

不完备信息系统中基于相似关系的知识约简^{*})

杨习贝¹ 於东军¹ 吴 陈^{1,2} 杨静宇¹

(南京理工大学计算机科学与技术学院 南京 210094)¹

(德瑞索大学信息科学与技术学院 费城 PA19104)²

摘 要 以具有丢失型未知属性值的不完备信息系统为研究对象,根据非对称相似关系,讨论了知识约简问题。在不完备决策系统中,引入了近似、粗糙分布约简以及广义决策约简,讨论了它们之间的相互关系,给出了近似分布约简的判定定理、可辨识矩阵以及约简公式,并进行了实例分析,为从不完备信息系统中获取知识提供了新的理论基础与操作手段。

关键词 不完备信息系统,非对称相似关系,近似分布约简,粗糙分布约简,广义决策约简

Knowledge Reductions in Incomplete Information Systems Based on Similarity Relation

YANG Xi-Bei¹ YU Dong-Jun¹ WU Chen^{1,2} YANG Jing-Yu¹

(School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094)

(College of Information Science and Technology, Drexel University, Philadelphia, PA 19104, USA)²

Abstract In this paper, the incomplete information system in which all unknown values are looked as lost is deeply investigated. In such incomplete information systems, approach to knowledge reduction based on similarity relation is studied. Moreover, the approximate and rough distribution reducts, generalized decision reduct are introduced into incomplete decision system, the judgment theorems, discernibility matrixes and reduction formulas with respect to approximate distribution reducts are obtained. Finally, an illustrative example is analyzed. These results are meaningful both in the theory and in applications for rules' acquisition in incomplete decision systems.

Keywords Incomplete information system, Non-symmetric similarity relation, Approximate distribution reduction, Rough distribution reduct, Generalized decision reduction

1 引言

粗糙集理论^[1,2]的主要目的是在不损失有用信息的基础上简化信息(决策)系统,亦称为知识约简。Pawlak 提出的经典粗糙集是用来处理具有离散属性值的完备系统的,因此他使用不可分辨关系(等价关系)进行分类。在此基础上,通过研究条件属性划分与决策属性划分之间的关系,已经有一些学者给出了相应的知识约简的概念。如 Kryszkiewicz^[3]给出了 5 种不同的知识约简形式,张文修^[4]等人提出了(最大)分配约简、分布约简(上近似约简),米据生^[5]基于可变精度粗糙模型给出了上、下近似分布约简的概念,管涛^[6]在模糊目标信息系统中推广了几种知识约简形式。

然而,以上的所有约简都是用来处理完备决策系统的。但在现实世界中,由于众多原因,所面临的信息系统往往是不完备的^[7],因此进行不完备信息(决策)系统中知识约简方法的研究是有其实际意义的。目前,在不完备系统中的知识约简研究主要有以下成果:Kryszkiewicz^[8]根据容差关系给出了广义决策约简的判别矩阵方法;周献中^[9]将分布、最大分布、分配约简引入到不完备决策系统中,并提出了分配序约简;管延勇^[10]等人使用最大全相容类技术^[11]用于获取不完备信息系统中的最优可信规则;吴伟志等人研究了不完备模糊信息系统^[12]中的知识约简方法。但不难发现,所有这些工作都是

假设信息系统中的未知属性值仅仅是被遗漏的,但又是确实存在的^[13],也就是说所使用的粗糙模型是建立在容差关系(自反性、对称性)的基础上的。然而,对于不完备信息系统中未知属性值的语义解释还有另外一种情况,即所有的未知属性值是被认为是丢失的,是不允许被比较的^[14]。针对这种情况,Stefanowski^[15]等人构建了相似关系(自反性、传递性)。遗憾的是,目前对于具有丢失型未知属性值的不完备信息系统(决策)中的知识约简方法的研究尚不多见,故笔者将此作为本文的研究内容。

在经典粗糙模型中,下、上近似分布约简是保持所有决策类的下、上近似集不变化的最小属性子集。一般来说,下近似分布约简可用于获取这样的规则:“若 $y \in [x]_B$, 则必定 $y \in D_j$ ”,其中 $[x]_B$ 是对象 x 基于属性集合 B 的等价类, D_j 是某个决策类,这是确定性规则;上近似分布约简可用于获取这样的规则:“若 $y \in [x]_B$, 则 y 可能属于 D_j ”, y 属于 D_j 的可信度由粗糙隶属度决定^[4]。近似分布约简的最大优点在于由约简系统产生的命题规则与由原系统产生的命题规则是相容的,即约简不会改变由对象所产生的规则的决策结果^[5]。因此,笔者将下、上近似分布约简、粗糙约简、广义决策约简引入到具有丢失型未知属性值的不完备决策系统中,讨论了它们之间的关系,并且给出了近似分布约简的判定定理和可辨识属性矩阵。

