$, $\Delta' \in \text{Ext}(\Delta)$ and $\Gamma \subseteq \Delta \subseteq \Gamma'$. Then $\Gamma' = \text{Cn}(\Gamma)$, $\Delta' = \text{Cn}(\Delta)$ and $\Gamma \subseteq \Delta \subseteq \Gamma'$. Hence, by cumulative monotonicity and cumulative transitivity, $\text{Cn}(\Gamma) = \text{Cn}(\Delta)$ and thus $\Gamma' = \Delta'$. Since $\Gamma' \in \text{Ext}(\Gamma)$, it follows that $\Delta' \in \text{Ext}(\Gamma)$. ■

PROPOSITION 5.6 Cn is a cumulative consequence operation for **L** iff $\text{Cn} = \text{Cn}_{>}$, for some stoppered **L**-faithful Makinson model $\langle \mathfrak{M}, >, \models \rangle$.

Proof. It is shown in Makinson (1989) that for any stoppered Makinson model $\langle \mathfrak{M}, >, \models \rangle$, $\text{Cn}_{>}$ satisfies inclusion, cumulative transitivity and cumulative monotonicity. It is clear that $\text{Cn}_{>}$ satisfies **L**-invariance whenever $\langle \mathfrak{M}, >, \models \rangle$ is **L**-faithful. That any cumulative consequence operation for **L** is $\text{Cn}_{>}$, for some stoppered **L**-faithful Makinson model $\langle \mathfrak{M}, >, \models \rangle$ follows from proposition 5.8 below, since stoppered epistemic preference models are stoppered **L**-faithful Makinson models. ■

---

* Notice that we use the same name for a different, though closely related, condition on *consequence* operations, namely $\Gamma \subseteq \Delta \subseteq \text{Cn}(\Gamma) \Rightarrow \text{Cn}(\Gamma) \subseteq \text{Cn}(\Delta)$.

Since epistemic preference models are a special kind of Makinson models, definition 5.2 also defines stoppered epistemic preference models. The following proposition gives a slightly different characterization of these models.

PROPOSITION 5.7 An epistemic preference model $\langle WD_L, \succ, \models \rangle$ is stoppered iff for every $\mathcal{M} \in WD_L$ each element of the set $\{\mathcal{N} \in WD_L \mid \mathcal{N} \subseteq \mathcal{M}\}$ is $\succeq$ some $\prec$-minimal element of that set.

Proof. The result follows immediately from the facts that (1) $\mathcal{M} \in WD_L$ iff $\mathcal{M} = Mod_L(\Gamma)$, for some $\Gamma \subseteq \mathcal{L}_L$ and (2) for all $\mathcal{N} \in WD_L$, $\mathcal{N} \models \Gamma$ iff $\mathcal{N} \subseteq \mathcal{M}$. ∎

At first sight, the above characterization of being stoppered seems to be much neater than definition 5.2, since no reference is made to the satisfaction relation. (Notice that in modal logic interesting results are obtained by relating modal formulas to *frame* properties, which do not refer to the satisfaction relations of the models.) However, as remarked before, an epistemic preference model is completely determined by WD_L and $\succ$. Hence the reference to $\models$ is still implicitly present.

Cumulative consequence operations for **L** are not only characterized by stoppered **L**-faithful Makinson models, but also by stoppered epistemic preference models for **L**.

PROPOSITION 5.8 Let **L** be a standard logic. Cn is a cumulative consequence operation for **L** iff Cn = $Cn_\succ$, for some stoppered epistemic preference model $\langle WD_L, \succ, \models \rangle$.

