第 45 卷 第 4 期 西南师范大学学报(自然科学版) 2020 年 4 月
Vol. 45 No. 4 Journal of Southwest China Normal University (Natural Science Edition) Apr. 2020

DOI:10.13718/j.cnki.xsxb.2020.04.012

基于启发式二分策略的属性约简方法①

黄治国[1], 杨清琳[2]

1. 河南工程学院 计算机学院,郑州 451191;2. 广西财经学院 现代教育技术部,南宁 530003

摘要:分辨矩阵为属性约简与求核运算提供了一种规范精确的数学模型,通过分辨矩阵模型可方便地获取决策系统全体属性约简.本文基于分辨矩阵运用启发式信息与二分策略扩展结点,设计了一种构造约简树求取全体属性约简的有效算法,该约简树从根结点到叶结点全体路径构成的析取范式与系统分辨函数等价,其极小析取范式所有析取项即为决策系统全体属性约简.该方法适用于满足任意约简准则的分辨矩阵,能够显著地减少搜索空间和保证全体约简求解的完备性,理论分析与实验结果说明了算法的可行性与有效性.

关 键 词:属性约简;启发式信息;二分策略;分辨矩阵

中图分类号:**TP301.6** **文献标志码**:**A** **文章编号**:1000-5471(2020)04-0059-09

粗糙集理论一直受到众多学者的高度关注,已广泛应用于数据挖掘、知识发现、模式识别等诸多领域[1].属性约简从高维数据中剔除与决策类别无关的属性子集,通过删除冗余属性实现数据降维和压缩数据容量,一直是粗糙集理论研究中的热点内容[2].

在基于粗糙集的知识发现研究中,许多学者对属性约简算法进行了大量的研究.文献[3]分析了相对正域的渐增式特征,改进相对正域的计算方法,每次选择使其变化最大的属性,直至求得一个约简.文献[4]采用基数排序方法计算正域,定义属性重要度计算公式,每次选择最重要的属性,直至生成一个约简.文献[5]将分治法思想融入粗糙集算法中,设计基于分治策略的属性约简算法,求得给定属性序下的唯一约简.文献[6]提出了一种基于矩阵简化的约简构造方法,重复选择属性并执行矩阵简化操作,直至矩阵中仅存在单属性元素,最终所有单属性的集合即为求得的一个约简.文献[7]将不一致决策表转化为一致决策表,给出了5种一致决策表约简的定义,利用一致决策表中属性重要度函数的单调性求得决策表属性约简.以上方法均仅是求得决策系统的一个属性约简,且不能保证其一定为最小属性约简.文献[8]针对多种常见的属性约简准则,分析其本质内涵并将之分层归类,为进一步实现面向不同约简准则的高效属性约简算法奠定理论基础.文献[9]提出了一种基于容差关系的属性约简方法,为不完备决策表求取全体属性约简给出了一种解决方案.文献[10]依据条件等价类定义决策向量,在此基础上设计了满足不同属性约简准则的分辨矩阵构造方法,为高效属性约简算法研究奠定了分辨矩阵构造基础.但这些研究并没有就如何设计高效的约简算法作进一步讨论.文献[11]基于扩展法则变换分辨函数,为决策系统求取全体属性约简提供了一种完备高效的计算方法.

本文深入分析分辨矩阵的一些有用特性,吸收扩展法则思想设计一种操作分辨集搜索全体约简的有效方法,并给出了详细的算法描述和时间复杂度分析,理论分析与实验结果说明了该算法的高效性与完备性.

---

① 收稿日期:2018-04-24
基金项目:河南省高等学校重点科研项目(17A520027);河南工程学院博士基金项目(D2013003).
作者简介:黄治国(1978-),男,博士,副教授,主要从事数据挖掘、智能计算等方面的研究.

1 分辨矩阵与分辨函数

$DS=\langle U, C \cup D, V, f\rangle$ 为决策系统的形式化表示，其中 U 为非空有限论域；C 为条件属性集，D 为决策属性集，$C \cap D=\varnothing$；$V=\bigcup_{a \in (C \cup D)} V_a$ 为属性值域，V_a 为属性 a 的值域；$f: U \times (C \cup D) \rightarrow V$ 为信息函数，$f(x,a)$ 为对象 $x \in U$ 在属性 a 上的取值.

定义 1[6] 给定决策系统 $DS=\langle U, C \cup D, V, f\rangle$，令 $C=\{a_1,\cdots,a_m\}$，$U=\{x_1, \cdots, x_n\}$，分辨矩阵 $\boldsymbol{M}$ 是一个 $|U| \times |U|$ 的矩阵，矩阵中每一元素定义为

$$m_{ij}=\begin{cases}\{a \in C: f(x_i, a) \neq f(x_j, a), 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant n\} & i \neq j \wedge f(x_i, D) \neq f(x_j, D) \\ \varnothing & \text{否则}\end{cases}$$

分辨矩阵满足对称性且对于 $1 \leqslant \forall i \leqslant n$ 均有 $c_{ii}=\varnothing$，因此可仅通过下三角($1 \leqslant j < i \leqslant n$)元素来表示矩阵. 为叙述方便，令 M_{DS} 为决策系统 DS 的分辨矩阵 $\boldsymbol{M}$ 中非空元素构成的集合，即 $M_{DS}=\{m_{ij}: m_{ij} \neq \varnothing\}$.

定义 2[6] 给定决策系统 $DS=\langle U, C \cup D, V, f\rangle$，$M_{DS}=\{m_{ij}\} \neq \varnothing$，其分辨函数定义为

$$F_{DS}(C)=\wedge\{\vee m_{ij}: m_{ij} \in M_{DS} \wedge m_{ij} \neq \varnothing\}$$

其中，$\vee m_{ij}$ 为 m_{ij} 中所有属性的析取式，$\wedge\{\vee m_{ij}\}$ 为所有 $\vee m_{ij}$ 的合取式. 后续描述中在不影响理解时，将 $F_{DS}(C)$ 简写为 F_{DS}.

定义 3[12] 定义 $a \in C$ 关于属性集 $B \subseteq C$ 的布尔映射函数为

$$\rho_B(a)=\begin{cases}0 & a \notin B \\ 1 & a \in B\end{cases}$$

分辨函数关于属性集 $B \subseteq C$ 的布尔映射函数为 $\rho_B(F_{DS}(C))=\wedge\{\vee \rho_B(m_{ij}): m_{ij} \in M_{DS} \wedge m_{ij} \neq \varnothing\}$.