^{*})国家自然科学基金(No. 60472060, 60572034)、国家自然科学基金重点项目(No. 60632050)、江苏省自然科学基金(No. BK2006081)、南京理工大学科研发展基金(No. AB96125)资助项目。杨习贝 博士研究生,研究方向为智能信息处理、粗糙集理论。

综上,本文的主要内容安排如下:第2节简要介绍了不完备信息系统,研究了其中基于相似关系的知识约简方法,第3节主要讨论了不完备决策系统中的几种知识约简形式,第4节进行实例分析,最后总结全文。

2 不完备信息系统与知识约简

一个不完备信息系统为四元组: $IIS = \langle U, AT, V, f \rangle$ 。其中, U 是一个被称为论域的非空有限的对象集合; AT 是非空有限的属性集合; 对于 $\forall a \in AT$, 有 $a: U \rightarrow V_a$, V_a 是属性 a 的值域(可包丢失型空值,文中用“?”表示),全体属性值域集合 $V = \bigcup_{a \in AT} V_a$; 定义 f 为信息函数,对于 $\forall a \in AT, \forall x \in U$, 有 $f(x, a) \in V_a$ 。

定义1 在不完备信息系统 IIS 中,对于 $\forall A \subseteq AT$, 由 A 决定的相似关系记为 $S(A)$ 且

$$S(A) = \{(x, y) \in U^2 : \forall a \in A, f(x, a) = ? \vee f(y, a) = f(x, a)\} \quad (2.1)$$

很明显,定义1中所示的相似关系是自反和传递的,但并不一定是对称的,因此 Stefanowski 给出了非对称相似于 x 的集合 $S_A(x) = \{y \in U : (y, x) \in S(A)\}$, x 与之非对称相似的集合 $S_A^{-1}(x) = \{y \in U : (x, y) \in S(A)\}$ 。

定理1 在不完备信息系统 IIS 中,对于 $\forall A \subseteq AT$, 有 $S(AT) \subseteq S(A)$; (2.2)

$$y \in S_A^{-1}(x) \Leftrightarrow S_A^{-1}(y) \subseteq S_A^{-1}(x); \quad (2.3)$$

$$S_A^{-1}(x) = \bigcup \{S_{AT}^{-1}(y) : y \in S_A^{-1}(x)\}. \quad (2.4)$$

证明:式(2.2)显然成立,故只需证式(2.3)与(2.4)。

$S_A^{-1}(y) \subseteq S_A^{-1}(x) \Rightarrow y \in S_A^{-1}(x)$ 显然成立,故只需证 $y \in S_A^{-1}(x) \Rightarrow S_A^{-1}(y) \subseteq S_A^{-1}(x)$ 即可。对于 $\forall z \in S_A^{-1}(y)$, 有 $(y, z) \in S(A)$, 又因为 $y \in S_A^{-1}(x)$, 所以 $(x, y) \in S(A)$ 。由于 $S(A)$ 具有传递性,所以 $(x, z) \in S(A)$, 即 $z \in S_A^{-1}(x)$, 式(2.3)证毕。

由于 $S(A)$ 具有自反性,因此很明显 $S_A^{-1}(x) \subseteq \bigcup \{S_{AT}^{-1}(y) : y \in S_A^{-1}(x)\}$ 成立。对于 $\forall S_{AT}^{-1}(y) \in \bigcup \{S_{AT}^{-1}(y) : y \in S_A^{-1}(x)\}$, 因为 $y \in S_A^{-1}(x)$, 所以 $S_{AT}^{-1}(y) \subseteq S_A^{-1}(y) \subseteq S_A^{-1}(x)$, 即 $\bigcup \{S_{AT}^{-1}(y) : y \in S_A^{-1}(x)\} \subseteq S_A^{-1}(x)$ 。综上, $S_A^{-1}(x) = \bigcup \{S_{AT}^{-1}(y) : y \in S_A^{-1}(x)\}$, 式(2.4)证毕。

