High Level Fuzzy Labels for Vague Concepts

Zengchang Qin* and Jonathan Lawry

Artificial Intelligence Group, Department of Engineering Mathematics, University of Bristol, BS8 1TR, UK.
{z.qin; j.lawry}@bris.ac.uk

1 Introduction

Vague or imprecise concepts are fundamental to natural language. Human beings are constantly using imprecise language to communicate each other. We usually say 'John is tall and strong' but not 'John is exactly 1.85 meters in height and he can lift 100kg weights'. Humans have a remarkable capability to perform a wide variety of physical and mental tasks without any measurements. This capability partitionsof objects into granules, with a granule being a clump of objects drawn together by indistinguishability, similarity, proximity or function [8]. We will focus on developing an understanding of how we can use vague concepts to convey information and meaning as part of a general strategy for practical reasoning and decision making.

We may notice that *labels* are used in natural language to describe what we see, hear and feel. Such labels may have different degrees of vagueness. For example, when we say Mary is *young* and she is *female*, the label *young* is more vague than the label *female* because people may have more widely different opinions on being *young* than being *female*. For a particular concept, there could be more than one label that is appropriate for describing this concept, and some labels could be more appropriate than others. A random set framework, *Label Semantics*, was proposed to interpret these facts [3]. In such a framework, linguistic expressions or labels such as *small*, *medium* and *large* are used for modelling. These labels are usually defined by overlapping fuzzy sets which are used to cover the universes of continuous variables. Different from Computing with Words [9], fuzzy labels are usually predefined and used for building intelligent systems such as decision tree [4, 5], naive Bayes learning [7] and rule induction systems [6] without involving the computing of semantic meanings of these labels.

In this paper, we extended the label semantics framework with high level fuzzy labels. In previous research of label semantics, fuzzy labels are used to describe a numerical data element and the corresponding appropriateness degree for using a

* Current Address: Berkeley Initiative in Soft Computing, Electrical Engineering and Computer Sciences Department, University of California, Berkeley, CA 94720, USA.

Z. Qin and J. Lawry: *High Level Fuzzy Labels for Vague Concepts*, Advances in Soft Computing **6**, 317–324 (2006)
www.springerlink.com

particular fuzzy label is just the membership of this data element belonging to the fuzzy label. Due to the vagueness and impreciseness of the real-world, numerical values are not always available. Here, we extend the label semantics framework to use higher level labels to describe some vague concepts which are also defined by intervals or fuzzy sets. The rest of the paper is structured as follows. Section 2 introduces the label semantics framework, based on which, the idea of high level fuzzy labels is disussed and supported with an example in section 3.

2 Label Semantics For Uncertainty Modeling

Label semantics is a methodology of using linguistic expressions or fuzzy labels to describe numerical values. For a variable x into a domain of discourse Ω we identify a finite set of fuzzy labels $\mathscr{L} = \{L_1, \cdots, L_n\}$ with which to label the values of x. Then for a specific value $x \in \Omega$ an individual I identifies a subset of $\mathscr{L}$, denoted D_x^I to stand for the description of x given by I, as the set of labels with which it is appropriate to label x. If we allow I to vary across a population V with prior distribution P_V, then D_x^I will also vary and generate a random set denoted D_x into the power set of $\mathscr{L}$ denoted by $\mathscr{S}$. We can view the random set D_x as a description of the variable x in terms of the labels in $\mathscr{L}$. The frequency of occurrence of a particular label, say S, for D_x across the population then gives a distribution on D_x referred to as a mass assignment on labels[2]. More formally,

Definition 1 (*Label Description*) *For $x \in \Omega$ the label description of x is a random set from V into the power set of $\mathscr{L}$, denoted D_x, with associated distribution m_x, which is referred to as mass assignment:*

$$\forall S \subseteq \mathscr{L}, \quad m_x(S) = P_V(\{I \in V | D_x^I = S\}) \tag{1}$$

where $m_x(S)$ is called the mass associated with a set of labels S and

$$\sum_{S \subseteq \mathscr{L}} m_x(S) = 1 \tag{2}$$

Intuitively mass assignment is a distribution on appropriate label sets and $m_x(S)$ quantifies the evidence that S is the set of appropriate labels for x.

