# On the Representation of Nonmonotonic Relations in the Theory of Evidence

Ronald R. Yager
Machine Intelligence Institute
Iona College
New Rochelle, NY 10801

## Abstract

A Dempster-Shafer belief structure provides a mechanism for representing uncertain knowledge about a variable. A compatibility relation provides a structure for obtaining information about one variable based upon our knowledge about a second variable. An inference scheme in the theory of evidence involves the use of a belief structure on one variable, called the antecedent, and a compatibility relationship to infer a belief structure on the second variable, called the consequent. The concept of monotonicity in this situation is related to change in the specificity of the consequent belief structure as the antecedent belief structure becomes more specific. We show that the usual compatibility relations, type 1, are always monotonic. We introduce type II compatibility relations and show that a special class of these, which we call irregular, are needed to represent nonmonotonic relations between variables. We discuss a special class of nonmonotonic relations called default relations.

## 1. Introduction

In many reasoning environments we are required to represent our knowledge about a variable in a probabilistic manner. However it is often the case that our knowledge about the probability distribution of these random variables is not complete, we are not sure of the assignment of the probability weights to the individual outcomes of the variable. The mathematical theory of evidence [Dempster, 1967; Dempster, 1968; Shafer, 1976; Shafer, 1987; Lowrance and Garvey, 1982; Gordon and Shortliffe, 1984; Smets, 1988; Yager et al., 1994] allows for the representation and manipulation of this nonspecific type of knowledge. A number of different semantics can be associated with this theory, Dempster [1967], Shafer [1976] and Smets and Kennes [1994]. While in this paper we shall favor the one introduced by Dempster the basic ideas presented in this work are valid in the other interpretations.

Within this theory, a belief structure, a basic probability assignment function (bpa), provides a knowledge scheme for representing our incomplete knowledge about the probability distribution of a variable. One manifestation of this situation is that our knowledge of the probabilities of these type of variables is usually an interval rather then a point, it lacks specificity. Increased knowledge results in a narrowing of these intervals. A second concept which plays a central role in this theory is the idea of a compatibility relation. A compatibility relation provides information about the allowable solutions to <u>one variable given information about a second variable</u>. It is closely related to the concept of a rule in expert systems.

The inference scheme in this theory involves the inputting of a belief structure on one of the variables, called the antecedent or primary variable into compatibility relation and to obtain a belief structure on the secondary variable. A compatibility relation is called monotonic if as we gain more information about the distribution of the primary variable we don't lose information about the secondary variable. The originally compatibility relations introduced by Shafer only allowed the representation of monotonic relationships. In this work we propose a representation of compatibility relation to allow for the representation of nonmonotonic knowledge in the framework of the theory of evidence.

## 2. Dempster-Shafer Structure and Information

In this section we briefly review some ideas from the theory of evidence necessary for our development. Assume V is a random variable which can take its values in the set X. A belief structure or Dempster-Shafer structure (D-S structure) on X is a mapping m, called a basic probability assignment, (bpa), where m: $2^X \rightarrow [0, 1]$ such that (1) $\sum_{A \subset X} m(A) = 1$ and (2) $m(0) = 0$.

Assume B is any subset of X. Two important measures are associated with B in the framework of this theory. The first measure, called the measure of plausibility, denoted P1(B), is defined as

$Pl(B) = \sum_{A \subset X} Poss[B/A] * m(A)$, where Poss[B/A] = 1 if the set $A \cap B \neq \phi$ and Poss[B/A] = 0 if the set $A \cap B = \phi$. The second measure is called the measure of belief, denoted Bel[B], and defined as $Bel(B) = \sum_{A \subset X} Cert[B/A] * m(A)$ where $Cert[B/A] = 1 - Poss[\bar{B}/A]$. Cert[B/A] is a measure of the degree of inclusion of A in B, if $A \subset B$ then Cert[B/A] = 1 otherwise it is zero.