定义2 令 IIS 为一不完备信息系统,若 $A \subseteq AT$, 则 A 是 IIS 的约简当且仅当 $S(A) = S(AT)$ 且对于 $\forall B \subset A$, 都有 $S(A) \neq S(B)$ 。

为简便起见,在不完备信息系统 IIS 中,对于 $\forall x, y \in U$, 令 $D_{AT}(x, y) = \{a \in AT : (x, y) \notin S(a)\} = \{a \in M_{AT}(x) : f(x, a) \neq f(y, a) \vee f(y, a) = ?\}$, 其中 $M_{AT}(x) = \{a \in AT : f(x, a) \neq ?\}$ 。

定理2 令 IIS 为一不完备信息系统,若有 $A \subseteq AT$, 则 $S(A) = S(AT) \Leftrightarrow A \cap D_{AT}(x, y) \neq \emptyset (\forall D_{AT}(x, y)) \neq \emptyset$ 。

证明:(1)必要性:因为 $S(A) = S(AT)$, 所以就有 $S_A^{-1}(x) = S_{AT}^{-1}(x) (\forall x \in U)$, $y \notin S_A^{-1}(x) \Leftrightarrow y \notin S_{AT}^{-1}(x) (\forall y \in U)$, 即存在 $a \in M_A(x) \subseteq M_{AT}(x)$, 使得 $f(x, a) \neq f(y, a) \vee f(y, a) = ?$, 因此 $a \in D_{AT}(x, y)$ 。又由于 $a \in M_A(x) \subseteq A$, 因此 $A \cap D_{AT}(x, y) \neq \emptyset (\forall D_{AT}(x, y) \neq \emptyset)$ 。充分性: $A \subseteq AT \Rightarrow S(AT) \subseteq S(A)$, 故只需证 $S(AT) \supseteq S(A)$ 即可。对于 $\forall (x, y) \notin S(AT)$, 有 $D_{AT}(x, y) \neq \emptyset$, 又由于 $A \cap D_{AT}(x, y) \neq \emptyset$, 因此必定存在 $a \in A$ 使得 $f(x, a) \neq f(y, a) \vee f(y, a) = ?$, 即 $(x, y) \notin S(A)$, 所以 $(x, y) \in S(A) \Rightarrow (x, y) \in S(AT)$, 即 $S(AT) \supseteq S(A)$ 成立。

IIS 中的知识约简是保持非对称相似关系不变的最小属性子集。在 IIS 中,对于 $\forall x, y \in U$, 若 $(x, y) \in S(AT)$, 则 $D_{AT}(x, y) = \emptyset$ 。令 $\bigvee D_{AT}(x, y)$ 为布尔表达式,若 $D_{AT}(x, y) = \emptyset$, 则 $\bigvee D_{AT}(x, y) = 1$, 否则 $\bigvee D_{AT}(x, y)$ 是包含在 $D_{AT}(x, y)$ 中属性对应变量的析取。

令 $\Delta = \bigwedge_{(x, y) \in U \times U} \bigvee D_{AT}(x, y)$, $\Delta(x) = \bigwedge_{y \in U} \bigvee D_{AT}(x, y)$, Δ 称为 IIS 的分辨函数, $\Delta(x)$ 称为 IIS 中对象 x 的分辨函数,利用布尔代数的吸收率将分辨函数化为析取范式,就可以得到 IIS 的约简。

3 不完备决策系统中的知识约简

3.1 知识约简概念

不完备决策系统是不完备信息系统一种特殊形式。记不完备决策系统 $IDS = \langle U, AT \cup D, V, f \rangle$, 其中 AT 是非空有限的条件属性集合, D 为非空有限的决策属性集合且 $AT \cap D = \emptyset$ 。由于专家在信息不完备的情况下也能给出决策,因此 V_D 中没有未知属性值, $U/D = \{D_1, D_2, \dots, D_m\}$ 表示论域上的划分,即 $\bigcup_{1 \leq i \leq m} D_i = U$ 且对于 $\forall D_i, D_j \in U/D$, 有 $D_i \cap D_j = \emptyset$ 。

定义3 设 IDS 为一不完备决策系统,其中 $A \subseteq AT$, 对于 $\forall D_i \in U/D$, D_i 基于 $S(A)$ 的下、上近似集分别记为 $A_*(D_i)$, $A^*(D_i)$ 且

$$A_*(D_i) = \{x \in U : S_A^{-1}(x) \subseteq D_i\}; A^*(D_i) = \bigcup_{y \in D_i} \{S_A(y) : y \in D_i\} \quad (3.1)$$

