Preference-based semantics for nonmonotonic logics*

Frans Voorbraak
Department of Philosophy
Utrecht University
P.O. Box 80126
3508 TC Utrecht
The Netherlands

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 arc aggregated into one preference order. This allows the capturing of *all* main nonmonotonic formalisms, including Default Logic oi Reiter [1980]. The preferential models introduced in this paper are motivated from an cpistemic point *of* view, and are therefore called cpistemic preference models. The consequence operations induced by cpistemic preference models are characterized. Further, the view is defended that the rationality of cumulative monotonicity does not imply thai 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.

1 Introduction

Shoham [1987; 1988] introduces preferential semantics as a possible unifying framework of nonmonotonic logics. In this 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 T preferentially entails φ iff (p is true in every model M such that (1) *V* is true in M and (2) r has no model N which is preferred to M.

In this paper we propose a generalization of Shoham's framework which is obtained by allowing (not necessarily transitive) preference *relations* between sets of *models.* 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 preference.

The investigations were supported by the Foundation for Philosophical Research (SWON), which is subsidized by the Netherlands Organization for Scientific Research (NWO).

A consequence of the additional freedom is that in our approach Default Logic of [Reiter, 1980), which is one of the major 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 Default Logic, but it follows from our interpretation of the preference relation.

We interpret the preferences between (sets of) models as cpistemic preferences (of an ideally rational agent) between world descriptions. Since an agent usually has only partial information about the world, cpistemic preferences are most naturally captured by preferences between *sets* of models, corresponding to *partial* world descriptions. In our view, an agent has in general cpistemic 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 beliefs (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 ram. 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 rams 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 *oi* strict partial orders, which itself is in general not a strict partial order. (See section 2 below.) 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 *el al.,* 1990], the boundedness condition of [Makinson, 1989), etcetera. We return to this issue in section 5.

The rest of this paper is organized as follows. In section 2 some preliminary definitions and results concerning preference relations and preferential semantics are treated. In section 3, Default Logic is given a preferential semantics. Section 4 contains a description of the consequence operation characterized by the preferential models of this paper. Cumulativity and other rationality requirements are discussed in section 5. Finally, in section 6, the main conclusions are mentioned.

Proofs of the results can be found in the full paper [Voorbraak, 1993].

2 Preliminaries

Throughout this paper, $>$ denotes an asymmetric, and therefore irreflexive, binary relation

Definition 2.1 For any asymmetric $>$,

(i)	$x \sim y$	$\Leftrightarrow_{def}$	not $x > y$ and not $y > x$.
(ii)	$x \approx y$	$\Leftrightarrow_{def}$	$x \sim z$ iff $y \sim z$, for all $z \in X$.
(iii)	$x \gtrapprox y$	$\Leftrightarrow_{def}$	$x > y$ or $x \approx y$.
(iv)	$x \geqq y$	$\Leftrightarrow_{def}$	$x > y$ or $x = y$.

In addition to definition 2.1 we use standard notation, such as $<$ for the converse of $>$ and $\not>$ for the complement of $>$. Notice that $\approx$ is an equivalence relation, that $> \cap \approx = \varnothing$, and that $\approx = \gtrapprox \cap \lessapprox$. It follows that instead of $>$ one may choose $\gtrapprox$ as primitive notion, and define $>$ as $\gtrapprox \cap \not\lessapprox$. Of course, also $\geqq$ can be chosen as primitive.

The intuitive reading of $x > y$ is "y is preferred to x". The intuitive reading of the other relations mentioned in definition 2.1 is as follows:

- $x \sim y$: x and y are preferentially unrelated. ($\sim$ is called *indifference* relation.)
- $x \approx y$: x and y are equally preferred.
- $x \gtrapprox y$: y is at least as preferred as x.
- $x \geqq y$: y is identical to x or y is preferred to x

We use the preference relations to model the (nonmonotonic) reasoning of a *rational* agent and it is usually assumed that a relation expressing *rational* preference is a strict partial order, i.e., a relation which is not only asymmetric, but also transitive. However, we explicitly allow intransitive preference relations, since, as mentioned in the introduction, it is perfectly reasonable for an agent to have a preference for less specific world descriptions, unless there he has some information to the contrary.

