Belief functions and default reasoning

S. Benferhat\textsuperscript{1}  
benferha@irit.fr  

A. Saffiotti\textsuperscript{2}  
asaffiio@ulb.ac.be  

P. Smets\textsuperscript{2}  
psmets@ulb.ac.be

\textsuperscript{1}IRIT (UPS)  
118 Route de Narbonne  
Toulouse 31062 Cedex, France  

\textsuperscript{2}IRIDIA, Université Libre de Bruxelles  
50 av. F. Roosevelt  
B-1050 Bruxelles, Belgium

Abstract. We present a new approach to dealing with default information based on the theory of belief functions. Our semantic structures, inspired by Adams' \(\varepsilon\)-semantics, are epsilon-belief assignments, where values committed to focal elements are either close to 0 or close to 1. We define two systems based on these structures, and relate them to other non-monotonic systems presented in the literature. We show that our second system correctly addresses the well-known problems of specificity, irrelevance, blocking of inheritance, ambiguity, and redundancy.

1. Introduction

Default reasoning is the process of drawing conclusions from i) a set of general rules which may have exceptions, and ii) a set of facts representing the available information (which is often incomplete). The conclusions so drawn are only plausible and can be revised in the light of the new information. The desirable properties for a consequence relation that capture default reasoning have been discussed at length in the AI literature. They can be summarized as follows.

- \textbf{Rationality:} the consequence operator used to generate plausible conclusions from a knowledge base should satisfy the rationality postulates proposed by Kraus, Lehmann and Magidor (1990).
- \textbf{Specificity:} results obtained from sub-classes should override those obtained from super-classes (Touretzky, 1984). For example, from the knowledge base \(\Delta = \{\text{"Birds fly", "Penguins do not fly", "Penguins are birds"}\}\), one should deduce that birds which are penguins do not fly, since penguins are a subclass of birds.
- \textbf{Property inheritance:} objects should inherit properties from super-classes unless there is contradiction on that property. For example, from the previous \(\Delta\), one should deduce that birds that are red fly, since being red is irrelevant to flying. Also, if we add the rule "Birds have legs", then one should deduce that penguins have legs too, since having legs is not a conflicting property. Failure to perform these deductions is referred to as the problem of "irrelevance" and of "inheritance blocking", respectively.
- \textbf{Ambiguity preservation:} in a situation where we have an argument in favor of a proposition, and an independent argument in favor of its negation, we should not conclude anything about that proposition. The most popular example is the so-called Nixon-diamond: knowing that "Quakers are pacifists", "Republicans are not pacifists", and Nixon is both a Quaker and a republican, one should not deduce that Nixon is a pacifist, nor that he is not.\footnote{Note that this is different from the situation of inconsistency, where we have an argument which supports both a conclusion and its contrary, as in "\(\alpha \rightarrow \beta\)" and "\(\alpha \rightarrow \neg \beta\)".}

In the last decade there have been several proposals for reasoning with default information. Some of them are based on the use of uncertainty models such as probability theory (Adams, 1975; Pearl, 1988), or possibility theory (Dubois and Prade, 1988; Benferhat et al., 1992). Up to now, however, no single system has been reported that correctly addresses all of the desiderata above. In this paper, we show how we can use belief functions, originally developed for modeling uncertainty (Shafer, 1976; Smets, 1988), to build a non-monotonic system that gives a satisfactory answer to all of the above issues.

There have already been a few works on representing default information with belief functions (e.g., Hsia, 1991; Smets and Hsia, 1991). These works require the assessment of numeric values, whose origin is often an open question. Finding a solution free from such assessments would somehow avoid the problem of the origin of the numbers. In this paper, we give another interpretation of default information by using a class of "extreme" belief functions, called epsilon-belief functions, whose non-null masses are either close to 0 or close to 1. The idea of using extreme values is not new to plausible reasoning: for instance, Adams (1975) and Pearl (1990) use extreme probabilities to encode default information; and De Kleer (1990) and Poole (1993) use extreme probabilities for diagnosis.

The rest of this paper is organized as follows. In the next section, we give a short reminder on Adams' \(\varepsilon\)-semantics and Pearl' System Z, and recall a few notions of the theory of belief functions (see Shafer, 1976; Smets, 1988) for a complete exposition. In section 3, we introduce \(\varepsilon\)-belief functions, and show how to use them to define a non-monotonic consequence relation. This relation turns out to be too cautious, and we define two more relations in sections 4 and 5. The first one is based on the least-commitment principle, and is equivalent to system Z. The second one uses Dempster's rule of combination, and is incomparable with the current systems. In section 6, we study this relation in more detail and show that it correctly addresses all of the issues above.
2. Background

We are interested in default rules of the form "generally, if we have $\alpha$ then we have $\beta$", where $\alpha$ and $\beta$ are formulae of some underlying language $L$. For the goals of this paper, we assume that $L$ is a classical propositional language. An interpretation for $L$ is an assignment of a truth value in $\{T, F\}$ to each formula of $L$, in accordance with the classical rules of propositional calculus; we denote by $\Omega$ the set of all such interpretations (also called worlds). We say that an interpretation $\omega$ is a model of a formula $\alpha$, and write $\omega = \alpha$ iff $\omega(\alpha) = T$, and denote by $[\alpha]$ the set of all the models of $\alpha$.

We write a default rule "generally, if $\alpha$ then $\beta$" as $\alpha \rightarrow \beta$, where $\alpha$ and $\beta$ are formulae of $L$. Note that "$\rightarrow$" is a non-classical arrow, and it should not be confused with material implication. Given a default rule $d = \alpha \rightarrow \beta$, we denote by $\phi_d$ the formula of $L$ obtained by replacing $\rightarrow$ by material implication, namely, $\phi_d = \neg \alpha \lor \beta$. A default base is a multiset $\Delta = \{\alpha_1 \rightarrow \beta_1, \ldots, \alpha_m \rightarrow \beta_m\}$ of default rules. We emphasize that a base is a multiset rather than a set, i.e., $\Delta = (\alpha \rightarrow \beta)$ is different from $\Delta = (\alpha - \beta, \alpha - \beta)$. We use default bases to represent background knowledge about what normally is the case. Given a default base $\Delta$, we are interested in defining a consequence relation $\vdash$ between formulae of $L$ that tells us which consequences we can reasonably drawn from the known facts. We would like $\vdash$ to fulfill the desiderata listed above. For example, given the base $\Delta = \{b \rightarrow f; p \rightarrow b, p \rightarrow f\}$ (where "$b$" stands for "bird", "f" for "flies", and "$p$" for "penguin"), we would like to have $b \vdash f$ and $b \lor p \vdash -f$, but not $b \lor p \vdash f$.