定理3 设 IDS 为一不完备决策系统, $A \subseteq AT$, 对于 $\forall D_i \in U/D$, 有 $A^*(D_i) = \{x \in U : S_A^{-1}(x) \cap D_i \neq \emptyset\} = \bigcup \{S_A(y) : y \in D_i\}$ 。

证明:对于 $\forall x \in U$, 若 $S_A^{-1}(x) \cap D_i \neq \emptyset$, 则必定存在 $y \in D_i$ 且 $y \in S_A^{-1}(x)$, 即 $(x, y) \in S(A)$, 因此 $y \in D_i$ 且 $x \in S_A(y)$, $x \in \bigcup \{S_A(y) : y \in D_i\}$ 成立。对于 $\forall x \in \bigcup \{S_A(y) : y \in D_i\}$, 必定存在 $y \in D_i$ 且 $x \in S_A(y)$, 因此 $y \in S_A^{-1}(x)$ 成立,即 $S_A^{-1}(x) \cap D_i \neq \emptyset$ 。

定义4 设 IDS 为一不完备决策系统,若 $S(AT) \subseteq R$ (D), 则称 IDS 是协调的,否则就是不协调的,其中 $R(D)$ 为等价关系,即 $R(D) = \{(x, y) \in U^2 : \forall d \in D, f(x, d) = f(y, d)\}$ 。

定义5 设 IDS 为一不完备决策系统,其中 $A \subseteq AT$, 记 $L_A = \{A_*(D_1), A_*(D_2), \dots, A_*(D_m)\}$ (3.2)

$$H_A = \{A^*(D_1), A^*(D_2), \dots, A^*(D_m)\} \quad (3.3)$$

$$\partial_A(x) = \{D_i : S_A^{-1}(x) \cap D_i \neq \emptyset\} \quad (3.4)$$

(1)若 $L_A = L_{AT}$, 则称 A 是下近似分布协调集。若 A 是下近似分布协调集,且 A 的任意真子集都不是下近似分布协调集,则称 A 是下近似分布约简。

(2)若 $H_A = H_{AT}$, 则称 A 是上近似分布协调集。若 A 是上近似分布协调集,且 A 的任意真子集都不是上近似分布协调集,则称 A 是上近似分布约简。

(3)若对于 $\forall x \in U$, 有 $\partial_A(x) = \partial_{AT}(x)$, 则称 A 是广义决策协调集。若 A 是广义决策协调集,且 A 的任意真子集都不是广义决策协调集,则称 A 是广义决策约简。

(4)若 $L_A = L_{AT}$ 且 $H_A = H_{AT}$, 则称 A 是粗糙分布协调集。若 A 是粗糙分布协调集,且 A 的任意真子集都不是粗糙分布协调集,则称 A 是粗糙分布约简。

定理4 令 IDS 为一不完备决策系统,其中 $B, A \subseteq AT$,

若 B 是下近似分布约简, A 是上近似分布约简, 则 $A \cup B$ 为粗糙近似分布约简。

证明: 因为 $B \subseteq A \cup B \subseteq AT$, 所以对于 $\forall D_i \in U/D$, 有 $B \cdot (D_i) \subseteq (A \cup B) \cdot (D_i) \subseteq AT \cdot (D_i)$ 。又因为 B 是下近似分布约简, 即 $LB = L_{AT}$, 所以 $L_{A \cup B} = L_{AT}$ 。同理可证 $H_{A \cup B} = H_{AT}$ 。此时, 若在 $A \cup B$ 中去掉一个元素形成 $(A \cup B)'$, 则对于 $\forall C \subseteq (A \cup B)'$ 必定有 $L_C \neq L_{AT}$ 或者 $H_C \neq H_{AT}$, 因此 $A \cup B$ 为粗糙分布约简。

下近似分布约简是保持所有决策类的下近似都不发生变化的最小属性子集, 上近似分布约简是保持所有决策类的上近似都不发生变化的最小属性子集, 而粗糙分布约简则是保持所有决策类的下、上近似集同时都不发生变化的最小属性子集。

定理 5 设 IDS 为一不完备决策系统, 其中 $A \subseteq AT$, 对于 $\forall x \in U$, 记

$$P_A(x) = \{D_i : x \in A \cdot (D_i)\}, Q_A(x) = \{D_i : x \in A^* \cdot (D_i)\} \quad (3.5)$$

