知识表达系统的简化与集族的极小子集（Ⅰ）*）

李小霞　陈绵云

（华中科技大学控制科学与工程系　武汉430074）

摘　要　论文定义集族的Ⅰ型极小子集和Ⅱ型极小子集，揭示集族的Ⅰ型极小子集问题和Ⅱ型极小子集问题的转化关系，Ⅰ型极小子集族与Ⅰ型核以及Ⅱ型极小子集族与Ⅱ型核之间的关系，描述了生成集族近似Ⅰ型极小子集的贪心方法和生成所有的Ⅰ型极小子集的幂集树方法。

关键词　Ⅰ型极小子集，Ⅱ型极小子集

Reduct in a Knowledge Representation System and Teeny Subset Related to a Subset Family（Ⅰ）

LI Xiao-Xia　CHEN Mian-Yun

(Dept. of Control Sci. & Eng, Huazhong Univ. of Sci. & Tech., Wuhan 430074)

Abstract　The paper defines two new concepts, including type Ⅰ teeny subset and type Ⅱ teeny subset related to a subset family. Then it reveals that how the two problems are converted to one another. And it discovers the relation between the family consisting of all type Ⅰ teeny subsets and type Ⅰ core, and the relation between the family consisting of all type Ⅱ teeny subsets and type Ⅱ core. It also describes a greedy method to produce an approximate type Ⅰ teeny subset and a power tree method to produce all the type Ⅰ teeny subsets.

Keywords　Type Ⅰ teeny subset, Type Ⅱ teeny subset

粗集理论由波兰的 Pawlak 于1982年提出[1]，它的应用多数涉及决策表条件属性的简化和决策表条件属性值的简化。基本粗集理论[2,3]定义了知识表达系统的三种简化并研究了核与简化的关系。为了降低问题的复杂度，Skowron[4]定义了知识表达系统的分辨矩阵并用它来指导知识表达系统属性的简化。这种方法自提出后得到诸多的应用、发展与改造，因此有必要对它的基础进行研究。《知识表达系统的简化与集族的极小子集（Ⅰ）》定义了集族的Ⅰ型极小子集与Ⅱ型极小子集，研究它们的性质，探讨它们的生成方法。《知识表达系统的简化与集族的极小子集（Ⅱ）》则定义知识表达系统的不可分辨矩阵，并且通过不可分辨矩阵，知识表达系统的三大属性简化问题被转化为生成集族的Ⅱ型极小子集问题，由此可导出如何通过分辨矩阵将这三大属性简化问题转化为生成集族的Ⅰ型极小子集问题。

1. 集族的Ⅰ型极小子集

定义1（集族的Ⅰ型极小子集）　U 是非空有限集，$F \subseteq (P(U)-\{\varnothing\})$，$P(U)$ 表示 U 的幂集，$F \neq \varnothing$，A 是 U 的非空子集，如果 A 满足下面的两个条件：*a*）对任意的 $X \in F$，$X \cap A \neq \varnothing$；*b*）不存在 A 的非空真子集 A^*，满足对于任意的 $X \in F$，$X \cap A^* \neq \varnothing$，则称 A 为 (U,F) 的Ⅰ型极小子集。x 是 U 中任意一个元素，如果 $\{x\} \in F$，则称 x 不能去掉。定义 $\mathrm{MinSub}^1(U,F)=\{Y \mid Y$ 是 (U,F) 的Ⅰ型极小子集$\}$，称为 (U,F) 的Ⅰ型极小子集族，定义 $\mathrm{Core}^1(U,F)=\{x \mid x \in U$，并且 x 不能去掉$\}$，称为 (U,F) 的Ⅰ型核。

定理1　U 是非空有限集，$F \subseteq (P(U)-\{\varnothing\})$，$F \neq \varnothing$，*a*）$Y$ 是 U 的非空子集，Y 为Ⅰ型极小子集的充要条件是：对任意的 $X \in F$，$X \cap Y \neq \varnothing$，并且对任意的 $y \in Y$，存在 $A \in F$，使得 $A \in Y=\{y\}$。*b*）$\mathrm{Core}^1(U,F)=\bigcap \mathrm{MinSub}^1(U,F)$。

