Finding Soft Relations in Granular Information Hierarchies

Trevor Martin, Yun Shen and Ben Azvine

Abstract— **When faced with large volumes of information, it is natural to adopt a granular approach by grouping together related items. Frequently, this is extended to a granular hierarchy, with progressively finer division as one moves down the hierarchy. The widespread use of hierarchical organisation shows that this is a natural approach for humans, as is the use of fuzzy granules rather than inflexible category specifications.**

Care is needed when information systems use fuzzy sets in this way - they are not disjunctive possibility distributions, but must be interpreted conjunctively. We clarify this distinction and show how an extended mass assignment framework can be used to extract relations between granules. These relations are association rules and are useful when integrating multiple information sources categorised according to different hierarchies. Our association rules do not suffer from problems associated with use of fuzzy cardinalities.

I. Introduction

Zadeh [1] suggested that granulation, organisation and causation were three important aspects underlying human cognition. Hierarchical categorisation combines the first two of these. The organisation of information into hierarchical conceptual structures is a common approach, providing a multi-level description (see e.g. [2, 3]). There is evidence that the human brain organises and classifies sensory input in a hierarchical fashion (see for example Jeff Hawkins Why can't a computer be more like a brain", IEEE Spectrum April 2007, pp17-22) and many familiar ways to access information such as books, libraries, computer file structures, the world wide web and other computer networks provide backing for the idea that hierarchical categorisation is an efficient way to organise and access information. For instance, books may be divided into fiction and non-fiction; within the non-fiction category we may then have further sub-division by genre such as historical fiction, mystery, humour, etc. In a similar way, one might divide the population according to age or income, as is frequently the case in socio-economic groupings. Again, the highest level split (say 0-18, 18-65, over 65) may be further sub-divided, so that (for example) 0-18 might be divided into 0-5, 5-11 and 11-18. We can regard a hierarchical categorisation as a series of progressively finer granulations. Clearly in most cases, the imposition of a crisp dividing line between categories might lead to anomalous behaviour near the borderlines and a fuzzy approach avoids such problems.

Trevor Martin (corresponding author), Artificial Intelligence Group, University of Bristol BS8 1TR UK, trevor.martin@ bris.ac.uk

Ben Azvine, Computational Intelligence Group, Intelligent Systems Laboratory, BT Research and Venturing, Ipswich IP5 3RE UK, ben.azvine@bt.com ..

Typically, such granulations are treated as if the categories were crisp. A library organising its book collection has to put each book in precisely one location. A humorous historical mystery novel has to be classified as (predominantly) belonging to one genre, and placed on the appropriate shelf in a library. Thus, even when a granulation is fuzzy, circumstances may dictate that it must be defuzzified, often in a somewhat arbitrary way. Information systems often face the same problem. The case of an instance (e.g. a novel) belonging to multiple categories is less of an issue when one is not dealing with physical objects – so an online bookstore could include our humorous historical mystery novel in each category to enhance its likelihood of being browsed. However, there is no way of indicating that its membership in one category should be stronger than in others. The use of graded membership (fuzziness) in categories enhances their expressive power and usefulness.

A related problem arises when trying to combine multiple sources of information that have been categorised in some way (often hierarchically). For example, the category of "vintage wine" has a different (but objective) definition, depending on the country of origin. To a purist, vintage wines are made from grapes harvested in a single year – however, the European Union allows up to 5% of the grapes to be harvested in a different year, the USA allows 15% in some cases and 5% in others, while other countries such as Chile and South Africa may allow up to 25%. Thus even taking a simple crisp granulation of wines into vintage an non-vintage categories can lead to problems if we try to integrate different sources.

A. iPHI - Intelligent Personal Hierarchies of Information

The iPHI system (intelligent Personal Hierarchies for Information) [4] aims to combine and integrate multiple sources of information and to configure access to the information based on an individual's personal categories. It is now widely accepted that taxonomic information, and in particular a class hierarchy, is an important part of a knowledge base. This is not just because objects in the real world are naturally associated with classes, but also because taxonomic information helps to organise knowledge through classification and provide efficient computation through inheritance. One of the key issues in the process of integration is determining when categories from different sources are the same, or approximately the same. For example, a classified directory may contain a category "Pubs

Yun Shen, Artificial Intelligence Group, University of Bristol BS8 1TR UK, yun.shen@bris.ac.uk

0-7695-3032-X/07 $25.00 © 2007 IEEE
DOI 10.1109/GrC.2007.30

and restaurants" whereas an online travel listing may categorise the same businesses as "places to eat". For a number of reasons, the two sets may not correspond exactly, but there may be considerable overlap. Such additional information may be useful in the integration of the two sources as well as in detection of "identical" instances from the two sources. For example we have used information about film genre (categorisation) to enhance the instance matching process [5, 6].

