第 32 卷第 1 期
2015 年 1 月

计算机应用与软件
Computer Applications and Software

Vol. 32 No. 1
Jan. 2015

属性约简中的范式转换算法研究

俞雪平[1,2] 胡云安[3]

[1]（海军航空工程学院研究生管理大队 山东 烟台 264000）
[2]（92914 部队 海南 临高 571800）
[3]（海军航空工程学院控制工程系 山东 烟台 264000）

摘 要 通过研究属性约简中合取范式到析取范式的转换过程,发现减少冗余项和重复计算可以适当提高转换效率。同时考虑到范式的动态变化,设计一种边转换边化简的增量转换算法,可以利用已有结果直接进行计算。对于减量情况,抽象出范式转换的数学模型,给出相应转换的构造形式和分析过程,并提出一种近似减量转换算法,从而实现了不同变化情况下生成析取范式的动态计算。最后通过仿真实验验证了算法的可行性和高效性。

关键词 约简 增量 析取范式 合取范式 范式转换

中图分类号 TP18 **文献标识码** A DOI:10.3969/j.issn.1000-386x.2015.01.068

ON NORMAL FORM CONVERSION ALGORITHM IN ATTRIBUTE REDUCTION

Yu Xueping[1,2] Hu Yun'an[3]

[1] (*Graduate Students' Brigade, Naval Aeronautical Engineering Institute, Yantai 264000, Shandong, China*)
[2] (*92914 Troop, Lin'gao 571800, Hainan, China*)
[3] (*Department of Control Engineering, Naval Aeronautical Engineering Institute, Yantai 264000, Shandong, China*)

Abstract By studying the conversion process from conjunctive normal form to disjunctive normal form in attribute reduction, we find that the conversion efficiency can be properly improved by reducing the redundant formula and the iterative calculations. Taking into account the dynamic change of the normal form, we design an incremental conversion algorithm which converts and simplifies at the same time and can directly calculate by using existing results. For decrement situation, we extract the mathematical model of normal form conversion, give the construction form of corresponding conversion and the analysis process, and also propose an approximate reduction conversion algorithm, therefore realises the dynamic calculation in regard to generating the disjunctive normal form in different variation circumstances. Finally, through simulation experiments we verify the feasibility and efficiency of the algorithm.

Keywords Reduction Increment Disjunctive normal form Conjunctive normal form Normal form conversion

0 引 言

信息系统中的属性之间相互关联,其中某些属性是冗余的。冗余属性的存在,既是对存储和处理资源的浪费,也影响人们作出正确的判断和简洁的决策。利用粗糙集理论[1],这种处理不精确、不完全与不一致数据的新型数学工具,可以在保持系统分类或决策能力不变的前提下进行属性约简,删除其中不必要或不重要的属性,从而使得推导的规则简便明了,可解释性强。

关于属性约简的研究主要包括代数表示与信息表示的方法[2]。在代数表示下,大体上又可分为基于正区域的属性约简算法、基于分辨矩阵及在此基础上改进的属性约简算法等。一般来讲,一个信息系统的属性约简不是唯一的。同一问题在不同的知识表示下,其求解难度是不同的,已经证明了寻找一个决策表中具有最少条件属性的最小约简是 NP-hard 问题。同样是属性组合爆炸的原因,计算决策表所有全部约简的时间复杂度是指数级的。虽然属性完全约简计算代价较高,但它是属性最小约简的参考,并且可以提供更多的选择。由 A. Skowron 提出的基于分辨矩阵的方法[3]通过引入代数知识和矩阵运算,把信息系统决策表的属性约简这一复杂的组合搜索问题转化为简单的基于符号的逻辑推理过程,从而可以得到所有可能的属性约简结果。但在应用基于分辨矩阵的属性完全约简时,一般采用人工转换的方法或基于机器的搜索方法实现分辨函数的合取范式向析取范式的转换,比直接寻找最小约简要繁琐许多,存在时空浪费以及效率瓶颈等问题[4],限制了该方法的应用范围。通过矩阵化简和启发式算法可以省去不必要的重复计算,适当缩减求解时间[5,6]。然而,现有的这些研究几乎都是针对静态数据的,而在实际应用中,信息系统通常都是动态变化的,这将改变原先的分辨矩阵并导致约简结果的变化。增量属性约简可以直接利用先前的约简和当前的变化更新结果,而不用重新约简,减少了计算量,适合解决类似问题[7-9]。它亦是其他约简算法的基础,但相应的减量算法则研究较少。

