TOLERANT FUZZY PATTERN MATCHING: AN INTRODUCTION

DIDIER DUBOIS and HENRI PRADE

Institut de Recherche en Informatique de Toulouse, Université Paul Sabatier,
118 route de Narbonne, 31062 Toulouse Cedex, France
Email: {dubois, prade}@irit.irit.fr

ABSTRACT: The fuzzy pattern matching technique has been developed in the framework of fuzzy set and possibility theory in order to take into account the imprecision and the uncertainty pervading values which have to be compared to requirements (which may be fuzzy) in a pattern matching process. This paper restates the basic principles and extends them to situations where (sub)patterns are only required to be satisfied up to a given tolerance (which may be fuzzy), or where the different subparts of a compound pattern may have various levels of importance. Both cases correspond to a relaxation of elementary patterns, which can be expressed by a fuzzy relation modelling an approximate equality or an uncertain strict equality respectively. We also study the more sophisticated case where some elementary patterns have not to be satisfied with the highest priority provided that weaker requirements remain satisfied. The fuzzy pattern matching technique applies in a variety of problems including the evaluation of soft queries with respect to a fuzzy database, the evaluation of the fuzzy condition parts of rules in approximate reasoning, or the evaluation of the membership of an ill-known object to a flexible class in classification problems.

KEY WORDS: Pattern matching, uncertainty, imprecision, possibility theory, tolerance, importance.

1. INTRODUCTION

Basically, a pattern matching process involves a pattern describing some requirements and a data base where items, that is, pieces of information about objects are gathered. In the simplest case, data which are identical to the specification expressed in the pattern are looked for and retrieved. Flexibilities can be introduced for instance in order to limit the identity requirement to components of the items or in order to check that a condition expressed in the pattern is satisfied (e.g. that a component of an item is an element of a set specified in the pattern). However in these cases, the pattern matching remains an all-or-nothing process. Indeed an item does or does not match the requirements of the pattern.

When data are pervaded with imprecision and uncertainty or when the pattern includes vague specifications, the matching between the pattern and an item becomes naturally a matter of degree. In that case a fuzzy requirement may be more or less satisfied, or we may be not completely sure that a crisp requirement holds if the item under consideration is fuzzy. Cayrol, Farreny and Prade [2], [3] have developed a fuzzy pattern matching technique which is based on [0,1]-valued possibility and necessity measures (Zadeh [19], Dubois and Prade [4]) in order to estimate the compatibility between an item and what is required by the pattern. In this approach each component of an item and each elementary requirement are respectively associated with a possibility distribution and a fuzzy set. A possibility distribution is a convenient tool for the representation of imprecise, uncertain or vague pieces of data pertaining to single-valued attributes, while a fuzzy set models a soft constraint whose satisfaction is a matter of degree.

The next section recalls the basic principles of the fuzzy pattern matching technique. Then, in Section 3 the model is extended in order to account for tolerance relations and levels of importance which can be used to relax the requirements expressed by the elementary patterns. We also deal with the case where the levels of importance of requirements are not constant over the attribute domains but also guarantee the satisfaction of weaker requisites. The concluding remarks briefly survey potential applications.

2. FUZZY PATTERN MATCHING

Fuzzy pattern matching takes into account the semantic representation attached to the components of the data and the patterns and provides an evaluation of the similarity between a pattern and an item. This evaluation is graded in the interval [0,1]. In this paper a pattern is supposed to have an and/or tree-like representation whose leaves correspond to elementary components called atoms. For simplicity we assume that these atoms are not variables. The reader is referred to Cayrol, Farreny and Prade [3] for the treatment of variables and control symbols in fuzzy pattern matching strategies.

2.1. *General Principle*

The basic idea is to attach to each atom of a pattern the membership function of a fuzzy set restricting the values which are more or less compatible with the meaning of the atom. These values belong to some prescribed scale corresponding to the range of the attribute which the atom refers to. For instance the symbol *tall* refers to a scale of heights, and corresponds to a fuzzy subset of this scale. In place of a numerical scale we may have a discrete set of typical elements as well. The data are represented by lists whose components are associated with possibility distributions which model the imprecision pervading the data. Such lists contain possibly ill-known attribute values pertaining to the description of objects. Namely, a component in a list refers to only one (ill-located) element of the scale or domain of the concerned attribute.

The basic dissymmetry of the pattern-data matching is preserved by this modeling convention. Indeed, a fuzzy pattern represents an ill-bounded class of objects, while a fuzzy item represents an ill-known object whose precise description is not available. Namely let P and D be respectively a pattern atom and an item component pertaining to the same single-valued attribute, which are to be compared. P and D refer to the same scale U conveying their meanings. Let μ_P be the membership function associated to atom P and π_D the possibility distribution attached to D. Both are mappings from U to [0,1]. Let u be an element of U. Then $\mu_P(u)$ is the grade of compatibility between the value u and the meaning of P. Namely $\mu_P(u) = 1$ means total compatibility with P and $\mu_P(u) = 0$ means total incompatibility with P. By contrast, $\pi_D(u)$ is the grade of possibility that u is the (unique) value of the attribute describing the object modelled by the item. D is a fuzzy set of *possible* values (only one of which is the genuine value of the ill-known attribute), while P is a fuzzy set of *more or less* compatible values. For instance $\pi_D(u) = 1$ means that u is totally possible, while $\pi_D(u) = 0$ means that u is totally impossible as an attribute value of the object to which the item pertains. In the following μ_P and π_D are always supposed to be normalized, i.e. there is always a value which is totally compatible with P, and a value totally possible in the range D.

