一种覆盖粗糙 Vague 集模型及其不确定性度量

王 伟[1] 彭进业[1,2] 李 展[1]
(西北大学信息科学与技术学院 西安 710127)[1] (西北工业大学电子信息学院 西安 710072)[2]

摘 要 针对文献[21]提出的覆盖粗糙 Vague 集模型中幂等性并不成立的问题,提出了一种新的基于近邻域的覆盖粗糙 Vague 集模型,并讨论了相关性质及与 I 型覆盖粗糙 Vague 集模型的关系;最后通过引入覆盖粒度空间下知识熵的概念,定义了一种 II 型覆盖粗糙 Vague 集模型的不确定性度量方法。算例分析表明,II 型覆盖粗糙 Vague 模型的不确定性程度随粒度减小而减小。

关键词 覆盖,粗糙 Vague 集,知识熵,不确定性

中图法分类号 TP18 **文献标识码** A

Covering Rough Vague Sets and Uncertainty Measurement

WANG Wei[1] PENG Jin-ye[1,2] LI Zhan[1]
(School of Information Science and Technology, Northwest University, Xi'an 710127, China)[1]
(School of Electronics and Information, Northwestern Poly Technical University, Xi'an 710072, China)[2]

Abstract Aiming at the error of idempotent property of the covering rough Vague sets in paper [21], a new covering rough Vague sets model based on neighbor domain was proposed. This paper discussed some related properties and relationship with the first covering rough Vague sets. Finally, we defined a uncertainty measurement method for covering rough Vague sets based on knowledge entropy in covering granular space. Example analysis shows that the uncertainty degree of the second covering rough Vague sets decreases with the decrease of the granularity.

Keywords Covering, Rough Vague sets, Knowledge entropy, Uncertainty

1 引言

近几年在机器学习领域,基于覆盖关系建立的覆盖粗糙集模型由于其较好的实际应用价值,得到了研究者的广泛关注,涌现出大量的研究成果[5-12]。许多学着将其和其他处理不确定问题的理论相融合,提出了覆盖粗糙模糊集模型等,并研究了相应的不确定度量问题等[13-15]。Vague 集理论作为模糊集的推广,在处理不确定性信息时比传统的模糊集有更强的表达能力且更具灵活性,是一种新型的处理模糊性问题的数学分析工具。有许多学者将其和粗糙集模型结合,提出了新的粗糙 Vague 集模型并研究了相关性质[16-19]。徐久成等人在覆盖粗糙集模型研究的基础上,进一步推广了粗糙 Vague 集模型,提出了新的覆盖粗糙 Vague 集模型[20,21]。该模型是对经典粗糙集理论的推广,可以处理更加模糊的数据,有很重要的实际应用价值。然而实例分析表明,该模型并不满足所提出的幂等性,因此有必要进行修正。

寻求合适的度量来刻画覆盖粗糙 Vague 集的不确定性是该领域研究的一个重要方向。研究发现,不确定程度与覆盖粒度空间粒度大小和集合的覆盖粗糙 Vague 集边界域大小有关。目前覆盖粗糙 Vague 集边界域大小一般用粗糙度来刻画。对覆盖粒度空间粒度大小,不同学者提出了不同的度量方式,包括粒度熵[22]、知识含量测度[23,24]等。处理知识的不确定性的方法往往用香农(Shannon)信息熵[4]来刻画,信息系统的信息熵越大,系统的不确定性越大,而且信息系统的知识粒度越大,系统的不确定性越大。本文从知识熵的角度来度量覆盖粒度空间的不确定性,并结合覆盖粗糙 Vague 集边界域大小,提出了一种基于熵的覆盖粗糙 Vague 集不确定性度量方法。研究表明,对于 II 型覆盖粗糙 Vague 模型,其粒度越小,该度量值越小,表明其不确定性程度越小。

2 现有覆盖粗糙 Vague 集模型及其局限性

为讨论方便,首先介绍相关基本概念。

定义 1[5] 设 U 为非空有限论域,$C=\{X|X\subseteq U\}$ 是 U 的子集簇,若对 $\forall X\in C, X\neq\emptyset$ 且 $\cup C=U$,则称 C 是 U 的一个覆盖,并称 (U,C) 为覆盖近似空间。显而易见,覆盖是划分的一种推广。