2.1. Probabilistic semantics for default rules

Adams (1975), and later Pearl (1988), have suggested a probabilistic interpretation where a default rule $\alpha \rightarrow \beta$ is read as a constraint $P(\beta|\alpha) > 1-\varepsilon$, with $P$ a probability distribution and $\varepsilon$ an infinitesimal positive number. Given a set of defaults $\Delta$, they construct a class of probability distributions $\mathcal{P}$ such that, for each distribution $P$ in $\mathcal{P}$ and each default $\alpha \rightarrow \beta$ in $\Delta$, $P(\beta|\alpha) > 1-\varepsilon$. A formula $\beta$ is said to be an $\varepsilon$-consequence of $\alpha$ with respect to $\Delta$, denoted by $\alpha \varepsilon \beta$, if for each $P$ in $\mathcal{P}$ there exists a real function $\omega$ such that $\lim_{\varepsilon \to 0} \omega(\varepsilon) = 0$ and $P(\beta|\alpha) > 1-\varepsilon$. Said differently, $\beta$ is a consequence of $\alpha$ with respect to $\Delta$ if the conditional probability $P(\beta|\alpha)$ is very high provided that the conditional probability of each default in $\Delta$ is very high.

Adams' $\varepsilon$-consequence is not entirely satisfactory. For example, it suffers from the problem of irrelevance mentioned above: from the default "Generally, birds fly", $\varepsilon$-consequence does not deduce that red birds fly also. To overcome this limitation, Pearl (1990) has proposed a default reasoning system, called $Z$, based on a ranking of default rules that respects the notion of specificity. Given a default rule $d = \alpha \rightarrow \beta$, Pearl gives a method to rank-order the rules in $\Delta$ such that the least specific rules (i.e. with most general antecedents) get the least priority. To do this, he defines the notion of tolerance: a rule $\alpha \rightarrow \beta$ is said to be tolerated by a base $\{\alpha_i \rightarrow \beta_i, i = 1, \ldots, m\}$ iff $(\alpha \rightarrow \beta, \alpha \rightarrow \beta_i, \ldots, \alpha \rightarrow \beta_{m+1})$ is consistent. Then, he partitions $\Delta$ into an ordered set $\{\Delta_1, \Delta_2, \ldots, \Delta_k\}$ such that rules in $\Delta_i$ are tolerated by all rules in $\Delta_1 \cup \ldots \cup \Delta_{i-1}$.

Roughly speaking, $\kappa(\omega)$ corresponds to the index of the lowest sub-base that contains a rule violated by $\omega$, and $z(\alpha) = \kappa(\omega)$ is the minimum rank of a model of $\alpha - \omega$, $z(\omega)$ can be read as a degree of "abnormality" of $\alpha$ with respect to the rules in $\Delta$. Finally, Pearl defines a non-monotonic inference relation, denoted here by $\rightarrow_Z$, as follows:

$$\alpha \rightarrow_Z \beta \iff z(\alpha \land \beta) < z(\alpha \land \neg \beta).$$

An equivalent treatment of default information has been done in the framework of possibility theory (Benferhat et al., 1992).

2.2. A reminder on belief functions

Let $\Omega$ be a finite set of worlds, one of them being the actual world. A basic assignment on $\Omega$ is a function $m: 2^\Omega \rightarrow [0, 1]$ that satisfies:

$$m(\emptyset) = 0, \quad \sum_B m(B) = 1, \quad B \subseteq \Omega$$

The term $m(\alpha)$, called the basic belief mass given to $\alpha$, represents the part of a total and finite amount of belief that supports the fact that the actual world belongs to $\alpha$ and does not support the fact that the actual world belongs to a strict subset of $\alpha$. Any subset $A$ of $\Omega$ for which $m(A) > 0$ is called a focal element.

An agent's belief can equivalently be represented by the function $\text{bel}: 2^\Omega \rightarrow [0, 1]$, called a belief function, defined by

$$\text{bel}(A) = \sum_{B \subseteq A} m(B).$$

The relation between $m$ and $\text{bel}$ is one-to-one. The term $\text{bel}(A)$ represents the degree of belief, of necessary support, that the actual world belongs to $A$. Related to $\text{bel}$ is another function $\text{pl}: 2^\Omega \rightarrow [0, 1]$, called a plausibility function, given by

$$\text{pl}(A) = \sum_{B \cap A \neq \emptyset} m(B).$$

The term $\text{pl}(A)$ quantifies the degree of plausibility, of potential support, that the agent could give to the fact that the actual world belongs to $A$. When the focal elements of a basic belief assignment are singletons, then $\text{bel} = \text{pl}$ is a probability measure. When the focal elements $A_1, \ldots, A_n$ are nested (that is, $A_1 \subseteq \ldots \subseteq A_n$), $\text{bel}$ is called a consonant belief function, and $\text{bel}$ is a necessity measure and $\text{pl}$ is a possibility measure (Zadeh, 1978; Dubois and Prade, 1988). When $m$ has at most one focal element $A \neq \emptyset$, it is called a simple support function.

When a new piece of evidence telling that the actual world belongs to $A$ becomes available to the agent, his/her belief is revised by the application of the so-called Dempster's rule of conditioning. The basic belief mass $m(X)$ that was supporting the subset $X$ of $\Omega$, now supports $X \cap A$. This transfer of belief is described by the following relation, where $\text{bel}(IA)$ denotes the conditional belief function.
The impact of a piece of evidence \( E \) that bears on \( \Omega \) is represented by a belief function \( \text{bel} \) that describes the agent's beliefs on \( \Omega \) given \( E \) and nothing else. Suppose the agent receives two distinct pieces of evidence \( E_1 \) and \( E_2 \), and let \( \text{bel}_1 \) and \( \text{bel}_2 \) be the belief functions induced by each evidence individually. The combined effect of \( E_1 \) and \( E_2 \) is represented by Dempster's rule of combination. The corresponding basic belief assignment, denoted by \( m_1 \circ m_2 \), is given by

\[
m_1 \circ m_2 (A) = \frac{\sum_{B \cap C = \emptyset} m_1 (B) \cdot m_2 (C)}{1 - \sum_{B \cap C = \emptyset} m_1 (B) \cdot m_2 (C)}.
\]

3. Epsilon-belief functions

In this section, we extend the definition of \( \varepsilon \)-semantics in a belief function framework. First, we introduce the notion of epsilon-belief functions, whose values are either close to 0 or close to 1. Next, we interpret a default rule \( \alpha \rightarrow \beta \) as meaning that the conditional belief \( \text{bel}([\beta]([\alpha])) \) is close to 1. Finally, we define a consequence relation in a natural way: \( \alpha \rightarrow \beta \) follows from a base of defaults \( \Delta \) if, for each epsilon-belief function which satisfies all the default rules in \( \Delta \), we have \( \text{bel}([\beta]([\alpha])) \) close to 1. It turns out that this definition gives us the same results as Adams' \( \varepsilon \)-consequence relation.

