软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2016,27(2):295−308 [doi: 10.13328/j.cnki.jos.004854]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

粗糙集的最优近似集*

张清华[1,2], 薛玉斌[1], 王国胤[2]

[1](重庆邮电大学 理学院,重庆 400065)

[2](计算智能重庆市重点实验室(重庆邮电大学),重庆 400065)

通讯作者: 张清华, E-mail: zhangqh@cqupt.edu.cn

摘 要: Pawlak 教授提出的粗糙集理论是解决集合边界不确定的重要手段,他构建了边界不确定集合的两条精确边界,但没有给出用已有知识基来精确或近似地构建目标概念(集合)X 的方法.在前期的研究中提出了寻找目标概念 X 的近似集方法,但并没有给出最优的近似集.首先,回顾了集合间的相似度概念和粗糙集的近似集 $R_\lambda(X)$的构建方法,提出并证明了 $R_\lambda(X)$所满足的运算性质.其次,找到了 $R_\lambda(X)$比上近似集 $\overline{R}(X)$ 和下近似集 $\underline{R}(X)$更近似于目标概念 X 的λ成立的区间.最后,提出了 $R_{0.5}(X)$作为目标概念的最优近似集所满足的条件.

关键词: 粗糙集;模糊集;近似集;相似度;粒计算

中图法分类号: TP18

中文引用格式: 张清华,薛玉斌,王国胤.粗糙集的最优近似集.软件学报,2016,27(2):295−308. http://www.jos.org.cn/1000-9825/4854.htm

英文引用格式: Zhang QH, Xue YB, Wang GY. Optimal approximation sets of rough sets. Ruan Jian Xue Bao/Journal of Software, 2016,27(2):295−308 (in Chinese). http://www.jos.org.cn/1000-9825/4854.htm

Optimal Approximation Sets of Rough Sets

ZHANG Qing-Hua[1,2], XUE Yu-Bin[1], WANG Guo-Yin[2]

[1](School of Science, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

[2](Chongqing Key Laboratory of Computational Intelligence (Chongqing University of Posts and Telecommunications), Chongqing 400065, China)

Abstract: Rough set theory proposed by professor Pawlak is an important mean to solve the problem of uncertain boundary region. Pawlak constructed two crisp boundaries for the set with uncertainty boundary but did not give any exact or approximate methods of using the existing knowledge base to build an approximation set of a target concept .In order to solve this problem, in the previous researches a method for looking for this kind of approximation target concept (set) is proposed. However, that method does not give out a kind of optimal approximation set. In this paper, firstly, the concept of the similarity between the target set and its approximation set and the method for constructing approximation set of rough set are reviewed, and the operation properties are proposed and proved respectively. Secondly, an interval of λ is found, and in this interval $R_\lambda(X)$ is more similar to the target concept X than the upper-approximation set $\overline{R}(X)$ or lower-approximation set $\underline{R}(X)$. Finally, the conditions of $R_{0.5}(X)$ as an optimal approximation set of the target concept X are proposed.

Key words: rough set; fuzzy set; approximation set; similarity; granular computing

目前,随着计算机技术和自动控制技术的不断发展,越来越多与此相关的研究工作关注到不确定性问题[1].如何对大量数据和不确定性问题进行有效处理,并从中寻找到知识和规律,是一个日渐重要的研究课题[2].自 1965 年

* 基金项目: 国家自然科学基金(61472056, 61272060); 重庆市自然科学基金(cstc2012jjA40032, cstc2013jcyjA40063)

Foundation item: National Natural Science Foundation of China (61472056, 61272060); Chongqing Natural Science Foundation of China (cstc2012jjA40032, cstc2013jcyjA40063)

收稿时间: 2014-07-29; 修改时间: 2015-02-09; 采用时间: 2015-05-08

Zadeh 教授提出模糊集(fuzzy set)理论[3]以来,不确定性问题的研究取得了重大突破,各种新的处理不确定性问题的理论如雨后春笋般不断涌现.如 1982 年 Pawlak 提出了粗糙集理论[4];1990 年,张铃、张钹提出了商空间理论[5];1993 年姚一豫等人提出了区间集及区间代数理论[6];1999 年 Molodtsov 提出了软集理论(soft set theory)[7]等等.这些理论都是处理不确定性问题的重要方法.目前这些理论之间相结合的研究是不确定性问题研究的热点.来源于简单信息模型的粗糙集理论,是一种能定量分析处理不完整、不一致、不精确信息的数学工具,其核心是通过不可分辨关系(等价关系)来建立一个划分空间(知识基),并在该划分空间上用两个精确的集合(上近似集和下近似集)来刻画目标概念 X.粗糙集理论思想新颖,方法独特,计算简便,现已成为一种重要的智能信息处理技术[8].模糊集理论的本质,是通过一个隶属函数将论域映射到 0 到 1 的区间,再通过设定域值(即截集)筛选出所期望的元素.目前,把粗糙集理论和模糊集理论结合起来研究不确定性问题,是一流行的趋势,并且取得了令人瞩目的成果[9−12].

在粗糙集理论研究中我们定义:在给定知识基(U,R)的条件下,若 X 即目标概念恰好是(U,R)中一些知识粒的并集,则 X 在(U,R)下是精确的;若 X 不能用(U,R)中的知识粒的并集表示,则称 X 是粗糙的.在 Pawlak 粗糙集模型中,上近似集和下近似集构成了边界不确定集合的两条精确边界.因此,论域 U 被这两条精确边界分为了 3 个部分:绝对属于 X 的部分,即正域部分;绝对不属于 X 的部分,即负域部分;以及可能属于 X 的部分,即边界域部分.在挖掘数据过程中我们发现正域部分和负域部分的概念是确定的,但边界域的概念并不确定.那么要如何挖掘边界区域的概念呢?很多研究者利用上近似集和下近似集来替代目标概念进行研究,并得到一些近似的规则和知识.特别地,用下近似集来作为目标概念的近似集来计算规则的精度和不确定性,是目前研究的主要方法.但这两种方法要么缩小了目标概念的研究范围,要么放大了目标概念的研究范围,在很多情况下用这两种方法得到的结果并不能真实反映目标概念的属性.

因此,有学者尝试构造目标概念的近似算子来解决该问题,得到了一类新的模型即粗糙集的扩展型.很多学者对此进行了讨论[13−15],Pawlak 等人也对此进行了分析和总结[16].目前得到的扩展模型中比较典型的模型有两类:变精度粗糙集模型和概率粗糙集模型.米据生等人讨论了变精度粗糙集模型,并利用它进行属性约简,取得了较好的效果[17−21].Yao 和 Ziarko 等人结合概率论和包含度提出了概率粗糙模型,也取得了较好的理论成果[22−25].但这些方法只是构建了扩展的 Pawlak 近似算子,他们并没有用现有的知识粒来构建目标概念的近似集.那么,利用现有的知识粒我们是否能够建立一个集合(当前知识基下的一些知识粒的并)作为 X 的近似集,它比上近似集和下近似集更加近似于 X?如果能够建立一个这样的近似集,我们将用它来提取规则,其相似度将会提高.基于此设想我们在文献[26]中提出了用模糊截集的手段来构建目标概念的近似集的方法,找到一个相对较好的近似集 $R_{0.5}(X)$并发现了它的一些性质.但文献[26]并没有给出目标概念的最优近似集也没有给出一般近似集的性质.为此本文提出和证明了一般近似集所满足的运算规则,找到了比上近似集和下近似集更近似于 X 的 λ 成立的区间.提出了 $R_{0.5}(X)$作为 X 的最优近似集所满足的条件.

本文第 1 节介绍相关的基本概念.第 2 节介绍粗糙集的相似度并证明粗糙集的近似集满足的运算规律.第 3 节给出最优近似集.第 4 节是结束语.

1 相关基本概念

为了更清楚地阐述本文思想,在介绍粗糙集的近似集之前,先回顾粗糙集的相似度等相关概念.

定义 1(模糊集[27]). 设在论域 U 上给定了一个映射:

$$A:U\to[0,1],$$
$$u\mapsto A(u),$$

则称 A 为 U 上的模糊集,称 $A(u)$ 为 A 的隶属函数(或称为 u 对 A 的隶属度).

由定义 1 不难看出,对于某 F 集 A,若 $A(u)$ 只取值为 0 或 1 这两个数时,模糊集 A 就退化成为一个普通集合.因此,普通集合只是模糊集的特殊情形.如果 $A(u)\equiv 0$, 则称 A 为空集$\varnothing$,如果 $A(u)\equiv 1$, 则称 A 为全集 U 在给定的论域 U 上可以有多个 F 集,记 U 上的 F 集的全体为$\mathbb{F}(U)$,即 $\mathbb{F}(U)=\{A\mid A:U\to[0,1]\}$.

定义 2(截集[27]). 设 $A\in\mathbb{F}(U),\lambda\in[0,1]$, 记

$$A_\lambda = \{u \mid u \in U, A(u) \geqslant \lambda\},$$

称 A_λ 为 A 的一个λ截集,λ称为域值(或置信水平);

$$A_{\underset{\cdot}{\lambda}} = \{u \mid u \in U, A(u) > \lambda\},$$

称 $A_{\underset{\cdot}{\lambda}}$ 为 A 的λ强截集.

