文章编号:1000－2367(2015)05－0153－06 DOI:10.16366/j.cnki.1000－2367.2015.05.029

基于粒计算的融合性贴近度方法

吕 康[1],魏培文[2],张 辉[3]

(1.河南教育学院 信息技术系,郑州 450002;2.河南师范大学 新联学院,河南 新乡 453007;
3.河南师范大学 新联学院,郑州 450002)

摘 要:在粒计算的基础上,将粒进行了形式化表示,详细地讨论了粒如何进行度量,进一步对粒在相似度度量方面加以阐述.对不同的事物相同标准和相同事物不同度量标准这两种情况的相似度度量进行探讨,并提出了一种新的基于粒计算的融合性贴近度方法对不同相似性度量方法进行了统一,减少了中间过程,避免了标准不同不能进行比较的问题.

关键词:粒计算;公式;度量;相似度

中图分类号:TP3－0 **文献标志码**:A

人们看世界的方式一般都是从粒的角度开始的,并且人类认识世界和处理问题的方式和方法都是由粒的方式进行的.因此,可以认为粒广泛存在于现实世界中,是对现实世界的一种抽象.那么我们就要考虑,如何能让计算机模拟人的思维对待现实世界和现实问题,并能利用计算机计算、解决并管理相应的项目和内容呢?由于粒计算的基本思想是在问题求解中使用粒,它通过现实问题的抽象、粒之间的关系、粒的分解和合成及粒或者粒集间的转换来描述和解决问题[1-3].因此,考虑到使用粒来解决计算机中的一些不易解决和处理的问题,模拟人类思维对复杂问题进行相应求解,进而达到使用计算机解决现实世界中复杂问题的目的.

粒计算的模型有多种,概括起来:模糊集、商空间、粗糙集、或概念格等,对粒的定义也是多种多样,这其实并不利于粒的长远发展,因此,如何为粒提供一种统一的表示形式和含义解释,如何为粒计算的发展建立一个统一的理论框架等已经成为粒计算研究的热点.目前在这方面进行研究的人很多,大家始终不能达成统一,但一般还是比较认可以粒计算的公式化表达形式来统一各个模型的[4-9]].当然,一个理论体系的形成也不是一朝一夕的,需要各个方面知识的积累和构造.本文就使用公式对粒进行形式化表示进行研究,并就粒与粒之间的相各种贴近度进行对比性研究.

1 粒的形式化表示

基于公式的粒计算形式化方法是粒计算中最基础但也是最关键的问题,只有统一的表示出粒及粒计算才能从理论上统一各个模型.如何能够将粒形式化表示出来,又能涵盖粒计算的所有框架是个关键问题,这里我们考虑使用公式对粒和粒值等进行合理描述,具体表示如下:

首先,设论域 U 为有限集合.考虑 n 个 U 的笛卡儿积 $U\times\cdots\times U$,并简记为 $U_n(n\geq 1)$.这里 U_n 中的元素称为 n 元组,记作 $<a_1,\cdots,a_n>$,其中 $a_1,\cdots,a_n\in U$.当 $H\subseteq U_n$ 时,H 称为 U 上 n 元关系.

定义 1[5] (原子公式)设 U 为论域,$P\subseteq U_n$ 是 n 元关系,$t_1,\cdots,t_n$ 是 n 个项,称 $P(t_1,\cdots,t_n)$ 为 U 上的原子公式,记作 ϕ.

由原子公式和逻辑联结词可定义更复杂的合式公式,而合式公式可以用于解决更加复杂的问题,表述更加准确的信息.

收稿日期:2015-01-22;**修回日期**:2015-06-15.

基金项目:国家自然科学基金(U1204609);河南省教育厅科学技术研究重点项目(13A520223;13B520087).

第 1 作者简介(通信作者):吕 康(1982－),男,河南新乡人,河南教育学院讲师,研究方向为数据挖掘、粗糙集,E-mail:24295376@qq.com.

定义 2(合式公式)[5] 设 U 为论域，信息系统中，论域 U 上的合式公式可归纳定义如下：

1)U 上的原子公式 ϕ 是 U 上的公式；

2) 若 ϕ 是 U 上的公式，则 $\neg\phi$ 也是 U 上的公式；

3) 若 ϕ,φ 是 U 上的公式，则 $\phi \wedge \varphi, \phi \vee \varphi, \phi \rightarrow \varphi$ 均为 U 上的公式；

4) 有限步使用 1),2) 或 3) 得到的式子均是 U 上的合式公式.

