Nonmonotonicity and Compatibility Relations in Belief Structures

Ronald R. Yager
Machine Intelligence Institute
Iona College
New Rochelle, NY 10801
E-mail: yager@panix.com

Abstract: We concern ourselves with the situation in which we use the Dempster-Shafer belief structure to provide a representation of a random variables in which our knowledge of the probability distribution is imprecise. We discuss the role of compatibility relations as a means of enabling inference about one variable, the secondary variable, based upon knowledge about another variable, the primary. We define monotonicity as a condition in which an increase in information about the primary variable in an inference should not result in a decrease in information about the secondary variable. We show what are the conditions required of a compatibility relation to lead to monotonic and nonmonotonic inferences. We provide some examples of nonmonotonic relations.

1. Introduction

The pioneering work of Dempster and Shafer [1, 2] has resulted in the development of an uncertainty modeling framework which can be used to provide a generalization of probability theory by allowing the imprecise knowledge of the probabilities. Here our knowledge about the probabilities of the various events is usually expressed in terms of interval in which it lies, it lacks specificity. Increase in information about the underlying probabilities results in a narrowing of these intervals.

Within this D-S framework a concept which plays a central role is the compatibility relation, it provides knowledge about the allowable solutions to one variable given information about a second variable [3]. It is closely related to the concept of a rule in knowledge based systems. A fundamental inference schema within the Dempster-Shafer theory involves the use of a compatibility relation and a belief structure on one of the variables, called the primary variable, to infer a belief structure on the secondary variable. A compatibility relation is called monotonic if an increase in information about the primary variable can't result in a loss of information about the secondary variable. Here we investigate an extension of the basic concept of a compatibility relation to allow for the representation of non-monotonic knowledge in the framework of the D-S theory..

2. Dempster-Shafer Structure and Information

A D-S belief structure [2] on a set X is a mapping $m: 2^X \rightarrow [0, 1]$, called a basic probability assignment (bpa), such

$$(1) \quad \sum_{A \subset X} m(A) = 1$$

$$(2) \quad m(\emptyset) = 0.$$

The subsets of X for which $m(A) \neq 0$ are called the focal elements of m. We shall find it convenient to denote these focal elements as A_j.

One use of this structure is that of representing information about the probability distribution of a variable V taking its value in the set X in the case in which the probability distribution is imprecisely known. In this interpretation the assignment $m(A_j) = a_j$ is meant to indicate that in some unknown manner a_j units of probability are to be divided among the elements in A_j.

Example: A simple example will illustrate type of representation. Assume we are about to have a presidential election and are concerned about the interest rates that will result. Assume there are three candidates for president, Tom, Dick and Harry. The polls tell us the probability of each candidate winning: Prob(Tom) = 0.5, Prob(Dick) = 0.3 and Prob(Dick) = 0.2. In addition each candidate has specified his policy regarding interest rates, Tom said he wants low interest rates, between 4% and 6%, Dick said he wants interest rates between 5% and 8% and Harry said he wants interest rates between 7% and 10%. This information induces a D-S belief structure on the set X of possible interest rates. Here our focal elements are $A_1 = \{x/ 4\% \leq x \leq 6\%\}$, $A_2 = \{x/ 5\% \leq x \leq 8\%\}$ and $A_3 = \{x/ 7\% \leq x \leq 10\%\}$ and

$m(A_1)$ = Prob(Tom) = 0.5, $m(A_2)$ = Prob(Dick) = 0.3 and $m(A_3)$ = Prob(Harry) = 0.3.

Assume m is a D-S belief structure on X, two important measures associated with this are plausibility and belief. Let B be any subset of X the plausibility of B, denoted Pl(B), is defined as $Pl(B) = \sum_{A \subset X} Poss[B/A]\, m(A)$ where $Poss[B/A] = Max_x[B(x) \wedge A(x)]$, B(x) and A(x) being the characteristic functions of the sets A and B and $\wedge$ = min. The second measure, the belief of B, denoted Bel(B), is defined as $Bel(B) = \sum_{A \subset X} Cert[B/A]\, m(A)$ where $Cert[B/A] = 1 - Poss[\bar{B}/A]$.

Observation: If A_j are the focal elements of m then $Pl(B) = \sum_{j,\, A_j \cap B \neq \varnothing} m(A_j)$ and $Bel(B) = \sum_{j,\, A_j \subseteq B} m(A_j)$

Observation: If $B_1 \subset B_2$ then $Pl(B_1) \leq Pl(B_2)$ and $Bel[B_1] \leq Bel[B_2]$.