An important aspect of these two measures is that they provide an upper and lower bound on the probability of B, that is $Bel(B) \leq Prob(B) \leq Pl(B)$.[Dempster, 1967; Strat, 1984] For each B, we define R(B) = [Bel(B), Pl(B)], as our range of indefiniteness about the probability of the set B. The smaller Pl(B) – Bel(B) the more we know about the probability of B.

**Definition:** Assume $m_1$ and $m_2$ are two D-S structures on X if $R_1(B) \subset R_2(B)$ for all $B \subset X$, we shall say $m_1$ is more specific than $m_2$ and denote this as $m_1 \; S \; m_2$.

Thus a bpa being more specific than another indicates the first provides better knowledge of the probabilities involved in the situation.

If $a' \leq a$ and $b' \geq b$ and if $Prob(B) \in [a, b]$ then we can infer that $Prob(B) \in [a', b']$. This observation allows one to introduce a logical entailment principle associated with D-S granules [Yager, 1985; Yager, 1986]. In particular, if we know that $m_1$ is a valid representation of the probability structure on X and if $m_1 \; S \; m_2$ then $m_2$ also presents a true, although less informative, picture of the probability structure on X. We can capture this idea in the form of a logical entailment principle

$$m_1 \; \& \; m_1 \; S \; m_2 \vdash m_2.$$

Yager [1986] has introduced an idea of containment of two D-S structures.

**Definition:** Assume $m_1$ and $m_2$ are two D-S structures such that $A_1, \ldots, A_q$ are the focal elements of $m_1$, with their weights $m_1(A_i) = a_i$. Let $B_1, \ldots, B_n$ be the focal elements of $m_2$ with their weights $m_2(B_j) = b_j$. If a set of values $c_{ij}$, $0 \leq c_{ij} \leq 1$, $j = 1, \ldots, n$, can be found which have the following properties, then we say $m_1 \subset m_2$,

$$\sum_{j=1}^{n} c_{ij} = a_i \quad i = 1, \ldots, q$$

$$\sum_{i=1}^{q} c_{ij} = b_j \quad j = 1, \ldots, n$$

and $c_{ij} > 0$ only if $A_i \subset B_j$. The condition $c_{ij} > 0$ if $A_i \subset B_j$ is equivalent to

$$c_{ij} \leq Cert[B_j/A_i].$$

It can be easily seen that if $m_1 \subset m_2$ then $m_1 \; S \; m_2$ and if $m_1 \subset m_2$ then $m_1 \vdash m_2$.♦

Closely related ideas have been suggested by other authors. [Dubois and Prade, 1986; Lamata and Moral, 1989; Kruse and Schwecke, 1990; Klawonn and Smets, 1992]

Assume m is a bpa on X with focal elements $A_1$ and $A_2$ with $m(A_1) = a_1$ and $m(A_2) = a_2$. Let m' also be a bpa on X with focal elements $B_1$, $B_2$ and $B_3$ where $m(B_i) = b_i$ and $B_1 = A_1$, $B_2 = A_2$ and $B_3 = A_2$. Furthermore assume $b_1 = a_1$, $a_2 = b_2 + b_3$. It can be shown for any subset D of X Pl'(D) = Pl(D) and Bel(D) = Bel'(D). Thus m and m' are effectively the same bpa. We say m' is an **a ternative** <u>view</u> of m.

Assume $m_1$ is a bpa on X with focal elements $A_1, \ldots, A_q$ with $m_1(A_i) = a_i$. Assume $m_2$ is a bpa on X with focal elements $B_1, \ldots, B_n$ where $m_2(B_j) = b_j$. Assume $m_1 \subset m_2$. The following useful alternative view $m_3$ can be made of $m_2$. $m_3$ has focal elements $Q_{ij}$ where $Q_{ij} = B_j$ for all i and $m_3(Q_{ij}) = c_{ij}$ satisfy the properties in the definition of containment

$$\sum_{j=1}^{n} c_{ij} = a_i \quad i = 1, \ldots, q$$

$$\sum_{i=1}^{q} c_{ij} = b_j \quad j = 1, \ldots, n$$

$c_{ij} > 0$ only if $A_i \subset B_j$.

