串行概率粗糙集近似

马建敏 姚红娟 潘笑晨

(长安大学理学院数学与信息科学系 西安 710064)

摘 要 经典的概率粗糙集模型是基于等价关系和条件概率提出的。但在实际应用中,知识库存在多种不确定性因素,使得对象间的关系未必满足等价关系。因此在保证条件概率有意义的情况下,将等价关系推广到串行二元关系,讨论了串行关系下的概率粗糙集近似;研究了当目标概念发生变化时,串行概率粗糙下、上近似的性质;进一步,通过调整两个阈值,给出了对应的串行概率粗糙下、上近似的变化趋势。

关键词 概率粗糙集,串行概率近似空间,串行概率粗糙集,单调性

中图法分类号 TP182 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.01.012

Serial Probabilistic Rough Set Approximations

MA Jian-min YAO Hong-juan PAN Xiao-chen

(Department of Mathematics and Information Sciences, Faculty of Science, Chang'an University, Xi'an 710064, China)

Abstract The classical probabilistic rough set model was proposed based on an equivalence relation and a conditional probability. However, uncertainty in knowledge base makes it difficult to satisfy the equivalence relation between any two objects. This paper considered the serial binary relation instead of an equivalence relation, making the conditional probability meaningful. Then the serial probabilistic rough set approximations were introduced based on a serial relation. Properties of the serial probabilistic rough lower and upper approximations were discussed when the target concepts are variable. Furthermore, by adjusting the two thresholds, the corresponding serial probabilistic rough lower and upper approximations were also investigated.

Keywords Probabilistic rough set, Serial probabilistic approximation space, Serial probabilistic rough set, Monotonicity

1 引言

波兰学者 Pawlak 于 1982 年提出的粗糙集[1-2]是一种数据分析理论,常被用来处理不精确、不协调和不完全数据等问题。粗糙集理论由于能从大量的数据中挖掘潜在的、有利用价值的知识而被广泛应用于信息系统分析、人工智能、知识工程与数据挖掘、决策支持系统、模式识别与分类等领域[3-4]。

Pawlak 粗糙集理论的基本思想是利用等价关系在论域上产生的划分作为知识,通过这些知识中的基本元(等价类)引入上近似集与下近似集来表示目标概念。经过 30 多年的发展,粗糙集无论在理论上还是在应用领域等方面都取得了很大成功。在等价关系方面,Pawlak 粗糙集可以扩展到不同二元关系下的粗糙集模型[5-11]。在定义方面,Pawlak 粗糙集模型中下近似要求等价类必须完全包含于目标概念的要求太强,从而导致信息丢失,上近似要求等价类与目标概念的交非空的条件太弱而使上近似包含过多的信息。于是,Pawlak 和 Wong 提出了精度为 0.5 的基于条件概率的概率粗糙集[12];Yao 通过引入两个阈值,利用条件概率提出了概率粗糙集[13];Ziarko 提出了基于错误分类率的变精度粗糙集[14]。Yao 将基于条件概率的粗糙隶属函数引入到粗糙集中,并给出了概率粗糙集的统一框架[15]。将概率方法应用到粗糙集中,还产生了决策理论粗糙集模型、Bayesian 粗糙集模型、概率规则归纳模型等。由于概率粗糙集是基于条件概率提出的,而条件概率要求作为条件的事件的概率不能为 0,因此本文考虑串行二元关系下的概率粗糙集。

本文在概率粗糙集的基础上提出了串行概率粗糙集近似,研究了它们的性质,讨论了目标概念发生变化时,串行概率粗糙下、上近似的单调性,以及阈值的变化对串行概率粗糙下、上近似的影响。

2 概率粗糙集近似

设 U 为非空有限论域,$R\subseteq U\times U$ 是 U 上的二元关系。对于任意的 $x,y\in U$,若 x 与 y 具有关系 R,则记作 xRy 或 $(x,y)\in R$。$R_s(x)=\{y\in U\mid xRy\}$ 为 x 的后继领域[16]。

若对于任意的 $x\in U$,$R_s(x)\neq\emptyset$,则称 R 为串行的;若对于任意的 $x\in U$,$x\in R_s(x)$,则称 R 为自反的;若对于任意的 $x,y\in U$,$y\in R_s(x)$ 蕴涵 $x\in R_s(y)$,则称 R 为对称的;若对于任意的 $x,y\in U$,$y\in R_s(x)$ 蕴涵 $R_s(y)\subseteq R_s(x)$,则称 R 为传

到稿日期:2017-03-03 返修日期:2017-07-08 本文受国家自然科学基金项目(10901025,11501048)资助。

马建敏(1978-),女,博士,教授,硕士生导师,主要研究方向为粗糙集、粒计算与概念格,E-mail:cjm-zm@126.com(通信作者);**姚红娟**(1988-),女,硕士,主要研究方向为粗糙集与三支决策;**潘笑晨**(1992-),女,硕士,主要研究方向为粗糙集与粒计算。

递的。显然，自反二元关系必为串行二元关系。满足自反、对称和传递的二元关系 R 为等价关系。当 R 是等价关系时，$R_s(x)$ 就是对象 x 所在的关于 R 的等价类 $[x]_R$。