**Definition 1.** An epsilon-mass assignment on \( \Omega \) is a function \( m_\varepsilon : 2^\Omega \rightarrow [0,1] \) such that, for each \( A \subseteq \Omega \), either \( m_\varepsilon (A) = 0 \) or \( m_\varepsilon (A) = 1 - \varepsilon_A \), where \( \varepsilon_A \) is an infinitesimal. The vector \( \varepsilon = (\varepsilon_1, \ldots, \varepsilon_k) \) of the infinitesimals appearing in \( m_\varepsilon \) is called the parameter of \( m \). The belief function induced from \( m_\varepsilon \) is called an epsilon-belief function (epsilon-bf).

Throughout this paper, we denote by \( \text{bel}_\varepsilon \) and \( \text{pl}_\varepsilon \) the belief and plausibility functions corresponding to a given epsilon-mass assignment \( m_\varepsilon \).

**Definition 2.** Let \( \text{bel}_\varepsilon \) be an \( \varepsilon \)-bf with parameter \( \varepsilon \). We say that \( \text{bel}_\varepsilon \) is an \( \varepsilon \)-bf-model of a default rule \( \alpha \rightarrow \beta \), and write \( \text{bel}_\varepsilon \models \alpha \rightarrow \beta \), if \( \lim_{\varepsilon \to 0} \text{bel}_\varepsilon ([\beta]([\alpha])) = 1 \), where the limit is taken with respect to all the elements of \( \varepsilon \) going to 0.

When working with values that depend on \( \varepsilon \), we will often use the notion of one value being infinitely larger than another, written \( a \gg b \). We say that \( a \gg b \) if \( \lim_{\varepsilon \to 0} \text{bel}_\varepsilon ([\beta]([\alpha])) = 0 \). We also say that \( a \vDash \beta \) and \( a \vDash \beta \) are of the same order, written \( a = b \), if \( \lim_{\varepsilon \to 0} \text{bel}_\varepsilon ([\beta]([\alpha])) = c \), with \( c \neq 0 \) and finite. The following properties will be useful for working with epsilon-bfs.

**Lemma 1.** Let \( \text{bel}_\varepsilon \) be an \( \varepsilon \)-bf. For any \( \alpha, \beta \in \mathcal{L} \),

\[
\text{a) } \lim_{\varepsilon \to 0} \text{bel}_\varepsilon ([\beta]([\alpha])) = 1 \iff \text{pl}_\varepsilon ([\alpha \land \beta]) = \max \{ \text{pl}_\varepsilon ([\alpha]), \text{pl}_\varepsilon ([\beta]) \}.
\]

**4. Entailment based on least-commitment**

One way to select some of the \( \varepsilon \)-bf-models of \( \Delta \) is by using the notion of being minimally informative: intuitively, we want to look at the consequences of "only knowing" \( \Delta \) (and nothing more). A similar approach has been taken, for the case of possibility measures, by Benferhat et al. (1992). We recall the following

**Definition 4.** Let \( \text{bel}_1 \) and \( \text{bel}_2 \) two (epsilon-) belief functions over \( \Omega \). Then, \( \text{bel}_1 \) is less committed that \( \text{bel}_2 \) iff, for any \( A \subseteq \Omega \), \( \text{pl}_1(A) \geq \text{pl}_2(A) \).
The least-commitment principle (Smets, 1988) states that, in order to model an item of information by a belief function, we should use the least committed belief function that is compatible with the information. Note that the least committed belief function representing a formula \( \alpha \) is given by the simple support function that gives mass 1 to \([\alpha]\) and 0 anywhere else. We show how to build an e-bf-model of \( \Delta \) based on this principle. We start by allocating a quasi-unitary mass to the set \( \{a\} \) of the worlds where (the propositional equivalent of) all the defaults in \( \Delta \) are satisfied, and the remaining mass \( \varepsilon \) to \( \Omega \). If there were no conflicts in the defaults, this allocation would be an e-bf-model of \( \Delta \). When there are conflicts, however, this e-bf will not satisfy some of the defaults in \( \Delta \)—namely, those that inherit a conflicting property from a more general class. Then, we put aside the defaults that are already satisfied, and put almost all of the free mass \( \varepsilon \) on the set \( \{\alpha\} \) corresponding to the still unsatisfied defaults, leaving a smaller \( \varepsilon \)' on \( \Omega \). This new e-bf is an e-bf-model of \( \Delta \) (and of \( \Delta \)) if we have no conflicts in \( \Delta \). Otherwise, we iterate the procedure until the e-bf will satisfy all the defaults in \( \Delta \).

More precisely, let \( \varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \) be a vector of infinitesimals such that

\[
(\varepsilon_{LC}) \quad \varepsilon_i \gg \varepsilon_{i+1} \text{ for any } i = 1, \ldots, n-1,
\]

where \( n \) is the cardinality of \( \Delta \). We build an e-bf in the following way.

**Step 0.** Let \( i = 0, \delta_0 = \emptyset, \text{sat}_0 = \emptyset, m_0 \) s.t. \( m_0(\emptyset) = 1 \) and \( m_0(A) = 0 \) otherwise.

**Step 1.** Repeat until \( \delta_i = \emptyset \)

1a. Let \( i = i+1 \)

1b. Let \( \delta_i = \delta_{i-1} - \text{sat}_{i-1} \)

1c. Let \( bel_i \) be the e-bf given by:

\[
m_i(\emptyset) = \varepsilon_i; \quad m_i([\alpha]) = m_{i-1}(\emptyset) - \varepsilon_i;
\]

\[
m_i(A) = m_{i-1}(A) \text{ otherwise.}
\]

1d. Let \( \text{sat}_i = \{ \delta_i | bel_i \models \delta \} \)

1e. If \( \text{sat}_i = \emptyset \) and \( \delta_i \neq \emptyset \), then Fail.