For example, an expression such as '*the score on a dice is small*', as asserted by individual I, is interpreted to mean $D_{SCORE}^I = \{small\}$, where $SCORE$ denotes the value of the score given by a single throw of a particular dice. When I varies across a population V, different sets of labels could be given to describe the variable $SCORE$, so that we obtain the random set of D_{SCORE} into the power set of $\mathscr{L}$.

[2] Since $\mathscr{S}$ is the power set of L, the logical representation $S \in \mathscr{S}$ can be written as $S \subseteq \mathscr{L}$. The latter representation will be used through out this thesis. For example, given $\mathscr{L} = \{L_1, L_2\}$, we can obtain $\mathscr{S} = \{\emptyset, \{L_1\}, \{L_2\}, \{L_1, L_2\}\}$. For every element in $\mathscr{S}$: $S \in \mathscr{S}$, the relation $S \subseteq \mathscr{L}$ will hold.

In this framework, *appropriateness degrees* are used to evaluate how appropriate a label is for describing a particular value of variable x. Simply, given a particular value α of variable x, the appropriateness degree for labeling this value with the label L, which is defined by fuzzy set F, is the membership value of α in F. The reason we use the new term 'appropriateness degrees' is partly because it more accurately reflects the underlying semantics and partly to highlight the quite distinct calculus based on this framework [3]. This definition provides a relationship between mass assignments and appropriateness degrees.

Definition 2 (*Appropriateness Degrees*)

$$\forall x \in \Omega,\ \forall L \in \mathscr{L} \quad \mu_L(x) = \sum_{S \subseteq \mathscr{L}: L \in S} m_x(S)$$

For example, given a set of labels defined on the temperature outside: $\mathscr{L}_{Temp} = \{low, medium, high\}$. Suppose 3 of 10 people agree that '*medium* is the only appropriate label for the temperature of 15° and 7 agree 'both *low* and *medium* are appropriate labels'. According to def. 1, the mass assignment for 15° is $m_{15}(medium) = 0.3$, and $m_{15}(low, medium) = 0.7$ or formally:

$$m_{15} = \{medium\} : 0.3,\ \{low, medium\} : 0.7$$

More details about the theory of mass assignment can be found in [1]. In this example, we have that the appropriateness of *medium* as a description of 15° is $\mu_{medium}(15) = 0.7 + 0.3 = 1$, and that of *low* is $\mu_{low}(15) = 0.7$.

It is certainly true that a mass assignment on D_x determines a unique appropriateness degree for μ_L for any $L \in \mathscr{L}$, but generally the converse does not hold. For example, given $\mathscr{L} = \{L_1, L_2, L_3\}$ and $\mu_{L_1} = 0.3$ and $\mu_{L_2} = 1$. We could obtain an infinite family of mass assignments:

$$\{L_1, L_2\} : \alpha,\ \{L_2\} : \beta,\ \{L_2, L_3\} : 0.7 - \beta,\ \{L_1, L_2, L_3\} : 0.3 - \alpha$$

for any α and β satisfying: $0 \leq \alpha \leq 0.3, \quad 0 \leq \beta \leq 0.7$. Hence, the first assumption we make is that the mass assignment m_x are consonant and this allows us to determine m_x uniquely from the appropriateness degrees on labels as follows:

Definition 3 (*Consonant Mass Assignments on Labels*) *Let* $\{\beta_1, \cdots, \beta_k\} = \{\mu_L(x) | L \in \mathscr{L}, \mu_L(x) > 0\}$ *ordered such that* $\beta_t > \beta_{t+1}$ *for* $t = 1, 2, \cdots, k-1$ *then:*

$$m_x = M_t : \beta_t - \beta_{t-1},\ for\ t = 1, 2, \cdots, k-1,$$
$$M_k : \beta_k, \quad M_0 : 1 - \beta_1$$

where $M_0 = \emptyset$ *and* $M_t = \{L \in \mathscr{L} | \mu_L(x) \geq \beta_t\}$ *for* $t = 1, 2 \ldots, k$.

