基于Local约简的序贯三支分类器

鞠恒荣[1] 李华雄[1,2] 周献中[1,2] 黄 兵[3] 杨习贝[4]

(南京大学工程管理学院 南京 210093)[1] (南京大学智能装备新技术研究中心 南京 210093)[2]

(南京审计大学审计科学研究院 南京 211815)[3] (江苏科技大学计算机科学与工程学院 镇江 212003)[4]

摘 要 序贯三支决策是三支决策理论近年发展起来的一种新型决策方法。传统的序贯三支决策方法鲜有针对序贯信息粒的构建和其在分类学习中的应用的研究。针对这两个问题,研究了Local约简与Global约简之间的内在序贯性,并以此构建了具有约简特性的序贯信息粒。在此基础上设计了一种序贯三支分类器。实验结果表明,该序贯三支分类器不仅能很好地在合适信息粒上进行分类,而且较传统的分类算法提高了数据集的分类精度。

关键词 分类器,Local约简,序贯,粗糙集,三支决策

中图法分类号 TP18 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.09.006

Sequential Three-way Classifier with Local Reduction

JU Heng-rong[1] LI Hua-xiong[1,2] ZHOU Xian-zhong[1,2] HUANG Bing[3] YANG Xi-bei[4]

(School of Management and Engineering, Nanjing University, Nanjing 210093, China)[1]

(Research Center for Novel Technology of Intelligent Equipments, Nanjing University, Nanjing 210093, China)[2]

(Institute of Auditing Science, Nanjing Audit University, Nanjing 211815, China)[3]

(School of Computer Science and Engineering, Jiangsu University of Science and Technology, Zhenjiang 212003, China)[4]

Abstract Sequential three-way decision is a novel decision approach of three-way decision theory in recent years. However, in classical sequential three-way decision research, little attention was paid to two important issues, one is the construction of sequential information granule, and the other is the application in classification learning. To address such issues, the intrinsic sequential properties of local and global reductions were studied firstly in this paper. Based on such properties, the sequential information granule was constructed with reduct's property. Furthermore, a sequential three-way classifier was proposed and designed. The experimental results show that, the proposed classifier is not only good at making classification at an appropriate information granule, but it can also improves the classification accuracy when compared with several classical classifiers.

Keywords Classifier, Local reduction, Sequential, Rough set, Three-way decision

1 引言

三支决策理论是近年发展起来的一种新的用于处理不精确、不完整信息的决策分析方法。作为传统二支决策理论的一种重要推广,三支决策理论考虑了决策过程中存在的不确定性因素,将延迟决策或不承诺决策作为当信息不足以决定接受或拒绝时的第三种决策行为。目前,三支决策理论已被广泛应用于医疗诊断[1]、推荐系统[2]、论文评审[3]等多个学科和领域中。

三支决策理论最初是由加拿大华人学者Yao Yiyu教授提出[4-5]的,它的基本思想起源于粗糙集和决策粗糙集理论的研究,其主要目的是赋予粗糙集正域、边界域和负域一定的语义解释。Yao将Bayes风险决策方法引入粗糙集理论模型中,通过分析各种决策的风险代价,找出最小风险代价的决策,以此作为把对象划分到正域、负域和边界区域的依据,从而形成了接受决策、拒绝决策和延迟决策的三支决策语义[6]。决策粗糙集理论及三支决策理论的最新研究进展可参见文献[7-12]。最初对粗糙集和三支决策的研究主要集中于静态的决策信息系统。基于某一确定的决策信息系统,决策者一次性做出接受决策、拒绝决策和延迟决策。然而,在现实决策分

到稿日期:2016-07-16 返修日期:2016-09-17 本文受国家自然科学基金(71671086,61473157,61572242,61503160,71171107,71201076),江苏省普通高校学术学位研究生科研创新计划项目(KYLX16_0021)资助。

鞠恒荣(1989—),男,博士生,主要研究方向为粗糙集、粒计算、代价敏感,E-mail:justjuhengrong@126.com;**李华雄**(1977—),男,博士,讲师,主要研究方向为粗糙集、数据挖掘、机器学习,E-mail:huaxiongli@nju.edu.cn;**周献中**(1962—),男,博士,教授,博士生导师,主要研究方向为粗糙集、智能信息处理、系统工程理论及应用,E-mail:zhouxz@nju.edu.cn;**黄 兵**(1972—),男,博士,教授,主要研究方向为粗糙集、直觉模糊集理论方法,E-mail:hbhuangbing@126.com;**杨习贝**(1980—),男,博士后,副教授,主要研究方向为粗糙集、粒计算、知识发现,E-mail:zhenjiangyangxibei@163.com。

析中，决策者初始获得的信息往往是不充分的，获取新的有效信息需要一个过程，因此人们的决策也应该随着信息的更新和补充逐步给出。基于此考虑，在原有研究成果的基础上，Yao 和 Deng 提出了序贯三支决策方法[13-14]，给出了一种序贯三支决策算法框架。在此思想的引导下，Li 等人针对代价敏感序贯三支决策进行了进一步的研究[15-17]。

在现有的序贯三支决策研究成果中，专门针对序贯信息粒构建的研究相对较少，研究者大多根据属性重要度将重要度最大的属性作为起始序列，而将所有属性集作为终止序列。Yao 在序贯三支决策方法提出的过程中指出了粒计算的一条基本法则，即：

"…examine the problem at a finer granulation level with more detailed information when there is a need or benefit for doing so."