若 R 为 U 上的串行二元关系，则称 (U,R) 为串行近似空间。若 R 为 U 上的等价关系，则称 (U,R) 为 Pawlak 近似空间。

定义 1[2,17] 设 (U,R) 为 Pawlak 近似空间。对于任意的 $X\subseteq U$，X 关于 Pawlak 近似空间 (U,R) 的下近似和上近似分别定义为：

$$\underline{apr}(X)=\{x\in U\mid [x]_R\subseteq X\}$$

$$\overline{apr}(X)=\{x\in U\mid [x]_R\cap X\neq\emptyset\}$$

X 关于 (U,R) 的正域、负域和边界域分别定义为：

$$POS(X)=\underline{apr}(X)$$

$$NEG(X)=U-\overline{apr}(X)=\{x\in U\mid [x]_R\cap X=\emptyset\}$$

$$BND(X)=\overline{apr}(X)-\underline{apr}(X)$$

在定义 1 给出的 Pawlak 粗糙集近似中，下近似要求等价类完全包含于目标概念 X，这个要求过强从而丢失了对认识目标概念 X 作用很大的等价类；上近似要求等价类与目标概念 X 交非空，这个要求太弱而包含了对认识目标概念作用不大的等价类。Yao 在文献[12-14]的基础上，利用条件概率和两个阈值来控制等价类和目标概念 X 之间的关系，提出了概率粗糙集的统一框架[15]。

设 U 是非空有限论域，Pr 是 U 上的概率测度。对于任意的 $A,B\subseteq U$，$\Pr(A\mid B)=\dfrac{\Pr(A\cap B)}{\Pr(B)}$ $(\Pr(B)>0)$ 表示事件 B 发生的条件下，事件 A 发生的条件概率。

定义 2[15] 设 (U,R) 为 Pawlak 近似空间，Pr 是 U 上的概率测度。对于任意的 $X\subseteq U$，$0\leqslant\beta<\alpha\leqslant 1$，$X$ 关于 (U,R) 的概率粗糙下、上近似分别定义为：

$$\underline{apr}_\alpha(X)=\{x\in U\mid \Pr(X\mid [x]_R)\geqslant\alpha\}$$

$$\overline{apr}_\beta(X)=\{x\in U\mid \Pr(X\mid [x]_R)>\beta\}$$

则 $\underline{apr}_\alpha$ 和 $\overline{apr}_\beta$ 定义了 U 的幂集 2^U 上的一对算子，分别称为概率粗糙下、上近似算子。

对于任意的 $X\subseteq U$，$0\leqslant\beta<\alpha\leqslant 1$，$X$ 关于 (U,R) 的概率正域、概率负域和概率边界域分别定义为：

$$POS_\alpha(X)=\underline{apr}_\alpha(X)=\{x\in U\mid \Pr(X\mid [x]_R)\geqslant\alpha\}$$

$$NEG_\beta(X)=U-\overline{apr}_\beta(X)=\{x\in U\mid \Pr(X\mid [x]_R)\leqslant\beta\}$$

$$BND_{(\alpha,\beta)}(X)=\overline{apr}_\beta(X)-\underline{apr}_\alpha(X)=\{x\in U\mid \beta<\Pr(X\mid [x]_R)<\alpha\}$$

性质 1[15] 设 (U,R) 为 Pawlak 近似空间，Pr 是 U 上的概率测度。则对于任意的 $X,Y\subseteq U$，$0\leqslant\beta<\alpha\leqslant 1$，$\alpha_1\in(0,1]$，且 $\beta_1\in[0,1)$，下列性质成立：

(PLU0) $\underline{apr}_\alpha(X)\subseteq\overline{apr}_\beta(X)$；

(PL1) $\underline{apr}_\alpha(U)=\overline{apr}_\beta(U)=U$；

(PU1) $\underline{apr}_\alpha(\emptyset)=\overline{apr}_\beta(\emptyset)=\emptyset$；

(PL2) $\underline{apr}_\alpha(\sim X)=\sim\overline{apr}_{1-\alpha}(X)$；

(PU2) $\overline{apr}_\beta(\sim X)=\sim\underline{apr}_{1-\beta}(X)$；

(PL3) $\underline{apr}_\alpha(X\cap Y)\subseteq\underline{apr}_\alpha(X)\cap\underline{apr}_\alpha(Y)$；

(PU3) $\overline{apr}_\beta(X\cap Y)\subseteq\overline{apr}_\beta(X)\cap\overline{apr}_\beta(Y)$；

(PL4) $\underline{apr}_\alpha(X\cup Y)\supseteq\underline{apr}_\alpha(X)\cup\underline{apr}_\alpha(Y)$；

(PU4) $\overline{apr}_\beta(X\cup Y)\supseteq\overline{apr}_\beta(X)\cup\overline{apr}_\beta(Y)$；

(PL5) $X\subseteq Y\Rightarrow\underline{apr}_\alpha(X)\subseteq\underline{apr}_\alpha(Y)$；

(PU5) $X\subseteq Y\Rightarrow\overline{apr}_\beta(X)\subseteq\overline{apr}_\beta(Y)$；

(PL6) $\alpha\leqslant\alpha_1\Rightarrow\underline{apr}_{\alpha_1}(X)\subseteq\underline{apr}_\alpha(X)$；

(PU6) $\beta\leqslant\beta_1\Rightarrow\overline{apr}_{\beta_1}(X)\subseteq\overline{apr}_\beta(X)$；

(PL7) $\underline{apr}_\alpha(X)=\overline{apr}_\beta(\underline{apr}_\alpha(X))$；

(PU7) $\overline{apr}_\beta(X)=\underline{apr}_\alpha(\overline{apr}_\beta(X))$；

(PL8) $\underline{apr}_\alpha(X)=\underline{apr}_{\alpha_1}(\underline{apr}_\alpha(X))$；

(PU8) $\overline{apr}_\beta(X)=\overline{apr}_{\beta_1}(\overline{apr}_\beta(X))$。