Observation: If $B \neq \varnothing$ then $Pl(B) \geq Bel(B)$

In interpreting a D-S belief structure as providing imprecise information about the probability of a variable V the significance of these two measures is that they provide upper and lower bounds on the probability of B [1],

$$Bel(B) \leq Prob(V \in B) \leq Pl(B).$$

Assuming that V is a random variable, its generated by a probability distribution, the best knowledge we can have about V is specific knowledge of the probability of each subset. For any subset B we shall define R(B) = [Bel(B), Pl(B)], our range of indefiniteness about the probability of the subset B. The smaller R(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 $R_1(B) \subseteq R_2(B)$ for all $B \subseteq X$, we shall say m_1 is more **specific** than m_2 and denote this as m_1 ***S*** m_2.

It is clear that the more specific a D-S structure the more information we have about the probability distribution.

Observation: If $a' \leq a$ and $b' \geq b$ and it is known that $Prob(B) \in [a, b]$ and then we can infer that $Prob(B) \in [a', b']$

This observation allows us to introduce a logical entailment principle associated with D-S structures [4]: 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 represents a valid, although less informative, picture of the probability structure on X.

In [4] Yager has introduced an idea of containment of two D-S structures.

Definition: Assume m_1 and m_2 are two D-S structures on X. Let A_i be the q focal elements of m_1, where $m_1(A_i) = a_i$. Let B_j be the n focal elements of m_2, where $m_2(B_j) = b_j$. If a set of values c_{ij} for i = 1 to q and j = 1 to n, can be found which have the following properties,

$$0 \leq c_{ij} \leq 1$$
$$\sum_{j=1}^{n} c_{ij} = a_i \quad i = 1 \text{ to } q$$
$$\sum_{j=1}^{q} c_{ij} = b_j \quad j = 1 \text{ to } m$$
$$c_{ij} > 0 \text{ only if } A_i \subset B_j$$

then we say $m_1 \subset m_2$

Yager [4] has proved the following theorem,

Theorem: $m_1 \subset m_2 \rightarrow m_1$ ***S*** m_2.

Let us briefly look at the issue of combining belief structures. Assume m_1 and m_2 are two bpa on X providing information about V. The effect of both pieces of information, obtained via Dempster's rule [1, 2], is a conjuncted bpa m on X denoted $m = m_1 \cap m_2$ where m for each $A \neq \varnothing$

$$m(A) = \frac{1}{1 - K} \sum_{A_i \cap B_j = A} m(A_i)\, m_2(B_j).$$

here A_i and B_j are the focal elements of m_1 and m_2 and $K = \sum_{A_i \cap B_j = \varnothing} m(A_i)\, m_2(B_j)$.

In above K is interpreted as the degree of conflict between m_1 and m_2, if K = 1 they are completely conflicting while if K = 0 there is no conflict.

3. Compatibility Relations and Inference in D-S Structures

Information about one variable can be obtained using information about another variable if we have some knowledge about how they are related, this forms the basis of most inference systems. Assume V and U are two variables taking their values in the sets X and Y respectively. In D-S theory compatibility relations are used to represent this knowledge.

Definition: A type I **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 $(x, y) \in C$. If $A_i = \{y \mid C(x_i, y) = 1\}$ and $B_j = \{x \mid C(x, y_j) = 1\}$ then we require $A_i \neq \varnothing$ and $B_j \neq \varnothing$.

From a knowledge-base perspective we can use a compatibility relation to represent the knowledge that if $V = x_i$ then A_i is the subset of values of Y that are possible solutions for the variable U.

We now describe some examples of compatibility relations which represent some familiar forms of knowledge. Let $X = \{x_1, ..., x_q\}$, $Y = \{y_1, ..., y_n\}$ and $B = \{y_1, ..., y_e\}$ $e \leq n$.