该基本法则指出，只有在有决策需要或者决策有益的情况下，决策者才应该在更精细的粒度下进行决策。例如，在图像识别中，如果决策者的决策任务只是鉴定出图像中是否有人像，那么如果能在模糊的图像中完成决策，则没必要在高清图像中进行。虽然高清图像能够提供更清晰的信息，但是对于任务驱动的决策者而言，为完成这一简单任务而耗资获取高清图像则易造成资源的浪费。因此，如何构建、选择序贯信息粒是三支决策研究的基础。

属性约简是粗糙集理论研究的一个重要内容[18-20]。传统意义上的属性约简要求在保证所有正域不发生变化的前提下删除一些冗余属性，从而简化决策信息系统[21]。本文将保证所有正域不发生变化的属性约简统称为 Global 约简。然而在实际应用中，一些特殊的决策类往往得到决策者的青睐，决策者也许只关心众多正域中的某个正域不发生变化即能满足决策需求。基于此，Chen 等人提出了 Local 约简的概念[21]。研究表明，Local 约简和 Global 约简并不是相互排斥的互斥关系，它们是相互紧密联系的。简言之，Local 约简和 Global 约简之间存在一种内在的序贯关系，这为如何构建序贯信息粒提供了一种解决方案。

与此同时，如何将序贯三支决策思想应用于分类问题也是本文思考的一个方面。在分类问题研究中，对于某一个测试样本，分类器往往只能基于某一固定的属性集立即做出属于该类和不属于该类的判断，这在信息不充分的情况下极易造成错分的情况，进而造成分类器泛化性能不足。依据三支决策思想，若在信息不充分的情况下无法判别类别标记，则可将该样本下放到更细的一个信息层进行判别，直至在最终信息层上做出二分类判别。

本文第 2 节简要介绍 Pawlak 粗糙集和三支决策；第 3 节介绍 Local 约简的定义，并基于此构建一个序贯三支分类器；第 4 节利用 5 组 UCI 上的数据对所提算法进行了对比分析；最后总结全文。

2　Pawlak 粗糙集与三支决策

一般而言，决策信息系统可被定义为一个二元组 $S=\langle U, AT\cup D\rangle$，其中 U 表示所有对象的集合，称为论域；AT 表示所有条件属性的集合；D 为决策属性集合。

对于 $\forall a\in AT\cup D$，定义映射：$U\to V_a$，V_a 表示属性 a 的值域，即 $a(x)\in V_a(\forall x\in U)$。

在决策信息系统 S 中，根据属性集合 AT，可得到一个不可分辨关系，形如：

$$IND(AT)=\{(x,y)\in U^2:\forall a\in AT, a(x)=a(y)\} \quad (1)$$

显然，不可分辨关系满足自反性、对称性和传递性，因而是一个等价关系。

定义 1[18]　令 S 为一决策信息系统，对于 $\forall X\subseteq U$，X 的下近似集合 $\underline{AT}(X)$ 与上近似集合 $\overline{AT}(X)$ 分别定义为：

$$\underline{AT}(X)=\{x\in U:[x]_{AT}\subseteq X\} \quad (2)$$

$$\overline{AT}(X)=\{x\in U:[x]_{AT}\cap X\neq\varnothing\} \quad (3)$$

其中，$[x]_{AT}=\{y\in U:(x,y)\in IND(AT)\}$ 表示 U 中所有与 x 具有不可分辨关系 $IND(AT)$ 的对象的集合，即 x 的等价类。

根据子集 X 的下、上近似集，可将论域 U 划分为 3 个互不相交的区域：正域 $POS(X)$、边界域 $BND(X)$ 和负域 $NEG(X)$，将其分别定义如下：

$$POS_{AT}(X)=\underline{AT}(X)=\{x\in U:[x]_{AT}\subseteq X\} \quad (4)$$

$$BND_{AT}(X)=\overline{AT}(X)-\underline{AT}(X)$$
$$=\{x\in U:[x]_{AT}\cap X\neq\varnothing\wedge[x]_{AT}\not\subset X\} \quad (5)$$

$$NEG_{AT}(X)=U-\overline{AT}(X)$$
$$=\{x\in U:[x]_{AT}\cap X=\varnothing\} \quad (6)$$

Pawlak 和 Skowron 建议采用粗糙隶属度函数来重新定义上、下近似集。粗糙隶属度 μ_{AT} 可定义为：

$$\mu_{AT}=P(X|[x]_{AT})=\frac{|X\cap[x]_{AT}|}{|[x]_{AT}|} \quad (7)$$

其中，$|\cdot|$ 表示一个集合的基数；$P(X|[x]_{AT})$ 表示分类的条件概率。对应粗糙隶属度，Pawlak 3 个成对分离的区域定义等价于：

$$POS_{AT}(X)=\underline{AT}(X)=\{x\in U:P(X|[x]_{AT})=1\} \quad (8)$$

$$BND_{AT}(X)=\overline{AT}(X)-\underline{AT}(X)$$
$$=\{x\in U:0<P(X|[x]_{AT})<1\} \quad (9)$$

$$NEG_{AT}(X)=U-\overline{AT}(X)=\{x\in U:P(X|[x]_{AT})=0\} \quad (10)$$

基于以上 3 个区域，Yao 引入了三支决策的概念，即正规则、边界规则和负规则。正规则、负规则和边界规则分别对应于接受决策、拒绝决策和延迟决策。具体表述如下：

(P)规则：$\forall x\in U$，假如 $P(X|[x]_{AT})=1$，则选择 $x\in POS_{AT}(X)$；

(B)规则：$\forall x\in U$，假如 $0<P(X|[x]_{AT})<1$，则选择 $x\in BND_{AT}(X)$；

(N)规则：$\forall x\in U$，假如 $P(X|[x]_{AT})=0$，则选择 $x\in NEG_{AT}(X)$。

在很长一段时间里，对粗糙集和三支决策的研究主要集中于静态的决策信息系统。然而，现实决策分析中，决策者初始获得的信息往往是不充分的，获取新的有效信息需要一个过程，而人们的决策也是随着信息的更新和补充逐步给出的。基于此考虑，Yao 和 Deng 提出了序贯三支决策方法。