定义 3(粗糙集的上、下近似集[4,28]). 给定信息表知识表达系统 $S=\langle U,A,V,f\rangle$ 对于任一对象集合 $X\subseteq U$ 和属性集合 $R\subseteq A$, X 关于 R 的上近似集 $\overline{R}(X)$ 和下近似集 $\underline{R}(X)$ 分别定义如下:

$$\overline{R}(X) = \cup\{Y_i \mid Y_i \in U / IND(R) \wedge Y_i \cap X \neq \varnothing\},$$

$$\underline{R}(X) = \cup\{Y_i \mid Y_i \in U / IND(R) \wedge Y_i \subseteq X\},$$

其中, $U / IND(R) = \{X \mid (X \subseteq U \wedge \forall_{x\in X, y\in X, b\in R}(b(x)=b(y)))\}$ 是不分明关系 R 在 U 上的划分.

$\underline{R}(X)$ 表示知识空间(U,R)中一定能归入目标概念 X 的所有知识粒的并; $\overline{R}(X)$ 表示知识空间(U,R)中可能归入目标概念 X 的所有知识粒的并.集合 $BN_R(X)=\overline{R}(X)-\underline{R}(X)$ 称为目标概念 X 关于属性集合 R 的边界域(boundary region); $POS_R(X)=\underline{R}(X)$ 称为目标概念 X 关于属性集合 R 的正域(positive region); $NEG_R(X)=U-\overline{R}(X)$ 称为目标概念 X 的负域(negative region). $BN_R(X)$ 是知识空间(U,R)中既不能肯定归入目标概念 X,又不能肯定归入 $\overline{X}$ (X 的补集)的所有知识粒的并集.

定义 4(粗糙集[4,28]). 给定信息表知识表达系统 $S=\langle U,A,V,f\rangle$,对于任一对象集合 $X\subseteq U$ 和属性集合 $R\subseteq A$,当且仅当 $\overline{R}(X)\neq\underline{R}(X)$ 时,称集合 X 是 R 粗糙集(rough 集).

在知识表达系统 $S=\langle U,A,V,f\rangle$ 给定的情况下,若给出的目标概念 X 是粗糙的,则记为 $R(X)$.为方便叙述,我们令 $R(X)=[\underline{R}(X),\overline{R}(X)]$, 此时, $\underline{R}(X)\subseteq X\subseteq\overline{R}(X)$. 即我们给出了目标概念 X 的两个边界集合.在现有的知识空间(U,R)下,我们无法用(U,R)中已有的知识粒的并集来得到 X,故目标概念 X 是不确定的.

定义 5(集合的相似度[26]). 设 A,B 是有限论域 U 上的两个子集,即 $A\subseteq U, B\subseteq U$, 定义映射: $S:U\times U\to[0,1]$, 即 $(A,B)\to S(A,B)$. 称 $S(A,B)$是集合 A,B 的相似度,如果 $S(A,B)$满足如下条件:

(1) 对任意的 $A,B\in U, 0\leqslant S(A,B)\leqslant 1$ (有界性);

(2) 对任意的 $A,B\in U, S(A,B)=S(B,A)$ (对称性);

(3) 对任意的 $A,B\in U, S(A,A)=1; S(A,B)=0$ 的充要条件是 $A\cap B=\varnothing$.

对于任意满足(1)、(2)和(3)这 3 个条件的公式都是集合 A,B 的相似度公式.

这里我们构造相似度公式为 $S(A,B)=\dfrac{|A\cap B|}{|A\cup B|}$, 这里,$|\cdot|$表示元素个数,显然,我们构造的公式满足定义 5.

2 粗糙集的近似集及其性质

在知识表达系统 $S=\langle U,A,V,f\rangle$ 中,如果给定的概念 X 是粗糙的,那么必有 $\underline{R}(X)\subset X\subset\overline{R}(X)$. 因此,如何寻找目标概念 X 的最优近似集,是一个值得探讨的问题.文献[29,30]中给出了论域 U 中元素 x 隶属于目标概念 X 的隶属函数公式,文献[26]基于这个隶属函数再结合截集概念,构建了目标概念 X 的近似集.

设论域为 U,知识空间为 $U/IND(R)$, 对象子集为 $X\subseteq U, \forall x(x\in U)$, x 属于集 X 的隶属函数为

$$\mu_X^R(x)=\frac{|X\cap[x]_R|}{|[x]_R|},$$

显然, $0\leqslant\mu_X^R(x)\leqslant 1$,它可以用来表示论域 U 中任意一个元素 x 隶属于目标概念 X 的程度.

设 $U=\{x_1,x_2,\dots,x_n\}$,令 $X=\{x_{j_1},x_{j_2},\dots,x_{j_k}\}$ 为论域 U 上的一个子集,若令 $F_X^R=\{\mu_X^R(x_1),\mu_X^R(x_2),\dots,\mu_X^R(x_n)\}$, 显然 F_X^R 是集合 U 上的一个模糊集(即 $F_X^R\in\mathbb{F}(U)$), 且在这个模糊集上,位于同一个等价类的不同元素的隶属度相同.

根据上面的定义,再结合粗糙集的上下近似集,我们可以得到:

$$\underline{R}(X)=\{x\mid x\in U\wedge\mu_X^R(x)=1\},$$

$$\overline{R}(X)=\{x\mid x\in U\wedge 0<\mu_X^R(x)\leqslant 1\},$$

$$BN_R(X)=\overline{R}(X)-\underline{R}(X)=\{x\mid x\in U\wedge 0<\mu_X^R(x)<1\}.$$

从上面的定义我们可以看出,通过映射我们把论域实际上分成了 3 个部分.第 1 部分元素全部属于 X,也就是 $\underline{R}(X)$ 部分,它的元素隶属度全为 1.第 2 部分元素全部不属于 X,也就是 $U/\overline{R}(X)$ 部分,它的隶属度全为 0.最后部分元素是边界部分,即 $BN_R(X)=\overline{R}(X)-\underline{R}(X)$,它是否属于 X 是不清晰的,也就是说,它的元素的隶属度在 0~1 之间.文献[26]基于此提出了通过模糊截集的方法来构建了一个在上近似集和下近似集之间的集合 $R_{\lambda}(X)$.

定义 6(X 的λ近似集[26]). 设目标概念 X 是论域 U 的一个子集,令

$$R_{\lambda}(X)=\{x\in U\mid \mu_X^R(x)\geqslant\lambda\},1\geqslant\lambda>0,$$

称 $R_{\lambda}(X)$ 为 X 的λ近似集.

定义 7(X 的λ强近似集[26]). 设目标概念 X 是论域 U 的一个子集,令

$$\underset{\cdot}{R}_{\lambda}(X)=\{x\in U\mid \mu_X^R(x)>\lambda\},1>\lambda>0,$$

称 $\underset{\cdot}{R}_{\lambda}(X)$ 为 X 的λ强近似集.

显然,粗糙集的上近似集和下近似集只是 $R_{\lambda}(X)$ 的两个特例,且满足 $\underline{R}(X)\subseteq R_{\lambda}(X)\subseteq\overline{R}(X)$. 文献[26]给出了 $R_{\lambda}(X)$ 和 $\underset{\cdot}{R}_{\lambda}(X)$ 的一些简单性质,但 $R_{\lambda}(X)$ 的性质不止这一些.

通常情况下,粗糙集的上近似集和下近似集可满足以下性质[22,31]:

(1) $\underline{R}(\sim X)=\sim\overline{R}(X),\overline{R}(\sim X)=\sim\underline{R}(X)$.

(2) 若$X\subseteq Y$,则$\underline{R}(X)\subseteq\underline{R}(Y),\overline{R}(X)\subseteq\overline{R}(Y)$.

(3) $\underline{R}(X\cap Y)\subseteq\underline{R}(X)\cap\underline{R}(Y),\overline{R}(X\cap Y)\subseteq\overline{R}(X)\cap\overline{R}(Y)$.

(4) $\underline{R}(X\cup Y)\supseteq\underline{R}(X)\cup\underline{R}(Y),\overline{R}(X\cup Y)\supseteq\overline{R}(X)\cup\overline{R}(Y)$.

粗糙集的近似集 $R_{\lambda}(X)$ 与强近似集 $\underset{\cdot}{R}_{\lambda}(X)$ 也有类似的性质.

性质 1. 设 X,Y 是论域 U 上的两个子集(概念),则,

(1) $R_{\lambda}(\sim X)=\sim\underset{\cdot}{R}_{1-\lambda}(X),\underset{\cdot}{R}_{\lambda}(\sim X)=\sim R_{1-\lambda}(X)$.

(2) 若$X\subseteq Y$,则$R_{\lambda}(X)\subseteq R_{\lambda}(Y),\underset{\cdot}{R}_{\lambda}(X)\subseteq\underset{\cdot}{R}_{\lambda}(Y)$.

(3) $R_{\lambda}(X\cup Y)\subseteq R_{\lambda}(X)\cup R_{\lambda}(Y),\underset{\cdot}{R}_{\lambda}(X\cup Y)\subseteq\underset{\cdot}{R}_{\lambda}(X)\cup\underset{\cdot}{R}_{\lambda}(Y)$.

(4) $R_{\lambda}(X\cap Y)\supseteq R_{\lambda}(X)\cap R_{\lambda}(Y),\underset{\cdot}{R}_{\lambda}(X\cap Y)\supseteq\underset{\cdot}{R}_{\lambda}(X)\cap\underset{\cdot}{R}_{\lambda}(Y)$.