定义 2 可以语义解释为：令 $F(U)$ 表示 U 上所有公式的集合，那么，由于经典命题逻辑公式中并没有出现 $n(n \geq 1)$ 元关系，而经典一阶逻辑公式中又含有量词，所以定义 1 中的公式其实是介于命题逻辑和一阶逻辑之间的一种原子公式[4-9]. 如果公式 ϕ 中出现 n 个变量 $x_1,\cdots,x_n$，则称 ϕ 为 U 上的 n 元合式公式. 有时为了强调这 n 个变量，可以将合式公式 ϕ 表示为 $\phi(x_1,\cdots,x_n)$.

定义 3(粒值) 粒值 v 分为原子粒值和合式粒值. 粒值是在论域 U 中，使用公式 ϕ 进行计算所得值 v，即为粒值. 原子粒值的类型可以是实数型、字符型和布尔型等基本的数据类型.

定义 4(粒) 设 U 为论域，ϕ 为论域上公式，v 为在论域 U 中由公式 ϕ 计算所得粒值，m 为 ϕ 计算后形成的粒集合，那么二元对 $(\phi,m(\phi,v))$ 称为信息系统 IS 中的粒. 简记为 m.

定义 5(粒集)[10-11] 设公式 $\psi = \{\phi_1,\phi_2,\cdots,\phi_r\}(i = 1,2,\cdots,n)$ 是由一系列原子公式 ϕ_i 构成的公式集合. 集合 $\widetilde{P} \subseteq U, \widetilde{P} = \{P_1, P_2,\cdots, P_m\}$，其中 $P_k = \{p_{k1},p_{k2},\cdots,p_{ks}\}(k = 1,2,\cdots,m$，$s$ 表示每个 P_k 中包含 s 个特征信息)，C_j 表示层次 j，$P_k' \in C_j$ 表示值 P_k' 达到层次 j，那么集合 $\widetilde{P}$ 中的各向量对象的特征信息通过公式 ϕ_i 计算后所得值达到层次 j 的相应的向量对象的集合 $m_{\widetilde{P}}^{j}(\phi_i)$：

$m_{\widetilde{P}}^{j}(\phi_i) = \{P_k \mid P_k' \in m_{\widetilde{P}}^{j}(\phi_i), P_k' \in C_j, P_k \in \widetilde{P}, P_k \rightarrow P_k', k = 1,2,\cdots,m\}, j = 1,2,\cdots,h$；$h$ 表示层次的个数，这里层次之间假定不交叉.

当 $i = 1$ 时，表示一维情况下的粒集合. 当 $i = n$ 时，表示高维(n 维) 情况下的粒集合.

定义 6(隶属度)[10-11] 设论域 U 上的向量对象的集合 $\widetilde{P} = \{P_1,P_2,\cdots, P_m\}(\widetilde{P} \subseteq U)$，$\forall \phi_i \in \Psi$，$\Psi$ 是公式集合$(i = 1,2,\cdots,n)$，$m_{\widetilde{P}}^{j}(\phi_i)$ 隶属于集合 $\widetilde{P}$ 的程度定义为

$$\mu_{\widetilde{P}}^{j}(\phi_i) = \frac{|m_{\widetilde{P}}^{j}(\phi_i) \cap \widetilde{P}|}{|\widetilde{P}|}(i = 1,2,\cdots,n;j = 1,2,\cdots,h) \text{ 显然，} 0 \leq \mu_{\widetilde{P}}^{j}(\phi_i) \leq 1, \sum_{j=1}^{h}\mu_{\widetilde{P}}^{j}(\phi_i) = 1.$$

2 粒的度量

2.1 粒的相似度度量

粒的度量是一个相对的概念也是一个基本问题. 考察某一论域 U 中用一个公式做的度量是否精确就要看是否有比该公式划分得更细的度量，还值得注意的是要考虑这个更细的度量是否具有实际意义. 如何衡量这个度量“粗细”呢?当一个公式划分的度量集合代表的单位尺度是论域中对对象属性测度的有意义的最小尺度时，那么它对应的度量就不是一个粗的度量，反之同理.