**Step 2.** Return \( bel_{i-1} \).

Note that all the focal elements are nested—the inner one being \( \{\alpha\} \)—and then the final e-bf returned by Step 2 is a consonant belief function. The procedure fails to find an e-bf if \( \Delta \) is inconsistent; in this case, we have \( \delta_0 = \emptyset \) and \( \delta_i \neq \emptyset \).

**Example 1.** Let \( \Delta = \{ b \rightarrow f, p \rightarrow b, p \rightarrow \neg f \} \) (where "b" stands for "bird", "f" for "flies", and "p" for "peanut"), and let us apply the previous algorithm. We have \( \delta_1 = \emptyset, \) and \( bel_1 \) given by:

\[
m_1([\emptyset]) = 1 - \varepsilon_1; \quad m_1([p]) = \varepsilon_1; \quad m_1([f]) = 0 \text{ otherwise.}
\]

We compute the set \( \text{sat}_1 \) of defaults which are satisfied by \( bel_1 \). We have:

\[
pl_1([b\land p\land f]) = 1 - \varepsilon_1 + \varepsilon_1 = 1,
\]

\[
pl_1([b\land p]) = \varepsilon_1,
\]

\[
pl_1([p\land f]) = \varepsilon_1,
\]

\[
pl_1([p\land \neg f]) = \varepsilon_1.
\]

Therefore, \( \text{sat}_1 = \{b\land f\} \). We iterate, and get \( \delta_2 \) by removing \( b\land f \) from \( \delta_1 \). Then:

\[
m_2([\emptyset]) = \varepsilon_2; \quad m_2([\alpha]) = \varepsilon_1 - \varepsilon_2;
\]

\[
m_2([\delta_1]) = 1 - \varepsilon_1; \quad m_2([\emptyset]) = 0 \text{ otherwise}
\]

with \( \varepsilon_1 \) infinitely larger than \( \varepsilon_2 \). We have:

\[
pl_2([p\land \emptyset]) = \varepsilon_1,
\]

\[
pl_2([p\land b]) = \varepsilon_1,
\]

\[
pl_2([p\land \neg f]) = \varepsilon_1,
\]

\[
pl_2([p\land f]) = \varepsilon_2.
\]

All the defaults in \( \delta_2 \) are satisfied and the algorithm ends by returning \( bel_2 \).

We denote by \( Bel_{LC}(\Delta) \) the family of e-bf's built by the procedure above — the elements of this family differ in the choice of \( \varepsilon \), provided that \( (\varepsilon_{LC}) \) is satisfied. This family "behaves well" for our goals: it is a subset of \( Bel_{L}(\Delta) \), and it induces a unique ordering \( \preceq_{LC} \) on the worlds in \( \Omega \). The latter property means that we can decide entailment by just looking at one element of \( Bel_{LC}(\Delta) \).

**Lemma 3.** Let \( \Delta \) be a default base. Then:

a) Any element of \( Bel_{LC}(\Delta) \) is an e-bf-model of \( \Delta \).

b) Let \( bel_1 \) and \( bel_2 \) be two elements of \( Bel_{LC}(\Delta) \), and \( \preceq_1 \) and \( \preceq_2 \) the corresponding orderings induced on \( \Omega \). Then, \( \preceq_1 \models \preceq_2 \).

It is interesting to compare \( Bel_{LC}(\Delta) \) with the result of the stratification proposed by Pearl (1988). We can show that the focal elements of the partitions of \( Bel_{LC}(\Delta) \) are directly related to the elements of the partition of \( \Delta \) obtained by \( Z \).

**Lemma 4.** Let \( \Delta = \Delta_1 \cup \ldots \cup \Delta_n \) be the stratification given by system \( Z \), and let \( bel_1 \) the e-bf built by our algorithm at step 1. Then, for any default \( \alpha \models \beta \) in \( \Delta \),

\[
a) \alpha \models \beta \text{ is tolerated by } \Delta \text{ iff } pl_1([\alpha]) = 1
\]

\[
b) \alpha \models \beta \in \Delta_i \text{ iff } bel_1 \models \alpha \models \beta.
\]

So, our algorithm produces the same ranking over \( \Delta \) than \( Z \). However, our approach does not require an a priori definition of the notion of "tolerance", but relies on the notion of being "less committed". Which of these two notions provides a more natural starting point is a matter of opinion.

We now use the set \( Bel_{LC}(\Delta) \) to give our second definition of entailment. It is similar to (BF), but we restrict the attention to the ebf-models that are in \( Bel_{LC}(\Delta) \).

\[
(\varepsilon_{LC}) \alpha \vDash_{LC} \beta \text{ iff for any } bel \in Bel_{LC}(\Delta), \text{ } \beta \models \alpha.
\]

**Example 2.** We can use the e-bf \( bel_2 \) built in Example 1 to check that we have \( b \land p \vDash_{LC} f \). In fact, we have:

\[
pl_2([b\land p\land f]) = \varepsilon_1,
\]

\[
pl_2([b\land p\land \neg f]) = \varepsilon_2,
\]

\[
\varepsilon_1 > \varepsilon_2.
\]

Hence, by lemma 1, \( \lim_{\varepsilon} bel_2([\neg f][b\land p]) = 1 \), and then \( b \land p \vDash_{LC} f \).

As \( Bel_{LC}(\Delta) \) is a subset of \( Bel(\Delta) \), \( \vDash_{LC} \)-consequences include \( \vDash \)-consequences; in particular, they include the consequences of system \( P \). As it turns out, the \( LC \) consequence relation is strictly larger than \( \vDash_p \), and precisely coincides with Pearl's system \( Z \), as it appears from lemma 4 above.

**Theorem 2.** For a given \( \Delta \), \( \alpha \vDash_{LC} \beta \) iff \( \alpha \vDash_{Z} \beta \).

### 5. Entailment based on Dempster's rule

The second way that we propose to strengthen the entailment relation (BF) is by considering only the ebf-models of \( \Delta \) that can be built by using Dempster's rule of combination. The intuitive argument goes as follows. Suppose we regard each default in \( \Delta \) as being one item of evidence provided by one of several distinct sources of information. Then, it makes sense to represent each default individually by one belief function.

\[\text{The construction given here should extend to the case where each source } S_i \text{ of information provides an entire base } \Delta_i \text{ of defaults: we could use the least-commitment principle to build a representative belief function for each } \Delta_i \text{ as above, and then combine these representatives by Dempster's rule.} \]
Belief functions and default reasoning

and combine these belief functions by Dempster’s rule to obtain a representation of the aggregate effect of all the defaults. We can then define entailment by looking at the conditionals that are satisfied by the combined belief function.