性质 1[18]　令 S 为一决策信息系统，$\forall A_1, A_2\subseteq AT$，$A_1\subset A_2$，对于任意 $X\subseteq U$，可得到：

(1)$[x]_{A_2}\subseteq[x]_{A_1}$;

(2)$POS_{A_1}(X)\subseteq POS_{A_2}(X)$;

(3)$BND_{A_1}(X)\supseteq BND_{A_2}(X)$;

(4)$NEG_{A_1}(X)\subseteq NEG_{A_2}(X)$。

证明:根据粗糙下、上近似集的定义,性质 1 易证。

由性质 1 可知,更大的属性子集能够在论域 U 上得到更细的划分,更大的属性子集能够提升正域和负域的区域,压缩由边界域带来的不确定性。对于任意的对象 $x\in U$,其可以分别由属性子集 A_1 和 A_2 进行描述,本文将 x 在不同属性子集上的描述简记为 $Des_{A_1}(x)$ 和 $Des_{A_2}(x)$。

现考虑一系列属性子集 $\{A_1,A_2,\cdots,A_k\}$,其满足如下条件:

$$A_1\subset A_2\subset\cdots\subset A_k\subseteq AT \tag{11}$$

根据等价类的定义,易得:

$$[x]_{A_k}\subseteq\cdots\subseteq[x]_{A_2}\subseteq[x]_{A_1} \tag{12}$$

$$POS_{A_1}(X)\subseteq POS_{A_2}(X)\subseteq\cdots\subseteq POS_{A_k}(X) \tag{13}$$

$$BND_{A_1}(X)\supseteq BND_{A_2}(X)\supseteq\cdots\supseteq BND_{A_k}(X) \tag{14}$$

$$NEG_{A_1}(X)\subseteq NEG_{A_2}(X)\subseteq\cdots\subseteq NEG_{A_k}(X) \tag{15}$$

由此便可得到同一对象 x 在不同属性层次上的描述:

$$Des_{A_1}(x),Des_{A_2}(x),\cdots,Des_{A_k}(x) \tag{16}$$

3 Local 约简与序贯三支分类

3.1 Local 约简及其定义

在一个模式识别过程中,无论是计算机还是人,都要首先找出一些具有代表性的特征(属性),然后根据这些特征去识别,可见特征选择是识别分类和分类器设计的前期准备工作。在粗糙集理论中,这一过程被称为属性约简。作为一种独特的模式识别方法,属性约简是在保持原系统某些度量不变的情况下,通过删除冗余属性来得到简化的数据。保持正域不发生变化是经常采用的一种度量准则,传统意义上的正域是所有决策类下近似集的并集,这样的一种约简方法也可视为 Global 约简。其描述见定义 2。

定义 2[10] 令 S 为一决策信息系统,$U/IND(D)=\{X_1,X_2,\cdots,X_n\}$ 是由决策属性 D 诱导的划分,对于任意的 $A\subseteq AT$,A 被称为 S 的 Global 约简当且仅当 A 满足以下条件:

1)对于任意的 $X_j\in U/IND(D)$,$POS_A(X_j)=POS_{AT}(X_j)$;

2)对于任意的真子集 $B\subset A$,$POS_B(X_j)\neq POS_A(X_j)$,$X_j\in U/IND(D)$。

然而在实际应用中,一些特殊的决策类往往得到研究者的关注。例如,在临床检查过程中,耳科医生只会检查病人耳部是否患有疾病。由此可见,与这一特殊决策类紧密关联的条件属性集就显得尤为重要,因为基于这些信息,决策者更能针对这一决策类进行判断。基于这一考虑,Chen 等人提出了 Local 约简方法[21],在 Local 约简中,决策者只关注决策类中的某一类进行相应的属性约简,其描述如定义 3 所示。

定义 3[21] 令 S 为一决策信息系统,$U/IND(D)=\{X_1,X_2,\cdots,X_n\}$ 是由决策属性 D 诱导的划分,对于任意的 $A\subseteq AT$,A 被称为依赖于决策类 X_j 的 Local 约简当且仅当 A 满足以下条件:

1)$POS_A(X_j)=POS_{AT}(X_j)$;

2)对于任意的真子集 $B\subset A$,$POS_B(X_j)\neq POS_A(X_j)$。

与定义 2 中的 Global 约简定义不同,定义 3 中的 Local 约简只要求所得到的约简保证某个决策类 X_j 的正域不发生变化即可。Local 约简较 Global 约简拥有极大的优势。理论上,Global 约简中的属性(包括核属性)都可以根据相关性能划分到一些特定的决策类的约简中;实践中,决策者所关注的特定类的关键属性能够被快速地提取出来,同时也能在一定程度上减少所需要的属性个数。

性质 2 令 S 为一决策信息系统,假设 $U/IND(D)=\{X_1,X_2\}$,即 $n=2$,对于任意的 $A\subseteq AT$,可得到 $Core_A(U/IND(D))=Core_A(X_1)\cup Core_A(X_2)$。

证明:由于 $X_1\cap X_2=\varnothing$,根据粗糙集下近似集的定义可得到 $POS_A(X_1)\cap POS_A(X_2)=\varnothing$,由此可得到 $POS_A(U/IND(D))=POS_A(X_1)\cup POS_A(X_2)$。对于任意的 $a\in Core_A(U/IND(D))$,当且仅当 1)$POS_{A-\{a\}}(X_1)<POS_A(X_1)$ 或 2)$POS_{A-\{a\}}(X_2)<POS_A(X_2)$ 或 $POS_{A-\{a\}}(X_1)<POS_A(X_1)\wedge POS_{A-\{a\}}(X_2)<POS_A(X_2)$。由此可得 1)$a\in Core_A(X_1)$ 或 2)$a\in Core_A(X_2)$ 或 3)$a\in Core_A(X_1)\wedge a(Core_A(X_2)$,因此 $a\in Core_A(X_1)\cup Core_A(X_2)$ 显然可得。类似地对于任意的 $a\in Core_A(X_1)\cup Core_A(X_2)$,可得 $a\in Core_A(U/IND(D))$。综上所述,$Core_A(U/IND(D))=Core_A(X_1)\cup Core_A(X_2)$。