在同一个论域 U 中，可以利用各种原子或者合式公式构造相应粒集. 于是，一个公式 ϕ，就可以得到一个相应的粒集 $m(\phi,v)$. 如果有两个公式 ϕ,φ，用符号 $\phi \Rightarrow \varphi$ 来表示 ϕ,φ 之间的联系，则 $\phi \Rightarrow \varphi$ 的强度可以用下面两种测度定量描述：

$$m(\phi) \cap m(\varphi) = m(\phi \wedge \varphi), m(\phi) \cup m(\varphi) = m(\phi \vee \varphi).$$

一个公式 ϕ 的单粒 $m(\phi)$ 的度量是指粒 $m(\phi)$ 的相对大小，用 $G(\phi) = \dfrac{|m(\phi)|}{|U|}$ 表示. 这种方式可以用来比较不同公式计算后所得粒的大小度量. 但粒的度量包含的还有很多方面其他的方面，比如度量论域中事物的相似性[10-11].

判断两个事物的相似程度，往往不是一个简单的事情，事物的属性多种多样，属性之间的关系也千变万化，相同的属性个数不同的组合得到的语义往往也不尽相同. 首先，属性间的关系及其组合的度量办法已是目前迫切需要解决的问题. 其次，可以对两种组合分别进行相似性度量，判断这组数据的真实性，等等. 因此，如何在海量数据中选择多特征属性进行行之有效的组合是数据挖掘的重要目标.

2.2　基于粒计算的贴近度

定义 7[10-11]　集合 $\widetilde{P}$ 和 $\widetilde{Q}$ 之间的多层次距离定义为:

$$d(\widetilde{P},\widetilde{Q}) = \Big(\sum_{j=1}^{h}\sum_{i=1}^{r} \mid \mu_{\widetilde{P}}^{j}(\phi_i) - \mu_{\widetilde{Q}}^{j}(\phi_i) \mid^{\lambda}\Big)^{\frac{1}{\lambda}}, \tag{1}$$

其中 $m_{\widetilde{P}}^{j}(\phi_i)$ 为隶属于集合与集合的程度,且 $0 < m_{\widetilde{P}}^{j}(\phi_i) < 1$,$\lambda$ 为适当参数,r 为公式个数,h 为层次个数.

距离反映了对象间的差距程度,也就是不相似的程度.

定义 8[10-11]　设 U 为论域,$\widetilde{P},\widetilde{Q} \in U$,称

$$n(\hat{P},\hat{Q}) = 1 - c[d(\hat{P},\hat{Q})^{\alpha}] \tag{2}$$

为 $\widetilde{P},\widetilde{Q}$ 的贴近度,其中 c 和 α 为适当选择的参数,$d(\widetilde{P},\widetilde{Q})$ 为 $\widetilde{P}$ 与 $\widetilde{Q}$ 间的多层次距离.

若取 $\alpha = \lambda = 1, c = \dfrac{1}{(r\times h)^{\frac{1}{\lambda}}}$ 时就得到多层次海明距离贴近度:

$$n(\widetilde{P},\widetilde{Q}) = 1 - \frac{1}{(r\times h)}\sum_{j=1}^{h}\sum_{i=1}^{r} \mid \mu_{\widetilde{P}}^{j}(\phi_i) - \mu_{\widetilde{Q}}^{j}(\phi_i) \mid . \tag{3}$$

使用这种方法进行相似性度量,有效地解决了各种公式标准不统一的复杂情况,任何公式情况的组合都可以使用这种度量相似度的方法,避免了标准不同不能进行比较的问题.并且多种属性组合,若语义相互关联,可通过计算判断该数据的真实程度.

2.3　基于粒计算的格贴近度

数据分为连续型和离散型两种,很多数据不适合使用简单的加减乘除来计算,因此提出了另外一种相似性的度量标准,即格贴近度,它的计算方法采用交并方式,在某些计算方面有突出优势.其计算方法如下.