Two scalar measures are used in order to estimate the compatibility between a pattern atom P and its counterpart D in the item list, namely a degree of possibility of matching $\Pi(P ; D)$ and a degree of necessity of matching $N(P ; D)$ which are respectively defined by (see Zadeh [19], and Dubois and Prade [4]):

$$\Pi(P ; D) = \sup_{u \in U} \min(\mu_P(u), \pi_D(u)), \tag{1}$$
$$N(P ; D) = \inf_{u \in U} \max(\mu_P(u), 1 - \pi_D(u)). \tag{2}$$

The limiting cases where $\Pi(P ; D)$ and $N(P ; D)$ take values 0 and 1 are useful to study in order to lay bare the semantics of these indices. For any fuzzy set, F on U, let $F^\circ = \{u \in U \mid \mu_F(u) = 1\}$ be the core of F, and $s(F) = \{u \in U, \mu_F(u) > 0\}$ its support. Then it can be checked that

(i) $\Pi(P ; D) = 0$ if and only if $s(P) \cap s(D) = \emptyset$,
(ii) $\Pi(P ; D) = 1$ if and only if $P^\circ \cap D^\circ \neq \emptyset$,
(iii) $N(P ; D) = 1$ if and only if $s(D) \subseteq P^\circ$
(iv) $N(P ; D) > 0$ if and only if $D^\circ \subset s(P)$ (strict inclusion)

The measure $\Pi(P ; D)$ estimates to what extent it is possible that P and D refer to the same value u; in other words, $\Pi(P ; D)$ is a degree of overlapping of the fuzzy set of values compatible with P, with the fuzzy set of possible values of D. The measure $N(P ; D)$ estimates to what extent it is necessary (i.e. certain) that the value to which D refers is among the ones compatible with P; in other words $N(P ; D)$ is a degree of inclusion of the set of possible values of D into the set of values compatible with P. However (2) is based on a stronger view of inclusion between fuzzy sets than the usual one ($\pi_D \leq \mu_P$); the latter only implies N(P ; D) ≥ 0.5, while here, N(P ; D) = 1 if and only if all the more or less possible values restricted by π_D are compatible with P to degree 1. The duality of possibility and necessity, i.e. the necessity of an event corresponds to the impossibility of the opposite event, is expressed here by the relation

$$N(P ; D) = 1 - \prod(\bar{P} ; D) \tag{3}$$

where $\mu_{\bar{P}} = 1 - \mu_P$ is the membership function of the complement of the fuzzy set of values compatible with P. Clearly we always have

$$\prod(P ; D) \geq N(P ; D).$$

But the property max(∏(P ; D), 1 – N(P ; D)) which holds if P is not fuzzy, is no longer valid, generally.

Besides it is worth noticing that N(F ; F) = 1 if and only if μ_F is the membership function of an ordinary subset of U which stands for μ_P and π_D. Indeed, in this case we are completely certain that an ill-known value which is only known to belong to D = F satisfies the requirement, expressed by the pattern P, of belonging to F, whatever its precise value. Otherwise we only have N(F ; F) ≥ 1/2. Indeed when P = D have a fuzzy meaning, we cannot be sure that the degree of matching $\mu_P(x)$ of the component item x fuzzily restricted by $\pi_D = \mu_P$, is high. The indices N(P ; D) and ∏(P ; D) do not intend to estimate to what extent the functions μ_P and π_D are equal, but to what extent an attribute value whose more or less possible values are described by π_D belongs for sure to P with a high degree. In any case we have N(s(F) ; F) = 1 since an ill-known value restricted by $\pi_D = \mu_F$ is not allowed to have its precise value outside the support s(F) of F. Note that when D is precise, i.e. ∃d, $\pi_D(d) = 1$, and ∀ u ≠ d, $\pi_D(u) = 0$ which can be written D = {d} then

$$\prod(P ; \{d\}) = N(P ; \{d\}) = \mu_P(d). \tag{4}$$

Table 1 summarizes the respective locations of ∏(P ; D) and N(P ; D) in [0,1] when P and/or D is precise, imprecise (but clear-cut) or fuzzy. Note also that

$$\prod(P \cup P' ; D) = \max(\prod(P ; D), \prod(P' ; D)); \tag{5}$$
$$N(P \cap P' ; D) = \min(N(P ; D), N(P' ; D)) \tag{6}$$

where fuzzy set union and intersection are defined by means of max and min operations respectively. This is the characteristic properties of possibility and necessity measures [19], [4]. But in general we only have

$$\prod(P \cap P' ; D) \leq \min(\prod(P ; D), \prod(P' ; D)); \tag{7}$$
$$N(P \cup P' ; D) \geq \max(N(P ; D), N(P' ; D)). \tag{8}$$

Table 1

P	D		
	Precise $D = \{d\}$	Imprecise (non-fuzzy)	Fuzzy
Precise $P = \{p\}$	$\Pi = N = \begin{cases} 1 \text{ if } p = d \\ 0 \text{ if } p \neq d \end{cases}$	$N = 0$ $\Pi = \begin{cases} 1 \text{ if } p \in D \\ 0 \text{ otherwise} \end{cases}$	$N < 1$ $N = 0$ if $\exists d \neq p, \pi_D(d) = 1$ or π_D continuous $\Pi = \pi_D(p)$
Imprecise (non-fuzzy)	$\Pi = N = \begin{cases} 1 \text{ if } d \in P \\ 0 \text{ if } d \notin P \end{cases}$	$N = \Pi = \begin{cases} 1 \text{ if } D \subseteq P \\ 0 \text{ if } D \cap P = \emptyset \end{cases}$	$N > 0 \Rightarrow \Pi = 1$
Fuzzy	$\Pi = N = \mu_P(d)$	$\Pi \geq N$	$\Pi \geq N$

2.2. *The Fuzzy Pattern Matching Procedure (Outline)*

2.2.1. Knowledge Encoding and Elementary Matching Computation

Fuzzy pattern matching, i.e. formulas (1) and (2), apply to fuzzy sets and possibility distributions defined on any kind of universes U, finite or not, ordered or not. When U is a part of the real line, it is convenient to use parametrized representations, like trapezoids. For practical applications, trapezoidal membership functions are generally sufficient; indeed, slight modifications of the shape of a membership function do not affect the evaluation of results significantly. A trapezoidal function is represented by a 4-tuple (a, b, c, d) of parameters whose meanings are explained in Figure 1. c and d are non-negative.

