Proceedings of the Ninth International Conference on Machine Learning and Cybernetics, Qingdao, 11-14 July 2010

Dissimilarity Measure of Logical Expressions

WEIFENG ZHANG[1], ZENGCHANG QIN[1]

[1]Intelligent Computing and Machine Learning Laboratory
School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China
E-MAIL: zwf.zhang@gmail.com, zengchang.qin@gmail.com

Abstract

The concept of label semantics[6] proposed by Lawry is a random set framework for modeling with words. It proposes two fundamental and inter-related measures of the a appropriateness of labels as descriptions of an object or value. In this paper, based on an introduction of logical expressions based on label semantics, we present a new measure to evaluate the distance between two logical expressions of fuzzy labels. Two fundamental and important properties of this distance are discussed. It is theoretically proved and tested by experiments that the new measure can explain the dissimilarity of linguistic rules well.

Keywords:

Label semantics; Fuzzy labels; Dissimilarity measure.

1. Introduction

In order to understand the use of natural language for information and knowledge processing in computer systems, Zadeh proposed the concept of computing with words paradigm. He suggested a form of precisiated natural language [12] based on the theory of generalized constraints and linguistic variables. Label semantics, introduced by Lawry [6, 5, 3, 2], provides an alternative representation for modeling with words. In contrast to Zadeh's methodology this approach is based on measures of an agent's subjective belief that a logical expression is appropriate to describe a particular object or value.

The theory of fuzzy sets proposed by Zadeh in 1965 [13] is an efficient method for representing and processing vague information. In order to use fuzzy set in practice, many similarity/dissimilarity measures [1, 4, 7, 10] have been proposed for measuring the degree of similarity between fuzzy sets. But those measures are not proper to deal with the similarity/dissimilarity measures between logical expressions which are concepts based on given linguistic variables. Label semantics focus on the decision making process an intelligent agent must go through in order to identify which labels or logical expressions can actually be used to describe an object or value. For this reason, Lawry has introduced the mass assignment and appropriateness degree which measures the appropriateness of using a particular subset of labels to describe an object or value. One step further, we present a measure to evaluate the dissimilarity of logical expressions based on the mass assignments and appropriateness degrees which can quantize the divergences between logical expressions in this paper.

This paper is organized as follows: Section 2 gives a general introduction to label semantics. In Section 3, logical expression of labels is defined formally. In section 4, we first define the distance between two variables on the discourse and then the formal definition of distance between logical expressions is given. And some important properties of this distance are discussed and proved as well. In section 5, an experiment on a vague concept, color, is done to practice the distances defined in section 4 and show the significant properties of these distances vividly.

2 Label semantics

Label semantics is a random set framework for modeling with words. The fundamental notion underlying label semantics is that when individuals make assertion of the form '**x** is θ', they are essentially providing information about what labels are appropriate for the value of the underlying variable **x** [5].

Definition 1 (Mass assignment on labels) *For $x \in \Omega$ the label description of x is a random set from V into the power set of $\mathcal{L}$,denoted by $\mathcal{D}_x$, with associated distribution m_x, which is referred to as mass assignment:*

$$\forall S \subseteq \mathcal{L}, m_x(S) = P(I \in V : D_x^I = S). \qquad (1)$$

978-1-4244-6525-5/10/$26.00 ©2010 IEEE

To evaluate how appropriate a label is for describing a particular value of variable x, $appropiatenessdegrees$ are defined.

