广义变精度粗糙集模型中近似算子研究*)

孙士保[1,2] 普杰信[1] 秦克云[2]
(河南科技大学电子信息工程学院 洛阳 471003)[1] (西南交通大学智能控制开发中心 成都 610031)[2]

摘 要 定义了多数包含关系;借助引入的误差参数 $\beta(0\leqslant\beta<0.5)$,提出了基于后继邻域的广义变精度粗糙集模型的 β 上近似 $\overline{apr}_\beta X$、β 下近似 $\underline{apr}_\beta X$、β 边界 $bnr_\beta X$ 和 β 负域 $negr_\beta X$ 的定义;详细讨论了 β 上、下近似算子 $\underline{apr}_\beta X$ 与 $\overline{apr}_\beta X$ 的性质;从对偶性角度出发推广了 β 上近似、β 下近似算子 $\underline{apr}_\beta X$ 与 $\overline{apr}_\beta X$,得到了两对对偶的上、下近似算子 $\underline{apr}'_\beta X$ 与 $\overline{apr}'_\beta X$ 和 $\underline{apr}''_\beta X$ 与 $\overline{apr}''_\beta X$;最后全面讨论了推广后的两对上、下近似算子 $\underline{apr}'_\beta X$ 与 $\overline{apr}'_\beta X$ 和 $\underline{apr}''_\beta X$ 与 $\overline{apr}''_\beta X$ 的性质,详细分析了它们同广义变精度粗糙集模型中上、下近似算子 $\underline{apr}_\beta X$ 与 $\overline{apr}_\beta X$ 和一般关系下的变精度粗糙集模型中上、下近似算子 $\underline{R}_\beta X$ 与 $\overline{R}_\beta X$ 的关系。

关键词 变精度粗糙集模型,近似算子,二元关系,后继邻域

Research on Approximation Operator in Generalized Variable Precision Rough Set Model

SUN Shi-Bao[1,2] PU Jie-Xin[1] QIN Ke-Yun[2]
(Electronic Information Engineering College, Henan University of Science and Technology, Luoyang 471003)[1]
(Intelligent Control Development Center, Southwest Jiaotong University, Chengdu 610031)[2]

Abstract The majority inclusion relation is defined. A generalized variable precision rough set model based on successor neighbor is proposed by means of the error parameter $\beta(0\leqslant\beta<0.5)$. β upper approximation $\overline{apr}_\beta X$, β lower approximation $\underline{apr}_\beta X$, β, boundary regions $bnr_\beta X$ and β negative field $negr_\beta X$ are defined. The properties of $\overline{apr}_\beta X$ and $\underline{apr}_\beta X$ are discussed in detail. Two pair of dual approximation operators $\underline{apr}'_\beta X$ and $\overline{apr}'_\beta X$ together with $\underline{apr}''_\beta X$ and $\overline{apr}''_\beta X$ are obtained from generalizing $\underline{apr}_\beta X$ and $\overline{apr}_\beta X$. Finally, the properties of two pair of dual approximation operators are discussed, the relation between two pair of dual approximation operators and $\underline{apr}_\beta X$ and $\overline{apr}_\beta X$ together $\underline{R}_\beta X$ and $\overline{R}_\beta X$ are analyzed in detail.

Keywords Variable precision rough set model, Approximation operators, Binary relation, Successor neighbor

1 引言

粗糙集(Rough Sets)理论是 Z. Pawlak 教授于 1982 年提出的一种研究不完整、不确定知识和数据的表达、学习、归纳的理论方法[1]。该理论已经成为智能计算领域的研究热点,并在信息处理、数据挖掘(DM)和数据库知识发现(KDD)等认知领域有成功的应用[2,3]。

粗糙集理论将分类与知识联系在一起,根据已知数据自身的不可分辨关系,通过一对近似算子,对某一给定概念进行近似表示。它是一种数据驱动的方法,本质上不需要任何关于数据和相应问题的先验知识和附加信息,因此特别适合应用于知识发现与数据挖掘领域。Pawlak 粗糙集模型的一个局限性是它所处理的分类必须是完全正确的或肯定的,因而它的分类是精确的,亦即只考虑完全“包含”和“不包含”,而没有某种程度上的“包含”和“属于”。Pawlak 粗糙集模型的另一个局限性是它所处理的对象是已知的,且从模型中得到的结论仅适用于这些对象。但在实际应用中,往往需要把从小规模对象集中得到的结论应用于大规模对象集上去。Pawlak 粗糙集模型的局限性限制了它的应用。为了克服这些局限性,Ziarko 提出了变精度粗糙集模型[4](VPRS 模型,Variable Precision Rough Set Model)。它是 Pawlak 粗糙集模型的扩展,其基本思想是在 Pawlak 粗糙集模型中引入参数 $\beta(0\leqslant\beta<0.5)$,即允许一定程度的错误分类率存在,它可以解决属性间无函数关系的数据分类问题。当 $\beta=0$ 时,变精度粗糙集就退化为 Pawlak 粗糙集。这种推广的模型有利于从看似不相关的数据中发现潜在的相关数据。目前变精度粗糙集模型已经在很多行业得到了广泛的应用[5]。