Consider the knowledge: *if* $V = x_1$ *then* $U \in B$ this can be represented by the relation C

$C(x_1, y) = 1$ for $y \in B$

$C(x_1, y) = 0$ for $y \notin B$

$C(x, y) = 1$ for $x \neq x_1$,

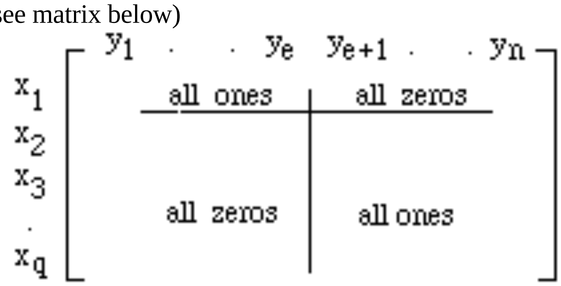
Consider now the knowledge

if $V = x_1$ then $U \in B$ **and** if $V \neq x_1$ then $U \in \overline{B}$.

This is represented by the compatibility relation C where

$C(x_1, y) = 1$	$y \in B$
$C(x_1, y) = 0$	$y \notin B$
$C(x, y) = 1$	$x \neq x_1$ and $y \in \overline{B}$
$C(x, y) = 0$	$x \neq x_1$ and $y \notin \overline{B}$

(see matrix below)



More generally if $A = \{x_1, . . . , x_k\}$ $(k \leq q)$ the compatibility relation

$C(x, y) = 1$	$x \in A$ and $y \in B$
$C(x, y) = 0$	$x \in A$ and $y \notin B$
$C(x, y) = 1$	$x \notin A$.

is representative of the knowledge "if $V \in A$ then $U \in B$".

Assume K_j are r pieces of knowledge about the relationship between V and U each of which is representable as a compatibility relation C_j. The effect of all of these pieces of knowledge is the conjunction of the individual pieces of knowledge *"K_1 and K_2, . . . and K_r"* which results in an overall compatibility relation C, where $C = C_1 \cap C_2 .... \cap C_r$.

An important implication of this is that if C is the effective compatibility relation under K_1, ..., K_r and if we get additional information about the relationship between V and U in terms of another piece of information, K_{r+1}, this results in an effective relation $C^* = C \cap C_{r+1}$ where $C^* \subseteq C$, $C^*(x, y) \leq C(x, y)$ for all x, y. Thus more information usually results in a smaller compatibility relation in the sense of containment. The combining of individual compatibility relations can result in combined compatibility relationship C for which there exists some x such that there exists no $y \in Y$ for which $C(x, y) = 1$. In this case we shall say that our knowledge about the relationship between U and V is **conflicting**. If the conjunction of knowledge leads to a conflict, we must resolve this conflict before proceeding to use the resulting relationship. We shall not address this issue here and assume all conjunctions are non-conflicting.

A second, though less pernicious problem that arises when we combine compatibility relations is degeneracy. We shall say a relation C is degenerate if there exists a $y \in Y$ for which there exists no x such that $C(x, y) = 1$. Degeneracy doesn't preclude our using the resulting C.

An important type of inference occurs in the situation in which we have some information about a variable V in terms of a D-S belief structure m on X, additionally 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 U in terms of a D-S belief structure m* on Y. Once realizing that a compatibility relationship can be viewed as D-S belief structure on $X \times Y$ having one focal element C, the procedure for obtaining this information involves an application of Dempster's rule. Assuming A_i are the focal elements of m our inference procedure is as follows:

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

2. Apply Dempster's rule to C and m to obtain a D-S structure m^+ on $X \times Y$ with focal elements $E_i = (A_i \times Y) \cap C$ where $m^+(E_i) = m(A_i)$.

3. Project m^+ onto Y obtaining the D-S structure m* with focal elements $F_i = Proj_y[E_i] = \{y \mid \text{if } E_i(x, y) = 1 \text{ for some } x\}$ and where $m^*(F_i) = m(A_i)$. (Effectively $F_i(y) = Max_x[E_i(x, y)]$.

Symbolically we shall denote the process of inferring m* from C and m as $m^* = C \otimes m$. Informally a compatibility relation can be seen as being monotonic if knowing more about the distribution on V allows us to know more about the distribution on U.

Definition: Let C be a compatibility relation on $X \times Y$ and let m_1 and m_2 be any two bpa on X such that m_1 $\boldsymbol{S}$ m_2. If $m_1^* = C \otimes m_1$ and $m_2^*\ C \otimes m_2$ then C is said to be monotonic if m_1^* $\boldsymbol{S}$ m_2^*.

The following theorem addresses the monotonicity property for type I compatibility functions.