We shall call $m_3$ an <u>alternative view of $m_2$ tagged to $m_1$</u>. Alternatively, we can say that if $m_1$ and $m_2$ are two bpa on X and if there exists an alternative view, $m_3$, of $m_2$ tagged to $m_1$ then $m_1 \subset m_2$ and $m_1 \subset m_3$ and $m_2 = m_3$.

Assume $m_1$ and $m_2$ are two bpa on X relating information about V. The effect of both these pieces of information is a conjuncted bpa m on X denoted

$$m = m_1 \cap m_2$$

where m is obtained via Dempster's rule [Dempster, 1967] such that for each $A \neq \phi$

$$m(A) = (1/1 - k) \sum_{A_i \cap B_j = A} m(A_i) * m_2(B_j).$$

where $A_i$ and $B_j$ are the focal elements of $m_1$ and $m_2$ and

$$K = \sum_{A_i \cap B_j = \phi} m(A_i) * m_2(B_j).$$

If $K \neq 1$, we call $m_1$ and $m_2$ combinable or non-conflicting.

## 3. Compatibility Relations and Inference in D-S Structures

Assume V and U are two variables taking their values in the sets X and Y, respectively. A compatibility relation C between V and U is a relation on $X \times Y$ such that for each $x \in X$ there exists at least one y such that $(x, y) \in C$ and for each $y \in Y$ there exists at least one $x \in X$ such that $(x, y) \in C$. Formally, we shall call this a type I compatibility relation. If $A_i = \{y \mid C(x_i, y) = 1\}$ and $B_j = \{x \mid C(x, y_j) = 1\}$ then we require $A_i \neq \phi$ and $B_j \neq \phi$. Intuitively, we shall use a compatibility relation to represent the idea that if we know that $V = x_i$ then $A_i$ equals the subset of values of Y that are possible solutions for the variable U.

We note that any C can be naturally represented in matrix form,

$$\begin{array}{c} \\ x_1 \\ x_2 \\ \\ x_q \end{array} \begin{array}{c} y_1 \ldots\ldots\ldots y_n \\ \left[ \begin{array}{ccc} & & \\ & C_{ij} & \\ & & \end{array} \right] \end{array}$$

where $C_{ij} = 1$ means $(x_i, y_j) \in C$ and $C_{ij} = 0$ means $(x_i, y_j) \notin C$

Compatibility relations are effectively rules. Let A and B be subsets of X and Y. The rule *if $V \in A$ then $U \in B$* is represented by the compatibility relation $C(x, y) = 1$ for $x \in A$ and $y \in B$; $C(x, y) = 0$ for $x \in A$ and $y \notin B$ and $C(x, y) = 1$ for $x \notin A$.

A general principle for combining these type I compatibility relations can be established. Assume $K_1, K_2, \ldots, K_r$ are r pieces of knowledge about the relationship between V and U each of which is representable as a compatibility relation $C_1, \ldots, C_r$. The effect of all of these pieces of knowledge is a conjunction of the individual pieces of knowledge "*$K_1$ and $K_2, \ldots$ and $K_r$*" which results in an overall compatibility relation C, where $C = C_1 \cap C_2, \ldots, \cap C_r$. An important implication of this principle is that if C is the effective compatibility relation under $K_1, \ldots, K_r$ and if we get further information about the relationship between V and U in terms of another piece of information, $K_{r+1}$, resulting in an effective relation $C^* = C \cap C_{r+1}$ where $C^* \subset C$. We recall $C^* \subset C$ if $C^*(x, y) \leq C(x, y)$ for all x, y. Thus more information usually results in a smaller compatibility relation in the sense of cardinality.

An important class of inference problems involves the situation in which we have some knowledge about the probability structure of the variable V in terms of a bpa m on X, we have knowledge about the relationship between V and U in terms of a compatibility function C on $X \times Y$ and we are interested in obtaining knowledge about the probability structure on U in terms of a bpa $m^*$ on Y. The procedure for securing this information in essence involves an application of Dempster's rule.