然而,基于变精度粗糙集模型的推广工作没有得到进一步的研究。到目前为止,有关变精度粗糙集模型的研究主要集中在等价关系下的约简方法和模型推广两个方面。等价关系下的约简方法有 β 约简[6]、β 上(下)近似(分布)约简[7]、基于结构的约简[8]等,它们使得应用变精度粗糙集理论进行信息处理成为可能。但在很多实际问题中,对象之间的等价关系很难构造,或者对象之间本质上就没有等价关系,因此需要对等价关系进行推广。变精度粗糙集模型的推广是把等价关系推广为论域上的模糊关系[9,10]或一般二元关系[11]。文[9,10]主要是把变精度方法引入到模糊粗糙集,探讨模糊粗糙集的理论和应用;文[11]把变精度粗糙集模型中的等价关系推

*)国家自然科学基金资助项目(批准号:60474022)。孙士保 副教授,博士研究生,主要研究领域包括智能信息处理、机器学习和数据挖掘等;普杰信 教授,博士,主要研究领域包括智能信息处理和数据挖掘等;秦克云 教授,博士生导师,主要研究领域包括智能信息处理、机器学习和逻辑与不确定性推理等。

广为论域U上的一般二元关系R，(U,R)为广义近似空间，对于任意的$X\subseteq U$，X关于广义近似空间(U,R)的β下近似和β上近似为

$$\underline{R}_\beta X=\{x\in U:C(R_S(x),X)\leqslant\beta\}$$

$$\overline{R}_\beta X=\{x\in U:C(R_S(x),X)<1-\beta\} \qquad (1)$$

并在此基础上定义了X关于广义近似空间(U,R)的β边界域和β负域，然后讨论了所给模型的一些性质，最后讨论了集合的相对可辨性、近似依赖和属性约简等。但是文[11]中的模型有其固有的局限性：近似算子缺乏良好的性质，如传递性、欧几里得性等。本文提出从x的邻域出发来定义广义变精度粗糙集模型，通过详细讨论这种模型知它比文[11]中的模型具有更好的性质；另外，从满足对偶性的要求出发对这种广义变精度粗糙集模型进行进一步的推广，即从下近似出发找它的对偶上近似和从上近似出发找它的对偶下近似，这样就形成了两对对偶的上、下近似算子$\underline{apr}'_\beta X$与$\overline{apr}'_\beta X$和$\underline{apr}''_\beta X$与$\overline{apr}''_\beta X$；最后全面分析推广后的两对上、下近似算子$\underline{apr}'_\beta X$与$\overline{apr}'_\beta X$和$\underline{apr}''_\beta X$与$\overline{apr}''_\beta X$的性质及它们同本文中定义的广义变精度粗糙集模型中上、下近似算子$\underline{apr}_\beta X$与$\overline{apr}_\beta X$和文[11]中的一般关系下的变精度粗糙集模型中上、下近似$\underline{R}_\beta X$与$\overline{R}_\beta X$的关系。

2 广义变精度粗糙集模型

定义 1[12] 设U是有限非空的论域，$R\subseteq U\times U$是U上一个任意二元关系。对于$\forall x,y\in U$，若xRy，即$(x,y)\in R$，则称y是x的后继(successor)，记$R_p(x)=\{y\in U\mid yRx\}$为$x$的后继邻域；称$x$是$y$的前继(predecessor)，记$R_p(x)=\{y\in U\mid yRx\}$为$x$的前继邻域。称$R$是串行的，若$\forall x\in U,R_S(x)\neq\varnothing$；称$R$是逆串行的，若$\forall x\in U,R_p(x)\neq\varnothing$；称$R$是自反的，若$\forall x\in U,x\in R_s(x)$；称$R$是对称的，若$\forall x,y\in U,x\in R_s(y)$蕴涵$y\in R_s(x)$；称$R$是传递的，若$\forall x,y\in U,y\in R_s(x)$蕴涵$R_s(y)\subseteq R_s(x)$；称$R$是欧几里得的，若$\forall x,y\in U,y\in R_s(x)$蕴涵$R_s(x)\subseteq R_s(y)$。

定义 2[12] 设X和Y表示有限论域U的非空子集，如果对于每一个$e\in X$有$e\in Y$，则称Y包含X，记作$Y\supseteq X$。令

$$C(X,Y)=\begin{cases}1-|X\cap Y|/|X|, & |X|>0\\ 0, & |X|=0\end{cases}$$

其中$|X|$表示集合X的基数。称$C(X,Y)$为集合X关于集合Y的相对错误分类率。即如果我们将集合X中的元素分到集合Y中，则出现分类错误的比例为$C(X,Y)\times 100\%$。

令$0\leqslant\beta<0.5$，多数包含关系定义为

$$Y\overset{\beta}{\supseteq}X\Leftrightarrow C(X,Y)\leqslant\beta$$

显然，$Y\supseteq X$当且仅当$C(X,Y)=0$。

定义 3 设U是有限非空的论域，$R\subseteq U\times U$是U上一个任意二元关系，称(U,R)为广义近似空间。对于任意的子集$X\subseteq U,0\leqslant\beta<0.5$，$X$关于广义近似空间$(U,R)$的$\beta$下近似$\underline{apr}_\beta X$和上近似$\overline{apr}_\beta X$分别定义为

$$\underline{apr}_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)\leqslant\beta\},x\in U$$

$$\overline{apr}_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)<1-\beta\},x\in U \qquad (2)$$