对于多类决策($n>2$)问题,性质 2 同样成立。性质 2 表明,相对于某一个特定的决策类而言,Local 核中的每个属性都是不可或缺的。为了简化问题,本文仅考虑二分类问题。

性质 3 令 S 为一决策信息系统,$U/IND(D)=\{X_1,X_2\}$,假设 R_{L_1} 和 R_{L_2} 分别是基于 X_1 和 X_2 得到的 Local 约简,R_G 为决策信息系统的一个 Global 约简。根据 Pawlak 粗糙理论的单调性性质,可得:

$$R_{L_1}\subseteq R_G;R_{L_2}\subseteq R_G \tag{17}$$

证明:对于任意的 $a\in R_{L_1}$,根据 Local 约简的定义可知 $POS_{R_{L_1}-\{a\}}(X_1)<POS_{R_{L_1}}(X_1)$,因为 R_{L_1} 为决策信息系统基于决策类 X_1 的 Local 约简,所以 $POS_{R_{L_1}}(X_1)=POS_{AT}(X_1)$ 且 R_{L_1} 中每个元素(属性)都是相互独立的。由此可得到,$POS_{AT-\{a\}}(X_1)<POS_{AT}(X_1)$。根据 Global 约简的定义,可断定属性 a 在 Global 约简中也是不能缺少的,因此 $a\in R_G$。所以 $R_{L_1}\subseteq R_G$ 可证。同理可证 $R_{L_2}\subseteq R_G$。

根据定义 3 和性质 2 可给出 Local 约简中属性的重要度定义。

令 S 为一决策信息系统,$\forall A\subseteq AT$,$\forall a\subseteq A$,a 相对于决策类 X_j 的重要度定义为:

$$Sig_{in}(a,A)=\frac{|POS_A(X_j)-POS_{A-\{a\}}(X_j)|}{|U|} \tag{18}$$

由上式可以看出,$Sig_{in}(a,A)$ 反映了将 a 从当前条件属性集 A 中删除后正域的变化程度。根据性质 1 可知 $|POS_A(X_j)-POS_{A-\{a\}}(X_j)|\geqslant 0$,所以 $Sig_{in}(a,A)\geqslant 0$。相应地,也可定义:

$$Sig_{out}(a,A)=\frac{|POS_{A\cup\{a\}}(X_j)-POS_A(X_j)|}{|U|} \tag{19}$$

其中,$a\in AT-A$,$Sig_{out}(a,A)$ 用以度量向属性集 A 增加属性

a 后正域的变化程度。同理，根据性质 1 可得 $Sig_{out}(a,A)\geqslant 0$。根据上述属性的重要度，可设计启发式属性约简算法，如算法 1 所示。

算法 1　基于启发式的 Local 属性约简算法

输入：决策信息系统 S=〈U，AT∪D〉，决策类 X_j

输出：约简 R_{L_j}

Step1　计算 $POS_{AT}(X_j)$。

Step2　red←∅；

Step3　∀a⊆AT，计算属性 a 的重要度 $Sig_{in}(a,A)$。

Step4　若 a_i 满足 $Sig_{in}(a_i,A)=\max\{(a\in AT:Sig_{in}(a,A)\}$，则 red←$a_i$，计算 $POS_{red}(X_j)$。

Step5　若 $POS_{red}(X_j)\neq POS_{AT}(X_j)$，重复以下循环，否则转 Step6。

Step5.1　∀a∈AT−red，计算 $Sig_{out}(a,red)$；

Step5.2　若 $Sig_{out}(a_i,red)=\max\{Sig_{out}(a,red):a\in AT-red\}$，则 red=red∪{$a_i$}；

Step5.3　计算 $POS_{red}(X_j)$；

Step6　∀a∈red，若 $POS_{red-\{a\}}(X_j)=POS_{red}(X_j)$，则 red=red−a。

Step7　R_{L_j}←red。

Step8　输出 R_{L_j}。

由于本文只考虑二分类问题，因此算法 1 中 X_j 可具体为 X_1 或 X_2。由此可得到基于 X_1 和 X_2 的 Local 约简 R_{L1} 和 R_{L2}。由性质 3 可知，决策信息系统的 Local 约简是 Global 约简的子集。基于此性质，可根据 Local 约简得到 Global 约简。即若 $POS_{R_{L_1}}(U/IND(D))=POS_{AT}(U/IND(D))$，则 $R_G=R_{L_1}$；若 $POS_{R_{L_1}}(U/IND(D))<POS_{AT}(U/IND(D))$，则根据属性重要度向 R_{L_1} 中添加属性直至相等，由此可得到 Global 约简[21]。Global 约简的求解过程如算法 2 所示。

算法 2　基于 Local 约简的 Global 约简求解算法

输入：决策系统 S=〈U，AT∪D〉，Local 约简 R_L

输出：Global 约简 R_G

Step1　R_G←R_L。

Step2　计算 $POS_{R_G}(U/IND(D))$。

Step3　若 $POS_{R_G}(U/IND(D))\neq POS_{AT}(U/IND(D))$，重复以下循环，否则转 Step4。

Step3.1　∀a∈AT−R_G，计算 $Sig_{out}(a,R_G)$。

Step3.2　若 $Sig_{out}(a_i,R_G)=\max\{Sig_{out}(a,R_G):a\in AT-R_G\}$，则 $R_G=R_G\cup\{a_i\}$。