除了上述基于分辨函数的属性约简外,还有许多关于命题

收稿日期:2013-03-12。**俞雪平**,博士生,主研领域:模式识别与智能系统。**胡云安**,教授。

的可满足性问题都可以归结为范式的转换问题,因此研究范式的高效转换算法具有相当广泛的应用价值[10]。本文抽象出范式转换的数学模型,通过研究发现合取范式与析取范式之间存在的内在规律,提出了相应的范式增量转换算法,并充分运用合取和析取运算的吸收率,减少冗余数据,提高计算效率。算法原理简单,易于实现,仿真实验表明算法是高效可行的。

1 相关背景

1.1 分辨函数与布尔矩阵

知识系统 $S=(U,A,V,f)$,其中 U 为论域,x_i 为论域中的对象,$U=\{x_1,x_2,\cdots,x_n\}$。A 为非空的属性集合,$A=C\cup D$,$C=\{c_1,c_2,\cdots,c_m\}$ 是条件属性集合,$D=\{d\}$ 是决策属性,且 $C\cap D=\phi$。$V=\bigcup_{a\in A}V_a$ 是属性值的集合,V_a 表示属性 $a\in A$ 的值域。$f:U\times A\rightarrow V$ 是一个信息函数,它指定 U 中每个对象 x 的属性值。信息系统的分辨矩阵只考虑条件属性,主要用于区分论域中的不同对象,而决策系统的分辨矩阵还要考虑决策属性,目的是分辨具有不同决策属性值的论域对象。

对于系统 S 的决策分辨矩阵可表示为 $M(S)=[m_{ij}]_{n\times n}$,$m_{ij}$ 是能够区分两个对象 x_i 和 x_j 的所有条件属性的集合,即:

$$m_{ij}=\begin{cases}\{c_k\in C\mid f(x_i,c_k)\neq f(x_j,c_k)\} & f(x_i,d)\neq f(x_j,d)\\ \phi & f(x_i,d)=f(x_j,d)\end{cases}$$

根据分辨矩阵可以得到相应的分辨函数 $f_{DM}=\wedge\{\vee m_{ij}\mid 1\leqslant j<i\leqslant n,m_{ij}\neq\phi\}$,其形式上是一个合取范式,将其转换为析取范式后,其中每一个合取式便是一个属性约简结果。

1.2 合取范式与析取范式

下面引入一个布尔函数,称其为区分函数,用 Δ 表示。对每个属性 $a\in A$,指定一个布尔变量“a”。若 $a(x,y)=\{a_1,a_2,\cdots,a_k\}\neq\varnothing$,则指定一个布尔函数 $a_1\vee a_2\vee\cdots\vee a_k$,用 $\sum a(x,y)$ 来表示;若 $a(x,y)=\varnothing$,则指定布尔常量 1。(布尔)区分函数 Δ 可定义如下:$\Delta=\prod_{(x,y)\in U\times U}\sum a(x,y)$。 现在介绍几个与布尔函数有关的概念[11]。

(1) 布尔表达式 E 是一个范式,如果 E 仅仅由布尔变量和常量通过析取与合取运算来表达。其中,变元 x 或变元的否定 $\neg x$ 称为文字,记为 l。

(2) 布尔表达式 E 是一个合取范式(CNF),如果 E 是由一些析取式组成的合取所构成的范式。其中,一组文字的析取 $l_1\vee\cdots\vee l_m$ 称为子句(clause),记为 C。

(3) 布尔表达式 E 是一个析取范式(DNF),如果 E 是由一些合取式组成的析取所构成的范式。其中,一组文字的合取 $l_1\wedge\cdots\wedge l_m$ 称为子方(cube),记为 D。