$\underline{apr}_\beta X$也称为β正区域，记为$posr_\beta X$。

相应地，X关于广义近似空间(U,R)的β边界域定义为

$$bnr_\beta X=\bigcup\{R_s(x)\mid\beta<C(R_s(x),X)<1-\beta\},x\in U$$

X关于广义近似空间(U,R)的β负区域定义为

$$negr_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)\geqslant 1-\beta\},x\in U$$

X的β正区域(或X的β下近似)可理解为将U中对象的邻域以不大于β的分类误差分于X的集合。X的β负区域相应理解为将U中对象的邻域以不大于β的分类误差分于X的补集(即$\sim X$)的集合。

定理 1 设(U,R)是广义近似空间，对于$\forall X\subseteq U$，$posr_\beta X=negr_\beta(\sim X)$，其中$\sim X=U-X$。

证明：$posr_\beta X=\underline{apr}_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)\leqslant\beta\}$

$=\bigcup\{R_s(x)\mid 1-\dfrac{|R_s(x)\cap X|}{|R_s(x)|}\leqslant\beta\}=\bigcup\{R_s(x)\mid\dfrac{|R_s(x)\cap X|}{|R_s(x)|}\geqslant 1-\beta\}$

$=\bigcup\{R_s(x)\mid\dfrac{|R_s(x)\cap(\sim X)|}{|R_s(x)|}\leqslant\beta\}=\bigcup\{R_s(x)\mid 1-\dfrac{|R_s(x)\cap(\sim X)|}{|R_s(x)|}\geqslant 1-\beta\}$

$=\bigcup\{R_s(x)\mid C(R_s(x),\sim X)\geqslant 1-\beta\}=negr_\beta(\sim X)$

X的β边界域是由那些以不大于β的分类误差既不能分类于X又不能分类于$\sim X$的U中对象的邻域所构成的集合。如果$bnr_\beta X=\varnothing$，则$posr_\beta X\cup negr_\beta X=U$。

X的β上近似是由那些以不大于β的分类误差不能分类于$\sim X$的U中对象的邻域所构成的集合。

定理 2 设(U,R)是广义近似空间，$0\leqslant\beta<0.5$，则下近似算子$\underline{apr}_\beta X$和上近似算子$\overline{apr}_\beta X$满足的性质为

(1)$\underline{apr}_\beta\varnothing=\overline{apr}_\beta\varnothing=\varnothing$

证明：$\forall x\in U$，当$R_s(x)=\varnothing$时，$C(R_s(x),\varnothing)=0\leqslant\beta$，$\underline{apr}_\beta\varnothing=\overline{apr}_\beta\varnothing=\varnothing$；当存在$R_s(x)\neq\varnothing$，$C(R_s(x),\varnothing)=1>\beta$且$C(R_s(x),\varnothing)=1>1-\beta$，从而$\underline{apr}_\beta\varnothing=\overline{apr}_\beta\varnothing=\varnothing$。

(2)$X\subseteq Y\Rightarrow\underline{apr}_\beta X\subseteq\underline{apr}_\beta Y$和$\overline{apr}_\beta X\subseteq\overline{apr}_\beta Y$

证明：由于$X\subseteq Y$，$C(R_s(x),X)\geqslant C(R_s(x),Y)$，因此$\underline{apr}_\beta X\subseteq\underline{apr}_\beta Y$和$\overline{apr}_\beta X\subseteq\overline{apr}_\beta Y$。

(3)$\underline{apr}_\beta(X\cap Y)\subseteq\underline{apr}_\beta X\cap\underline{apr}_\beta Y$；$\overline{apr}_\beta(X\cup Y)\supseteq\overline{apr}_\beta X\cup\overline{apr}_\beta Y$；$\underline{apr}_\beta(X\cup Y)\supseteq\underline{apr}_\beta X\cup\underline{apr}_\beta Y$；$\overline{apr}_\beta(X\cap Y)\subseteq\overline{apr}_\beta X\cap\overline{apr}_\beta Y$

证明：只证明第一个式子，对于$\forall X,Y\subseteq U$，$C(R_s(x),X\cap Y)\geqslant C(R_s(x),X)$和$C(R_s(x)),X\cap Y\geqslant C(R_s(x),X)$，因此$\underline{apr}_\beta(X\cap Y)\subseteq\underline{apr}_\beta X$且$\underline{apr}_\beta(X\cap Y)\subseteq\underline{apr}_\beta Y$，所以$\underline{apr}_\beta(X\cap Y)\subseteq\underline{apr}_\beta X\cap\underline{apr}_\beta Y$。其余各式证明类似。

(4)$\underline{apr}_\beta X\subseteq\overline{apr}_\beta X$

证明：由上、下近似的定义可直接推出。

(5)$\underline{apr}_\beta X\subseteq\underline{apr}_\beta(\underline{apr}_\beta X)$和$\overline{apr}_\beta X\subseteq\overline{apr}_\beta(\overline{apr}_\beta X)$