In this paper we describe a new method for calculating association rules to find correspondences between fuzzy granules in different hierarchies. We discuss the semantics of fuzzy sets when used to describe granules, and introduce a mass assignment-based method to rank association rules and show that the new method gives more satisfactory results than approaches based on fuzzy cardinalities.

II. Background

A. Fuzzy Sets in Information Systems

Many authors (e.g. [7]) have proposed the use of fuzzy sets to model uncertain values in databases and other knowledge based applications . The standard interpretation of a fuzzy set in this context is as a *possibility distribution* - that is to say it represents a single valued attribute which is not known exactly. For example we might use the fuzzy set *tall* to represent the height of a specific person or *low* to represent the value shown on a dice. The fuzzy sets tall and low admit a range of values, to a greater or lesser degree; the actual value is taken from the range. Knowing that a dice value *val* is *even* restricts the possible values to *val=2 XOR val=4 XOR val=6* (where *XOR* is an exclusive or). If a fuzzy set on the same universe is defined as *low = {1/1, 2/1, 3/0.4}* then knowing the value *val* is *low* restricts the possible values to *val=1 XOR val=2 XOR val=3* with corresponding memberships.

The conjunctive interpretation of a fuzzy set occurs when the attribute can have multiple values. For example, a person may be able to speak in several languages; we could model this as a fuzzy set of languages, where membership would depend on the degree of fluency. This is formally a relation rather than a function on the underlying domains. Our position in this paper is to make a distinction between the conjunctive interpretation - modelled by a fuzzy relation – and the disjunctive interpretation – modelled by a possibility distribution. To emphasise the distinction, we use the notation

$F(a) = \{x/\mu(x) \mid x \in U\}$

to denote a single valued attribute F of some object *a* (i.e. a possibility distribution over a universe *U*) and

$R(a) = [x/\chi(x) \mid x \in U]$

to denote a multi-valued attribute (relation). Granules represent the latter case, since we have multiple values that satisfy the predicate to a greater or lesser degree.

B. Association Rules

In creating association rules within transaction databases (e.g. [8]), the standard approach is to consider a table in which columns correspond to items and each row is a transaction. A column contains 1 if the item was bought, and 0 otherwise. The aim of association rule mining is to determine whether or not there are links between two disjoint subsets of items – for example, do customers generally buy biscuits and cheese when beer, lager and wine are bought.

Let us denote the set of items by X, so that any transaction can be represented as $tr \subseteq X$ and we have a multiset *Tr* of transactions.

An association rule is of the form $S \Rightarrow T$

where *s, t* are subsets of *X* such that *s* and *t* are non-overlapping and *S* (resp *T*) is the set of transactions containing the items *s* (resp *t*). This is interpreted as stating that when the items in *s* appear in a transaction, it is likely that the items in *t* will also appear i.e. this is not an implication in the formal logical sense.

Most authors use two measures to assess the significance of association rules.

The support of a rule is the fraction of transactions in which both *S* and *T* appear, and the confidence of a rule is the conditional probability of *T* given *S*

$$Support(S,T) = |S \cap T|$$

and

$$Conf(S,T) = \frac{|S \cap T|}{|S|}$$

where we operate on multisets rather than sets.

Typically a threshold is set for support, so that only frequently occurring sets of items *s* and *t* are considered; a second threshold filters out rules of low confidence.

Various approaches to fuzzifying association rules have been proposed, including [9-11]. The standard extension to the fuzzy case is to treat the (multi-) sets *S, T* as fuzzy and generalise the intersection and cardinality using a t-norm and sigma-count respectively.

$$Conf(S,T) = \frac{\sum_{x \in X} \mu_{S \cap T}(x)}{\sum_{x \in X} \mu_S(x)}$$

Note that many authors just refer to fuzzy sets, rather than multisets.

