View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Xiamen University Institutional Repository

第 28 卷　第 6 期
2015 年 6 月
模式识别与人工智能
PR & AI
Vol. 28　No. 6
Jun.　2015

# 图支配集问题的粗糙集属性约简方法*

谭安辉[1,2]　李进金[2]　陈锦坤[2]　林国平[2]

[1]( 厦门大学 数学科学学院　厦门 361005)
[2]( 闽南师范大学 数学与统计学院　漳州 363000)

**摘　要**　探讨粗糙集的属性约简和图的支配集问题之间的联系. 通过构造信息系统,将粗糙集的属性约简问题与图的支配集问题相联系,从而把图的支配集问题转化为粗糙集的属性约简问题. 首先证明图的极小支配集恰是其构造的信息系统的属性约简,然后提出一种基于信息熵的最小支配集算法,最后通过实例验证该算法的可行性和有效性.

**关键词**　粗糙集,信息系统,属性约简,图,支配集,信息熵
**中图法分类号**　TP 18　　**DOI**　10.16451/j.cnki.issn1003-6059.201506004

# An Attribute Reduction Method Based on Rough Sets for Dominating Sets of Graph

TAN An-Hui[1,2], LI Jin-Jin[2], CHEN Jin-Kun[2], LIN Guo-Ping[2]
[1](*School of Mathematical Sciences*, *Xiamen University*, *Xiamen* 361005)
[2](*School of Mathematics and Statistics*, *Minnan Normal University*, *Zhangzhou* 363000)

## ABSTRACT

The relationship between attribute reduction problem in rough sets and dominating set problem in graph is discussed. By constructing an information system, the attribute reduction problem in rough sets is associated with the dominating set problem in graph, so as to transformed the dominating set problem into the attribute reduction problem. Firstly, it is proved that the minimal dominating set of a graph is exactly the attribute reduction of the constructed information system. Then, a minimum dominating set algorithm based on information entropy is proposed. Finally, A practical example illustrates the feasibility and efficiency of the proposed algorithm.

**Key Words**　Rough Set, Information System, Attribute Reduction, Graph, Dominating Set, Information Entropy

* 国家自然科学基金项目(No. 61379021,11301367,11061004)资助
收稿日期:2014-03-03;修回日期:2014-06-20
作者简介　谭安辉,男,1986 年生,博士研究生,主要研究方向为粗糙集、概念格、图论. E-mail: shujujiegouwang@126.com. 李进金(通讯作者),男,1960 年生,教授,博士生导师,主要研究方向为人工智能、粗糙集、拓扑学. E-mail: Jinjinli@munu.edu.cn. 陈锦坤,男,1978 年生,硕士,讲师,主要研究方向为人工智能、粗糙集、图论. 林国平,女,1978 年生,博士研究生,副教授,主要研究方向为人工智能、机器学习、粒计算.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

# 1 引 言

属性约简问题和支配集问题分别是粗糙集理论和图论的热点研究课题[1-2]，并且都有广泛的实际应用[2-6]. 目前针对这两个问题，已存在众多的求解算法[2,7-11]. 布尔逻辑运算是它们共同的基础方法[12-14]，即通过布尔逻辑运算，求得粗糙集的所有约简和图的所有极小支配集，分别对粗糙集的辨识集和图的邻接集进行析取、合取运算，得到粗糙集的属性约简和图的极小支配集.

但是，针对这两个问题的研究历来都是各自展开，很少考虑它们内在的更多联系. 因此，本文从粗糙集的辨识集和图的邻接集的关系出发，在布尔逻辑运算基础上，构造图在粗糙集理论中的一个信息系统表示，从而将粗糙集的属性约简问题与图的支配集问题相联系. 基于这一联系，证明图的极小支配集恰是其构造的信息系统的属性约简. 自然地，粗糙集的属性约简算法便可用来求解图的支配集. 特别地，引入粗糙集理论中一种基于信息熵的约简方法，并通过实例验证其可行性和有效性. 本文的工作有望对这两个问题带来新的研究思路.

# 2 基本概念

## 2.1 信息系统的属性约简

粗糙集的数据表示通常为一个信息系统. 下文引入信息系统及其属性约简的定义.