Step3.3　计算 $POS_{R_G}(U/IND(D))$。

Step4　∀a∈R_G，若 $POS_{R_G-\{a\}}(U/IND(D))=POS_{R_G}(U/IND(D))$，则 $R_G=R_G-a$。

Step5　输出 R_G。

与传统算法[20]求解 Global 约简不同，基于 Local 约简的 Global 约简求解算法充分利用了 Local 和 Global 之间的内在序贯性。传统方法在求解过程中，约简集合从空集（或全集）出发，再依据属性重要度向集合中添加（或删除）属性。算法 2 中 Global 约简集合则从 Local 集出发进行判断，这一机制必将大大提高算法的效率。

3.2　序贯三支分类器

在序贯三支决策的研究过程中，如何界定属性集序列和如何应用序贯三支决策的思想是研究者值得考虑的两大问题。一方面，在已有的研究成果中，大多以第一个属性为起始序列，以所有的属性集为终止序列。另一方面，在分类问题的研究中，对于某一个测试样本，分类器往往只能基于某一固定的属性集做出属于该类和不属于该类的判断，这在信息不充分的情况下极易造成错分的情况，进而造成分类器性能不足。基于以上讨论，考虑到 Local 约简和 Global 约简内在的序贯性，本文以 Local 属性约简形成的属性集为起始序列、以 Global 属性约简形式的属性集为终止序列设计一个序贯三支分类器，其算法流程如算法 3 所示。

算法 3　基于 Local 约简的序贯三支分类器（S3CLR）

输入：决策信息系统 S=〈U，AT∪D〉，测试样本 s，Local 约简 R_{L_j}，Global 约简 R_G

输出：s 的类别

Step1　计算 $M=R_G-R_{L_j}$，得到 $M=\{a_1,a_2,\cdots,a_l\}$。

Step2　由最近邻规则在 $S'=\langle U,R_{L_j}\cup D\rangle$ 中找到与 s 最近邻的样本 x。

Step3　根据 x 的情况对 s 进行标记：

若 $x\in POS_{R_{L_j}}(X_j)$，则将 s 标记为 d_j；

若 $x\in NEG_{R_{L_j}}(X_j)$，则将 s 标记为 $\neg d_j$；

若 $x\in BND_{R_{L_j}}(X_j)$，则转 Step4。

Step4　若 M≠∅，i 从 1 取到 l，进行以下循环，否则转 Step5。

Step4.1　$R_{L_j}\leftarrow\{R_{L_j},a_i\}$；

Step4.2　在 $S'=\langle U,R_{L_j}\cup D\rangle$ 中找到与 s 最近邻的样本 x；

Step4.3　根据 x 的情况对 s 进行标记：

若 $x\in POS_{R_{L_j}}(X_j)$，则将 s 标记为 d_j，跳出循环；

若 $x\in NEG_{R_{L_j}}(X_j)$，则将 s 标记为 $\neg d_j$，跳出循环；

若 $x\in BND_{R_{L_j}}(X_j)$，i=i+1。

Step5　若 s 未标记类别，则进行如下操作：

若 $P(X_j|[x]_{R_{L_j}})>\gamma$，则将 s 标记为 d_j，否则将其标记为 $\neg d_j$。

Step6　输出 s 的类别标记。

与算法 1 类似，由于本文只考虑二分类问题，因此算法 3 中 X_j 可具体为 X_1 或 X_2。当 $j=1$ 时，d_j 即为 d_1，$\neg d_j$ 实为 d_2；当 $j=2$ 时，d_j 即为 d_2，$\neg d_j$ 为 d_1。在该算法中，测试样本 s 先基于 Local 约简属性集进行判断，若在该属性层无法给出 s 的类别标记，再进入更细的属性层进行判别。由于未纳入 Global 约简属性集的属性为冗余属性，因此这些属性构成的属性层可不予以考虑，这在一定程度上压缩了算法的时间复杂度和空间复杂度。

4　实验分析

本节将通过实验对比分析本文所提算法的性能。本文选择了 UCI 上的 5 组数据进行实验，这 5 组数据的描述如表 1 所列。实验运行环境为 Windows 7 & Matlab R2012b，运行实验的处理器为英特尔第三代酷睿 i5-3470，4.00GHz CPU，内存为 8GB。

表 1　数据描述

序号	数据集	对象个数	属性个数	决策分布（X_1：X_2）
1	Adult	1605	14	(1210:395)
2	Dermatology	366	34	(294:72)
3	Ionosphere	351	34	(225:126)
4	Wdbc	569	30	(212:357)
5	Zoo	101	16	(58:43)

对于二分类问题，可将样例根据其真实类别与分类器预测类别的组合划分为真正例(true positive)、假正例(false positive)、真反例(true negative)和假反例(false negative)4 种情形。令 TP,FP,TN,FN 分别表示其对应的样例数，则分类结果的混淆矩阵如表 2 所列。

表 2 分类结果的混淆矩阵

真实情况	预测结果	
	正例	反例
正例	TP	FN
反例	FP	TN

根据混淆矩阵，可将查准率 P、查全率 R 和精度(Acc)分别定义为：

$$P=\frac{TP}{TP+FP} \tag{20}$$

$$R=\frac{TP}{TP+FN} \tag{21}$$

$$Acc=\frac{TP+TN}{TP+FN+FP+TN} \tag{22}$$

表 3－表 6 分别列出了属性个数、查准率、查全率和分类精度的对比结果。由于本文采用了十倍交叉验证法，因此实验结果均以平均值的形式给出。

表 3 列出了 5 组数据集在不同决策类下得到的 Local 约简、Global 约简的属性个数以及本文所提动态分类算法所需要的属性个数，及其与初始属性个数的比较。

表 3 属性个数的比较

数据集	原始	X_1			X_2		
		Local	Global	Dynamic	Local	Global	Dynamic
Adult	14	10.1	10.1	10.1	10	10.2	10.005
Dermatology	34	4.7	4.7	4.7	4	4	4
Ionosphere	34	19.6	19.8	19.6629	21.2	26.9	22.4540
Wdbc	30	16.6	20.7	17.1611	17.1	19.8	17.4763
Zoo	16	11.2	11.6	11.2	5.8	7.5	5.85