The preferences of such an agent can be described by the lexicographic aggregation of two preferential criteria, which lets a second preference criterion apply if (and only if) the first criterion is indifferent with respect to the choices at hand. For weak orders, the notion of lexicographic aggregation is well-known. (A *weak order* is a strict partial order for which $\gtrapprox$ is total or complete, in the sense that either $x \gtrapprox y$ or $y \gtrapprox x$ holds.

Since for weak orders $\sim = \approx$, while in general $\sim$ and $\approx$ are distinct, there exist at least two reasonable notions of lexicographic aggregation in the general case. We will use only one of these, so we do not bother to distinguish between weak and strong variants.

Definition 2.2 Let for all positive integers $i \leq n$, $>_i$ be an asymmetric binary relation on X. $\lambda(\langle >_1,\ldots,>_n\rangle)$, the *lexicographic aggregation* of $>_1,\ldots,>_n$, is defined inductively as follows: $\lambda(\langle\,\rangle) = \varnothing$, and $\lambda(\langle >_1,\ldots,>_{m+1}\rangle) = \lambda(\langle >_1,\ldots,>_m\rangle) \cup (\sim_1 \cap \ldots \cap \sim_m \cap >_{m+1})$

Let us write $\lambda(>_1,\ldots,>_n)$ instead of $\lambda(\langle >_1,\ldots,>_n\rangle)$. Notice that $\lambda(>) = >$ and that $(x,y) \in \lambda(>_1,>_2)$ iff y is 1-preferred to x, or x and y are 1-indifferent and y is 2-preferred to x. One easily shows that for weak orders $>_1,\ldots,>_n$ on X, $\lambda(>_1,\ldots,>_n)$ is again a weak order on X. However, the following example shows that the lexicographic aggregation of two strict partial orders on X is itself not necessarily a strict partial order on X.

Example 2.3 Let $>_1$ be the strict partial order $\{(x_3,x_1)\}$ and let $>_2$ be the weak order $\{(x_1,x_2),(x_2,x_3),(x_1,x_3)\}$. Then $> =_{def} \lambda(>_1,>_2)$ is not transitive, since $x_1 > x_2$ and $x_2 > x_3$, but not $x_1 > x_3$.

We will 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 relation, we will divide out the indistinguishability relation.

Throughout this paper, **L** denotes a standard logic. That is, the consequence operation Cn_L of **L** satisfies inclusion ($\Gamma \subseteq Cn_L(\Gamma)$), idempotency ($Cn_L(\Gamma) = Cn_L(Cn_L(\Gamma))$), and monotonicity ($\Gamma \subseteq \Delta \Rightarrow Cn_L(\Gamma) \subseteq Cn_L(\Delta)$). $\mathcal{L}_L$ denotes the language of **L** and Mod_L is the class of **L**-models. The preference relation is defined on $\{\mathcal{M} \subseteq Mod_L \mid$ for some $\Sigma \subseteq \mathcal{L}_L, \mathcal{M} = \{M \mid M \models_L \Sigma\}\}$. Hence each relevant set $\mathcal{M}$ of **L**-models is characterized by some subset of $\mathcal{L}_L$. Equivalently, one can define the preferences on $\{\Sigma \subseteq \mathcal{L}_L \mid Cn_L(\Sigma) = \Sigma\}$, the set of **L**-theories.

In [Kraus *et al.*, 1990] a state is merely labelled and *not* identified with a set of worlds. (Cf. def. 2.9 below.) We have taken a similar position in [Voorbraak, 1992], but there we describe the same (epistemic) states in different formal languages, whereas in this paper we restrict ourselves to a single language, viz. $\mathcal{L}_L$, which makes it reasonable to identify objects which cannot be distinguished in that language.