Definition 2 (Appropriateness degrees)

$$\forall x \in \Omega, \forall L \in \mathcal{L}, \mu_L(x) = \sum_{S \subseteq \mathcal{L}: L \in S} m_x(S). \quad (2)$$

Example 1 *Given a set of labels defined on the height of an adult:* $\mathcal{L}_{Height} = \{short, medium, tall\}$*. Suppose 4 of 10 people agree that ‘ medium is the only appropriate label for the height of 173cm ’ and 6 support that ‘ both short and medium are appropriate labels ’. According to Definition 1, the mass assignment for 173cm is:*

$$m_{173} = \{medium\} : 0.4, \{short, medium\} : 0.6. \quad (3)$$

And according to Definition 2, the appropriateness degree ofmedium as a description of 173cm is $\mu_{medium}(173)=0.4+0.6=1$ *and that of short is* $\mu_{short}(173)=0.6$*.*

3 Logical expressions of fuzzy labels

Given a universe of discourse Ω containing a set of objects or instances to be described, it is assumed that all relevant expression can be generated recursively from a finite set of basic labels $\mathcal{L} = \{L_1, L_2, ..., L_n\}$. Operators for combing expressions are restricted to the standard logical connectives of negation ‘$\neg$’, conjunction ‘$\wedge$’, disjunction ‘$\vee$’ and implication ‘$\rightarrow$’. Hence, the set of logical expressions of labels can be formally defined as follows:

Definition 3 (Logical expressions of labels) *The set of logical expressions, LE, is defined recursively as follows:*
(i) $L_i \in LE$ *for i=1,2,...,n.*
(ii)If $\theta, \varphi \in LE$ *then* $\neg\theta, \theta \wedge \varphi, \theta \vee \varphi, \theta \rightarrow \varphi \in LE$*.*

Basically, we interpret the main logical connectives as follows: $\neg L$ means that L is not an appropriate label, $L_1 \wedge L_2$ means both L_1 and L_2 are appropriate labels, $L_1 \vee L_2$ means that either L_1 or L_2 are appropriate labels and $L_1 \rightarrow L_2$ means that L_2 is an appropriate label whenever L_1 is. As well as labels for a single variable, we may want to evaluate the appropriateness degree of a complex logical expression $\theta \in LE$. Consider the set of logical expressions LE obtained by recursive application of the standard logical connectives in $\mathcal{L}$. In order to evaluate the appropriateness degrees of such expressions we must identify what information they provide regarding the appropriateness of labels. In general, for any label expression θ we should be able to identify a maximal set of label sets, $\lambda(\theta)$, that are consistent with θ so that the meaning of θ can be interpreted as the constraint $D_x \in \lambda(\theta)$.

Definition 4 (λ-Function) *Let θ and φ be expressions generated by recursive application of the connectives $\neg, \wedge, \vee$ and $\rightarrow$ to the elements of $\mathcal{L}$(i.e. $\theta \in LE$).Then the set of possible label sets defined by a linguistic expression can be determined recursively as follows:*
(i) $\lambda(L_i(x)) = \{S \subseteq \mathcal{F} | \{L_i\} \subseteq S\}$.
(ii) $\lambda(\neg\theta) = \lambda(\theta)$.
(iii) $\lambda(\theta \vee \varphi) = \lambda(\theta) \bigcap \lambda(\varphi)$.
(iv) $\lambda(\theta \wedge \varphi) = \lambda(\theta) \bigcup \lambda(\varphi)$.
(v) $\lambda(\theta \rightarrow \varphi) = \overline{\lambda(\theta)} \bigcap \lambda(\varphi)$.

Then $\lambda(\theta)$ corresponds to those subsets of $\mathcal{F}$ identified as being possible values of D_x by the expression θ. λ function will be a useful tool in our next work.

4 Distance of Logical Expressions

From last section we know that variables on the universe which can be expressed by different mass assignments have different logical meanings obviously. Thus a measure which can measure the difference between the logical meanings of variables is needed.