(4) 布尔表达式 E 是一个极小析取范式,如果 E 是一个析取范式且包含最小数目的合取式。

区分函数 Δ 有如下性质:函数 Δ 的极小析取范式中的所有合取式是属性集 A 的所有约简。换句话说,约简是满足能区别由整个属性集区别的所有对象的属性极小子集。

1.3 模型与推导

属性约简其实就是分辨函数的逻辑运算过程。可以证明,有限布尔代数上的任何布尔表达式都可以唯一地表示为析取范式或合取范式。任何一个合取范式都可以通过演算转换为一个析取范式。考虑单步增量计算的简化模型为 $A\wedge B=C$,其中 A 和 C 均为析取范式,B 为合取范式的子句。

假设现已知 B 和 C 求 A,则单步减量计算转化为求逆问题。通常,$DNF(C)\rightarrow CNF(C)\rightarrow CNF(A)\rightarrow DNF(A)$ 的求解过程较为复杂,如果能直接从析取范式 C 入手,应该更为方便。这里把 A,B,C 分别看作广义空间中的一个整体,由方程 $A\wedge B=C$ 可知,$C\subseteq A,C\subseteq B$,显然,$A$ 是存在的,C 便是方程的一个解,但 A 不唯一。对于属性约简来说,由于 A 和 C 是简化后的极小析取范式,析取范式的子方之间是不可比的,A 自身便是方程的一个最大解 $A=C\vee A$,即 A 是包含所有约简的极大极小集,它是唯一的。

考虑 A,B,C 的形式和含义以及合取运算性质,由:

$$\begin{aligned}A&=A\wedge(B\vee B')=(A\wedge B)\vee(A\wedge B')=C\vee(A\wedge B')\\&=(C\vee A)\wedge(C\vee B')=A\wedge(C\vee B')\end{aligned}$$

可以利用 B' 和 C 来构造 D,即 $D=C\vee B'$,其中 B' 为 B 的差集,满足 $B'\subseteq A\subseteq B\vee B'$。下面通过集合的包含关系讨论方程的解 A 的情况,并给出 A 的通解形式和构造过程。

定义 如果集合 L 有两个二元运算 $\vee$ 和 $\wedge$,它们满足幂等律、交换律和结合律,并且满足吸收律,那么 L 是一个格。如果除此之外还满足分配律,那么 L 称为分配格。

引理 在任意分配格中,由 $a\vee x=a\vee y$ 和 $a\wedge x=a\wedge y$ 一起可推出 $x=y$。

定理 如果 B' 满足条件 $B'\subseteq A\subseteq B\vee B'$,则 $D=C\vee B'=A$。

证明 由条件可知 $B\vee B'\subseteq A\vee B\subseteq B\vee B'$,则 $A\vee B=B\vee B'$,这样 $D\vee B=(C\vee B')\vee B=B\vee B'=A\vee B$,此外 $B\wedge B'\subseteq A\wedge B=C$,则还有 $D\wedge B=(C\vee B')\wedge B=(C\wedge B)\vee(B\wedge B')=C=A\wedge B$,那么根据引理命题得证。

通过以上分析可以看出,如何构造满足条件的极小 B' 成为解决问题的关键。但条件中的 A 是未知的,只能利用 B 和 C 的信息以及定理证明中的关系 $B\wedge B'\subseteq C$ 进行试探。设条件属性集合 $B_0=\{b_1,\cdots,b_n\}$,$B=\bigvee_{i}^{s}b_i$,记为 $\{b_1,\cdots,b_s\}$,则差集 B' 表示不在 B 中的所有属性集合,即 $B'=B_0-B=\{b_{s+1},\cdots,b_n\}$,而 $C=\bigvee_{i}^{t}(\bigwedge_{j}^{t_i}c_{ij})$。取 $m=card(B')$,在 B' 的幂集中搜索,若存在 $[(\bigwedge_{j}^{k}b'_j)\wedge B]\subseteq C$,其中 $b'_j\in B'$,$k<\max(t_i)$,则 $A=C\vee(\bigwedge^{k}b'_j)$,否则 $A=C$。