定义 2[22] 设 (U,C) 为覆盖近似空间,x 为 U 中的一个对象,则在覆盖近似空间下,x 的描述集为 $Des_c(x)=\{K|$

到稿日期:2011-09-09 返修日期:2011-11-23 本文受高等学校博士学科点专项科研基金(20096102110025),国家自然基金(61075014/F030303)资助。

王 伟(1979-),男,博士生,主要研究方向为机器学习、Vague 集相关理论及应用等,E-mail:szbwangw@126.com;彭进业(1964-),男,博士,教授,博士生导师,主要研究方向为图像处理、人工智能等;李 展(1973-),男,博士生,讲师,主要研究方向为模式识别、SVM 分类器应用环境迁移等。

$K\in C\wedge x\in K\}$。

定义 3[22] 设(U,C)为覆盖近似空间，x为U中的一个对象，则在覆盖近似空间下，x的不可分辨集为$Ind_c(x)=\bigcap_{K\in Des_c(x)} K$。

定义 4[22] 设(U,C)为覆盖近似空间，$x\in U$，称$Md(x)=\{K\in C|x\in K\wedge(\forall S\in C\wedge x\in S\wedge S\subseteq K\Rightarrow K=S)\}$为$x$的最小描述。

定义 5[9] 设(U,C)为覆盖近似空间，$x\in U$，任意$K\in C$，如果$K\in Md(x)$，则称K为x的一个邻域。$\bigcup\{K|K\in Md(x)\}$称为x的全邻域，记为$AN(x)$。$\bigcap\{K|K\in Md(x)\}$称为x的近邻域，记为$CN(x)$。

定义 6[16] 设论域$U=\{u_1,u_2,\cdots,u_n\}$，其中元素$u_i(i=1,2,\cdots,n)$是所讨论的对象。U上一个 Vague 集A由真隶属度函数t_A和假隶属度函数f_A所描述：$t_A:U\to[0,1]$，$f_A:U\to[0,1]$。其中，$t_A(u_i)$是由支持u_i的证据所导出的肯定隶属度的下界，$f_A(u_i)$是由反对u_i的证据所导出的否定隶属度的下界，且$t_A(u_i)+f_A(u_i)\leqslant 1$。元素$u_i$在 Vague 集$A$中的隶属度被区间$[0,1]$的一个子区间$[t_A(u_i),1-f_A(u_i)]$所界定，称该区间为$u_i$在$A$中的 Vague 值。

定义 7[16] 一个 Vague 集A的补集A^c定义为：$t_{A^c}(x)=f_A(x)$，$1-f_{A^c}(x)=1-t_A(x)$。

性质 1[16] Vague 集A为B所包含，即$A\subseteq B$，当且仅当$t_A(x)\leqslant t_B(x)$且$1-f_A(x)\leqslant 1-f_B(x)$。

定义 8[16] 两个 Vague 集A和B相等，即$A=B$，当且仅当$A\subseteq B$和$B\subseteq A$，即$t_A(x)=t_B(x)$，$1-f_A(x)=1-f_B(x)$。

性质 2[19] 设有 Vague 集A、B、C，有如下性质：

(1)$A\cup B=B\cup A$，$A\cap B=B\cap A$；

(2)$A\cup A=A$，$A\cap A=A$；

(3)$A\cup(B\cup C)=(A\cup B)\cup C$，$A\cap(B\cap C)=(A\cap B)\cap C$；

(4)$A\cup[1,1]=[1,1]$，$A\cap\varnothing=\varnothing$。

为了论述方便，将文献[21]提出的覆盖粗糙 Vague 集模型称为 I 型覆盖粗糙 Vague 集模型。

定义 9[21] 设(U,C)为覆盖近似空间，其中U为非空有限论域，C为U上的一个覆盖，对$\forall x\in U$，Vague 集关于覆盖近似空间(U,C)的上、下近似分别为$\overline{C}(A)$，$\underline{C}(A)$，且有：

$$t_{\overline{C}(A)}(x)=\sup\{t_A(y)|y\in\bigcup Md(x)\}$$

$$f_{\overline{C}(A)}(x)=\inf\{f_A(y)|y\in\bigcup Md(x)\}$$

$$t_{\underline{C}(A)}(x)=\inf\{t_A(y)|y\in\bigcup Md(x)\}$$

$$f_{\underline{C}(A)}(x)=\sup\{f_A(y)|y\in\bigcup Md(x)\}$$

则称序对$CV=(\underline{C}(A),\overline{C}(A))$为基于覆盖的粗糙 Vague 集，简记为$(\underline{CF}(A),\overline{CF}(A))$。