证明：*a*）必要性：$\forall\, y \in Y$，根据前提可知，存在 $A^* \in F$，使得 $(Y-\{y\}) \cap A^* = \varnothing$，因为 $Y \cap A^* \neq \varnothing$，所以 $(Y-\{y\}) \cap A^* = \{y\}$。充分性：假设 Y^* 是 Y 的真子集，任取 $y^* \in Y-Y^*$，则 $Y^* \subseteq Y-\{y^*\}$，根据前提可知：对 y^*，存在 $A^* \in F$，使得 $Y \cap A^* = \{y^*\}$，因此 $(Y-\{y^*\}) \cap A^* = \varnothing$，由此可推得 $Y^* \cap A^* = \varnothing$。*b*）容易推得 $\mathrm{Core}^1(U,F) \subseteq \bigcap \mathrm{MinSub}^1(U,F)$，下面证明 $\mathrm{Core}^1(U,F) \supseteq \bigcap \mathrm{MinSub}^1(U,F)$。对任意的 $y \in \bigcap \mathrm{MinSub}^1(U,F)$，假设 $\{y\} \notin F$，则 $U-\{y\}$ 必有非空子集是Ⅰ型极小子集，这与 $y \in \bigcap \mathrm{MinSub}^1(U,F)$ 矛盾，因而假设不成立，$\{y\} \in F$，即 $y \in \mathrm{Core}^1(U,F)$。下面证明，如果 $\{y\} \notin F$，则 $U-\{y\}$ 必有非空子集是Ⅰ型极小子集：如果 $U=\{y\}$，则 $P(U)=\{\{y\},\varnothing\}$，此时 $F=\{\{y\}\}$，这与 $\{y\} \notin F$ 矛盾，因此 $U \neq \{y\}$，即 $U-\{y\} \neq \varnothing$；对任意的 $A \in F$，因为 $\{y\} \notin F$，所以 $A-\{y\} \neq \varnothing$，即存在 $a \in A-\{y\}$，因此 $(U-\{y\}) \cap A \neq \varnothing$；假设 $U-\{y\}$ 的任何非空子集都不是Ⅰ型极小子集，既然 $U-\{y\} \neq \varnothing$，所以 $U-\{y\}$ 不是Ⅰ型极小子集，根据极小子集的定义，可推得存在 $U-\{y\}$ 的非空真子集 Y^1，使得对任意的 $A \in F$，$Y^1 \cap A \neq \varnothing$；如果存在 $U-\{y\}$ 的非空真子集 Y^k，使得对任意的 $A \in F$，$Y^k \cap A \neq \varnothing$，既然 Y^k 是 $U-\{y\}$ 的非空子集，所以 Y^k 不是Ⅰ型极小子集，根据极小子集的定义可推得存在 Y^k 的非空真子集 Y^{k+1}，使得对任意的 $A \in F$，$Y^{k+1} \cap A \neq \varnothing$；根据数学归纳法可推得 $U-\{y\}$ 至少有自然数多个非空真子集，这与 U 是有限集矛盾，所以假设不成立，$U-\{y\}$ 的非空子集中存在Ⅰ型极小子集。证毕。

*）自然科学基金（79970025）。**李小霞**　博士生，研究方向包括知识发现、粗集理论及其应用；**陈绵云**　教授，博导，研究方向包括灰色动态建模与控制、模糊控制与决策、一般系统论及应用。

2. I 型极小子集的生成方法

使用下面的贪心方法可以生成(U,F)的一个近似 I 型极小子集。

输入:$U=\{a_1,a_2,\cdots,a_n\}$,$F=\{F_1,F_2,\cdots,F_m\}$

输出:(U,F)的近似 I 型极小子集

方法:

Step1:Fam 表示(U,F)的近似 I 型极小子集,将 Fam 清空;依次扫描 F 中的元素,设当前被扫描的元素为 F_i,对 F_i 作下述的操作:若 F_i 是单元素集,则 $Fam=Fam\cup F_i$,并将 F 中包含 F_i 的元素删除,将 F_i 中的元素从 U 中删除。

Step2:如果 F 为空,则进入 Step4,否则进入 Step3。

Step3:对任意的 $a\in U$,生成 Weight(a),方法如下:Weight(a)=0,依次扫描 F 中的元素,如果被扫描的元素中含有 a,则 Weight(a)=Weight(a)+1;从 U 中选择 a^*,使得对于任意的 $a'\in U$,Weight(a^*)≥Weight(a'),将 a^* 添加到 Fam 中,将 a^* 从 U 中删除,将 F 中含 a^* 的元素删除,返回 Step2。

Step4:输出 Fam。

注:这只是一个近似解法,所得 Fam 满足对任意的 $A\in F$,$Fam\cap A\neq\varnothing$,但是有可能存在 Fam 的真子集 Fam^*,满足对任意的 $A\in F$,$Fam^*\cap A\neq\varnothing$。