There are two technical choices to be made. The first one is which belief function to use to represent each default. Sticking to the arguments used in the last section, we propose to use the least committed belief function that satisfies that default: for a default $d = \alpha \rightarrow \beta$, this is the simple support function that allocates a mass of 1 in $d$ to $[\emptyset, d]$, and the remaining mass to $\emptyset$. We note by ssfd this function. Given a default base $\Delta = \{d_1, ..., d_n\}$, then, we consider the e-bf obtained by combining all the ssfd’s by Dempster’s rule:

$$bel_{\Delta} = \Theta\{ssfd \; d \in \Delta\}.$$  

The e-bf’s obtained in this way may be complex: the focal elements are not nested, and the mass values may include products of several $\epsilon_d$ (or $1-\epsilon_d$) for different $d$. Luckily, the plausibility of each world can be approximated in a simple way.

**Lemma 5.** Let $bel_{\Delta}$ be an e-bf built from $\Delta$ as above. Then, for any world $\omega$ in $\Omega$,

$$pl_{\Delta}(\omega) = \prod_{d: \omega \notin \emptyset} \epsilon_d \; \prod_{d: \omega \notin \emptyset} (1-\epsilon_d).$$  

If $\omega$ satisfies all formulas of $\Delta$, then $pl_{\Delta}(\omega) = 1$. The second question is how to choose the infinitesimals $\epsilon_d$’s. We impose constraints on the $\epsilon_d$’s based on the two following principles:

- **Auto-deduction principle.** We require $bel_{\Delta} = \alpha \rightarrow \beta$ for each default $\alpha \rightarrow \beta$ in $\Delta$ (i.e., $bel_{\Delta}$ is an ebf-model of $\Delta$).

- **Least commitment principle.** We want each ssfd to be as un-committed as possible (hence, each $\epsilon_d$ should be as large as possible).

For each $d = \alpha \rightarrow \beta$ in $\Delta$, the first principle generates the constraint (cf. lemma 1)

$$\max \left\{ pl_{\Delta}(\omega) \; \omega = \alpha \land \beta \right\} \geq \max \left\{ pl_{\Delta}(\omega) \; \omega = \alpha \land \beta \right\}$$

which is equivalent to (cf. Lemma 5)

$$(C_d) \max \left\{ \prod_{\omega = \alpha \beta} \epsilon_d \right\} \geq \max \left\{ \prod_{\omega = \alpha \beta} \epsilon_d \right\}$$

By solving the system given by all the $C_d$’s, we can get constraints on the form $\epsilon_i >_{\infty} \epsilon_j$ between some of the elements of $\epsilon$ (and between products of elements). The second principle can be used to avoid equilibria in conjunction constraints of $\epsilon$ by the following argument. Suppose that, for some $\epsilon_i$ and $\epsilon_j$, neither $\epsilon_i >_{\infty} \epsilon_j$ nor $\epsilon_j >_{\infty} \epsilon_i$ is the case; as both $\epsilon_i$ and $\epsilon_j$ should be made as large as possible, no one can be larger than the other, and then $\epsilon_i = \epsilon_j$.

We now describe an algorithm to solve a set $C$ of constraints of the form $(C_d)$. We call term any product of elements of $\epsilon$, and complex term a term containing at least two elements of $\epsilon$. The algorithm returns a set $\xi = \{\xi_0, ..., \xi_l\}$ of equivalence classes of terms such that: 1) all the terms in a class are of the same order; 2) any term in $\xi_j$ is infinitely larger than any term in $\xi_i$ when $i < j$; and 3) any element of $\epsilon$ is in some class.

**Step 0.** Let $l = 0$, $A = \{\xi_0\}$ (1) is a term and $1 >_{\infty} \xi_0 \in C$).  

0.a. Let $\xi_0$ be a set of $\epsilon_1$ such that there exists a complex term $t$ in $A_1$ which contains $\epsilon_1$.

0.b. If $\xi_0 = \emptyset$ then $\xi_0 = A_1$; else $\xi_1 = A_1 \cup \xi_0$ and $i = 1$.

**Step 1.** Repeat until $C = \emptyset$.

1.a. Let $i = i + 1$. Remove from $C$ any satisfied constraint.

1.b. Let $A_i$ be a set of terms in $C$ which does not appear in the right side of any constraint of $C$.

1.c. Let $\xi_i$ be a set of $\epsilon_j$ which does not appear in any $\xi_j$ and where there exists a complex term $t$ in $A_i$ which contains $\epsilon_j$.

1.d. If $\xi_i \neq \emptyset$ then $A_i = A_i \cup \xi_i$ and $i = i + 1$.

1.e. Let $A_\emptyset$ be the set of $\epsilon_j$ which does not appear neither in any constraint of $C$ nor in any $\xi_j < i$.

1.f. Let $\xi_j = A_\emptyset + A_i$.

**Step 2.** Return the sets $\xi_j = 1, i$.

**Example 3.** Let again $\Delta = \{b \rightarrow f, p \rightarrow f, p \rightarrow b\}$. The simple support functions corresponding to the three defaults in $\Delta$ are given by

$$m_1(\neg b \lor f) = 1 - \epsilon_1, \quad m_2(\neg p \lor \neg f) = 1 - \epsilon_2, \quad m_3(\neg p \lor b) = 1 - \epsilon_3.$$

The requirement of auto-deductivity gives us the following three constraints:

$$\max \left\{ pl_{\Delta}(\omega) \; \omega = b \lor f \right\} >_{\infty} \max \left\{ pl_{\Delta}(\omega) \; \omega = b \land \neg f \right\}$$

i.e., $1 >_{\infty} \epsilon_1$.

$$\max \left\{ pl\Delta(\omega) \; \omega = p \lor \neg f \right\} >_{\infty} \max \left\{ pl\Delta(\omega) \; \omega = p \land f \right\}$$

i.e., $\epsilon_2 >_{\infty} \epsilon_1$.

$$\max \left\{ pl\Delta(\omega) \; \omega = p \lor b \right\} >_{\infty} \max \left\{ pl\Delta(\omega) \; \omega = p \land b \right\}$$

i.e., $\epsilon_3 >_{\infty} \epsilon_2$.

Let us apply the previous algorithm. The set $\xi_0$ contains exactly one element $\epsilon_1$. When we put $\epsilon_0$ in the highest value then all the constraints will be satisfied. Therefore the result is $\xi_0 = \{\epsilon_1\}$, $\xi_1 = \{\epsilon_2, \epsilon_3\}$, and $\epsilon_1 >_{\infty} \epsilon_2 = \epsilon_3$.