当 $\alpha=\beta$ 时，定义 2 给出的概率粗糙下、上近似即为 Yao 在文献[19]中给出的下、上近似。且当 $\alpha=1$，$\beta=0$ 时，有：

$$\underline{apr}(X)=\underline{apr}_1(X),\ \overline{apr}(X)=\overline{apr}_0(X)$$

即 Pawlak 粗糙下、上近似是概率粗糙集下、上近似中阈值 $\alpha=1$，$\beta=0$ 时的特殊情况。

3 串行概率粗糙集近似

定义 2 给出的概率粗糙集是基于等价关系和条件概率提出来的。但在实际应用中，对象之间的关系未必满足等价性，弱化概率粗糙集模型中的等价关系是推广概率粗糙集的一种方法[19,23]。此外，条件概率要求做分母的事件的概率必须严格大于零。因此研究串行二元关系下的条件概率是有意义的。下面考虑串行二元关系下的概率粗糙集近似。

设 (U,R) 为串行近似空间，Pr 是 U 上的概率测度，称三元组 $(U,R,\Pr)$ 为串行概率近似空间。

定义 3 设 $(U,R,\Pr)$ 为串行概率近似空间。对于任意的 $X\subseteq U$，$0\leqslant\beta<\alpha\leqslant 1$，$X$ 关于串行概率近似空间 $(U,R,\Pr)$ 的串行概率粗糙下、上近似分别定义为：

$$\underline{Sapr}_\alpha(X)=\{x\in U\mid \Pr(X\mid R_s(x))\geqslant\alpha\}$$

$$\overline{Sapr}_\beta(X)=\{x\in U\mid \Pr(X\mid R_s(x))>\beta\}$$

则 $\underline{Sapr}_\alpha$ 和 $\overline{Sapr}_\beta$ 定义了幂集 2^U 上的一对算子，分别称为串行概率粗糙下、上近似算子。

对于任意的 $X\subseteq U$，$0\leqslant\beta<\alpha\leqslant 1$，$X$ 关于串行概率近似空间 $(U,R,\Pr)$ 的串行概率正域、串行概率负域和串行概率边界域分别定义为：

$$SPOS_\alpha(X)=\underline{Sapr}_\alpha(X)=\{x\in U\mid \Pr(X\mid R_s(x))\geqslant\alpha\}$$

$$SNEG_\beta(X)=U-\overline{Sapr}_\beta(X)=\{x\in U\mid \Pr(X\mid R_s(x))\leqslant\beta\}$$

$$SBND_{(\alpha,\beta)}(X)=\overline{Sapr}_\beta(X)-\underline{Sapr}_\alpha(X)=\{x\in U\mid \beta<\Pr(X\mid R_s(x))<\alpha\}$$

性质 2 设 $(U,R,\Pr)$ 为串行概率近似空间。对于任意的 $X,Y\subseteq U$，$0\leqslant\beta<\alpha\leqslant 1$，$\alpha_1\in(0,1]$，$\beta_1\in[0,1)$，串行概率粗糙上、下近似满足下列性质：

(SLU0) $\underline{Sapr}_\alpha(X)\subseteq\overline{Sapr}_\beta(X)$；

(SL1) $\underline{Sapr}_\alpha(U)=\overline{Sapr}_\beta(U)=U$；

(SU1) $\underline{Sapr}_\alpha(\emptyset)=\overline{Sapr}_\beta(\emptyset)=\emptyset$；

(SL2) $\underline{Sapr}_\alpha(\sim X)=\sim\overline{Sapr}_{1-\alpha}(X)$；

(SU2) $\overline{Sapr}_\beta(\sim X)=\sim\underline{Sapr}_{1-\beta}(X)$；

(SL3) $\underline{Sapr}_{\alpha}(X\cap Y)\subseteq\underline{Sapr}_{\alpha}(X)\cap\underline{Sapr}_{\alpha}(Y)$;

(SU3) $\overline{Sapr}_{\beta}(X\cap Y)\subseteq\overline{Sapr}_{\beta}(X)\cap\overline{Sapr}_{\beta}(Y)$;

(SL4) $\underline{Sapr}_{\alpha}(X\cup Y)\supseteq\underline{Sapr}_{\alpha}(X)\cup\underline{Sapr}_{\alpha}(Y)$;

(SU4) $\overline{Sapr}_{\beta}(X\cup Y)\supseteq\overline{Sapr}_{\beta}(X)\cup\overline{Sapr}_{\beta}(Y)$;

(SL5) $X\subseteq Y\Rightarrow\underline{Sapr}_{\alpha}(X)\subseteq\underline{Sapr}_{\alpha}(Y)$;

(SU5) $X\subseteq Y\Rightarrow\overline{Sapr}_{\beta}(X)\subseteq\overline{Sapr}_{\beta}(Y)$;

(SL6) $\alpha\leqslant\alpha_1\Rightarrow\underline{Sapr}_{\alpha_1}(X)\subseteq\underline{Sapr}_{\alpha}(X)$;

(PU6) $\beta\leqslant\beta_1\Rightarrow\overline{Sapr}_{\beta_1}(X)\subseteq\overline{Sapr}_{\beta}(X)$。

性质 3　设(U,R,Pr)为串行概率近似空间。对于任意的 $X,Y\subseteq U$, $0\leqslant\beta<\alpha\leqslant 1$, $\alpha_1\in(0,1]$, $\beta_1\in[0,1)$,串行概率正域、串行概率负域和串行概率边界域满足下列性质:

(S1) $SPOS_{\alpha}(\sim X)=SNEG_{1-\alpha}(X)$;

(S2) $SNEG_{\beta}(\sim X)=SPOS_{1-\beta}(X)$;

