第6卷第2期　　空军工程大学学报(自然科学版)　　Vol.6 No.2
2005年4月　　JOURNAL OF AIR FORCE ENGINEERING UNIVERSITY(NATURAL SCIENCE EDITION)　　Apr. 2005

直觉模糊语义匹配的相似性度量

雷英杰，　赵　晔，　王　涛，　王　坚，　申晓勇

(空军工程大学导弹学院，陕西三原 713800)

摘　要：首先引用Atanassov直觉模糊集的基本概念和运算，综合考虑其隶属度函数与非隶属度函数两个因素的影响，通过对Zadeh模糊集相似度的计算模型进行扩展，给出了直觉模糊语义匹配的相似度的几种扩展计算方法，即最大最小法、算术平均最小法、几何平均最小法、相关系数法、指数法等．并举例阐明其应用，从而使直觉模糊语义匹配度的计算方法得到进一步拓广。最后，研究‘了直觉模糊相似关系，证明了直觉模糊相似矩阵的一个定理。

关键词：直觉模糊集合；直觉模糊逻辑；语义匹配；相似度

中图分类号：TP182　文献标识码：A　文章编号：1009-3516(2005)02-0083-04

在语义描述上，经典的康托尔(Cantor)集合论只能描述“非此即彼”的“分明概念”。Zadeh的模糊集理论[1]则可以扩展描述外延不分明的“亦此亦彼”的“模糊概念”。Atanassov直觉模糊集(Intuitionistic Fuzzy Sets,IFS)[2-5]，是对Zadeh模糊集理论最有影响的一种扩充和发展。IFS增加了一个新的属性参数——非隶属度函数，进而还可以描述“非此非彼”的中立状态，更加细腻地刻画客观世界的模糊性本质，因而引起众多学者的关注。

1　直觉模糊集

Atanassov对直觉模糊集给出下述定义。

定义1(直觉模糊集)　设X是一个给定论域，则X上的一个直觉模糊集A为

$$A=\{\langle x,\mu_A(x),\gamma_A(x)\rangle \mid x\in X\}$$

其中$\mu_A(x):X\to[0,1]$和$\gamma_A(x)$分别代表A的隶属函数$\mu_A(x)$和非隶属函数$\gamma_A(x)$，且对于A上的所有$x\in X$，$0\leqslant\mu_A(x)+\gamma_A(x)\leqslant 1$成立。

当X为连续空间时，$A=\int_X\langle\mu_A(x),\gamma_A(x)\rangle/x, x\in X$；当$X=\{x_1,x_2,\cdots,x_n\}$为离散空间时，$A=\sum_{i=1}^{n}\langle\mu_A(x),\gamma_A(x)\rangle/x, x\in X$；当$X=\{1,2,\cdots,n\}$。直觉模糊集$A$有时可以简记作$A=\langle x,\mu_A,\gamma_A\rangle$。显然，每一个一般模糊子集对应于下列直觉模糊子集$A=\{\langle x,\mu_A(x),1-\mu_A(x)\rangle \mid x\in X\}$。

对于X中的每一个直觉模糊子集，称$\pi_A(x)=1-\mu_A(x)-\gamma_A(x)$为$A$中$x$的直觉指数(Intuitionistic Index)，它是x对A的犹豫程度(Hesitancy degree)的一种测度。显然，对于每一个$x\in X$，$0\leqslant\pi_A(x)\leqslant 1$。对于$X$中的每一个一般模糊子集$A$，$\pi_A(x)=1-\mu_A(x)-[1-\mu_A(x)]=0,\forall x\in X$。

2　直觉模糊相似度

在模糊推理中，由于知识的前提条件中的A与证据中的A'不一定完全相同，因此在决定选用哪条知识进行推理时必须首先考虑哪条知识的前提条件A与证据中的前提条件A'近似匹配的问题，即它们的相似程度

收稿日期：2004-09-10
基金项目：国防科技预研基金资助项目(51406030104DZ0120)
作者简介：雷英杰(1956-)，男，陕西渭南人，教授，博士生导师，主要从事智能信息处理与智能决策等研究.