于是就有

(1) A 是下近似分布协调集 $\Leftrightarrow$ 对于 $\forall x \in U$, 有 $P_A(x) = P_{AT}(x)$;

(2) A 是上近似分布协调集 $\Leftrightarrow$ 对于 $\forall x \in U$, 有 $Q_A(x) = Q_{AT}(x)$ 。

定理 6 设 IDS 为一不完备决策系统, 其中 $A \subseteq AT$, 则 A 是广义决策协调集; $\Leftrightarrow A$ 是上近似分布协调集; $\Rightarrow A$ 是下近似分布协调集。

证明: 很明显 $\partial_A(x) = \partial_{AT}(x) \Leftrightarrow Q_A(x) = Q_{AT}(x)$ 成立, 所以就有 A 是广义决策协调集; $\Leftrightarrow A$ 是上近似分布协调集。设 A 是上近似分布协调集, 则对于 $X_i \in U/D$, $S_A^{-1}(x) \cap D_i \neq \emptyset \Leftrightarrow S_{AT}^{-1}(x) \cap D_i \neq \emptyset$, 又由于 U/D 表示论域上的划分, 因此就有 $S_A^{-1}(x) \subseteq D_i \Leftrightarrow S_{AT}^{-1}(x) \subseteq D_i$, 即 A 是下近似分布协调集。

定理 7 设 IDS 为一不完备决策系统, 若 $A \subseteq AT$, 则 A 是广义决策约简; $\Leftrightarrow A$ 是上近似分布约简。

3.2 知识约简方法

定义 6 设 IDS 为一不完备决策系统, 其中 $A \subseteq AT$, 记 $D_1 = \{(x, y) : \forall X_i \in U/D, \forall x \in AT \cdot (X_i), \forall y \in U - X_i\}$ (3.6)

$D_2 = \{(y, x) : \forall X_i \in U/D, \forall x \in X_i, \forall y \in U - AT^* \cdot (X_i)\}$ (3.7)

定义

$$D_1(x, y) = \begin{cases} a \in AT : (x, y) \notin S(a), (x, y) \in D_1 \\ AT, & (x, y) \notin D_1 \end{cases}$$

$$D_2(y, x) = \begin{cases} a \in AT : (y, x) \notin S(a), (y, x) \in D_2 \\ AT, & (y, x) \notin D_2 \end{cases}$$

为 IDS 的下(上)近似分布的可辨识属性矩阵。

定理 8 令 IDS 为一不完备决策系统, 其中 $A \subseteq AT$, 则

(1) $L_A = L_{AT} \Leftrightarrow$ 对于 $\forall (x, y) \in D_1, A \cap D_1(x, y) \neq \emptyset$,

(2) $H_A = H_{AT} \Leftrightarrow$ 对于 $\forall (y, x) \in D_2, A \cap D_2(y, x) \neq \emptyset$

证明: (1) 必要性: 假设 $\exists x \in AT^* \cdot (X_i), \exists y \in U - X_i$ 使得 $A \cap D_1(x, y) = \emptyset$, 于是就有 $(x, y) \in S(A), y \in S_A^{-1}(x)$, 又由于 $L_A = L_{AT}$, 因此 $AT \cdot (X_i) = A \cdot (X_i)$, 即 $S_{AT}^{-1}(x) \subseteq X_i \Leftrightarrow S_A^{-1}(x) \subseteq X_i$, 于是得到 $y \in X_i$, 这与 $y \in U - X_i$ 矛盾。充分性: $A \subseteq AT \Rightarrow AT \cdot (X_i) \supseteq A \cdot (X_i)$, 故只需证 $AT \cdot (X_i) \subseteq A \cdot (X_i)$ 即可。对于 $\forall x \in AT \cdot (X_i)$, 有 $S_{AT}^{-1}(x) \subseteq X_i$, 对于 $\forall y \in U - X_i$ 有 $(x, y) \notin S(AT)$, 又因为 $A \cap D_1(x, y) \neq \emptyset$, 所以必

定存在 $a \in A$ 使得 $(x, y) \notin S(a)$, 即 $(x, y) \notin S(A)$, 所以就有 $(x, y) \in S(A) \Rightarrow (x, y) \in S(AT), S_A^{-1}(x) \subseteq S_{AT}^{-1}(x), AT \cdot (X_i) \subseteq A \cdot (X_i)$ 。