Before we give the formal definition of our models for preferential entailment, we introduce some notation. $\models_{h,L} \subseteq \mathcal{P}(Mod_L) \times \mathcal{L}_L$ is the satisfaction relation given by $\mathcal{M} \models_{h,L} \varphi$ iff for all $M \in \mathcal{M}$, $M \models_L \varphi$. (The notation derives from [Voorbraak, 1991], where sets of models are called hypervaluations.) Notice that $\models_{h,L}$ induces a standard consequence operation. For any set $\mathcal{M} \subseteq Mod_L$, $//\mathcal{M}//_L$ abbreviates $\{\varphi \in \mathcal{L}_L \mid \mathcal{M} \models_{h,L} \varphi\}$ and for any $\Sigma \subseteq \mathcal{L}_L$, $\backslash\backslash\Sigma\backslash\backslash_L =_{def} \{M \in Mod_L \mid M \models_L \Sigma\}$. $\backslash\backslash\Sigma\backslash\backslash_L$ is also written as $MOD_L(\Sigma)$.

$\mathcal{M}^+ =_{def} \{M \in Mod_L \mid$ for all $\varphi \in //\mathcal{M}//\ M \models_L \varphi\}$. 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 $\mathcal{M}$. The set of *world descriptions* of **L**, $WD_L =_{def} \{\backslash\backslash\Sigma\backslash\backslash_L \mid \Sigma \subseteq \mathcal{L}_L\}$.

The subscript **L** will sometimes be omitted when confusion is unlikely to occur. The following proposition lists some properties of the introduced notions.

Proposition 2.4
(i) $//(\backslash\backslash\Sigma\backslash\backslash)// = Cn_{\mathbf{L}}(\Sigma)$
(ii) $\backslash\backslash(//\mathcal{M}//)\backslash\backslash = \mathcal{M}^+$
(iii) If $Cn_{\mathbf{L}}(\Sigma) = //\mathcal{M}//$, then $\backslash\backslash\Sigma\backslash\backslash = \mathcal{M}^+$
(iv) For all $\mathcal{M} \in WD_{\mathbf{L}}$, $\mathcal{M} = \mathcal{M}^+$.

We are now ready to define our models for preferential entailment, which we call *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.

Definition 2.5 An *epistemic preference model* for **L** is a triple $\langle WD_{\mathbf{L}}, >, \models \rangle$, where $>$ is an irreflexive relation on the set of world descriptions $WD_{\mathbf{L}}$, and $\models$ is $\models_{h,\mathbf{L}}$ restricted to $WD_{\mathbf{L}} \times \mathcal{L}_{\mathbf{L}}$.

Definition 2.6 Let $\langle WD_{\mathbf{L}}, >, \models \rangle$ be an epistemic preference model. $\mathcal{M} \in WD_{\mathbf{L}}$ *preferentially satisfies* φ, notation $\mathcal{M} \models_{\mathbf{L},>} \varphi$, iff $\mathcal{M} \models_{\mathbf{L}} \varphi$ and for all $\mathcal{N} < \mathcal{M}$, $\mathcal{N} \not\models \varphi$. φ is a *preferential consequence* of Γ, notation $\Gamma \models_{\mathbf{L},>} \varphi$, iff for all $\mathcal{M} \in WD_{\mathbf{L}}$, $\mathcal{M} \models_{\mathbf{L},>} \Gamma$ implies $\mathcal{M} \models \varphi$.

We write $Cn_{\mathbf{L},>}(\Gamma)$ for $\{\varphi \in \mathcal{L}_{\mathbf{L}} \mid \Gamma \models_{\mathbf{L},>} \varphi\}$ and we omit the subscript **L** if it is clear which standard logic is used. Notice that a epistemic preference model is determined by **L** and the preference relation $>$ on $WD_{\mathbf{L}}$.

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 $\langle WD_{\mathbf{L}}, >, \models \rangle$ is called *proper* iff set inclusion is contained in $\succcurlyeq \cup \preccurlyeq$, i.e., for all $\mathcal{M}, \mathcal{N} \in WD_{\mathbf{L}}$, $\mathcal{M} \subseteq \mathcal{N}$ 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 $\mathcal{M} \subsetneq \mathcal{N}$.

For comparison, we also provide here the models for preferential entailment that were proposed in [Shoham, 1987; 1988], [Makinson, 1989], and [Kraus *et al.*, 1990].