如同经典粗糙集的定义一样，在上述定义中，若对$\forall x\in U$，有$\underline{CF}(A)(x)=\overline{CF}(A)(x)$，则称 Vague 集$A$关于覆盖近似空间$(U,C)$是可定义的，否则称 Vague 集$A$关于覆盖近似空间$(U,C)$是粗糙的。$\underline{CF}$，$\overline{CF}$分别为基于覆盖的粗糙 Vague 集下近似算子和上近似算子，并规定基于覆盖的粗糙 Vague 集$CV=\varnothing$，当且仅当$\underline{CF}(A)=\varnothing$且$\overline{CF}(A)=\varnothing$。上述定义表明，任何 Vague 集模型在覆盖近似空间中，都可以用两个 Vague 集来逼近，即上近似$\overline{CF}(A)$和下近似$\underline{CF}(A)$。

下面举例说明 I 型覆盖粗糙 Vague 集并不满足文献[21]所提出的幂等性。

例 1 设(U,C)为一覆盖近似空间，论域$U=\{x_1,x_2,\cdots,x_5\}$，已知U上的一个覆盖$C=\{\{x_1,x_2\},\{x_2,x_3\},\{x_3,x_4\},\{x_4,x_5\}\}$，有 Vague 集：

$$A=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

则根据定义有：

$$\underline{CF}(A)=\{\frac{[0.7,0.8]}{x_1},\frac{[0.1,0.2]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.1,0.2]}{x_4},\frac{[0.3,0.4]}{x_5}\}$$

$$\underline{CF}(\underline{CF}(A))=\{\frac{[0.1,0.2]}{x_1},\frac{[0.1,0.2]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.1,0.2]}{x_4},\frac{[0.1,0.2]}{x_5}\}$$

$$\overline{CF}(A)=\{\frac{[0.8,0.9]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.8,0.9]}{x_3},\frac{[0.5,0.6]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

$$\overline{CF}(\overline{CF}(A))=\{\frac{[0.8,0.9]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.8,0.9]}{x_3},\frac{[0.8,0.9]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

显然，$\underline{CF}(A)\neq\underline{CF}(\underline{CF}(A))$，$\overline{CF}(A)\neq\overline{CF}(\overline{CF}(A))$。因此文献[21]中所定义的覆盖粗糙 Vague 集模型并不满足其所给出的幂等性。

3 一种新的覆盖粗糙 Vague 集模型

为了与前文的论述保持一致，将本文提出的覆盖粗糙模糊集模型称为 II 型覆盖粗糙 Vague 集模型。

定义 10 设(U,C)为覆盖近似空间，其中U为非空有限论域，C为U上的一个覆盖，对$\forall x\in U$，Vague 集关于覆盖近似空间(U,C)的上、下近似分别为$\overline{C}(A)$，$\underline{C}(A)$，且有：

$$t_{\overline{C}(A)}(x)=\sup\{t_A(y)|y\in\bigcap Md(x)\}$$

$$f_{\overline{C}(A)}(x)=\inf\{f_A(y)|y\in\bigcap Md(x)\}$$

$$t_{\underline{C}(A)}(x)=\inf\{t_A(y)|y\in\bigcap Md(x)\}$$

$$f_{\underline{C}(A)}(x)=\sup\{f_A(y)|y\in\bigcap Md(x)\}$$

则称序对$CV=(\underline{C}(A),\overline{C}(A))$为 II 型覆盖粗糙 Vague 集，记为$(\underline{CS}(A),\overline{CS}(A))$。

与 I 型覆盖粗糙 Vague 集模型相比，II 型覆盖粗糙 Vague 集模型缩小了对象不可分辨的集合，即只考察对象近邻域中的元素。显然，任何对象x的近邻域包含于它的全邻域中。因此，有$\underline{CF}(A)\subseteq\underline{CS}(A)\subseteq A\subseteq\overline{CS}(A)\subseteq\overline{CF}(A)$。其直观意义就是，II 型覆盖粗糙 Vague 集上、下近似比 I 型覆盖粗糙 Vague 集模型更加逼近所讨论的 Vague 集对象集合，从而可以更好地发现模糊信息系统中一些隐藏的知识。

性质 3 设(U,C)为一覆盖近似空间，A，B分别为论域U上 Vague 集，则 II 型覆盖粗糙 Vague 集下近似算子$\underline{CS}$和上近似算子$\overline{CS}$具有如下性质：