(2) 必要性: 假设 $\exists x \in X_i, \exists y \in U - AT^* \cdot (X_i)$ 使得 $A \cap D_2(y, x) = \emptyset$, 于是就有 $(y, x) \in S(A), x \in S_A^{-1}(y)$, 又由于 $H_A = H_{AT}$, 因此 $AT^* \cdot (X_i) = A^* \cdot (X_i)$, 即 $S_{AT}^{-1}(y) \cap X_i = \emptyset \Leftrightarrow S_A^{-1}(y) \cap X_i = \emptyset$, 于是得到 $x \notin X_i$, 这与 $x \in X_i$ 矛盾。充分性: $A \subseteq AT \Rightarrow AT^* \cdot (X_i) \subseteq A^* \cdot (X_i)$, 故只需证 $AT^* \cdot (X_i) \supseteq A^* \cdot (X_i)$ 即可。对于 $\forall y \in U - AT^* \cdot (X_i)$, 有 $S_{AT}^{-1}(y) \cap X_i = \emptyset$, 对于 $\forall x \in X_i$, 有 $(y, x) \notin S(AT)$, 又因为 $A \cap D_2(y, x) \neq \emptyset$, 所以必定存在 $a \in A$ 使得 $(y, x) \notin S(a)$, 即 $(y, x) \notin S(A)$, 所以 $(y, x) \in S(A) \Rightarrow (y, x) \in R(AT)$, 即 $S_A^{-1}(y) \subseteq S_{AT}^{-1}(y), AT^* \cdot (X_i) \supseteq A^* \cdot (X_i)$ 。

定义 7 设 IDS 为一不完备决策系统, $D_1(x, y), D_2(y, x)$ 分别为下、上近似分布的可辨识属性矩阵, 则称 $F_1 = \bigwedge \{ \bigvee \{ a : a \in D_1(x, y) \} : (\forall D_1(x, y) \neq \emptyset) \}, F_2 = \bigwedge \{ \bigvee \{ a : a \in D_2(y, x) \} : (\forall D_2(y, x) \neq \emptyset) \}$ 为下、上近似分布的辨识公式。

定理 9 设 IDS 为一不完备决策系统, 辨识公式 $F_l (l=1, 2)$ 的极小析取范式为 $F_l = \bigvee_{k=1}^t (\bigwedge_{s=1}^{q_k} a_{ls})$, 记 $B_k = \{a_{ls} : s=1, 2, \dots, q_k\}$, 则 $\{B_k : k=1, 2, \dots, t\}$ 分别是 IDS 中下、上近似分布约简的全体。

证明: 类似于文[4]中的证明过程。

4 实例分析

表 1 不完备决策系统

U	a	b	c	d	e
O_1	3	2	1	0	Φ
O_2	2	3	2	0	Φ
O_3	2	3	2	0	Ψ
O_4	?	2	?	1	Φ
O_5	?	2	?	1	Ψ
O_6	2	3	2	1	Ψ
O_7	3	?	?	3	Φ
O_8	?	0	0	?	Ψ
O_9	3	2	1	3	Ψ
O_{10}	1	?	?	?	Φ
O_{11}	?	2	?	?	Ψ
O_{12}	3	2	1	?	Φ

表 1 是文[15]用来分析的一个不完备决策系统, 其中 $AT = \{a, b, c, d\}$ 为条件属性集合, 决策属性为 e 且 $U/e = \{D_\Phi, D_\Psi\} = \{\{O_1, O_2, O_4, O_7, O_{10}, O_{12}\}, \{O_3, O_5, O_6, O_8, O_9, O_{11}\}\}$, 因此就有 $A * (D_\Phi) = \{O_1, O_{10}\}, A * (D_\Phi) = \{O_1, O_2, O_3, O_4, O_5, O_7, O_{10}, O_{11}, O_{12}\}, A * (D_\Psi) = \{O_6, O_8, O_9\}, A * (D_\Psi) = \{O_2, O_3, O_4, O_5, O_6, O_7, O_8, O_9, O_{11}, O_{12}\}$ 。

根据定义 6, 可以得到 $D_1 = \{(O_1, O_3), (O_1, O_5), (O_1, O_6), (O_1, O_8), (O_1, O_9), (O_1, O_{11}), (O_{10}, O_3), (O_{10}, O_5), (O_{10}, O_6), (O_{10}, O_8), (O_{10}, O_9), (O_{10}, O_{11}), (O_6, O_1), (O_6, O_2), (O_6, O_4), (O_6, O_7), (O_6, O_{10}), (O_6, O_{12}), (O_8, O_1), (O_8, O_2), (O_8, O_4), (O_8, O_7), (O_8, O_{10}), (O_8, O_{12}), (O_9, O_1), (O_9, O_2), (O_9, O_4), (O_9, O_7), (O_9, O_{10}), (O_9, O_{12})\}$ 。