Theorem: Every type I compatibility relation is monotonic

We note this theorem tells us that it is impossible to represent a non-monotonic relationship with these type I compatibility relations. Another interesting result which relates to the concept of monotonicity can be obtained using the results of the following theorem.

Theorem: Assume C and $\hat{C}$ are two compatibility relations on $X \times Y$ such that $\hat{C} \subseteq C$. Let m be a D-S structure m on X. If $m^* = C \otimes m$ and $\widehat{m^*} = \hat{C} \otimes m$ then $\widehat{m^*} \subseteq m^*$ and hence $\widehat{m^*}$ $\boldsymbol{S}$ m^*.

This theorem says that the smaller the compatibility relationship the more information obtained, this of course makes sense in that the smaller the relationship the less possible U values for a given V value. There are however some further implications of this theorem for knowledge based systems. Let C_j for j = 1 to q be a collection nonconflicting compatibility relations relating V and U. Effectively these implies in a compatibility relation $C = \bigcap_{j=1}^{q} C_j$. Consider now the acquisition of another piece of non-conflicting knowledge C_{q+1}, this results in a new effective compatibility relation $\hat{C} = C \cap C_{q+1}$ since $\hat{C} \subseteq C$, the above therefore says we get better information using $\hat{C}$ then using C. The implication here is that adding nonconflicting information in terms additional compatibility relations generally results in a increase, can't cause a decrease, in our knowledge about the inferred belief structure.

Another implication of this theorem is the following. Assume C_1 is the compatibility between V and U and C_2 is another relation such that $C_1 \subset C_2$. If we use C_2 instead of C_1 to do our inferring, we shall get a correct inference although it may be less specific than one we would have gotten had we used C_1.

4. 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 on $\mathcal{X} \times Y$ is a relation such that for each $T \in \mathcal{X}$ there exists at least one $y \in Y$ such that R(T, y) = 1.

In our framework the understanding to be accorded a type II compatibility relation is that if V *is* T, V is known certainly to be one of the elements in T, V the condition R(T, y) = 1 implies that y is a possible value for U. We shall find useful to denote V *is* T as $V \in_E T$.

When $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2, y_3\}$ an example of a type II compatibility relation is

	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

In this example if we know that $V = x_1$ then the only possible solution for U is is y_1. If we know that $V = x_3$ then possible values for U are y_2 and y_3. If we know that $V \in \{x_1, x_2\}$ then the possible solutions of U are y_1 and y_2.

The following terminology and definitions shall be useful in discussing type II compatibility relations. We shall call the subsets of X which are singletons the **principle elements** of $\mathcal{X}$. We shall denote these principle elements as T_i where $T_i = \{x_i\}$. For any $T \in \mathcal{X}$ we shall let W be the subset of Y which are possible values of U when $V \in_E T$, W = {y | if R(T, y) = 1}. We call W the **associated set** of T and denote this pair as $T \to W$. For the principle elements we denote the associated sets as W_i, $T_i \to W_i$.

We now provide a classification of these type II relations

Definition: Assume R is a compatibility relation of type II:

i. It shall be called **regular** if for all T_a, T_b and T_c where $T_c = T_a \cup T_b$ we have $W_c = W_a \cup W_b$.

ii. It shall be called **irregular** if there exists a triple T_a, T_b and T_c where $T_c = T_a \cup T_b$ such that W_c is strictly contained in $W_a \cup W_b$, $W_c \subset W_a \cup W_b$.

iii. It shall be called **super-regular** if there exists a triple T_a, T_b and T_c where $T_c = T_a \cup T_b$ such that $W_c \supset W_a \cup W_b$.

We note these three definitions exhaustively cover all possible type II relations. While regularity excludes it from being any of the other two, it is possible for a relationship to be both irregular and super regular.

A quality that should be inherent in any rational representation of relational knowledge is captured by the following principle of rationality.

Principle of Rationality: Since the knowledge $V \in_E T$ implies that the value of V must be some $x \in T$, any solution for U that is possible under T, any $y \in W$, must be possible under some $x_i \in T$.

The concept of a normal compatibility relations plays a important role in rationality.

Definition: We call a type II compatibility relation R **normal** if for every pair $T \rightarrow W$

$$W \subseteq \bigcup_{i \text{ s.t. } x_i \in T} W_i.$$

A relation is normal if the associated set of T is contained in the union of the associated sets of principle elements making up T.

Theorem: All rational compatibility relations must be normal.