2 算法实现

2.1 增量算法

任何合取范式都存在唯一的一个族集与之一一对应。当这个族集的基数较大时,可能存在相关项,有必要首先进行化简以降低计算复杂度。这里先考虑单步增量情况,通过设计合理的范式转换算法来减少算法的运行时间,实现快速的范式转换。

单步增量算法

输入:$A=\{A_1,A_2,\cdots,A_r\}$,$B=\{b_1,\cdots,b_s\}$;

输出:$C=\{C_1,C_2,\cdots,C_t\}$,其中 $r,s,t\in$ ¥。

步骤 1 初始化 $C=\{\}$,$C'=\{\}$。

步骤 2 如果 A 中某个集合与 B 有相同元素,则 A_i 不更新,即若存在 $b_j\in B$,使得 $b_j\in A_i$,则 $C=C\cup A_i$,$C'=C'\cup A_i$,并且将 A_i 从 A 中删除,即 $A=A-A_i$。

步骤 3 对于 A 中每一个集合以及 B 中的每一个元素,若存在 $A_k\in A$ 及 $b_j\in B$,使得 $C'\not\subset\{A_k,b_j\}$,则 $C=C\cup(\bigcup_{k,j}\{A_k,b_j\})$。

步骤 4　返回结果 C,算法结束。

对于子句的数目大于 2 的合取范式可以利用结合律分别对其析取式进行合取运算,通过将输出 C 赋给输入 A,将 B 更新为下一个子句,再调用单步增量算法,如此反复可得最终的析取范式。

增量算法　输入:$A=\{A_1,A_2,\cdots,A_r\}$,$B=\{B_1,\cdots,B_s\}$;输出:$C=\{C_1,C_2,\cdots,C_t\}$,其中 $r,s,t\in$ ¥。

步骤 1　初始化 $C=A$,$j=1$,若 $A=\varnothing$,则 $C=\{\{b_{11}\},\cdots,\{b_{1i}\}\}$,$j=2$。

步骤 2　将 C 和 B_j 作为输入,调用单步增量算法,用输出更新 C。同时,$j=j+1$。

步骤 3　若 $j\neq s$,执行步骤 2,否则,返回结果 C,算法结束。

由于合取与析取是对偶运算,析取范式到合取范式的转换与合取范式到析取范式的转换是相同的操作过程,均可用增量算法求解。对于伪合取范式到析取范式的转换,本文的增量算法可以直接计算,而文献[10]中的算法需要首先将含有命题变元的合取式的子式化为合取范式,多了一个化析取范式为合取范式的过程。对于含有否定变元的合取范式的转换,只需在步骤 3 中稍加改动即可,但应该将文献[10]算法中的操作:若存在 $A_k\in A$ 及 $b_j\in B$,使得$\neg b_j\in A_k$,则集合$\{A_k,b_j\}$化简为$\{A_k-b_j\}$,修改为删除集合$\{A_k,b_j\}$,此时结果可能不再满足覆盖条件。例如,已知 $A=(a\wedge b)\vee(c\wedge d)$,$B=\neg c\vee\neg d$,则 $C=A\wedge B=(a\wedge b\wedge\neg c)\vee(a\wedge b\wedge\neg d)$。

2.2　减量算法

根据前面的讨论,考虑到 B 为子句,B'和 C 均为析取范式,寻找极小 B'需要遍历所有 B 的差集构成的幂集,计算代价过高。为此,从 C 的子方出发以减少最终的计算量,给出相应的减量算法,可以得到 DNF(A)的近似解。

单步减量算法　输入:$C=\{C_1,C_2,\cdots,C_t\}$,$B=\{b_1,\cdots,b_s\}$,条件属性集合 $B_0=\{b_1,\cdots,b_n\}$;输出:$A=\{A_1,A_2,\cdots,A_r\}$,其中 $n,r,s,t\in$ ¥。