Definition 5 (Distance between two variables) *Given two variables x_1 and x_2 from an universe Ω who is fully covered by sets of labels $\{S_1, S_2, ..., S_n\}$, then the distance between x_1 and x_2 is:*

$$\mathcal{D}(x_1, x_2) = \lim_{M \to \infty} \sum_{m=1}^{M} D(u, v) \quad (4)$$

where

$$D(u, v) = \sum_{i=1}^{n} (m_u(S_i) - m_v(S_i))^2 \quad (5)$$

$$u = x_1 + \frac{m}{M}(x_2 - x_1) \quad (6)$$

$$v = x_1 + \frac{m-1}{M}(x_2 - x_1) \quad (7)$$

This definition has two important properties:

Theorem 1 *The distance $\mathcal{D}$ defined by Equation* 4 *is symmetric.*

Proof 1

$$\mathcal{D}(x_2, x_1) = \lim_{M \to \infty} \sum_{m'=1}^{M} D(u', v')$$

where

$$u' = x_2 + \frac{m'}{M}(x_1 - x_2)$$
$$= x_1 + \frac{M - m'}{M}(x_2 - x_1)$$

$$v' = x_2 + \frac{m' - 1}{M}(x_1 - x_2)$$
$$= x_1 + \frac{M - m' + 1}{M}(x_2 - x_1)$$

$M - m' + 1$ has the same value domain. Thus

$$D(u', v') = D(v, u)$$

In addition, because of the Symmetry of D defined by Equation 5, the following conclusion can be deduced.

$$\mathcal{D}(x_2, x_1) = \mathcal{D}(x_1, x_2)$$

So the distance between variables is symmetric. □

Theorem 2 *If given $x_1 < x_2 < x_3$, then*

$$\mathcal{D}(x_1, x_3) \geq \mathcal{D}(x_1, x_2) \tag{8}$$

Theorem 2 demonstrates that this measure has non-negative correlation with the distance of x which can be easily proved as follows:

Proof 2

$$\mathcal{D}(x_1, x_2) = \lim_{M\to\infty} \sum_{m=1}^{M} D(u, v)$$

where

$$u = x_1 + \frac{m}{M}(x_2 - x_1)$$
$$v = x_1 + \frac{m-1}{M}(x_2 - x_1)$$
$$\mathcal{D}(x_1, x_3) = \lim_{M\to\infty} \sum_{m=1}^{M} D(u', v')$$

where

$$u' = x_1 + \frac{m}{M}(x_3 - x_1)$$
$$v' = x_1 + \frac{m-1}{M}(x_3 - x_1)$$

obviously

$$\forall m, u' - v' > u - v$$

because $D(u', v') >= D(u, v)$,thus

$$\mathcal{D}(x_1, x_3) \geq \mathcal{D}(x_1, x_2)$$

Now proof of Theorem 2 has been completed. □

Above definition express the logical distance between two variables on the discourse. Now we consider this question: how can we design a measure to calculate the distance between a point and a fussy set. Previous works have given a logical explanation about the relationship of a variable on the discourse and a fussy set defined on the same discourse. Basically the relationship can be expressed by using membership function: $\mu_x(S) = \alpha$, where $\alpha \in [0, 1]$. Now we present a more logical, reasonable and also more complex measure based on it.

Definition 6 (Distance between variable and set of labels) *Given a point x_0 and a set of labels S covers a continuous extent [a,b] on the universe Ω, the distance between x_0 and S is defined as follows:*

$$\mathcal{D}(x_0, S) = \frac{\int_a^b \mathcal{D}(x_0, x)dx}{b - a} \tag{9}$$

Further more,two sets of labels also have different logical meanings. Distance of sets of labels which is used to measure the divergence between them can be defined as follows:

Definition 7 (Distance of sets of labels) *Given two sets of labels S_i and S_j with S_i covering a continuous extent [a,b] and S_j covering a continuous extent [c,d]. Then distance between these two sets is defined as:*

$$\mathcal{D}(S_i, S_j) = \frac{\int_a^b \int_c^d \mathcal{D}(x, y)dxdy}{(b - a)(d - c)}, x \in S_i, y \in S_j \tag{10}$$

when $i = j$

$$\mathcal{D}(S_i, S_j) = \mathcal{D}(S_i, S_i) = 0 \tag{11}$$

For the symmetry of the distance between two variables, it is obviously that the distance of sets of labels is also symmetric.