例如 $U=\{a,b,c,d,e,f\}$,$F=\{F_1,F_2,F_3,F_4,F_5,F_6\}$,$F_1=\{a,d,f\}$,$F_2=\{a,b,d,e\}$,$F_3=\{a,f\}$,$F_4=\{c,d,e\}$,$F_5=\{b,e\}$,$F_6=\{c,f\}$,对(U,F)执行上述的方法:Step1:$Fam=\varnothing$。Step2:$F\neq\varnothing$,进入 Step3。Step3:Weight(a)=Weight(d)=Weight(e)=Weight(f)=3,Weight(b)=Weight(c)=2,$a^*=a$,$Fam=\{a\}$,$U=\{b,c,d,e,f\}$,$F=\{F_4,F_5,F_6\}$。Step2:$F\neq\varnothing$,进入 Step3。Step3:Weight(b)=Weight(d)=Weight(f)=1,Weight(c)=Weight(e)=2,$a^*=e$,$Fam=\{a,e\}$,$U=\{b,c,d,f\}$,$F=\{F_6\}$。Step2:$F\neq\varnothing$,进入 Step3。Step3:Weight(b)=Weight(d)=0,Weight(c)=Weight(f)=1,$a^*=f$,$Fam=\{a,e,f\}$,$U=\{b,c,d\}$,$F=\varnothing$。Step2:$F=\varnothing$,进入 Step4。Step4:输出 $Fam=\{a,e,f\}$。通过上述的贪心算法得到的{a,e,f}满足对任意的 $A\in F$,$\{a,e,f\}\cap A\neq\varnothing$,但是它的真子集{e,f}也满足这个条件,所以{a,e,f}不是 I 型极小子集。

除了上述的贪心方法,也可以用幂集树方法生成所有的 I 型极小子集。U 是非空有限集,将它所有的非空子集组织在一棵树中,称为 U 的幂集树,记为 T(U),简记为 T,T 的顶点是 U 的所有非空子集,对 U 的任意非空子集 A 而言,如果 A=U,则 A 处于树 T 的第0层,如果 $|A|=|U|-k$,则 A 处于树 T 的第 k 层。将树 T 每一层的结点都排成一列,树 T 的边存在并且仅存在于树 T 相邻两层的结点之间。设树 T 第 k 层的元素序列为 $A_1^k,A_2^k,\cdots,A_l^k$,第 k+1层元素序列为 $A_1^{k+1},A_2^{k+1},\cdots,A_m^{k+1}$,则结点 A_i^{k+1}(i=1,2,…,m)的父结点为 $A_1^k,A_2^k,\cdots,A_l^k$ 中第一个能够包含它的元素。下述的方法被称为生成 $MinSub^1(U,F)$的幂集树方法。

输入:(U,F)及 U 的幂集树 T(U)

输出:$MinSub^1(U,F)$

Step1:生成 $F^{min}=\min(F)$,min(F)定义为$\{B|B\in F$,并且不存在 $B^*\in F$,使得 $B^*\subset B\}$。A^* 表示当前正在被扫描的 T(U)的结点,将 A^* 初始化为根结点 U,进入 Step2。

Step2:if(A^* 与 F^{min}中某些元素的交集为空集)

then{将当前幂集树中以 A^* 为根结点的子树剪枝}

进入 Step3。

Step3:if(A^* 所在层还有结点 B 排在 A^* 后面)

then{$A^*=B$}

else if (A^* 所在层的下一层非空)

then {令 A^* 为 A^* 所在层下一层的第一个结点}

else {$A^*=\varnothing$}

Step4:if(A^* 不为空集)

then { goto Step2}

else {将当前幂集树称为剪枝后的幂集树,并将它的所有的叶结点收集到集合 L 中,输出 $\min(L)$},方法完毕。

定理2 U 是非空有限集,$F\subseteq(P(U)-\{\varnothing\})$,$F\neq\varnothing$,则有下面的结论成立:*a*)定义 $F^{min}=\min(F)$,A 是 U 的任意非空子集,A 与 F 中任意一个元素的交集不为空集等价于 A 与 F^{min}中任意一个元素的交集不为空集。*b*)生成 $MinSub^1(U,F)$的幂集树方法输出的 $\min(L)$等于 $MinSub^1(U,F)$。