性质 1[12] 给定决策系统 $DS=\langle U, C \cup D, V, f\rangle$，$M_{DS}=\{m_{ij}\} \neq \varnothing$，则 $\rho_C(F_{DS}(C))=1$.

由定义 3 可知对分辨矩阵任意非空元素 m_{ij}，$\vee m_{ij}$ 关于属性集 $B \subseteq C$ 的布尔映射函数为

$$\rho_B(\vee m_{ij})=\begin{cases}0 & m_{ij} \cap B=\varnothing \\ 1 & m_{ij} \cap B \neq \varnothing\end{cases}$$

对于 $\forall m_{ij} \in M_{DS} \wedge m_{ij} \neq \varnothing$ 均有 $m_{ij} \cap C \neq \varnothing$，故对于分辨矩阵中任意非空元素 m_{ij} 恒有 $\rho_C(\vee m_{ij})=1$，因此分辨函数 F_{DS} 关于属性集 C 的布尔映射函数 $\rho_C(F_{DS}(C))=\wedge\{\rho_C(\vee m_{ij}): m_{ij} \in M_{DS} \wedge m_{ij} \neq \varnothing\}=1$.

定义 4[8] 给定决策系统 $DS=\langle U, C \cup D, V, f\rangle$，$R(R \subseteq C)$ 关于 D 的可辨识关系定义为：$DIS(R, D)=\{(x_i, x_j) \in U \times U: f(x_i, R) \neq f(x_j, R) \wedge f(x_i, D) \neq f(x_j, D) \wedge (|f([x_i]_{IND(R)}, D)|=1 \vee |f([x_j]_{IND(R)}, D)|=1)\}$. 其中 $[x_i]_{IND(R)}$ 是不可分辨关系 R 下由对象 x_i 定义的等价类，定义为 $[x_i]_{IND(R)}=\{x: x \in U \wedge f(x, R)=f(x_i, R)\}$.

性质 2 给定决策系统 $DS=\langle U, C \cup D, V, f\rangle$，$R \subseteq C$，则 $DIS(R, D) \subseteq DIS(C, D)$.

证 针对 $\forall(x_i, x_j) \in DIS(R, D)$ 必有 $f(x_i, R) \neq f(x_j, R)$ 且 $f(x_i, D) \neq f(x_j, D)$ 且 $|f([x_i]_{IND(R)}, D)|=1 \vee |f([x_j]_{IND(R)}, D)|=1$. 由 $f(x_i, R) \neq f(x_j, R)$ 可推出 $f(x_j, C) \neq f(x_j, C)$；由 $R \subseteq C$ 可得 $[x_i]_{IND(C)} \subseteq [x_i]_{IND(R)}$，$[x_j]_{IND(C)} \subseteq [x_j]_{IND(R)}$，又 $|f([x_i]_{IND(R)}, D)|=1 \vee |f([x_j]_{IND(R)}, D)|=1$，故有 $|f([x_i]_{IND(C)}, D)|=1 \vee |f([x_j]_{IND(C)}, D)|=1$. 因此针对 $\forall(x_i, x_j) \in DIS(R, D)$ 必有 $f(x_i, C) \neq f(x_j, C)$ 且 $f(x_i, D) \neq f(x_j, D)$ 且 $|f([x_i]_{IND(C)}, D)|=1 \vee |f([x_j]_{IND(C)}, D)|=1$，即 $(x_i, x_j) \in DIS(C, D)$，性质 $DIS(R, D) \subseteq DIS(C, D)$ 得证.

定义 5[8] 对于决策系统 $DS=\langle U, C \cup D, V, f\rangle$，给定目标概念对象集合 $X \subseteq U$，X 关于 $R(R \subseteq C)$ 的近似精度定义为

$$\partial_R(X)=\frac{|\bigcup\{[x]_{IND(R)}: [x]_{IND(R)} \subseteq X\}|}{|\bigcup\{[x]_{IND(R)}: [x]_{IND(R)} \cap X \neq \varnothing\}|}$$

其中，$\bigcup\{[x]_{IND(R)}: [x]_{IND(R)} \subseteq X\}$ 是论域 U 中通过知识 R 判断确定属于集合 X 的元素构成的集合，记为正域 $POS_R(X)$，显然有 $0 \leqslant \partial_R(X) \leqslant 1$. 近似精度反映了通过现有知识向目标概念能够逼近的程度，当

$\partial_R(X)=1$ 时，X 可通过已有知识被精确描述；当 $\partial_R(X)<1$ 时，X 仅可通过已有知识被粗糙定义.

定义 6[8]　给定决策系统 $DS=\langle U, C\cup D, V, f\rangle$，若 $R(R\subseteq C)$ 满足：

1) $\partial_R(D)=\partial_C(D)$；

2) $\forall a\in R$，$\partial_{R-\{a\}}(D)\neq\partial_R(D)$

则称 R 为条件属性集 C 关于决策属性集 D 的相对属性约简. 相对属性约简的本质在于保持决策系统近似质量不发生改变.

定理 1　给定决策系统 $DS=\langle U, C\cup D, V, f\rangle$，$R\subseteq C$，$DIS(R,D)=DIS(C,D)$ 的充要条件是 $\partial_R(D)=\partial_C(D)$.

证明：

1) 证明：若 $DIS(R,D)=DIS(C,D)$，则 $\partial_R(D)=\partial_C(D)$，利用反证法.

假设 $\partial_R(D)\neq\partial_C(D)$，则 $\exists x_i\in U$ 使得$[x_i]_{IND(C)}\subseteq POS_C(D)$ 但$[x_i]_{IND(R)}\nsubseteq POS_C(D)$，则 $\exists x_j\in[x_i]_{IND(R)}$ 使得 $x_j\notin[x_i]_{IND(C)}$ 且 $f(x_i,D)\neq f(x_j,D)$. 由$[x_i]_{IND(C)}\subseteq POS_C(D)$ 可得 $|f([x_i]_{IND(C)},D)|=1$，由 $x_j\notin[x_i]_{IND(C)}$ 可得 $f(x_i,C)\neq f(x_j,C)$，因此有$(f(x_i,C)\neq f(x_j,C))\wedge(f(x_i,D)\neq f(x_j,D))\wedge(|f([x_i]_{IND(C)},D)|=1\vee|f([x_j]_{IND(C)},D)|=1)$ 成立，即$(x_i,x_j)\in DIS(C,D)$. 由 $x_j\in[x_i]_{IND(R)}$ 可得 $f(x_i,R)=f(x_j,R)$，故$(x_i,x_j)\notin DIS(R,D)$.$(x_i,x_j)\in DIS(C,D)$ 但$(x_i,x_j)\notin DIS(R,D)$，这与前提条件 $DIS(R,D)=DIS(C,D)$ 矛盾，因此假设不成立，所以 $\partial_R(D)=\partial_C(D)$.