Definition 2.7 A *Shoham model* for **L** is a triple $\langle Mod_{\mathbf{L}}, >, \models_{\mathbf{L}} \rangle$, where $>$ is a strict partial order on $Mod_{\mathbf{L}}$.

Definition 2.8 A *Makinson model* 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 $\{\varphi \in \mathcal{L}_{\mathbf{L}} \mid \mathcal{M} \models_{\mathbf{L}} \varphi\}$ is closed under $Cn_{\mathbf{L}}$.

Definition 2.9 A *KLM model* is a quadruple $\langle S, \ell, >, \models \rangle$, where S is an arbitrary set of states, $\ell : S \to \mathcal{P}(Mod_{\mathbf{L}})$ 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 $\ell(s) \models_{h,\mathbf{L}} \varphi$.

Preferential entailment for the above models is defined completely analogous to 2.6. 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 the subject of [Dix and Makinson, 1992). For any Shoham model there exists an equivalent epistemic preference model.

From a technical point of view, KLM models or Makinson models are perhaps better suited than epistemic preference models to play the role of basic preference models. However, since often epistemic intuitions arc used to informally justify the preferential models, it is advisable to lake these intuitions serious and consider preferences between "real" epistemic states. In the following section we show that epistemic preference models arc more general than Shoham models.

3 Preference-based semantics for Default Logic

In the previous section we mentioned that any Shoham model has an equivalent epistemic preference model. The following proposition guarantees that whenever a logic has a preferential semantics in terms of Shoham models it also has one in terms of epistemic preference models. It follows that all nonmonotonic logics treated in [Shoham, 1987; Shoham, 1988] can also be captured in our approach. This includes Predicate Circumscription [McCarthy, 1980|, the "minimal knowledge" approach of [Halpern and Moses, 1985], and some variants of Reiter's Default Logic, but *not* Reiter's Default Logic itself. However, Default Logic *can* be captured by epistemic preference models.

For convenience, we repeat some basic definitions of [Reiter, 1980]. The underlying standard logic **L** is assumed to be ordinary first-order logic. A *default (rule)* is an expression of the form $\alpha : \beta_1,\ldots,\beta_n / \omega$ $(n \geq 1)$, where α (the *prerequisite*), $\beta_1,\ldots,\beta_n$ (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 $\mathcal{T} = \langle D, \Gamma \rangle$, where D is a set of defaults and Γ is a set of closed formulas of **L**. E is called an *extension* of $\langle D, \Gamma \rangle$ iff E is a fixed point of the function $f : \mathcal{P}\mathcal{L} \to \mathcal{P}\mathcal{L}$, where $f(\Sigma)$ is defined to be the smallest **L**-theory containing Γ and every ω such that $\alpha : \beta_1,\ldots,\beta_n / \omega \in D$, $\alpha \in f(\Sigma)$ and for all $i \in \{1,\ldots,n\}$, $\neg\beta_i \notin \Sigma$.

As a basis for our preferential semantics for Default Logic we use the semantics proposed in [Etherington, 1988], which already makes use of preference relations.

Definition 3.1 Let $\delta = \alpha : \beta_1,\ldots,\beta_n / \omega$ be a default. $<_\delta$, the *preference relation corresponding to* δ, is defined as follows: For any $\mathcal{M}, \mathcal{N} \subseteq Mod_{\mathbf{L}}$, $\mathcal{M} <_\delta \mathcal{N}$ iff (1) $\mathcal{N} \models_{h,\mathbf{L}} \alpha$, (2) $\neg\exists i\; \mathcal{N} \models_{h,\mathbf{L}} \neg\beta_i$, and (3) $\mathcal{M} = \mathcal{N} - \{M \mid M \not\models_{\mathbf{L}} \omega\} \neq \mathcal{N}$.

Intuitively, $\mathcal{M} <_\delta \mathcal{N}$ means that on account of δ the world description $\mathcal{M}$ is preferred to the world description $\mathcal{N}$. In case the set of defaults is finite, the preference relation $<_D$ corresponding to a set of defaults D is simply the transitive closure of the union of the preference relations corresponding to the elements of D. For default theories with infinitely many default rules, a slightly more complicated definition is called for.