定义 9[10-11]　设 U 为论域,$\widetilde{P},\widetilde{Q} \in U$,$\Psi$ 为 U 上的公式集合,$\phi_i \in \Psi$,记

$$m_{\widetilde{P}}^{j}(\phi_i) \circ m_{\widetilde{Q}}^{j}(\phi_i) \triangleq \frac{1}{n}\sum_{i=1}^{n}[\bigvee_{j=1}^{h}(\mu_{\widetilde{P}}^{j}(\phi_i) \wedge \mu_{\widetilde{Q}}^{j}(\phi_i))], \tag{4}$$

$$m_{\widetilde{P}}^{j}(\phi_i) \otimes m_{\widetilde{Q}}^{j}(\phi_i) \triangleq \frac{1}{n}\sum_{i=1}^{n}[\bigwedge_{j=1}^{h}(\mu_{\widetilde{P}}^{j}(\phi_i) \vee \mu_{\widetilde{Q}}^{j}(\phi_i))], \tag{5}$$

其中,$\wedge$ 表示取小,$\vee$ 表示取大,称 $m_{\widetilde{P}}^{j}(\phi_i) \circ m_{\widetilde{Q}}^{j}(\phi_i)$,$m_{\widetilde{P}}^{j}(\phi_i) \otimes m_{\widetilde{Q}}^{j}(\phi_i)$ 分别为 U 上粒集 $m_{\widetilde{P}}^{j}(\phi_i)$ 和 $m_{\widetilde{Q}}^{j}(\phi_i)$ 的内积和外积,为方便,分别简记为$(\widetilde{P} \circ \widetilde{Q})$ 和$(\widetilde{P} \otimes \widetilde{Q})$.

定义 10[10-11]　设粒集 $m_{\widetilde{P}}^{j}(\phi_i)$ 和 $m_{\widetilde{Q}}^{j}(\phi_i) \in U$,则称

$$n(m_{\widetilde{P}}^{j}(\phi_i), m_{\widetilde{Q}}^{j}(\phi_i))_L = \left\{\frac{1}{n}\sum_{i=1}^{n}[\bigvee_{j=1}^{h}(\mu_{\widetilde{P}}^{j}(\phi_i) \wedge \mu m_{\widetilde{Q}}^{j}(\phi_i))] \wedge \left[\frac{1}{n}\sum_{i=1}^{n}[\bigwedge_{j=1}^{h}(\mu_{\widetilde{P}}^{j}(\phi_i) \vee \mu_{\widetilde{Q}}^{j}(\phi_i))]\right]^{c}\right\}$$

为粒集 $m_{\widetilde{P}}^{j}(\phi_i)$ 和 $m_{\widetilde{Q}}^{j}(\phi_i)$ 之间的格贴近度.

定义 10 也可简记为

$$n(\widetilde{P},\widetilde{Q})_L = (\widetilde{P} \circ \widetilde{Q}) \wedge (\widetilde{P} \otimes \widetilde{Q})^{c}. \tag{6}$$

2.4　基于粒计算的融合的贴近度方法

2.2 和 2.3 部分都是用于度量两个粒集合间的相似程度的,但是它们是两种不同贴近度.它们之间很难做出比较,不能简单地认为孰优孰劣.

定义 11[7]　若映射 $n:\widetilde{P}(U) \times \widetilde{Q}(U) \to [0,1]$,$\forall \widetilde{P}, \forall \widetilde{Q} \in P(U)$ 满足下列条件:

1) $n(\widetilde{P},\widetilde{P}) = 1$;2) $n(\widetilde{P},\widetilde{Q}) = n(\widetilde{Q},\widetilde{P})$;3) $n(U,\Phi) = 0$(Φ 为空集);4) 若 $\widetilde{P} \subseteq \widetilde{Q} \subseteq \widetilde{R}$,则 $n(\widetilde{P},\widetilde{R}) \leq n(\widetilde{P},\widetilde{Q}) \wedge n(\widetilde{Q},\widetilde{R})$,则称 $n(\widetilde{P},\widetilde{Q})$ 为 $\widetilde{P}$ 和 $\widetilde{Q}$ 的贴近度,这里 $P(U)$ 表示 U 的幂集.

两者就理论层面来讲,基于粒计算的贴近度符合贴近度的公理化定义 11 的条件 1)、2)、3)、4);而格贴近度不同于这种贴近度,它并不完全符合贴近度功利化定义 11 中的条件 1.于是这就要求在实际应用中对两个粒集相似度要求极为严格时,即同一集合相似度值必须等于 1 时,就不能使用这种贴近度,只能使用 2.2 中的基于粒计算的贴近度进行计算.但是也并不是说基于粒计算的格贴近度就不能使用或者不便于使用,恰恰相反,因为它计算较为简便反而更加容易被采用.但是不管两者计算结果的值相差多大,每种贴近度对于衡量多个对象间的相似效果相同,对相似的趋势描述的也非常一致和准确.在度量相同事物相似度时可以采取