2) 证明：若 $\partial_R(D)=\partial_C(D)$，则 $DIS(R,D)=DIS(C,D)$，利用反证法.

假设 $DIS(R,D)\neq DIS(C,D)$，则由性质 1 可推出 $\exists x_i,x_j\in U$ 使得$(x_i,x_j)\in DIS(C,D)$ 但$(x_i,x_j)\notin DIS(R,D)$. 由$(x_i,x_j)\in DIS(C,D)$ 可知有 $f(x_i,C)\neq f(x_j,C)$ 且 $f(x_i,D)\neq f(x_j,D)$ 且 $|f([x_i]_{IND(C)},D)|=1\vee|f([x_j]_{IND(C)},D)|=1$ 成立，可推出 x_i，x_j 中至少一个属于 $POS_C(D)$；由$(x_i,x_j)\notin DIS(R,D)$ 可推出有 $f(x_i,R)=f(x_j,R)$ 或者 $f(x_i,D)=f(x_j,D)$ 或者 $|f([x_i]_{IND(R)},D)|>1\wedge|f([x_j]_{IND(R)},D)|>1$ 成立. 因为 $f(x_i,D)\neq f(x_j,D)$，所以只可能有 $f(x_i,R)=f(x_j,R)$ 或者 $|f([x_i]_{IND(R)},D)|>1\wedge|f([x_j]_{IND(R)},D)|>1$ 成立，由 $f(x_i,D)\neq f(x_j,D)$ 与 $f(x_i,R)=f(x_j,R)$ 可推出 x_i，x_j 两对象均不属于 $POS_R(D)$，由 $|f([x_i]_{IND(R)},D)|>1\wedge|f([x_j]_{IND(R)},D)|>1$ 也可推出 x_i，x_j 两对象均不属于 $POS_R(D)$，因此 $POS_R(D)\neq POS_C(D)$，因而有 $\partial_R(D)\neq\partial_C(D)$，与前提 $\partial_R(D)=\partial_C(D)$ 相矛盾，因此假设不成立，所以 $DIS(R,D)=DIS(C,D)$.

根据定理 1，可将相对约简定义转换为对应于可辨识关系的等价描述. 给定决策系统 $DS=\langle U, C\cup D, V, f\rangle$，若 $R(R\subseteq C)$ 为 C 相对于 D 的属性约简当且仅当 R 满足：

1) $DIS(R,D)=DIS(C,D)$；

2) $\forall a\in R$，$DIS(R-\{a\},D)\neq DIS(C,D)$.

因此，可依据系统的可辨识关系是否变化，作为 $R(R\subseteq C)$ 是否为约简的判定标准.

定理 2　给定决策系统 $DS=\langle U, C\cup D, V, f\rangle$，$M_{DS}=\{m_{ij}\}\neq\emptyset$，$R\subseteq C$，则 $DIS(R,D)=DIS(C,D)$ 当且仅当 $\rho_R(F_{DS})=\rho_C(F_{DS})$.

证明：

1) 证明：若 $DIS(R,D)=DIS(C,D)$ 则 $\rho_R(F_{DS})=\rho_C(F_{DS})$.

对 $\forall m_{ij}\in M_{DS}\wedge m_{ij}\neq\emptyset$，必存在 U 中对象 $x'_i\in[x_i]_{IND(C)}$，$x'_j\in[x_j]_{IND(C)}$，使得$(x'_i,x'_j)\in DIS(C,D)$，此时 $m_{ij}=\{a: a\in C\wedge f(x_i,a)\neq f(x_j,a)\}=\{a: a\in C\wedge f(x'_i,a)\neq f(x'_j,a)\}$. 又已知 $DIS(R,D)=DIS(C,D)$，故当$(x'_i,x'_j)\in DIS(R,D)$，则必 $\exists a\in R$，$f(x'_j,a)\neq f(x'_j,a)$ 成立，因而有 $R\cap m_{ij}\neq\emptyset$ 故 $\rho_R(\vee m_{ij})=1$. 针对 $\forall m_{ij}\in M_{DS}\wedge m_{ij}\neq\emptyset$ 均有 $\rho_R(\vee m_{ij})=1$，因此 $\rho_R(F_{DS})=\wedge\{\rho_R(\vee m_{ij}): m_{ij}\in M_{DS}\wedge m_{ij}\neq\emptyset\}=1$. 又由性质 2 得 $\rho_C(F_{DS}(a_1,\cdots a_m))=1$，因此 $\rho_R(F_{DS})=\rho_C(F_{DS})$.

2) 证明：若 $\rho_R(F_{DS})=\rho_C(F_{DS})$ 则 $DIS(R,D)=DIS(C,D)$.

先证 $DIS(C, D) \subseteq DIS(R, D)$. 针对 $\forall (x_i, x_j) \in DIS(C, D)$，必存在 $m = \{a: a \in C \wedge f(x_i, a) \neq f(x_j, a)\} \neq \varnothing$ 使得 $m \in M_{DS}$. 由条件 $\rho_R(F_{DS}) = \rho_C(F_{DS})$ 与性质1可得 $\rho_R(F_{DS}) = 1$，因此对 M_{DS} 中任意非空元素 m 恒有 $\rho_R(\vee m) = 1$，即 $R \cap m \neq \varnothing$，因此有 $f(x_i, R) \neq f(x_j, R)$. 又由 $(x_i, x_j) \in DIS(C, D)$ 可得 $f(x_i, D) \neq f(x_j, D)$ 且 $|f([x_i]_{IND(C)}, D)| = 1 \vee |f([x_j]_{IND(C)}, D)| = 1$，故针对 $\forall (x_i, x_j) \in DIS(C, D)$ 均有 $f(x_i, R) \neq f(x_j, R)$ 且 $|f([x_i]_{IND(C)}, D)| = 1 \vee |f([x_j]_{IND(C)}, D)| = 1$.