**定义1**[1] 称 $IS=(U,A,I,f)$ 为一个信息系统，其中，$U$ 是非空有限对象集，称为论域；$A$ 是非空有限属性集；$I=\bigcup_{a\in A} I_a$，$I_a$ 表示属性 $a$ 的值域；$f: U\times A\rightarrow I$ 为信息函数，对 $\forall x\in U, a\in A$，有 $f(x,a)\in I_a$.

任意属性子集 $B\subseteq A$ 决定一个二元不可辨识关系：

$$IND(B)=\{(x,y)\in U\times U \mid \forall a\in B, f(x,a)=f(y,a)\}.$$

$$U/IND(B)=\{[x]_B \mid x\in U\}$$

构成 $U$ 的一个划分，其中

$$[x]_B=\{y\in U \mid (x,y)\in IND(B)\}$$

为 $x$ 关于 $B$ 的等价类.

**定义2**[1,14] 设 $IS=(U,A,I,f)$ 是一个信息系统，称属性集 $B\subseteq A$ 是 $IS$ 的协调集，$IND(B)=IND(A)$. 进一步，若

$$IND(B-\{b\})\neq IND(B),\ \forall b\in B,$$

则称 $B$ 是 $IS$ 的约简.

如果一个属性属于所有的约简，则称这个属性为核心属性[1,6].

针对信息系统的属性约简问题，至今已有众多解决方法. 特别地，引入基于辨识矩阵的方法.

**定义3**[1,14] 设 $IS=(U,A,I,f)$ 为一个信息系统，对 $\forall x\in U, y\in U$，

$$d(x,y)=\{a\in A \mid f(x,a)\neq f(y,a)\}$$

称为 $(x,y)$ 的辨识集.

$$\boldsymbol{D}=\{d(x,y) \mid (x,y)\in U\times U\}$$

称为信息系统 $IS$ 的辨识矩阵. 对 $(x,y)\in U\times U$，若 $d(x,y)\neq\varnothing$，称 $(x,y)$ 为可区分对象对.

**引理1**[12,14] 设 $IS=(U,A,I,f)$ 是一个信息系统. 对 $\forall x\in U, y\in U$，有

1) $d(x,x)=\varnothing$，

2) $d(x,y)=d(y,x)$，

3) $y\notin [x]_A \Leftrightarrow d(x,y)\neq\varnothing$.

证明请参见文献[12]、文献[14].

**定义4**[12,14] 设 $IS=(U,A,I,f)$ 是一个信息系统，

$$f_D(a_1^*,a_2^*,\cdots,a_m^*)=\wedge\{\vee d(x,y) \mid d(x,y)\in D, d(x,y)\neq\varnothing\}$$

称为信息系统 $IS$ 的布尔函数或辨识函数，其中，布尔变量 $a_1^*,a_2^*,\cdots,a_m^*$ 分别对应属性 $a_1,a_2,\cdots,a_m$，$\vee d(x,y)$ 表示所有变量 $a^*\in d(x,y)$ 的析取式，$\wedge\{\vee d(x,y)\}$ 表示所有 $\vee d(x,y)$ 的合取式.

**定理1**[12,14] 设 $IS=(U,A,I,f)$ 是一个信息系统，$B\subseteq A$ 是 $IS$ 的一个约简当且仅当 $\bigwedge_{a\in B} a^*$ 是布尔函数中极小析取式中的一个合取式.

证明请参见文献[12]、文献[14].

由定理1，若

$$\begin{aligned}&f_D(a_1^*,a_2^*,\cdots,a_m^*)\\&=\wedge\{\vee d(x,y) \mid d(x,y)\in D, d(x,y)\neq\varnothing\}\\&=\bigvee_{i=1}^{t}\Big(\bigwedge_{j=1}^{s_i} a_j^*\Big),\end{aligned}$$

则

$$B_i=\{a_j \mid j\leqslant s_i\},\ i\leqslant t$$

是 $IS$ 所有的约简，其中 $\bigwedge_{j=1}^{s_i} a_j^*$，$i\leqslant t$ 是辨识函数 $f_D(a_1^*,a_2^*,\cdots,a_m^*)$ 所有极小析取式中的合取式.

**引理2**[12,14] 设 $IS=(U,A,I,f)$ 是一个信息系统，属性 $a\in A$ 是核心属性当且仅当存在 $(x,y)\in U\times U$，使得 $d(x,y)=\{a\}$.

证明请参见文献[12]、文献[14].