证明:

(1) 因为

$$R_{\lambda}(\sim X)=\{x\mid\mu_{\tilde{X}}^R(x)\geqslant\lambda\}=\left\{x\mid\frac{|[x]_R\cap(\sim X)|}{|[x]_R|}\geqslant\lambda\right\}=\left\{x\mid\frac{|[x]_R\cap(U-X)|}{|[x]_R|}\geqslant\lambda\right\}$$

$$=\left\{x\mid 1-\frac{|[x]_R\cap X|}{|[x]_R|}\geqslant\lambda\right\}=\left\{x\mid\frac{|[x]_R\cap X|}{|[x]_R|}\leqslant 1-\lambda\right\}=\sim\underset{\cdot}{R}_{1-\lambda}(X),$$

$$\underset{\cdot}{R}_{\lambda}(\sim X)=\left\{x\mid\mu_{\tilde{X}}^R(x)>\lambda\right\}=\left\{x\mid\frac{|[x]_R\cap(\sim X)|}{|[x]_R|}>\lambda\right\}=\left\{x\mid\frac{|[x]_R\cap(U-X)|}{|[x]_R|}>\lambda\right\}$$

$$=\left\{x\mid 1-\frac{|[x]_R\cap X|}{|[x]_R|}>\lambda\right\}=\left\{x\mid\frac{|[x]_R\cap X|}{|[x]_R|}<1-\lambda\right\}=\sim R_{1-\lambda}(X),$$

因此,性质(1)成立.

(2) 对任意元素 $x\in R_{\lambda}(X)$, 有 $x\in\left\{x\mid\frac{|[x]_R\cap X|}{|[x]_R|}\geqslant\lambda\right\}\subseteq X$; 因为 $X\subseteq Y$, 故 $\frac{|[x]_R\cap X|}{|[x]_R|}\leqslant\frac{|[x]_R\cap Y|}{|[x]_R|}$. 所以 $x\in\left\{x\mid\frac{|[x]_R\cap X|}{|[x]_R|}\geqslant\lambda\right\}\subseteq Y$. 因此 $x\in R_{\lambda}(Y)$, 故 $R_{\lambda}(X)\subseteq R_{\lambda}(Y)$. 同理,对任意元素 $x\in\underset{\cdot}{R}_{\lambda}(X)$, 有 $x\in\left\{x\mid\frac{|[x]_R\cap X|}{|[x]_R|}>\lambda\right\}\subseteq X$; 因为 $X\subseteq Y$,故 $\frac{|[x]_R\cap X|}{|[x]_R|}\leqslant\frac{|[x]_R\cap Y|}{|[x]_R|}$. 所以, $x\in\left\{x\mid\frac{|[x]_R\cap X|}{|[x]_R|}>\lambda\right\}\subseteq Y$. 因此, $x\in\underset{\cdot}{R}_{\lambda}(Y)$, 所以 $\underset{\cdot}{R}_{\lambda}(X)\subseteq\underset{\cdot}{R}_{\lambda}(Y)$.

故性质(2)成立.

(3) 因为 $X\subseteq X\cup Y$ 且 $Y\subseteq X\cup Y$, 由性质(2)可知, $R_{\lambda}(X)\subseteq R_{\lambda}(X\cup Y)$ 并且 $R_{\lambda}(Y)\subseteq R_{\lambda}(X\cup Y)$, 因此 $R_{\lambda}(X\cup Y)\supseteq R_{\lambda}(X)\cup R_{\lambda}(Y)$.同理,因为 $X\subseteq X\cup Y$ 且 $Y\subseteq X\cup Y$, 由性质(2)可知, $\underset{\cdot}{R}_{\lambda}(X)\subseteq \underset{\cdot}{R}_{\lambda}(X\cup Y)$ 并且 $\underset{\cdot}{R}_{\lambda}(Y)\subseteq \underset{\cdot}{R}_{\lambda}(X\cup Y)$, 因此 $\underset{\cdot}{R}_{\lambda}(X\cup Y)\supseteq \underset{\cdot}{R}_{\lambda}(X)\cup \underset{\cdot}{R}_{\lambda}(Y)$.故性质(3)成立.值得一提的是, $R_{\lambda}(X\cup Y)\subseteq R_{\lambda}(X)\cup R_{\lambda}(Y)$ 不一定成立, $\underset{\cdot}{R}_{\lambda}(X\cup Y)\subseteq \underset{\cdot}{R}_{\lambda}(X)\cup \underset{\cdot}{R}_{\lambda}(Y)$ 也不一定成立.

(4) 因为 $X\cap Y\subseteq X$ 且 $X\cap Y\subseteq Y$, 由性质(2)可知, $R_{\lambda}(X\cap Y)\subseteq R_{\lambda}(X)$ 且 $R_{\lambda}(X\cap Y)\subseteq R_{\lambda}(Y)$,因此 $R_{\lambda}(X\cap Y)\subseteq R_{\lambda}(X)\cap R_{\lambda}(Y)$. 同理,因为 $X\cap Y\subseteq X$ 且 $X\cap Y\subseteq Y$, 由性质(2)可知, $\underset{\cdot}{R}_{\lambda}(X\cap Y)\subseteq \underset{\cdot}{R}_{\lambda}(X)$ 并且 $\underset{\cdot}{R}_{\lambda}(X\cap Y)\subseteq \underset{\cdot}{R}_{\lambda}(X\cap Y)\subseteq \underset{\cdot}{R}_{\lambda}(Y)$, 因此 $\underset{\cdot}{R}_{\lambda}(X\cap Y)\subseteq \underset{\cdot}{R}_{\lambda}(X)\cap \underset{\cdot}{R}_{\lambda}(Y)$, 故性质(4)成立.值得一提的是, $R_{\lambda}(X\cap Y)\supseteq R_{\lambda}(X)\cap R_{\lambda}(Y)$ 不一定成立, $\underset{\cdot}{R}_{\lambda}(X\cap Y)\supseteq \underset{\cdot}{R}_{\lambda}(X)\cap \underset{\cdot}{R}_{\lambda}(Y)$ 也不一定成立. □

显然,当我们取 λ=0.5 时就得到文献[26]中提到的 $R_{0.5}(X)$的性质.

(1) $\underset{\cdot}{R}_{0.5}(\sim X)=\sim R_{0.5}(X)$.

(2) 若 $X\subseteq Y$,则 $R_{0.5}(X)\subseteq R_{0.5}(Y)$.

(3) $R_{0.5}(X\cup Y)\supseteq R_{0.5}(X)\cup R_{0.5}(Y)$.

(4) $R_{0.5}(X\cap Y)\supseteq R_{0.5}(X)\cap R_{0.5}(Y)$.

故 $R_{0.5}(X)$ 的性质只是 $R_{\lambda}(X)$ 的性质的一个特例.

3 粗糙集的最优近似集

按照本文定义如果直接用下近似集 $\underline{R}(X)$ 作为目标概念 X 的近似集,则相似度为

$$S(X,\underline{R}(X))=\frac{|X\cap \underline{R}(X)|}{|X\cup \underline{R}(X)|}=\frac{|\underline{R}(X)|}{|X|}.$$

如果直接用上近似集 $\overline{R}(X)$ 作为目标概念 X 的近似集,则相似度为

$$S(X,\overline{R}(X))=\frac{|X\cap \overline{R}(X)|}{|X\cup \overline{R}(X)|}=\frac{|X|}{|\overline{R}(X)|},$$

则 $R_{\lambda}(X)$ 中有没有比 $\overline{R}(X)$ 和 $\underline{R}(X)$ 更好的近似集呢?如果有,则应满足 $S(X,\overline{R}(X))\leqslant S(X,R_{\lambda}(X))$ 且 $S(X,\underline{R}(X))\leqslant S(X,R_{\lambda}(X))$.

在知识表达系统 $S=\langle U,A,V,f\rangle$ 下,当目标概念 X 给定时,$|X|$是一个定值,$|\underline{R}(X)|$也是一个定值,为书写方便,我们令 $\dfrac{|\underline{R}(X)|}{|X|+|\underline{R}(X)|}=\sigma_X$.下面讨论 $R_{\lambda}(X)$ 应满足的条件.

定理 1. 设论域 U 是一个有限论域,R 是 U 上的等价关系,若 $\lambda\geqslant\sigma_X$, 则 $S(X,\underline{R}(X))\leqslant S(X,R_{\lambda}(X))$.特别地,当 $\lambda=\sigma_X$ 时,有 $S(X,\underline{R}(X))=S(X,R_{\lambda}(X))$.

证明:对任意的 $x\in R_{\lambda}(X)$, 有 $\mu_X^R(x)\geqslant\lambda$, 即 $\mu_X^R(x)=\dfrac{|[x]_R\cap X|}{|[x]_R|}\geqslant\lambda$.因为 R 是 U 上的等价关系,由 R 在 U 上形成的等价类记为 $[x_1]_R,[x_2]_R,\dots,[x_n]_R$, 则

$$R_{\lambda}(X)=\{x\mid\mu_X^R(x)\geqslant\lambda\}=\{x\mid\mu_X^R(x)=1\}\cup\{x\mid\lambda\leqslant\mu_X^R(x)<1\}.$$

显然, $\{x\mid\mu_X^R(x)=1\}=\underline{R}(X)$. 现令 $\{x\mid\lambda\leqslant\mu_X^R(x)<1\}=[x_{i_1}]_R,[x_{i_2}]_R,\dots,[x_{i_k}]_R$, 因此,

$$X\cap\underline{R}(X)=X\cap(\underline{R}(X)\cup[x_{i_1}]_R\cup[x_{i_2}]_R\cup\dots\cup[x_{i_k}]_R).$$

又因为 $\underline{R}(X)$, $[x_{i_1}]_R,[x_{i_2}]_R,\dots,[x_{i_k}]_R$ 中任意两个相交均为空集,故

$$\begin{aligned}|X\cap R_{\lambda}(X)|&=|X\cap\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\dots+|X\cap[x_{i_k}]_R|\\&=|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\dots+|X\cap[x_{i_k}]_R|.\end{aligned}$$