假设 $|f([x_i]_{IND(R)}, D)| > 1 \wedge |f([x_j]_{IND(R)}, D)| > 1$ 成立，由 $|f(;x_i]_{IND(C)}, D)| = 1 \vee |f([x_j]_{IND(C)}, D)| = 1$ 可知 x_i，x_j 中至少一个属于 $POS_C(D)$，可令 $x_i \in POS_C(D)$(同理可令 $x_j \in POS_C(D)$，证明过程与 $x_i \in POS_C(D)$ 类似)，则 $\exists x_k \in U$ 使得 $x_k \in [x_i]_{IND(R)} \wedge x_k \notin [x_i]_{IND(C)}$ 且 $f(x_i, D) \neq f(x_k, D)$，那么必 $\exists m' \in M_{DS} \wedge m' \neq \varnothing$ 其行对应 x_i 列对应 $[x_k]_{IND(C)}$ 中一对象，使得 $m' = \{a: a \in C \wedge f(x_i, a) \neq f(x_k, a)\} \neq \varnothing$ 且 $R \cap m' = \varnothing$，则有 $\rho_R(\vee m') = 0$ 故 $\rho_R(f_{DS}) = 0$，与前提 $\rho_R(F_{DS}) = \rho_C(F_{DS}) = 1$ 矛盾，故 $|f([x_i]_{IND(R)}, D)| > 1 \wedge |f([x_j]_{IND(R)}, D)| > 1$ 不成立，即有 $|f([x_i]_{IND(R)}, D)| = 1 \vee |f([x_j]_{IND(R)}, D)| = 1$ 成立. 因此，针对 $\forall (x_i, x_j) \in DIS(C, D)$ 均有 $f(x_i, R) \neq f(x_j, R) \wedge f(x_i, D) \neq f(x_j, D)$ 且 $|f([x_i]_{IND(R)}, D)| = 1 \vee |f([x_j]_{IND(R)}, D)| = 1$ 成立，即 $(x_i, x_j) \in DIS(R, D)$，故 $DIS(C, D) \subseteq DIS(R, D)$.

又由性质2有 $DIS(R, D) \subseteq DIS(C, D)$，因此有 $DIS(R, D) = DIS(C, D)$.

综上所述，当 $M_{DS} = \{m_{ij}\} \neq \varnothing$ 时，$DIS(R, D) = DIS(C, D)$ 的充分必要条件是 $\rho_R(F_{DS}) = \rho_C(F_{DS})$，即 $DIS(R, D) = DIS(C, D)$ 当且仅当 $\rho_R(F_{DS}) = \rho_C(F_{DS})$.

联立定理1、定理2可得，给定决策系统 $DS = \langle U, C \cup D, V, f\rangle$，$M_{DS} = \{m_{ij}\} \neq \varnothing$，则 $R(R \subseteq C)$ 为 C 相对于 D 的属性约简当且仅当 R 满足：①$\rho_R(F_{DS}) = 1$；② $\forall a \in R$，$\rho_{R-\{a\}}(F_{DS}) = 0$. 换而言之，若 R 是保持分辨函数值为真的极小条件属性集，则 R 为 C 相对于 D 的一个属性约简.

性质3 给定决策系统 $DS = \langle U, C \cup D, V, f\rangle$，$M_{DS} = \{m_{ij}\} \neq \varnothing$，$R \subseteq C$，则$\rho_R(F_{DS}) = 1$ 当且仅当对 $\forall m_{ij} \neq \varnothing$ 均有 $R \cap m_{ij} \neq \varnothing$.

证 若 $\rho_R(F_{DS}) = 1$，则对 $\forall m_{ij} \neq \varnothing$ 均有 $\rho_R(\vee m_{ij}) = 1$，即 $R \cap m_{ij} \neq \varnothing$. 若对 $\forall m_{ij} \neq \varnothing$ 均有 $R \cap m_{ij} \neq \varnothing$，则对 $\forall m_{ij} \neq \varnothing$ 均有 $\rho_R(\vee m_{ij}) = 1$，即 $\rho_R(f_{DS}) = 1$. 因此，当 $M_{DS} = \{m_{ij}\} \neq \varnothing$ 时，给定 $R \subseteq C$，$\rho_R(F_{DS}) = 1$ 的充要条件是对 $\forall m_{ij} \neq \varnothing$ 均有 $R \cap m_{ij} \neq \varnothing$.

根据性质3，可将相对约简定义转换为对应于 M_{DS} 的等价描述. 给定决策系统 $DS = \langle U, C \cup D, V, f\rangle$，$M_{DS} = \{m_{ij}\} \neq \varnothing$，则 $R(R \subseteq C)$ 为 C 相对于 D 的属性约简当且仅当 R 满足：

1) $\forall m_{ij} \in M_{DS}(m_{ij} \neq \varnothing \Rightarrow R \cap m_{ij} \neq \varnothing)$；

2) $\forall a \in R$，$\exists m_{ij} \in M_{DS}(m_{ij} \neq \varnothing \wedge (R - \{a\}) \cap m_{ij} = \varnothing)$.

2 基于启发式二分策略的属性约简方法

定理2表明，约简对应使分辨函数映射值为真的极小属性集. 本节进一步分析分辨集、分辨函数与约简的对应关系，给出相关的定理及证明，并在此基础上设计求取决策系统全体约简的有效算法.

定理3 令 G_{DS} 是将分辨函数 F_{DS} 运用命题逻辑吸收律和分配律等值变换后的极小析取范式(即析取范式中任意两项简单合取式之间不存在属性集包含关系)，那么存在 h 和 $X_i \subseteq C(1 \leqslant i \leqslant h)$ 使得 $G_{DS} = (\wedge X_1) \vee \cdots \vee (\wedge X_h)$，则决策系统 DS 的所有约简 $RED(DS) = \{X_1, \cdots X_h\}$.