**例1** 给定一个信息系统 $IS=(U,A,I,f)$ 如表1所示，其中

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

$U = \{x_1, x_2, x_3, x_4, x_5, x_6\}$,

$A = \{v_1, v_2, v_3, v_4, v_5\}$.

利用布尔运算求解信息系统 $IS$ 的所有约简.

信息系统 $IS$ 的辨识矩阵 $\boldsymbol{D}$ 为

$$\boldsymbol{D} = \begin{pmatrix} \varnothing & \{v_1, v_2, v_3, v_4\} & \{v_1, v_2, v_3, v_4\} & \{v_1, v_2, v_3, v_4\} & \{v_1, v_2, v_3, v_4, v_5\} & \{v_1, v_2, v_3, v_4\} \\ & \varnothing & \{v_1, v_2, v_3\} & \{v_1, v_2, v_3, v_4\} & \{v_1, v_2, v_3, v_5\} & \{v_1, v_2, v_3\} \\ & & \varnothing & \{v_1, v_2, v_3, v_4\} & \{v_1, v_2, v_3, v_5\} & \{v_1, v_2, v_3\} \\ & & & \varnothing & \{v_1, v_4, v_5\} & \{v_1, v_4\} \\ & & & & \varnothing & \{v_5\} \\ & & & & & \varnothing \end{pmatrix}.$$

由辨识矩阵可看出,该信息系统的辨识集满足

$$d(x_i, x_j) \supseteq d(x_i, x_6), \ \forall 1 \leqslant i \leqslant 6, \ 1 \leqslant j \leqslant 6.$$

因此 $IS$ 的布尔函数为

$$\begin{aligned} & f_D(v_1^*, v_2^*, \cdots, v_5^*) \\ & = \wedge\{\vee d(x_i, x_j) \mid d(x_i, x_j) \neq \varnothing, 1 \leqslant i \leqslant 6, \\ & \quad 1 \leqslant j \leqslant 6\} \\ & = \bigwedge_{i=1}^{6} (\{\vee d(x_i, x_6) \mid d(x_i, x_6) \neq \varnothing\}) \\ & = (v_1 \wedge v_5) \vee (v_2 \wedge v_4 \wedge v_5) \vee (v_3 \wedge v_4 \wedge v_5), \end{aligned}$$

所以

$$\{v_1, v_5\}, \{v_2, v_4, v_5\}, \{v_3, v_4, v_5\}$$

是信息系统的所有约简,其中最小约简为$\{v_1, v_5\}$.

**表 1　信息系统 *IS***

Table 1　Information system *IS*

| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ |
|---|---|---|---|---|---|
| $x_1$ | 1 | 1 | 1 | 1 | 0 |
| $x_2$ | 2 | 2 | 2 | 0 | 0 |
| $x_3$ | 3 | 3 | 3 | 0 | 0 |
| $x_4$ | 4 | 0 | 0 | 4 | 0 |
| $x_5$ | 0 | 0 | 0 | 0 | 5 |
| $x_6$ | 0 | 0 | 0 | 0 | 0 |

由例 1 可知,对于一个信息系统,其布尔函数的运算结果实质上由较小的辨识集决定.

## 2.2　图的支配集

给定一个图 $G = (V, E)$,其中,$V$ 表示图的顶点集,$E$ 表示图的边集,对 $\forall v_1 \in V, v_2 \in V$,若$(v_1, v_2) \in E$,则称 $v_1, v_2$ 相邻. 顶点 $v \in V$ 称为 $G$ 的孤立点,当且仅当 $V - \{v\}$ 中的任意顶点和 $v$ 都不相邻. 对于顶点 $v \in V$,记 $N(v)$ 为所有与 $v$ 相邻的顶点集,

$$N[v] = N(v) \cup \{v\}$$

为 $v$ 的邻接集,显然 $v \in N[v]$,$N[v] \neq \varnothing$.