Definition: Assume R is a compatibility relation with $T_i \rightarrow W_i$. We call R* its *normalizing relation* if R* is defined such that the principle elements of R* have the same associated sets as R and for all T not principle elements $T \rightarrow W$ where $W = \bigcup_{i \text{ s.t. } x_i \in T} W_i$.

For any R we define R_N to be its normalized version where

$$R_N(T, y) = R(T, y) \wedge R^*(T, y). \; (\wedge \text{ is the min}).$$

This process is called normalization. If R is normal than $R_N = R$.

Since the principle of rationality provides a form of universal knowledge that can always be applied to any compatibility relation.

Imposition of Rationality – Any non-normal compatibility relation can be replaced by its normalized version.

It should be noted that normality doesn't preclude either super regularity or irregularity as the following relationship illustrates

	y1	y2	y3
{x1}	1	0	0
{x2}	0	1	0
{x3}	0	0	1
{x1, x2}	1	0	0
{x1, x3}	1	0	1
{x2, x3}	0	0	1
{x1, x2, x3}	1	1	1

The following theorem indicates that under normality there exists some relationship between super-regularity and irregularity.

Theorem: Every normal super-regular type II compatibility relation is irregular.

Proof: Assume R is a normal super-regular relation. Let $T_a \rightarrow W_a$, $T_b \rightarrow W_b$ and $T \rightarrow W$ be a super-regular triple, ie. $T = T_a \cup T_a$ and $W_a \cup W_b \subset W$. Let $y^* \in W$ but $y^* \notin W_a$ and $y^* \notin W_b$. Because of normality there exists some $x^* \in T$ such that $\{x^*\} \rightarrow Z^*$ and $y^* \in Z^*$. Furthermore since $T = T_a \cup T_b$, x* must be contained in at least one of the two sets. Without loss of generality assume $x^* \in T_a$. Consider the triple $\{x^*\} \rightarrow Z^*$, $T_a \rightarrow W_a$ and $T_a \rightarrow W_a$. We note here $\{x^*\} \cup T_a = T_a$. Let $W_a \cup Z^* = W$ since $y^* \in Z^*$ but $y^* \notin W_a$ then $W_a \subset W$ thus $W_a \subset W_a \cup Z^*$ and R is irregular.

Thus if $\mathfrak{R}$ is the set of all normal type II compatibility relations it can be broken up into two mutually exclusive classes $\mathfrak{R}_1$ and $\mathfrak{R}_2$ such that $\mathfrak{R} = \mathfrak{R}_1 \cup \mathfrak{R}_2$ where $\mathfrak{R}_1$ are the regular relations and $\mathfrak{R}_2$ are the irregular relations. We note that while all super-regular relations must be irregular, there exist irregular relations that are not super-regular. We shall call an irregular relation that is not super regular strictly irregular.

Based upon the definition of regular, we observe that a regular type compatibility relation is completely characterized by the associated sets of the principle elements.

Observation: Assume R is a regular type two compatibility relation on $X \times Y$. Let T_i be the primary elements of X and W_i their associated sets. If T is any other element of X the W associated with T is the union of the W_i's such that $x_i \in T$.

5. Nonmonotonicity and type II Compatibility Relations

We shall now describe the reasoning process used when we have a type II compatibility relation. Let V and U be variables taking value in X and Y. Assume our information about V is given in terms of a D-S belief structure m on X where A_i are the focal elements and $m(A_i) = a_i$. Assume our knowledge of the relationship between V and U is given by a type II compatibility relation $R: X \times Y \rightarrow [0, 1]$. We are 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 $\widehat{m}$ on X where $\widehat{m}$ has focal elements $\widehat{A}_i \in$ X where $\widehat{A}_i = \{A_i\}$ and $\widehat{m}(\widehat{A}_i) = m(A_i) = a_i$. (The focal elements of $\widehat{m}$ are singleton subsets of X whose element is a focal element in X, if $A_i = \{x_1, x_2\}$ then $\widehat{A}_i = \{\{x_1, x_2\}\}$).

1. Extend $\widehat{m}$ to be a bpa on $X \times Y$ where $\widehat{A}_i \times Y$ are the focal elements and $\widehat{m}(\widehat{A}_i \times Y) = a_i$

2. Conjunct R and $\widehat{m}$. This results in a bpa m^+ on $X \times Y$ with focal elements $E_i = (\widehat{A}_i \times Y) \cap R$ and $m^+(E_i) = a_i$

3. Project m^+ onto Y. This results in the bpa m* on Y with focal elements $F_i = Proj_Y(E_i)$, that is $F_i(y) = Max_{T \in X}[E_i(T, y)]$, and where $m^*(F_i) = m(A_i)$.