又 $X\cup R_{\lambda}(X)=X\cup([x_{i_1}]_R-X)\cup([x_{i_2}]_R-X)\cup\dots\cup([x_{i_k}]_R-X)$, 且 $X,([x_{i_1}]_R-X),([x_{i_2}]_R-X),\dots,([x_{i_k}]_R-X)$ 中任意两个集合的交集均为空集,因此,

$$|X\cup R_\lambda(X)|=|X|+|[x_{i_1}]_R-X|+|[x_{i_2}]_R-X|+\ldots+|[x_{i_k}]_R-X|.$$

故 $$S(X,R_\lambda(X))=\frac{|X\cap R_\lambda(X)|}{|X\cup R_\lambda(X)|}=\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|}{|X|+|[x_{i_1}]_R-X|+|[x_{i_2}]_R-X|+\ldots+|[x_{i_k}]_R-X|}.$$

又因为 $\mu_X^R(x)\geqslant\lambda$,所以对于任意的 j=1,2,…,k 有

$$\frac{|[x_{i_j}]_R\cap X|}{|[x_{i_j}]_R|}=\frac{|[x_{i_j}]_R\cap X|}{|[x_{i_j}]_R\cap X|+|[x_{i_j}]_R-X|}\geqslant\lambda,$$

所以,$\dfrac{|[x_{i_j}]_R\cap X|}{|[x_{i_j}]_R-X|}\geqslant\dfrac{\lambda}{1-\lambda}$,即 $|[x_{i_j}]_R-X|\leqslant\dfrac{1-\lambda}{\lambda}|[x_{i_j}]_R\cap X|$,带入 $S(X,R_\lambda(X))$ 有,

$$\begin{aligned}S(X,R_\lambda(X))&=\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|}{|X|+|[x_{i_1}]_R-X|+|[x_{i_2}]_R-X|+\ldots+|[x_{i_k}]_R-X|}\\&\geqslant\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|}{|X|+\frac{1-\lambda}{\lambda}(|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|)}.\end{aligned}$$

现在令 $\dfrac{|\underline{R}(X)|}{|X|}=t$, $|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|=k$,因为 $\lambda\geqslant\sigma_X=\dfrac{|\underline{R}(X)|}{|X|+|\underline{R}(X)|}$,故 $\dfrac{1-\lambda}{\lambda}\leqslant\dfrac{|X|}{|\underline{R}(X)|}=\dfrac{1}{t}$,因此,

$$\begin{aligned}S(X,R_\lambda(X))&\geqslant\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|}{|X|+\frac{1-\lambda}{\lambda}(|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|)}\\&=\frac{t|X|+k}{|X|+\frac{1-\lambda}{\lambda}k}\geqslant\frac{t|X|+k}{|X|+\frac{1}{t}k}=t=S(X,\underline{R}(X)).\end{aligned}$$

故当 $\lambda\geqslant\sigma_X$ 成立时有 $S(X,\underline{R}(X))\leqslant S(X,R_\lambda(X))$,并且当 $\lambda=\sigma_X$ 时,$S(X,\underline{R}(X))=S(X,R_\lambda(X))$. □

例 1:设论域 $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\}$,论域 U 在属性集 R 下被划分为 $U/R=\{\{x_1,x_2,x_3\},\{x_4,x_5,x_6\},\{x_7,x_8\}\}$,令目标概念 $X=\{x_1,x_2,x_3,x_4,x_7\}$,显然论域 U 中每个元素的隶属度为 $F_X^R=\left\{1,1,1,\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{2},\frac{1}{2}\right\}$,因此,$R_{0.5}(X)=\{x_1,x_2,x_3,x_7,x_8\},\underline{R}(X)=\{x_1,x_2,x_3\}$.因为 $\sigma_X=\dfrac{|\underline{R}(X)|}{|X|+|\underline{R}(X)|}=\dfrac{3}{8}<\dfrac{1}{2}$,而 $S(X,\underline{R}(X))=\dfrac{3}{5},S(X,R_{0.5}(X))=\dfrac{2}{3}$,显然,当 $\lambda\geqslant\sigma_X$ 时,有 $S(X,\underline{R}(X))\leqslant S(X,R_\lambda(X))$.

推论 1. 设论域 U 是一个有限论域,R 是 U 上的等价关系,若 $\lambda\geqslant\sigma_X$,则

$$S(X,\underline{R}(X))\leqslant S(X,\underset{\cdot}{R}_\lambda(X)).$$

定理 1 和推论 1 表明,在区间 $[\sigma_X,1]$ 上,$R_\lambda(X)$ 和 $\underset{\cdot}{R}_\lambda(X)$ 总比 $\underline{R}(X)$ 更接近于目标概念 X.显然文献[26]中定理 1 的 $R_{0.5}(X)$ 只是这里的一个特例.下一步讨论 $R_\lambda(X)$ 比 $\overline{R}(X)$ 更近似于 X 的条件.

引理 1[26]. 设 a,b,c 和 d 是实数,且 $0<c<a,0<d<b$,若 $\dfrac{a}{b}\geqslant\dfrac{c}{d}$,则 $\dfrac{a}{b}\leqslant\dfrac{a-c}{b-d}$;若 $\dfrac{a}{b}\leqslant\dfrac{c}{d}$,则 $\dfrac{a}{b}\geqslant\dfrac{a-c}{b-d}$.

在 $\overline{R}(X)$ 中存在一些信息粒,它们都不包含于 $R_\lambda(X)$.不妨设 $\overline{R}(X)-R_\lambda(X)=[x_{j_1}]_R\cup[x_{j_2}]_R\cup\ldots\cup[x_{j_s}]_R$,这里 $[x_{j_1}]_R,[x_{j_2}]_R,\ldots,[x_{j_s}]_R$ 是一族两两不相交的集合,且与 X 都有非空交集,因此,

$$([x_{j_1}]_R\cap X)\cup([x_{j_2}]_R\cap X)\cup\ldots\cup([x_{j_s}]_R\cap X)=X-R_\lambda(X),$$

$$([x_{j_1}]_R-X)\cup(x_{j_2}]_R-X)\cup\ldots\cup([x_{j_s}]_R-X)=\overline{R}(X)-R_\lambda(X)-X.$$

定理 2. 论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关系,若 $\frac{|X|}{|\bar{R}(X)|} \geqslant \frac{|X-R_\lambda(X)|}{|\bar{R}(X)-R_\lambda(X)-X|}$,则 $S(X,\bar{R}(X)) \leqslant S(X,R_\lambda(X))$.

证明:令 $X-R_\lambda(X)=([x_{j_1}]_R\cap X)\cup([x_{j_2}]_R\cap X)\cup\ldots\cup([x_{j_s}]_R\cap X)$,这里的 $[x_{j_1}]_R,[x_{j_2}]_R,\ldots,[x_{j_s}]_R$ 是一族两两不相交的集合.因为 $0<\mu_X^R(x_{j_k})=\frac{|[x_{j_k}]_R\cap X|}{|[x_{j_k}]_R|}<\lambda$, k=1,2,...,s,故 $|[x_{i_k}]_R\cap X|\neq\varnothing$,所以,

$$X\cap R_\lambda(X)=X-([x_{j_1}]_R\cap X)-([x_{j_2}]_R\cap X)-\ldots-([x_{j_s}]_R\cap X)=X-(X-R_\lambda(X)).$$

又因为 $[x_{j_1}]_R,[x_{j_2}]_R,\ldots,[x_{j_s}]_R$ 两两均不相交,因此 $[x_{j_1}]_R\cap X,[x_{j_2}]_R\cap X,\ldots,[x_{j_s}]_R\cap X$ 也两两不相交,所以,

$$X\cap R_\lambda(X)=X-(X-R_\lambda(X)),|X\cap R_\lambda(X)|=|X|-|X-R_\lambda(X)|.$$

与此同时, $X\cup R_\lambda(X)=\bar{R}(X)-(([x_{j_1}]_R-X)\cup([x_{j_2}]_R-X)\cup\ldots\cup([x_{j_s}]_R-X))$. 而 $([x_{j_1}]_R-X),([x_{j_2}]_R-X),\ldots,([x_{j_s}]_R-X)$ 也两两不相交,因此,

$$X\cup R_\lambda(X)=\bar{R}(X)-(\bar{R}(X)-R_\lambda(X)-X),|X\cup R_\lambda(X)|=|\bar{R}(X)|-|\bar{R}(X)-R_\lambda(X)-X|.$$

故
$$S(X,R_\lambda(X))=\frac{|X\cap R_\lambda(X)|}{|X\cup R_\lambda(X)|}=\frac{|X|-|X-R_\lambda(X)|}{|\bar{R}(X)|-|\bar{R}(X)-R_\lambda(X)-X|},$$

又因为 $\frac{|X|}{|\bar{R}(X)|}\geqslant\frac{|X-R_\lambda(X)|}{|\bar{R}(X)-R_\lambda(X)-X|}$,则根据引理 1 可知:

$$\frac{|X|-|X-R_\lambda(X)|}{|\bar{R}(X)|-|\bar{R}(X)-R_\lambda(X)-X|}\geqslant\frac{|X|}{|\bar{R}(X)|}.$$

因此,$S(X,\bar{R}(X))\leqslant S(X,R_\lambda(X))$. □

例 2:设论域 $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\}$,论域 U 在属性集 R 下被划分为 $U/R=\{\{x_1,x_2,x_3\},\{x_4,x_5,x_6\},\{x_7,x_8\}\}$,令目标概念 $X=\{x_1,x_2,x_3,x_4,x_7\}$,显然,论域 U 中每个元素的隶属度为 $F_X^R=\left\{1,1,1,\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{2},\frac{1}{2}\right\}$,由计算可知 $S(X,\bar{R}(X))=\frac{5}{8},S(X,R_{0.5}(X))=\frac{2}{3}$,显然当 $\frac{|X|}{|\bar{R}(X)|}>\frac{|X-R_{1/2}(X)|}{|\bar{R}(X)-R_{1/2}(X)-X|}$ 时,有 $S(X,\bar{R}(X))<S(X,R_{0.5}(X))$.