**定义 5**[2]　给定一个图 $G = (V, E)$,$S \subseteq V$ 为 $G$ 的顶点子集. 若对 $\forall v \in V$,则 $v \in S$ 或 $N(v) \cap S \neq \varnothing$,称 $S$ 为 $G$ 的支配集. 若 $S$ 是支配集,且 $S$ 的任何真子集都不是支配集,则称 $S$ 为 $G$ 的极小支配集; 若不存在支配集 $S'$,使得 $|S'| < |S|$,则称 $S$ 为 $G$ 的最小支配集,其中 $|S|$ 为 $S$ 的基数.

在图 $G = (V, E)$ 中,设

$$f_G(v_1^*, v_2^*, \cdots, v_m^*) = \wedge\{\vee N[v] \mid v \in V\},$$

其中,布尔变量 $v_1^*, v_2^*, \cdots, v_m^*$ 分别对应顶点 $v_1, v_2, \cdots, v_m$,$\vee N[v]$ 表示所有变量 $v_i^* \in N[v]$ 的析取式,$\wedge\{\vee N[v]\}$ 表示所有 $\vee N[v]$ 的合取式.

**定理 2**[2]　给定一个图 $G = (V, E)$,顶点集 $S \subseteq V$ 是 $G$ 的一个极小支配集,当且仅当 $\bigvee_{v \in S} v^*$ 是 $f_G(v_1^*, v_2^*, \cdots, v_m^*)$ 的极小析取式中的一个合取式.

证明请参见文献 [2].

由定理 2 得,若

$$\begin{aligned} & f_G(v_1^*, v_2^*, \cdots, v_m^*) \\ & = \wedge\{\vee N[v] \mid v \in V\} = \bigvee_{i=1}^{t} (\bigwedge_{j=1}^{s_i} v_j^*), \end{aligned}$$

则

$$S_i = \{v_j \mid j \leqslant s_i\}, \ i \leqslant t$$

是 $G$ 所有的极小支配集,其中 $\bigwedge_{j=1}^{s_i} v_j^*$, $i \leqslant t$ 是 $f_G(v_1^*, v_2^*, \cdots, v_m^*)$ 的所有的极小析取式中的合取式.

**例 2**　给定一个图 $G = (V, E)$,如图 1 所示,其中

$$V = \{v_1, v_2, v_3, v_4, v_5\}, \ E = \{e_1, e_2, e_3, e_4, e_5\}.$$

利用布尔运算求解该图的所有极小支配集过程如下:

$N[v_1] = \{v_1, v_2, v_3, v_4\}$, $N[v_2] = \{v_1, v_2, v_3\}$,

$N[v_3] = \{v_1, v_2, v_3\}$, $N[v_4] = \{v_1, v_4\}$, $N[v_5] = \{v_5\}$.

则有

$$\begin{aligned} & f_G(v_1^*, v_2^*, \cdots, v_5^*) \\ & = \wedge\{\vee N[v] \mid v \in V\} \\ & = (v_1 \vee v_2 \vee v_3 \vee v_4) \wedge (v_1 \vee v_2 \vee v_3) \wedge \\ & \quad (v_1 \vee v_2 \vee v_3) \wedge (v_1 \vee v_4) \wedge (v_5) \\ & = (v_1 \wedge v_5) \vee (v_2 \wedge v_4 \wedge v_5) \vee (v_3 \wedge v_4 \wedge v_5). \end{aligned}$$

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

可得图 $G$ 的所有极小支配集为

$$\{v_1,v_5\},\{v_2,v_4,v_5\},\{v_3,v_4,v_5\},$$

最小支配集为$\{v_1,v_5\}$.

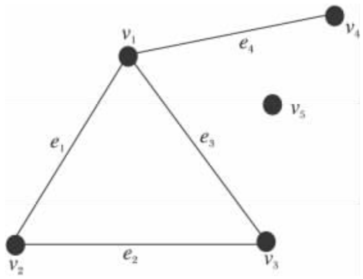


图 1 $G=(V,E)$

Fig. 1 $G=(V,E)$

## 3 图的诱导信息系统

可看出例 1 中的信息系统和例 2 中的图具有相同的布尔运算结果,原因在于该信息系统的较小辨识集恰好是该图的邻接集. 基于此,本节从图的邻接集和信息系统的辨识集之间的关系出发,构造图的一个信息系统,并进一步讨论图的极小支配集和所构造信息系统的属性约简之间的关系.