We may have a = b, c = 0, or d = 0. Thus, precise as well as imprecise values can be represented by 4-tuples. Let P and D be two fuzzy values whose membership

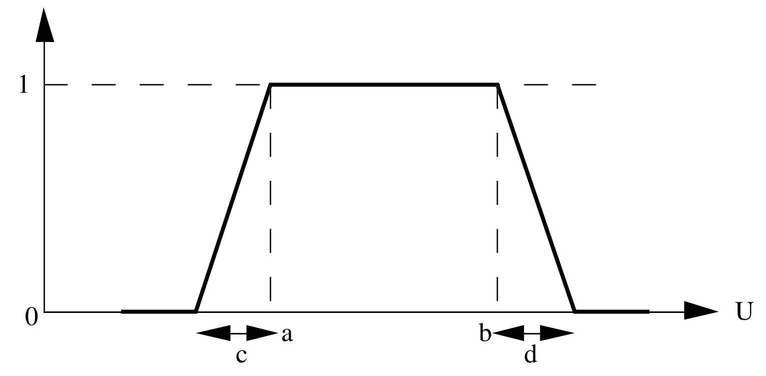

Figure 1

functions are trapezoidal and represented by the 4-tuples (a_1, b_1, c_1, d_1) and (a_2, b_2, c_2, d_2) respectively. The possibility and necessity measures (formulae (1) and (2)) are directly computed from practical representations of values, as intersection of straight lines. See Figures 2 and 3.

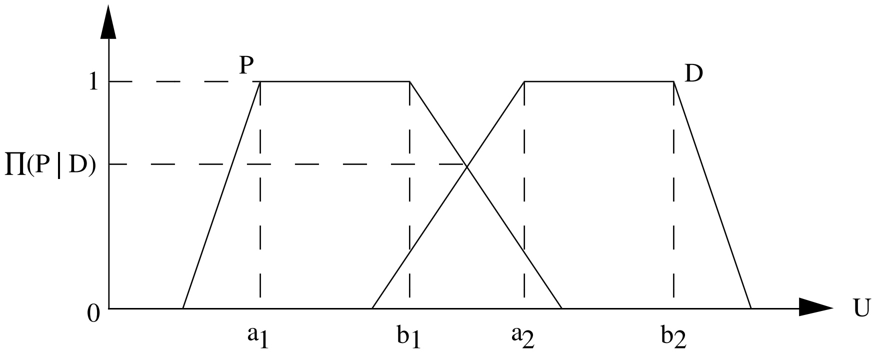

Figure 2

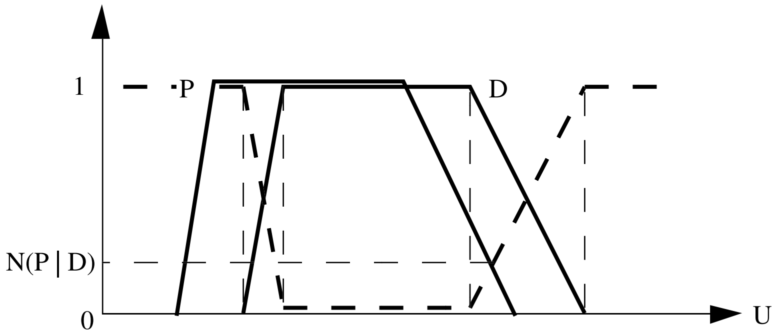

Figure 3

Note that the typically fuzzy situation whereby $0 < N(P ; D) \leq \prod(P ; D) < 1$ occurs provided that both $P° \cap D° = \emptyset$ and $D° \subset s(P)$ hold. This means that the core of D completely lies in the borderline area of P, that is, if trapezoidal representations are used, the core of D entirely lies on one side or another of P°, the core of P.

2.2.2. Global Matching Evaluation

The atomic measures of possibility and of necessity are aggregated separately in order to obtain two global measures between the whole pattern and the whole item. When the pattern expresses a conjunction of elementary requirements "P_1 and… P_n", this aggregation is performed using the 'min' operation and preserves the respective semantics of the measures in terms of possibility and necessity. Indeed we have (for instance, see Dubois and Prade [4]):

$$\prod(P_1 \times \ldots \times P_n \,;\, D_1 \times \ldots \times D_n) = \min_{i=1,\ldots,n} \prod(P_i \,;\, D_i), \quad (9)$$

$$N(P_1 \times \ldots \times P_n \,;\, D_1 \times \ldots \times D_n) = \min_{i=1,\ldots,n} N(P_i \,;\, D_i), \quad (10)$$

where P_i and D_i are supposed to be defined on the same domain U_i, and where $\times$ denotes the Cartesian product defined for two fuzzy sets F_i and F_j by

$$\forall\, u_i \in U_i, \forall\, u_j \in U_j, \mu_{F_i \times F_j}(u_i,u_j) = \min(\mu_{F_i}(u_i), \mu_{F_j}(u_j)). \quad (11)$$

The attributes appearing in the pattern and the item are supposed to be non-interactive (Zadeh [19]). This means that a restriction on the possible values of item D_i expressed by π_{D_i} has absolutely no influence on the possibility distribution π_{D_j} restricting the possible values of the item D_j; there is no constraint linking the values of D_i and D_j. Note that (11) supposes that all the parts of the pattern (which expresses what is required) have an equal importance from the user's point of view. Disjunctive requirements "P_1 or … P_n" are dealt with using formulae similar to (9)-(10) changing min into max, and $P_1 \times \ldots \times P_n$ into $P_1 + \ldots + P_n$ with $\mu_{P_i + P_j}(u_i,u_j) = \max(\mu_{P_i}(u_i), \mu_{P_j}(u_j))$.