证明：只证明第一个式子，$\underline{apr}_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)\leqslant\beta\}$，而对于$\forall x\in U$，若$x\in\underline{apr}_\beta X$，则存在$y\in U$，使$x\in R_s(y)$且$C(R_s(y),X)\leqslant\beta$，从而$R_s(y)\subseteq\underline{apr}_\beta X$，则$C((R_s(y),\underline{apr}_\beta X)=0\leqslant\beta$，从而$R_s(y)\subseteq\underline{apr}_\beta(\underline{apr}_\beta X)$，故$x\in\underline{apr}_\beta(\underline{apr}_\beta X)$，所以$\underline{apr}_\beta X\subseteq\underline{apr}_\beta(\underline{apr}_\beta X)$，证毕。另一式证明类似。

(6)$\overline{apr}_\beta X\subseteq\underline{apr}_\beta(\overline{apr}_\beta X)$和$\underline{apr}_\beta X\subseteq\overline{apr}_\beta(\underline{apr}_\beta X)$

证明：只证明第一个式子，$\underline{apr}_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)<1-\beta\}$，而对于$\forall x\in U$，若$x\in\underline{apr}_\beta X$，则存在$y\in U$，使$x\in R_s(y)$且$C(R_s(y),X)<1-\beta$，从而$R_s(y)\subseteq\overline{apr}_\beta X$，则$C(R_s(y),\overline{apr}_\beta X)=0\leqslant\beta$，从而$R_s(y)\subseteq\underline{apr}_\beta(\underline{apr}_\beta X)$，故$x\in\underline{apr}_\beta(\underline{apr}_\beta X)$，所以$\overline{apr}_\beta X\subseteq\underline{apr}_\beta(\underline{apr}_\beta X)$，证毕。另一式证明类似。

(7)$\alpha\geqslant\beta\Rightarrow\underline{apr}_\alpha X\supseteq\underline{apr}_\beta X$和$\overline{apr}_\alpha X\subseteq\overline{apr}_\beta X$

证明：因为 $\alpha\geqslant\beta$，$\forall C(R_s(x),X)\leqslant\beta$ 必有 $C(R_s(x),X)\leqslant\alpha$，所以 $\underline{apr}_\alpha X\supseteq\underline{apr}_\beta X$ 和 $\overline{apr}_\alpha X\subseteq\overline{apr}_\beta X$。

(8)若 R 是自反的，$\underline{apr}_\beta U=\overline{apr}_\beta U=U$。

证明：当 R 是自反的时，$C(R_s(x),U)\equiv 0$，从而 $\underline{apr}_\beta U=\overline{apr}_\beta U=U$。

3 近似算子的其它定义形式及其性质

由上述性质可以看出，(2)式中的近似算子不是对偶的，这对于近似计算来说是很不方便的。为了解决这对近似算子不能构成对偶的缺陷，可以用两种方法去推广(2)式中的近似算子。

①$\underline{apr}'_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)\leqslant\beta\},x\in U$

$=\{x\in U\mid\exists y[x\in R_s(y),C(R_s(y),X)\leqslant\beta]\}$，

$\overline{apr}'_\beta X=\sim\underline{apr}'_\beta(\sim X)=\sim\{x\in U\mid\exists y[x\in R_s(y),C(R_s(y),\sim X)\leqslant\beta]\}$

$=\{x\in U\mid\forall y[x\in R_s(y)\Rightarrow C(R_s(y),X)\leqslant 1-\beta]\}\cup(\sim R_s(U))$。

②$\underline{apr}''_\beta X=\sim\overline{apr}''_\beta(\sim X)=\sim\{x\in U\mid\exists y[x\in R_s(y),C(R_s(y),\sim X)<1-\beta]\}$

$=\{x\in U\mid\forall y[x\in R_s(y)\Rightarrow C(R_s(y),X)\leqslant\beta]\}\cup(\sim R_s(U))$，

$\overline{apr}''_\beta X=\bigcup\{R_s(x)\mid C(R_s(x),X)<1-\beta\},x\in U$

$=\{x\in U\mid\exists y[x\in R_s(y),C(R_s(y),X)<1-\beta]\}$。

定理 3 设 U 是有限非空的论域，$R\subseteq U\times U$ 是 U 上一个任意二元关系，称 (U,R) 为广义近似空间。对于任意的子集 $X\subseteq U$，$0\leqslant\beta<0.5$，X 关于广义近似空间 (U,R) 的 β 下近似 $\underline{apr}'_\beta X$ 和 β 上近似 $\overline{apr}'_\beta X$ 分别满足下列性质：

(1)$\overline{apr}'_\beta X=\sim\underline{apr}'_\beta(\sim X)$ 和 $\underline{apr}'_\beta X=\sim\overline{apr}'_\beta(\sim X)$；

证明：由定义直接可得。

(2)$X\subseteq Y\Rightarrow\underline{apr}'_\beta X\subseteq\underline{apr}'_\beta Y,\overline{apr}'_\beta X\subseteq\overline{apr}'_\beta Y$；

证明：设 $x\in\underline{apr}'_\beta X$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y),X)\leqslant\beta$。显然对于 $X\subseteq Y,C(R_s(y),Y)\leqslant\beta$，结合 $x\in R_s(y)$，可知 $x\in\underline{apr}'_\beta Y$，于是 $\underline{apr}'_\beta X\subseteq\underline{apr}'_\beta X$。

同理由对偶性可证 $\overline{apr}'_\beta X\subseteq\overline{apr}'_\beta Y$。

(3)$\underline{apr}'_\beta(X\cap Y)\subseteq\underline{apr}'_\beta X\cap\underline{apr}'_\beta Y,\overline{apr}'_\beta(X\cup Y)\supseteq\overline{apr}'_\beta X\cup\overline{apr}'_\beta Y$；