We denote by $Bel_{\Delta}(\Delta)$ the family of e-bf’s built by Dempster’s rule and whose parameter $\epsilon$ satisfy the constraints above. As we did for the $Bel_{\Delta}(\Delta)$ family, we make sure that the elements of $Bel_{\Delta}(\Delta)$ have the right properties for our goals: they are ebf-models of $\Delta$, and they induce a unique ordering on $\Omega$.

**Lemma 6.** Let $\Delta$ be a default base. Then:

a) Any element of $Bel_{\Delta}(\Delta)$ is an ebf-model of $\Delta$.

b) Let $bel_1$ and $bel_2$ be elements of $Bel_{\Delta}(\Delta)$, and $\epsilon_1$ and $\epsilon_2$ the corresponding orderings induced on $\Omega$. Then, $\epsilon_1 < c \epsilon_2$.

Our third and last definition of entailment, called LCD (Least-Commitment plus Dempster’s rule), is obtained by focusing on the ebf-models that are in $Bel_{\Delta}(\Delta)$.

$$LCD \; \alpha \rightarrow_{LCD} \beta \text{ iff for any } bel_\epsilon \text{ in } Bel_{\Delta}(\Delta), \; bel_\epsilon \models \alpha \rightarrow \beta$$

Note that, as all the elements of $Bel_{\Delta}(\Delta)$ are ebf-models of $\Delta$, the $\rightarrow_{LCD}$ relation is as least as strong as $\rightarrow_{bf}$. In particular, $\rightarrow_{LCD}$ satisfies the KLM properties for system P (Krause et al., 1990). In fact, LCD is strictly stronger than P. For example, LCD correctly addresses the irrelevance problem, as shown by the following example.

**Example 4.** We first show how to use the result of the previous example to verify that $b \land p \rightarrow_{bf} \neg f$. In fact, by applying lemmas 1(b) and 5, we have $pl_{\Delta}(b \land p \land \neg f) = \epsilon_1$, and we know that $\epsilon_1 >_{\infty} \epsilon_2$. Next, con-
Consider a new property "red" (r) unrelated to b, p and f. We expect that red birds fly (note that this is not the case in system P). For any bel in Belα and its corresponding pl, we have
\[
\begin{align*}
pl((bArA-f)) &= \max\{pl((bArA-f)), pl((bArA-f-n))\} \\
&= \max\{\varepsilon_2, 1\} = 1 \\
pl((bArA-f-n)) &= \max\{pl((bArA-f-n)), pl((bArA-f-n-p))\} \\
&= \max\{\varepsilon_1, \varepsilon_1\} = \varepsilon_1
\end{align*}
\]
Hence, we have pl((bArA-f)) = \varepsilon_1, which implies bArA \tvd _{LCD} f, as was desired.

The following theorem summarizes the relation between (LCD) and system P.

**Theorem 3.** For a given \( \Delta \), if \( \alpha \vdash \beta \) then \( \alpha \vdash _{LCD} \beta \). 

The converse is not true.

### 6 Analysis of LCD-consequence

We have seen that the LCD consequence relation gives us strictly more than system P; in particular, it correctly answers the problem of irrelevance. In this section, we study in more detail the patterns of reasoning that are captured by LCD. To do this, we consider the desiderata listed in the introduction, and show how LCD addresses them. We also contrast the LCD solution with the one obtained by other existing systems that go beyond system P.

We start by property inheritance. Several systems, including Pearl's system Z, suffer from the problem of inheritance blocking. The canonical example is built by adding to the usual penguin problem the default b \( \rightarrow \) l (generally, birds have legs). From this, system Z cannot deduce that penguins have legs also, i.e., \( p \vdash _Z Z \).

Another desiderata listed in the introduction was the ability to stay uncommitted in cases of ambiguity. The following example shows a case of ambiguity where system Z would deduce an undesired result, while LCD does not.

**Example 6.** Let \( \Delta = \{b \rightarrow f, p \rightarrow f, p \rightarrow b, m \rightarrow f\} \), where the last default means "Generally, objects with metal-wings fly."

The simple support functions are again those in Example 3, plus the following one:
\[
\begin{align*}
m_4(l_{-m}f) &= 1-\varepsilon_4, \\
m_4(\Omega) &= \varepsilon_4.
\end{align*}
\]
The constraints that pl_β must satisfy are the same as in Example 3, plus
\[
\begin{align*}
\max\{pl_\beta(\{l_{-m}f\})\} &= \max\{pl_\beta(\{l_{-m}f\})\} = 1-\varepsilon_4, \\
\max\{pl_\beta(\{l_{-m}f\})\} &= \max\{pl_\beta(\{l_{-m}f\})\} = \varepsilon_4.
\end{align*}
\]
As the ordering above says nothing of the relative magnitude of \( \varepsilon_4 \) and \( \varepsilon_1 \), we do not have neither bApA \( \vdash _{LCD} f \) nor bApA \( \vdash _{Z} f \). Notice, by contrast, that Z would give us the arbitrary result bApA \( \vdash _Z f \).

The following theorem summarizes the relation between LCD and system Z.

**Theorem 4.** The consequence relations \( \vdash _{LCD} \) and \( \vdash _Z \) are incomparable.

The last desiderata in the introduction was syntax-independence. The following example shows that LCD is not sensitive to duplicates of rules in the default base.

**Example 7.** Consider a variant of the Quaker-Republican problem where the rule "Generally, Quaker are pacifists" has been duplicated: \( \Delta = \{q \rightarrow p, q \rightarrow p, r \rightarrow p\} \). By using a lexicographic approach, we would deduce q \( \vdash _{P} p \), while we would prefer to acknowledge the ambiguity and deduce nothing. In LCD, we have
\[
\begin{align*}
m_1(\{l_{-q}p\}) &= 1-\varepsilon_1, \\
m_1(\Omega) &= \varepsilon_1, \\
m_2(\{l_{-q}p\}) &= 1-\varepsilon_2, \\
m_2(\Omega) &= \varepsilon_2, \\
m_3(\{l_{-r}q\}) &= 1-\varepsilon_3, \\
m_3(\Omega) &= \varepsilon_3.
\end{align*}
\]
Together with the constraints
\[
\begin{align*}
\varepsilon_1 &> \varepsilon_2, \\
\varepsilon_1 &> \varepsilon_2, \\
\varepsilon_1 &> \varepsilon_3
\end{align*}
\]
By applying our algorithm, we get one single class, and so \( \varepsilon_3 \neq \varepsilon_1 \). Then
\[
\begin{align*}
pl(\{q \rightarrow p\}) &= \varepsilon_3 \text{ and } pl(\{q \rightarrow p\} \rightarrow l_{-q}p) = \varepsilon_2,
\end{align*}
\]
and hence, as desired, we have neither q \( \vdash _{LCD} p \) nor q \( \vdash _{Z} p \).