由表 3 可得到如下结论：

(1)与初始数据集相比，Local 约简和 Global 约简的属性个数以及文中所提动态分类算法所需要的属性个数远少于初始的属性个数，表明原始数据中存在大量的冗余属性。利用属性约简可将这些冗余属性删除，以压缩数据的存储空间，从而使决策者得到有效的分析数据。

(2)从决策类角度进行分析可以发现，同一数据在不同决策类下得到的 Local 约简和 Global 约简并不相等，从而使得基于不同类别的动态分类算法需要的属性个数也不相同。

(3)从算法角度进行分析可以发现，Global 约简的属性个数多于或等于 Local 约简的属性个数。而动态分类算法所需的属性个数介于两者之间或者等于 Local 约简的属性个数，由此可见，多数测试样本能够基于 Local 约简进行相应的分类，只有少数测试样本需要补充信息后再进行分类。

表 4 和表 5 分别列出了 5 组数据集在不同属性集和不同分类算法下的查准率和查全率结果。为了体现本文所提算法的有效性，实验中选取了 KNN 分类算法和贝叶斯分类算法，在表中将其简记为 KNN 和 BAY。

表 4 查准率 P 的对比结果

数据集	原始		决策类	Local		Global		Dynamic
	KNN	BAY		KNN	BAY	KNN	BAY	S3CLR
Adult	0.6900	0.5921	X_1	0.6656	0.6758	0.6656	0.6758	**0.6902**
			X_2	0.6651	0.6899	0.6634	0.6934	**0.7016**
Dermatology	0.9507	0.8503	X_1	0.9234	**0.9607**	0.9234	**0.9607**	0.9234
			X_2	0.9648	0.9541	0.9648	0.9541	**0.9731**
Ionosphere	0.9007	0.8660	X_1	**0.9266**	0.9097	**0.9266**	0.9097	0.9198
			X_2	**0.9246**	0.9159	0.9137	0.9095	**0.9246**
Wdbc	0.7492	0.7246	X_1	0.7284	0.8379	0.7258	0.8209	**0.8807**
			X_2	0.7288	0.8727	0.7368	0.8583	**0.8798**
Zoo	0.9474	0.9144	X_1	0.9107	0.9144	0.9124	0.9144	**0.9169**
			X_2	0.8988	**0.9199**	0.8988	0.9112	0.9079

表 5 查全率 R 的对比结果

数据集	原始		决策类	Local		Global		Dynamic
	KNN	BAY		KNN	BAY	KNN	BAY	S3CLR
Adult	0.6969	0.5208	X_1	0.6703	0.5956	0.6703	0.5956	**0.6718**
			X_2	**0.6712**	0.5926	0.6697	0.6054	0.6688
Dermatology	0.9220	0.9419	X_1	0.8108	0.8908	0.8108	0.8908	**0.9108**
			X_2	0.9489	0.8640	0.9489	0.8640	**0.9726**
Ionosphere	0.8896	0.8129	X_1	0.9197	0.8755	0.9140	0.8698	**0.9209**
			X_2	**0.9112**	0.8904	0.9068	0.8731	**0.9112**
Wdbc	0.7966	0.6956	X_1	0.7749	0.8664	0.7696	**0.8571**	0.8237
			X_2	0.7741	0.8573	0.7836	**0.8695**	0.8111
Zoo	0.9245	0.9023	X_1	0.9290	0.9023	0.9234	0.9023	**0.9481**
			X_2	0.8970	0.9273	0.8970	0.9440	**0.9495**

综合表 4 和表 5 的实验结果可知，本文所提算法 S3CLR 获取最高的查准率时对应的查全率并不一定也是最高的；

类似地，S3CLR 获取最高的查全率时对应的查准率也不一定是最高的。这是由于查准率和查全率是一对矛盾的度量。一般而言，查准率高时，查全率往往偏低；而查全率高时，查准率往往偏低。只有在少数分类任务中，查全率和查准率都很高。因此，总体而言，本文所提算法的查准率和查全率比 KNN 和贝叶斯分类的更好。

表 6 列出了 5 组数据集在不同属性集和不同分类算法下的分类精度对比结果。相比较于 KNN 分类器和贝叶斯分类器基于 Local 约简集和 Global 约简集的分类精度，本文所提动态分类算法 S3CLR 能够获得更高的分类精度。

表 6　分类精度 Acc 的对比结果

数据集	原始		决策类	Local		Global		Dynamic
	KNN	BAY		KNN	BAY	KNN	BAY	S3CLR
Adult	0.7707±0.031	0.7495±0.034	X_1	0.7502±0.047	0.7645±0.036	0.7502±0.047	0.7645±0.036	**0.7758**±0.041
			X_2	0.7495±0.041	0.7682±0.037	0.7483±0.039	0.7732±0.031	**0.7858**±0.040
Dermatology	0.9589±0.030	0.9041±0.087	X_1	0.9179±0.035	**0.9454**±0.056	0.9179±0.034	**0.9454**±0.056	0.9179±0.034
			X_2	0.9738±0.034	0.9402±0.058	0.9738±0.034	0.9402±0.058	**0.9811**±0.033
Ionosphere	0.8890±0.091	0.8149±0.073	X_1	**0.9202**±0.107	0.8773±0.103	0.9145±0.107	0.8716±0.103	**0.9202**±0.103
			X_2	**0.9119**±0.114	0.8917±0.161	0.9061±0.114	0.8744±0.161	**0.9119**±0.115
Wdbc	0.7627±0.051	0.6063±0.021	X_1	0.7435±0.073	0.8101±0.075	0.7363±0.073	0.8008±0.075	**0.8121**±0.066
			X_2	0.7381±0.073	0.7753±0.075	0.7487±0.073	0.7700±0.075	**0.8138**±0.066
Zoo	0.9400±0.053	0.9200±0.082	X_1	0.9400±0.071	0.9200±0.108	0.9300±0.071	0.9200±0.097	**0.9500**±0.082
			X_2	0.9200±0.069	0.9300±0.082	0.9200±0.069	**0.9400**±0.071	**0.9400**±0.070