Proof. The "if" part follows from the "if" part of proposition 5.6. For the "only if" part we proceed as in proposition 4.5. Let Cn be a cumulative consequence operation for **L**. Define $\succ = \lambda(>, \subset)$, where $>$ is the asymmetric version of $\geq = \{\langle \mathcal{N}, \mathcal{M} \rangle \mid$ for some $\Sigma \subseteq \mathcal{L}_L$, $\mathcal{N} \subseteq Mod(\Sigma)$ and $\mathcal{M} = Mod(Cn(\Sigma))\}$. Then Cn = $Cn_\succ$, since $\mathcal{M} \models_\succ \Sigma$ iff for some $\Delta \subseteq \Sigma$, $\mathcal{M} = Mod(Cn(\Delta)) \subseteq Mod(Cn(\Sigma))$. It remains to show that $\langle WD_L, \succ, \models \rangle$ is stoppered. Assume $\mathcal{M} \models \Gamma$. Consider $\mathcal{N} = Mod(Cn(\Gamma))$. Since $\mathcal{M} \subseteq Mod(\Gamma)$, either $\mathcal{M} \succ \mathcal{N} \models_\succ \Gamma$, or for some $\Sigma \subseteq \mathcal{L}_L$, $\mathcal{N} \subseteq Mod(\Sigma)$ and $\mathcal{M} = Mod(Cn(\Sigma))$. From the second possibility we can derive $\Gamma \subseteq Cn(\Sigma)$ and $\Sigma \subseteq Cn(\Gamma)$, which for any cumulative consequence operation Cn implies $Cn(\Sigma) = Cn(\Gamma)$. Hence we obtain either $\mathcal{M} \succ \mathcal{N} \models_\succ \Gamma$, or $\mathcal{M} = \mathcal{N}$. Therefore, $\langle WD_L, \succ, \models \rangle$ is stoppered. ∎

The obtained characterizations of consequence operations in terms of classes of preference models are summarized in the picture below.

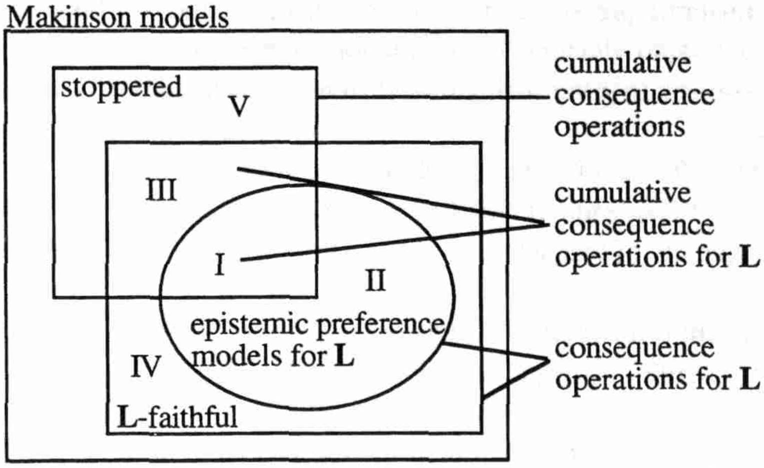

Figure 5.1. Summary of obtained characterization results.
Cumulative consequence operations are characterized by the models in I∪II∪V. Cumlative consequence operations for **L** correspond to the models in I∪II or the models in I∪II∪III∪IV. Consequence operations for **L** correspond to the models in I(∪III).

Cumulative monotonicity is considered to be a rationality requirement, and there is a tendency to disqualify non-cumulative nonmonotonic logics, such as **DL**. For example, in Brewka (1990) **DL** is modified to satisfy cumulative monotonicity. However, we are not prepared to conclude that the consequence operation of a nonmonotonic logic has to be cumulative. First of all, nonmonotonic logics might be used to formalize the reasoning of agents or systems which are not ideally rational and which, for example, reason nonmonotonically by "firing" default rules in appropriate circumstances.

Further, it cannot be inferred from the rationality of cumulative monotonicity that any nonmonotonic logic formalizing the nonmonotonic reasoning of an ideally rational agent has to be cumulative. There is an analogy here with consistency: Although an ideally rational agent only believes a consistent set of formulas, we do not have to require that the logic **L** under which the beliefs are closed is consistent in the sense that for all $\Sigma \subseteq \mathcal{L}_{\mathbf{L}}$, $Cn_{\mathbf{L}}(\Sigma) \neq \mathcal{L}_{\mathbf{L}}$.* An inconsistent set Σ will be revised before it will become accepted by a rational agent, and this revision process is not described by **L**, but by operations as studied in Gärdenfors (1988). Similarly, a rational agent will revise his default beliefs if they do not give rise to rational preferences, and this revision process does not have to be described by the nonmonotonic consequence operation.