Assume $A_i$ are the focal elements of m such that $m(A_i) = a_i$. We proceed as follows

1. Extend m to be a bpa on $X \times Y$ where $m(A_i \times Y) = m(A_i)$.

2. Apply Dempster's rule to C and m to obtain a conjuncted structure $m^+$ such that

$$m^+(E_i) = m(A_i \times Y) = m(A_i)$$

where $E_i = (A_i \times Y) \cap C$.

3. Obtain the bpa on Y, $m^*$ by projecting $E_i$ onto Y. That is $m^*(F_i) = m(A_i)$ where $F_i = Proj_Y[E_i] = \{y \mid$ if $E_i(x, y) = 1$ for some $x\}$. Effectively, $F_i(y) = Max_x E_i(x, y)$.

Since it is possible for two different E's to project into the same F,

$$m^*(F_j) = \sum_{\substack{\text{over } E_i \text{ s.t.} \\ F_j = \text{Proj } E_i}} m(E_i).$$

Essentially we call a compatibility relation monotonic if an increase in knowledge about the antecedent does not result in a decrease in knowledge about the consequent. In the following definition we formalize this idea.

**Definition:** Assume C is a compatibility relation on $X \times Y$. Let $m_1$ and $m_2$ be two bpa on X such that $m_1 \, S \, m_2$. If $m_1$ and C allows us to infer $m_1^*$ on Y and $m_2$ and C allows us to infer $m_2^*$ on Y then C is said to be **monotonic** if $m_1^* \, S \, m_2^*$.

It can be shown [Yager, 1988] that every type I compatibility relation is monotonic and thus it is impossible to represent a nonmonotonic relationship with these type I compatibility relations. Since many kinds of commonsense knowledge have a nonmonotonic aspect there is a need to provide compatibility structures to represent this type of knowledge. In the following we introduce such a structure.

## 4. Non-monotonicity and Type II Compatibility Relations

Assume V and U are two variables taking values in the set X and Y respectively. Let $\mathcal{X}$ be the power set of X minus the null element, thus $T \in \mathcal{X}$ is a non-null subset of X.

**Definition:** A type II compatibility relation R is a relation on $\mathcal{X} \times Y$ such that for each $T \in \mathcal{X}$ there exists a $y \in Y$ such that $R(T, y) = 1$.

The understanding to be accorded a type II compatibility relation is that if $R(T, y) = 1$ then if V is known to be some element in T then y is a possible value for U, thus $R(T, y)$ implies $(x, y)$, for all $x \in T$, are possible solution pairs for V and U.

The following terminology and definitions shall be useful in our future discussions. We shall call the subsets of X which are singletons the principle elements of $\mathcal{X}$. With each $T \in \mathcal{X}$ we shall denote W to be the subset of Y that are the possible values of U when $x \in T$, $W = \{y \mid \text{if } R(T, y) = 1\}$. W is called the associated set in Y of T and sometimes we shall denote this pair as $T \rightarrow W$.

With $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2, y_3\}$ a type II compatibility relation would be as shown in the following matrix.

| | $y_1$ | $y_2$ | $y_3$ |
|---|---|---|---|
| $\{x_1\}$ | 1 | 0 | 0 |
| $\{x_2\}$ | 1 | 0 | 0 |
| $\{x_3\}$ | 0 | 1 | 1 |
| $\{x_1, x_2\}$ | 1 | 1 | 0 |
| $\{x_1, x_3\}$ | 1 | 1 | 1 |
| $\{x_2, x_3\}$ | 1 | 1 | 1 |
| $\{x_1, x_2, x_3\}$ | 1 | 1 | 1 |

**Definition:** Assume $T_i$ are the principle elements of $\mathcal{X}$, $T_i = \{x_i\}$ and let $T_i \to W_i$. Assume T is some arbitrary non-principle element in X where $T \to W$. We call a relation R **normal** if

$$W \subset \bigcup_{\substack{i \\ x_i \in T}} W_i.$$

Thus we see that a relation is normal if an associated set is contained in the union of the associated sets of principle elements making it up.