We shall denote the process of inferring m* from R and m as $m^* = R \otimes m$. We should note that the following simple expression of this inference process can be obtained. Assume m is a bpa on X focal elements A_i. Let R be a type II compatibility relation such that W_{A_i} is the set associated with A_i, $A_i \rightarrow W_{A_i}$. Then if $m^* = R \otimes m$, the focal elements of m* are W_{A_i} and $m^*(W_{A_i}) = m(A_i)$.

Definition: We shall say that a type I compatibility relation C is **equivalent** to a type II compatibility relation R, denoted $R \Leftrightarrow C$, if for every knowledge structure m about V their inferred values for V are the same, for all m we have $R \otimes m = C \otimes m$.

Definition: Assume C is a type I compatibility relation on $X \times Y$, we call R its **regular extension** to type II if R is a type II relation defined as follows

(1) For any primary element, $T_i = \{x_i\}$, $R(T_i, y) = C(x_i, y)$, that is $W_i = \{y \mid C(x_i, y) = 1\}$, the set associated with T_i under R is the same as that associated with x_i under C.

(2) For any non-primary element T of X, its associated set W is the union of the W_i's for $x_i \in T$.

Observation: If R is the regular extension of C then $R \Leftrightarrow C$.

It can be easily shown that every regular type II compatibility relation is an extension of a unique type I relation. More specifically there exists a one to one correspondence between type II regular relations and type I compatibility relations obtained by the regular extension. Thus to each type I compatibility relation there exists a unique regular type II relation, its regular extension, and each regular type II relation is unique to one type I relation. We shall call a pair C and its regular extension R, a type I-II representation pair. Since we have previously shown that the two elements in a representation pair C-R are equivalent as far as inferences and we have also indicated that every type I relation is monotonic, we obtain the following theorem.

Theorem: Every regular type II relation is monotonic.

As we have previously indicated the set of all normal (rational) type II compatibility relations $\mathfrak{R} = \mathfrak{R}_1 \cup \mathfrak{R}_2$.where $\mathfrak{R}_1$ is the set of regular relations and $\mathfrak{R}_2$ is the set of irregular relations and where $\mathfrak{R}_1$ and $\mathfrak{R}_2$ are disjoint. We have just shown that all relations in $\mathfrak{R}_1$ are monotonic. We will now proceed to show that all the relations in $\mathfrak{R}_2$ are nonmonotonic, all irregular relations are nonmonotonic.

We recall that a relation R is monotonic if for every m_1 and m_2 on X such that $m_1 \subset m_2$ we have $m_1^* \subset m_2^*$ where $m_1^* = R \otimes m_1$ and $m_2^* = R \otimes m_2$ We say that R is non-monotonic if there exists a pair m_1 and m_2 on X where $m_1 \subset m_2$ such that $R \otimes m_1 \not\subset R \otimes m_2$. We now proceed to show that all irregular are non-monotonic.

Theorem: All irregular relations are nonmonotonic.

Proof: Assume R is an irregular relation. This implies that there exists at least three elements T_a, T_b and T_c of X such that $T_a \cup T_b = T_c$ and $W_c \subset W_a \cup W_b$, that is there exists one y* such that $y^* \in W_a$ or $y^* \in W_b$ but $y^* \notin W$. Consider now the two bpa m_1 and m_2 on X such that m_1 is defined by $m_1(T_a) = a$ and $m_1(T_b) = 1 - a$ where $a > 0$ and where m_2 is defined by $m_2(T_c) = 1$. It is obvious that $m_1 \subset m_2$, since $T_a \subset T_c$ and $T_b \subset T_c$. If $m_1^* = R \otimes m_1$ then $m_1^*(W_a) = a$ and $m_1^*(W_b) = 1 - a$. If $m_2^* = R \otimes m_2$ then $m_2^*(W_c) = 1$. Consider the subset $D = \{y^*\}$ of Y,

$$Pl_1(D) = a \ Poss[D/W_a] + (1 - a) \ Poss[D/W_b]$$

$$Pl_2(D) = Poss[D/W_c].$$

Since $y^* \notin W_c$ then $Poss[D/W_c] = 0$ and $Pl_2(D) = 0$. Since y is in at least one of W_a or W_b, without loss of generality, assume it is definitely in W_a then $Poss[D/W_a] = 1$ 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 results indicates that not only are all irregular relations non-monotonic but the only way to represent a non-monotonic compatibility relation in terms of normal relations is via an irregular relation.