证 针对 $\forall X_i \in \{X_1, \cdots, X_h\}(1 \leqslant i \leqslant h)$，由布尔映射函数定义均有 $\rho_{X_i}(\wedge X_i) = 1$，因此 $\rho_{X_i}(G_{DS}) = 1$；因 G_{DS} 中任意2项简单合取式间不会存在属性集包含关系，即对 $\forall X_i, X_j \in \{X_1, \cdots, X_h\}$ $(1 \leqslant i, j \leqslant h, j \neq i)$ 均有 $X_j - X_i \neq \varnothing$，则对 $\forall X_i, X_j \in \{X_1, \cdots X_h\}(1 \leqslant i, j \leqslant h)$ 与 $\forall a \in X_i$ 均有 $X_j - (X_i - \{a\}) \neq \varnothing$，那么 $\rho_{X_i - \{a\}}(\wedge X_j) = 0$，因此对 $\forall a \in X_i$ 均有 $\rho_{X_i - \{a\}}(G_{DS}) = 0$. 又因 G_{DS} 是将分辨函数 F_{DS} 运用吸收律和分配律等值变换后的极小析取范式，根据命题逻辑可知对 $\forall B \subseteq C$ 均有 $\rho_B(F_{DS}) = \rho_B(G_{DS})$，因此可将上述性质描述为：对 $\forall X_i \in \{X_1, \cdots, X_h\}(1 \leqslant i \leqslant h)$ 均有 $\rho_{X_i}(F_{DS}) = 1$；对

$\forall a \in X_i$ 均有 $\rho_{X_i-\{a\}}(F_{DS})=0$. 因此 $\forall X_i \in \{X_1,\cdots,X_h\}(1\leqslant i\leqslant h)$ 均是一个约简，即 $\{X_1,\cdots,X_h\}\subseteq RED(DS)$.

依据定理 2，$\forall R \in RED(DS)$ 均满足：$\rho_R(F_{DS})=1$；$\forall a\in R$，$\rho_{R-\{a\}}(F_{DS})=0$. 又因 G_{DS} 是将分辨函数 F_{DS} 运用吸收律和分配律等值变换后的极小析取范式，根据命题逻辑可知对 $\forall B\subseteq C$ 均有 $\rho_B(F_{DS})=\rho_B(G_{DS})$. 因此 $\forall R\in RED(DS)$ 均满足：$\rho_R(G_{DS})=1$；$\forall a\in R$，$\rho_{R-\{a\}}(G_{DS})=0$. 由 $\rho_R(G_{DS})=1$ 可知至少存在某项 $X_i\in\{X_1,\cdots X_h\}(1\leqslant i\leqslant h)$ 使得 $R\supseteq X_i$，假设存在 2 项以上则 $\exists X_i, X_j\in\{X_1,\cdots X_h\}$ $(1\leqslant i\leqslant h, 1\leqslant j\leqslant h, i\neq j)$ 使得 $R\supseteq X_i \wedge R\supseteq X_j$，又 X_i 与 X_j 互不包含，则 $R-X_i\neq\varnothing$ 那么必 $\exists a\in R$ 使得 $\rho_{R-\{a\}}(\wedge X_i)=1$ 即 $\exists a\in R(\rho_{R-\{a\}}(G_{DS})=1)$，与 $\forall a\in R$，$\rho_{R-\{a\}}(D_{DS})=0$ 相矛盾，因此当前仅存在一项 $X_i\in\{X_1,\cdots,X_h\}(1\leqslant i\leqslant h)$ 使得 $R\supseteq X_i$，故此时有 $\rho_R(G_{DS})=\rho_R(\wedge X_i)$. 又因为 $\rho_R(G_{DS})=1$ 且 $\forall a\in R$，$\rho_{R-\{a\}}(G_{DS})=0$，故有 $\rho_R(\wedge X_i)=1$ 且 $\forall a\in R$，$\rho_{R-\{a\}}(\wedge X_i)=0$，则必有 $R=X_i$. 因此针对 $\forall R\in RED(DS)$ 均存在且仅存在一项 $X_i\in\{X_1,\cdots,X_h\}(1\leqslant i\leqslant h)$ 使得 $R=X_i$，即对 $\forall R\in RED(DS)$ 均有 $R\in\{X_1,\cdots,X_h\}$，即 $RED(DS)\subseteq\{X_1,\cdots,X_h\}$.

综上所述，若 $G_{DS}=(\wedge X_1)\vee\cdots\vee(\wedge X_h)$ 是将 F_{DS} 运用吸收律和分配律等值变换后的极小析取范式，则决策系统 DS 的所有约简 $RED(DS)=\{X_1,\cdots,X_h\}$.

由定理 3 可知，求取所有约简问题可转化为将分辨函数 F_{DS} 转化为等价的极小析取范式 G_{DS}，G_{DS} 中每一析取项(极小析取范式中的每一简单合取式 $\wedge X_i(1\leqslant i\leqslant h)$)对应一个约简，$G_{DS}$ 中全体析取项对应决策系统全体约简. 为生成分辨函数的极小析取范式，需在等价变换过程中运用命题逻辑吸收律和分配律. 相应地，可基于布尔函数操作对求取所有约简集问题建模.

定义 7　给定决策系统 $DS=\langle U, C\cup D, V, f\rangle$，$M_{DS}=\{m_{ij}\}\neq\varnothing$，其分辨集 ϵ 定义为 M_{DS} 中所有非空元素构成的集合，即分辨集 $\epsilon=\{m_{ij}: m_{ij}\in M_{DS}\wedge m_{ij}\neq\varnothing\}$，分辨函数可表示为 $F_{DS}=F(\epsilon)=\wedge\{\vee\epsilon_k: \epsilon_k\in\epsilon\}$.

定义 8　给定决策系统 $DS=\langle U, C\cup D, V, f\rangle$，分辨集 $\epsilon=\{m_{ij}: m_{ij}\in M_{DS}\wedge m_{ij}\neq\varnothing\}$，其极小分辨集定义为 $\min(\epsilon)=\{\epsilon_k: \epsilon_k\in\epsilon\wedge((\forall\epsilon_l\in\epsilon\wedge k\neq l)\rightarrow(\epsilon_l\not\subset\epsilon_k))\}$，即 $\min(\epsilon)$ 是 ϵ 中删除所有超集元素后形成的集合(若某元素 $\epsilon_l\supset\epsilon_k(k\neq l)$，则 ϵ_l 为超集元素).