推论 2. 设论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关系,若 $\frac{|X|}{|\bar{R}(X)|}\geqslant\frac{|X-\underset{\cdot}{R}_\lambda(X)|}{|\bar{R}(X)-\underset{\cdot}{R}_\lambda(X)-X|}$,则 $S(X,\bar{R}(X))\leqslant S(X,\underset{\cdot}{R}_\lambda(X))$.

由定理 2 和推论 2 可得 $R_\lambda(X)$ 和 $\underset{\cdot}{R}_\lambda(X)$ 比 $\bar{R}(X)$ 更近似于目标概念 X 的条件.显然,文献[26]中定理 2 的 $R_{0.5}(X)$ 也只是这里的一个特例.

上面证明得到比 $\underline{R}(X)$ 更近似于 X 的λ成立区间,下面接着找寻最近似于 X 的λ的取值或者条件.

定理 3[26]. 设论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关系,对任意的 $0.5<\lambda\leqslant1$,有

$$S(X,\underline{R}(X))\leqslant S(X,R_\lambda(X))\leqslant S(X,R_{0.5}(X)).$$

定理 4[26]. 设论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关系,对任意的 $0.5\leqslant\lambda_1\leqslant\lambda_2\leqslant1$,则

$$S(X,R_{\lambda_2}(X))\leqslant S(X,R_{\lambda_1}(X)).$$

定理 3 和定理 4 说明,当 l 在区间[0.5,1]上时,$R_\lambda(X)$ 一定比 $\underline{R}(X)$ 更接近于目标概念 X,并且 $R_\lambda(X)$ 与 X 的相似度函数 $S(X,R_\lambda(X))$ 是λ的单调递减函数,此时 $R_{0.5}(X)$ 是目标概念(集合)X 的最优近似集.那么,$R_{0.5}(X)$ 是区间 $[\sigma_X,1]$ 上的最优近似集么?这一结果不一定成立,下面举例说明.

例 3:在决策信息表(表 1)中,论域 $U=\{x_1,x_2,\ldots,x_{22}\}$,条件属性 $C=\{a,b,c,d,e\}$,决策属性 $D=\{f\}$.

Table 1 Decision information table

表 1 决策信息表

	a	b	c	d	e	f
x_1	1	0	0	0	0	1
x_2	1	0	0	0	0	1
x_3	1	0	0	0	0	0
x_4	1	0	0	0	0	0
x_5	1	0	0	0	0	0
x_6	0	1	0	0	0	1
x_7	0	1	0	0	0	1
x_8	0	1	0	0	0	0
x_9	0	1	0	0	0	0
x_{10}	0	1	0	0	0	0
x_{11}	0	0	1	0	0	1
x_{12}	0	0	1	0	0	1
x_{13}	0	0	1	0	0	0
x_{14}	0	0	1	0	0	0
x_{15}	0	0	1	0	0	0
x_{16}	0	0	0	1	0	1
x_{17}	0	0	0	1	0	1
x_{18}	0	0	0	1	0	0
x_{19}	0	0	0	1	0	0
x_{20}	0	0	0	1	0	0
x_{21}	0	0	0	0	1	1
x_{22}	0	0	0	0	1	1

根据粗糙集理论可知:

$$U/IND(C)=\{\{x_1,x_2,x_3,x_4,x_5\},\{x_6,x_7,x_8,x_9,x_{10}\},\{x_{11},x_{12},x_{13},x_{14},x_{15}\},\{x_{16},x_{17},x_{18},x_{19},x_{20}\},\{x_{21},x_{22}\}\},$$

$$U/IND(D)=\{\{x_1,x_2,x_6,x_7,x_{11},x_{12},x_{16},x_{17},x_{21},x_{22}\},\{x_3,x_4,x_5,x_8,x_9,x_{10},x_{13},x_{14},x_{15},x_{18},x_{19},x_{20}\}\}.$$

令决策属性形成的 2 个概念分别为

$$X_1=\{x_1,x_2,x_6,x_7,x_{11},x_{12},x_{16},x_{17},x_{21},x_{22}\},$$

$$X_2=\{x_3,x_4,x_5,x_8,x_9,x_{10},x_{13},x_{14},x_{15},x_{18},x_{19},x_{20}\}.$$

根据定义 $\mu_X^R(x)=\dfrac{|[x]_R\cap X|}{|[x]_R|}$ 可知:

$$F_{X_1}(U)=\left\{\frac{0.4}{x_1},\frac{0.4}{x_2},\ldots,\frac{0.4}{x_{20}},\frac{1}{x_{21}},\frac{1}{x_{22}}\right\},$$

此时,

$$\sigma_{X_1}=\frac{1}{6},R_{0.5}(X_1)=\{x_{21},x_{22}\},R_{0.4}(X_1)=U,$$

而 $S(X_1,R_{0.5}(X_1))=\dfrac{|X_1\cap R_{0.5}(X_1)|}{|X_1\cup R_{0.5}(X_1)|}=0.2$, $S(X_1,R_{0.4}(X_1))=\dfrac{|X_1\cap R_{0.4}(X_1)|}{|X_1\cup R_{0.4}(X_1)|}=0.5$, 显然, $S(X_1,R_{0.4}(X_1))>S(X_1,R_{0.5}(X_1))$.

同理,

$$F_{X_2}(U)=\left\{\frac{0.6}{x_1},\frac{0.6}{x_2},\ldots,\frac{0.6}{x_{20}},\frac{0}{x_{21}},\frac{0}{x_{22}}\right\}.$$

此时有,

$$\sigma_{X_2}=0,R_{0.5}(X_2)=\{x_1,x_2,\ldots,x_{20}\},R_{0.4}(X_2)=\{x_1,x_2,\ldots,x_{20}\}.$$

因此, $R_{0.5}(X_2)=R_{0.4}(X_2)$, 故 $S(X_2,R_{0.4}(X_2))=S(X_2,R_{0.5}(X_2))$. 综上可知,这里的 $R_{0.5}(X_1)$ 不是 X_1 的最优近似集但 $R_{0.5}(X_2)$ 是 X_2 的最优近似集.那么,满足什么条件的 $R_{0.5}(X)$ 才是最优的近似集呢?

当 $\lambda\in[\sigma_X,0.5]$ 时, $R_\lambda(X)$ 中存在一些知识粒,它们都不包含于 $R_{0.5}(X)$ 中,因此不妨设 $R_\lambda(X)-R_{0.5}(X)=[x_{j_1}]_R\cup[x_{j_2}]_R\cup\ldots\cup[x_{j_s}]_R$, 这里 $[x_{j_1}]_R,[x_{j_2}]_R,\ldots,[x_{j_s}]_R$ 是一族两两不相交的集合.

定理 5. 设论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关系,对任意的 $\sigma_X\leqslant\lambda\leqslant 0.5$, 令 $\beta_k=|X\cap[x_{j_k}]_R|/|[x_{j_k}]_R-X|$, k=1,2,…,s,若 $\beta_k\leqslant S(X,R_{0.5}(X))$, 则

$$S(X,R_{\sigma_X}(X))\leqslant S(X,R_\lambda(X))\leqslant S(X,R_{0.5}(X)).$$

证明:

(1) 由定理 1 可知,

$$S(X,R_\lambda(X))\geqslant S(X,\underline{R}(X))=S(X,R_{\sigma_X}(X)).$$

故

$$S(X,R_{\sigma_X}(X))\leqslant S(X,R_\lambda(X)),$$

下面只需证明 $S(X,R_\lambda(X))\leqslant S(X,R_{0.5}(X))$.

(2) 对任意 $x\in R_{0.5}(X)$ 有 $\mu_X^R(x)=\dfrac{|[x]_R\cap X|}{|[x]_R|}\geqslant 0.5$, 令 $\{x\mid\mu_X^R(x)\geqslant 0.5\}=\{x\mid\mu_X^R(x)=1\}\cup\{x\mid 0.5\leqslant\mu_X^R(x)<1\}$, $\{x\mid 0.5\leqslant\mu_X^R(x)<1\}=[x_{i_1}]_R\cup[x_{i_2}]_R\cup\ldots\cup[x_{i_k}]_R$, 而 $\underline{R}(X)=\{x\mid\mu_X^R(x)=1\}$, 由此可得,

$$R_\lambda(X)=\{x\mid\mu_X^R(x)\geqslant\lambda\}=\{x\mid\mu_X^R(x)=1\}\cup\{x\mid 0.5\leqslant\mu_X^R(x)<1\}\cup(R_\lambda(X)-R_{0.5}(X))$$
$$=\underline{R}(X)\cup([x_{i_1}]_R\cup[x_{i_2}]_R\cup\ldots\cup[x_{i_k}]_R)\cup([x_{j_1}]_R\cup[x_{j_2}]_R\cup\ldots\cup[x_{j_s}]_R).$$

因此 $X\cap R_\lambda(X)=X\cap\{\underline{R}(X)\cup([x_{i_1}]_R\cup\ldots\cup[x_{i_k}]_R)\cup([x_{j_1}]_R\cup\ldots\cup[x_{j_s}]_R)\}$, 因为 $\underline{R}(X),[x_{i_1}]_R,\ldots,[x_{i_k}]_R,[x_{j_1}]_R,\ldots,[x_{j_s}]_R$ 是一族两两不相交的集合,故

$$|X\cap R_{0.5}(X)|=|X\cap\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|$$
$$=|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|,$$
$$|X\cap R_\lambda(X)|=|X\cap\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_s}]_R|$$
$$=|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_s}]_R|.$$