(S3) $SPOS_{\alpha}(X\cap Y)\subseteq SPOS_{\alpha}(X)\cap SPOS_{\alpha}(Y)$;

(S4) $SNEG_{\beta}(X\cap Y)\supseteq SNEG_{\beta}(X)\cup SNEG_{\beta}(Y)$;

(S5) $SPOS_{\alpha}(X\cup Y)\supseteq SPOS_{\alpha}(X)\cup SPOS_{\alpha}(Y)$;

(S6) $SNEG_{\beta}(X\cup Y)\subseteq SNEG_{\beta}(X)\cap SNEG_{\beta}(Y)$;

(S7) $X\subseteq Y\Rightarrow SPOS_{\alpha}(X)\subseteq SPOS_{\alpha}(Y)$,
　　$SNEG_{\beta}(Y)\subseteq SNEG_{\beta}(X)$;

(S8) $\alpha\leqslant\alpha_1\Rightarrow SPOS_{\alpha_1}(X)\subseteq SPOS_{\alpha}(X)$;

(S9) $\beta\leqslant\beta_1\Rightarrow SNEG_{\beta}(X)\subseteq SNEG_{\beta_1}(X)$;

(S10) $SBND_{(\alpha,\beta)}(\sim X)=SBND_{(1-\beta,1-\alpha)}(X)$。

串行概率正域 $SPOS_{\alpha}(X)$、串行概率负域 $SNEG_{\beta}(X)$ 以及串行概率边界域 $SBND_{(\alpha,\beta)}(X)$ 构成了 U 上的一个三划分。特别地,当 $X=U$ 时,由性质 2 及性质 3 可知:

$$SPOS_{\alpha}(U)=U,SNEG_{\beta}(U)=\varnothing,SBND_{(\alpha,\beta)}(U)=\varnothing$$

而当 $X=\varnothing$ 时,可知:

$$SPOS_{\alpha}(\varnothing)=\varnothing,SNEG_{\beta}(\varnothing)=U,SBND_{(\alpha,\beta)}(\varnothing)=\varnothing$$

4　串行概率粗糙集近似的单调性

在串行概率近似空间(U,R,Pr)中,当阈值 α,β 发生变化或目标概念 X 发生变化时,相应地,串行概率粗糙上、下近似中的对象也随之增加或减少。下面讨论当阈值 α,β 发生变化或目标概念 X 发生变化时,对应的串行概率粗糙上、下近似的变化趋势。本节讨论无限论域 U 构成的串行概率近似空间。

由集合序列上、下极限的定义和性质[20-22]可得:

引理 1　设(U,R,Pr)为串行概率近似空间,$\alpha\in(0.5,1]$。任取 U 上的子集序列 $\{X_n\}_{n=1}^{\infty}$,有:

$$\underline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)$$

$$\overline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)$$

引理 2　设(U,R,Pr)为串行概率近似空间,$\beta\in[0,0.5)$。任取 U 上的子集序列 $\{X_n\}_{n=1}^{\infty}$,有:

$$\underline{\lim_{n\to\infty}}\,\overline{Sapr}_{\beta}(X_n)=\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}\overline{Sapr}_{\beta}(X_m)$$

$$\overline{\lim_{n\to\infty}}\,\overline{Sapr}_{\beta}(X_n)=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\overline{Sapr}_{\beta}(X_m)$$

定理 1　设(U,R,Pr)为串行概率近似空间,$\alpha\in(0.5,1]$。任取 U 上的子集序列 $\{X_n\}_{n=1}^{\infty}$。

(1)若 $\{X_n\}_{n=1}^{\infty}$ 单调增加,则有:

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)=\bigcup_{n=1}^{\infty}\underline{Sapr}_{\alpha}(X_n)$$

(2)若 $\{X_n\}_{n=1}^{\infty}$ 单调减少,则有:

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)=\bigcap_{n=1}^{\infty}\underline{Sapr}_{\alpha}(X_n)$$

证明:根据极限存在定理[22]可知,$\lim\limits_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)$存在当且仅当$\overline{\lim\limits_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\underline{\lim\limits_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)$,即:

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)\,\exists\Leftrightarrow\underline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\overline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\lim_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)\quad(1)$$

(1)设 $\{X_n\}_{n=1}^{\infty}$ 单调增加。由性质 2(SL5)可知下近似集合序列 $\{\underline{Sapr}_{\alpha}(X_n)\}_{n=1}^{\infty}$ 是单调增加的。则对于任意的 $n\geqslant 1$,有 $\bigcap\limits_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)=\underline{Sapr}_{\alpha}(X_n)$。故由引理 1 可知:

$$\underline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\bigcup_{n=1}^{\infty}\bigcap_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)=\bigcup_{n=1}^{\infty}\underline{Sapr}_{\alpha}(X_n)$$

又由集合序列 $\{\underline{Sapr}_{\alpha}(X_n)\}_{n=1}^{\infty}$ 单调增加可得:

$$\bigcup_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)=\bigcup_{m=1}^{\infty}\underline{Sapr}_{\alpha}(X_m)$$

于是

$$\overline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)=\bigcap_{n=1}^{\infty}\bigcup_{m=1}^{\infty}\underline{Sapr}_{\alpha}(X_m)$$
$$=\bigcup_{m=1}^{\infty}\underline{Sapr}_{\alpha}(X_m)=\bigcup_{n=1}^{\infty}\underline{Sapr}_{\alpha}(X_n)$$

故$\overline{\lim\limits_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\underline{\lim\limits_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)$。由式(1)可知:

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)=\underline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\bigcup_{n=1}^{\infty}\underline{Sapr}_{\alpha}(X_n)$$