We have shown that the LCD consequence relation behaves well with respect to all of the desiderata stated in the introduction. Unfortunately, there are cases where LCD gives us results whose intuitive acceptability is questionable.

**Example 8.** Consider the Quaker-Republican problem with the extra rule "Generally, ecologists are pacifists": \( \Delta = \{q \rightarrow p, e \rightarrow p, r \rightarrow p\} \). We have
\[
\begin{align*}
m_1(\{l_{-q}p\}) &= 1-\varepsilon_1, \\
m_1(\Omega) &= \varepsilon_1, \\
m_2(\{l_{-q}p\}) &= 1-\varepsilon_2, \\
m_2(\Omega) &= \varepsilon_2, \\
m_3(\{l_{-r}q\}) &= 1-\varepsilon_3, \\
m_3(\Omega) &= \varepsilon_3.
\end{align*}
\]
Together with the constraints
The last example may be disappointing, in that we may consider that ecologists who are Quakers and republicans are pacifists.

The reason for the answer given by LCD is to be found in the multi-source interpretation of the approach: having two sources to independently justify a conclusion is regarded as a stronger reason to accept that conclusion. The assumption of independence between the sources is essential to use Dempster’s rule of combination; it may be interesting to study variants of LCD based on different rules of combination.

7. Three special cases

The LCD consequence relation has been built by giving each default \( d \) an infinitesimal weight \( \varepsilon_d \), and combining these weights by Dempster’s rule of combination. We then used the auto-deduction and least-commitment principles to constrain the possible values of the \( \varepsilon_d \)’s. As an alternative, we could have imposed some a-priori relation between the \( \varepsilon_d \)’s.

In this section, we show three special cases of consequence relations obtained in the second way. The relations we get are already known in the literature: in the first case, we get the so-called “penalty logic” (Pinkas, 1992; Dupin et al., 1994); in the second case, we get the lexicographic approaches (Dubois et al., 1992; Benferhat et al., 1993; Lehmann, 1993); and the last case leads to the preferred sub-theories of Brewka (1989).

Consider a default base \( \Delta \) and let \( \{\Delta_0, \Delta_1, \ldots, \Delta_n\} \) be the partition of \( \Delta \) given by System Z. As we did for LCD, we associate each default \( d \) in \( \Delta \) with the following simple support function

\[
m_d(I) = 1 - \varepsilon_d, \quad m_d(\Omega) = \varepsilon_d, \quad m_d(\text{elsewhere}) = 0.
\]

The three special cases we want to analyze are obtained by fixing the value of the \( \varepsilon_d \)’s infinitesimally as follows.

For the first case, we choose one “global” infinitesimal \( \varepsilon \), and we let, for each \( d \in \Delta_i \), \( \varepsilon_d = \varepsilon^i \). Then by applying Dempster’s rule of combination we get for each interpretation:

\[
\text{Pl}_d(\omega) = \prod_{d \in \Delta_i} \varepsilon_d | \omega \neq \phi_d,
\]

where \( k_i(\omega) \) is the number of defaults of \( \Delta_i \) which are not satisfied by the interpretation \( \omega \). We can interpret the resulting mass assignment \( m_\omega \) in terms of costs: each piece of information \( d \) of \( \Delta_i \) associated to the cost \( c(d) = m_\omega(d) \), interpreted as the price to pay if \( d \) is not satisfied. Then each interpretation \( \omega \) is associated to the cost of pieces of information of \( \Delta \) which are falsified by \( \omega \), namely:

\[
C(\omega) = \Sigma_{d \in \Delta_i} c(d).
\]

We can show that

\[
C(\omega) < C(\omega') \iff \text{Pl}_d(\omega) > \text{Pl}_d(\omega').
\]

(Obvious, since \( \text{pl}_d(\omega) = \varepsilon^{\Sigma_{i=1,n} k_i(\omega)} \).)

Hence, the non-monotonic consequence relation obtained by using Dempster’s combination and the \( \varepsilon_d \)’s defined above generates the same results as the inference relation based on penalty logic (Pinkas, 1991; Dupin et al., 1994).

For the second case, we equate all the \( \varepsilon_d \)’s of the defaults \( d \) that belong to the same layer \( \Delta_i \). More precisely, we associate an infinitesimal \( \varepsilon_d \) to each layer \( \Delta_i \), and, for each \( d \in \Delta_i \), let \( \varepsilon_d = \varepsilon_i \). We then ask that \( \varepsilon_i \) is a positive real number infinitely small, and that, for any \( i > 1 \),

\[
\prod_{j=1}^{i-1} \varepsilon_j | d_j | \gg \varepsilon_i .
\]

where \( |d_j| \) is number of default rules in the stratum \( \Delta_j \). We can show that the order on different interpretations obtained by using these \( \varepsilon_j \)’s can be characterized as follows, where \( |d| \) is the number of defaults of \( \Delta_i \) satisfied by \( \omega \).

Proposition 1. \( \text{pl}(\omega) >\text{pl}(\omega') \) if and only if there exist a positive number \( 1 \leq \kappa < \infty \) such that:

\[
\forall j>i, |d| \equiv |d'|, \quad \text{and} \quad |d'| > \kappa |d|. \]

This ordering corresponds to the so-called Lexicographical ordering defined in (Dubois et al., 1992; Benferhat et al., 1993; Lehmann, 1993), and hence the consequence relation obtained by using Dempster’s combination and the \( \varepsilon_d \)’s defined above is the same as the one obtained by the lexicographic approaches. This ordering has also been considered in diagnosis by De Kleer (1990) and Lang (1994).

For the last case, let \( d_{ij} \) be the \( j \)-th default (according to some arbitrary enumeration) in the \( i \)-th layer \( \Delta_i \). Then, we associate \( d_{ij} \) to an infinitesimal \( \varepsilon_{ij} \) such that:

\[
\forall i, \varepsilon_{ij} \equiv \varepsilon_{ik}, \quad \text{for} \quad k \neq j.
\]

For the last case, let \( d_{ij} \) be the \( j \)-th default (according to some arbitrary enumeration) in the \( i \)-th layer \( \Delta_i \). Then, we associate \( d_{ij} \) to an infinitesimal \( \varepsilon_{ij} \) such that:

\[
\forall i, \varepsilon_{ij} \equiv \varepsilon_{ik}, \quad \text{for} \quad k \neq j.
\]

This ordering corresponds to the Preferred sub-theories originally proposed by Brewka (1989), and later independently introduced in (Dubois et al., 1992) in the setting of possibilistic logic. 6 Hence, the consequence relation obtained by using Dempster’s combination and the \( \varepsilon_d \)’s defined above is the same as the one obtained by the system of Brewka.