(1)$\underline{CS}(\emptyset)=\overline{CS}(\emptyset)=\emptyset,\underline{CS}(U)=\overline{CS}(U)=U$；

(2)$\underline{CS}(A)\subseteq A\subseteq\overline{CS}(A)$；

(3)$\underline{CS}(A\cap B)=\underline{CS}(A)\cap\underline{C}(B)$；

(4)$\overline{CS}(A\cup B)=\overline{CS}(A)\cup\overline{CS}(B)$；

(5)$A\subseteq B\Rightarrow\underline{CS}(A)\subseteq\underline{CS}(B),\overline{CS}(A)\subseteq\overline{CS}(B)$；

(6)$\underline{CS}(A^c)=(\overline{CS}(A))^c,\overline{CS}(A^c)=(\underline{CS}(A))^c$；

(7)$\underline{CS}(A)=\underline{CS}(\underline{CS}(A))$；

根据定义 9 和定义 10 可得：

性质(1)—性质(6)的证明较为简单，从略。这里仅对性质(7)进行证明。

证明：由定义 10 可得：

$$\begin{aligned}t_{\underline{CS}(\underline{CS}(A))}&=\inf\{t_{\underline{CS}(A)}(y)\mid y\in\cap Md(x)\}\\&=\inf\{\inf\{t_A(y)\mid y\in\cap Md(x)\}\}\\&=\inf\{t_A(y)\mid y\in\cap Md(x)\}=t_{\underline{CS}(A)}\end{aligned}$$

$$\begin{aligned}f_{\underline{CS}(\underline{CS}(A))}&=\sup\{f_{\underline{CS}(\underline{CS}(A))}(y)\mid y\in\cap Md(x)\}\\&=\sup\{\sup\{f_A(y)\mid y\in\cap Md(x)\}\}\\&=\sup\{f_A(y)\mid y\in\cap Md(x)\}=f_{\underline{CS}(A)}\end{aligned}$$

所以，$\underline{CS}(A)=\underline{CS}(\underline{CS}(A))$。同理可证，$\overline{CS}(A)=\overline{CS}(\overline{CS}(A))$。

下面举例说明型 II 覆盖粗糙 Vague 集满足幂等性。为了说明问题，仍使用与例 1 相同的算例。

例 2　设(U,C)为一覆盖近似空间，论域$U=\{x_1,x_2,\cdots,x_5\}$，已知U上的一个覆盖$C=\{\{x_1,x_2\},\{x_2,x_3\},\{x_3,x_4\},\{x_4,x_5\}\}$，有 Vague 集：

$$A=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

则根据定义 10 有：

$$\underline{CS}(A)=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.3,0.4]}{x_5}\}$$

$$\underline{CS}(\underline{CS}(A))=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.3,0.4]}{x_5}\}$$

$$\overline{CS}(A)=\{\frac{[0.8,0.9]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

$$\overline{CS}(\overline{CS}(A))=\{\frac{[0.8,0.9]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

显然有，$\underline{CS}(A)=\underline{CS}(\underline{CS}(A))$；$\overline{CS}(A)=\overline{CS}(\overline{CS}(A))$。因此 II 覆盖粗糙 Vague 集满足幂等性。

4　二种覆盖粗糙 Vague 集模型的关系

下面讨论一般情况下二种覆盖粗糙 Vague 集的关系。

定理 1　设(U,C)为一覆盖近似空间，A为论域U上 Vague 集，则有$\underline{CF}(A)\subseteq\underline{CS}(A)\subseteq A\subseteq\overline{CS}(A)\subseteq\overline{CF}(A)$。

证明：设$\forall x\in U,\forall K\in Md(x)$，有$\cap Md(x)\subseteq K\subseteq\cup Md(x)$，又根据性质 3 可知，$\underline{CS}(A)\subseteq A\subseteq\overline{CS}(A)$，所以只需证明$\underline{CF}(A)\subseteq\underline{CS}(A)$及$\overline{CS}(A)\subseteq\overline{CF}(A)$即可。因为：

$$\inf\{t_A(y)\mid y\in\cup Md(x)\}\leqslant\inf\{t_A(y)\mid y\in\cap Md(x)\}$$

$$\sup\{f_A(y)\mid y\in\cap Md(x)\}\leqslant\sup\{f_A(y)\mid y\in\cup Md(x)\}$$

即$t_{\underline{CF}}(A)\leqslant t_{\underline{CS}(A)},1-f_{\underline{CF}(A)}\leqslant 1-f_{\underline{CS}(A)}$，所以$\underline{CF}(A)\subseteq\underline{CS}(A)$。