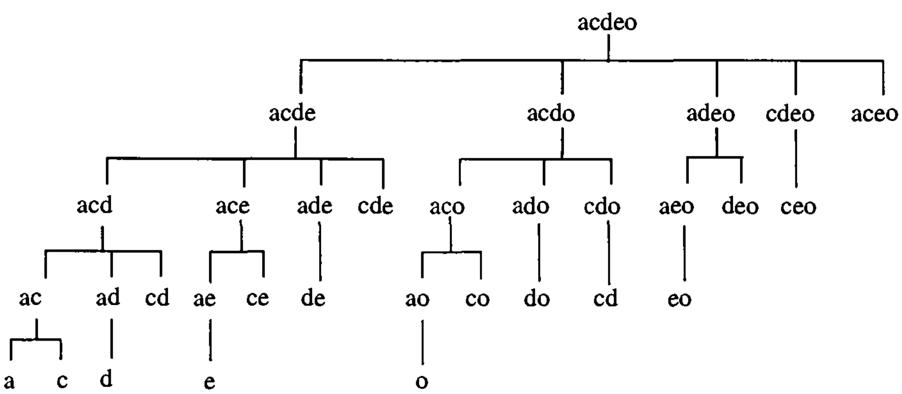

图1 T({a,c,d,e,o})

证明:*a*)必要性显然成立。充分性:X 是 F 中任意一个元素,则存在 $X^*\in F^{min}$,使得 $X\supseteq X^*$;因为 A 与 F^{min}中任意一个元素的交集非空,所以 $A\cap X^*\neq\varnothing$;因为 $X^*\subseteq X$,所以 $A\cap X\neq\varnothing$,充分性得证。*b*)首先证明对 U 的任意非空子集 A 而言,A 在剪枝后的幂集树中,等价于 A 和 F^{min}中任意元素的交集非空。必要性:由幂集树方法的描述可知剪枝后的幂集树中所有的结点都是被扫描过的,A 在剪枝后的幂集树中,所以 A 被扫描过并且未被剪枝,根据幂集树方法的描述可知 A 与 F^{min}中任意元素的交集非空。充分性:定义 $=R^A\{A^0,A^1,\cdots,A^k,\}$,$A^0=U$,$A^k=A$,$A^0,A^1,\cdots,A^k$,是剪枝前的幂集树中自

根结点至 A 的路径上的所有结点，由幂集树的构造过程，可知这些结点都包含 A。因为 A 与 F^{min} 中所有元素的交集非空，所以 R^A 中的元素与 F^{min} 中任意元素的交集非空。因此如果这些结点被扫描，将不会导致剪枝操作。由此可得出结论，A 在剪枝后的幂集树中依然存在。其次，根据此定理的第一个结论，可知对 U 的任意非空子集 A 而言，A 在剪枝后的幂集树中，等价于 A 和 F 中任意元素的交集非空，进而根据极小 I 型子集的定义，可推得 $\min(L)=\mathrm{MinSub}^1(U,F)$。证毕。

例如：U={a,c,d,e,o}，F={F_1,F_2,F_3}，F_1={a,c,d,e}，F_2={a,c,d,o}，F_3={c,d,e,o}，U 的幂集树 T(U)如图1所示，用幂集树方法生成(U,F)所有的 I 型极小子集，剪枝后的幂集树如图2所示，根据定理2，可得 $\mathrm{MinSub}^1(U,F)=\min(L)$ = min{{c}{d}{c,d}{a,e}{c,e}{d,e}{c,d,e}{a,o} {c,o}{d,o}{c,d}{e,o}{d,e,o}{c,e,o} {a,c,e,o}}={{c}{d}{a,e}{a,o}{e,o}}，这是最坏的一种情况，需要扫描31个结点，即幂集树中所有的结点。G={{o}{e}{c}{a,d} }，用幂集树方法生成(U,G)所有的 I 型极小子集，剪枝后的幂集树如图3所示，根据定理2，可得 $\mathrm{MinSub}^1(U,G)=\min(L)$=min{{c,d,e,o}{a,c,e,o}}={{c,d,e,o}{a,c,e,o}}，这是比较好的一种情况，只需要扫描6个结点。