2.2.3. Ranking of Best-Matching Data

Usually, a single pattern, expressing some requirement, is matched against a whole database, i.e. a collection of data. From a user's point of view, it is interesting to get the data which yield the best compatibility measures with respect to the pattern. Let us describe the behaviour of our fuzzy pattern matching-based system in that case. Several situations may happen (see Table 1): i) no item matches the pattern; the possibility (and necessity) measures are zero. ii) One or several items match the pattern with positive possibility or necessity degrees. In that case, we are faced to the problem of ranking the items, each of them being characterized by a pair of degrees. In order to establish a classification of several items with respect to a pattern, we may take the following remarks into account:

- The necessity degree is of greater importance than the possibility degree, because when the necessity degree is positive, we are (more or less) certain that the item matches the requirement.
- If we obtain a necessity degree equal to zero and a possibility degree equal to 1 for each item, it means that what is known about the data is too imprecise or vague with respect to the pattern. We may then try to improve the knowledge about the data.
- The more precise D, the closer to zero $\prod(P \,;\, D) - N(P \,;\, D)$ is. However the converse is false (since $\prod(P \,;\, D) = N(P \,;\, D)$ for non-fuzzy subsets $P = D$).
- Generally, the Pareto ordering is used to rank pairs of degrees, according to the following way: $(\prod,N)$ is greater than $(\prod',N')$ if and only if $\prod > \prod'$ and $N \geq N'$, or $\prod \geq \prod'$ and $N > N'$. However, there are situations where $\prod(P \,;\, D) > \prod(P \,;\, D')$ and $N(P \,;\, D) < N(P \,;\, D')$. Here P and D stand for compound pattern or item. The

Pareto ordering is only partial. Thus we may obtain *several* data which somewhat best match the pattern. It is worth noticing that if D and D' are maximal elements for the Pareto ordering, we have

$$N(P ; D) < N(P ; D') < \Pi(P ; D') < \Pi(P ; D)$$

which indicates that D' is the best item from the point of view of certainty of matching. Note that we may have the above incomparability situation between D and D' while nothing can be said about the respective precision of D and D'. Namely if the four indices strictly lie in the unit interval, the two differences $\Pi(P ; D) - N(P ; D)$ and $\Pi(P ; D') - N(P ; D')$ can be totally independent of one another if D and D' lie on different sides of the core of P (see Figure 4). On the whole it sounds reasonable to rank-order the items in terms of N(P ; D) and to use $\Pi(P ; D)$ in order to break ties.

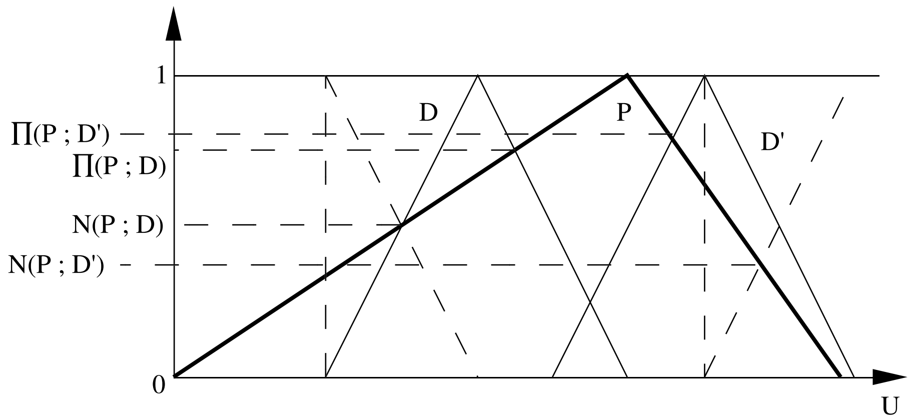

Figure 4

Providing a weighted set of retrieved data may not be informative enough for a user. It can be useful to also explain the reasons for uncertainty if any. For instance explaining that the certainty of matching for an item D such that $\Pi(P ; D) > 0$ could be improved by adding more information about the value of some attribute i, i.e. restricting this value to a strict subset D'_i of D_i, so that $N(P_i ; D'_i) > 0$.

2.3. *Example*

This example (Dubois, Prade and Testemale [6]) deals with (more or less approximate) descriptions of second-hand cars for sale in a garage. The different involved attributes are the age of the car, the purchase price, the gas consumption and the speed of the car. The available data can be represented in a so-called relational table (Table 2). Each row in the table corresponds to a list of pairs (<(fuzzy) value> <name of attribute>); for instance the first row corresponds to the item

V1: ((new age) (expensive price) (economical consumption) (rather-fast speed)).

A domain is attached to each attribute; this domain is the set of all the values which may be taken by the attribute. Here for instance the domain attached to the attribute 'age-of-the-car' is the interval [0,20]. All the considered domains in this example are continuous scales.

Table 2

	Age of the car (years)	Purchase price (FF)	Consumption (liter/100 km)	Maximum Speed (km/h)
V1	new	expensive	economical	rather-fast
V2	less-than-3	around-45000	rather-economical	180-200
V3	years	between-50000-and-60000	heavy	fast
V4	very-recent	less-than-20000	8-9	180
V5	around-5 years	around-10000	heavy	rather-fast
V6	5-10	cheap	economical	not-very-fast
V7	old	32000-40000	very-economical	between-140-
	new			and-160

Let us now consider the following problem: a person who wants to buy a car looks for a car corresponding to various criteria. For instance, we may have the following request:

P: ((new age) AND (moderate price) AND
(very-economical consumption) AND (fast speed)).