We see that the non-monotonicity resides in triples T_a, T_b, and T_c such that $T_c = T_a \cup T_b$ and where $W_c \subset W_a \cup W_b$. We shall call these non-monotonic triples. Assume R is an irregular relation. Let $\{S_1, ..., S_n\}$ be the set of all elements in X which participate in non-monotonic triples. Let $Card(S_i)$ be the number of elements in S_i and let $L_R = Min_i\ Card(S_i)$. We shall call L_R the nonmonotonicity level of R. If $L_R = 1$ we called it a primary non-monotonicity.

Let m be a bpa on X with focal elements A_j.. We let $L_m = Min_j[Card(A_j)]$ and call it the imprecision level of m. Assume m_1 and m_2 are bpa on X such that $m_1 \subset m_2$, it follows from the definition of containment that $m_1 \subset m_2 \rightarrow L_{m_1} \leq L_{m_2}$.

Observation: Assume R is an irregular relation with level of non-monotonicity L_R. Let $T_i = \{x_i\}$ be the primary elements of R. Let T be a subset of X such that $Card(T) \leq L_R$. Then the associated W set of T satisfies $W = \bigcup_{i \text{ s.t. } x_i \in T} W_i$.

In essence R appears regular for all T such that $Card(T) \leq L_R$.

Theorem: Assume R is an irregular relation with non-monotonicity level L_R. Let m_2 be a bpa on X such that $L_{m_2} \leq L_R$. Let m_1 be another bpa on X such that $m_1 \subset m_2$. If $m_1^* = R \otimes m_1$ and $m_2^* = R \otimes m_2$, that $m_1^* \subset m_2^*$.

Proof: If $\{A_k\}$ are the focal elements of m_1 and B_j are the focal elements of m_2 then $Card(A_k) \leq L_R$ and Card $B_j \leq L_R$. In this situation all our operations are performed within the monotonic or regular range of R and hence we get monotonicity property associated with regular relations. For a given R and m, if $L_m \leq L_R$ we shall say that m is the regular range or under the non-monotonicity of R.

We recall that if we have some information about V in terms of a bpa m_1 and if we get another piece of information m_2, this results in an effective bpa $m_3 = m_1 \cap m_2$ where $m_3 \subset m_1$. Using this we can now see an important implication of the previous result. If we reached a state of our knowledge about V which is under the non-monotonicity level of R then the gaining of more information about V will allow us to move monotonically along in our knowledge about U.

6. Dispositional and Hierarchical Compatibility Relations

A important class of non-monotonic compatibility relations are those representing default rules [5]. In this section we discuss some classes on nonmonotonic compatibility relations in the spirit of default rules. Consider an ordinary rule such as if a then b, it has an antecedent and consequent. Three conditions can be in effect regarding this rule, we know the antecedent to be true, we know it to be false or we don't know whether it is true or false. If we know the antecedent is satisfied then we infer the consequent. If the antecedent is false then we can't infer anything, similarly if the truth of the antecedent is unknown we don't infer consequent. In the case of a what we shall call a **dispositional** rule we act differently, in the case in which the truth of the antecedent is unknown, we infer the consequent. In essence in these dispositional rules if the antecedent is possible, not known to be unsatisfied, then we infer the consequent. These rules are useful in that they allow us to act in situations in which we don't have all the information.

In the following we shall introduce a general structure for these dispositional relations. Let V and U be variables taking their values in X and Y. Let A be a subset of X, the antecedent . Let B_1 and B_2 be subsets of Y where $B_2 \not\subset B_1$. Let R be a type II compatibility relationship between V and U, $R: X \rightarrow Y$. In the case of what we call as dispositional rules the relation R must satisfy the following conditions. For any $T \in X$ and any bpa m_T having the single focal element T it must be the case that

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

Thus we see that R is a function of Poss[A/T]. We note that $Poss[A/T] = 1$ if $A \cap B \neq \emptyset$ and $Poss[A/T] = 0$ otherwise. It should be emphasized that the structure of R is such for any $T \in X$ its associated value, $W = B_1$ if $Poss[A/T] = 1$ and $W = B_2$ if $Poss[A/T] = 0$. We shall indicate this as $R = D(A, B_1, B_2)$ and call it a standard dispositional rule. An important special case occurs when $B_2 = Y$.