不同的度量标准,更能精确判断事物间的相似程度.简单地说,基于粒计算的贴近度计算粒集间的相似程度时对于粒集和隶属度的计算结果要求更加严格精准,而基于粒计算的格贴近度只需要计算出两个粒集间的值较大粒集和隶属度就可以,无形中简化了计算的复杂度,一定程度上可以忽略一些缺失信息,避免计算的不准确性.因此,当数据集中一些不关键数据缺失时,可以考虑使用基于粒计算的格贴近度进行度量粒集间的相似度,当数据集对数值要求更加精准时,可以考虑使用基于粒计算的贴近度进行度量粒集间的相似度.

为了统一度量结果,提出了一种新的度量方法,使用求平均的方式融合两种贴近度的结果.

定义 12

$$N(\tilde{P},\tilde{Q}) = \frac{1}{2}[n(\tilde{P},\tilde{Q}) + n(\tilde{P},\tilde{Q})_L]. \tag{7}$$

其中,N 代表基于粒计算的融合后的贴近度,n 代表基于粒计算的贴近度,n_L 代表基于粒计算的格贴近度.

3　算法及实例说明

3.1　粒化算法

输入:一个有限集合 U 及粒计算公式;输出:论域 U 中各个对象间的融合贴近度 N.

算法步骤:

步骤 1　在论域 U 中,使用公式 ϕ 进行计算,得到相应对象的粒值和粒集;

步骤 2　对已得向量对象集合中,使用隶属度公式计算各公式计算出的各个对象的某一层次隶属于该向量对象的隶属度,得 $\mu_{\mathrm{P}}^{j}(\varphi_i)$;

步骤 3　根据基于粒计算的海明距离贴近度公式(3),计算出两两对象集合间的贴近度,并输出该贴近度值;

步骤 4　根据基于粒计算的格贴近度的内外积公式(4)、(5) 计算出两两对象间的内外积;

步骤 5　根据基于粒计算的格贴近度的公式(6) 计算出两两对象间的格贴近度;

步骤 6　根据基于粒计算的融合贴近度的公式(7) 计算出两两对象间的融合后的贴近度;

步骤 7　循环步骤 3 — 步骤 7 直到两两对象间的融合贴近度计算完毕,退出循环;

步骤 8　结束.

3.2　实例分析

设 $\tilde{P},\tilde{Q},\tilde{R}$ 分别为某一高校的 3 个班的学生,学生人数分别为 50、20 和 80.现计算各班级学生身体素质是否相似,可以有针对性地合班训练从而提高学生体能.

因高校班级人数及男女生比例差距较大,男女生体能测试项目不同等原因,一般的度量方式不能计算出综合情况的结果,不利于合班训练,使用本文计算方法可以在这种复杂的环境下进行计算.对相应论域及各个对象的粒值、粒集及隶属度计算后得到表 1.

计算各集合之间的相似度的一个原子公式和一个合式公式两个不同标准分别使用本文定义的距离贴近度和格贴近度两种相似度度量标准,根据算法,得到两个对象相似度的计算过程如下:

$d(\tilde{P},\tilde{Q}) = 0.1, n(\tilde{P},\tilde{Q}) = 0.9; \tilde{P} \circ \tilde{Q} = 0.7, (\tilde{P} \otimes \tilde{Q}) = 0.2, n(\tilde{P},\tilde{Q})_L = 0.7.$

$d(\tilde{P},\tilde{R}) = 0.15, n(\tilde{P},\tilde{R}) = 0.85; \tilde{P} \circ \tilde{R} = 0.7, (\tilde{P} \otimes \tilde{R}) = 0.125, n(\tilde{P},\tilde{R})_L = 0.7.$

$d(\tilde{Q},\tilde{R}) = 0.3, n(\tilde{Q},\tilde{R}) = 1 - 2.2/6 = 0.7; \tilde{Q} \circ \tilde{R} = 0.75, (\tilde{Q} \otimes \tilde{R}) = 0.125, n(\tilde{Q},\tilde{R})_L = 0.75.$

基于粒计算的融合后的贴近度结果如下:

$N(\tilde{P},\tilde{Q}) = \frac{1}{2}[n(\tilde{P},\tilde{Q}) + n(\tilde{P},\tilde{Q})_L] = 0.8; N(\tilde{P},\tilde{R}) = \frac{1}{2}[n(\tilde{P},\tilde{R}) + n(\tilde{P},\tilde{R})_L] = 0.775;$

$N(\tilde{Q},\tilde{R}) = \frac{1}{2}[n(\tilde{Q},\tilde{R}) + n(\tilde{Q},\tilde{R})_L] = 0.725.$

从这个例子可以看出相同事物使用不同的标准原子公式和合式公式计算，所得结果能够更加全面地说明问题.使用不同的相似性度量公式来衡量相同事物，其计算结果是有一定差距的，但一般不会偏离实际情况.而且从不同层次上对学生的身体素质进行多个方面综合评价，评价结果客观真实地反映了集合间的相似程度，前两个班的相似程度最大，可以合班训练，计算结果具有全面性和准确性.

表 1 粒集及隶属度计算对照表

		优	良	差
$\tilde{P}$	$m_{\tilde{P}}^{j}(\phi_i)$	10	35	5
	$\mu_{\tilde{P}}^{j}(\phi_i)$	0.2	0.7	0.1
$\tilde{Q}$	$m_{\tilde{Q}}^{j}(\phi_i)$	2	15	3
	$\mu_{\tilde{Q}}^{j}(\phi_i)$	0.1	0.75	0.15
$\tilde{R}$	$m_{\tilde{R}}^{j}(\phi_i)$	10	60	10
	$\mu_{\tilde{R}}^{j}(\phi_i)$	0.125	0.75	0.125

我们不难从中分析出，相似性的判断往往从两个方面考虑：第一，两个事物之间若按照相同标准如何衡量其相似程度.第二，一个事物有两个衡量标准，如何衡量这两个标准的相似程度，如果出现冲突如何解决.这里将两个事物设为对象集合，将衡量标准设为公式.

第一，不同的事物，相同标准.当两个事物按照相同标准进行度量，若度量结果相似，则这两个事物相似.但是，有时，两个事物只在某一方面相似，但在其他很多方面和整体均不相似，若只用一种标准衡量这两个事物的相似度未免片面.例如，两个班级的学生人数相同，体能优秀人数也相同，但良好和差的人数相差较大，若只根据体能优秀人数就判断两个班级的体能情况相似，就会做出错误的判断导致错误的训练结果.解决该问题可以使用多个度量标准进行综合度量.这里需要在事物的多个方面和层次上对两个事物进行综合方面的相似性度量.使用多种原子公式或合式公式进行计算，可以提高相似性度量的精度和准确度，并且通过公式计算来衡量两个事物之间的相似度还可以判断描述该事物的数据的支持程度.

第二，相同事物，不同度量标准.公式作为相似度的衡量标准，其在语义方面既可以是属性也可以是属性组合或者是属性在语义上的关联关系.例如，属性"年龄"和"职称"的组合与属性"职业"和"身体状况"的组合在语义上的关联关系就比较大.这两种属性组合在语义方面达到的效果比较相似，就称这两个公式(度量标准)相似.

为了在计算时避免重复性工作和浪费时间，只需选取其中一组属性组合进行度量，这样做可以降低属性维度，减少工作量，缩短工作时间.然而如何比较这两个公式(两个属性组合)相似程度，需要我们给出基于公式的粒计算相似性的度量方法.这其中需要考虑的情况非常多也较为复杂.

当两个衡量标准均为原子公式时，如何计算相似度；当两个衡量标准均为合式公式时，如何计算相似度；当两个衡量标准一个为原子公式一个为合式公式时，如何计算相似度.以上各种情况下相似性判断方法都采用了一种办法进行计算——基于粒计算的融合贴近度，这种方法使得计算过程简单、方便、有效，更好地解决了冲突问题，更利于决策者快速做出决策.

4 结束语

判断两个事物的相似程度，往往不是一个简单的事情，事物的属性多种多样，属性之间的关系也千变万化，相同的属性个数不同的组合得到的语义往往大相径庭.使用计算机进行判断更是难上加难，不但存在标准不统一，而且存在问题描述不准确等问题.为了能够使用计算机准确且合理地度量事物间的相似程度，本文就不同的事物相同标准和相同事物不同度量标准这两种情况进行探讨，举了贴近度和格贴近度两种相似度的度量标准，对原子公式和合式公式的定义、异同和使用给出了详细说明，并对不同公式的度量方法进行了统一——基于粒计算的融合性贴近度，减少了中间过程，避免了标准不同不能进行比较的问题.最后使用不同的度量方式进行计算并给出了相应算法及计算步骤和结果分析.粒的相似性度量是粒计算研究就的基础也是重点，为以后粒计算的统一和发展打下坚实基础.而粒的相似性度量则是对不同事物的相似性进行讨论，为粒的度量方面做了前期研究工作.