The membership functions of the fuzzy values are given in Table 3, using the encoding of 2.2.1. The results obtained in our example with pattern P are

$$\Pi(P\ ;\ V2) = 0.3 \quad ; \quad \Pi(P\ ;\ V7) = 0.5,$$

and all other degrees are equal to zero.

Thus in this example only two cars, V2 and V7, are proposed to the buyer whose request is expressed by pattern P. Since the data are too imprecise or vague, there are no cars which more or less certainly meet the requirements. In such a case, we need more information about the cars if we want to have some certainty that the requirements are satisfied or, as we shall see in Section 3, we may also diminish the relative importance of some of the requirements in the pattern.

3. TOLERANCE AND IMPORTANCE ASSIGNMENT

In case of a precise pattern $P = \{p\}$ and a precise item $D = \{d\}$, we have pointed out in Table 1 that

$$\Pi(\{p\}\ ;\ \{d\}) = N(\{p\}\ ;\ \{d\}) = 0 \Leftrightarrow p \neq d.$$

Moreover in the more general case where P is a crisp pattern, we have $\Pi(P\ ;\ \{d\}) = N(P\ ;\ \{d\}) = 0$ as soon as $d \notin P$ even if d is "close" to some value in P. In this section we investigate two ways of weakening the requirement of a strict identity

Table 3

(1) Attribute: age-of-the-car; domain [0,20].				
new	(0	1	0	1)
very-recent	(1	2	1	1)
around-5	(5	5	1	1)
old	(10	15	2	5)
less-than-3	(0	3	0	0)
(2) Attribute: selling-price; domain: [5000, 100000].				
expensive	(80000	100000	5000	0)
around-45000	(44000	46000	1000	1000)
around-10000	(9000	11000	1000	1000)
cheap	(5000	10000	0	5000)
between-50000-and-60000	(50000	60000	5000	5000)
moderate	(35000	60000	5000	5000)
(3) Attribute: consumption; domain [5, 15].				
economical	(6	7	0.5	1)
rather-economical	(7	8	1	1)
very-economical	(5	6	0	0.5)
heavy	(9	15	0.5	0)
(4) Attribute: (maximum) speed; domain [100, 250].				
fast	(180	200	20	20)
rather-fast	(150	180	20	20)
not-very-fast	(120	140	10	10)
between-140-and-160	(140	160	20	20)

between the precise value of the item (precise value which may be ill-known) and a value compatible with the pattern P. These two ways amount to modify, in fact to enlarge, the pattern P into a pattern P*, such that $\mu_{P*} \geq \mu_P$. P* will be obtained by composing P with a fuzzy relation R which in one case models an approximate equality, i.e. an equality with some tolerance, and in the other case an "uncertain" equality, i.e. a strict equality which may not be completely important to satisfy.

3.1. *Modeling Tolerance*

Two values u_1 and u_2 belonging to a domain U may be considered as approximately equal even if they are not identical. For instance if the pattern requires somebody who is 40 years old, an item corresponding to a person who is 39 may be considered in some cases as approximately matching the request. An approximate equality can be conveniently modelled by means of a fuzzy relation R which is reflexive (i.e. $\forall$ u $\in$ U, $\mu_R(u,u) = 1$) and symmetrical (i.e. $\forall u_1 \in U$, $\forall u_2 \in U$, $\mu_R(u_1,u_2) = \mu_R(u_2,u_1)$). The closer u_1 and u_2 are, the closer to 1 $\mu_R(u_1,u_2)$ must be. The quantity $\mu_R(u_1,u_2)$ can be viewed as a grade of approximate equality of u_1 with u_2. R is then called a proximity or a tolerance relation. When the pattern is any subset

P of U (P may be fuzzy) but the item remains a (precise) constant d, the tolerance R can be taken into account in the degree of matching (which is equal to $\prod(P ; \{d\}) = N(P ; \{d\}) = \mu_P(d)$ if R is ignored) by replacing P by the enlarged subset $P \circ R$, defined by

$$\forall\ u \in U, \mu_{P \circ R}(u) = \sup_{u' \in U} \min(\mu_P(u'), \mu_R(u,u')). \tag{12}$$

The inequality $\forall\ u \in U, \mu_{P \circ R}(u) \geq \mu_P(u)$ expresses that $P \circ R$ contains P. Roughly speaking $P \circ R$ gathers the elements in P and the elements outside of P which are somewhat close to an element in P. In the general case, the matching of D with respect to P, taking into account R is graded by $\prod(P \circ R ; D)$ and by $N(P \circ R ; D)$. The use of tolerance relations in the fuzzy pattern matching was already suggested in Cayrol, Farreny and Prade [3]. Obviously, a tolerance relation is attached to the domain of an attribute variable, and different tolerance relations can be involved in the evaluation of the matching of an item with a compound pattern.

On a universe U which is a subset of the real line an approximate equality can be conveniently modelled by a fuzzy relation R of the form $\mu_R(x,y) = \mu_L(|x - y|)$, for instance

$$\forall x, \forall y, \mu_R(x,y) = \max\left(0, \min\left(1, \frac{\delta + \varepsilon - |x - y|}{\varepsilon}\right)\right) = \begin{cases} 1 \text{ if } |x - y| \leq \delta \\ 0 \text{ if } |x - y| > \delta + \varepsilon \\ (\delta + \varepsilon - |x - y|) / \varepsilon \text{ otherwise} \end{cases} \tag{13}$$

where δ and ε are respectively positive and strictly positive parameters which modulate the approximate equality. When μ_P has a trapezoidal shape and R is given by (13), $P \circ R$ is trapezoidal also and $P \circ R$ is very easy to compute. Namely if P is represented by (a, b, c, d) as in Figure 1, $P \circ R$ will be represented by $(a - \delta, b + \delta, c + \varepsilon, d + \varepsilon)$.