$D_2 = \{(O_6, O_1), (O_6, O_2), (O_6, O_4), (O_6, O_7), (O_6, O_{10}), (O_6, O_{12}), (O_8, O_1), (O_8, O_2), (O_8, O_4), (O_8, O_7), (O_8, O_{10}),$

(下转第 177 页)

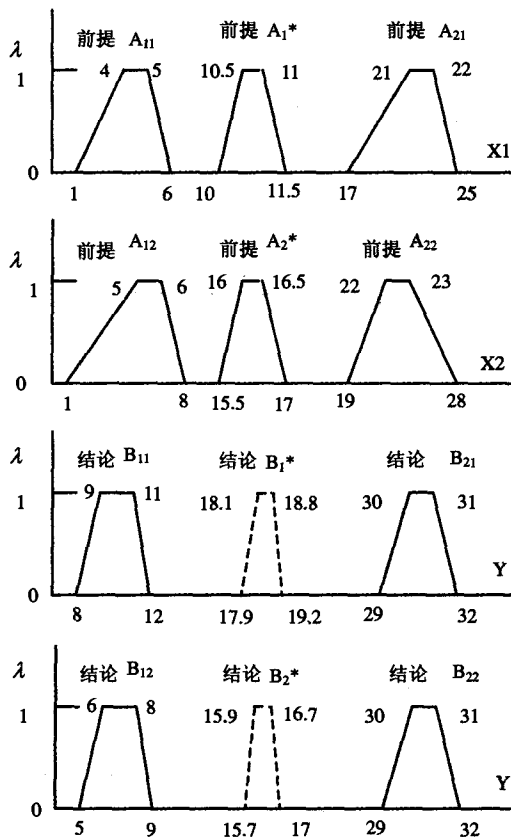


图5 本文方法推理结果

(上接第165页)

$(O_8, O_{12}), (O_9, O_1), (O_9, O_2), (O_9, O_4), (O_9, O_7), (O_9, O_{10}), (O_9, O_{12}), (O_1, O_3), (O_1, O_5), (O_1, O_6), (O_1, O_8), (O_1, O_9), (O_1, O_{11}), (O_{10}, O_3), (O_{10}, O_5), (O_{10}, O_6), (O_{10}, O_8), (O_{10}, O_9), (O_{10}, O_{11})$ 。

根据定义7可以求得 $F_1 = F_2 = a \wedge b \wedge d$, 由定理9可知 $\{a, b, d\}$ 是表1所示不完备决策系统的下近似分布约简和上近似分布约简。根据定理4, $\{a, b, d\}$ 也是表1的粗糙分布约简。

结束语 粗糙集理论与方法在不完备信息系统中的扩充是粗糙集理论研究的一个热点问题。目前, 不完备信息系统中的未知属性值可以有两种不同的语义解释, 即遗漏型或丢失型。由于以往很多学者在不完备决策系统中进行知识约简方法的研究都是基于容差关系的(遗漏型未知属性值), 因此笔者的出发点是不完备决策系统中基于非对称相似关系(丢失型未知属性值)的粗糙集模型。接着, 本文将几种不同的约简形式引入到不完备决策系统中并给出了相应的判定定理与辨识矩阵。这些工作都为从不完备系统中获取知识提供了新的理论方法和技术手段。

笔者下一步将对形式更为复杂的不完备决策系统中知识约简方法进行研究, 如决策系统中同时出现遗漏型和丢失型未知属性值^[16]。

参考文献

- Pawlak Z. Rough set theory and its applications to data analysis [J]. Cybernetics and Systems, 1998, 29: 661~688
- Pawlak Z. Rough sets and intelligent data analysis [J]. Information Sciences, 2002, 147: 1~12
- Kryszkiewicz M. Comparative study of alternative types of knowledge reduction in inconsistent systems [J]. Information Sci-