As it is the case for our LCD system, all the three particular cases discussed in this section are strictly stronger than system F. In general, we state the following:

Theorem 5. LCD-consequence is incomparable with all of: penalty logic; the lexicographic approaches; and the preferred sub-theories approach.

6 Brewka’s preferred subtheories have also been used by Boutilier (1992) in system Z to define a nonmonotonic inference relation, and by Baral (1992) to combine belief bases.
8. Conclusions

We have detailed a new approach to deal with default information based on a special class of belief functions, and have used it to define three non-monotonic consequence relations. The last one, LCD-consequence, appears to be particularly attractive. We have proved that LCD is stronger than system P, and thus it satisfies the rationality postulates of Kraus, Lehmann and Magidor (1990). Moreover, we have given examples showing that LCD correctly addresses the well-known problems of irrelevance (example 4); of blocking of inheritance (example 5); of ambiguity (example 6); and of redundancy (example 7). In this, LCD has a distinctive advantage over all currently existing approaches. Finally, we have shown that the construction used to define LCD can be used to build alternative definitions of several existing systems of default reasoning, including penalty logic, the lexicographic approach, and Brewka's preferred sub-theories. Hence, our belief-function based semantics for defaults appears to be able to capture many of the current systems in a common framework.

It is interesting to notice that, despite its good behaviour, LCD does not satisfy rational monotonicity. We are currently studying different rules that can characterize LCD-consequence. Moreover, although LCD is insensitive to the number of repetitions of the same default rule, it is sensitive to the number of different rules supporting the same conclusion. This is not surprising given that LCD is based on the interpretation of defaults as items of information provided by independent sources. We plan to explore variants of LCD that abandon the assumption of independence.

Acknowledgements. Work by the second and third author has been partially supported by the ARC "BELON", founded by the Communauté Française de Belgique.

References

Adams, E. W. (1975) The logic of conditionals. Reidel, Dordrecht.

Baral, S., Kraus, J., Minker, V.S. Subrahmanian (1992) Combining knowledge bases consisting in first order theories. Computational Intelligence, 8(1), 45-71.

Benferhat, S., Dubois, D. and Prade, H. (1992) Representing default rules in possibilistic logic. Proc. of the 3rd Conf. on Knowledge Representation and Reasoning KR'92.

Benferhat, S., Cayrol, C., Dubois, D., Lang, J. and Prade, H. (1993) Inconsistency management and prioritized syntax-based entailment. Proc. of IJCAI'93.

Benferhat, S., Saffiotti, A. and Smets, Ph (1995) Belief functions and default reasoning. Tech. Rep. TR/IRIDIA/95-5, Université Libre de Bruxelles, Belgium. Available on: ftp://iridia.ulb.ac.be/pub/saffiotti/uncertainty/tr_95_5.ps.

Boutilier, G. (1992). What is a Default priority?.

De Kleer, J. (1990) Using Crude Probability Estimates to Guide Diagnosis. Artificial Intelligence, 45, 381-391.

Dubois, D., Lang, J. and Prade, H. (1992) Inconsistency in possibilistic knowledge bases — To live or not live with it. In: Fuzzy Logic for the Management of Uncertainty (L.A. Zadeh, J. Kacprzyk, eds.), Wiley, 335-351.

Dubois, D. and Prade, H. (1988) (with the collaboration of Farreny H., Martin-Clouaire R., Testemale C.) Possibility Theory – An Approach to Computerized Processing of Uncertainty. Plenum Press, New York

Dupin de Saint Cyr, F., Lang, J., and Schiex, N. (1994) Penalty logic and its link with Dempster-Shafer theory. Proc. of UAI'94, 204-211.

Gabbay, D. (1985) Theoretical foundations for non-monotonic reasoning in expert systems. In: K.R. Apt (Ed.) Logics and models of Concurrent Systems. Springer Verlag, Berlin, 439-457.

Goldszmidt, M. and Pearl, J. (1991) System Z+: A formalism for reasoning with variable-strength defaults. Procs. of the AAAI Conference, Anaheim, CA.

Hsia, Y-T. (1991) A belief-Function Semantics for Cautious Non-Monotonicity, Technical Report TR/IRIDIA/91-3, Universitë Libre de Bruxelles, Belgium.

Kraus, S., Lehmann, D. and Magidor, M. (1990) Non-monotonic reasoning, preferential models and cumulative logics. Artificial Intelligence 44, 167-207.

Lang, J. (1994) Syntax-based default reasoning as probabilistic model-based diagnosis. In Proc. of UAI'94, 391-398.

Lehmann, D. (1993) Another perspective on Default Reasoning, Technical report, Hebrew University.

Lehmann, D. and Magidor, M. (1992) What does a conditional knowledge base entail? Artificial Intelligence, 55, 1-60.

Pearl, J. (1988) Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, CA.

Pearl J. (1990) System Z: A natural ordering of defaults with tractable applications to default reasoning. Proc. of the Conf. on Theoretical Aspects of Reasoning about Knowledge, 121-135.

Pinkas, G. (1991). Propositional nonmonotonic reasoning and inconsistency in symmetric neural networks. Proc. of the 12th IJCAI, 525-530, Sydney, Australia.

Poole, D. (1993) Average-case analysis of a search algorithm for estimating prior and posterior probabilities in Bayesian networks with extreme probabilities. Proc. of IJCAI'93, 606-612.

Shafer, G. (1976) A Mathematical Theory of Evidence. Princeton University Press, New Jersey.

Shoham, Y. (1988) Reasoning About Change — Time and causation from the standpoint of Artificial Intelligence. MIT Press, Cambridge, MA.

Smets, Ph. (1988) Belief functions. In: Ph. Smets, E.H. Mandani, D. Dubois, H. Prade (Eds.) Non-Standard Logics for Automated Reasoning. Academic Press, New York, 253-286.

Smets, Ph and Hsia, Y-T. (1991) Default Reasoning and the Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model. In: P. Bonissone, M. Henrion, L. Kanal and J. Lemmer (Eds.) Transferable Belief Model.