For the previous example, given $\mu_{L_1}(x) = 0.3$ and $\mu_{L_2}(x) = 1$, we can calculate the consonant mass assignments as follows: The appropriateness degrees are ordered as $\{\beta_1, \beta_2\} = \{1, 0.3\}$ and $M_1 = \{L_2\}$, $M_2 = \{L_1, L_2\}$. We then can obtain

$$m_x = \{L_2\} : \beta_1 - \beta_2, \{L_1, L_2\} : \beta_2 = \{L_2\} : 0.7, \{L_1, L_2\} : 0.3$$

Because the appropriateness degrees are sorted in def. 3 the resulting mass assignments are "nested". Clearly then, there is a unique consonant mapping to mass assignments for a given set of appropriateness degree values. The justification of the consonance assumption can be found in [1, 3]. Notice that in some cases we may have non-zero mass associated with the empty set This means that some voters believe that x cannot be described by any labels in $\mathscr{L}$. For example, if we are given $\mu_{L_1}(x) = 0.3$ and $\mu_{L_2}(x) = 0.8$, then the corresponding mass assignment is:

$$\{L_2\} : 0.5, \{L_1, L_2\} : 0.3, \emptyset : 0.2$$

where the associated mass for the empty set is obtained by $1 - \beta_1 = 0.2$.

3 High Level Label Description

In this section, we will consider how to use a high level fuzzy label to describe another fuzzy label. Here the term *high level* does not mean a hieracrhial structure. We will actually consider two set of fuzzy labels which are independently defined on the same universe. If the cardinality of a set of labels $\mathscr{L}$ is denoted by $|\mathscr{L}|$. We then can say $\mathscr{L}_1$ higher level labels of $\mathscr{L}_2$ if $\mathscr{L}_1 < \mathscr{L}_2$. We will acutally consider the methodology of using one set of fuzzy labels to represent the other set of fuzzy labels.

For example, a fuzzy concept *about_m* is defined by an interval on [a, b] (see the left-hand side figure of Fig. 1), so that the appropriateness degree of using fuzzy label *small* to label *about_m* is:

$$\mu_{small}(about\_m) = \frac{1}{b-a} \int_a^b \mu_{small}(u)du \tag{3}$$

If the vagueness of the concept *about_m* depends on the interval denoted by δ where the length of the interval $|\delta| = b - a$. We then can obtain:

$$\mu_{small}(about\_m) = \frac{1}{|\delta|} \int_{u \in \delta} \mu_{small}(u)du \tag{4}$$

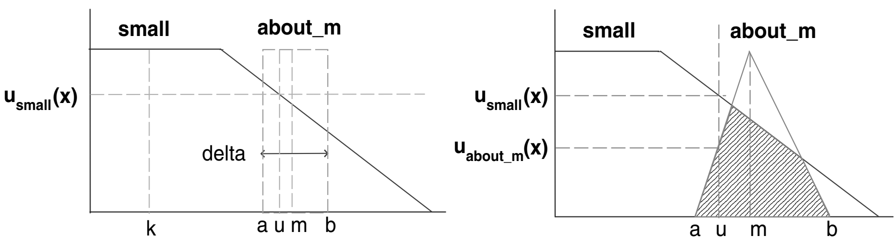

Fig. 1. The appropriateness degree of using *small* to label vague concept *about_m* is defined by the ratio of the area covered by both labels to the area covered by *about_m* only.

If *about_m* is defined by other fuzzy labels rather than an interval, for example, a triangular fuzzy set (e.g., the right-hand side figure of Fig. 1). How can we define the appropriateness degrees?

We begin by considering a data element $x \in [a,b]$, the function $\mu_{about\_m}(x)$ represents the degree of x belonging to the fuzzy label F. Function $\mu_{small}(x)$ defines the appropriateness degrees of using label *small* to describe x [3]. We essentially hope to obtain the appropriateness degrees of using *small* to label *about_m*. We then consider the each elements belonging to *about_m*. If $\mu_{about\_m}(x) = 1$, which means x is absolutely belonging to *about_m*, then the appropriateness degree is just $\mu_{small}(x)$. However, if $\mu_{about\_m} < \mu_{small}(x)$, we can only say it is belonging to *about_m* in certain degrees. Logically, fuzzy operation AND is used, and in practical calculation, the $\min(\cdot)$ function is employed. The appropriateness is then defined by:

$$\mu_{small}(about\_m) = \frac{\int_{u\in\delta} \min(\mu_{small}(u), \mu_{about\_m}(u))du}{\int_{u'\in\delta} \mu_{about\_m}(u')du'} \tag{5}$$

where function $\min(x,y)$ returns the minimum value between x and y. Equation 4 is a special case of equation 5 where the following equations always hold:

$$\mu_{small}(u) = \min(\mu_{small}(u), \mu_{about\_m}(u))$$

$$|\delta| = \int_{u\in\delta} \mu_{about\_m}(u)du$$

Definition 4 *Given a vague concept (or a fuzzy label) F and a set of labels $\mathscr{L} = \{L_1, \ldots, L_m\}$ defined on a continuous universe Ω. The appropriateness degrees of using label L ($L \in \mathscr{L}$) to describe F is:*

$$\mu_L(F) = \frac{\int_{u\in\delta} \min(\mu_L(u), \mu_F(u))du}{\int_{u'\in\delta} \mu_F(u')du'} \tag{6}$$

where δ is the universe covered by fuzzy label F.

Given appropriateness degrees, the mass assignment can be obtained from the appropriateness degrees by the consonance assumption. Equation 5 is a general form for all kinds of fuzzy sets which are not limited to an interval or a triangular fuzzy sets.

Example 1. Figure 1 gives a set of isosceles triangular fuzzy labels $F_1, \ldots, F_8$ and two high level fuzzy label *small* and *large* defined on the same universe. The membership functions (the non-zero part) for F_5, F_6 and *small* are defined as follows:

$$PS \rightarrow y = \frac{5}{2}x - 3, \;\; PT \rightarrow y = -\frac{5}{2}x + 5$$

$$QR \rightarrow y = \frac{5}{2}(x - 0.4) - 3, \;\; QU \rightarrow y = -\frac{5}{2}(x - 0.4) + 5$$

$$OU \rightarrow y = -\frac{5}{6}x + 2$$

[3] Here we interpret $\mu(\cdot)$ in different manners: membership function and appropriateness degrees, though they are mathematically the same.

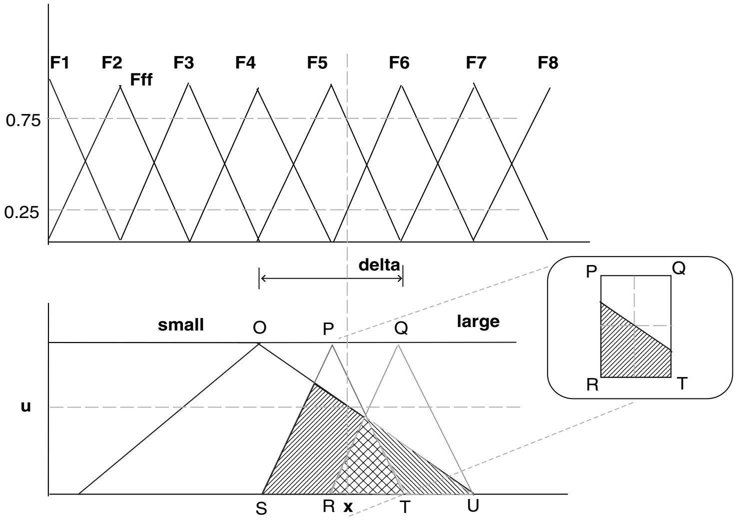

Fig. 2. The relations between fuzzy labels.

As we can see from Fig. 2: $\mu_{F_5}(x) = 0.75$ and $\mu_{F_6}(x) = 0.25$ given $x = 1.7$. According to definition 4 we can obtain:

$$\mu_{small}(F_5) = 0.8,\ \mu_{large}(F_5) = 1$$

$$\mu_{small}(F_6) = 0.5,\ \mu_{large}(F_6) = 1$$

So that the corresponding consonant mass assignments (see definition 3) are:

$$m_{F_5} = \{small, large\} : 0.8, \{large\} : 0.2$$

$$m_{F_6} = \{small, large\} : 0.5, \{large\} : 0.5$$

High level labels *small* and *large* can be used to describe $x = 1.7$ by the following steps.