若 $\epsilon_k\subseteq\epsilon_l$ 则据命题逻辑吸收律可得 $(\vee\epsilon_k)\wedge(\vee\epsilon_l)=(\vee\epsilon_k)$，故有 $F(\epsilon)=F(\epsilon-\{\epsilon_l\})$. 通过多次使用吸收律，可将分辨集 ϵ 转换为其极小分辨集，因而 $F_{DS}=F(\epsilon)=F(\min(\epsilon))$. 如此，则可按如下步骤直接得到极小分辨集：① 分辨集 ϵ 初始值置空；② 若分辨矩阵所有元素提取完毕则转 ⑤，否则提取分辨矩阵当前元素 $M(i,j)$；③ 若 ϵ 中存在某元素包含于 $M(i,j)$ 转 ④，否则删除 ϵ 中所有包含 $M(i,j)$ 的元素，然后将 $M(i,j)$ 加入 ϵ；④ 生成下一当前元素 $M(i,j)$ 转 ③；⑤ 结束. 操作结束，ϵ 即为极小分辨集.

直接生成极小分辨集流程遵循了合取范式的消解原理，因此与原分辨集在求取约简结果上是等价的，且最大限度减少了存储分辨矩阵的元素数目.

为实现基于分辨矩阵求取全体属性约简的启发式算法，以下先给出分辨集核、分辨集关于特定属性的排除集定义，再给出算法的实现原理及证明.

定义 9　分辨集 ϵ 的核定义为 ϵ 中所有单属性元素构成的集合，即 $CORE(\epsilon)=\{\epsilon_k: \epsilon_k\in\epsilon_\wedge|\epsilon_k|=1\}$.

定理 4　$CORE(\epsilon)=\bigcap RED(DS)$，其中 $RED(DS)$ 为决策系统 DS 全体约简.

证　① 对 $\forall a\in CORE(\epsilon)$ 可知单属性 a 对应分辨函数 F_{DS} 中的一个合取项，由性质 3 对 $\forall R\in RED(DS)$，$R\subseteq C$ 均有 $\rho_R(F_{DS})=1$，则必有 $\rho_R(a)=1$ 即 $a\in R$. 若 $a\in CORE(\epsilon)$，则对 $\forall R\in RED(DS)$ 均有 $a\in R$，所以 $CORE(\epsilon)\subseteq\bigcap RED(DS)$.

② 对 $\forall a\in\bigcap RED(DS)$，令 $RED(DS)=\{X_1,\cdots,X_h\}$，则 $a\in X_i(1\leqslant i\leqslant h)$. 若决策系统 DS 仅存在一个约简则 $G_{DS}=\wedge X_1$，$a\in X_1$ 则单属性 a 必对应 F_{DS} 中一合取项，即 $a\in\epsilon$ 则必有 $a\in CORE(\epsilon)$；若 DS 存在多个约简则可将 G_{DS} 转化为 $G_{DS}=(\wedge X_1)\vee\cdots\vee(\wedge X_h)=a\wedge(\wedge(X_1-\{a\}))\vee\cdots\vee(\wedge(X_h-\{a\}))$，又因为 G_{DS} 是将合取范式 F_{DS} 等值变换后的极小析取范式，故单属性 a 必对应 F_{DS} 中一合取

项，即单属性 $a \in \epsilon$，故 $a \in CORE(\epsilon)$. 因此对 $\forall a \in \bigcap RED(DS)$ 均有 $a \in CORE(\epsilon)$，即 $\bigcap RED(DS) \subseteq CORE(\epsilon)$.

综合 ①② 可得 $CORE(\epsilon) = \bigcap RED(DS)$.

定义 10 给定分辨集$\epsilon \neq \emptyset$，分辨集ϵ关于属性 $a(a \in C)$ 的排除集定义为：$\epsilon^{-}(a) = \{\epsilon_k - \{a\}: (\epsilon_k \in \epsilon) \wedge ((\epsilon_k - \{a\}) \neq \emptyset)\}$.

性质 4 给定极小分辨集$\epsilon \neq \emptyset$，a 为ϵ中任一非核属性，$\epsilon' = \epsilon - \{\epsilon_k: (\epsilon_k \in \epsilon) \wedge (\epsilon_k \cap \{a\} \neq \emptyset)\}$. 若$\epsilon' \neq \emptyset$则$F(\epsilon) = (a \wedge F(\epsilon')) \vee (F(\min(\epsilon^{-}(a))))$；若$\epsilon' = \emptyset$，则 $F(\epsilon) = (a) \vee (F(\min(\epsilon^{-}(a))))$.

证明：

① 若$\epsilon' \neq \emptyset$，令$\epsilon = \{\epsilon_1, \cdots, \epsilon_r, \epsilon_{r+1}, \cdots, \epsilon_s\}$且非核属性 $a \in \epsilon_k (1 \leqslant k \leqslant r)$，$a \notin \epsilon_k (r+1 \leqslant k \leqslant s)$，则

$$
\begin{aligned}
F(\epsilon) = & F(\epsilon_1, \cdots, \epsilon_r) \wedge F(\epsilon_{r+1}, \cdots, \epsilon_s) = \\
& (a \vee F(\epsilon_1^{-}(a), \cdots, \epsilon_r^{-}(a))) \wedge F(\epsilon_{r+1}, \cdots, \epsilon_s) = \\
& (a \wedge F(\epsilon_{r+1}, \cdots, \epsilon_s)) \vee (F(\epsilon_1^{-}(a), \cdots, \epsilon_r^{-}(a)) \wedge F(\epsilon_{r+1}, \cdots, \epsilon_s)) = \\
& (a \wedge F(\epsilon')) \vee (F(\epsilon^{-}(a))) = \\
& (a \wedge F(\epsilon')) \vee (F(\min(\epsilon^{-}(a))))
\end{aligned}
$$

② 若$\epsilon' = \emptyset$，令$\epsilon = \{\epsilon_1, \cdots, \epsilon_r\}$，则对 $\forall\ \epsilon_k \in \epsilon$，$1 \leqslant k \leqslant r$ 均有非核属性$a \in \epsilon_k$，故

$$
\begin{aligned}
F(\epsilon) = & F(\epsilon_1, \cdots, \epsilon_r) = \\
& (a) \vee F(\epsilon_1^{-}(a), \cdots, \epsilon_r^{-}(a)) = \\
& (a) \vee (F(\epsilon^{-}(a))) = \\
& (a) \vee (F(\min(\epsilon^{-}(a))))
\end{aligned}
$$

给定极小分辨集$\epsilon \neq \emptyset$，若其子集$\epsilon' \neq \emptyset$，则ϵ'也是极小分辨集；$\epsilon^{-}(a)$不一定为极小分辨集，可将其转化为极小分辨集，且根据命题逻辑可知 $F(\epsilon^{-}(a))$ 与 $F(\min(\epsilon^{-}(a)))$ 等值. 对 $F(\epsilon')$ 与 $F(\min(\epsilon^{-}(a)))$ 递归运用同样分解过程，$F(\epsilon)$ 必能等价转化为析取范式，再应用命题逻辑吸收律可将其化为形如$(\wedge X_1) \vee \cdots \vee (\wedge X_h)$的极小析取范式，依据定理 3，$\{X_1, \cdots, X_h\}$ 即为决策系统全体约简.