In point of fact, all reasonable compatibility relations can be seen to be normal. Assume there exists $y \in W$ but $y \notin \bigcup_{\substack{i \\ x_i \in T}} W_i$. Since the knowledge $V \in T$ implies that the value of V must be some $x \in T$, the value for U must be possible under an $x \in T$, but if y is not associated with any $x \in T$ then it can't be a solution.

**Definition:** A normal compatibility relation R of type II shall be called **regular** if for all $T_1 \to W_1$, $T_2 \to W_2$, $T_3 \to W_3$ such that $T_3 = T_1 \cup T_2$ then $W_3 = W_1 \cup W_2$. It is called **irregular** if there exists such a triple where $W \subset W_1 \cup W_2$.

We shall now describe the reasoning process used when we have a type II compatibility relation. Assume R is a type two compatibility relation on $\mathcal{X} \times Y$ Let our information about V be given in terms of a bpa m on X such that $A_1, \ldots, A_p$ are the focal elements of m where $m(A_i) = a_i$. We are here interested in obtaining a bpa m* on Y providing the information about U. The procedure is essentially the same as in the case of a type I compatibility relation after an initial space transformation is made.

**0.** Transform m to an equivalent bpa m' on $\mathcal{X}$ as follows:

m' is a bpa on $\mathcal{X}$ with focal elements $A_i' \in \mathcal{X}$ where

$$A_i' = \{A_i\} \text{ and } m'(A_i') = m(A_i) = a_i.$$

That is the focal elements of m are singleton subsets of $\mathcal{X}$ whose element is the corresponding focal element in X. Thus if $A_i = \{x_1, x_2\}$ then $A_i' = \{\{x_1, x_2\}\}$.

**1.** Extend m' to be a bpa on $\mathcal{X} \times Y$ such that $m'(A_i' \times Y) = m'(A_i') = m(A_i) = a_i$

2. Conjunct R and m' to obtain a bpa $m^+$ also on $\mathcal{X} \times Y$ such that $m^+(E_i') = a_i$ where

$$E_i' = (A_i' \times Y) \cap R$$

3. Obtain the bpa m* on Y by projecting $E_i'$ onto Y. That is $m^*(F_i') = a_i$ where

$$F_i' = \text{proj}_Y(E_i'), \text{ that is}$$

$$F_i'(y) = \underset{T \in \mathcal{X}}{\text{Max}}[E_i'(T, y)].$$

Using this inference mechanism we can show that *every* *__regular__* *type II relation is monotonic.* We will now proceed to show that all all irregular relations are non-monotonic. We recall that a relation R is non-monotone if there exists a pair $m_1$ and $m_2$ on X where $m_1 \subset m_2$ such that, $R \circ m_1 \not\subset R \circ m_2$.

**Theorem:** All irregular relations are non-monotone.

**Proof:** Assume R is an irregular relation. This implies that there exists at least three elements $T_1$, $T_2$ and T of $\mathcal{X}$ such that $T_1 \cup T_2 = T$ with $W_1$, $W_2$ and W being their associated sets in Y such that W is strictly less than $W_1 \cup W_2$, ie. there exists one $y^* \in y$ such that $y^* \notin W$ and but where $y^* \in W_1$ or $y^* \in W_2$. Consider the two bpa $m_1$ and $m_2$ on X such that $m_1$ is defined by $m_1(T_1) = a$ and $m_1(T_2) = 1 - a$ where $a > 0$ and where $m_2$ is defined by $m_2(T) = 1$. It is obvious that $m_1 \subset m_2$, since $T_1 \subset T$ and $T_2 \subset T$. We first note that if $m_1^* = R \circ m_1$ then $m_1^*(W_1) = a$ and $m_1^*(W_2) = 1 - a$ and if $m_2^* = R \circ m_2$ then $m_2^*(W) = 1$. Consider the subset $D = \{y^*\}$ of Y, $Pl_1(D) = a \text{ Poss}[D/W_1] + (1 - a) \text{ Poss}[D/W_i]$ and $Pl_2(D) = \text{Poss}[D/W]$. Since $y^* \notin W$ then $\text{Poss}[D/W] = 0$ and $Pl_2(D) = 0$. Since y is in at least one of $W_1$ or $W_2$ without loss of generality assume it is definitely in $W_1$ then $\text{Poss}[D/W_1] = 1$ and hence $Pl_1(D) \geq a$. Therefore