综合表 3—表 6 的实验结果可知，本文所提动态分类算法基于动态属性集合，能够获取最高的分类精度，同时保证了较好的查准率和查全率。

结束语　如何界定序贯三支决策中序贯信息粒和如何将序贯三支思想应用于分类学习是本文的出发点和落脚点。本文通过分析 Local 约简和 Global 约简之间存在的序贯性，以 Local 约简为序贯信息粒的起始序列、Global 约简为终止序列构建了一种基于 Local 约简的动态序贯三支分类器。实验结果表明，本文提出的分类算法能够获得高分类精度，同时保证了较好的查准率和查全率。

本文只是针对序贯三支分类器的初步探索，下一步将对以下几个问题进行深入的探讨：

(1)本文基于 Local 约简构建序贯信息粒，构建其他更普遍形式的信息粒值得进一步考虑；

(2)多分类问题是常见的分类方法，如何设计多分类序贯三支分类器是值得探讨的科学问题。

参 考 文 献

[1] LURIE J D, SOX H C. Principles of medical decision making [J]. Spine, 1999, 24(5): 493-498.

[2] ZHANG H R, MIN F. Three-way recommender systems based on random forests [J]. Knowledge Based Systems, 2016, 91: 275-286.

[3] WELLER A C. Editorial peer review: Its strengths and weakness[M]. Medford, NJ: Information Today, 2001.

[4] YAO Y Y, WONG S K M. A decision theoretic framework for approximating concepts [J]. International Journal of Man-machine Studies, 1992, 37: 793-809.

[5] YAO Y Y. The superiority of three-way decisions in probabilistic rough set models[J]. Information Sciences, 2011, 181: 1080-1096.

[6] JU H R, YANG X B, YU H L, et al. Research on attribute reduction criteria in decision-theoretic rough set [J]. Journal of Nanjing Normal University (Natural Science Edition), 2015, 38 (1): 41-47. (in Chinese)
鞠恒荣，杨习贝，于化龙，等. 决策粗糙集的属性约简准则研究 [J]. 南京师范大学学报(自然科学版)，2015，38(1)：41-47.

[7] JIA X Y, LIAO W H, TANG Z M, et al. Minimum cost attribute reduction in decision-theoretic rough set models[J]. Information Sciences, 2013, 219: 151-167.

[8] JIA X Y, TANG Z M, LIAO W H, et al. On an optimization representation of decision-theoretic rough set model [J]. International Journal of Approximate Reasoning, 2014, 55: 156-166.

[9] DOU H L, YANG X B, SONG X N, et al. Decision-theoretic rough set: A multicost strategy[J]. Knowledge Based Systems, 2016, 91: 71-83.

[10] JU H R, YANG X B, YU H L, et al. Cost-sensitive rough set approach[J]. Information Sciences, 2016, 355-356: 282-298.

[11] LIANG D C, LIU D. Deriving three-way decisions from intuitionistic fuzzy decision-theoretic rough sets [J]. Information Sciences, 2015, 300: 28-48.

[12] LIU D, LIANG D C, WANG C C. A novel three-way decision model based on incomplete information system[J]. Knowledge Based Systems, 2016, 91: 32-45.

[13] YAO Y Y. Granular computing and sequential three-way decisions[M]// Lingras P, Wolski M, Cornelis C, et al., eds. Rough Sets and Knowledge Technology. Springer Berlin Heidelberg, 2013: 16-27.

[14] YAO Y Y, DENG X F. Sequential three-way decisions with probabilistic rough sets[C]// Wang Y, Celikyilmaz A, Kinsner W, et al., eds. IEEE International Conference on Cognitive Information & Cognitive Computing. IEEE, 2011: 120-125.

[15] LI H X, ZHOU X Z, HUANG B, et al. Cost-sensitive three-way decision: A sequential strategy[M]// Lingras P, Wolski M, Cornelis C, et al., eds. Rough Sets and Knowledge Technology. Springer Berlin Heidelberg, 2013: 325-337.

[16] LI H X, ZHANG L B, HUANG B, et al. Sequential three-way decision and granulation for cost-sensitive face recognition[J]. Knowledge Based Systems, 2016, 91: 241-251.

[17] ZHANG L B, LI H X, ZHOU X Z, et al. Cost-sensitive sequential three-way decision for face recognition[M]// Kryszkiewicz M, Cornelis C, Ciucci C, et al., eds. Rough Sets and Intelligent Systems Paradigms. Springer International Publishing, 2014: 375-383.

(下转第 57 页)

参 考 文 献

[1] GANTER B, WILLE R. Formal Concept Analysis-Mathematical Foundations[M]. New York: Springer Berlin Heidelberg, 1999.

[2] WILLE R. Restructuring lattice theory: an approach based on hierarchies of concepts[M]// Rival I. Ordered Sets. Dordrecht: Reidel, 1982: 445-470.

[3] CH S. Peirce: Collected Papers[M]. Cambridge: Harvard Univ. Press, 1931-1935.

[4] LEHMANN F, WILLE R. A triadic approach to formal concept analysis[M]// Ellis G, Levinson R, Rich W, et al. Conceptual Structures: Applications, Implementation and Theory (LNCS 954). Heidelberg: Springer, 1995: 32-43.

[5] WILLE R. The basic theorem of triadic concept analysis[J]. Order, 1995, 12(2): 149-158.

[6] BIEDERMANN K. Triadic Galois connections [M]// Denecke K, Lders. General algebra and applications in discrete mathematics. Aachen: Shaker Verlag, 1997: 23-33.