是否大于某个预先设定的阈值λ。由于A与A'都是用相应的直觉模糊集及其属性函数刻画的,因此对其相似程度的计算就转化为对其相应直觉模糊集的计算。

两个直觉模糊集所表示的模糊概念的相似程度称为匹配度。目前,常用的计算匹配度的方法主要有贴近度、语义距离及相似度。下面集中讨论相似度的计算。

相似度的含义是相似度越大说明两者越匹配。计算相似度通常有最大最小法、算术平均最小法、几何平均最小法、相关系数法、指数法等方法,它们分别适用于不同的场合。同样,在计算直觉模糊集的相似度时,需要考虑其隶属度函数与非隶属度函数两个因素的合成,因而也必须对一般模糊集的相似度的计算进行相应扩展。下面给出扩展的直觉模糊集的几种相似度计算模型。

设$A, B \in IFS(U), x_i \in U, i = 1, 2, \cdots n$,且

$$A = \langle \mu_A(x_1), \gamma_A(x_1) \rangle / x_1 + \langle \mu_A(x_2), \gamma_A(x_2) \rangle / x_2 + \cdots + \langle \mu_A(x_n), \gamma_A(x_n) \rangle / x_n$$

$$B = \langle \mu_B(x_1), \gamma_B(x_1) \rangle / x_1 + \langle \mu_B(x_2), \gamma_B(x_2) \rangle / x_2 + \cdots + \langle \mu_B(x_n), \gamma_B(x_n) \rangle / x_n$$

A与B之间的相似度记为$r(A, B)$。

1) 最大最小法

$$r(A, B) = \frac{1}{2}\left\{ \frac{\sum_{i=1}^{n} \min(\mu_A(x_i), \mu_B(x_i))}{\sum_{i=1}^{n} \max\{\mu_A(x_i), \mu_B(x_i)\}} + \frac{\sum_{i=1}^{n} (1 - \max(\gamma_A(x_i), \gamma_B(x_i)))}{\sum_{i=1}^{n} (1 - \min\{\gamma_A(x_i), \gamma_B(x_i)))} \right\}$$

例如,设$U = \{a, b, c, d\}$,且

$$A = <0.3, 0.7>/a + <0.4, 0.6>/b + <0.6, 0.4>/c + <0.8, 0.2>/d$$

$$B = <0.2, 0.8>/a + <0.5, 0.5>/b + <0.6, 0.4>/c + <0.7, 0.3>/d$$

则

$$r(A,B) = \frac{0.3 \wedge 0.2 + 0.4 \wedge 0.5 + 0.6 \wedge 0.6 + 0.8 \wedge 0.7}{2(0.3 \vee 0.2 + 0.4 \vee 0.5 + 0.6 \vee 0.6 + 0.8 \vee 0.7)} + \frac{4 - (0.7 \vee 0.8 + 0.6 \vee 0.5 + 0.4 \vee 0.4 + 0.2 \vee 0.3)}{2(4 - (0.7 \wedge 0.8 + 0.6 \wedge 0.5 + 0.4 \wedge 0.4 + 0.2 \wedge 0.3))} = 0.86$$

2) 算术平均最小法

$$r(A, B) = \frac{\sum_{i=1}^{n} \min[\mu_A(x_i), \mu_B(x_i)]}{\sum_{i=1}^{n} [\mu_A(x_i) + \mu_B(x_i)]} + \frac{\sum_{i=1}^{n} \min[1 - \gamma_A(x_i), \gamma_B(x_i)]}{\sum_{i=1}^{n} [(1 - \gamma_A(x_i)) + (1 - \gamma_B(x_i))]}$$

对于上例,$r(A, B) = 0.93$。

3) 几何平均最小法

$$r(A, B) = \frac{\sum_{i=1}^{n} \min(\mu_A(x_i), \mu_B(x_i))}{2 \times \sum_{i=1}^{n} \sqrt{(\mu_A(x_i) \times \mu_B(x_i)}} + \frac{\sum_{i=1}^{n} \min(1 - \gamma_A(x_i), 1 - \gamma_B(x_i))}{2 \times \sum_{i=1}^{n} \sqrt{((1 - \gamma_A(x_i)) + (1 - \gamma_B(x_i)))}}$$

对于上例,$r(A, B) = 0.93$。

4) 相关系数法

$$r(A, B) = \frac{\sum_{i=1}^{n} |\mu_A(x_i) - \bar{\mu}_A| \times |\mu_B(x_i) - \bar{\mu}_B|}{2 \times \sqrt{[\sum_{i=1}^{n} (\mu_A(x_i) - \bar{\mu}_A)^2] \times [\sum_{i=1}^{n} (\mu_B(x_i) - \bar{\mu}_B)^2]}} + \frac{\sum_{i=1}^{n} |\varphi_A(x_i) - \bar{\varphi}_A| \times |\varphi_B(x_i) - \bar{\varphi}_B|}{2 \times \sqrt{[\sum_{i=1}^{n} (\varphi_A(x_i) - \bar{\varphi}_A)^2] \times [\sum_{i=1}^{n} (\varphi_B(x_i) - \bar{\varphi}_B)^2]}}$$