3.2. *Importance Assignment*

One possible way for taking into account the relative importance of elementary requirements in a compound pattern is again to use tolerance relations. More precisely, the elementary patterns P are all the more enlarged into $P \circ R$ (using a tolerance relation R) as they are considered as less important. In this sub-section another method using weights is proposed for modelling importance. This method corresponds to the use of a particular fuzzy relation R, still reflexive and symmetrical, but which now models the idea of an uncertain strict equality rather than an approximate equality (allowing for some tolerance when values are close). Namely R is now of the form

$$\mu_R(u_1,u_2) = \begin{cases} 1 \text{ if } u_1 = u_2 \\ 1 - w \text{ if } u_1 \neq u_2. \end{cases} \tag{14}$$

When w = 1, this is the usual strict equality and $P \circ R = P$. When w = 0, $P \circ R = U$ and the requirement expressed by the pattern will be trivially satisfied. The larger w, the more important the requirement expressed by the pattern. Indeed, the introduction of the weight w amounts to modifying the pattern P into $P^* = P \circ R$ such that

$$\mu_{P^*}(u) = \max(\mu_P(u), 1 - w) \tag{15}$$

As we can see, P* considers as acceptable any value outside of the support of P with the degree 1 – w. It means that the larger w, the smaller the degree of acceptability of a value outside of the support of P and the more important the requirement expressed by P.

In Section 2, we recalled that the atomic measures of possibility and of necessity must be aggregated separately using the min operation in case of a compound conjunctive pattern (see formulas (9), (10)). This implicitly supposes the equal importance of the requirements corresponding to the atomic patterns. Let us now combine the result of the pattern matching obtained with the P_i^* rather than with the P_i, using a symmetrical aggregation operation which is again the minimum operation, where each P^*_i is associated with a weight w_i in the sense of (15). Then, the following identities are valid and can be established by straightforward calculations:

$$\min_{i=1,\ldots,n} \Pi(P^*_i ; D_i) = \min_{i=1,\ldots,n} \max(1 - w_i, \Pi(P_i ; D_i)),$$
$$\min_{i=1,\ldots,n} N(P^*_i ; D_i) = \min_{i=1,\ldots,n} \max(1 - w_i, N(P_i ; D_i)). \tag{16}$$

Clearly when $w_i = 0$, the degrees of matching $N(P_i ; D_i)$ and $\Pi(P_i ; D_i)$ are ignored in the combination, then P_i has absolutely no importance; the larger w_i, the smaller the degrees of matching concerning P_i which are effectively taken into account in the aggregation.

In order to keep the results normalized, the following condition should be satisfied by the w_i:

$$\max_{i=1,n} w_i = 1.$$

It expresses that the most important requirement has the maximum weight.

Example (continued)

Let us now consider the following pattern which is a weighted version of the previously considered pattern

P*: ((new age) 0.8) AND ((moderate price) 0.5) AND
((very-economical consumption) 1) AND ((fast speed) 0.2).

This pattern expresses that the most important feature is a low gas consumption (weight equal to 1), and in a decreasing order of importance, the age, the price and the speed. The results obtained in the example of Section 2 with the new pattern P* are

$\Pi(P^* ; V1) = 0.5$, $\Pi(P^* ; V2) = 0.3$, $\Pi(P^* ; V6) = 0.2$,
$\Pi(P^* ; V7) = 0.8$ and $N(P^* ; V7) = 0.5$.

All others $\prod(P^* ; V_i)$ or $N(P^* ; V_j)$ are equal to zero.

Four cars are now retrieved (instead of 2 in the non-weighted case). Indeed in the weighted case, car V1 may be of interest to the buyer because he or she somehow neglects the purchase price. Moreover, vehicle V7 is strongly preferred in the weighted case ($N(P^* ; V7) = 0.5$). This is because the buyer does not really bother about the speed. This type of information, which helps in the discriminating process, cannot be expressed by means of the standard fuzzy pattern matching.

N.B.: Weighted disjunctive aggregation can be also defined (see Dubois, Prade and Testemale [6]) by the following combination laws which are dual to the ones in the conjunctive case:

$$\max_{i=1,\ldots,n} \min(w_i, \prod(P_i ; D_i)) \text{ and } \max_{i=1,\ldots,n} \min(w_i, N(P_i ; D_i)).$$

When all the w_i are equal to 1 we recover the ordinary disjunctive aggregation; when $w_i = 0$, the requirement P_i is ignored.

3.3. *Variable Weights and Rejection Sets*

In the above example, each weight of importance is a constant and thus does not depend upon the value taken by the concerned attribute for the considered object. This limitation may create some unnatural behaviour of the matching procedure. For instance, the price may be of a limited importance only within a certain range of values; when this price becomes very high, this criterion alone should cause the rejection of the considered car, in spite of the rather low importance weight.

To cope with this limitation we have proposed (Dubois, Prade and Testemale [6]) that the weight of importance become a function of the concerned attribute value. Formally, let s(P) be the support of the fuzzy set associated with the atom P. The weighting function w is supposed to be constant at least on s(P) and to increase for values outside of P, possibly reaching 1 outside s(P). See Figure 5. Let us show that the fuzzy pattern matching computations can still be easily handled in this case.