如图5所示, 本文方法也适用于多维变量模糊差值推理, 得出的结论模糊集(虚线表示的)保证了模糊集的凸性和正规。

结论 Koczy 和 Hirota 提出的 KH 线性插值推理方法的推理结果在许多情况下是非凸的, 甚至不是模糊集, 这使得 KH 方法在实际应用中受到很多限制。本文提出一种基于核集和宽度相似性模糊插值推理方法, 首先利用拉格朗日插值确定模糊集的核心, 然后根据线性插值构造一新的规则, 在模糊集之间的宽度相似性原理的基础上, 使新规则的前件与观察值之间的相似性等于后件与结论模糊集之间的相似性, 从而得到结论模糊集的形状宽度, 保证推理结果的凸性和正规性, 并把此方法推广到多维变量的情况。此方法简单直观, 通过仿真试验并且与 KH 方法进行了比较, 证明该方法能较好地保证推理结果隶属函数的凸性和正规性, 为智能系统中的模糊推理提供了一个十分有用的工具。

参考文献

- Kóczy L T, Hirota K. Interpolative reasoning with insufficient evidence in sparse fuzzy rules bases Inform. Sci, 1993, 71:169~201
- Koczy L T, Hirota K. Approximate reasoning by linear rule interpolation and general approximation [J]. International Journal of Approximate Reasoning, 1993, 9(3):197~225
- Shi Y M M. Some Considerations on Koczy's fuzzy Interpolative Reasoning Method. In: Proceedings of the International Joint Conference of the Fourth IEEE International Fuzzy Engineering Symposium, 1995. 2117~2122
- Chen Shi-Jay, Chen Shyi-Ming. A New Method to Measure the Similarity between Fuzzy Numbers. In: The IEEE International Conference on Fuzzy Systems, 2001. 1123~1126
- 王天江, 卢正鼎, 李凡. 基于几何相似的模糊推理. 计算机科学, 2004, 31(9):169~175
- Li Hong-xing. Fuzzy control interpolative principle [J]. China Science, 1998, 28(3): 259~267
- ences, 1999, 113: 271~292
- 张文修, 米据生, 吴伟志. 不协调目标信息系统的知识约简[J]. 计算机学报, 2003, 26(1): 12~18
- Mi Jusheng, Wu Weizhi, Zhang Wenxiu. Approaches to knowledge reduction based on variable precision rough set model [J]. Information Sciences, 2004, 159 (3-4): 255~272
- 管涛, 冯博琴. 模糊目标信息系统上的知识约简方法[J]. 软件学报, 2004, 15(10): 1470~1478
- 王国胤. Rough 集理论在不完备信息系统中的扩充[J]. 计算机研究与发展, 2002, 39(10): 1238~1243
- Kryszkiewicz M. Rough set approach to incomplete information systems [J]. Information Sciences, 1998, 112: 39~49
- 周献中, 黄兵. 基于粗糙的不完备信息系统属性约简[J]. 南京理工大学学报(自然科学版), 2003, 27(5): 630~635
- Guan Yanyong, Wang Hongkai. Set-valued information systems [J]. International Journal of Information Sciences, 2006, 176 (17): 2507~2525
- Leung Yee, Li Deyu. Maximal consistent block technique for rule acquisition in incomplete information systems [J]. Information Sciences, 2003, 15: 85~106
- Wu Weizhi, Zhang Wenxiu, Li Huaizu. Knowledge acquisition in incomplete fuzzy information systems via the rough set approach [J]. Expert Systems, 2003, 20(5): 280~286
- Grzymala-Busse J W, Wang A Y. Modified algorithms LEM1 and LEM2 for rule induction from data with missing attribute values. In: Proceeding of the Fifth International Workshop on Rough Sets and Soft Computing (RSSC'97) at the Third Joint Conference on Information Sciences (JCIS'97), Research TrianglePark, NC, March 1997. 69~72
- Grzymala-Busse J W. On the unknown attribute values in learning from examples. In: Proceeding of the 6th International Symposium on Methodologies for Intelligent Systems (ISMIS-91), Charlotte, North Carolina, October 1991. Lecture Notes in Artificial Intelligence, vol 542. Springer-Verlag, Berlin, Heidelberg, New York, 1991. 368~377
- Stefanowski J, Tsoukias A. Incomplete information tables and rough classification [J]. Computational Intelligence, 2001, 17: 545~566
- Grzymala-Busse J W. Characteristic Relations for Incomplete Data: A Generalization of the Indiscernibility Relation. In: Proceeding of the Third International Conference on Rough Sets and Current Trends in Computing (RSCTC 2004), Lecture Notes in Artificial Intelligence, vol 3066. Springer-Verlag, Berlin, Heidelberg, New York, 2004. 244~253