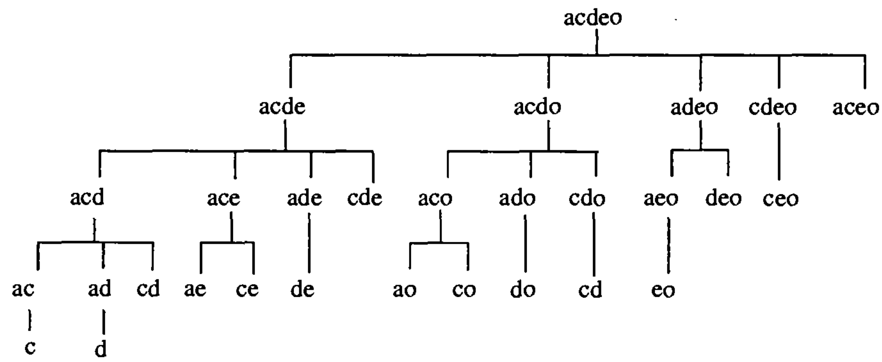

图2 根据({a,c,d,e,o},{{a,c,d,e}{a,c,d,o}{c,d,e,o}剪枝后的幂集树

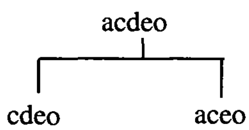

图3 根据({a,c,d,e,o},{{o}{e}{c}{a,d}})剪枝后的幂集树

3. 集族的 I 型极小子集

定义2(集族的 I 型极小子集) U 是非空有限集，F⊆(P(U)−{∅})，F≠∅，A 是 U 的非空子集，如果 A 满足下面的两个条件：*a* 对任意的 X∈F 而言，A 不包含在 X 中；*b* 不存在 A 的非空真子集 A^*，满足对于任意的 X∈F，A^* 不包含在 X 中，则称 A 为(U,F)的 I 型极小子集。

定理3 U 是非空有限集，F⊆(P(U)−{∅})，F≠∅，A 是 U 的非空子集，定义 $F'=\{X' \mid X\in F\}$(X' 表示子集 X 的补集)，则 A 是(U,F)的 I 型极小子集等价于 A 是(U,F')的 I 型极小子集。

证明：容易证明对任意的 U 的子集 B、C 而言，$B\cap C\neq\emptyset$ 等价于 B 不包含在 C' 中。A 是(U,F)的 I 型极小子集，等价于$\forall X\in F, A\cap X\neq\emptyset$并且如果 A^* 是 A 的非空真子集，则存在 $X^*\in F$，使得 $A^*\cap X^*=\emptyset$，等价于$\forall X\in F$，A 不包含在 X' 中并且如果 A^* 是 A 的非空真子集，则存在 $X^*\in F$，使得 $A^*\subseteq(X^*)'$，等价于 A 是(U,F')的 I 型极小子集。证毕。

定理4 U 是非空有限集，$F\subseteq$P(U)−{∅})，$F\neq\emptyset$，定义 $\mathrm{MinSub}^2(U,F)$={A|A 是(U,F)的 I 型极小子集}，称为(U,F)的 I 型极小子集族，u 是 U 中任意一个元素，如果存在 $X^*\in F$，使得 $X^*=\{u\}'$，则称 u 不能去掉，定义 $\mathrm{Core}^2(U,F)=\{u \mid u\in U$，并且 u 不能去掉$\}$，称为(U,F)的 I 型核，则 $\bigcap\mathrm{MinSub}^2(U,F)=\mathrm{Core}^2(U,F)$。

证明：根据定理1，$\bigcap\mathrm{MinSub}^1(U,F)=\mathrm{Core}^1(U,F')$，根据定理3，可推得$\bigcap\mathrm{MinSub}^2(U,F)=\mathrm{Core}^1(U,F)$，根据 $\mathrm{Core}^1(U,F')$的定义可推得$\bigcap\mathrm{MinSub}^2(U,F)=\{u \mid u\in U$，并且$\{u\}\in F'\}$，根据 F' 的定义可推得$\bigcap\mathrm{MinSub}^2(U,F)=\{u \mid u\in U$，并且$\{u'\}\in F\}$，进而推得 $\bigcap\mathrm{MinSub}^2(U,F)=\mathrm{Core}^2(U,F)$。证毕。

参考文献

1 Pawlak Z. Rough Sets. International J. of Information and Computer Sciences, 1982, 11(5): 341～356

2 Pawlak Z. Rough Sets—Theoretical Aspects of Reasoning about Data. Dordrecht: Kluwer Academic , 1991

3 曾黄麟. 粗集理论及其应用—关于数据推理的新方法(修订版). 重庆大学出版社, 1998

4 Skowron A, Rauszer C. The discernibility matrices and functions in information systems. In: Slowinski R, ed. Intelligent Decision Support—Handbook of Applications and Advances of the Rough Sets Theory. Dordrecht, Kluwer: Academic Publishers, 1992. 331～362