参 考 文 献

[1] Zadeh L A. Fuzzy logic-computing with words[J]. IEEE Transactions on Fuzzy Systems, 1996, 4: 103-110.

[2] Zadeh L A. Towards a theory of fuzzy information granulation and its centrality I human reasoning and fuzzy logic[J]. Fuzzy Sets and Systems,1997,90:111-121.

[3] Zadeh L A. Announcement of GrC[EB/OL]. [2015-01-10]. http://www. cs. uregina. ca/～yyao/GrC/.

[4] 刘 清,孙 辉,王洪发.粒计算研究现状及基于 Rough 逻辑语义的粒计算研究[J].计算机学报,2008,31(4):543-555.

[5] 闫 林,闫 硕.粒计算下的决策系统分解与决策转换[J].计算机工程与应用,2012,48(4):13-39.

[6] 陈 光,钟宁,姚一豫,等.粒计算中粒度转换的运算符[J].计算机科学,2012,33(11A):114-115.

[7] 汪培庄.模糊集合论及其应用[M].上海:上海科学技术出版社,1983.

[8] 李 鸿,马小平.粒的特征及形式化表示研究[J].计算机工程与应用,2011,47(11):3-6.

[9] 李 鸿.粒计算的基本原理研究[J].宿州学院学报,2012,38(12):209-212.

[10] 马媛媛,徐久成,孙 林.基于粒计算贴近度的理论研究[J].计算机科学,2006,33(11A):114-115.

[11] 马媛媛,徐久成,孙 林.基于粒计算格贴近度的理论研究[J].河南师范大学学报(自然科学版),2007,35(01):48-50.

Fusion Closeness Degree Method Based on Granular Computing

LYU Kang[1], WEI Peiwen[2], ZHANG Hui[3]

(1. Department of Information Technology, Henan Institute of Education, Zhengzhou 450002, China; 2. Xinlian College, Henan Normal University, Xinxiang 453007, China; 3. Xinlian College, Henan Normal University, Zhengzhou 450002, China)

Abstract: In this paper, the formalize express of grains that based on the granular computing. It discusses how to measure grain so that the similar between grains is expounded. How to measure the similar both the different thing the same standard and the same thing the different standard is discussed. A new method of fusion closeness degree based on granular computing is proposed to unify the different similarity measurement methods, the pilot process is reduce. So the problem that the different standard can't to be compared is avoided.

Keywords: granular computing; formula; measure; close-degree

(上接第 81 页)

Separation of Photogenerated Charges and Photoelectric Properties in Solid State Dye-Sensitized Solar Cells with Zn_2SnO_4 Nanowires

ZHAO Taotao

(School of Physics & Electronics, Henan University, Kaifeng 475004, China)

Abstract: Zn_2SnO_4 nanowires were synthesized on stainless steel filter with hydrothermal method. Zn_2SnO_4 nanowires/$Cu_4Bi_4S_9$ heterojunction was first prepared to increase the separation efficiency of photogenerated charges in composite system. Here, the overall light-to-electricity efficiencies of several Zn_2SnO_4 nanowires/$Cu_4Bi_4S_9$ dye-sensitized solar cells were studied systematically with changing $Cu_4Bi_4S_9$ thickness gradually. With $Cu_4Bi_4S_9$ of 1.0 μm thickness, there are the highest steady state and electric field induced surface photovoltage, and the corresponding Zn_2SnO_4 nanowires/$Cu_4Bi_4S_9$ solar cells exhibits the highest photoelectric conversion efficiency of 4.12%. From the absorption, thickness of film, build-in electric field and energy level matching, etc., the separation process and transport mechanism of photogenerated charges in Zn_2SnO_4 nanowires/$ZnFe_2O_4$ heterojunctions and solid state dye-sensitized solar cells were analysed in detail.

Keywords: surface photovoltage spectroscopy; photovoltaic response; heterojunction; dye-sensitized solar cells.