In addition to cumulativity, there is another possible rationality requirement on default beliefs and their induced nonmonotonic consequence operations,

---

* In fact, this requirement is inconsistent with the inclusion property, satisfied by standard logics.

namely that $Cn(\Gamma)$ is inconsistent only if Γ is. Let us call an extension operation Ext for **L** *consistent* iff for all $\Gamma \subseteq \mathscr{L}_{\mathbf{L}}$, if $Cn_{\mathbf{L}}(\Gamma) \neq \mathscr{L}_{\mathbf{L}}$, then $Ext(\Gamma) \neq \{\mathscr{L}_{\mathbf{L}}\}$. It is easy to see that consistent nonmonotonic consequence operations Cn for **L** are characterized by the following condition: for all $\Gamma \subseteq \mathscr{L}_{\mathbf{L}}$, if $Cn_{\mathbf{L}}(\Gamma) \neq \mathscr{L}_{\mathbf{L}}$, then $Cn(\Gamma) \neq \mathscr{L}_{\mathbf{L}}$.

In general, it is of course not easy to guarantee that a set of default beliefs induces a consistent consequence operation. However, there exist interesting special cases for which simple sufficient conditions can be formulated. For example, it is known that a default theory $\langle D,\Gamma\rangle$ has an inconsistent extension iff the set Γ of facts is inconsistent. Hence every set D of defaults which guarantees the existence of extensions, such as a set of normal defaults, induces a consistent consequence operation.

Not every set of normal defaults induces a cumulative consequence operation. For example, the set $\{: p / p, p \vee q : \neg p / \neg p\}$, which is used by Makinson (1989) to show that default logic is not cumulative, consists of normal defaults. But it is quite simple to revise a set of normal defaults into a set that yields a cumulative consequence operation. Besnard (1989) proposes the following translation of normal defaults into *free defaults*, i.e., defaults without prerequisites: $Tr_{Bes}(\alpha : \beta / \beta) = : \alpha \supset \beta / \alpha \supset \beta$. It is shown in the next section that sets of thus obtained defaults induce a cumulative consequence operation. However, since Tr_{Bes} lacks a proper justification and gives rise to some counterinuitive results, we also consider an alternative, more general translation Tr_{free} defined by $Tr_{free}(\alpha : \beta_1,\ldots,\beta_n / \omega) =_{def} : \alpha\wedge\beta_1,\ldots,\alpha\wedge\beta_n / \alpha \supset \omega$.

Unfortunately, Tr_{free} does not necessarily revise a set of (normal) defaults into a set which induces a cumulative consequence operation. But Tr_{free} can be motivated by a famous rationality property of preference relations that thus far has received little or no attention in the literature on preferential semantics, namely the *sure-thing principle* of Savage (1972), which is informally stated thus:

> If the person would not prefer **f** to **g**, either knowing that the event B obtained, or knowing that the event ~B obtained, then he does not prefer **f** to **g**. –Savage (1972, p. 21)

Another formulation of the principle is:

> If **f**, **g**, and **f'**, **g'** are such that:
> 1. in ~B, **f** agrees with **g**, and **f'** agrees with **g'**,
> 2. in B, **f** agrees with **f'**, and **g** agrees with **g'**,
> 3. $\mathbf{f} \leq \mathbf{g}$;
>
> then $\mathbf{f'} \leq \mathbf{g'}$. –Savage (1972, p. 23)

A reasonable formalization of the sure-thing principle in epistemic preference models is given below.

DEFINITION 5.9 An epistemic preference model $\langle WD_L, >, \models \rangle$ is said to *satisfy the sure-thing principle* iff for all world descriptions $\mathcal{M}$, $\mathcal{N}$, $\mathcal{M}'$, $\mathcal{N}'$, such that $\mathcal{N} \subseteq \mathcal{N}'$ and $\mathcal{M} \cap \mathcal{N} \geqslant \mathcal{M}' \cap \mathcal{N}$, it is the case that $(\mathcal{M} \cap \mathcal{N}) \cup (\mathcal{N}' \backslash \mathcal{N}) \geqslant (\mathcal{M}' \cap \mathcal{N}) \cup (\mathcal{N}' \backslash \mathcal{N})$.