(2)设 $\{X_n\}_{n=1}^{\infty}$ 单调减少。由性质 2(SL5)可知下近似集合序列 $\{\underline{Sapr}_{\alpha}(X_n)\}_{n=1}^{\infty}$ 单调减少。故对于任意的 $n\geqslant 1$,有 $\bigcup\limits_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)=\underline{Sapr}_{\alpha}(X_n)$。由引理 1 可知:

$$\overline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\underline{Sapr}_{\alpha}(X_m)=\bigcap_{n=1}^{\infty}\underline{Sapr}_{\alpha}(X_n)$$

由 $\{\underline{Sapr}_{\alpha}(X_n)\}_{n=1}^{\infty}$ 单调减少且 $\varnothing\subseteq\underline{Sapr}_{\alpha}(X_n)$ 有下界可知,$\lim\limits_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)$存在。于是由式(1)可得:

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha}(X_n)=\overline{\lim_{n\to\infty}}\,\underline{Sapr}_{\alpha}(X_n)=\bigcap_{n=1}^{\infty}\underline{Sapr}_{\alpha}(X_n)$$

定理 2　设(U,R,Pr)为串行概率近似空间,$\beta\in[0,0.5)$。任取 U 上的子集序列 $\{X_n\}_{n=1}^{\infty}$。

(1)若 $\{X_n\}_{n=1}^{\infty}$ 单调增加,则有:

$$\lim_{n\to\infty}\overline{Sapr}_{\beta}(X_n)=\bigcup_{n=1}^{\infty}\overline{Sapr}_{\beta}(X_n)$$

(2)若 $\{X_n\}_{n=1}^{\infty}$ 单调减少,则有:

$$\lim_{n\to\infty}\overline{Sapr}_{\beta}(X_n)=\bigcap_{n=1}^{\infty}\overline{Sapr}_{\beta}(X_n)$$

证明:根据性质 2 和引理 2,类似定理 1 的证明可证结论成立。

性质 2 的(SL5)和(SU5)说明当目标概念集合序列 $\{X_n\}_{n=1}^{\infty}$ 单调增加(或减少)时,串行概率下近似集合序列 $\{\underline{Sapr}_{\alpha}(X_n)\}_{n=1}^{\infty}$ 和串行概率上近似集合序列 $\{\overline{Sapr}_{\beta}(X_n)\}_{n=1}^{\infty}$ 也是单调增加(或减少)的。定理 1 和定理 2 说明,当目标概念集合序列单调时,串行概率下、上近似集是收敛的,且当目标概念集合序列单调增加(或减少)时,对应的串行概率下近似集合序列 $\{\underline{Sapr}_{\alpha}(X_n)\}_{n=1}^{\infty}$ 和串行概率上近似集合序列

$\{\overline{Sapr}_{\beta}(X_n)\}_{n=1}^{\infty}$分别收敛到串行概率下近似集合序列和串行概率上近似集合序列的并(或交)。

定理3 设$(U,R,\Pr)$为串行概率近似空间，$\{\alpha_n\}_{n=1}^{\infty}$是$(0.5,1]$上的实数数列。则对于任意的$X\subseteq U$，

(1)若$\{\alpha_n\}_{n=1}^{\infty}$单调增加，则有：

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha_n}(X)=\bigcap_{n=1}^{\infty}\underline{Sapr}_{\alpha_n}(X)$$

(2)若$\{\alpha_n\}_{n=1}^{\infty}$单调减少，则有：

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha_n}(X)=\bigcup_{n=1}^{\infty}\underline{Sapr}_{\alpha_n}(X)$$

证明：(1)设$\{\alpha_n\}_{n=1}^{\infty}$单调增加。由性质2(SL6)可知串行概率下近似集合序列$\{\underline{Sapr}_{\alpha_n}(X)\}_{n=1}^{\infty}$是单调减少的，且$\emptyset\subseteq\underline{Sapr}_{\alpha_n}(X)$。即$\{\underline{Sapr}_{\alpha_n}(X)\}_{n=1}^{\infty}$单调减少且有下界，故集合序列$\lim\limits_{n\to\infty}\underline{Sapr}_{\alpha_n}(X)$存在，且

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha_n}(X)=\overline{\lim_{n\to\infty}}\underline{Sapr}_{\alpha_n}(X)=\underline{\lim_{n\to\infty}}\underline{Sapr}_{\alpha_n}(X)$$

由串行概率下近似序列$\{\underline{Sapr}_{\alpha_n}(X)\}_{n=1}^{\infty}$单调减少可得$\bigcup\limits_{m=n}^{\infty}\underline{Sapr}_{\alpha_m}(X)=\underline{Sapr}_{\alpha_n}(X)$。故由集合序列上下极限性质可得：

$$\overline{\lim_{n\to\infty}}\underline{Sapr}_{\alpha_n}(X)=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\underline{Sapr}_{\alpha_m}(X)=\bigcap_{n=1}^{\infty}\underline{Sapr}_{\alpha_n}(X)$$

于是

$$\lim_{n\to\infty}\underline{Sapr}_{\alpha_n}(X)=\overline{\lim_{n\to\infty}}\underline{Sapr}_{\alpha_n}(X)=\bigcap_{n=1}^{\infty}\underline{Sapr}_{\alpha_n}(X)$$

(2)类似(1)的证明可证结论成立。

定理4 设$(U,R,\Pr)$为串行概率近似空间，$\{\beta_n\}_{n=1}^{\infty}$是$[0,0.5)$上的一个实数数列。则对于任意的$X\subseteq U$，