As pointed out by [10], using min and the sigma count for cardinality can be unsatisfactory because it does not distinguish between several tuples with low memberships and few tuples with high memberships - for example,

$$S = [x_1/1, x_2/0.01, x_3/0.01, \ldots, x_{1000}/0.01]$$

$$T = [x_1/0.01, x_2/1, x_3/0.01, \ldots, x_{1000}/0.01]$$

leads to

$$Conf(S,T) = \frac{1000 \times 0.01}{1 + 999 \times 0.01} \approx 0.91$$

which is extremely high for two almost disjoint sets. Using a fuzzy cardinality (i.e. a fuzzy set over the possible cardinality values) is also potentially problematic.

For these reasons, we propose the use of mass assignment theory in calculating the support and confidence of association rules between fuzzy categories.

The fuzziness in our approach arises because we allow partial membership in categories – for example, instead of looking for an association between biscuits and beer, we might look for an association between alcoholic drinks and snack foods. It is important to note that we are dealing with conjunctive fuzzy sets (monadic fuzzy relations) here. Mass assignment theory is normally applied to fuzzy set representing possibility distributions and the operation of finding the conditional probability of one fuzzy sets given another is known as semantic unification [12]. This rests on the underlying assumption of a single valued attribute – a different approach is required to find the conditional probability when we are dealing with set-valued attributes.

C. Mass Assignments

A mass assignment [13] (see also [14]) is a distribution over a power set, representing disjunctive uncertainty about a value. For a universe U

$$m : P(U) \rightarrow [0,1] \qquad \sum_{X \subseteq U} m(X) = 1 \tag{1}$$

It is related to a fuzzy set (possibility distribution) A as follows:
Let μ_A be the membership function of A with range

$$R(\mu_A) = \{\mu_A^1, \mu_A^2, \ldots, \mu_A^m\}$$

$$\textit{such that} \quad \mu_A^1 > \mu_A^2 > \ldots > \mu_A^m$$

and Ai be the alpha-cuts at these values i.e.

$$A_i = \{x | \mu_A(x) \geq \mu_A^i\}$$

(also known as the focal elements)
Then

$$m_A(A_i) = \mu_A^i - \mu_A^{i+1} \tag{2}$$

Given a fuzzy set A, the corresponding mass assignment can be written as

$$M(A) = \{A_i : m_A(A_i) \mid A_i \subseteq A\}$$

where conventionally only the focal elements (non-zero masses) are listed in the mass assignment.
The mass assignment represents a family of probability distributions on U, with the restrictions

$$p : U \rightarrow [0,1]$$

$$\sum_{x \in U} p(x) = 1$$

$$m(\{x\}) \leq p(x) \leq \sum_{x \in X} m(X)$$

For example, if
$X = \{a, b, c, d\}$
and A is the fuzzy set
$\{a/1, b/0.8, c/0.3, d/0.2\}$
then

$$M(A) = \{\{a\} : 0.2, \{a, b\} : 0.5, \{a, b, c\} : 0.1, \{a, b, c, d\} : 0.2\}$$

This also leads to a mass assignment definition of the cardinality of a fuzzy set as a distribution over integers

$$\Pr(|A| = n) = \sum_{\substack{A_i \subseteq A \\ |A_i| = n}} m_A(A_i)$$

for $0 \leq n \leq |U|$
Clearly in this framework, the cardinality of a fuzzy set can be given as a distribution over integer values, or an expected value can be produced from this distribution in the usual way.

Baldwin introduced the least prejudiced distribution (lpd) which is a specific distribution satisfying eq (1) above but also obeying

$$L_A(x) = \sum_{x \in A_i} \frac{m(A_i)}{|A_i|} \tag{3}$$

where $|A|$ indicates the cardinality of the set A and the summation is over all focal elements containing x.

Informally, wherever mass is associated with a non-singleton focal element, it is shared equally between the members of the set. Clearly a least prejudiced distribution is a restriction of the original assignment.

The steps from lpd to mass assignment and then to fuzzy set can be reversed, so that we can derive a unique fuzzy set for any frequency distribution on a finite universe, by assuming the relative frequencies are the least prejudiced distribution. (proof in [15])

If the relative frequencies are written

$$L_A = \{L_A(x_1), L_A(x_2), \ldots, L_A(x_n)\}$$

such that

$$L_A(x_1) > L_A(x_2) > \ldots > L_A(x_n)$$

then we can define

$$A_i = \{x | x \in U \wedge L_A(x) \geq L_A(x_i)\}$$

and the fuzzy set memberships are given by

$$\mu_A(x_i) = |A_i| \times L_A(x_i) + \sum_{j=i+1}^{n} \left(|A_j| - |A_{j-1}|\right) \times L_A(x_j)$$

D. Fuzzy relations and mass assignments

In standard mathematics, a relation is a conjunctive set of ordered *n*-tuples i.e. it represents a conjunction of *n* ground clauses.