给定一个图 $G=(V,E)$,$V=\{v_1,v_2,\cdots,v_m\}$,当且仅当 $\boldsymbol{M}_G=(m_{ij})_{m\times m}$ 满足:

1) $m_{ij}=1$,若 $v_j\in N[v_i]$,

2) $m_{ij}=0$,若 $v_j\notin N[v_i]$,

则称 $\boldsymbol{M}_G$ 为图的邻接矩阵,显然,$\boldsymbol{M}_G$ 是对角线为 1 的对称方阵.

**定义 6** 给定一个图 $G=(V,E)$,$V=\{v_1,v_2,\cdots,v_m\}$,$\boldsymbol{M}_G=(m_{ij})_{m\times m}$ 是图的邻接矩阵,当且仅当 $IS_G$ 满足:

1) $U=\{x_1,x_2,\cdots,x_m,x_{m+1}\}$,

2) $A=\{v_1,v_2,\cdots,v_m\}$,

3) 信息函数为

$$f(x_i,v_j)=\begin{cases} i\cdot m_{ij}, & 1\leqslant i\leqslant m,1\leqslant j\leqslant m \\ -0, & i=m+1,1\leqslant j\leqslant m \end{cases}$$

则称信息系统 $IS_G=(U,A,I,f)$ 为图 $G$ 的诱导信息系统.

在定义 6 中,$x_1,x_2,\cdots,x_m$ 分别对应顶点 $v_1,v_2,\cdots,v_m$,而 $v_{m+1}$ 为新增添的对象,其作用可在下文的分析中体现.

**例 3** 给定一个图 $G=(V,E)$ 如例 2 所示. 根据定义 6,给出该图的邻接矩阵和诱导信息系统.

图的顶点邻接矩阵如表 2 所示.

**表 2 邻接矩阵 $\boldsymbol{M}_G$**

Table 2 Adjacency matrix $\boldsymbol{M}_G$

| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ |
|---|---|---|---|---|---|
| $v_1$ | 1 | 1 | 1 | 1 | 0 |
| $v_2$ | 1 | 1 | 1 | 0 | 0 |
| $v_3$ | 1 | 1 | 1 | 0 | 0 |
| $v_4$ | 1 | 0 | 0 | 1 | 0 |
| $v_5$ | 0 | 0 | 0 | 0 | 1 |

图的诱导信息系统如表 3 所示.

**表 3 诱导信息系统 $IS_G$**

Table 3 Induced information system $IS_G$

| | $v_1$ | $v_2$ | $v_3$ | $v_4$ | $v_5$ |
|---|---|---|---|---|---|
| $x_1$ | 1 | 1 | 1 | 1 | 0 |
| $x_2$ | 2 | 2 | 2 | 0 | 0 |
| $x_3$ | 3 | 3 | 3 | 0 | 0 |
| $x_4$ | 4 | 0 | 0 | 4 | 0 |
| $x_5$ | 0 | 0 | 0 | 0 | 5 |
| $x_6$ | 0 | 0 | 0 | 0 | 0 |

**性质 1** 给定一个图 $G=(V,E)$,

$$V=\{v_1,v_2,\cdots,v_m\},$$

$IS_G=(U,A,I,f)$ 为该图的诱导信息系统. 则任意对象对$(x_i,x_j)\in U\times U$ 的辨识集,满足

$$d(x_i,x_j)=\begin{cases} N[v_i]\cup N[v_j], & 1\leqslant i\leqslant m,1\leqslant j\leqslant m \\ N[v_i], & 1\leqslant i\leqslant m,j=m+1 \\ \varnothing, & 1\leqslant i\leqslant m+1,i=j \end{cases}$$

**证明** 对 $\forall v_k\in A$,$1\leqslant i\leqslant m$,$IS_G$ 的信息函数满足

$$f(x_i,v_k)=\begin{cases} i, & v_k\in N[v_i] \\ 0, & v_k\notin N[v_i] \end{cases}$$

1) 若 $1\leqslant i\leqslant m$,$1\leqslant j\leqslant m$,则

$$f(x_i,v_k)\neq f(x_j,v_k)\Leftrightarrow v_k\in(N[v_i]\cup N[v_j]),$$

因此有

$$d(x_i,x_j)=N[v_i]\cup N[v_j].$$

2) 若 $j=m+1$,则信息函数满足: 对 $\forall v_k\in A$,有 $f(x_j,v_k)=0$.