(1)若$\{\beta_n\}_{n=1}^{\infty}$单调增加，则有：

$$\lim_{n\to\infty}\overline{Sapr}_{\beta_n}(X)=\bigcap_{n=1}^{\infty}\overline{Sapr}_{\beta_n}(X)$$

(2)若$\{\beta_n\}_{n=1}^{\infty}$单调减少，则有：

$$\lim_{n\to\infty}\overline{Sapr}_{\beta_n}(X)=\bigcup_{n=1}^{\infty}\overline{Sapr}_{\beta_n}(X)$$

证明：(1)若$\{\beta_n\}_{n=1}^{\infty}$单调增加，由性质2(SU6)可知串行概率上近似集合序列$\{\overline{Sapr}_{\beta_n}(X)\}_{n=1}^{\infty}$是单调减少的。由串行概率上近似集合序列$\{\overline{Sapr}_{\beta_n}(X)\}_{n=1}^{\infty}$单调减少且有下界可知$\lim\limits_{n\to\infty}\overline{Sapr}_{\beta_n}(X)$存在。于是：

$$\lim_{n\to\infty}\overline{Sapr}_{\beta_n}(X)=\overline{\lim_{n\to\infty}}\overline{Sapr}_{\beta_n}(X)=\underline{\lim_{n\to\infty}}\overline{Sapr}_{\beta_n}(X)$$

由串行概率上近似集合序列$\{\overline{Sapr}_{\beta_n}(X)\}_{n=1}^{\infty}$单调减少可得$\bigcup\limits_{m=n}^{\infty}\overline{Sapr}_{\beta_m}(X)=\overline{Sapr}_{\beta_n}(X)$。于是：

$$\overline{\lim_{n\to\infty}}\overline{Sapr}_{\beta_n}(X)=\bigcap_{n=1}^{\infty}\bigcup_{m=n}^{\infty}\overline{Sapr}_{\beta_m}(X)=\bigcap_{n=1}^{\infty}\overline{Sapr}_{\beta_n}(X)$$

因此：

$$\lim_{n\to\infty}\overline{Sapr}_{\beta_n}(X)=\overline{\lim_{n\to\infty}}\overline{Sapr}_{\beta_n}(X)=\bigcap_{n=1}^{\infty}\overline{Sapr}_{\beta_n}(X)$$

(2)类似(1)的证明可证结论成立。

性质2的(SL6)和(SU6)说明，随着阈值α,β的单调增加(或减少)，相应的串行概率粗糙上、下近似集是单调减少(或增加)的。定理3和定理4进一步说明，当阈值α,β单调增加(或减少)时，相应的串行概率粗糙下、上近似集是收敛的。且当阈值数列$\{\alpha_n\}_{n=1}^{\infty}$增加(或减少)时，串行概率粗糙下近似集合序列收敛到串行概率粗糙下近似集合序列的交(或并)；当阈值数列$\{\beta_n\}_{n=1}^{\infty}$增加(或减少)时，串行概率粗糙上近似集合序列收敛到串行概率粗糙上近似集合序列的交(或并)。

由串行概率正域和串行概率负域的定义及定理1—定理4可知，当目标概念X单调变化或刻画目标概念与后继领域之间包含程度的阈值α,β单调时，相应的串行概率正域和串行概率负域的集合序列也是收敛的。

推论1 设$(U,R,\Pr)$为串行概率近似空间。对于U的子集序列$\{X_n\}_{n=1}^{\infty}$，$(0.5,1]$上的数列为$\{\alpha_n\}_{n=1}^{\infty}$，$[0,0.5)$上的数列为$\{\beta_n\}_{n=1}^{\infty}$，$n$为自然数。

(1)若$\{X_n\}_{n=1}^{\infty}$单调增加，则有：

$$\lim_{n\to\infty}SPOS_{\alpha}(X_n)=\bigcup_{n=1}^{\infty}SPOS_{\alpha}(X_n)$$

$$\lim_{n\to\infty}SNEG_{\beta}(X_n)=\bigcap_{n=1}^{\infty}SNEG_{\beta}(X_n)$$

(2)若$\{X_n\}_{n=1}^{\infty}$单调减少，则有：

$$\lim_{n\to\infty}SPOS_{\alpha}(X_n)=\bigcap_{n=1}^{\infty}SPOS_{\alpha}(X_n)$$

$$\lim_{n\to\infty}SNEG_{\beta}(X_n)=\bigcup_{n=1}^{\infty}SNEG_{\beta}(X_n)$$

(3)若$\{\alpha_n\}_{n=1}^{\infty}$单调增加，则有：

$$\lim_{n\to\infty}SPOS_{\alpha_n}(X)=\bigcap_{n=1}^{\infty}SPOS_{\alpha_n}(X)$$

若$\{\alpha_n\}_{n=1}^{\infty}$单调减少，则有：

$$\lim_{n\to\infty}SPOS_{\alpha_n}(X)=\bigcup_{n=1}^{\infty}SPOS_{\alpha_n}(X)$$

(4)若$\{\beta_n\}_{n=1}^{\infty}$单调增加，则有：

$$\lim_{n\to\infty}SNEG_{\beta_n}(X)=\bigcup_{n=1}^{\infty}SNEG_{\beta_n}(X)$$

若$\{\beta_n\}_{n=1}^{\infty}$单调减少，则有：

$$\lim_{n\to\infty}SNEG_{\beta_n}(X)=\bigcap_{n=1}^{\infty}SNEG_{\beta_n}(X)$$

当目标概念增加或减少时，串行概率边界域$SBND_{(\alpha,\beta)}(X)$未必单调增加或减少；而当阈值α,β同时增加或减少时，串行概率边界域$SBND_{(\alpha,\beta)}(X)$也未必单调变化。