Definition 3.2 Let D be a set of defaults and let $\mathfrak{M}$ be a set of models. $<_D$, the preference relation corresponding to D over $\mathcal{P}\mathfrak{M}$ is defined as follows: $\mathcal{M} <_D \mathcal{N}$ iff there exist $\delta_1, \delta_2, \ldots \in$ D and subsets $\mathcal{N}_0, \mathcal{N}_1, \mathcal{N}_2, \ldots$ of $\mathfrak{M}$ such that $\mathcal{N} = \mathcal{N}_0 >_{\delta_1} \mathcal{N}_1 >_{\delta_2} \mathcal{N}_2 >_{\delta_3} \ldots$, and $\mathcal{M} = \bigcap \mathcal{N}_i$.

If $\mathcal{D} = \langle D, \Gamma \rangle$ is a default theory, then $<_{\mathcal{D}}$ denotes the restriction of $<_D$ to the sets of worlds in which Γ is valid. In [Etherington, 1988] it is shown that extensions of a default theory $\langle D, \Gamma \rangle$ correspond to the formulas valid in the $<_{\mathcal{D}}$-minimal elements which are $\mathcal{D}$-stable, where $\mathcal{M}$ is called $\mathcal{D}$-*stable* iff $\exists D' \subseteq D$ such that $\mathcal{M} <_{D'} \backslash\backslash\Gamma\backslash\backslash_L$ and every justification β of a default $\delta \in D'$ is true in some $M \in \mathcal{M}$.

The preferential semantics for Default Logic will be given in two stages. Before the models corresponding to sets of defaults are given, we first define preference models for default theories. The preference relation associated with a default theory $\mathcal{D}$ is a variant of $<_{\mathcal{D}}$ which takes the role of stability into account.

Definition 3.3 Let $\mathcal{D} = \langle D, \Gamma \rangle$ be a default theory. The *epistemic preference model associated with* $\mathcal{D}$ is the model $\langle WD_L, >_{\mathcal{D}}, \models \rangle$, where $>_{\mathcal{D}} = \lambda(>_0, >_{\mathcal{D}})$ and $>_0$ is defined as follows: $\mathcal{N} >_0 \mathcal{M}$ iff $\mathcal{M} = \emptyset \neq \mathcal{N}$ and there is no $<_{\mathcal{D}}$-minimal and $\mathcal{D}$-stable $\mathcal{N}' <_{\mathcal{D}} \mathcal{N}$. To avoid stacked subscripts, we write $\models_{\mathcal{D}}$ and $Cn_{\mathcal{D}}$ for the preferential consequence relation and operation of $\langle WD_L, >_{\mathcal{D}}, \models \rangle$.

Proposition 3.4 If $\langle WD_L, >_{\mathcal{D}}, \models \rangle$ is the epistemic preference model associated with $\mathcal{D} = \langle D, \Gamma \rangle$, then for any $\Delta \subseteq \mathcal{L}_L$, $Cn_{\mathcal{D}}(\Delta) = \bigcap \{E \mid E$ is an extension of $\mathcal{D}$ and $\Delta \subseteq E\}$.

An immediate corollary of proposition 3.4 is the monotonicity of $Cn_{\mathcal{D}}$. 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 3.5 Let D be a set of defaults. The *epistemic preference model associated with* D is the model $\langle WD_L, >_D, \models \rangle$, where $>_D = \lambda(\rhd_D, \subset)$ and $\rhd_D$ is defined as follows: $\mathcal{N} \rhd_D \mathcal{M}$ iff $\mathcal{N} >_{\langle D, //\mathcal{N}// \rangle} \mathcal{M}$. To avoid stacked subscripts, we write $\models_D$ and Cn_D for the preferential consequence relation and operation of $\langle WD_L, >_D, \models \rangle$.

Proposition 3.6 If $\langle WD_L, >_D, \models \rangle$ is the epistemic preference model associated with a set D of defaults, then for all $\Gamma \subseteq \mathcal{L}_L$, $Cn_D(\Gamma) = \bigcap \{E \mid E$ is an extension of $\langle D, \Gamma \rangle\}$.