因此对 $1\leqslant i\leqslant m$, $j=m+1$,有

$$\begin{aligned} d(x_i,x_{m+1}) &=\{v_k\in A\mid f(x_i,v_k)\neq f(x_{m+1},v_k)\} \\ &=\{v_k\in A\mid f(x_i,v_k)\neq 0\} \\ &=N[v_i]. \end{aligned}$$

3) 若 $1\leqslant i\leqslant m+1$,$i=j$,显然 $d(x_i,x_j)=\varnothing$.

基于性质 1,可得到图的支配集和其诱导信息系统的约简之间的关系.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

**定理 3** 给定一个图 $G=(V,E)$,

$$V=\{v_1,v_2,\cdots,v_m\},$$

$IS_G=(U,A,I,f)$ 为图 $G$ 的诱导信息系统,则 $S\subseteq V$ 是图 $G$ 的极小支配集,当且仅当 $S$ 是信息系统 $IS_G$ 的约简.

**证明** 由定理 1 和定理 2 可知,只需证明

$$f_D(v_1^*,v_2^*,\cdots,v_m^*)=f_G(v_1^*,v_2^*,\cdots,v_m^*),$$

其中

$$f_D(v_1^*,v_2^*,\cdots,v_m^*)=\wedge\{\vee d(x_i,x_j)\mid d(x_i,x_j)\neq\varnothing,x_i\in U,x_j\in U\},$$

$$f_G(v_1^*,v_2^*,\cdots,v_m^*)=\wedge\{\vee N[v_i]\mid v_i\in V,1\leqslant i\leqslant m\}.$$

由性质 1,对 $\forall 1\leqslant i\leqslant m$,有

$$d(x_i,x_{m+1})=N[v_i];$$

对 $\forall 1\leqslant i\leqslant m,1\leqslant j\leqslant m$,有

$$d(x_i,x_j)=d(x_j,x_i)=N[v_i]\cup N[v_j].$$

因此

$$\begin{aligned}
&f_D(v_1^*,v_2^*,\cdots,v_m^*)\\
&=\wedge\{\vee d(x_i,x_j)\mid d(x_i,x_j)\neq\varnothing,x_i\in U,x_j\in U\}\\
&=\wedge\{\vee d(x_i,x_j)\mid d(x_i,x_j)\neq\varnothing,x_i\in U,x_j\in U,\\
&\quad 1\leqslant i<j\leqslant m+1\}\\
&=\wedge\{\vee d(x_i,x_{m+1})\mid d(x_i,x_{m+1})\neq\varnothing,x_i\in U,\\
&\quad 1\leqslant i\leqslant m\}\wedge\{\vee d(x_i,x_j)\mid d(x_i,x_j)\neq\varnothing,\\
&\quad x_i\in U,x_j\in U,1\leqslant i<j\leqslant m\}\\
&=\wedge\{\vee N[v_i]\mid v_i\in V,1\leqslant i\leqslant m\}\wedge\\
&\quad\{\vee(N[v_i]\cup N[v_j])\mid v_i\in V,v_j\in V,\\
&\quad 1\leqslant i<j\leqslant m\}\\
&=\wedge\{\vee N[v_i]\mid v_i\in V,1\leqslant i\leqslant m\}\\
&=f_G(v_1^*,v_2^*,\cdots,v_m^*).
\end{aligned}$$

**推论 1** 给定一个图 $G=(V,E)$,

$$V=\{v_1,v_2,\cdots,v_m\},$$

$IS_G=(U,A,I,f)$ 为图 $G$ 的诱导信息系统,则 $S\subseteq V$ 是图 $G$ 的最小支配集,当且仅当 $S$ 是 $IS_G$ 的最小约简.

**推论 2** 给定一个图 $G=(V,E)$,

$$V=\{v_1,v_2,\cdots,v_m\},$$

$IS_G=(U,A,I,f)$ 为图 $G$ 的诱导信息系统,$v\in A$ 是 $IS_G$ 的核心属性,当且仅当 $v$ 是 $G$ 的孤立点.

**证明** $v$ 是孤立点等价于 $N[v]=\{v\}$,由引理 1 和性质 1 可证.

# 4 基于信息熵的最小支配集算法

通过上述分析,得出图的最小支配集恰是其诱导信息系统的最小约简,因此信息系统的属性约简算法可用来求解图的最小支配集. 特别地,引入一个基于信息熵的约简算法.