In a similar way, a fuzzy relation represents a set of *n*-tuples that satisfy a predicate to a specified degree. For example, the set of dice scores then we could define a predicate *differBy4or5* on $U \times U$ as the set of pairs

[(1, 6), (1, 5), (2, 6), (5, 1), (6, 1), (6, 2)]

This is a conjunctive set in that each pair satisfies the predicate. In a similar way, a fuzzy relation represents a set of *n*-tuples that satisfy a predicate to a specified degree. Thus *differByLargeAmount* could be represented by

[(1, 6)/1, (1, 5)/0.6, (2, 6)/0.6, (5, 1)/0.6, (6, 1)/1, (6, 2)/0.6]

III. MASS-BASED ASSOCIATION RULES

We consider two granules, represented as monadic fuzzy relations *S* and *T* on the same domain, and wish to calculate the degree of association between them. For example, consider a database of sales employees, salaries and sales figures. We can categorise employees according to whether their salaries are high, medium or low and also according to whether their sales figures are good, moderate or poor. A mining task might be to find out whether the good sales figures are achieved by the highly paid employees. For example, given the table

name	sales	salary
a	100	1000
b	80	400
c	50	800
d	20	700

we might define

S = goodSales = [a/1, b/0.8, c/0.5, d/0.2]

and

T = highSalary = [a/1, b/0.4, c/0.8, d/0.7]

The confidence in an association rule can be calculated as follows

For a source granule

$$S = \left[x_1/\chi_S(x_1),\ x_2/\chi_S(x_2), \ldots,\ x_{|S|}/\chi_S\left(x_{|S|}\right)\right]$$

and a target granule

$$T = \left[x_1/\chi_T(x_1),\ x_2/\chi_T(x_2), \ldots,\ x_{|T|}/\chi_T\left(x_{|T|}\right)\right]$$

we can define the corresponding mass assignments as follows. Let the set of distinct memberships in S be

$$\left\{\chi_S^{(1)}, \chi_S^{(2)}, \ldots, \chi_S^{(n_S)}\right\}$$

where $\chi_S^{(1)} > \chi_S^{(2)} > \ldots > \chi_S^{(n_S)}$ and $n_S \le |S|$

Let

$$S_1 = \left\{\left[x \mid \chi_S(x) = \chi_S^{(1)}\right]\right\}$$

$$S_i = \left\{\left[x \mid \chi_S(x) \ge \chi_S^{(i)}\right]\right\} \cup S_{i-1} \qquad 1 < i \le n_S$$

and

Then the mass assignment corresponding to S is

$$\left\{S_i : m_S(S_i)\right\}, \quad 1 \le i \le n_S$$

where $m_S(S_k) = \chi_S^{(k)} - \chi_S^{(k+1)}$

and we define

$$\chi_S^{(i)} = 0 \quad if \quad i > n_S$$

For example, the fuzzy relation

S = [a/1, b/0.8, c/0.5, d/0.2]

has the corresponding mass assignment

$$M_S = \left\{\begin{array}{l}\{[a]\} : 0.2, \{[a],[a,b]\} : 0.3, \{[a],[a,b],[a,b,c]\} : 0.3, \\ \quad \{[a],[a,b],[a,b,c],[a,b,c,d]\} : 0.2\end{array}\right\}$$

The mass assignment corresponds to a distribution on the power set of relations, and we can define the least prejudiced distribution in the same way as for the standard mass assignment. In the example above

$$L_S = \{[a] : 0.5, [a,b] : 0.3, [a,b,c] : 0.15, [a,b,c,d] : 0.05\}$$

We can now calculate the confidence in the association between the granules *S* and *T* using mass assignment theory. In general, this will be an interval as we are free to move mass (consistently) between subsets of the granules.