In contrast to $>_{\mathcal{D}}$ of 3.4, $>_D$ of 3.6 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, 1992] it is proved that transitive preference models can capture Default Logic, provided the notion of preferred model is strengthened to $\mathcal{M} \models_{L,>} \Gamma$ iff $\mathcal{M} \vdash \Gamma$, for all $\mathcal{N} < \mathcal{M}$, $\mathcal{N} \nvdash \Gamma$, and for all $\mathcal{N}'$ ($\mathcal{M} \subset \mathcal{N}' \subseteq \backslash\backslash\Gamma\backslash\backslash_L$ implies $\mathcal{M} < \mathcal{N}'$). A similar result has been obtained independently by Lin and Shoham [1992].

As far as we know, our preferential semantics for Default Logic is the first that uses the original definition of preferred model (at the cost of allowing intransitive preference relations).

4 Nonmonotonic consequence operations

We prefer the sceptical interpretation of nonmonotonic logics and consider the the intersection of extensions to represent the consequences of a default theory, even though this intersection is not necessarily itself an extension. However, we compromise to the more brave or credulous authors by also considering extension operations next to the nonmonotonic consequence operations.

Definition 4.1 An *extension operation* Ext for **L** is a function $\mathcal{P}(\mathcal{L}_L) \rightarrow \mathcal{P}(\mathcal{P}(\mathcal{L}_L))$ satisfying the following three conditions: (for all $\Gamma, \Delta \subseteq \mathcal{L}_L$)

(i) If $\Delta \in Ext(\Gamma)$, then $\Gamma \subseteq \Delta$. (*inclusion*)

(ii) If $\Gamma' \in Ext(\Gamma)$ and $\Gamma \subseteq \Delta \subseteq \Gamma'$, then $\bigcap Ext(\Delta) \subseteq \Gamma'$. (*cumulative transitivity*)

(iii) $Ext(\Gamma) = Ext(Cn_L(\Gamma))$. (**L**-*invariance*)

Definition 4.2 $Cn : \mathcal{P}(\mathcal{L}_L) \rightarrow \mathcal{P}(\mathcal{L}_L)$ is called a *nonmonotonic consequence operation* for **L** iff for some extension operation Ext for **L**, for all $\Gamma \subseteq \mathcal{L}_L$, $Cn(\Gamma) = \bigcap Ext(\Gamma)$.

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, $Cn_L(\Gamma) = \Gamma$. 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 $\Gamma \subseteq \mathcal{L}_L$ $Ext'(\Gamma) \neq \emptyset$. (Just replace all values $\emptyset$ with $\{\mathcal{L}_L\}$.) We will not distinguish between having no extensions and having only the inconsistent extension, and we simply assume from now on that only nonempty sets are in the range of extension operations.

Proposition 4.3 Cn is a nonmonotonic consequence operation for **L** iff Cn satisfies: (for all $\Gamma, \Delta \subseteq \mathcal{L}_L$)

(i) $\Gamma \subseteq Cn(\Gamma)$ (*inclusion*)

(ii) $\Gamma \subseteq \Delta \subseteq Cn(\Gamma)$, then $Cn(\Delta) \subseteq Cn(\Gamma)$ (*cumulative transitivity*)

(iii) $Cn(\Gamma) = Cn(Cn_L(\Gamma))$ (**L**-*invariance*)

Some properties that can be derived from those mentioned above are $Cn(Cn(\Gamma)) = Cn(\Gamma)$ (idempotency), $Cn_L(Cn(\Gamma)) = Cn(\Gamma)$ and $Cn_L(\Gamma) \subseteq Cn(\Gamma)$. Moreover, if Cn is **L**-invariant, and Cn_L is **L'**-invariant, then Cn is **L'**-invariant.

Completely analogous to the representation theorems for cumulative consequence relations obtained in [Kraus *et al.*, 1990] and [Makinson, 1989], we have the following result for nonmonotonic consequence operations:

Proposition 4.4 Cn is a nonmonotonic consequence operation for **L** iff $Cn = Cn_{>}$, for some epistemic preference model $\langle WD_L, >, \models \rangle$.