同理, $X,([x_{i_1}]_R-X),\ldots,([x_{i_k}]_R-X),([x_{j_1}]_R-X),\ldots,([x_{j_s}]_R-X)$ 也是一族两两不相交的集合,故

$$|X\cup R_{0.5}(X)|=|X\cup(\underline{R}(X)\cup[x_{i_1}]_R\cup[x_{i_1}]_R\cup\ldots\cup[x_{i_k}]_R)|$$
$$=|X|+|[x_{i_1}]_R-X|+|[x_{i_2}]_R-X|+\ldots+|[x_{i_k}]_R-X|,$$
$$|X\cup R_\lambda(X)|=|X\cup(\underline{R}(X)\cup[x_{i_1}]_R\cup\ldots\cup[x_{i_k}]_R\cup[x_{j_1}]_R\cup\ldots\cup[x_{j_s}]_R)|$$
$$=|X|+|[x_{i_1}]_R-X|+\ldots+|[x_{i_k}]_R-X|+|[x_{j_1}]_R-X|+\ldots+|[x_{j_s}]_R-X|.$$

因此,

$$S(X,R_{0.5}(X))=\frac{|X\cap R_{0.5}(X)|}{|X\cup R_{0.5}(X)|}=\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+|X\cap[x_{i_2}]_R|+\ldots+|X\cap[x_{i_k}]_R|}{|X|+|[x_{i_1}]_R-X|+|[x_{i_2}]_R-X|+\ldots+|[x_{i_k}]_R-X|},$$

$$S(X,R_\lambda(X))=\frac{|X\cap R_\lambda(X)|}{|X\cup R_\lambda(X)|}=\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_s}]_R|}{|X|+|[x_{i_1}]_R-X|+\ldots+|[x_{i_k}]_R-X|+|[x_{j_1}]_R-X|+\ldots+|[x_{j_s}]_R-X|}.$$

令 $|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|=k_1$, $|X|+|[x_{i_1}]_R-X|+|[x_{i_2}]_R-X|+\ldots+|[x_{i_k}]_R-X|=k_2$, $S(X,R_{0.5}(X))=k$, 则 $S(X,R_{0.5}(X))=\dfrac{k_1}{k_2}=k$.

再令 $t_1=|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_s}]_R|$, $t_2=|[x_{j_1}]_R-X|+\ldots+|[x_{j_s}]_R-X|$, 因为 $\beta_k=|X\cap[x_{j_k}]_R|/|[x_{j_k}]_R-X|$, 且 $\beta_k\leqslant S(X,R_{0.5}(X))$, 故

$$S(X,R_\lambda(X))=\frac{k_1+t_1}{k_2+t_2}<\frac{k_1+t_1}{k_2+\dfrac{t_1}{k}}=\frac{k_1+t_1}{k_2+\dfrac{k_2}{k_1}t_1}=\frac{k_1}{k_2}=S(X,R_{0.5}(X)).$$

因此, $S(X,R_{\sigma_X}(X))\leqslant S(X,R_\lambda(X))\leqslant S(X,R_{0.5}(X))$. □

例 4:设论域 $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}$, 论域 U 在属性集 R 下被划分为 $U/R=\{\{x_1,x_2,\}\{x_3,x_4,x_5,x_6\},\{x_7,x_8,x_9\},\{x_{10}\}\}$, 令目标概念 $X=\{x_5,x_7,x_8,x_{10}\}$, 显然,论域 U 中每个元素的隶属度为 $F_X^R=\Big\{0,0,\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{2}{3},\frac{2}{3},$

$\frac{2}{3},1\}$. 因此 $\underline{R}(X)=\{x_{10}\}$, $R_{0.5}(X)=\{x_7,x_8,x_9,x_{10}\}$, $R_{1/4}(X)=\{x_3,x_4,x_5,x_6,x_7,x_8,x_9,x_{10}\}$. 因为 $\sigma_X=\frac{|\underline{R}(X)|}{|X|+|\underline{R}(X)|}=\frac{1}{5}$, $S(X,R_{0.5}(X))=\frac{3}{5}$ 而 $\beta=\frac{|X\cap[x_3]_R|}{|[x_3]_R-X|}=\frac{1}{3}<\frac{3}{5}$, 此时 $S(X,R_{\sigma_X}(X))=S(X,R_{1/4}(X))=\frac{1}{2}$, 显然,当 $\beta<S(X,R_{0.5}(X))$ 时, 有 $S(X,R_{\sigma_X}(X))=S(X,R_{1/4}(X))<S(X,R_{0.5}(X))$.

推论 3. 设论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关系,对任意的 $\sigma_X\leqslant\lambda\leqslant 0.5$, 令 $\beta_k=|X\cap[x_{j_k}]_R|/|[x_{j_k}]_R-X|$, k=1,2,…,s,若 $\beta_k\leqslant S(X,R_{0.5}(X))$, 则

$$S(X,R_{\sigma_X}(X))\leqslant S(X,R_\lambda(X))\leqslant S(X,R_{0.5}(X)).$$

当 $\lambda\in[\sigma_X,0.5]$ 时, $R_{\sigma_X}(X)$ 中存在一些知识粒,它们都不包含于 $R_{0.5}(X)$ 中,因此不妨设 $R_{\sigma_X}(X)-R_{0.5}(X)=x_1]_R\cup[x_2]_R\cup\ldots\cup[x_n]_R$, 这里 $[x_1]_R,[x_2]_R,\ldots,[x_n]_R$ 是一族两两不相交的集合.

定理 6. 设论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关令系,对任意的 $\sigma_X\leqslant\lambda_1<\lambda_2\leqslant 0.5$, 令 $\phi_k=|X\cap[x_k]_R|/|[x_k]_R-X|$, k=1,2,…,n,若 $\phi_k\leqslant S(X,\underline{R}(X))$, 则

$$S(X,R_{\lambda_1}(X))\leqslant S(X,R_{\lambda_2}(X)).$$

证明:任取 $\lambda\in[\sigma_X,0.5]$, 对任意的 $x\in R_\lambda(X)$ 有 $\mu_X^R(x)=|[x]_R\cap X|/|[x]_R|\geqslant\lambda$, 可令

$$\{x\mid\mu_X^R(x)\geqslant\lambda\}=\{x\mid\mu_X^R(x)=1\}\cup\{x\mid 0.5\leqslant\mu_X^R(x)<1\}\cup\{x\mid\lambda\leqslant\mu_X^R(x)<0.5\},$$

令 $\{x\mid 0.5\leqslant\mu_X^R(x)<1\}=[x_{i_1}]_R\cup[x_{i_2}]_R\cup\ldots\cup[x_{i_k}]_R$, $\{x\mid\lambda\leqslant\mu_X^R(x)<0.5\}=[x_{j_1}]_R\cup[x_{j_2}]_R\cup\ldots\cup[x_{j_s}]_R$, 显然,$\{j_1,j_2,\ldots,j_s\}\subseteq\{1,2,\ldots,n\}$,而 $\{x\mid\mu_X^R(x)=1\}=\underline{R}(X)$, 故,

$$R_\lambda(X)=\underline{R}(X)\cup([x_{i_1}]_R\cup[x_{i_2}]_R\cup\ldots\cup[x_{i_k}]_R)\cup([x_{j_1}]_R\cup[x_{j_2}]_R\cup\ldots\cup[x_{j_s}]_R).$$

因为 $\underline{R}(X),[x_{i_1}]_R,[x_{i_2}]_R,\ldots,[x_{i_k}]_R,[x_{j_1}]_R,[x_{j_2}]_R,\ldots,[x_{j_s}]_R$ 是一族两两不相交的集合,因此,

$$\begin{aligned}|X\cap R_{\lambda_1}(X)|&=|X\cap\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_{p+t}}]_R|\\&=|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_{p+t}}]_R|,\\|X\cap R_{\lambda_2}(X)|&=|X\cap\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_p}]_R|\\&=|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_p}]_R|.\end{aligned}$$

同理, $X,([x_{i_1}]_R-X),\ldots,([x_{i_k}]_R-X),([x_{j_1}]_R-X),\ldots,([x_{j_s}]_R-X)$ 也是一族两两不相交的集合,故

$$\begin{aligned}|X\cup R_{\lambda_1}(X)|&=|X\cup(\underline{R}(X)\cup[x_{i_1}]_R\cup\ldots\cup[x_{i_k}]_R\cup[x_{j_1}]_R\cup\ldots\cup[x_{j_{p+t}}]_R)|\\&=|X|+|[x_{i_1}]_R-X|+\ldots+|[x_{i_k}]_R-X|+|[x_{j_1}]_R-X|+\ldots+|[x_{j_{p+t}}]_R-X|,\\|X\cup R_{\lambda_2}(X)|&=|X\cup(\underline{R}(X)\cup[x_{i_1}]_R\cup\ldots\cup[x_{i_k}]_R\cup[x_{j_1}]_R\cup\ldots\cup[x_{j_p}]_R)|\\&=|X|+|[x_{i_1}]_R-X|+\ldots+|[x_{i_k}]_R-X|+|[x_{j_1}]_R-X|+\ldots+|[x_{j_p}]_R-X|.\end{aligned}$$