同理可证，$\overline{CS}(A)\subseteq\overline{CF}(A)$。所以$\underline{CF}(A)\subseteq\underline{CS}(A)\subseteq A\subseteq\overline{CS}(A)\subseteq\overline{CF}(A)$成立。

例 3　根据例 1 和例 2 可知：

$$A=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

$$\underline{CF}(A)=\{\frac{[0.7,0.8]}{x_1},\frac{[0.1,0.2]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.1,0.2]}{x_4},\frac{[0.3,0.4]}{x_5}\}$$

$$\overline{CF}(A)=\{\frac{[0.8,0.9]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.8,0.9]}{x_3},\frac{[0.5,0.6]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

$$\underline{CS}(A)=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.3,0.4]}{x_5}\}$$

$$\overline{CS}(A)=\{\frac{[0.8,0.9]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.5,0.6]}{x_5}\}$$

显然有，$\underline{CF}(A)\subseteq\underline{CS}(A)\subseteq A\subseteq\overline{CS}(A)\subseteq\overline{CF}(A)$成立。

5　基于熵的覆盖粗糙 Vague 集不确定性度量方法

定义 11[22]　设U为非空有限论域，C为U上的一个覆盖，则覆盖近似空间(U,C)的知识熵定义如下：

$$E(U,C)=-\sum_{x_I\in U}\frac{1}{|U|}\log\frac{|Ind_C(x_i)|}{|U|}$$

性质 4[22]　设U为非空有限论域，C为U上的一个覆盖，则覆盖近似空间(U,C)的知识熵具有如下性质：

(1)$0\leqslant E(U,C)\leqslant\log|U|$；

(2)当$\forall_{x_i\in U}(Ind_C(x_i)=U)$时，$E(U,C)$取最小值 0；

(3)当$\forall_{x_i\in U}(Ind_C(x_i)=x_i)$时，$E(U,C)$取最大值$\log|U|$。

定理 2[22]　设(U,C_1)和(U,C_2)为覆盖近似空间，若$C_1\leqslant C_2$，则有$E(U,C_1)\geqslant E(U,C_2)$。证明参见文献[22]。

定理 2 说明覆盖粒度空间的知识具有颗粒型。在具有较细关系的两个覆盖粒度空间上，其知识熵也具有偏序关系，即粒度空间越细，知识熵越大，覆盖粒度空间不确定性越小；反之，粒度空间越粗，知识熵越小，覆盖粒度空间不确定性越大。另一个引起覆盖粗糙 Vague 集不确定性的因素是其本身边界域的大小。类比经典粗糙集，通过覆盖粗糙 Vague 集本身的代数特征来定义 II 型覆盖粗糙 Vague 集本身的粗糙度。

定义 12 设 $CV=(\underline{CS}(A),\overline{CS}(A))$ 是给定论域 U 上的 II 型覆盖粗糙 Vague 集，$0<\beta\leqslant\alpha\leqslant 1$，则有 A 的 α 下近似：

$$\underline{CS}(A)_\alpha=\{y\mid y\in\bigcap Md(x_i),t_{\underline{C}(A)}(y)\geqslant\alpha,f_{\underline{C}(A)}(y)<\alpha,\ i=1,2,\cdots,n\}$$

A 的 β 上近似：

$$\overline{CS}(A)_\beta=\{y\mid y\in\bigcap Md(x_i),t_{\overline{C}(A)}(y)\geqslant\beta,f_{\overline{C}(A)}(y)<\beta,\ i=1,2,\cdots,n\}$$

定义 13 II 型覆盖粗糙 Vague 集的精度和粗糙度分别定义如下：

$$\gamma_{CV}^{\alpha,\beta}=\frac{|\underline{CS}(A)_\alpha|}{|\overline{CS}(A)^\beta|},\rho_{CV}^{\alpha,\beta}=1-\gamma_{CV}^{\alpha,\beta}$$