定理 4 说明 $CORE(\epsilon)$ 包含于每一约简，性质 4 说明，将决策系统约简可分为 2 类：一类包含非核属性 a；另一类不包含非核属性 a，那么生成所有约简的整个过程对应由极小分辨集ϵ递归构造一棵二叉树的过程：① 根结点赋值为$CORE(\epsilon)$，若ϵ中不存在非核属性该结点停止扩展，否则执行 ② 与 ③；② 选择非核属性 a，左子树赋值为由分辨集$\{a\} \cup (\epsilon' - CORE(\epsilon))$所构造的二叉树；③ 右子树赋值为由分辨集 $\min(\epsilon^{-}(a) - CORE(\epsilon))$ 构造的二叉树. 一个包含属性 a 的约简对应左子树中某条从根结点到叶结点的路径，一个不包含属性 a 的约简对应右子树中某条从根结点到叶结点的路径，此即为本文基于极小分辨集构造约简树求取全体约简的二分策略；此外，在针对约简树中每一结点向下二分扩展时，以本结点所对应极小分辨集中属性的出现频度作为启发式信息，选取频度最高的非核属性a，可加快约简树的生成过程. 基于启发式二分策略的完全约简获取算法具体描述见算法 1.

算法 1 基于启发式二分策略的完全约简获取算法

输入：非空极小分辨集ϵ

输出：决策系统所有约简 $RED(\epsilon)$

初始化：$RED(\epsilon) = \emptyset$($RED(\epsilon)$ 为全局变量)

Void $RED\_HBS(\epsilon)$

$\{cur\_red = \emptyset$；

if $(\epsilon = \text{NULL})$ return；

if $((|\ \forall\ \epsilon_k \in \epsilon\ | = 1)$

  $\{cur\_red = \epsilon$；

将 cur_red 以吸收方式加入 $RED(\in)$；

}

else

{对$\in$中每一非核属性计算 $SIG(b)$；

$a = \max\{SIG(b)\}$；

$\in' = \in - \{\in_k: (\in_k \in \in) \wedge (\in_k \cap \{a\} \neq \varnothing)\}$；

$\in' = \in' \cup \{a\}$；

$RED\_HBS(\in')$；

$\in'' = \in^-(a)$ 的极小分辨集；

$RED\_HBS(\in'')$；

}//end if

}//end RED_HBS

设极小分辨集为$\in = \{\in_1, \cdots, \in_t\}$，按算法 1 生成的二叉树结点数为 N. 算法对每一内部结点，均要计算核属性集 $CORE(\in)$（时间复杂度为 $O(t)$），计算所有非核属性 b 在$\in$中的频度信息 $SIG(b, \in)$（时间复杂度不超过 $O(t \cdot |C|)$，选择 $a = \max\{SIG(b, \in)\}$ 的属性 a（时间复杂度为 $O(|C|)$），计算$\in^-(a)$的极小分辨集（时间复杂度不超过 $O(t^2)$），又由二叉树性质可得内部结点数为$(N-1)/2$，故算法对所有内部结点操作的时间复杂度不超过 $O(N \cdot (t \cdot |C| + t^2))$，一般情形下 $|C| \ll t$，因此算法对内部结点操作的时间复杂度不超过 $O(N \cdot t^2)$；对每个叶子结点均存在一条从根结点到该叶结点的路径，由二叉树特征可得叶子结点数为$(N+1)/2$，故算法 1 对所有叶子结点收集候选约简的时间复杂度不超过 $O(N \cdot |C|)$，将候选约简以吸收方式加入约简集的时间复杂度不超过 $O(N^2)$. 因此，整个算法时间复杂度不超过 $O(N \cdot t^2 + N^2)$.

文献[11]所提出的扩展法则本质上可归结为按广度优先遍历求解，必须获取约简树全体路径并执行逻辑吸收操作后才能得到全体解，而算法 1 按深度优先遍历以吸收方式加入当前候选约简 cur_red（cur_red 是否为约简取决于其是否被吸收），有效地消除了部分非真实约简，因而其耗费存储空间要少于扩展法则.

3　实验分析

为进一步验证本文算法的性能，选用 UCI 机器学习数据库中的若干数据集，CPU 为 Intel 双核 2.5 GHz，内存容量为 4 G，操作系统为 Windows 10，使用 Visual Studio 2017 实现了相关实验，实验结果见表 1—表 3.

表 1　分辨集与极小分辨集实验结果对比

数据集	对象数	属性数	直接提取分辨矩阵元素		直接生成极小分辨集	
			元素数目	元素长度和	元素数目	元素长度和
soybean-small	47	35	810	10 111	99	947
zoo	101	16	3 873	30 022	14	50
monks-1	432	6	46 656	178 848	3	3
tic-tac-toe	958	9	207 832	1 227 844	36	72
kr-vs-kp	3 196	36	2 548 563	24 932 433	29	31
mushroom	8 124	22	16 478 528	209 012 760	30	198

表 1 给出直接提取分辨矩阵元素得到分辨集与直接生成极小分辨集的实验数据对比. 其中“元素数目”代表所生成分辨集中的所有元素的数目，“元素长度和”代表所有元素中包含的属性数目之和. 实验数据表明，相对于提取分辨矩阵所有元素而言，直接生成极小分辨集大大减少了存储空间.