尽管随着α,β的单调减少或增加，串行概率粗糙上、下近似逐渐增加或减少，但一般情况下，$\lim\limits_{\alpha\downarrow r}\underline{Sapr}_{\alpha}(X)=\underline{Sapr}_{r}(X)$，$\lim\limits_{\beta\uparrow r}\overline{Sapr}_{\beta}(X)=\overline{Sapr}_{r}(X)$，不一定成立。

下面给出当阈值α,β变化时，相应的串行概率粗糙下、上近似的极限。

定理5 设$(U,R,\Pr)$为串行概率近似空间，当$0<r<1$时，对于任意的$X\subseteq U$，有：

(1)$\lim\limits_{\alpha\uparrow r}\underline{Sapr}_{\alpha}(X)=\bigcap\limits_{\alpha<r}\underline{Sapr}_{\alpha}(X)=\underline{Sapr}_{r}(X)$

(2)$\lim\limits_{\alpha\downarrow r}\underline{Sapr}_{\alpha}(X)=\bigcup\limits_{\alpha>r}\underline{Sapr}_{\alpha}(X)=\overline{Sapr}_{r}(X)$

证明：(1)对于任意的$\alpha\in(0,1)$，当$\alpha<r$时，由性质2(SL6)可知$\underline{Sapr}_{\alpha}(X)\supseteq\underline{Sapr}_{r}(X)$。由定理3可得：

$$\lim_{\alpha\uparrow r}\underline{Sapr}_{\alpha}(X)=\bigcap_{\alpha<r}\underline{Sapr}_{\alpha}(X)\supseteq\underline{Sapr}_{r}(X)$$

若存在$x_0\in\bigcap\limits_{\alpha<r}\underline{Sapr}_{\alpha}(X)-\underline{Sapr}_{r}(X)$，则对于任意的$\alpha<r$，有$x_0\in\underline{Sapr}_{\alpha}(X)$，但是$x_0\notin\underline{Sapr}_{r}(X)$。即对于任意的$\alpha<r$，有$\Pr(X|R_s(x_0))\geqslant\alpha$，而$\Pr(X|R_s(x_0))<r$，矛盾。这是因为对于任意的$\alpha<r$，$\Pr(X|R_s(x_0))\geqslant\alpha$，由$\alpha$的任意性可得$\Pr(X|R_s(x_0))\geqslant r$，所以：

$$\bigcap_{\alpha<r}\underline{Sapr}_{\alpha}(X)=\underline{Sapr}_{r}(X)$$

因此，结论(1)成立。

(2)当 $\alpha>r$ 时,由性质 2 的(SL6)可知:

$\underline{Sapr}_{\alpha}(X)\subseteq\overline{Sapr}_r(X)$

故串行概率下近似集合序列$\{\underline{Sapr}_{\alpha}(X)\}$随着 α 的减少而增加,因此:

$$\lim_{\alpha\downarrow r}\underline{Sapr}_{\alpha}(X)=\bigcup_{\alpha>r}\underline{Sapr}_{\alpha}(X)\subseteq\overline{Sapr}_r(X)$$

若存在 $x_0\in\overline{Sapr}_r(X)-\bigcup_{\alpha>r}\underline{Sapr}_{\alpha}(X)$,则有 $\Pr(X|R_s(x_0))\geqslant r$,且对于任意的 $\alpha>r$ 有 $x_0\notin\underline{Sapr}_{\alpha}(X)$。故:

$\forall\alpha>r,\Pr(X|R_s(x_0))<\alpha$

可得 $\Pr(X|R_s(x_0))\leqslant r$,这与 $\Pr(X|R_s(x_0))>r$ 矛盾,所以

$$\bigcup_{\alpha>r}\underline{Sapr}_{\alpha}(X)=\overline{Sapr}_r(X)$$

因此,结论(2)成立。

定理 6　设$(U,R,\Pr)$为串行概率近似空间。对于任意的 $X\subseteq U$,当 $0<r<1$ 时,有:

(1)$\lim_{\beta\uparrow r}\overline{Sapr}_{\beta}(X)=\bigcap_{\beta<r}\overline{Sapr}_{\beta}(X)=\underline{Sapr}_r(X)$

(2)$\lim_{\beta\downarrow r}\overline{Sapr}_{\beta}(X)=\bigcup_{\beta>r}\overline{Sapr}_{\beta}(X)=\overline{Sapr}_r(X)$

证明:(1)当 $\beta<r$ 时,由性质 2(SU6)可知:

$\overline{Sapr}_{\beta}(X)\supseteq\underline{Sapr}_r(X)$

于是串行概率上近似集合序列$\{\overline{Sapr}_{\beta_n}(X)\}_{n=1}^{\infty}$关于 β 单调递减,所以:

$$\lim_{\beta\uparrow r}\overline{Sapr}_{\beta}(X)=\bigcap_{\beta<r}\overline{Sapr}_{\beta}(X)\supseteq\underline{Sapr}_r(X)$$

若存在 $y\in\bigcap_{\beta<r}\overline{Sapr}_{\beta}(X)-\underline{Sapr}_r(X)$,则对于任意的 $\beta<r$,$y\in\overline{Sapr}_{\beta}(X)$,但 $y\notin\overline{Sapr}_r(X)$。于是,由 $\Pr(X|R_s(y))>\beta(\forall\beta<r)$中 β 的任意性可知,$\Pr(X|R_s(y))\geqslant r$,这与 $\Pr(X|R_s(y))<r$矛盾。故$\bigcap_{\beta<r}\overline{Sapr}_{\beta}(X)=\underline{Sapr}_r(X)$,即结论(1)成立。