式中，$0<\beta\leqslant\alpha\leqslant 1,\overline{CS}(A)^\beta\neq\emptyset$。

下面定义基于熵的 II 型覆盖粗糙 Vague 集不确定性度量方法。

定义 14 设 (U,C) 为一覆盖近似空间，$CV=(\underline{CS}(A),\overline{CS}(A))$ 为论域 U 上的 II 型覆盖粗糙 Vague 集，$0<\beta\leqslant\alpha\leqslant 1$，定义 II 型覆盖粗糙 Vague 集一种不确定性度量为：

$$E_\beta^\alpha(A,C)=\rho_{CV}^{\alpha,\beta}(\log^{|U|}-E(U,C))$$

上述覆盖粗糙 Vague 集的不确定性度量方式，既反映了覆盖粗糙空间的粒度大小，也反映了覆盖粗糙 Vague 集本身的粗糙度，因此可以度量覆盖粗糙 Vague 集的不确定性。

6 算例分析

例 4 设 (U,C) 为一覆盖近似空间，论域 $U=\{x_1,x_2,\cdots,x_7\}$，已知 U 上的两个覆盖 $C_1=\{\{x_1,x_2\},\{x_3,x_6,x_7\},\{x_3,x_7\},\{x_4,x_5,x_6\}\}$，$C_2=\{\{x_1\},\{x_2\},\{x_3,x_7\},\{x_4,x_5\},\{x_6\}\}$。

有 Vague 集：

$$A=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.5,0.6]}{x_5},\frac{[0.3,0.4]}{x_6},\frac{[0.7,0.8]}{x_7}\}$$

由定义 5 知，$C_2\leqslant_3 C_1$；由定义 11 计算知，$E(U,C_1)=1.079$，$E(U,C_2)=1.5498$。显然 $E(U,C_1)<E(U,C_2)$，表明覆盖越细，知识熵值越大，提供的信息量越多，不确定性越小，这与定理 2 相符。

由 II 型覆盖粗糙 Vague 集定义可得：

$$\underline{CS_1}(A)=\{\frac{[0.7,0.8]}{x_1},\frac{[0.7,0.8]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.3,0.4]}{x_5},\frac{[0.3,0.4]}{x_6},\frac{[0.1,0.2]}{x_7}\}$$

$$\overline{CS_1}(A)=\{\frac{[0.8,0.9]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.7,0.8]}{x_3},\frac{[0.5,0.6]}{x_4},\frac{[0.5,0.6]}{x_5},\frac{[0.3,0.4]}{x_6},\frac{[0.7,0.8]}{x_7}\}$$

$$\underline{CS_2}(A)=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.1,0.2]}{x_3},\frac{[0.3,0.4]}{x_4},\frac{[0.3,0.4]}{x_5},\frac{[0.3,0.4]}{x_6},\frac{[0.1,0.2]}{x_7}\}$$

$$\overline{CS_2}(A)=\{\frac{[0.7,0.8]}{x_1},\frac{[0.8,0.9]}{x_2},\frac{[0.7,0.8]}{x_3},\frac{[0.5,0.6]}{x_4},\frac{[0.5,0.6]}{x_5},\frac{[0.3,0.4]}{x_6},\frac{[0.7,0.8]}{x_7}\}$$

这里，$CV_1=(\underline{CS_1}(A),\overline{CS_1}(A))$，$CV_2=(\underline{CS_2}(A),\overline{CS_2}(A))$ 分别是覆盖 C_1 和 C_2 下的 II 型粗糙 Vague 集。取 $\beta=0.3,\alpha=0.4$，有：

$$\rho_{CV_1}^{\alpha,\beta}=1-\gamma_{CV_1}^{\alpha,\beta}=1-\frac{|x_1,x_2|}{|x_1,x_2,x_3,x_7|}=\frac{1}{2}$$

$$\rho_{CV_2}^{\alpha,\beta}=1-\gamma_{CV_2}^{\alpha,\beta}=1-\frac{|x_1,x_2|}{|x_1,x_2,x_3,x_7|}=\frac{1}{2}$$

计算 II 型覆盖粗糙 Vague 集不确定性度量为：

$$E_\beta^\alpha(A,C_1)=\rho_{CV_1}^{\alpha,\beta}(\log^{|U|}-E(U,C_1))=0.5*(1.9459-1.079)=0.5*0.8669=0.43345$$

$$E_\beta^\alpha(A,C_2)=\rho_{CV_2}^{\alpha,\beta}(\log^{|U|}-E(U,C_2))=0.5*(1.9459-1.5498)=0.5*0.3961=0.1981$$