For two mass assignments

$$M_S = \left\{\left\{S_{p_i}\right\} : m_S(S_i)\right\}, \quad 1 \le p_i \le i \le n_S$$

$$M_T = \left\{\left\{T_{q_j}\right\} : m_T(S_j)\right\}, \quad 1 \le q_j \le j \le n_T$$

the composite mass assignment is

$$M_C = M_S \oplus M_T$$
$$= \{X : m_C(X)\}$$

where M_C is specified by the composite mass allocation function

$$C\left(i, j, S_{p_i}, T_{q_j}\right)$$

subject to

$$\sum_{j=1}^{n_T} \sum_{\substack{1 \le q_j \le j \\ 1 \le p_i \le i}} C\left(i, j, S_{p_i}, T_{q_j}\right) = m_S(S_i)$$

$$\sum_{i=1}^{n_S} \sum_{\substack{1 \le p_i \le i \\ 1 \le q_j \le j}} C\left(i, j, S_{p_i}, T_{q_j}\right) = m_T(T_j)$$

		0.2	0.1		0.3			0.4			
		a	a	ac	a	ac	acd	a	ac	acd	abcd
0.2	a	0.2									
0.3	a		0.1								
	ab										0.2
0.3	a				0.3						
	ab										
	abc										
0.2	a							0.2			
	ab										
	abc										
	abcd										

(a) $Conf(S \to T) = \dfrac{0.2\times 1 + 0.1\times 1 + 0.2\times 2 + 0.3\times 1 + 0.2\times 1}{0.2\times 1 + 0.1\times 1 + 0.2\times 2 + 0.3\times 1 + 0.2\times 1}$
$= 1$

		0.2	0.1		0.3			0.4			
		a	a	ac	a	ac	acd	a	ac	acd	abcd
0.2	a	0.2									
0.3	a										
	ab		0.1					0.2			
0.3	a										
	ab										
	abc				0.3						
0.2	a										
	ab										
	abc										
	abcd							0.2			

(b) $Conf(S \to T) = \dfrac{0.2\times 1 + 0.1\times 1 + 0.2\times 1 + 0.3\times 1 + 0.2\times 1}{0.2\times 1 + 0.1\times 2 + 0.2\times 2 + 0.3\times 3 + 0.2\times 4}$
$= 0.4$

Fig 1 - Mass allocation (a) maximising and (b) minimising the association rule confidence

$$Conf\left(S \to T\right) = \left(\frac{\sum_{i=1}^{n_S}\sum_{j=1}^{n_T} \sum_{\substack{1 \le q_j \le j \\ 1 \le p_i \le i}} C\left(i, j, S_{p_i}, T_{q_j}\right) \times \left|S_{p_i} \cap T_{q_j}\right|}{\sum_{i=1}^{n_S}\sum_{j=1}^{n_T} \sum_{\substack{1 \le q_j \le j \\ 1 \le p_i \le i}} C\left(i, j, S_{p_i}, T_{q_j}\right) \times \left|S_{p_i}\right|} \right) \quad (4)$$

This can be visualised using a mass tableau (see [13]) For example consider the granules

S = [a/1, b/0.8, c/0.5, d/0.2] and
T = [a/1, b/0.4, c/0.8, d/0.7]

These can be combined as shown in Fig 1. Clearly the mass can be allocated in many ways, subject to the column constraints and it is not always straightforward to find the minimum and maximum confidences arising from different composite mass allocations. Two extreme examples are shown, so that the confidence in the association rule between granules lies in the interval [0.4, 1]. In general there can be considerable computation involved in finding the maximum and minimum confidences for a rule. When ranking association rules it is preferable to have a single figure for confidence, rather than an interval which can lead to ambiguity in the ordering.

We can redistribute the mass according to the least prejudiced distribution i.e. split the mass in each row (column) equally between its sub-rows (sub-columns), i.e.

$$C\left(i, j, S_{p_i}, T_{q_j}\right) = \frac{m_S\left(S_i\right)}{\left|S_i\right|} \times \frac{m_T\left(T_j\right)}{\left|T_j\right|}$$

for each S_{pi}, T_{qj}

Using eq (3) and combining sub-rows (sub-columns) with the same label yields

$$Conf(S \to T) = \left(\frac{\sum_{i=1}^{n_S} \sum_{j=1}^{n_T} L_S(S_i) \times L_T(T_j) \times |S_i \cap T_j|}{\sum_{i=1}^{n_S} L_S(S_i) \times |S_i|} \right)$$

which can be further simplified into a single summation by making use of the nested structure of the sets S_i $1 \le i \le n_S$ and T_j $1 \le j \le n_T$ This enables us to calculate associations with roughly $O(n)$ complexity, rather than $O(n^4)$ where n is the number of focal elements in the source granule S.