表 2　启发式与否扩展结点生成全体约简对比

数据集	约简数	生成分辨集时间/s	启发式扩展结点		非启发式扩展结点	
			路径数	求取所有约简时间/s	路径数	求取所有约简时间/s
soybean-small	765	0.007	949	0.066	6 629	0.967
zoo	33	0.008	33	<0.001	90	<0.001
monks-1	1	0.075	1	<0.001	1	<0.001
tic-tac-toe	9	0.307	9	<0.001	9	<0.001
kr-vs-kp	4	3.931	4	<0.001	4	<0.001
mushroom	292	27.763	345	0.011	1 464	0.149

表2针对算法1给出其与不运用启发式信息扩展结点的实验结果对比．其中“生成分辨集时间”表示获取极小分辨集的时间(单位：s)，“路径数”表示所构造约简树从根到叶结点的路径数目，“启发式扩展结点”表示使用属性重要度作为启发式信息选择分支属性，“非启发式扩展结点”表示不运用启发式信息而仅按照预先设置的顺序选择分支属性．数据集“zoo”，“ monks-1”，“ tic-tac-toe”与“kr-vs-kp”在是否运用启发式信息求取所有约简的时间上难以体现差别，均小于0.001 s，主要原因在于此4个数据集所构造约简树的路径数目太少且相差不大．对于数据集“soybean-small”与“mushroom”，是否运用启发式信息扩展结点构造约简树的路径数目相差较大，因此前者求取全体约简时间耗费不到后者的1/10．

表 3　算法 1 与扩展法则算法所需存储空间对比

数据集	扩展法则算法		算法 1	
	存储元素数目	所需存储空间	存储元素数目	所需存储空间
soybean-small	949	5 057	765	4 016
zoo	33	199	33	199
monks-1	1	3	1	3
tic-tac-toe	9	72	9	72
kr-vs-kp	4	116	4	116
mushroom	345	1 956	292	1 640

表3给出了运用算法1求取所有属性约简与利用文献[11]中算法(简称扩展法则)求取全体约简所耗费空间(单位：B)的实验对比数据．数据集“zoo”，“ monks-1”，“ tic-tac-toe”与“kr-vs-kp”在2种算法上的初始候选约简集与最终约简集相等，因此所耗费的存储空间无法体现出区别．数据集“soybean-small”与“mushroom”上的实验结果表明算法1所耗费的存储空间比扩展法则少，主要原因在于算法1采取深度优先方式每当探索到候选约简均以吸收方式加入，而扩展法则通过广度优先方式必须找到所有候选约简再吸收生成全体约简．

4　结　　论

属性约简是粗糙集理论的重要研究内容，尤其是最小约简对于泛化能力具有重要意义．本文设计了一种构造约简树获取信息系统全体约简的有效算法：该算法提取分辨矩阵元素直接生成极小分辨集，根据启发式信息运用二分策略操作极小分辨集构造一棵约简树，该约简树从根结点到叶结点的每一条路径对应一个候选约简；在约简树的构造过程中采用深度优先遍历以吸收方式加入当前候选约简，当前候选约简是否为约简取决于其是否被吸收，有效地缩减了搜索空间和存储空间．该方法适用于满足任意约简准则的分辨矩阵，能够保证全体约简求解的完备性，理论分析与实验结果说明了算法的可行性与有效性．

参考文献：

[1]　黄卫华．多粒度粗糙集模型[J]．西南师范大学学报(自然科学版)，2017，42(5)：137-143．
[2]　项海飞．基于互信息的二进制区分矩阵特征约简方法[J]．西南师范大学学报(自然科学版)，2015，40(8)：33-38．
[3]　刘少辉，盛秋戬，吴　斌，等．Rough集高效算法的研究[J]．计算机学报，2003，26(5)：524-529．

[4] 徐章艳，刘作鹏，杨炳儒，等. 一个复杂度为 max(O(|C||U|)，O(|C|～2|U/C|))的快速属性约简算法 [J]. 计算机学报，2006，29(3)：391-399.

[5] 胡　峰，王国胤. 属性序下的快速约简算法 [J]. 计算机学报，2007，30(8)：1429-1435.

[6] YAO Y Y，ZHAO Y. Discernibility Matrix Simplification for Constructing Attribute Reducts [J]. Information Sciences，2009，179(7)：867-882.

[7] 常玉慧，郭庆军，钱　进. 不一致决策表快速知识约简算法研究 [J]. 小型微型计算机系统，2015，36(8)：1852-1856.

[8] 黄治国，程　浩. 不同约简准则的比较研究 [J]. 计算机工程与应用，2015，51(17)：134-139.

[9] 黄治国，王　淼. 不完备决策表中基于容差关系的属性约简方法 [J]. 微电子学与计算机，2016，33(6)：147-151，156.

[10] 黄治国，杨清琳. 基于决策向量的分辨矩阵构造方法 [J]. 湖南科技大学学报(自然科学版)，2017，32(2)：62-69.

[11] STARZYK J A，NELSON D E，STURTZ K. A Mathematical Foundation for Improved Reduct Generation in Information Systems [J]. Knowledge and Information Systems，2000，2(2)：131-146.

[12] 黄治国，杨晓骥. 基于改进分辨矩阵的属性约简方法 [J]. 计算机仿真，2014，31(9)：305-309.

On Method of Attribute Reduction Based on Heuristic Binary Strategy

HUANG Zhi-guo[1]，　YANG Qing-lin[2]

1. School of Computer Science，Henan University of Engineering，Zhengzhou 451191，China；

2. Modern Educational Technology Department，Guangxi University of Finance and Economics，Nanning 530003，China

Abstract：In the method of discernibility matrix，a kind of normative and precise mathematical model can be provided for finding the core and attribute reduction，which is capable of acquiring all attribute reductions conveniently. An efficient approach based on discernibility matrix is designed for constructing reduction tree by means of heuristic information and binary strategy. In this binary tree，all paths from root to leaf are corresponding to all disjunctors of disjunctive form which is equivalent to discernibility function，and all disjunctors of its extreme minimal disjunctive form are corresponding to all reductions of decision system. The proposed algorithm is applicable to the discernibility matrix satisfying any reduction criterion，and be capable of reducing the searching scope significantly，and finding out all reductions definitely. The theoretical analysis and experimental results show the feasibility and effectiveness of the proposed algorithm.

Key words：attribute reduction；heuristic information；binary strategy；discernibility matrix

责任编辑　崔玉洁