式中

$$\bar{\mu}_A = \frac{1}{n}\sum_{i=1}^{n} \mu_A(x_i), \bar{\varphi}_A = \frac{1}{n}\sum_{i=1}^{n} \varphi_A(x_i), \varphi_A(x_i) = 1 - \gamma_A(x_i)$$

$$\bar{\mu}_B = \frac{1}{n}\sum_{i=1}^{n}\mu_B(x_i), \bar{\varphi}_B = \frac{1}{n}\sum_{i=1}^{n}\varphi_B(x_i), \varphi_B(x_i) = 1 - \gamma_B(x_i)$$

对于上例,$r(A, B) = 0.9$。

5) 绝对值指数法

$$r(A, B) = \exp\{-\frac{1}{2}[\sum_{i=1}^{n}|\mu_A(x_i) - \mu_B(x_i)| + \sum_{i=1}^{n}|\gamma_A(x_i) - \gamma_B(x_i)|]\}$$

对于上例,$r(A, B) = 0.74$。

此外还有数量积法、夹角余弦法、绝对值倒数法、绝对值减数法等,都可以采用这种扩展方法,在考虑了非隶属度的影响之后,扩展为直觉模糊相似度计算模型。

无论用那种方法计算出的相似度,都可以将其转换为相应的匹配度。根据相似度的含义,我们不妨设匹配度

$$\delta_{\text{match}}(A, B) = r(A, B)$$

这里的相似度 $r(A, B)$ 应是扩展的直觉模糊集相似度。当相似度或匹配度大于某个预先指定的阈值 λ 时,就认为相应的模糊条件可与证据匹配。

3　直觉模糊相似关系

直觉模糊关系也是一种直觉模糊集合,但其论域是 N 个集合的叉积。在基于直觉模糊集理论的知识处理中,直觉模糊关系十分有用。

定义 2(直觉模糊关系)　设 X 和 Y 是普通、有限、非空集合或论域。定义在直积空间 $X\times Y$ 上的直觉模糊子集称为从 X 到 Y 之间的二元直觉模糊关系。记为

$$\boldsymbol{R} = \{\langle (x,y), \mu_R(x,y), \gamma_R(x,y)\rangle \mid x\in X, y\in Y\}$$

其中 $\mu_R: X\times Y\to[0,1]$ 和 $\gamma_R: X\times Y\to[0,1]$ 满足条件 $0\leqslant\mu_R(x,y)+\gamma_R(x,y)\leqslant 1$, $\forall (x,y)\in X\times Y$。

用 $IFR(X\times Y)$ 来表示 $X\times Y$ 上的直觉模糊子集的全体。若 X 和 Y 为有限集时,即 $X=\{x_1,x_2,\cdots,x_m\}$, $Y=\{y_1,y_2,\cdots,y_n\}$,则从 X 到 Y 之间的二元直觉模糊关系 R 可以用矩阵表示。对于 $\forall (x_i,y_j)\in X\times Y$ $(i=1,2,\cdots,m; j=1,2,\cdots,n)$,记为 $(\mu_{ij})_{m\times n}$ 和 $(\gamma_{ij})_{m\times n}$,其中 $\mu_{ij}=\mu_R(x_i,y_j)$, $\gamma_{ij}=\gamma_R(x_i,y_j)$, $0\leqslant\mu_{ij}\leqslant 1$, $0\leqslant\gamma_{ij}\leqslant 1$, $(i=1,2,\cdots,m; j=1,2,\cdots,n)$ 分别称为元素 x_i 与 y_j 之间关系 R 存在的程度和不存在的程度。

若 $X_1,X_2,\cdots,X_n$ 是 n 个集合,则所谓直积空间 $X_1\times X_2\times\cdots\times X_n$ 上的一个 n 元直觉模糊关系 R 是指 $X_1\times X_2\times\cdots\times X_n$ 上的一个直觉模糊子集。记为

$$\boldsymbol{R} = \{\langle (x_1,x_2,\cdots,x_n), \mu_R(x_1,x_2,\cdots,x_n), \gamma_R(x_1,x_2,\cdots,x_n)\rangle \mid x_i\in X_i, i=1,2,\cdots,n\}$$