证明：设 $x\in\underline{apr}'_\beta(X\cap Y)$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y)X\cap Y)\leqslant\beta$。显然 $C(R_s(y),X)\leqslant\beta$ 且 $C(R_s(y),Y)\leqslant\beta$，结合 $x\in R_s(y)$，可知 $x\in\underline{apr}'_\beta X$ 且 $x\in\underline{apr}'_\beta Y$，即 $x\in\underline{apr}'_\beta X\cap\underline{apr}'_\beta Y$，于是 $\underline{apr}'_\beta(X\cap Y)\subseteq\underline{apr}'_\beta X\cap\underline{apr}'_\beta Y$。

同理由对偶性可证 $\overline{apr}'_\beta(X\cup Y)\supseteq\overline{apr}'_\beta X\cup\overline{apr}'_\beta Y$。

(4)$\underline{apr}'_\beta(X\cup Y)\supseteq\underline{apr}'_\beta X\cup\underline{apr}'_\beta Y,\overline{apr}'_\beta(X\cap Y)\subseteq\overline{apr}'_\beta X\cap\overline{apr}'_\beta Y$；

证明：设 $x\in\underline{apr}'_\beta X\cup\underline{apr}'_\beta Y$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y),X)\leqslant\beta$ 或 $\exists z\in U,x\in R_s$ 且 $C(R_s(z),Y)\leqslant\beta$。若 $x\in R_s(y)$ 且 $C(R_s(y),X)\leqslant\beta$，显然 $C(R_s(y),X\cup Y)\leqslant\beta$，结合 $x\in R_s(y)$，可得 $x\in\underline{apr}'_\beta(X\cup Y)$；同理若 $x\in R_s(z)$ 且 $C(R_s(z),Y)\leqslant\beta$ 也可得 $x\in\underline{apr}'_\beta(X\cup Y)$。于是 $\underline{apr}'_\beta(X\cup Y)\supseteq\underline{apr}'_\beta X\cup\underline{apr}'_\beta Y$。

同理由对偶性可证 $\overline{apr}'_\beta(X\cap Y)\subseteq\overline{apr}'_\beta X\cap\overline{apr}'_\beta Y$。

(5)$\underline{apr}'_\beta X\subseteq\underline{apr}'_\beta\,\underline{apr}'_\beta X,\overline{apr}'_\beta X\supseteq\overline{apr}'_\beta\,\overline{apr}'_\beta X$；

证明：设 $x\in\underline{apr}'_\beta X$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y),X)\leqslant\beta$。显然对于任意的 $z\in R_s(y)$，有 $z\in\underline{apr}'_\beta X$，从而 $R_s(y)\subseteq\underline{apr}'_\beta X$，故 $C(R_s(y),\underline{apr}'_\beta X)=0\leqslant\beta$，可知 $R_s(y)\subseteq\underline{apr}'_\beta\,\underline{apr}'_\beta X$，结合 $x\in R_s(y)$，知 $x\in\underline{apr}'_\beta\,\underline{apr}'_\beta X$，于是 $\underline{apr}'_\beta X\subseteq\underline{apr}'_\beta\,\underline{apr}'_\beta X$。

同理由对偶性可证 $\overline{apr}'_\beta X\supseteq\overline{apr}'_\beta\,\overline{apr}'_\beta X$。

(6)当 R 是逆串行时，有 $\underline{apr}'_\beta U=U,\overline{apr}'_\beta\varnothing=\varnothing$；

证明：由条件知 $R_s(U)=U$，$\underline{apr}'_\beta U=\bigcup\{R_s(U)\mid C(R_s(U),U)=0\leqslant\beta\}=U$。

同理由对偶性可证 $\overline{apr}'_\beta\varnothing=\varnothing$。

(7)$\underline{apr}'_\beta\varnothing=\varnothing,\overline{apr}'_\beta U=U$；

证明：$\forall x\in U$，当存在 $R_s(x)\neq\varnothing,C(R_s(x),\varnothing)=1>\beta$，$\underline{apr}'_\beta\varnothing=\varnothing$；当存在 $R_s(x)=\varnothing$ 时，$\underline{apr}'_\beta\varnothing=\bigcup\{R_s(x)=\varnothing,C(R_s(x),\varnothing)=0\leqslant\beta\}=\varnothing$。

同理由对偶性可证 $\overline{apr}'_\beta U=U$。

定理 4 设 U 是有限非空的论域，$R\subseteq U\times U$ 是 U 上一个任意二元关系，称 (U,R) 为广义近似空间。对于任意的子集 $X\subseteq U$，$0\leqslant\beta<0.5$，X 关于广义近似空间 (U,R) 的 β 下近似 $\underline{apr}''_\beta X$ 和上近似 $\overline{apr}''_\beta X$ 分别满足下列性质：

(1)$\overline{apr}''_\beta X=\sim\underline{apr}''_\beta(\sim X)$ 和 $\underline{apr}''_\beta X=\sim\overline{apr}''_\beta(\sim X)$；

证明：由定义直接可得。

(2)$X\subseteq Y\Rightarrow\underline{apr}''_\beta X\subseteq\underline{apr}''_\beta Y,\overline{apr}''_\beta X\subseteq\overline{apr}''_\beta Y$；