步骤 1　初始化结果 $A=C$,差集 $B'=B_0-B=\{b_{s+1},\cdots,b_n\}$,族集 C 中最小集合的势为 $t_c=\min(t_i)$。

步骤 2　对于 B 中的每一个元素,若存在 $C_i\in C$,使得任意两个 $b_j\in C_i$,则将 C_i 从 C 中删除,即 $C=C-C_i$,$t=t-1$;若 $card(C_i)=t_c$,且存在 $b_j\in B$,使得 $b_j\in C_i$,则将 C_i 从 C 中删除。

步骤 3　对于 C 中每一个集合 C_k 以及 B 中的每一个元素,若存在 $b_j\in B$,使得 $b_j\in C_k$,并且对于 $C_h=C_k-b_j$,任意 $b_l\in B-b_j$,存在 $C'\in C$,使得 $C'\subseteq\bigcup_{h,l}\{C_h,b_l\}$,则 $A=(A-C_k)\cup C_h$。

步骤 4　返回结果 A,算法结束。

类似于增量算法,通过递归调用减量算法可以生成原析取范式。需要注意的是,此单步减量算法为近似算法,可以在大多数情况下得到正确结果,但也可能相比准确解遗漏部分子式。例如,已知方程 $A\wedge B=C$,其中析取范式 $C=(a\wedge b)\vee(c\wedge d)\vee(b\wedge e\wedge f)\vee(d\wedge e\wedge f)$,$B=b\vee c$,求析取范式 A。通过遍历搜索可知 $A=(a\wedge b)\vee(a\wedge c)\vee(c\wedge d)\vee(e\wedge f)$,而减量算法得到的解是 $A=(a\wedge b)\vee(c\wedge d)\vee(e\wedge f)$。

2.3　算法运用与复杂度分析

假设分辨矩阵变化前的属性约简结果为 C,分辨矩阵变化后的属性约简结果为 C_t;分辨矩阵变化项原先为 B,分辨矩阵变化项后来为 B_t。由方程 $A\wedge B=C$ 可知,若 $A=C$,则 B 不起作用,那么根据吸收律,比 B 大的项也不起作用。

添加操作　添加项 B_t,即 $C_t=C\wedge B_t$。若 $B_t\supseteq B$,则 $C_t=C$。可直接运用增量算法得到操作后的析取范式 C_t。

删除操作　删除项 B,即 $C_t\wedge B=C$。若 B 为析取范式,可直接推广减量算法得到操作后的析取范式 C_t;若 B 为合取范式,可反复运用减量算法得到操作后的析取范式 C_t。

修改操作　因为 $A\wedge B=C$,通用的方法是先由减量算法得到原析取范式 A,在与 B_t 合取得到修改后的析取范式 $C_t=A\wedge B_t$。若 $B_t\supseteq B$,则 $C_t\supseteq C$,即属性约简结果可能增加包含 B 中新添属性的组合,但若 $A=C$,则 $C_t=C$;若 $B_t\subseteq B$,则 $C_t\subseteq C$,即属性约简结果可能减少包含 B 中删去属性的组合,但包含 B_t 的属性约简结果不变;若 $B_t//B$,即它们之间互不包含,那么属性约简结果可能增加也可能减少,但若 $B_t\cap B\neq\phi$,则属性约简结果必包含 $B_t\cap B$ 项。

计算所有约简的时间复杂度是指数的。首先,需要考虑 C 的所有$|2^C|=2^{|C|}$个子集。其次,对于每个子集计算相对约简的时间复杂性是 $O(|C\|U|^2)$。因此,整个代价是 $O(2^{|C|}|C\|U|^2)$。基于分辨矩阵的约简需要通过对属于不同决策属性类别的论域对象进行两两比较才能建立分辨矩阵,这个过程的代价为$O(|C\|U|^2)$,因此其约简算法的时间复杂度一般为 $O(|C|^2|U|^2)$。令合取范式子句的个数为 p,析取范式子方的个数为 q,属性个数为$|C|$,单步减量算法的代价为 $O(|C|^2q^2)$,单步增量算法的代价为 $O(|C|^2q^2)$,则范式转换算法的时间复杂度为 $O(|C|^{2p}q^2)$,空间复杂度为 $s=O(|C|^pq)$。