其中 $\mu_R: X_1\times X_2\times\cdots\times X_n\to[0,1]$ 和 $\gamma_R: X_1\times X_2\times\cdots\times X_n\to[0,1]$ 满足条件

$$0\leqslant\mu_R(x_1,x_2,\cdots,x_n)+\gamma_R(x_1,x_2,\cdots,x_n)\leqslant 1, \forall (x_1,x_2,\cdots,x_n)\in X_1\times X_2\times\cdots\times X_n。$$

由以上定义可以看出,直觉模糊关系是一般模糊关系的一种推广。

定义 3　设直觉模糊关系 $\boldsymbol{R}\in IFR(X\times Y)$ 是

a) 自反的,若 $\forall x\in X, \mu_R(x,x)=1, \gamma_R(x,x)=0$

b) 逆自反的,若 $\forall x\in X, \mu_R(x,x)=0, \gamma_R(x,x)=1$。即 $\boldsymbol{R}^c$ 是自反的。

c) 对称的,若 $\boldsymbol{R}=\boldsymbol{R}^{-1}$。即 $\forall (x,y)\in X\times X$, $\mu_R(x,y)=\mu_R(y,x)$, $\gamma_R(x,y)=\gamma_R(y,x)$。

定义 4　设 $\boldsymbol{R}$ 是 X 上的直觉模糊关系 $\boldsymbol{R}\in IFR(X\times Y)$,

a) 若 $\boldsymbol{R}$ 是 X 上自反、对称关系,则称 $\boldsymbol{R}$ 是 X 上的直觉模糊相似关系,简称相似关系。当 X 为有限集时,称 $\boldsymbol{R}$ 为直觉模糊相似矩阵。

b) 若 $\boldsymbol{R}$ 是 X 上自反、对称、传递关系,则称 $\boldsymbol{R}$ 是 X 上的直觉模糊等价关系,简称等价关系。当 X 为有限集时,称 R 为直觉模糊等价矩阵。

引理 1　设直觉模糊关系 $\boldsymbol{R}\in IFR(X\times Y)$, $S\in IFR(Y\times Z)$, $T\in IFR(Z\times W)$,则对直觉模糊关系合成运算,结合律成立,即:$(\boldsymbol{R}\circ S)\circ T=\boldsymbol{R}\circ(S\circ T)$。

定理 1　若 $\boldsymbol{R}=[\langle\mu_{ij},\gamma_{ij}\rangle]_{m\times m}$ 是直觉模糊相似矩阵,则对任意自然数 k, $\boldsymbol{R}^k$ 也是直觉模糊相似矩阵。

证明　可用数学归纳法证明:

当 $k=1$ 时,$\boldsymbol{R}^k=\boldsymbol{R}$ 是直觉模糊相似矩阵;

假设当 $k=n$ 时,$\boldsymbol{R}^n$是直觉模糊相似矩阵。

以下证明当 $k=n+1$ 时,$\boldsymbol{R}^k=\boldsymbol{R}^{n+1}$也是直觉模糊相似矩阵。

记 $\boldsymbol{R}^n=[\langle\mu_{ij}^{(n)},\gamma_{ij}^{(n)}\rangle]_{m\times m}$,$\boldsymbol{R}^{n+1}=[\langle\mu_{ij}^{(n+1)},\gamma_{ij}^{(n+1)}\rangle]_{m\times m}$,因为 $\boldsymbol{R}^{n+1}=\boldsymbol{R}^n\circ\boldsymbol{R}$,且 $\boldsymbol{R}^n$ 与 $\boldsymbol{R}$ 都自反,所以

$$\mu_{ii}^{(n+1)}=\bigvee_{p=1}^{m}(\mu_{ip}^{(n+1)}\wedge\mu_{pi})\geqslant\mu_{ii}^{(n)}\wedge\mu_{ii}=1,$$

$$\gamma_{ii}^{(n+1)}=\bigwedge_{p=1}^{m}(\gamma_{ip}^{(n+1)}\vee\gamma_{pi})\geqslant\gamma_{ii}^{(n)}\vee\gamma_{ii}=0,$$

由定义 3 可知,$\boldsymbol{R}^{n+1}$具有自反性.