$$[Bel_1(D), Pl_1(D)] \not\subset [Bel_2(D), Pl_2(D)]$$

thus $m_1^* \not\subset m_2$ and the theorem is proven.

This result taken with our previous observation indicates that not only are all irregular relations nonmonotonic but the only way to represent a nonmonotonic compatibility in terms of normal relations is via an irregular relation. Thus these irregular type II compatibility relations provide the desired structure for representing nonmonotonic knowledge.

## 5. Default Compatibility Relations

A particularly important class of nonmonotonic relations are the so-called default relations [Reiter, 1980; McCarthy, 1986; McDermott and Doyle, 1980; Ginsberg, 1987]. Default type knowledge plays a crucial role in the representation of commonsense knowledge. Basically, a default rule is used to represent situations in which, when we are uncertain as to the value of V, we act as if the value of V was known to be the default value. For example, if we have a bird but we are sure of what type it is we act as if it was a flying bird. Another type of default rule is when you are uncertain as to the cause of some symptom and act as if it were the one simplest to fix. We now look at the representation of these default relations as irregular compatibility relations.

The essential structure of a default relation is captured in the following relation. Assume V and U are two variables taking their values in $X = \{x_1, x_2\}$ and $Y = \{y_1, y_2\}$ respectively. Assume that the relation between U and V is

| | $y_1$ | $y_2$ |
|---|---|---|
| $T = \{x_1\}$ | 1 | 0 |
| $T = \{x_2\}$ | 1 | 1 |
| $T = \{x_1, x_2\}$ | 1 | 0 |

Since $\{x_1\} \cup \{x_2\} = \{x_1, x_2\}$ we have $T_1 \cup T_2 = T_3 \rightarrow W \subset W_1 \cup W_2$.

A fruitful way of interpreting this relationship is by observing that we infer the default consequent $y_1$ when our antecedent knowledge allows for the possibility that the value of V can be $x_1$. Thus we infer $y_1$ for T such that $Poss[\{x_1\}/T] = 1$. A natural reading of this default relation would be

"if V is $x_1$ is possible then U is $y_1$".

A more general pattern for default relations can be introduced. Let $X = \{x_1, \ldots, x_q\}$ and let $Y = \{y_1, \ldots, y_p\}$. Let A be a subset of X and let $B_1$ and $B_2$ be subsets of Y where $B_2 \not\subset B_1$. Let the relationship R between V and U be characterized by the following, $R: X \rightarrow Y$ such that for any $T \in X$

$$R \circ T = B_1 \quad \text{if } Poss[A/T] = 1$$
$$R \circ T = B_2 \quad \text{if } Poss[A/T] = 0.$$

Thus we see that R is a function of Poss[A/T]. We shall indicate this as $R = D(A, B_1, B_2)$ and call it a standard default rule. We can read this as "if V is in A is possible then U is in $B_1$ else it is in $B_2$". An important special case occurs when $B_1 = Y$. This is called a simple default rule.

**Theorem:** A standard default rule generates an irregular compatibility relation.

**Proof:** 1) If $T_1, T_2 \subset A$ then
$T_1 \rightarrow W_1 = B_1$, $T_2 \rightarrow W_1 = B_1$ and
$T_1 \cup T_2 \subset T \rightarrow W_1 = B$ thus
$T_1, T_2 \subset T \rightarrow T_1 \cup T_2 = T_1 \rightarrow W_1 \cup W_2 = W$.