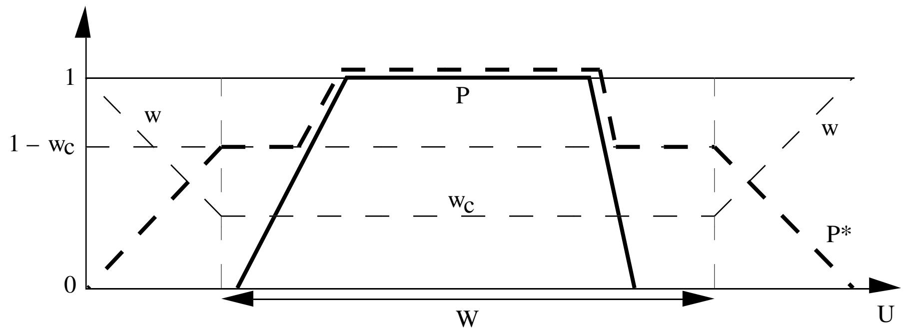

Figure 5

The weighted pattern P* is still defined by (15), i.e.

$$\mu_{P*}(u) = \max(\mu_P(u), 1 - w(u)).$$

Let $W \subseteq U$ be the subset where the weight w is constant, and equal to w_c (with $W \supseteq s(P)$). Let P' be the fuzzy set defined by

$$\mu_{P'}(u) = \begin{cases} 1 \text{ if } u \in W \\ 1 - w(u) \text{ if } u \notin W. \end{cases} \tag{17}$$

Then P* can be expressed in terms of P, P' and w_c in two different manners (see Figure 5), namely

$$\mu_{P*}(u) = \max(\mu_P(u), \min(1 - w_c, \mu_{P'}(u))) \tag{18}$$

and

$$\mu_{P*}(u) = \min(\max(\mu_P(u), 1 - w_c), \mu_{P'}(u)). \tag{19}$$

Applying (4)-(5) and (16), we obtain for $\Pi(P^* ; D)$ and $N(P^* ; D)$

$$\Pi(P^* ; D) = \max(\Pi(P ; D), \min(1 - w_c, \Pi(P' ; D))). \tag{20}$$

since $\sup_u \min(1 - w_c, \mu_{P'}(u), \pi_D(u)) = \min(1 - w_c, \Pi(P' ; D))$ and

$$\begin{aligned} N(P^* ; D) &= \min(\max(N(P ; D), 1 - w_c), N(P' ; D)) && (21) \\ &= \max[\min(N(P ; D), N(P' ; D)), \min(1 - w_c, N(P' ; D))] \\ &= \max(N(P ; D), \min(1 - w_c, N(P' ; D))) && (22) \end{aligned}$$

since $P \subseteq P'$. Note that formulas (18)-(19) and (20)-(22) still hold for slightly different definitions of P'. What is required for the membership function of P' is that

$$\begin{cases} \mu_{P'}(u) \geq \mu_P(u) & \text{if } \mu_P(u) \geq 1 - w_c \text{ (i.e. } u \in P_{1-w_c}) \\ \mu_{P'}(u) \geq 1 - w_c & \text{if } u \in W - P_{1-w_c} \\ \mu_{P'}(u) = 1 - w(u) \ (\leq 1 - w_c) \text{ if } u \notin W \end{cases}$$

since we should have $P \subseteq P'$, and only the part of $\mu_{P'}$ below $1 - w_c$ is taken into account effectively in the computation. Moreover if we apply (20) and (22) directly, we have still more freedom and we can keep $\mu_{P'}$ as a trapezoid (such that $\mu_{P'}(u) = 1 - w(u)$ if $u \notin W$) even if $P'^{\circ} \not\supseteq P^{\circ}$, when μ_P and π_D are trapezoidal themselves. Thus the handling of "variable" weights is not heavy from a computational point of view in practice. We basically have to handle a pair of patterns rather than a single one, and to weight one of the results. It can be easily seen that if the weight w is constant in (20) and (21) we recover the expressions of subsection 3.2 since then $N(P' ; D) = \Pi(P' ; D) = 1$ due to $P' = U$ using (17).

The notion of a variable weight as introducing rejection values can be interpreted in a slightly different manner. Namely P is the fuzzy subset of values that an item component must necessarily satisfy in order to ensure that the item is compatible with the requirement (compatibility set), the complement of P' can be viewed as the fuzzy set of values which make the considered item incompatible with the requirement, regardless of other elementary requirements (rejection set). The weight w_c acts on the influence of the rejection subset $\overline{P}'$. $w_c = 0$ suppresses the fuzzy set of compatible values P; only P' remain. On the contrary $w_c = 1$ inhibits the influence of P'.

When several attributes are involved, it can be established that the fuzzy sets P_i and P'_i can be separately used in the aggregation process. Assume $a_i \le b_i$ for i = 1,n. Then it is easy to see that

$$\min_{i=1,n} \max(a_i, \min(1 - w_{c_i}, b_i)) = \min(\min_i \max(a_i, 1 - w_{c_i}), \min_i b_i)$$

using distributivity of max over min.

Then if $a_i = \prod(P_i ; D_i)$, $b_i = \prod(P'_i ; D_i)$ we obtain, for non-interactive attributes

$$\prod(P^*_1 \times P^*_2 \times \ldots \times P^*_n ; D) = \min(\min_i \max(\prod(P_i ; D_i), 1 - w_{c_i}), \min_i \prod(P'_i ; D_i)) \quad (23)$$

where $D = D_1 \times D_2 \times \ldots \times D_n$, since (6) applies with equality. If $a_i = N(P_i ; D_i)$ and $b_i = N(P'_i ; D_i)$ we obtain, with (21),

$$N(P^*_1 \times P^*_2 \times \ldots \times P^*_n ; D) = \min(\min_i \max(N(P_i ; D_i), 1 - w_{c_i}), \min_i N(P'_i ; D_i)). \quad (24)$$

Hence (23) (respectively, (24)) expresses that an item D is possibly (respectively, certainly) compatible with n requirements if and only if 1) all important requirements are possibly (respectively, certainly) satisfied in the sense that the attribute values of the item possibly (respectively, certainly) lie in their compatibility sets P_i, 2) no attribute value of an item certainly (respectively, possibly) lies in the corresponding rejection set $\overline{P'_i}$ (since $\min_i \prod(P'_i ; D_i) = 1 - \max_i N(\overline{P}'_i ; D_i)$, for instance). Note that when compatibility sets are conjunctively combined, the rejection sets are disjunctively combined.