Let us give an informal justification of this formalization: Think of $\mathcal{N}$ as the set of worlds of $\mathcal{N}'$ in which B holds. $\mathcal{M} \cap \mathcal{N}$ and $\mathcal{M}' \cap \mathcal{N}$ play the role of **f** and **g**, respectively, and **f'** and **g'** are translated as $(\mathcal{M} \cap \mathcal{N}) \cup (\mathcal{N}' \backslash \mathcal{N})$ and $(\mathcal{M}' \cap \mathcal{N}) \cup (\mathcal{N}' \backslash \mathcal{N})$, respectively. World-descriptions are said to agree iff they are equal and $\leq$ is translated as $\geqslant$. See also the picture in figure 5.2.

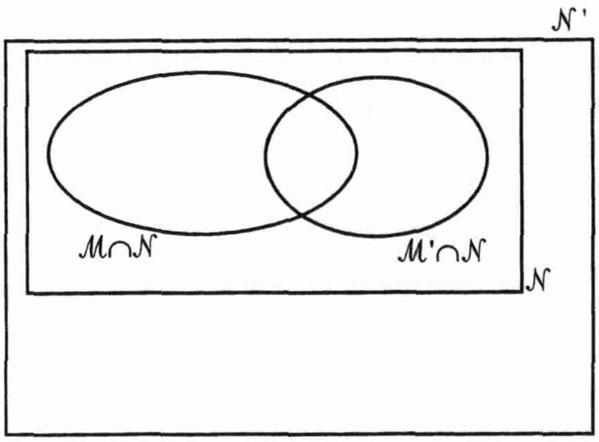

Figure 5.2. Illustration of the sure-thing principle.
If $\mathcal{M} \cap \mathcal{N} \geqslant \mathcal{M}' \cap \mathcal{N}$, then by the sure-thing principle it should also be the case that $(\mathcal{M} \cap \mathcal{N}) \cup (\mathcal{N}' \backslash \mathcal{N}) \geqslant (\mathcal{M}' \cap \mathcal{N}) \cup (\mathcal{N}' \backslash \mathcal{N})$.

However, just as we do not require the preference relation of an epistemic preference model to be transitive, we do not require the epistemic preference models to satisfy the sure-thing principle. We believe that this principle is best applied at the level of the individual preferential criteria, which are aggregated by the agent into a preference relation. In Voorbraak (1993), it is shown that the sure-thing principle allows the inference of a preference induced by a default δ, to the preference induced by $Tr_{free}(\delta)$.

An obvious objection against Tr_{free} and any other translation of defaults into free defaults is that a rule like BIRD(TWEETY) : FLY(TWEETY) / FLY(TWEETY) is supposed to "fire" only whenever it has become *known* that Tweety is a bird and it is supposed to be ignored in all other cases. But this objection implicitly interprets **DL** as the logic of some agent or system for which computational issues might matter, and not as the logic of an ideally rational agent.

The above suggests that an ideally rational agent should only believe free defaults. Therefore, we consider it worthwhile to look at the properties of free default theories in some detail. This is done in Voorbraak (1993).

6 Conclusion

A preference-based semantics for *all* main nonmonotonic logics (including Default Logic) can be given, provided one generalizes the framework originally proposed by Shoham, and allows not necessarily transitive preference relations between (sets of) models. This move can be epistemically motivated, since it is reasonable to prefer less specific world-descriptions unless one has evidence pointing towards more specific world-descriptions, and this results in a not necessarily transitive preference relation. Therefore, the preference models that we introduce are called epistemic preference models.

The consequence operations induced by epistemic preference models are called nonmonotonic consequence operations for **L**, where **L** is the underlying standard logic. Nonmonotonic consequence operations for **L** are characterized by the well-known properties inclusion and cumulative transitivity, together with **L**-invariance ($Cn(\Gamma) = Cn(Cn_L(\Gamma))$). Completely analogous to a result of Makinson (1989) it is shown that cumulative consequence operations for **L** are characterized by stoppered epistemic preference models.