Basically, λ-function provides a way of mapping from logical expressions of labels to random set descriptions of labels. So we also can define the distance between two logical expressions.

Definition 8 (Distance of logical expressions) *Given two logical expressions $\theta, \varphi \in LE$, then the distance between θ and φ is*

$$\mathcal{D}(\theta, \varphi) = \frac{1}{pr} \sum_{i=1}^{r} \sum_{j=1}^{p} \mathcal{D}(S_i^{\theta\wedge\neg\varphi}, S_j^{\varphi}) - \frac{1}{qt} \sum_{k=1}^{t} \sum_{l=1}^{q} \mathcal{D}(S_k^{\varphi\wedge\neg\theta}, S_l^{\theta}) \tag{12}$$

where p, q, r, t respectively means the count of elements of set $\mathbb{S}^{\varphi}$, $\mathbb{S}^{\theta}$, $\mathbb{S}^{\theta\wedge\neg\varphi}$, $\mathbb{S}^{\varphi\wedge\neg\theta}$

$S_i^{\theta} \in \mathbb{S}^{\theta} = \{S | S \in \lambda(\theta)\}, i = 1, 2, ..., q.$

$S_j^{\varphi} \in \mathbb{S}^{\varphi} = \{S | S \in \lambda(\varphi)\}, j = 1, 2, ..., p.$

$S_k^{\theta\wedge\neg\varphi} \in \mathbb{S}^{\theta\wedge\neg\varphi} = \{S | S \in \lambda(\theta) \bigcap \overline{\lambda(\varphi)}\}, k = 1, 2, ..., r.$

$S_l^{\varphi\wedge\neg\theta} \in \mathbb{S}^{\varphi\wedge\neg\theta} = \{S | S \in \lambda(\varphi) \bigcup \overline{\lambda(\theta)}\}, l = 1, 2, ..., t.$

When $\mathbb{S}^{\varphi\wedge\neg\theta} = \Phi$,

$$\mathcal{D}(\theta, \varphi) = \frac{1}{pr} \sum_{i=1}^{r} \sum_{j=1}^{p} \mathcal{D}(S_i^{\theta\wedge\neg\varphi}, S_j^{\varphi}) \tag{13}$$

When $\mathbb{S}^{\theta\wedge\neg\varphi} = \Phi$,

$$\mathcal{D}(\theta,\varphi) = \frac{1}{qt}\sum_{k=1}^{t}\sum_{l=1}^{q}\mathcal{D}(S_k^{\varphi\wedge\neg\theta}, S_l^{\theta}) \tag{14}$$

The above logical expression is one dimensional, which can be used by agents to describe one of the features of the object. Actually every object have more than one sides. Linguistic rule induction model based on label semantics is an effective approach. Based on Definition 3, a linguistic rule is a rule that can be represented as a multi-dimensional logical expressions of fuzzy labels.

Definition 9 (Multi-dimensional logical expressions of labels) *$MLE^{(n)}$ is the set of all multi-dimensional label expressions that can be generated from the logical label expression $LE_j : j = 1, ..., n$ and is defined recursively by*
(i) If $\theta \in LE_j$ for $j = 1, ..., n$ then $\theta \in MLE^{(n)}$.
(ii) If $\theta, \varphi \in MLE^{(n)}$ then $\neg\theta, \theta\wedge\varphi, \theta\vee\varphi, \theta\rightarrow\varphi \in MLE^{(n)}$

Similarly we could give the definition of distance between two $MLE^{(n)}$.