为解决信息的度量问题,Shannon 提出信息熵的概念,Ślezak [13] 将这种信息熵用于寻找信息系统的属性约简.

**定义 7** [13] 给定一信息系统 $IS=(U,A,I,f)$,Shannon 熵定义为

$$H(B)=-\sum_{i=1}^{r}\frac{|X_i|}{|U|}\log_2\left(\frac{|X_i|}{|U|}\right),$$

其中

$$U/IND(B)=\{X_1,X_2,\cdots,X_r\},$$

$|\cdot|$ 表示集合的基数.

Shannon 熵可度量属性的重要度,还可用来构建信息系统的约简算法. 对一个属性 $a\notin B\subseteq A$,$a$ 相对于属性集 $B$ 的重要度定义为 [13]

$$Sig(a,B)=H\{B\cup\{a\}\}-H(B).$$

属性 $a$ 是核心属性当且仅当 [13]

$$Sig(a,A-\{a\})>0.$$

**推论 3** 给定一个图 $G=(V,E)$,

$$V=\{v_1,v_2,\cdots,v_m\},$$

$IS_G=(U,A,I,f)$ 是图 $G$ 的诱导信息系统,则

$$H(A)=\log_2(|V|+1).$$

**证明** 由性质 1,对 $x_i\in U,x_j\in U$,若 $x_i\neq x_j$,则辨识集满足 $d(x_i,x_j)\neq\varnothing$. 另外由引理 1,对 $\forall x_i\in U$,有 $[x_i]_A=\{x_i\}$,则

$$U/IND(A)=\{\{x_1\},\{x_2\},\cdots,\{x_m\},\{x_{m+1}\}\},$$

$$\begin{aligned}
H(B)&=-\sum_{i=1}^{|U|}\frac{1}{|U|}\log_2\left(\frac{1}{|U|}\right)=\log_2|U|\\
&=\log_2(|V|+1).
\end{aligned}$$

结合定理 3、定理 4 及文献 [13] 中的约简方法,可构建图最小支配集的算法.

**算法** 基于信息熵的图最小支配集算法

**输入** 图 $G=(V,E)$

**输出** $G$ 的一个支配集

step 1 令 $Red$ 为 $G$ 的所有孤立点构成的集合;

step 2 计算 $G$ 的诱导信息系统

$$IS_G=(U,A,I,f);$$

step 3

While $H(Red)\neq H(A)$

$\{Red\leftarrow Red\cup\{v_0\}$,其中 $v_0$ 满足

$Sig(v_0,Red)=\max(Sig(v,Red),v\in A-Red)$

$\}$

step 4 返回一个支配集 $Red$.

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

step 1 的时间复杂度为 $O(|V|)$. step 2 的时间复杂度为 $O(|V|^2)$. 由文献[7]知,step 3 的时间复杂度为

$$O(|U|^2|A|) = O(|V|^3).$$

因此算法 1 的时间复杂度为 $O(|V|^3)$.

**例 4** 图 $G = (V,E)$ 如例 3 所示,利用算法 1 求解该图的支配集,过程如下.

step 1 该图的孤立点为 $v_5$,则令 $Red = \{v_5\}$.

step 2 该图的诱导信息系统 $IS_G = (U,A,I,f)$ 如例 3 所示.

step 3 计算得

$$H(\{v_5\}) = -\left(\frac{5}{6}\log_2\frac{5}{6} + \frac{1}{6}\log_2\frac{1}{6}\right), H(A) = \log_2 6,$$

则有 $H(\{v_5\}) \neq H(A)$,并且

$$H(\{v_5,v_1\}) = \log_2 6;$$

$$H(\{v_5,v_2\}) = -\left(\frac{4}{6}\log_2\frac{4}{6} + \frac{2}{6}\log_2\frac{2}{6}\right),$$

$$H(\{v_5,v_3\}) = -\left(\frac{4}{6}\log_2\frac{4}{6} + \frac{2}{6}\log_2\frac{2}{6}\right),$$

$$H(\{v_5,v_4\}) = \log_2 2.$$

则有

$$Sig(v_1,Red) = \max(Sig(v,Red), v \in A - Red),$$

因此,$Red = \{v_1,v_5\}$. 此时 $H(Red) = H(A)$,执行 step 4.

step 4 返回该图的一个支配集 $Red = \{v_1,v_5\}$.