It follows from this result that the nonmonotonic modal logics introduced in [McDermott, 1982] have a preferential semantics, and it is shown in [Shvarts, 1990] that Moore's autoepistemic logic is a special case of the nonmonotonic modal framework. Proposition 4.4 is also valid for **L**-faithful Makinson models.

5 Cumulativity and rationality

Gabbay [1985] introduces the notion of weakly monotonic, or cumulative, consequence, which is characterized in [Kraus *et al.*, 1990] and [Makinson, 1989] in terms of preferential models. Below we show that the characterization in terms of Makinson models can be restricted to epistemic preference models.

Definition 5.1 An extension operation Ext for **L** is called *cumulative* iff it satisfies the following condition: If $\Delta \in Ext(\Gamma)$, $\Delta' \in Ext(\Gamma')$ and $\Gamma' \subseteq \Gamma \subseteq \Delta$, then $\Delta' \in Ext(\Gamma)$. A *cumulative consequence operation for L* is a nonmonotonic consequence operation induced by a cumulative extension operation for **L**.

Proposition 5.2 Cumulative consequence operations are characterized by inclusion, cumulative transitivity, **L**-invariance, and cumulative monotonicity ($\Gamma \subseteq \Delta \subseteq Cn(\Gamma) \Rightarrow Cn(\Gamma) \subseteq Cn(\Delta)$.)

Definition 5.3 A Makinson model $\langle \mathfrak{M}, >, \models \rangle$ is called *stoppered* iff for all $\Gamma \subseteq \mathcal{L}_L$ and for all $\mathcal{M} \in \mathfrak{M}$, $\mathcal{M} \models \Gamma$ implies $\mathcal{N} \models_{>} \Gamma$, for some $\mathcal{N} \leq \mathcal{M}$.

The following result is essentially already obtained in [Makinson, 1989], but without the reference to **L** and without requiring the models to be **L**-faithful.

Proposition 5.4 Cn is a cumulative consequence operation for **L** iff $Cn = Cn_{>}$, for some stoppered **L**-faithful Makinson model $\langle \mathfrak{M}, >, \models \rangle$.

In the above proposition, the **L**-faithful Makinson models can be restricted to epistemic preference models for **L**:

Proposition 5.5 Cn is a cumulative consequence operation for **L** iff $Cn = Cn_{>}$, for some stoppered epistemic preference model $\langle WD_L, >, \models \rangle$.

Cumulative monotonicity is considered to be a rationality requirement, and in the literature there is a tendency to disqualify non-cumulative nonmonotonic logics, such as Default Logic. For example, in [Brewka, 1990], Default Logic 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 $\Sigma \subseteq \mathcal{L}_L$, $Cn_L(\Sigma) \neq \mathcal{L}_L$. (In fact, this requirement is inconsistent with the inclusion property, satisfied by standard logics.)

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 [9]. 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 obvious rationality requirement on default beliefs and their induced nonmonotonic consequence operations, namely that $Cn(\Gamma)$ is inconsistent only if Γ is. In other words, if $Cn_L(\Gamma) \neq \mathcal{L}_L$, then $Cn(\Gamma) \neq \mathcal{L}_L$. Let us call such consequence operation *consistent*. (The corresponding condition for Ext is: if $Cn_L(\Gamma) \neq \mathcal{L}_L$, then $Ext(\Gamma) \neq \{\mathcal{L}_L\}$.)

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 (D,Γ) 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: $Tr_{Bes}(\alpha : \beta / \beta) = : \alpha \supset \beta / \alpha \supset \beta$. Sets of thus obtained defaults induce a cumulative consequence operation. However, since Tr_{Bes} lacks a proper justification and gives rise to some counterintuitive results, we propose an alternative, more general translation Tr_{free} of arbitrary default into default without prerequisites. Tr_{free} is defined by means of $Tr_{free}(\alpha : \beta_1,\ldots,\beta_n / \omega) =_{def} : \alpha\wedge\beta_1,\ldots,\alpha\wedge\beta_n / \alpha \supset \omega$.