2) If $T_1, T_2 \subset \bar{A}$ we similarly show $T_1 \cup T_2 = T \rightarrow W_1 \cup W_2 = W$.

3) If it is not contained in either A or $\bar{A}$ then $T = T_1 \cup T_2$ where $T_1 \subset A$ and $T_2 \subset B$ since $T_1 \rightarrow B_1$ and $T_2 \rightarrow B_2$ and $T \rightarrow B$ then $T = T_1 \cup T_2 \rightarrow W \subset B_1 \cup B_2$.

Assume our knowledge of V is expressed by the bpa $A_1, \ldots, A_r$ where $m(A_i) = A_i$. Let our relation between V and U be expressed by the standard default rule $D(A, B_1 B_2)$. Then if m* is our expression of knowledge of U

$$m^*(W_i) = a_i$$

where $W_i$ is the associated Y set of $A_i$. Furthermore $A_i = B_1$ if $Poss[A/A_i] = 1$ and $A_i = B_2$ if $Poss[A/A_i] = 0$ thus m* is obtained as

$$m^*(B_1) = \sum_{i=1}^{r} Poss[A_i/A] * m(A_i)] = Pl(A)$$
$$m^*(B_2) = \sum_{i=1}^{r} (1 - Poss[A/A_i]) * m(A_i).$$
$$= \sum_{i=1}^{r} Cert[\bar{A}] * m(A_i) = Bel(\bar{A}).$$

An interesting type of default rule is what we shall call a hierarchical default rule. Assume $A_1, \ldots, A_k$ are a collection exclusive and exhaustive subsets of X, ie $A_1 \cap A_j = \phi$ and $\bigcup_{i=1}^{k} A_i = X$. Let $B_1, \ldots, B_k$ be a set of subsets of Y such that for $i > j$, $B_i \not\subset B_j$. We shall call R a hierarchical relation,

if $R: X \rightarrow Y$

such that

$R \circ T = B_1$ if $Poss[A_1/T] = 1$

$R \circ T = B_2$ if $Poss[A_1/T] = 0$ and $Poss[A_2/T] = 1$

$R \circ T = B_3$ if $Poss[A_1/T] = Poss[A_2/T] = 0$ and $Poss[A_3/T] = 1$

$R \circ T = B_e$ if $Poss[A_j/T] = 0 \quad j = 1, \ldots, e-1$ and $Poss[A_e/T] = 1$.

We shall denote such a relation as
$H(A_1, A_2, \ldots, A_k: B_1, \ldots, B_k)$.

We see that

$$D(A, B_1, B_2) = H(A, \bar{A}: B_1, B_2).$$

A prototypical rule which generates a hierarchical compatibility relation is the rule *always do the easiest thing*. Consider a device made up of three parts $q_1, q_2, q_3$. Let $x_i$ indicate the proposition "part i is busted" and let $y_i$ indicate the action "replace part i". Assume the parts are such that $q_1, q_2, q_3$ is the order in ascending difficulty of fixing the parts. Then the following hierarchical relation implements this rule

| | $Y_1$ | $Y_2$ | $Y_3$ |
|---|---|---|---|
| $\{x_1\}$ | 1 | 0 | 0 |
| $\{x_2\}$ | 0 | 1 | 0 |
| $\{x_3\}$ | 0 | 0 | 1 |
| $\{x_1, x_2\}$ | 1 | 0 | 0 |
| $\{x_1, x_3\}$ | 1 | 0 | 0 |
| $\{x_2, x_3\}$ | 0 | 1 | 0 |
| $\{x_1, x_2, x_3\}$ | 1 | 0 | 0 |

**We can easily show that these hierarchical default relations result in a irregular compatibility relation.**

## 6. References


[Dempster, 1967] Arthur P. Dempster. Upper and lower probabilities induced by a multi-valued mapping. Ann. of Mathematical Statistics 38: 325-339, 1967.