显然有结论：$C_2\leqslant_3 C_1,E_\beta^\alpha(A,C_2)<E_\beta^\alpha(A,C_1)$，表明覆盖越细，II 型覆盖粗糙 Vague 集不确定性度量值越小，不确定程度越小。因此可以采用本文的度量方式来度量 II 型覆盖粗糙 Vague 集的不确定性程度。

结束语 覆盖粗糙 Vague 集模型是机器学习领域应用前景非常广泛的一种粗糙集扩展模型。本文在覆盖粗糙集理论研究的基础上，具体分析了已有覆盖粗糙 Vague 集模型中幂等性不成立的局限性，提出了一种新的基于近邻域的覆盖粗糙 Vague 集模型，给出了两种模型的关系，并结合实例证明了新模型幂等性是成立的，完善了覆盖粗糙 Vague 集模型相关理论。另外，又结合覆盖近似空间下知识熵的概念，提出了 II 型覆盖粗糙 Vague 集模型的不确定性度量方法，为下一步进行属性约减和知识获取奠定了基础。

参考文献

[1] Pawlak Z. Rough set[J]. International Journal of Computer and Information Sciences, 1982, 11: 341-356

[2] Wang G Y, Li T R, Grzymala-Busse J, et al. Rough sets and knowledge technology [C] // Proceedings of RSKT2008, LNAI5009. Berlin, Germany: Springer-Verlag, 2008

[3] 王国胤，张清华，胡军. 粒计算研究综述[J]. 智能系统学报，2007, 2(6): 8-26

[4] 梁吉业，李德玉. 信息系统中的不确定性与知识获取[M]. 北京：科学出版社，2005

[5] Zhu W, Wang F Y. Reduction and axiomization of covering generalized rough sets[J]. Information Sciences, 2003, 152: 217-230

[6] Tsang E C C, Chen D G, Lee J W T, et al. On the upper approximations of covering generalized rough sets[C]// Proceedings of the 3rd International Conference Machine Learning and Cybernetics. Shanghai, China, 2004: 4200-4203

[7] Zhu W, Wang F Y. On Three Types of Covering-Based Rough Sets[J]. IEEE Transactions on Knowledge and Data Enginee-

ring,2007,19(8):1131-1144
[8] Zhu W. Relationship between generalized rough sets based on binary relation and covering[J]. Information Sciences,2009,179(3):210-225
[9] Zhu W,Wang F Y. A new type of covering rough set [C]// 3rd International IEEE Conference Intelligenty Stems. London:[s. n.],2006:444-449
[10] 胡军,王国胤. 覆盖粗糙集的模糊度[J]. 重庆邮电大学学报,2009,21(4):490-493
[11] 魏荣,刘保仓,史开泉. 基于覆盖广义粗集的模糊性[J]. 计算机科学,2007,34(1):153-155
[12] 徐伟华,张文修. 覆盖广义粗糙集的模糊性[J]. 模糊系统与数学,2006,20(6):115-121
[13] 徐菲菲,苗夺谦,李国道,等. 基于最简覆盖的粗糙模糊集的粗糙熵[J]. 计算机科学,2006,33(10):179-181
[14] 王健鹏,戴岱,周正春. 基于覆盖的模糊粗糙集模型[J]. 周口师范学院学报,2004,21(2):20-23
[15] 余美真,李进金. 覆盖粗糙模糊集的不确定性度量[J]. 模糊系统与数学,2009,23(6):145-148
[16] 李凡,徐章艳,饶勇. Vague集[J]. 计算机科学,2000,27(9):12-24,28
[17] 闫德勤,迟忠先. 粗糙集与 Vague 集[J]. 计算机科学,2004,31(8):133-135
[18] 刘金良,闫瑞霞,姚炳学. 粗糙 Vague 集的不确定性度量[J]. 系统工程与电子技术,2008,30(1):104-107
[19] 朱六兵. 粗糙集与 Vague 集的理论及应用研究[D]. 成都:西南交通大学,2006
[20] 张倩倩,徐久成,胡玉文. 基于覆盖的粗糙 Vague 集模型研究[J]. 广西大学学报,2009,35(5):653-657,662
[21] 徐久成,张倩倩. 覆盖粗糙 Vague 集的不确定性度量研究[J]. 计算机科学,2010,37(10):225-227
[22] 胡军,王国胤. 覆盖粒度空间的层次模型[J]. 南京大学学报,2008,44(5):551-558
[23] 刘纪芹,史开泉. 基于知识含量的粗糙集不确定性度量[J]. 计算机科学,2007,34(7):171-173
[24] 王国胤,张清华. 不同知识粒度下粗糙集的不确定性研究[J]. 计算机学报,2008,31(9):1588-1598
[25] 解滨,李磊军,米据生. 基于知识粒度的粗糙集的不确定性度量[J]. 计算机学报,2010,37(9):225-228
[26] 胡军,王国胤. 一种覆盖粗糙模糊集模型[J]. 软件学报,2010,21(5):968-977
[27] 胡军,王国胤. 粗糙集的不确定性度量准则[J]. 模式识别与人工智能,2010,23(5):606-615
[28] 李大湘,彭进业,李展. 基于半监督多示例学习的对象图像检索[J]. 控制与决策,2010(7):981-986
[29] 李大湘,彭进业,贺进芳. 基于视觉语义与 RSSVM 的图像检索[J]. 华南理工大学学报:自然科学版,2010(04):156-161