证明：只证右半部分，设 $x\in\overline{apr}''_\beta X$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y),X)<1-\beta$。显然对于 $X\subseteq Y,C(R_s(y),Y)<1-\beta$，结合 $x\in R_s(y)$，可知 $x\in\overline{apr}''_\beta Y$，于是 $\overline{apr}''_\beta X\subseteq\overline{apr}''_\beta Y$。

同理由对偶性可证 $\underline{apr}''_\beta X\subseteq\underline{apr}''_\beta Y$。

(3)$\overline{apr}''_\beta(X\cup Y)\supseteq\overline{apr}''_\beta X\cup\overline{apr}''_\beta Y,\underline{apr}''_\beta(X\cap Y)\subseteq\underline{apr}''_\beta X\cap\underline{apr}''_\beta Y$；

证明：设 $x\in\overline{apr}''_\beta X\cup\overline{apr}''_\beta Y$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y),X)<1-\beta$ 或 $\exists z\in U,x\in R_s(z)$ 且 $C(R_s(z),Y)<1-\beta$。若 $x\in R_s(y)$ 且 $C(R_s(y),X)<1-\beta$，显然 $C(R_s(y),X\cup Y)<1-\beta$，结合 $x\in R_s(y)$，可知 $x\in\overline{apr}''_\beta(X\cup Y)$；同理，若 $x\in R_s(z)$ 且 $C(R_s(z),Y)<1-\beta$，有 $x\in\overline{apr}''_\beta(X\cup Y)$。于是 $\overline{apr}''_\beta(X\cup Y)\supseteq\overline{apr}''_\beta X\cup\overline{apr}''_\beta Y$。

同理由对偶性可证 $\underline{apr}''_\beta(X\cap Y)\subseteq\underline{apr}''_\beta X\cap\underline{apr}''_\beta Y$。

(4)$\overline{apr}''_\beta(X\cap Y)\subseteq\overline{apr}''_\beta X\cap\overline{apr}''_\beta Y,\underline{apr}''_\beta(X\cup Y)\supseteq\underline{apr}''_\beta X\cup\underline{apr}''_\beta Y$；

证明：设 $x\in\underline{apr}''_\beta(X\cap Y)$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y),X\cap Y)<1-\beta$，显然 $C(R_s(y),X)<1-\beta$ 且 $C(R_s(y),Y)<1-\beta$，结合 $x\in R_s(y)$，可知 $x\in\underline{apr}''_\beta X$ 且 $x\in\underline{apr}''_\beta Y$，即 $x\in\underline{apr}''_\beta X\cap\underline{apr}''_\beta Y$，于是 $\overline{apr}''_\beta(X\cap Y)\subseteq\overline{apr}''_\beta X\cap\overline{apr}''_\beta Y$。

同理由对偶性可证 $\underline{apr}''_\beta(X\cap Y)\supseteq\underline{apr}''_\beta X\cup\underline{apr}''_\beta Y$。

(5)$\overline{apr}''_\beta X\subseteq\overline{apr}''_\beta\,\overline{apr}''_\beta X,\underline{apr}''_\beta X\supseteq\underline{apr}''_\beta\,\underline{apr}''_\beta X$；

证明：设 $x\in\overline{apr}''_\beta X$，由定义知 $\exists y\in U,x\in R_s(y)$ 且 $C(R_s(y),X)<1-\beta$。显然对于任意的 $z\in R_s(y)$，有 $z\in\overline{apr}''_\beta X$，从而 $R_s(y)\subseteq\overline{apr}''_\beta X$，故 $C(R_s(y),\overline{apr}''_\beta X)=0<1-\beta$，可知 $R_s(y)\subseteq\overline{apr}''_\beta\,\overline{apr}''_\beta X$，结合 $x\in R_s(y)$，可知 $x\in\overline{apr}''_\beta\,\overline{apr}''_\beta X$，于是 $\overline{apr}''_\beta X\subseteq\overline{apr}''_\beta\,\overline{apr}''_\beta X$。

同理由对偶性可证 $\underline{apr}''_\beta X\supseteq\underline{apr}''_\beta\,\underline{apr}''_\beta X$。

(6)当 R 是逆串行时，有 $\overline{apr}''_\beta U=U,\underline{apr}''_\beta\varnothing=\varnothing$；

证明：由条件知 $R_s(U)=U$，$\overline{apr}''_\beta U=\bigcup\{R_s(U)\mid C(R_s(U),U)=0<1-\beta\}=U$。

同理由对偶性可证$\underline{apr}''_\beta\varnothing=\varnothing$。

(7)$\overline{apr}''_\beta\varnothing=\varnothing$,$\underline{apr}''_\beta U=U$;

证明:$\forall x\in U$,当存在$R_s(x)\neq\varnothing$,$C(R_s(x),\varnothing)=1>1-\beta$,$\overline{apr}''_\beta\varnothing=\varnothing$;当存在$R_s(x)=\varnothing$时,$\overline{apr}''_\beta\varnothing=\bigcup\{R_s(x)=\varnothing,C(R_s(x),\varnothing)=0<1-\beta\}=\varnothing$。