[Dempster, 1968] Arthur P. Dempster. A generalization of Bayesian inference. Journal of the Royal Statistical Society, 205-247, 1968.

[Dubois and Prade, 1986] Didier Dubois and Henri Prade A set-theoretic view of belief functions. International Journal of General Systems 12: 193 - 226, 1986.

[Ginsberg, 1987] Matthew L. Ginsberg. Readings in Nonmonotonic Reasoning, Morgan Kaufmann, Los Altos, CA, 1987.

[Gordon and Shortliffe, 1984] Jean Gordon and Edward H. Shortliffe. The Dempster-Shafer theory of evidence. In Rule-Based Expert Systems: the MYCIN Experiments of the Stanford Heuristic Programming Project, Buchanan, B.G. & Shortliffe, E.H. (eds.), Reading, MA: Addison Wesley, 272-292, 1984.

[Klawonn and Smets, 1992] Frank Klawonn and Philippe Smets. The dynamic of belief in the transferable belief model and specialization-generalization matrices. In Uncertainty in AI 92, edited by Dubois, D., Wellman, M., D'ambrosio, B. and Smets, P., Morgan-Kaufmann: San Mateo, CA, 130-137, 1992.

lKruse and Schwecke, 1990]. Rudolf Kruse and E. Schwecke. Specialization: a new concept for uncertainty handling with belief functions. International Journal of General Systems 18: 49-60, 1990.

[Lamata and Moral, 1989] Maria T.Lamata and Serafin Moral. Classification of fuzzy measures. Fuzzy Sets and Systems 33: 243-243, 1989.

[Lowrance and Garvey, 1982] John D. Lowrance and Thomas D. Garvey. Evidential reasoning: A developing concept," In Proceedings of the IEEE International. Conf. on Cybernetics and Society, pages 6-9, Seattle, WA, 1982.

[McCarthy, 1986] John McCarthy. Applications of circumscription to formalizing common sense knowledge. Artificial Intelligence 28: 89-116, 1986.

[McDcrmott and Doyle, 1980]. Drew McDermot and Jon Doyle. Non-monotonic logic I. Artificial Intelligence 13: 41-72, 1980.

[Reiter, 1980] Raymond Reiter. A logic for default reasoning. Artificial Intelligence 13: 81-132,1980.

[Shafer, 1976] Glenn Shafer. A Mathematical Theory of Evidence, Princeton University Press, Princeton, N.J., 1976.

[Shafer, 1987] Glenn Shafer. Belief functions and possibility measures. In Analysis of Fuzzy Information, Vol 1: Mathematics and Logic, edited by Bezdek, J. C, CRC Press, Boca Raton: Florida, 1987.

[Smets, 1988]. Philippe Smets. Belief Functions. In Non-standard Logics for Automated Reasoning, Smets, P., Mamdani, E.H., Dubois, D. & Prade, H. (eds.), London: Academic Press, 253-277, 1988.

[Smets and Kennes, 1994] Philippe Smets and Robert Kennes. The transferable belief model. Artificial Intelligence 66: 191-234, 1994.

[Strat, 1984] Thomas M. Strat. Continuous belief functions for evidential reasoning. In Proceedings 4th American Association for Artificial Intelligence Conference pages 308-313, Austin, TX, 1984.

[Yager, 1985] Ronald R. Yager. Reasoning with uncertainty for expert systems. In the Proceedings of the Ninth International Joint Conference on Artificial Intelligence pages 1295-1297, Los Angeles, 1985.

[Yager, 1986] Ronald R. Yager. The entailment principle for Dempster-Shafer granules," Int. J. of Intelligent Systems 1: 247-262, 1986.

[Yager, 1988 Ronald R. Yager. Non-monotonic compatibility relations in the theory of evidence. Int. J. of Man-Machine Studies 29: 517-537, 1988.

[Yager et al., 1994] Ronald R. Yager, Janus Kacprzyk and Mario Fedrizzi. Advances in the Dempster-Shafer Theory of Evidence, John Wiley & Sons, New York, 1994.