Definition 10 (Distance between multi-dimensional logical expressions) *Given two n-dimensional logical expressions: Φ, Ψ with*

$$\Phi = \theta_{D_1}\wedge\theta_{D_2}\wedge ... \wedge\theta_{D_n}$$

$$\Psi = \varphi_{D_1}\wedge\varphi_{D_2}\wedge ... \wedge\varphi_{D_n}$$

where θ_{D_i}, φ_{D_i} respectively means the logical expressions in dimension D_i. hence, the distance between Φ and Ψ is defined as follows:

$$\mathcal{D}(\Phi,\Psi) = \sqrt{\sum_{i=1}^{n}|\mathcal{D}(\theta_{D_i},\varphi_{D_i})|^2} \tag{15}$$

5 Experiment

Color of an object is a vague concept and largely depends on the observer's subjective belief. But there is a general principle for human beings to decide using the most appropriate word such as yellow or red to describe the color of an object.And people know the difference among these colors, which is difficult to defined with some specific numerical value. We often make classifications based on these vague and subjective differences among objects. The above definition is try to make a measure to quantize these differences in order to let the agents have the ability to distinguish different objects.

HSV color space which can be well visualized by the conical representation model quite accords with human beings' visual features. Fig. 1 shows three fuzzy sets defined

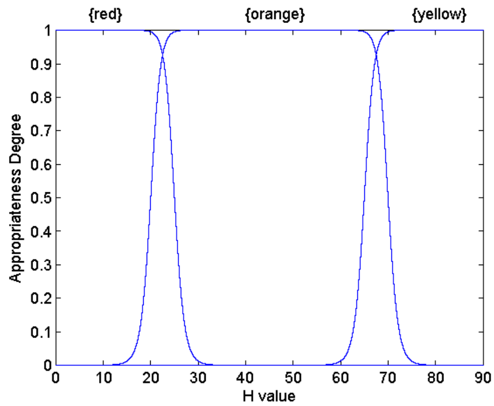

Figure 1. H values covered by three labels

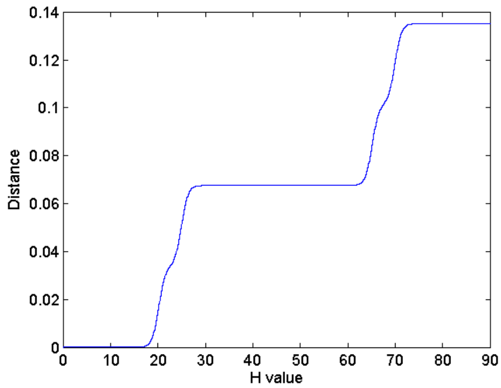

Figure 2. Distance between variables with $x_0 = 0$,x_1 varying from 0 to 90.

on the Hue value. h_0 and h_1 are two variables on this discourse. Fig. 2 shows the variation of $\mathcal{D}(h_0, h_1)$ with different values of h_0 and h_1. Here we let h_0 equals 0, and h_1 varies from 0 to 90. It is illustrated that where the mass assignments of h_1 have little change with its value varying, $\mathcal{D}(h_0, h_1)$ stays the same. In contrast, $\mathcal{D}(h_0, h_1)$ varies rapidly where the mass assignments of h_1 change with its value varying.

Table 1 shows the distances between these sets of labels and demonstrate that the greater the logical divergence of two labels is, the greater the distance we defined is. Further more, if person A holds the idea that a color with its H value equaling 20 is not orange, formally expressed as follows:

$$\theta_{20} = \neg orange$$

Person B thinks that this color is red or orange,also ex-

Table 1. Distances between sets of labels.

	$\{r\}$	$\{r,o\}$	$\{o\}$	$\{o,y\}$	$\{y\}$
$\{r\}$	0	0.0288	0.0626	0.0950	0.1246
$\{r,o\}$	0.0288	0	0.0354	0.0673	0.0968
$\{o\}$	0.0626	0.0354	0	0.0336	0.0626
$\{o,y\}$	0.0950	0.0673	0.0336	0	0.0304
$\{y\}$	0.1246	0.0968	0.0626	0.0304	0

Table 2. Distances between logical expressions.