[7] BIEDERMANN K. An equational theory for trilattices [J]. Algebra Universalis, 1999(42): 253-268.

[8] GANTER B, OBIEDKOV S. Implications in triadic formal contexts [M]// Wolff K E, Pfeiffer H D, Delugach H S. Conceptual Structures at Work (LNCS3127). Heidelberg: Springer, 2004: 186-195.

[9] MISSAOUI R, KWUIDA L. Mining triadic association rules from ternary relations [M]// Valtchev P, Jaschke R. International Conference on Formal Concept Analysis (LNCS6628). Heidelberg: Springer, 2011: 204-218.

[10] JASCHKE R, HOTHO A, SCHMITZ C, et al. TRIAS-an algorithm for mining iceberg tri-lattices[C]// Proceeding of the sixth international conference on data mining (ICDM'06). Piscataway, Null, IEEE, 2006: 907-911.

[11] IGNATOV D, KUZNETSOV S, MAGIZOV R, et al. From triconcepts to triclusters [M]// Kuznetsov S O, et al. Rough Sets, Fuzzy Sets, Data Mining and Granular Computing, RSFDGrC 2011(LNCS6743). Heidelberg: Springer, 2011: 257-264.

[12] KAYTOUE M, KUZNETSOV S, MAGIZOV J, et al. Mining Biclusters of similar values with triadic concept analysis [C]// Napoli A, Vychodil V. Proceedings of the 7th International Conference on Concept Lattices and Their Applications (CLA2011). 2011: 175-190.

[13] GNATYSHAK D, IGNATOV D, SEMENOV A, et al. Gaining insight in social networks with biclustering and triclustering [M]// Aseeva N, Babkin E, Kozyrev O. Perspectives in Business Informatics Research (LNBIP128). Heidelberg: Springer, 2012: 162-171.

[14] BELOHLAVEK R, VYCHODIL V. Optimal factorization of three-way binary data [C]// Hu X, Lin T Y, Raghavan V, et al. 2010 IEEE International Conference on Granular Computing. Piscataway, NJ: IEEE, 2010: 61-66.

[15] GLODEANU C. Factorization methods of binary, triadic, real and fuzzy data [J]. Studia Universitatis Babes-Bolyai Series Informatica, 2011, 56(2): 81-86.

[16] BELOHLAVEK R, GLODEANU C, VYCHODIL V. Optimal factorization of three-way binary data using triadic concepts [J]. Order, 2013, 30(2): 437-454.

[17] CYNTHIA G. Tri-ordinal factor analysis [M]// Cellie R P, Distel F, Ganter B. Formal Concept Analysis (LNCS7880). Heidelberg: Springer, 2013: 125-140.

[18] BELOHLAVEK R, OSICKA P. Triadic concept analysis of data with fuzzy attributes [C]// 2010 IEEE International Conference on Granular Computing. Piscataway, NJ: IEEE, 2010: 661-665.

[19] OSICKA P, KONECNY J. General approach to triadic concept analysis [C]// Kryszkiewicz M, Obiedkov S. Proceedings of the 7th International Conference on Concept Lattices and Their Applications (CLA2010). 2010: 116-126.

[20] BELOHLAVEK R, OSICKA P. Triadic concept lattices of data with graded attributes [J]. International Journal of General System, 2012, 41(2): 93-108.

[21] KONECNY J, OSICKA P. Triadic concept lattices in the framework of aggregation structures [J]. Information Science, 2014, 279: 512-527.

[22] GLODEANU C V. Fuzzy-Valued triadic implications [C]// Napoli A, Vychodil V. Proceedings of the 7th International Conference on Concept Lattices and Their Applications (CLA2011). 2011: 159-173.

[23] BELOHLAVEK R, OSICKA P. Triadic fuzzy Galois connections as ordinary connections [J]. Fuzzy Sets and Systems, 2014, 249: 83-99.

[24] OSICKA P. Algorithms for computation of concept trilattices of triadic fuzzy context [M]// Greco S, Meunier B B, Coletti G, et al. Advances in Computational Intelligence (CCIS 299). Heidelberg: Springer, 2012: 221-230.

[25] TRABELSI C, JELASSI N, YAHIA S. Scalable mining of frequent tri-concepts from folksonomies [M]// Tan P N, Chawla S, Ho C K, et al. Advances in Knowledge Discovery and Data Mining (LNCS7302). Heidelberg: Springer, 2012: 231-244.

[26] JELASSI M N, YAHIA S B, NGUIFO E M. A scalable mining of frequent quadratic concepts in d-folksonomies[J]. Computer Science, 2012. arXiv: 1212. 0087v1 [cs. SI].

[27] WEI L, WAN Q, QIAN T, et al. An Overview of Triadic Concept Analysis[J]. Journal of Northwest University (Natural Science Edition), 2014, 44(5): 689-699. (in Chinese)
魏玲，万青，钱婷，等. 三元概念分析综述[J]. 西北大学学报(自然科学版)，2014，44(5)：689-699.

(上接第 39 页)

[18] PAWLAK Z. Rough sets-theoretical aspects of reasoning about data[M]. Dordrecht: Kluwer Academic, 1991.

[19] JU H R, YANG X B, SONG X N, et al. Dynamic updating multigranulation fuzzy rough set: Approximations and reducts[J]. International Journal of Machine Learning and Cybernetics, 2014, 5(6): 981-990.

[20] JU H R, YANG X B, DOU H L, et al. Variable precision multigranulation rough set and attributes reduction[M]// Peters J F, Skowron A, Li T R, et al. , eds. Transactions on Rough Set XVII. Springer Berlin Heidelberg, 2014: 52-68.

[21] CHEN D G, ZHAO S Y. Local reduction of decision system with fuzzy rough sets [J]. Fuzzy Sets and Systems, 2010, 161: 1871-1883.