同理由对偶性可证$\underline{apr}''_\beta U=U$。

4 近似算子之间的关系

近似算子(2)和推广式①与②、(1)和推广式①与②虽然说它们的形式不同,但它们都是反映含有一定分类误差的包含关系,所以它们之间必然具有某种联系。我们经过研究发现近似算子之间具有以下关系。

定理5 设U是有限非空的论域,$R\subseteq U\times U$是U上一个任意二元关系,对于$\forall X\subseteq U$,$0\leqslant\beta<0.5$,若R是逆串行的,则近似算子(2)与推广式①和②满足关系式

$\underline{apr}''_\beta X\subseteq\underline{apr}_\beta X\subseteq\underline{apr}'_\beta X\subseteq\overline{apr}'_\beta X\ \overline{apr}_\beta\subseteq\overline{apr}''_\beta X$

证明:由定理2、定理3和定理4知$\underline{apr}_\beta X\subseteq\overline{apr}_\beta X$、$\underline{apr}'_\beta X\subseteq\overline{apr}'_\beta X$和$\underline{apr}''_\beta X\subseteq\overline{apr}''_\beta X$,所以由对偶性只须证明$\underline{apr}''_\beta X\subseteq\underline{apr}_\beta X\subseteq\underline{apr}'_\beta X$即可。设$\forall x\in\underline{apr}''_\beta X$,由于$R$是逆串行的,因此$R_s(U)=U$,由定义知$\exists y\in U$,$x\in R_s(y)$有$C(R_s(y),X)\leqslant\beta$,由近似算子(2)的定义知$R_s(y)\subseteq\underline{apr}_\beta X$,结合$x\in R_s(y)$,可得$x\in\underline{apr}_\beta X$,于是$\underline{apr}''_\beta X\subseteq\underline{apr}_\beta X$;其次,对于$\forall x\in\underline{apr}_\beta X$,由定义知$\exists y\in U$,$x\in R_s(y)$且$C(R_s(y),X)\leqslant\beta$,从而由定义式①知$\forall x\in\underline{apr}'_\beta X$,于是$\underline{apr}_\beta X\subseteq\underline{apr}'_\beta X$,所以有$\underline{apr}''_\beta X\subseteq\underline{apr}_\beta X\subseteq\underline{apr}'_\beta X$。

定理6 设U是有限非空的论域,$R\subseteq U\times U$是U上一个任意二元关系,对于$\forall X\subseteq U$,$0\leqslant\beta<0.5$,若近似算子(2)的推广式①和②定义的两对近似算子是等价的,则集合类$\{R_s(x)\neq\varnothing\mid x\in U\}$构成了$U$的一个覆盖。

证明:集合类$\{R_s(x)\neq\varnothing\mid x\in U\}$构成了$U$的一个覆盖,即是$R_s(U)=U$。反证,若不然,则$\exists x\in U$且$x\notin R_s(U)$,因此$x\in\underline{apr}''_\beta X$,但显然$x\notin\underline{apr}'_\beta X$,这与$\underline{apr}''_\beta X=\underline{apr}'_\beta X$,对$\forall X\subseteq U$成立相矛盾,从而$R_s(U)=U$。

定理7 设U是有限非空的论域,$R\subseteq U\times U$是U上一个任意二元关系,对于$\forall X\subseteq U$,$0\leqslant\beta<0.5$,若集合类$\{R_s(x)\neq\varnothing\mid x\in U\}$构成了$U$的一个划分,则近似算子(2)的推广式①和②定义的两对近似算子是等价的。

证明:由于集合类$\{R_s(x)\neq\varnothing\mid x\in U\}$构成了$U$的一个划分,因此$R$是逆串行的且$R_s(U)=U$,由定理5知$\forall X\subseteq U$,$0\leqslant\beta<0.5$,$\overline{apr}'_\beta X\subseteq\overline{apr}''_\beta X$。往证$\overline{apr}'_\beta X=\overline{apr}''_\beta X$。反证,若不然,$\exists X\subseteq U$使$\overline{apr}'_\beta X\subset\overline{apr}''_\beta X$,于是存在$x\in\overline{apr}''_\beta X\setminus\overline{apr}'_\beta X$。由$x\in\overline{apr}''_\beta X$知$\exists y\in U$使$x\in R_s(y)$,$C(R_s(y),X)<1-\beta$;而由$x\notin\overline{apr}'_\beta X$知$\exists z\in U$使$x\in R_s(z)$但$C(R_s(z),X)\geqslant1-\beta$。因为$x\in R_s(y)\cap R_s(z)\neq\varnothing$,由条件得$R_s(y)=R_s(z)$,这样$C(R_s(y),X)<1-\beta$与$C(R_s(z),X)\geqslant1-\beta$是相矛盾的,所以$\overline{apr}'_\beta X=\overline{apr}''_\beta X$。

同理由对偶性可证$\underline{apr}''_\beta X=\underline{apr}'_\beta X$。

定理8 设U是有限非空的论域,$R\subseteq U\times U$是U上一个任意二元关系,对于$\forall X\subseteq U$,$0\leqslant\beta<0.5$,若R是自反的,则近似算子(1)与近似算子(2)的推广式①满足关系式$\underline{R}_\beta X\subseteq\underline{apr}'_\beta X\subseteq\overline{apr}'_\beta X\subseteq\overline{R}_\beta X$。