3　实例验证

为了验证本文算法的有效性,分别在确定样本和随机数据上比较其与文献[12]算法的优劣,通过以下仿真实验测试算法的性能。实验的硬件平台是:CPU:T5300 1.73GHz,内存:1GB,操作系统:Windows XP。算法实现的软件是 Matlab 7.0。

3.1　确定样本测试

首先验证算法的正确性,选用 UCI 数据库中的数据集作为测试对象。各种范式转换算法用于属性约简的比较结果如表 1 所示。其中,$|U|$表示数据集中对象的数目;$|C|$表示数据集中条件属性的数目;m 表示最小约简包含的属性数目;n 表示所有全部约简集合的势;t 表示算法的运行时间,单位为 s。

表 1　不同范式转换算法在确定样本上的测试结果

样本	$\|U\|$	$\|C\|$	文献[12]算法			本文算法		
			m	n	t	m	n	t
soybean-small	47	35	2	765	249.03	2	765	5.531
zoo	101	17	10	4	0	10	4	0
Iris	150	4	4	4	0.453	4	4	0.453
Liver Disorders	345	6	3	9	1.578	3	9	1.578
Congressional Voting Records	435	16	8	3	2.22	8	3	2.047

由表 1 可知,本文算法与文献[12]算法运行结果相同,都可以得到样本的所有属性约简。由于第一个数据集的所有约简集合数量较大,本文算法运行时间明显少于文献[12]算法,而其他数据集的所有约简集合数量较小,本文算法与文献[12]算法耗时相差无几。这与两者算法的运行过程有关。

3.2　随机数据测试

为了比较算法的性能,利用自定义数据进行测试。随机生

成两组数据,分别以条件属性的个数与合取范式的子句个数为变量,考察各种算法在两个参数各自变化时的表现,测试比较结果如表 2 所示。其中,$|C|$ 表示范式中命题变量的数目;p 表示合取范式所对应集族的势;q 表示析取范式所对应集族的势;t 表示算法的运行时间。

表 2 不同范式转换算法在随机数据上的测试结果

样本	p	$\|C\|$	文献[12]算法		本文算法	
			q	t	q	t
p 变化	10	15	91	0.375	91	0.015
	100	15	443	48.906	443	1.031
	1000	15	928	479.92	928	13.141
$\|C\|$变化	100	10	53	0.266	53	0.031
	100	15	470	56.297	470	1.109
	100	20	2104	2 088.1	2104	39.641

由表 2 可知,本文算法与文献[12]算法在随机样本上的运行结果也相同。由于随机数据集的所有约简集合数量较大,本文算法耗时明显少于文献[12]算法。对于本文算法,其运行时间与所生成析取范式对应集族的势成正比,而这又与被转换合取范式对应集族的势和其中命题变量的数目有关,体现了算法的复杂度。

最后,测试增量算法和减量算法的性能,结果见表 3 和表 4。由表可知,对于样本动态变化的情况,本文的单步增减量算法运行时间较短,具有一定的优越性。增量算法可以胜任不同数目添加操作的情况,而删除操作对应的减量算法是近似计算,存在的误差一般是可接受的。

表 3 增量范式转换算法在随机数据上的测试结果

样本	p	$\|C\|$	常规算法		增量算法	
			q	t	q	t
添加合取范式的子句	100	15	443	1.031	443	0.032
	1000	15	928	13.141	928	0.094
	100	20	2104	39.641	2104	0.922

表 4 减量范式转换算法在随机数据上的测试结果

样本	p	$\|C\|$	常规算法		增量算法	
			q	t	q	t
添加合取范式的子句	100	15	428	2.609	428	0.734
	1000	15	923	22.969	923	3.375
	100	20	2101	264.38	2093	18.672