因为 $\boldsymbol{R}^{n+1}$与 $\boldsymbol{R}$ 是对称的,故

$$\mu_{ij}^{(n+1)}=\bigvee_{p=1}^{m}(\mu_{ip}^{(n)}\wedge\mu_{pj})=\bigvee_{p=1}^{m}(\mu_{pi}^{(n)}\wedge\mu_{jp})=\bigvee_{p=1}^{m}(\mu_{jp}\wedge\mu_{pi})=\mu_{ji}^{(n+1)}$$

$$\gamma_{ij}^{(n+1)}=\bigvee_{p=1}^{m}(\gamma_{ip}^{(n)}\wedge\gamma_{pj})=\bigvee_{p=1}^{m}(\gamma_{pi}^{(n)}\wedge\gamma_{jp})=\bigvee_{p=1}^{m}(\gamma_{jp}\wedge\gamma_{pi})=\gamma_{ji}^{(n+1)}$$

这里运用了直觉模糊合成运算的结合律(引理 1)

$$\boldsymbol{R}^{n+1}=\boldsymbol{R}^n\circ\boldsymbol{R}=\boldsymbol{R}\circ\boldsymbol{R}^n$$

即

$$\mu_R{}^{(n+1)}(x_i,x_j)=\bigvee_{p=1}^{m}(\mu_R{}^{(n)}(x_i,x_p)\wedge\mu_R(x_p,x_j))=\bigvee_{p=1}^{m}\mu_R(x_i,x_p)\wedge\mu_R{}^{(n)}(x_p,x_j))$$

$$\gamma_R{}^{(n+1)}(x_i,x_j)=\bigvee_{p=1}^{m}(\gamma_R{}^{(n)}(x_i,x_p)\wedge\gamma_R(x_p,x_j))=\bigvee_{p=1}^{m}\gamma_R(x_i,x_p)\wedge\gamma_R{}^{(n)}(x_p,x_j))$$

由定义 3 可知,$\boldsymbol{R}^{n+1}$具有对称性.

由定义 4 可知,$\boldsymbol{R}^{n+1}$是直模糊相似矩阵。证毕。

4 结束语

Zadeh 模糊集理论及应用,特别是在知识处理中的应用虽然也在进一步发展但已趋成熟[8],而 Atanassov 直觉模糊集理论用作知识处理领域,尚正在发展之中,且其数学描述较之 Zadeh 模糊集理论更加符合客观世界模糊对象的本质,因而形成新的研究热点。

本文在引用直觉模糊集基本概念基础上,综合考虑其隶属度函数与非隶属度函数两个因素的影响,通过对 Zadeh 模糊集相似度的计算模型进行扩展,给出了直觉模糊语义匹配相似度的几种扩展计算方法,包括最大最小法、算术平均最小法、几何平均最小法、相关系数法、指数法,并举例阐明其应用,从而使直觉模糊语义匹配度的计算方法得到进一步拓广。最后,研究了直觉模糊相似关系,证明了直觉模糊相似矩阵的一个定理。

参考文献:

[1] Zadeh L A. Fuzzy Sets[J]. Lnformation and Control, 1965, 8 (3) : 338 -353.

[2] Atanassov K. Intuitionistic Fuzzy Sets [J]. Fuzzy Sets and Systems, 1986, 20 (1) : 87 -96.

[3] Atanassov K. More on Intuitionistic Fuzzy Sets [J]. Fuzzy Sets and Systems, 1989, 33 (1) : 37 -46.

[4] Atanassov K. New Operations Defined Over the Intuitionistic Fuzzy Sets [J]. Fuzzy Sets and Systems, 1994, 61 (1) : 137 - 142.

[5] Atanassov Krassimir T, Kacprzyk Janusz, Szmidt Eulalia, et al. On Separability of Intuitionistic Fuzzy Sets[J]. Lecture Notes in Artificial Intelligence, 2003, 2715 : 285 -292.

[6] 雷英杰,王宝树. 直觉模糊逻辑的语义算子研究[J]. 计算机科学, 2004, 31 (11) : 4 -6.

[7] 雷英杰,孙金萍,王宝树. 模糊知识处理与模糊集理论的若干拓展[J]. 空军工程大学学报(自然科学版), 2004, 5(3) : 40 -44.

[8] 雷英杰,王宝树. 拓展模糊集之间的若干等价变换[J]. 系统工程与电子技术, 2004, 26 (10): 1414 -1417.

(下转第 91 页)