Unfortunately, Tr_{free} does not necessarily revise a set of (normal) defaults into a set that 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: 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 docs not prefer f to g. [Savage, 1972, p. 21]

The sure-thing principle is closely related to the ability to reason by cases, and if one does not allow (defaults with) prerequisites, then it is indeed possible to reason by cases in Default Logic.

An obvious objection againsi the 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 Default Logic as the logic of some agent or system for which computational issues might matter, and not as the logic of an ideally rational agent.

6 Conclusion

Default logic can be given a preferential semantics, provided Shoham's original requirement that the semantics has to be based on a strict partial order on the set of standard models is dropped. Default logic can be captured by episiemic preference models (for a some standard logic L), in which a (not necessarily transitive) preference relation on the set of (partial) world descriptions is defined. These epistemic preference models for L characterize the class of nonmonotonic consequence operations for L, which satisfy inclusion, cumulative transitivity, and GiL-invariancc.

Cumulative monotonicity can be considered to be a rationality requirement on consequence operations, but this docs not imply that noncumulativc logics arc necessarily inadequate formalizations of the (nonmonotonic) reasoning of an ideally rational agent. Instead, cumulative monotonicity can be seen as a constraint on the set of defaults a rational agent might believe, and in the context of Default Logic this constraint is for example satisfied sets of normal defaults without prerequisites.

References

[Besnard, 1989] P. Besnard, An Introduction to Default Logic Springer, Berlin, 1989.

[Brewka, 1990] G. Brewka, Cumulative Default Logic: in defense of nonmonotonic inference rules. Artificial Intelligence 50:183-205, 1990.

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

[Etherington, 1988] D. Etherington, Reasoning with Incomplete Information. Pitman, London, 1988.

[Gabbay, 1985] D. Gabbay, Theoretical foundations for nonmonotonic reasoning in expert systems. In Logics and Models of Concurrent Systems, pages 439-457, K.R. Apt (ed.), Springer, Berlin, 1985.

[Gardenfors, 1988] P. Giirdenfors, Knowledge in Flux: Modeling the Dynamics of Episiemic Stales. MIT Press, Cambridge Mass., 1988.

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

[Kraus et al, 1990] S. Kraus, D. Lehmann and M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logic. Artificial Intelligence 44:167-207, 1990.

[Lcvcsquc, 1990] H.J. Levesque, All I know: A study in autoepistcmic logic. Artificial Intelligence 42:263-309, 1990.

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

[Makinson, 1989] D. Makinson, General Theory of Cumulative Inference. In Non-Monotcnic Reasoning, pages 1-I8, M. Reinfrank et al. (eds.). Springer, Berlin, 1989.

[McCarthy, 19801 J. McCarthy, Circumscription - a form of nonmonotonic reasoning. Artificial Intelligence 13:27-39, 1980.

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

[Reiter, 198()| R. Re iter, A logic for default reasoning. Artificial Intelligence 13:81-132, 1980.

[Savage, 1972] L.J. Savage, The Foundations of Statistics Dover, New York, 1972 (2nd ed.; 1st cd. John Wiley, 1954)

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

[Shoham, 19881 Y. Shoham, Reasoning about Change. MIT Press, Cambridge MA, 1988.

[ShvarLs, 1990] G. Shvarts, Autoepistemic modal logic. In Theoretical Aspects of Reasoning About Knowledge: Proceedings of the Fourth Conference, , pages 97-109, R. Parikh (ed.), San Mateo CA, Morgan Kaufmann.

[Voorbraak, 1991) F. Voorbraak, A preferential model semantics for default logic. In Symbolic and Quantitative Approaches to Uncertainty, pages 344-351, R. Kruse and P. Siege! (eds.), Springer, Berlin, 199L

[Voorbraak, 1992] F. Voorbraak,Generalized Kripkc models for cpistcmic logic. In Theoretical Aspects of Reasoning About Knowledge: Proceedings of the Fourth Conference, pages 214-228, Y. Moses (ed.), Morgan Kaufmann, San Mateo CA ,1992

[Voorbraak, 1993] F. Voorbraak, Preference-based semantics for nonmonotonic logics. Logic Group Preprint Series, Dept. of Philosophy, Utrecht University.