4 结 语

分辨函数的转换过程就是运用分配律进行表达式展开,利用幂等律和吸收律消除冗余项,简化计算。本文通过研究范式转换过程,提出相应的增量与减量算法,实现范式生成的动态计算,解决了分辨函数变化情况下的属性约简问题。相比文献[12]中先转换再化简的方法,本文的算法同步转换与化简过程,提高了执行效率,实现了增量计算。文献[10]的算法没有预先扫描机制,化简时判断次数较多,也没给出算法的具体仿真验证,对于文献中的伪合取范式要先转换成合取范式形式才能计算,本文的算法通过将伪合取范式作为运行中间结果可以直接处理此类问题;对于包含否定项的范式转换,指出文献中算法的错误,通过给出反例说明其在某些情况下并不满足覆盖条件,但适当修改程序仍能得到正确的计算结果。尽管这样无法改变该类方法的时间复杂度,但结合实际情况,利用现有信息,仍然可以适当改善方法的性能。需要补充的是,范式转换算法还应与其他样本预处理技术相结合才能获得满意的属性约简结果。当数据规模较大时,求取属性完全约简的精确算法已不现实也没有必要,可以采用近似算法求取部分约简。

参 考 文 献

[1] Pawlak Zdzislaw. Rough Sets and Intelligent Data Analysis[J]. Information Sciences,2002,147(1-4): 1 - 12.

[2] Wang G Y,Zhao J,An J J,et al. A comparative study of algebra viewpoint and information viewpoint in attribute reduction[J]. Fundamenta Informaticae,2005,68(6): 289 - 301.

[3] Skowron A,Rauszer C. The Discernbility Matrices and Functions in Information Systems[C]//Slowinski R,ed. Intelligent Decision Support,Handbook of Applications and Advances of the Rough Sets Theory,Dordrecht: Kluwer Academic Publishers,1992: 331 - 362.

[4] 赵荣泳,张浩,李翠玲,等. 粗糙集理论中分辨函数的析取范式生成算法[J]. 计算机工程,2006,32(2): 183 - 185.

[5] Wang Jue,Wang Ju. Reduction algorithm based on discernibility matrix: the ordered attributes method[J]. Journal of Computer Science and Technology,2001,16(6): 489 - 504.

[6] 田卫东,周创德,胡学钢,等. 基于简化分辨矩阵的粗糙集属性约简算法[J]. 计算机科学,2008,35(3): 209 - 212.

[7] Ziarko W,Shan N. Data-based acquisition and incremental modification classfication rules[J]. Computational Intelligence,1995,11(2): 357 - 370.

[8] Hu Feng,Wang Guoyin,Huang Hai,et al. Incremental attribute reduction based on elementary sets[C]//Proceedings of the 10th International Conference on Rough Sets,Fuzzy Sets,Data Mining,and Granular Computing. Regina,Canada,2005: 185 - 193.

[9] 胡峰,代劲,王国胤. 一种决策表增量属性约简算法[J]. 控制与决策,2007,22(3): 268 - 272.

[10] 智慧来,智东杰,刘宗田. 从合取范式到析取范式的转换研究[J]. 计算机工程与应用,2012,48(2): 15 - 17.

[11] 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.

[12] 张德栋,李仁璞,赵永升. 一种高效的分辨函数范式转换算法[J]. 计算机应用研究,2010,27(3): 879 - 882.

(上接第 224 页)

[11] Damian V,Jose C,Victor M,et al. Reconstruction with criterion form labeled markers:New approach based on the morphological watershed[J]. Journal of Electronic Imaging,2010,19 (4):1 - 15.

[12] 李杰,苗长云,武志刚,等. 基于数学形态学的图像边缘检测算法的研究[J]. 计算机科学,2012,39(6A):546 - 548.

[13] Lu Qingwen,Chen Wufan. Unsupervised segmentation of medical image based on difference of mutual information[J]. Science in China Series F-Information Sciences,2006,49(4):484 - 493.

[14] Mathieu Bouchard,Anne-Laure Jousselme,Pierre-Emmanuel Doré. A proof for the positive definiteness of the Jaccard index matrix[J]. International Journal of Approximate Reasoning,2013,54(5):615 - 626.