# 5 结束语

本文探讨粗糙集的属性约简问题和图的支配集问题之间的联系. 通过构造图的一个信息系统,将图的支配集问题转化为粗糙集的属性约简问题. 这项研究使众多的粗糙集约简问题的理论和方法用于求解图支配集问题成为可能,另一方面也实现粗糙集方法在组合优化问题中的一个重要应用.

本文对粗糙集的属性约简问题和图的支配集问题组织间联系的探讨,现在只局限于理论的研究和方法的引进,怎样才能更高效地利用粗糙集的约简算法仍是值得进一步研究的课题.

## 参 考 文 献

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

[2] Haynes T W, Hedetniemi S T, Slater P J. Fundamentals of Domination in Graphs. New York, USA: Marcel Dekker, 1998

[3] Li J J. Topological Methods on the Theory of Covering Generalized Rough Sets. Pattern Recognition and Artificial Intelligence, 2004, 17(1): 7-10 (in Chinese)
(李进金. 覆盖广义粗集理论中的拓扑学方法. 模式识别与人工智能, 2004, 17(1): 7-10)

[4] Liu J N K, Hua Y X, He Y L. A Set Covering Based Approach to Find the Reduct of Variable Precision Rough Set. Information Sciences, 2014, 275: 83-100

[5] Chen C Y, Li Z G. A Study of Reduction of Attributes and Set Covering Problem. Computer Engineering and Applications, 2004, (2): 44-46, 84 (in Chinese)
(陈彩云,李治国. 关于属性约简和集合覆盖问题的探讨. 计算机工程与应用, 2004, (2): 44-46, 84)

[6] Chen J K, Li J J. An Application of Rough Sets to Graph Theory. Information Sciences, 2012, 201: 114-127

[7] Lu G, Zhou M T, Tang Y, *et al*. A Survey on Exact Algorithms for Dominating Set Related Problems in Arbitrary Graphs. Chinese Journal of Computers, 2010, 33(6): 1073-1087 (in Chinese)
(路 纲,周明天,唐 勇,等. 任意图支配集精确算法回顾. 计算机学报, 2010, 33(6): 1073-1087)

[8] Yu H, Yang D C. Approach to Solving Attribute Reductions with Ant Colony Optimization. Pattern Recognition and Artificial Intelligence, 2011, 24(2): 176-184 (in Chinese)
(于 洪,杨大春. 基于蚁群优化的多个属性约简的求解方法. 模式识别与人工智能, 2011, 24(2): 176-184)

[9] Liang J Y, Shi Z Z. The Information Entropy, Rough Entropy and Knowledge Granulation in Rough Set Theory. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2004, 12(1): 37-46

[10] Qian Y H, Liang J Y, Pedrycz W, *et al*. Positive Approximation: An Accelerator for Attribute Reduction in Rough Set Theory. Artificial Intelligence, 2010, 174(9/10): 597-618

[11] Miao D Q, Hu G R. A Heuristic Algorithm for Reduction of Knowledge. Journal of Computer Research & Development, 1999, 36(6): 681-684 (in Chinese)
(苗夺谦,胡桂荣. 知识约简的一种启发式算法. 计算机研究与发展, 1999, 36(6): 681-684)

[12] Skowron A, Rauszer C. The Discernibility Matrices and Functions in Information Systems // Słowiński R, ed. Intelligent Decision Support. Dordrecht, The Netherlands: Springer, 1992: 331-362

[13] Ślezak D. Approximate Entropy Reducts. Fundamenta Informaticae, 2002, 53(3/4): 365-390

[14] Yao Y Y, Zhao Y. Discernibility Matrix Simplification for Constructing Attribute Reducts. Information Sciences, 2009, 179(7): 867-882

[15] Wang G Y, Yu H, Yang D C. Decision Table Reduction Based on Conditional Information Entropy. Chinese Journal of Computers, 2002, 25(7): 759-766 (in Chinese)
(王国胤,于 洪,杨大春. 基于条件信息熵的决策表约简. 计算机学报, 2002, 25(7): 759-766)

?1994-2015 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net