(上接第 204 页)

低高维模糊粗糙集计算的复杂度问题,并在实际问题的处理中得到应用。

参 考 文 献

[1] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Seience,1982,11:341-356
[2] Pawlak Z. Rough set theory and its applications to data analysis [J]. Cybernetics and Systems,1998,29:661-688
[3] 张文修,吴志伟,梁吉业,等. 粗糙集理论与方法[M]. 北京:科学出版社,2001
[4] Wu W Z,Zhang W X,Li H Z. Knowledge acquisition in incomplete fuzzy information systems via the rough set approach[J]. Expert Systems,2003,20:280-286
[5] Yao Y Y. Relational interpretations of neighborhood operators and rough set approximation operators[J]. Information Sciences,1998,111:239 -259
[6] Dubois D,Prade H. Rough fuzzy sets and fuzzy rough sets[J]. International Journal of General System,1990,17:191-208
[7] Wu W Z,Mi J S,Zhang W X. Generalized fuzzy rough sets[J]. Information Sciences,2003,51:263-282
[8] Liu G L. Generalized rough sets over fuzzy lattices[J]. Information Sciences,2008,178:1651-1662
[9] Yao Y Y. Granular computing: Basic issues and possible solutions [C]// Paul P,ed. Proceedings of the 5thJoint Conference on Information Sciences. USA: Elsevier Publishing Company,2000:186-189
[10] Pawlak Z. Fuzzy sets and information granularity[C]// Gupta M,Ragade R,Yager R,et al. ,eds. Advances in Fuzzy Set Theory and Applications. North-Holland, Amsterdam Publishing Co. ,1979:3-18
[11] Pawlak Z. Fuzzy Logic=Computing with Words [J]. IEEE Trans. on Fuzzy Systems,1996,2:103-111
[12] Yao Y Y. Granular computing using neighborhood systems [C]// Roy R. Advances in soft computing: Engineering design and manufacturing. London: Springer-Verlag Company, 1999: 539-553
[13] 苗夺谦,范世栋. 知识的粒度计算及其应用[J]. 系统工程理论与实践,2002,22:48-56
[14] 张钹,张铃. 问题求解的理论及应用[M]. 北京:清华大学出版社,1990
[15] 张铃,张钹. 模糊商空间理论[J]. 软件学报,2003,14:770-776
[16] 吴明芬,沈挺,曹存根,等. 模糊商空间理论两个定理的补充[J]. 北京交通大学学报,2009,33(6):86-90
[17] Yao Y Y,Lingras P J. Interpretations of Belief Functions in the Theory of Rough Sets[J]. Information Sciences,1998,104:81-106
[18] Li T J. Rough approximation operators on two universes of discourse and their fuzzy extensions[J]. Fuzzy Sets and Systems,2008,159:3033-3050
[19] 邱雅竹,付蓉. 两个论域上的粗糙集模型及其应用[J]. 四川师范大学学报:自然科学版,2005,28:15-18
[20] 高明,王继成. 双论域下粗糙集数据约简方法研究[J]. 计算机工程与应用,2009,45:144-146
[21] 吴明芬,曹存根. 近似空间的笛卡尔积粗集模型及其可分解性[J]. 计算机科学,2011,38:225-228
[22] 陈卫东,张维明. 笛卡尔积运算对数据库数据质量的传递影响[J]. 计算机科学,2008,35:210-213