Theorem: A standard dispositional rule is a nonmonotonic compatibility relation.

Proof: 1. Let T be a subset of X which is contained neither in A or $\bar{A}$ then $T = T_a \cup T_b$ where $T_a \subset A$ and $T_b \subset \bar{A}$. If $T_a \subset A$ then Poss[A/T_a] = 1 and $T_a \to W_a = B_1$ and if $T_b \subset \bar{A}$ then Poss[A/T_b] = 0 and $T_b \to\supseteq W_b = B_2$. Since $A \cap T \neq \emptyset$ then $T \to W = B_1$. Since $W_a \cup W_b = B_1 \cup B_2 \supset B_1 = W$ then the R is irregular. Since irregularity imposes nonmonotonicity these standard dispositional rules are nonmonotonic.

Theorem: A standard dispositional rule D(A, B_1, B_2) has a primary non-monotonicity level.

Proof: Let $x_1 \in A$ and $x_2 \in \bar{A}$ then with $T_1 = \{x_1\}$, $T_2 = \{x_2\}$ and $T = \{x_1, x_2\}$, $T_1 \to B_1$, $T_2 \to B_2$ and $T \to B_1$ we have primary nonmonotonicity

Assume the relationship between V and U is expressed by R = D(A, B_1, B_2) and our knowledge of V is modeled by the bpa m with focal elements A_i, $m(A_i) = a_i$. If $m^* = R \otimes m$ then m^* has focal elements W_i, the associated sets of the A_i and $m^*(W_i) = a_i$. Since $W_i = B_1$ if Poss[A/A_i] = 1 and $W_i = B_2$ if Poss[A/A_i] = 0 thus m* is obtained as

$$m^*(B_1) = \sum_{i=1}^{r} \text{Poss}[A_i/A]\ m(A_i)] = \text{Pl}(A)$$

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

An useful type of rule is what we shall call a hierarchical dispositional rule. Assume A_i, i = 1 to k, are a collection of exclusive and exhaustive subsets of X, $A_i \cap A_j = \emptyset$ and $\bigcup_{i=1}^{k} A_i = X$. Let B_i, i = 1 to k, be subsets of Y such that for i > j, $B_i \not\subset B_j$. We call R: X $\to$ Y a hierarchical dispositional relation if for any $T \in X$

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

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

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

....

....

$R \otimes m_T = B_k$ if Poss[A_j/T] = 0, j = 1 to k – 1 and Poss[A_k/T] = 1.

We shall denote such a relation as H(A_1, A_2, . . . , A_k: B_1, ... , B_k). It can be show that these relations are nonmonotonic relations. We note D(A, B_1, B_2) = H(A, $\bar{A}$: B_1, B_2).

A prototypical imperative which generates a hierarchical compatibility relation is a rule "always fix the easiest thing". Consider a device made up of three parts q_1, q_2, q_3. Let x_i indicate the proposition "part q_i is busted" and let y_i indicate the action "replace part i". Furthermore assume the parts are such that $q_1 < q_2 < q_3$ where a < b mean b is more difficult to fix then a. Then the hierarchical relation R = H(A_1, A_2, A_3: B_1, B_2, B_3) with $A_i = \{x_i\}$ and $B_i = \{y_i\}$implements this rule (see matrix following)

	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

Finally if H(A_1, A_2, . . . , A_k: B_1, B_2, ...,B_k) is a compatibility relation between V and U and our knowledge about V is the bpa m having q focal elements F_i. If $m^* = R \otimes m$ then

$$m^*(B_j) = \sum_{i=1}^{q} \text{Poss}[A_j/F_i] * \text{Cert}[\bar{G}_{j-1}/F_i] * m(F_i)$$

where $G_0 = \emptyset$ and for j > 1 $G_j = \bigcup_{i=1}^{j} A_i$.

7. References

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

[2] Shafer, G. A Mathematical Theory of Evidence, Princeton University Press: Princeton, 1976.

[3]. Shafer, G., "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.

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

[5] Reiter, R., "A logic for default reasoning", Artificial Intelligence 13, 81–132, 1980.