证明:由文[9]及定理3知$\underline{R}_\beta X\subseteq\overline{R}_\beta X$和$\underline{apr}'_\beta X\subseteq\overline{apr}'_\beta X$。由对偶性只须证明$\underline{R}_\beta X\subseteq\underline{apr}'_\beta X$。由于$R$是自反的,因此$x\in R_s(x)$。设$\forall x\in\underline{R}_\beta X$,由定义知$C(R_s(x),X)\leqslant\beta$,从而$R_s(x)\subseteq\underline{apr}'_\beta X$,再由$x\in R_s(x)$可知$x\in\underline{apr}'_\beta X$,所以可以得到$\underline{R}_\beta X\subseteq\underline{apr}'_\beta X$。由对偶性可知另一半成立。

定理9 设U是有限非空的论域,$R\subseteq U\times U$是U上一个任意二元关系,对于$\forall X\subseteq U$,$0\leqslant\beta<0.5$,若R是等价的,则近似算子(1)与近似算子(2)的两个推广式满足关系式$\underline{apr}''_\beta X\subseteq\underline{R}_\beta X\subseteq\underline{apr}'_\beta X\subseteq\overline{apr}'_\beta X\subseteq\overline{R}_\beta X\subseteq\overline{apr}''_\beta X$。

证明:R是等价的,由定理8知$\underline{R}_\beta X\subseteq\underline{apr}'_\beta X\subseteq\overline{apr}'_\beta X\subseteq\overline{R}_\beta X$,又由对偶性只须证明$\underline{apr}''_\beta X\subseteq\underline{R}_\beta X$。设$\forall x\in\underline{apr}''_\beta X$,由于$R$是等价的,因此$R_s(U)=U$,由定义知$\exists y\in U$,$x\in R_s(y)$有$C(R_s(y),X)\leqslant\beta$。由$R$的传递性知$x\in R_s(y)$,有$R_s(y)\subseteq R_s(y)$;又由$R$的对称性知$x\in R_s(y)$,有$y\in R_s(x)$,同样由$R$的传递性知$y\in R_s(x)$,有$R_s(y)\subseteq R_s(x)$。故$R_s(x)=R_s(y)$,所以$C(R_s(x),X)\leqslant\beta$,因此$x\in\underline{R}_\beta X$。于是$\underline{apr}''_\beta X\subseteq\underline{R}_\beta X$。由对偶性可知另一半成立。

结论 本文针对一般关系下基本粗糙集模型的不足,定义了一般关系下的多数包含关系,在此基础上给出了广义变精度粗糙集模型。虽然该模型是广义粗糙集模型和Ziarko变精度粗糙集模型的一般形式,但它的上、下近似算子不满足对偶性,所以从对偶性的角度出发又对广义变精度粗糙集模型进一步进行推广,得到了两对对偶的上、下近似算子,经分析知它们在实际的数据处理中具有广阔的应用前景。但同时也发现,我们处理的数据对象不是最简化的,它是否还有更加简化的形式?如果有我们怎样去约简它?另外,被近似数据集还只是经典集合,存在于数据集上的关系也只是经典关系,当被近似对象变成数据集上的模糊子集或数据集上的经典关系变成模糊关系时,广义变精度粗糙集模型以及两种推广算子显然不能适用于这一类情形,如何在模糊关系下完善文[9,10]中的变精度模糊粗糙集模型,从而让它能够更好地应用于智能信息处理将是我们下一步要解决的内容。

参 考 文 献

1 刘清. Rough集及Rough推理. 北京:科学出版社,2003. 40~80

2 王国胤. 决策表核属性的计算方法. 计算机学报,2003, 26(5): 611~615

3 代建华,潘云鹤. 粗代数研究. 软件学报,2005, 16(7):1197~1204

4 Ziarko W. Variable precision rough set model. Journal of Computer System Science, 1993, 46(1): 39~59

5 陶志,许宝栋,汪定伟,等. 基于变精度粗糙集理论的粗糙规则挖掘算法. 信息与控制,2004,33(1):18~22

6 Beynon M. Reducts within the variable precision rough sets model: A further investigation. European Journal of Operational Research, 2001, 134: 592~605

7 张文修,梁怡,吴伟志,等. 信息系统与知识发现. 北京:科学出版社,2003. 56~67

8 Inuiguchi M. Structure-based Approaches to Attribute Reduction in Variable Precision Rough Set Models. In: Proceeding of 2005 IEEE International Conference on Granular Computing. Beijing China, IEEE GrC2005, 2005. 34~39

9 Mieszkowicz-Rolka A, Rolka L. Remarks on approximation quality in variable precision fuzzy rough sets model[A]. In: Rough Sets and Current Trends in Computing: 4th International Conference, RSCTC 2004[C], Heidelberg: Springer-Verlag, 2004. 402~411

10 Mieszkowicz-Rolka A, Rolka L. Fuzzy implication operators in variable precision fuzzy rough sets model [A]. In: Artifical Intelligence and Soft Computing- ICAISC 2004(volume 3070)[C]. Heidelberg: Springer-Verlag, 2004. 498~503

11 巩增泰,孙秉珍,邵亚斌,等. 一般关系下的变精度粗糙集模型. 兰州大学学报(自然科学版),2005,41(6):110~114

12 张文修,吴伟志,梁吉业,等. 粗糙集理论与方法. 北京:科学出版社,2003. 41~131