因此,

$$\begin{aligned}S(X,R_{\lambda_1}(X))&=|X\cap R_{\lambda_1}(X)|/|X\cup R_{\lambda_1}(X)|\\&=\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_{p+t}}]_R|}{|X|+|[x_{i_1}]_R-X|+\ldots+|[x_{i_k}]_R-X|+|[x_{j_1}]_R-X|+\ldots+|[x_{j_{p+t}}]_R-X|},\end{aligned}$$

$$\begin{aligned}S(X,R_{\lambda_2}(X))&=|X\cap R_{\lambda_2}(X)|/|X\cup R_{\lambda_2}(X)|\\&=\frac{|\underline{R}(X)|+|X\cap[x_{i_1}]_R|+\ldots+|X\cap[x_{i_k}]_R|+|X\cap[x_{j_1}]_R|+\ldots+|X\cap[x_{j_p}]_R|}{|X|+|[x_{i_1}]_R-X|+\ldots+|[x_{i_k}]_R-X|+|[x_{j_1}]_R-X|+\ldots+|[x_{j_p}]_R-X|}.\end{aligned}$$

令 $S(X,R_{\lambda_2}(X))=k,|X\cap R_{\lambda_2}(X)|=k_1,|X\cup R_{\lambda_2}(X)|=k_2$, 故

$$S(X,R_{\lambda_2}(X))=\frac{\left|X\cap R_{\lambda_2}(X)\right|}{\left|X\cup R_{\lambda_2}(X)\right|}=\frac{k_1}{k_2}=k,$$

$$S(X,R_{\lambda_1}(X))=\frac{\left|X\cap R_{\lambda_1}(X)\right|}{\left|X\cup R_{\lambda_1}(X)\right|}=\frac{k_1+\left|X\cap[x_{j_{p+1}}]_R\right|+\ldots+\left|X\cap[x_{j_{p+t}}]_R\right|}{k_2+\left|[x_{j_{p+1}}]_R-X\right|+\ldots+\left|[x_{j_{p+t}}]_R-X\right|}.$$

因为 $\{p+1,\ldots,p+t\}\subseteq\{1,2,\ldots,s\}$ 且 $\phi_k=\left|X\cap[x_{j_k}]_R\right|/\left|[x_{j_k}]_R-X\right|,\phi_k\leqslant S(X,\underline{R}(X)),k=1,2,\ldots,n$,故

$$\begin{aligned}S(X,R_{\lambda_1}(X))&=\frac{k_1+\left|X\cap[x_{j_{p+1}}]_R\right|+\ldots+\left|X\cap[x_{j_{p+t}}]_R\right|}{k_2+\left|[x_{j_{p+1}}]_R-X\right|+\ldots+\left|[x_{j_{p+t}}]_R-X\right|}\\&\leqslant\frac{k_1+S(X,\underline{R}(X))(\left|[x_{j_{p+1}}]_R-X\right|+\ldots+\left|[x_{j_{p+t}}]_R-X\right|)}{k_2+(\left|[x_{j_{p+1}}]_R-X\right|+\ldots+\left|[x_{j_{p+t}}]_R-X\right|)}.\end{aligned}$$

又由定理 1 可知, $S(X,\underline{R}(X))\leqslant S(X,R_{\lambda_2}(X))$, 再令 $\left|[x_{j_{p+1}}]_R-X\right|+\ldots+\left|[x_{j_{p+t}}]_R-X\right|=t$, 因此,

$$\begin{aligned}S(X,R_{\lambda_1}(X))&\leqslant\frac{k_1+S(X,\underline{R}(X))(\left|[x_{j_{p+1}}]_R-X\right|+\ldots+\left|[x_{j_{p+t}}]_R-X\right|)}{k_2+(\left|[x_{j_{p+1}}]_R-X\right|+\ldots+\left|[x_{j_{p+t}}]_R-X\right|)}\\&\leqslant\frac{k_1+S(X,R_{\lambda_2}(X))t}{k_2+t}=\frac{k_1+\frac{k_1}{k_2}t}{k_2t}\\&=\frac{k_1}{k_2}S(X,R_{\lambda_2}(X)).\end{aligned}$$

故 $S(X,R_{\lambda_1}(X))\leqslant S(X,R_{\lambda_2}(X))$. □

例 5:设论域 $U=\{x_1,x_2,x_3,x_4,x_5,x_6,x_7,x_8\}$, 论域 U 在属性集 R 下被划分为 $U/R=\{\{x_1,x_2,x_3\},\{x_4,x_5,x_6,x_7\},\{x_8\}\}$, 令目标概念 $X=\{x_1,x_4,x_8\}$, 显然论域 U 中每个元素的隶属度为 $F_X^R=\left\{\frac{1}{3},\frac{1}{3},\frac{1}{3},\frac{1}{4},\frac{1}{4},\frac{1}{4},\frac{1}{4},1\right\}$. 因此 $R_{0.5}(X)=\{x_{10}\},R_{1/4}(X)=U,R_{1/3}(X)=\{x_1,x_2,x_3,x_8\}$. 因为 $\sigma_X=\frac{\left|\underline{R}(X)\right|}{\left|X\right|+\left|\underline{R}(X)\right|}=\frac{1}{4}$. 而 $\phi_1=\frac{\left|X\cap[x_1]_R\right|}{\left|[x_3]_R-X\right|}=\frac{1}{2},\phi_2=\frac{\left|X\cap[x_4]_R\right|}{\left|[x_4]_R-X\right|}=\frac{1}{3}$. 此时, $S(X,R_{0.5}(X))=\frac{1}{3},S(X,R_{1/3}(X))=\frac{2}{5},S(X,R_{1/4}(X))=\frac{3}{8}$, 这里, $\phi_1\leqslant S(X,R_{0.5}(X))$ 且 $\phi_2\leqslant \mathrm{S}(X,R_{0.5}(X))$, 我们有 $S(X,R_{1/4}(X))<S(X,R_{1/3}(X))$.

推论 4. 设论域 U 是一个有限论域,X 是 U 上的概念,R 是 U 上的等价关系,对任意的 $\sigma_X\leqslant\lambda_1<\lambda_2\leqslant 0.5$, 令 $\phi_k=\left|X\cap[x_k]_R\right|/\left|[x_k]_R-X\right|$, $k=1,2,\ldots,n$,若 $\phi_k\leqslant S(X,\underline{R}(X))$, 则

$$S(X,R_{\lambda_1}(X))\leqslant S(X,R_{\lambda_2}(X)).$$

定理 5 和定理 6 表明,在区间 $[\sigma_X,0.5]$ 上 $R_{0.5}(X)$ 并不一定最优,但当满足一定条件时, $R_{0.5}(X)$ 可作为 X 的最优近似集.

4 结束语

粗糙集理论研究范围广阔,涉及领域众多.它在系统理论、系统开发以及计算机建模过程中发挥着重要作用[29,30],并已取得了许多令人瞩目的成果.在知识获取、决策分析、系统决策支持、数据库知识发现、模式识别、医疗诊断控制等应用领域上,粗糙集理论也发挥着重要作用[32−35].粗糙集理论也是粒计算研究的重要工具,它与模糊集理论,商空间理论一起构成了粒计算的三大数学基础.Pawlak 教授提出的粗糙集理论是描述边界不确定集合的重要手段.它主要是利用两个精确的集合即上近似集和下近似集来限制边界不确定集合 X 目前在应用领域上利用粗糙集的扩展模型做研究者很多,但直接利用 X 的现有的知识基来构建近似集的研究却很少.

我们在此方面做了探讨性的工作,提出了 X 的近似集构建方法,找到一个相对较好的近似集 $R_{0.5}(X)$并发现了它的一些性质.但我们在文献[26]中并没有给出一般近似集的性质以及最优的近似集.结合我们在文献[26]中给出的定义,我们发现了一般近似集所满足的性质.找到了一般近似集比下近似集更近似于 X 的λ成立的区间.最后找到了 $R_{0.5}(X)$作为最优近似集应满足的条件.近似集理论作为一种新方法在不确定信息的处理上有着重要的应用.本文在理论研究取得的成果可以与其他处理不确定信息的理论相结合,使这些理论在处理不确定信息上显得更加有效[36,37];在实际的应用中,它在数据挖掘和知识获取上面也有着相当大的优势[38−40],特别是在图像分割的边缘信息处理中,可能会有更好的效果.这些研究工作完善了我们前面提出的粗糙集的近似集理论,希望这些研究工作能丰富不确定信息度量理论以及增进人工智能的发展,扩展粗糙集理论的应用.

致谢 在此,我们向对本文工作给予修改建议和帮助的评审者、专家和帮助者表示深深的感谢.

References:

[1] Li DY, Liu CY, Du Y, Han X. Artificial intelligence with uncertainty. Ruan Jian Xue Bao/Journal of Software, 2004,15(11):1583−1594 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/1853.htm

[2] Liu J, Chen XP, Cai QS, Fan Y. Recognition structure of uncertainty: A unified framework for representation, reasoning and learning. Ruan Jian Xue Bao/Journal of Software, 2002,13(4):649−651(in Chinese with English abstract). http://www.jos.org.cn/1000-9825/13/649.htm

[3] Zadeh LA. Fuzzy sets. Information and Control, 1965,8(3):338−353. [doi: 10.1016/S0019-9958(65)90241-X]

[4] Pawlak Z. Rough sets. Int'l Journal of Computer & Information Sciences, 1982,11(5):341−356.

[5] Zhang L, Zhang B. The Theory and Applications of Problem Solving-Quotient Space Based Granular Computing. 2rd ed., Beijing: Tsinghua University Press, 2007 (in Chinese).