$$m_x = \{F_5, F_6\} : 0.25, \{F_5\} : 0.5, \emptyset : 0.25$$

F_5 and F_6 can be represented by the mass assignments of high level fuzzy labels: *small* and *large*. Considering the term $\{F_5, F_6\}$, it means that both two labels F_5 and F_6 are appropriate for labeling x with a certain degree. It defines a area covered both by F_5 and F_6 (see Fig. 2) which is an interval between R and T. Therefore, according to def. 4 we can obtain the mass assignment for $\{F_5, F_6\}$:

$$m_{\{F_5,F_6\}} = \{small, large\} : 0.5, \{large\} : 0.5$$

Finally, we obtain:

$$\begin{aligned} m_x &= (\{small, large\} : 0.5, \{large\} : 0.5) : 0.25, \\ &\quad (\{small, large\} : 0.8, \{large\} : 0.2) : 0.5, \emptyset : 0.25 \\ &= \{small, large\} : 0.525, \{large\} : 0.225, \emptyset : 0.25 \end{aligned}$$

From the above example, if we use *small* and *large* to describe x directly. By the function of *small* we can obtain $u = \frac{7}{12}$ so that the mass assignments are:

$$m_x = \{small, large\} : \frac{7}{12}, \{large\} : \frac{5}{12}$$

which is different from the result presented in example 1. It is because precision is lost by using two level of fuzzy labels. In our example, x is firstly repressed by F_5 and F_6 which is precise. However, the description of x by *small* and *large* through F_5 and F_6 is not precise any more, because F_5 and F_6 are not exact representation of x by involving uncertainties decided by δ. As we can see from the Fig. 3: the appropriateness degrees of using high level labels to describe low level concepts are depending on the uncertainty parameter δ. For example, given a data element m:

$$|\mu_{small}(F(\delta_1)) - \mu_{small}(m)| < |\mu_{small}(F(\delta_2)) - \mu_{small}(m)| < |\mu_{small}(F(\delta_3)) - \mu_{small}(m)|$$

So that:

$$\mu_{small}(m) = \lim_{\delta \to 0} \mu_{samll}(F(\delta))$$

where F is the function of the fuzzy label (a function of δ-either an interval, triangular fuzzy set or other type of fuzzy set) centered on m.

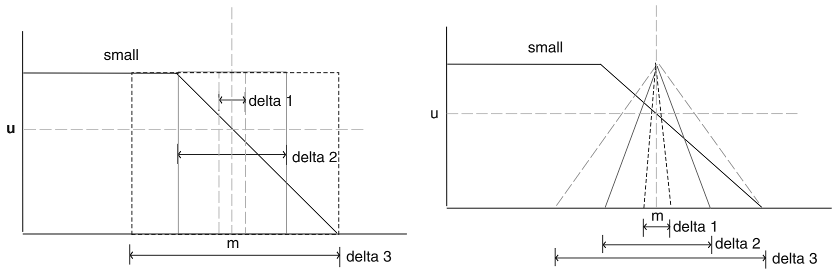

Fig. 3. The appropriateness degree of using *small* depends on the width of the vague concept of *about_m*.

4 Conclusions

In this paper, a methodology of using high level fuzzy labels to describe vague concepts or low level fuzzy labels is proposed based on label semantics framework. An example is given to show how to calcuate the mass assigments of high level fuzzy labels on a vague concept.

References

[1] J. F. Baldwin, T. P. Martin and B. W. Pilsworth, *Fril- Fuzzy and Evidential Reasoning in Artificial Intelligence*, John Wiley & Sons, Inc., 1995.

[2] George J. Klir and Bo Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice-Hall, Inc., Upper Saddle River, NJ, 1994.

[3] Jonathan Lawry, *Modelling and Reasoning with Vague Concepts*, Springer, 2006.

[4] Zengchang Qin and Jonathan Lawry, Decision tree learning with fuzzy Labels, *Information Sciences*, vol. 172: 91-129, June 2005.

[5] Zengchang Qin and Jonathan Lawry, Linguistic rule induction based on a random set semantics. *World Congress of International Fuzzy Systems Association* (IFSA-05), July 2005.

[6] Zengchang Qin and Jonathan Lawry, Prediction trees using linguistic modelling. *World Congress of International Fuzzy Systems Association* (IFSA-05), July 2005.

[7] Nicholas J. Randon and Jonathan Lawry, Classification and query evaluation using modelling with words, *Information Sciences* 176(4): 438-464, 2006.

[8] L. A. Zadeh, A new direction in AI - toward a computational theory of perceptions, *A.I. Magazine*, Spring 2001.

[9] L.A. Zadeh, Fuzzy logic = computing with words, *IEEE Trans. Fuzzy Systems*, Vol.2:103-111, 1996.