The idea of using compatibility and rejection sets and combining separately the information pertaining to each ot them separately has been independently suggested by Sanchez [15].

The rejection sets are also very much akin to the use of veto thresholds in relational approaches to multiple criteria aggregations whereby two items can become incomparable if item 1 is better than item 2 with regard to all criteria but one, but item 2 is significantly better than item 1 on the last criterion (see Roy and Bouyssou [13]).

4. CONCLUDING REMARKS

After restating the view of fuzzy pattern matching that the authors have developed in the framework of possibility theory since 1980, the paper has emphasized the introduction of tolerance either as approximate equality or as importance weighting, and has dealt with requirements of the form "this condition is not too important provided that this weaker (more tolerant) condition holds for sure". Another extension has been proposed for the evaluation of conditions involving fuzzy quantifiers in the presence of possibilistic data [11].

The problem of the comparison of imprecisely or fuzzily known pieces of data with a flexible condition, where fuzzy pattern matching plays a key role is encountered in a variety of problems:

- in advanced information systems where flexible queries are evaluated with respect to a database which may contain ill-known attribute values for some of the objects described in the base, e.g. [12], [17];
- in expert systems when condition parts of if… then… rules are fuzzy and/or facts are imprecise, uncertain or vague [9], [1];
- in case-based reasoning where the resemblance between approximately known objects or situations has to be evaluated [14];
- in classification systems where classes are fuzzily specified (for emphasizing the description of typical elements) and where objects are not always precisely known [8], [5]. See [10] for an example in medicine;
- in man-machine communication where linguistically described objects have to be identified from what is known about them [7], [18], [16].

ACKNOWLEDGEMENTS

The authors wish to thank Prof. K.W. Oh in Seoul who kindly published a preliminary version of the paper in the national Journal of Fuzzy Logic and Intelligent Systems of The Korea Fuzzy Mathematics and Systems Society (Vol. 3, N° 2, pp. 3-17, 1993).

REFERENCES

[1] Bensana E., Bel G., Dubois D. OPAL: A multi-knowledge-based system for industrial job-shop scheduling. Int. J. Prod. Research, 26(5), 1988, 795-819.

[2] Cayrol M., Farreny H., Prade H. Possibility and necessity in a pattern-matching process. Proc. of the 9th Inter. Congress on Cybernetics, Namur, Belgium, Sept. 1980, 53-65.

[3] Cayrol M., Farreny H., Prade H. Fuzzy pattern matching. Kybernetes, 11, 1982, 103-116.

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

[5] Dubois D., Prade H., Rossazza J.P. Vagueness, typicality, and uncertainty in class hierarchies. Int. J. of Intelligent Systems, 6, 1991, 167-183.

[6] Dubois D., Prade H., Testemale C. Weighted fuzzy pattern matching. Fuzzy Sets and Systems, 28, 1988, 313-331.

[7] Farreny H., Prade H. On the problem of identifying an object in a robotics scene from a verbal imprecise description. In: Advanced Software in Robotics (A. Danthine, M. Géradin, eds.), North-Holland, Amsterdam, 1984, 343-351.

[8] Granger C. An application of possibility theory to object recognition. Fuzzy Sets and Systems, 28, 1988, 351-362.

[9] Lebailly J., Martin-Clouaire R., Prade H. Use of fuzzy logic in a rule-based system in petroleum geology. In: Approximate Reasoning in Intelligent Systems, Decision and Control (E. Sanchez, L.A. Zadeh, eds.), Pergamon Press, 1986, 125-144.

[10] Pedrycz W., Bortolan G., Degani R. Classification of electrocardiographic signals: A fuzzy pattern matching approach. Artificial Intelligence in Medicine, 3, 1991, 211-226.

[11] Prade H. A two-layer fuzzy pattern matching procedure for the evaluation of conditions involving vague quantifiers. J. of Intelligent and Robotics Systems, 3, 1990, 93-101.

[12] Prade H., Testemale C. Generalizing data base relational algebra for a treatment of incomplete or uncertain information and vague queries. Information Sciences, 34, 1984, 115-143.

[13] Roy B., Bouyssou D. Aide Multicritère à la Décision: Méthodes et Cas. Economica, Paris, 1993.

[14] Salotti S. Représentation centrée objet et filtrage flou pour raisonner par analogie: Le système FLORAN. Actes du 7ème Congrès Reconnaissance des Formes et Intelligence Artificielle, Paris, Nov. 29-Dec. 1, 1989, AFCET Publ., 1695-1707.

[15] Sanchez E. Fuzzy logic and neural networks in Artificial Intelligence and Pattern Recognition. SPI, 1569, Stochastic and Neural Methods in Signal Processing, Image Processing and Computer Vision, 1991, 474-483.

[16] Vaina L.M., Jaulent M.C. Object structure and action requirements: A compatibility model for functional recognition. Int. J. of Intelligent Systems, 6, 1991, 313-336.

[17] Vandenberghe R., Van Schooten A., De Caluwe R., Kerre E.E. Some practical aspects of fuzzy database techniques: An example. Information Systems, 14(6), 1989, 465-472.

[18] Yager R.R. Robot planning with fuzzy sets. Robotica, 1, 1983, 41-50.

[19] Zadeh L.A. Fuzzy sets as a basis for a theory of possibility. Fuzzy Sets and Systems, 1, 1978, 3-28.