If we choose the least prejudiced distribution and re-arrange sub-rows into single rows with the same label (also columns) we obtain the following intersections

		0.45 a	0.25 ac	0.2 acd	0.1 abcd
0.5	a	a	a	a	a
0.3	ab	a	a	a	ab
0.15	abc	a	ac	ac	abc
0.05	abcd	a	ac	acd	abcd

This allows us to replace the calculation in eq 4 with straightforward calculations of the expected values of the cardinality of the source set and the intersection. In this case, the rule confidence is approximately 0.67 - lying in the interval shown in Fig 1 (obviously).

The example above gives a similar result to the cardinality-based method, but this is not always the case. For example if

$$S = [x_1/1, x_2/0.01, x_3/0.01, \ldots, x_{1000}/0.01]$$
$$T = [x_1/0.01, x_2/1, x_3/0.01, \ldots, x_{1000}/0.01]$$

then a fuzzy cardinality based approach gives a confidence of 10/10.99 ≈ 0.91 whereas our approach gives approximately 10^{-5}. Clearly this is a far more reasonable answer, as there are no elements with strong membership in both granules.

IV. Summary

We have described a new method for generating association rules between granules in different information hierarchies. These rules enable us to find related categories without leading to spurious relations suggested by association rules based on fuzzy cardinalities. The new method is currently being applied to discovery of links between film genres in different classification hierarchies, and links between categories in different classified business directories.

Acknowledgment

This work was partly funded by the Defence Technology Centre for Data and Information Fusion.

References

[1] Zadeh, L. A., "Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic," *Fuzzy Sets and Systems*, vol. 90, pp. 111-127, 1997.

[2] Sowa, J. F., *Conceptual Structures*: Addison Wesley, 1984.

[3] Yao, Y. Y., "Information Granulation and Rough Set Approximation," *International Journal of Intelligent Systems*, vol. 16, pp. 87-104, 2001.

[4] Martin, T. P. and B. Azvine, "Acquisition of Soft Taxonomies for Intelligent Personal Hierarchies and the Soft Semantic Web," *BT Technology Journal*, vol. 21, pp. 113-122, 2003.

[5] Martin, T. P. and Y. Shen, "Improving access to multimedia using multi-source hierarchical meta-data," in *Adaptive Multimedia Retrieval: User, Context, and Feedback*, vol. LNCS vol 3877, *LNCS*: Springer, 2006, pp. 266 - 278.

[6] Martin, T. P. and Y. Shen, "Improving Access to Multimedia Using Multi-source Hierarchical Meta-data," presented at Adaptive multimedia retrieval, Glasgow, Scotland, 2005.

[7] Bosc, P. and B. Bouchon-Meunier, "Databases and Fuzziness - Introduction," *International Journal of Intelligent Systems*, vol. 9, pp. 419, 1994.

[8] Agrawal, R. and R. Srikant, "Fast Algorithms for Mining Association Rules in Large Databases," presented at Very large data bases, Santiago, 1994.

[9] Bosc, P. and O. Pivert, "On Some Fuzzy Extensions of Association Rules," presented at IFSA world congress, Vancouver, Canada, 2001.

[10] Dubois, D., E. Hullermeier, and H. Prade, "A systematic approach to the assessment of fuzzy association rules," *Data Mining and Knowledge Discovery*, vol. 13, pp. 167-192, 2006.

[11] Kacprzyk, J. and S. Zadrozny, "Linguistic Summarization of Data Sets Using Association Rules," presented at Fuzzy systems; Exploring new frontiers, St Louis, MO, 2003.

[12] Baldwin, J. F., J. Lawry, and T. P. Martin, "Efficient Algorithms for Semantic Unification," presented at Information Processing and the Management of Uncertainty, Spain, 1996.

[13] Baldwin, J. F., "The Management of Fuzzy and Probabilistic Uncertainties for Knowledge Based Systems," in *Encyclopedia of AI*, S. A. Shapiro, Ed., 2nd ed: John Wiley, 1992, pp. 528-537.

[14] Dubois, D. and H. Prade, "On Several Representations of an Uncertain Body of Evidence," in *Fuzzy Information and Decision Processes*, M. M. Gupta and E. Sanchez, Eds.: North Holland, 1982.

[15] Baldwin, J. F., J. Lawry, and T. P. Martin, "A Mass Assignment Theory of the Probability of Fuzzy Events," *Fuzzy Sets and Systems*, vol. 83, pp. 353-367, 1996.