(2)对于任意的 $\beta>r$,由性质 2(SU6)可知:

$\overline{Sapr}_{\beta}(X)\subseteq\overline{Sapr}_r(X)$

于是,串行概率上近似集合序列$\{\overline{Sapr}_{\beta}(X)\}$关于 β 单调递减。由此可得:

$$\lim_{\beta\downarrow r}\overline{Sapr}_{\beta}(X)=\bigcup_{\beta>r}\overline{Sapr}_{\beta}(X)\subseteq\overline{Sapr}_r(X)$$

若存在 $y\in\overline{Sapr}_r(X)-\bigcup_{\beta>r}\overline{Sapr}_{\beta}(X)$,则 $y\in\overline{Sapr}_r(X)$,但对于任意的 $\beta>r$,$y\notin\overline{Sapr}_{\beta}(X)$。于是 $\Pr(X|R_s(y))>r$,由于对于任意的 $\beta>r$ 都有 $\Pr(X|R_s(y))\leqslant\beta$,于是 $\Pr(X|R_s(y))\leqslant r$,这与 $\Pr(X|R_s(y))>r$ 矛盾。故$\bigcup_{\beta>r}\overline{Sapr}_{\beta}(X)=\overline{Sapr}_r(X)$,即结论(2)成立。

结束语　概率粗糙集是基于等价关系的条件概率提出来的。但在实际问题中,知识库存在多种不确定性因素或者信息缺失等,使得讨论的关系往往达不到较高的标准。且在概率粗糙集中,条件概率要求作为条件的事件的概率必须大于0,因此讨论串行二元关系下的概率粗糙集是有意义的。本文主要研究了串行二元关系下的概率粗糙集模型,讨论了当目标概念 X 发生变化时,串行概率下、上近似的变化趋势,以及串行概率正域、负域及边界域随着目标概念变化的单调性;进一步讨论了当描述目标概念 X 和对象领域的包含程度的阈值 α,β 发生变化时,串行概率粗糙下、上近似的单调性。

参 考 文 献

[1] PAWLAK Z. Rough Sets [J]. International Journal of Computer and Information Science, 1982, 11(5): 341-356.

[2] PAWLAK Z. Rough Sets: Theoretical Aspects of Reasoning About Data [M]. Boston: Kluwer Academic Publishers, 1991.

[3] Chan C C. A rough set approach to attribute generalization in data mining [J]. Information Sciences, 1998, 107(1-4): 169-176.

[4] HU X. Knowledge discovery in database: an attribute-oriented rough set approach [D]. Canada: University of Regina, 1995.

[5] SKOWRON A, STEPANIUK J. Tolerance Approximation Spaces [J]. Fundamenta Information, 1996, 27(2/3): 245-253.

[6] YAO Y Y. Two Views of the Theory of Rough Sets in Finite Universes[J]. International Journal of Approximate Reasoning, 1996, 15(4): 291-318.

[7] YAO Y Y. Constructive and Algebraic Methods of the Theory of Rough Set[J]. Information Sciences, 1998, 109(1-4): 21-47.

[8] YAO Y Y. Relational Interpretations of Neighborhood Operators and Rough Set Approximation Operators[J]. Information Sciences, 1998, 111(1-4): 239-259.

[9] GRECO S, MATARAZZO B, SLOWINSKI R. Rough Approximation of a Preference Relation by Dominance Relations[J]. European Journal of Operational Research, 1999, 117(1): 63-83.

[10] ZHU W. Generalized Rough Sets based on Relations [J]. Information Sciences, 2007, 177(22): 4997-5011.

[11] PEI Z, PEI D W, ZHENG L. Topology vs Generalized Rough Sets[J]. International Journal of Approximate Reasoning, 2011, 52(2): 231-239.

[12] PAWLAK Z, WONG S K M, ZIARKO W. Rough Sets: Probabilistic Versus Deterministic Approach[J]. International Journal of Man-Machine Studies, 1988, 29(1): 81-95.

[13] YAO Y Y, WONG S K M. A Decision Theoretic Framework for Approximating Concepts[J]. International Journal of Man-Machine Studies, 1992, 37(6): 793-809.

[14] ZIARKO W. Variable Precision Rough Set Model[J]. Journal of Computer and System Science, 1993, 46(1): 39-59.

[15] YAO Y Y. Probabilistic Rough Set Approximations[J]. International Journal of Approximate Reasoning, 2008, 49(2): 255-271.

[16] 张文修,吴伟志,梁吉业,等. 粗糙集理论与方法 [M]. 北京:科学出版社,2001.

[17] 张文修,梁怡,吴伟志. 信息系统与知识发现 [M]. 北京:科学出版社,2003.

[18] 于洪. 三支决策:复杂问题求解方法与实践[M]. 北京:科学出版社,2015.

[19] YAO Y Y, LIN T Y. Generalization of rough sets using modal logic[J]. Intelligent Automation and Soft Computing, 1996, 2(2): 103-120.

[20] 朱成熹. 近世实分析基础[M]. 天津:南开大学出版社,1993.

[21] 匡继昌. 实分析与泛函分析[M]. 北京:高等教育出版社,2002.

[22] 薛昌兴. 实变函数与泛函分析[M]. 北京:高等教育出版社,1997.

[23] GONG Z T, SUN B Z, SHAO Y B, et al. The Variable Precision Rough Set Model based on the General Relationship [J]. Journal of Lanzhou University, 2005, 4(6): 110-114. (in Chinese)
巩曾泰,孙秉珍,邵亚斌,等. 一般关系下的变精度粗糙集模型[J]. 兰州大学学报(自科版),2005,4(6):110-114.