[6] Yao Y. Interval sets and interval-set algebras. In: Proc. of the 8th IEEE Int'l Conf. on.Cognitive Informatics (ICCI 2009). IEEE, 2009. 307−314. [doi: 10.1109/COGINF.2009.5250723]

[7] Molodtsov D. Soft set theory—first results. Computers & Mathematics with Applications, 1999,37(4-5):19−31. [doi: 10.1016/S0898-1221(99)00056-5]

[8] Wang GY, Yao YY, Yu H. A survey on rough set theory and its application. Chinese Journal of Computers, 2009,32(7):1229−1246 (in Chinese with English abstract). [doi: 10.3724/SP.J.1016.2009.01229]

[9] Sun BZ, Ma WM, Zhao HY. A fuzzy rough set approach to emergency material demand prediction over two universes. Applied Mathematical Modelling, 2013,37(10-11):7062−7070. [doi: 10.1016/j.apm.2013.02.008]

[10] Luo QJ, Wang GJ. Roughness and fuzziness in quantales. Information Science, 2014,271(7):14−30. [doi: 10.1016/j.ins.2014.02.105]

[11] Tiwari SP, Srivastava AK. Fuzzy rough sets, fuzzy preorders and fuzzy topologies. Fuzzy Set and Systems, 2013,210(4):63−68. [doi: 10.1016/j.fss.2012.06.001]

[12] Srinivas K, Rao GR, Govardhan A. Rough-Fuzzy classifier: A system to predict the heart disease by blending two different set theories. Arabian Journal for Science and Engineering, 2014,39(4):2857−2868. [doi: 10.1007/s13369-013-0943-1]

[13] Inuiguchi M, Yoshioka Y, Kusunoki Y. Variable-Precision dominance-based rough set approach and attribute reduction. Int'l Journal of Approximate Reasoning, 2009,50(8):1199−1214. [doi: 10.1016/j.ijar.2009.02.003]

[14] Xie G, Zhang JL, La KK, Yu L. Variable precision rough set for group decision-making: An application. Int'l Journal of Approximate Reasoning, 2007,49(2):331−343. [doi: 10.1016/j.ijar.2007.04.005]

[15] Liu JNK, Hu Y, He Y. A set covering based approach to find the reduct of variable precision rough set. Information Sciences, 2014, 275(3):83−100. [doi: 10.1016/j.ins.2014.02.023]

[16] Pawlak Z, Skowron A. Rough sets: Some extensions. Information Sciences, 2007,177(1):28−40. [doi: 10.1016/j.ins.2006.06.006]

[17] Mi JS, Wu WZ, Zhang WX. Approaches to knowledge reduction based on variable precision rough set model. Information Sciences, 2004,159(3-4):255−272. [doi: 10.1016/j.ins.2003.07.004]

[18] Zhang XY, Mo ZW. Variable precision rough sets. Pattern Recognition and Artificial Intelligence, 2004,17(2):151−155 (in Chinese with English abstract).
[19] Qian YH, Zhang H, Sang YL, Liang JY. Multigranulation decision-theoretic rough sets. Int'l Journal of Approximate Reasoning, 2014,55(1):225−237. [doi: 10.1016/j.ijar.2013.03.004]
[20] Zhang SG, Mi JS, Hu QH. Rough ε-support vector regression model. Journal of Nanjing University (Natural of Science), 2013,49 (5):650−654 (in Chinese with English abstract).
[21] Li SY, Qian YH. Attribute reduction algorithm based on multi-granularity rough decision. Journal of North University of China (Natural of Science), 2013,34(5):589−592 (in Chinese with English abstract).
[22] Ziarko W. Probabilistic approach to rough sets. Int'l Journal of Approximate Reasoning, 2008,49(2):272−284. [doi: 10.1016/j.ijar. 2007.06.014]
[23] Yao YY. Probabilistic rough set approximations. Int'l Journal of Approximate Reasoning, 2008,49(2):255−271. [doi: 10.1016/j.ijar. 2007.05.019]
[24] Yao YY, Li X. Comparison of rough-set and interval-set models for uncertain reasoning. Fundamenta Informaticae, 1996,27(2): 289−298.
[25] Yao YY. Two semantic issues in a probabilistic rough set model. Fundamenta Informaticae, 2011,108(3):249−265.
[26] Zhang QH, Wang GY, Xiao Y. Approximation sets and rough set. Ruan Jian Xue Bao/Journal of Software, 2012,23(7):1745−1759 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/4226.htm [doi: 10.3724/SP.J.1001.2012.04226]
[27] Yang LB, Gao YY. Principles and Applications of Fuzzy Mathematics. 4th ed., Guangzhou: South China University of Technology Press, 2011 (in Chinese).
[28] Wang GY. Rough Set Theory and Knowledge Discovery. Xi'an: Xi'an Jiaotong University Press, 2007 (in Chinese).
[29] Wang GY, Mao DQ, Wu WZ, Liang JY. Uncertain knowledge representation and processing based on rough set. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2010,22(5):541−544 (in Chinese with English abstract).
[30] Miao DQ, Li DY, Yao YY, Wang GY, Zhang YP, Wu ZW, Liang JY, Yu DR, Wang RZ. Uncertainty and Granular Computing. Beijing: Science Press, 2011 (in Chinese).
[31] Zhang WX, Wu WZ. An introduction and a survey for the studies of rough set theory. Fuzzy Systems and Mathematicas, 2000,14 (4):1−12 (in Chinese with English abstract).
[32] Zhang QH, Xiao Y. Fuzziness of rough set with different granularity levels. Journal of Information and Computational Science, 2011,8(3):385−392.
[33] Zhang QH, Xiao Y, Xing Y. The representation and processing of uncertain problems. Procedia Engineering, 2011,15:1958−1962. [doi: 10.1016/j.proeng.2011.08.365]
[34] Zhang QH, Shen W. Research on attribute reduction algorithm with weights. Journal of Intelligent and Fuzzy Systems, 2014,27(2): 1011−1019. [doi: 10.3233/IFS-131062]
[35] Zhang QH, Shen W, Yang QY. Attribute reduction in inconsistent disjunctive set-valued ordered decision information system. Journal of Computational Information Systems, 2012,8(12):5161−5167.
[36] Zhang QH, Wang J, Wang GY, Yu H. The approximation set of a vague set in rough approximation space. Information Sciences, 2015,300:1−19. [doi: 10.1016/j.ins.2014.12.023]
[37] Zhang QH, Wang J, Wang GY, Hu F. Approximation set of the interval set in Pawlak's Space. The Scientific World Journal, 2014, 2014:Article ID 317387. [doi: 10.1155/2014/317387]
[38] Zhang QH, Xiao Y, Shen W. Knowledge acquisition based on fuzzy entropy. Journal of Information & Computational Science, 2012,9(14):4099−4107.
[39] Zhang QH, Guo YL, Xiao Y. Attribute reduction based on approximation set of rough set. Journal of Computational Information Systems, 2014,10(16):6859−6866. [doi: 10.12733/jcis10998]
[40] Xiao Y. The uncertainty measure of concept and its application in knowledge acquisition [MS. Thesis]. Chongqing: Chongqing University of Posts and Telecommunications, 2013.

附中文参考文献:

[1] 李德毅,刘常昱,杜鹢,韩旭.不确定性人工智能.软件学报,2004,15(11):1583−1594. http://www.jos.org.cn/1000-9825/1853.htm

[2] 刘洁,陈小平,蔡庆生,范焱.不确定信息的认知结构表示、推理和学习.软件学报,2002,13(4):649−651. http://www.jos.org.cn/1000-9825/13/649.htm

[5] 张钹,张铃.问题求解理论及运用——商空间粒度计算理论及应用.第2版,北京:清华大学出版社,1990.

[8] 王国胤,姚一豫,于洪.粗糙集理论与应用研究综述.计算机学报,2009,32(7):1229−1246. [doi: 10.3274/SP.J.1016.2009.01229]

[18] 张贤勇,莫智文.变精度粗糙集.模式识别于人工智能,2004,17(2):151−155.

[20] 张仕光,米据生,胡清华.粗糙ε-支持向量回归模型.南京大学学报(自然科学版),2013,49(5):650−654.

[21] 李顺勇,钱宇华.基于多粒度粗糙决策下的属性约简算法.中北大学学报(自然科学版),2013,34(5):589−592.

[26] 张清华,王国胤,肖雨.粗糙集的近似集.软件学报,2012,23(7):1745−1759. http://www.jos.org.cn/1000-9825/4226.htm [doi: 10.3724/SP.J.1001.2012.04226]

[27] 杨纶标,高英仪.模糊数学原理及运用.第4版,广州:华南理工大学出版社,2011.

[28] 王国胤.Rough集理论与知识获取.西安:西安交通大学出版社,2011.

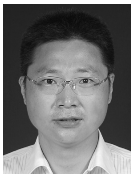
[29] 王国胤,苗夺谦,周志华,吴伟志,梁吉业,不确定信息的粗糙集表示和处理.重庆邮电大学学报(自然科学版),2010,22(5):541−544.

[30] 苗夺谦,李德毅,姚一豫,王国胤,张燕平,吴志伟,梁吉业,于达仁,王睿智.不确定性与粒计算.北京:科学出版社,2011.

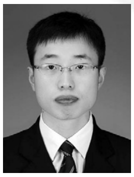
[31] 张文修,吴伟志.粗糙集理论介绍和研究综述.模糊系统与数学,2000,14(4):1−12.

[40] 肖雨.概念的不确定性度量及在知识获取中的应用[硕士学位论文].重庆:重庆邮电大学,2013.

张清华(1974－),男,重庆人,博士,教授,主要研究领域为不确定信息处理,粗糙集与粒计算理论.



王国胤(1970－),男,博士,教授,博士生导师,CCF杰出会员,主要研究领域为粗糙集与粒计算理论,数据挖掘,知识技术.



薛玉斌(1990－),男,硕士生,主要研究领域为不确定信息处理,粗糙集与粒计算理论.