We argue that cumulativity can be interpreted as a rationality constraint, but that this does not imply that noncumulative logics (such as Default Logic) should be dismissed. We prefer to let cumulativity (and other rationality principles) be a constraint on the set of defaults that a rational agent should believe. In this vein, we mention the close connection between defaults without prerequisites (free defaults) and Savage's sure-thing principle, which has thus far received little attention in the literature on nonmonotonic reasoning.

References

Besnard, P. 1989. *An Introduction to Default Logic*, Berlin: Springer.

Brewka, G. 1990. Cumulative default logic: in defense of nonmonotonic inference rules. *Artificial.Intelligence* 50:183-205.

Dix, J. and Makinson, D. 1992. The relationship between KLM and MAK models for nonmonotonic inference operations. *Journal of Logic, Language, and Information* 1: 131-140..

Etherington, D.W. 1988. *Reasoning with Incomplete Information,* Research Notes in Artificial Intelligence, London: Pitman.

Gabbay, D. 1985. Theoretical foundations for non-monotonic reasoning in expert systems. *Logics an Models of Concurrent Systems*, K.R. Apt (ed.), Berlin: Springer, pp.439-457.

Gärdenfors, P. 1988. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Cambridge Mass.: MIT Press.

Halpern, J.Y. and Moses, Y. 1985. Towards a theory of knowledge and ignorance: preliminary report. *Logics and Models of Concurrent Systems*, K.R. Apt (ed.), Berlin: Springer, pp. 459-476.

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

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

Lin, F. and Shoham, Y. 1992. A logic of knowledge and justified assumptions. *Artificial Intelligence* , 57: 271-289.

Łukaszewicz, L. 1985. Two results on default logic. *Proceedings IJCAI-85*, Los Altos: Morgan Kaufmann, pp. 459-461.

Makinson, D. 1989. General theory of cumulative inference. *Non-Monotonic Reasoning*, (Proc. 2nd international workshop), M. Reinfrank, J. de Kleer, M.L. Ginsberg, E. Sandewall (eds.), Berlin: Springer (1989), pp. 1-18.

McCarthy, J. 1980. Circumscription–A form of non-monotonic reasoning. *Artificial Intelligence* 13: 27-39.

McDermott, D. 1982. Nonmonotonic logic II: nonmonotonic modal theories *Journal of the ACM* 29: 33-57.

Moore, R.C. 1984. Possible-world semantics for autoepistemic logic. *Proc. 1984 Non-Monotonic Workshop*, New Paltz, NY, pp 344-354. Reprinted in *Readings in Nonmonotonic Reasoning*, M.L. Ginsberg (ed.), Los Altos: Morgan Kaufmann (1987), pp. 137-142.

Moore, R.C. 1985. Semantical considerations on nonmonotonic logic. *Artificial Intelligence* 25: 75-94.

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

Savage, L.J. 1972. *The Foundations of Statistics*, New York: Dover. (first edition: New York: John Wiley (1954))

Shoham, Y. 1987. A semantical approach to nonmonotonic logics. *Readings in Nonmonotonic Reasoning*, M.L. Ginsberg (ed.), Los Altos CA: Morgan Kaufmann, pp. 227-250.

Shoham, Y. 1988. *Reasoning about Change*, Cambridge MA: MIT Press.

Shoham, Y. and Moses, Y. 1989. Belief as defeasible knowledge. *Proc. IJCAI 89*, Los Altos: Morgan Kaufmann, pp. 1168-1173.

Voorbraak, F. 1991. A preferential model semantics for default logic, *Symbolic and Quantitative Approaches to Uncertainty*, R. Kruse and P. Siegel (eds.), Lecture Notes in Computer Science 548, Berlin: Springer, pp. 344-351.

Voorbraak, F. 1992. Generalized Kripke models for epistemic logic. *Theoretical Aspects of Reasoning About Knowledge: Proceedings of the Fourth Conference*, Y. Moses (ed.). San Mateo CA: Morgan Kaufmann, pp. 214-228.

Voorbraak, F. 1993. *As Far as I Know. Epistemic Logic and Uncertainty.* Dissertation, Utrecht University.