	θ_{20}	φ_{20}	γ_{20}
θ_{20}	0	0.1413	0.1735
φ_{20}	0.1413	0	0.0650
γ_{20}	0.1735	0.0650	0

pressed as follows:

$$\varphi_{20} = red \vee orange$$

And person C believes that this color is red.

$$\gamma_{20} = red$$

According to Definition 4, the possible label sets of the given logical expressions θ_{20}, γ_{20} and φ_{20} are calculated as follows:

$$\lambda(\neg o) = \{\{r\},\{y\}\}$$
$$\lambda(r) = \{\{r\},\{r,o\}\}$$
$$\lambda(o) = \{\{r,o\},\{o\}\{o,y\}\}$$

so that

$$\lambda(\theta_{20}) = \{\{r\},\{y\}\}$$
$$\lambda(\varphi_{20}) = \{\{r\},\{r,o\},\{o\}\{o,y\}\}$$
$$\lambda(\gamma_{20}) = \{\{r\},\{r,o\}\}$$

then according to Definition 8 and Table 1, we could calculate all the distances filled in Table 2. It is illustrated that distances between LEs are symmetric and could reasonably reflect the logical divergences between LEs.

6 Conclusion

In this paper we have defined the distance between variables on the discourse which focus on the difference of logical meanings they convey. Based on it we present the distance between variable and sets of labels, distance between sets. Finally the formal definition of distance between logical expression is proposed which is capable to measure the divergence of different logical expressions. It is proved that this measure is symmetric and could reflect the intuitionistic dissimilarity between logical expressions reasonably.

References

[1] L.K. Hyung, Y.S. Song, K.M. Lee, Similarity measure between fuzzy sets and between elements, *Fuzzy Sets and System* Vol. 62(1994), pp. 291-293.

[2] J. Lawry, *Modelling and Reasoning with Vague Concepts*, (2006), Springer.

[3] J. Lawry, A Framework for Linguistic Modelling, *Artificial Intelligence*, Vol. 155(2004), pp. 1-39

[4] Deng-Feng Li, Some measures of dissimilarity in intuitionistic fuzzy structures, *Journal of Computer and System Sciences* Vol. 8(2004), pp. 115-122.

[5] J. Lawry, J. Recasens, A Random Set Model for Fuzzy Labels, *Proceedings of The European Conference on Symbolic and Quantitative Approaches to Reasoning Under Uncertainty, Lecture Notes in Artificial Intelligence*, (Eds. T. Nielsen,N. Zhang), Vol. 2711(2003), pp. 357-369.

[6] J. Lawry, Label Semantics: A Formal Framework for Modeling with Words, *Proceedings of The European Conference on Symbolic and Quantitative Approaches to Reasoning Under Uncertainty, Lecture Notes in Artificial Intelligence*, (Eds. S. Benferhat, P. Besnard), Vol. 2143(2001), pp. 374-385.

[7] C.P. Pappis, N.I. Karacapilidis, A comparative assessment of measures of similarity of fuzzy values, *Fuzzy Sets and Systems* Vol. 56(1993), pp. 171-174.

[8] Z. Qin, J. Lawry, LFOIL:Linguistic rule induction in the label semantics framework, *Fuzzy sets and systems* Vol. 159(2008), pp. 435-448.

[9] Z. Qin, J. Lawry, Decision tree learning with fuzzy labels, *Inform. Sci.* 172(1-2)(2005), pp. 91-129.

[10] E. Szmidt, J. Kacprzyk, Distances between intuitionistic fuzzy sets, *Fuzzy Sets and Systems* Vol. 114(2000), pp. 505-518.

[11] L.A. Zadeh, Fuzzy Logic=Computing with Words, *IEEE Transanction on Fuzzy Systems* Vol. 4(1996), pp. 103-111.

[12] L.A. Zadeh, The Concept of Linguistic Variable and its Application to Approimate Reasoning Part2, *Information Science* Vol. 8(1975), pp. 301-357.

[13] L.A. Zadeh, Fuzzy sets, *Inform. Control* Vol. 8(1965), pp. 338-356.