控 制 与 决 策
Control and Decision

文章编号: 1001-0920 (2014) 11-2066-05 DOI: 10.13195/j.kzyjc.2013.0936

变精度不协调SFD信息系统中的拓扑学方法

李招文, 文国秋, 林福宁
(广西民族大学 理学院，南宁 530006)

摘 要: 通过介绍模糊决策精度的概念, 引入变精度不协调SFD信息系统; 利用拓扑学中的连通性, 在变精度不协调SFD信息系统中提出连通下(上)近似变精度协调集和连通下(上)近似变精度约简的概念; 给出通过辨识函数计算连通下(上)近似变精度约简的方法, 并给出了一个算例.

关键词: SFD信息系统；拓扑学；连通性；约简；辨识函数
中图分类号: TP18 **文献标志码:** A

Topological approaches in variable precision inconsistent SFD information systems

LI Zhao-wen, *WEN Guo-qiu*, *LIN Fu-ning*
(College of Science，Guangxi University for Nationalities，Nanning 530006，China．Correspondent：LI Zhao-wen, E-mail：lizhaowen8846@126.com)

Abstract: Variable precision inconsistent SFD information systems are proposed by introducing fuzzy decision accuracy. By applying the connectedness in topology, the conceptions of connected lower (upper) approximation variable precision consistent sets and connected lower (upper) approximation variable precision reductions are defined in variable precision inconsistent SFD information systems. The methods for computing a connected lower (upper) approximation variable precision reduction by the corresponding discernible function are given as well as an example.

Key words: SFD information system；topology；connectedness；reduction；discernible function

0 引 言

自粗糙集由Pawlak提出以来便在知识发现、决策分析和数据挖掘等领域获得巨大成功[1-2]. 属性约简是信息系统[3]中的重要研究内容, 通常是运用粗糙集理论中的辩识矩阵和辨识函数进行研究. 集值信息系统[4-5]和模糊决策信息系统[6-7]都是经典的信息系统的拓广, 它们的融合形成了集值模糊决策信息系统(简称SFD信息系统). SFD信息系统既有条件属性, 又有决策属性, 且条件属性对应集值, 决策属性对应隶属度. 文献[8-9]研究了这类信息系统的近似分配约简. 文献[10]考虑了非离散型信息系统的知识约简.

拓扑学是一个重要的数学分支, 是研究粗糙集理论及其应用的有力工具. 目前已有学者[11-12]将拓扑学思想引入到信息系统的研究之中.

本文结合拓扑学方法, 在变精度不协调SFD信息系统中, 通过最大相容类导出拓扑结构, 并联系拓扑学中的一些基本概念(如拓扑基、连通分支等)定义一对近似算子, 给出连通下(上)近似变精度协调集的概念、判定定理与约简方法, 为处理SFD信息系统的属性约简问题提供了一种新思路.

1 信息系统

在本文中, 2^X 表示集 X 的幂集, I 表示闭区间 $[0, 1]$, N 表示自然数集. 为方便讨论, 记 Y^X 为所有由集 X 到集 Y 的映射构成的集.

定义 1[3] 称 (U, A, F, D, G) 为集值模糊决策信息系统. 其中: U 为对象集; A 为条件属性集; $F = \{f_a \in (2^{V_a} - \{\varnothing\})^U : a \in A\}$ 为 U 与 A 的关系集, f_a 和 V_a 分别为 $a \in A$ 的信息函数和值域; D 为决策属性集;

收稿日期: 2013-07-10；**修回日期:** 2013-10-11.
基金项目: 国家自然科学基金项目(11061004, 11461005)；广西民族大学中国-东盟研究中心(广西科学实验中心) 2013年度开放课题项目(KT201310, KT201315)；广西自然科学基金项目(2014GXNSFAA118001)；广西高校科学技术研究项目(2013ZD020)；广西高校优化控制与工程计算重点实验室资助项目；广西哲学社会科学规划课题(13FJL004).
作者简介: 李招文(1962−), 男, 教授, 博士, 从事不确定数学和信息系统等研究；文国秋(1987−), 女, 硕士生, 从事信息系统和决策理论的研究.

$G=\{g_d \in I^U:\ d \in D\}$ 为 U 与 D 的关系集, g_d 为 $d \in D$ 的信息函数.

SFD 信息系统中对象的条件属性值未必是单一的, 若按等价关系来处理分类问题, 会出现划分过细的情况, 不便于提取知识. 为此, 本文考虑将等价关系弱化为文献 [3] 中的相容关系 (即满足自反性和对称性的关系).

定义 2　设 (U,A,F,D,G) 为 SFD 信息系统, 对 $B \subseteq A$, 定义 U 上的关于 B 的相容关系 T_B 如下:

$$T_B = \{(x,y) \in U \times U:\ f_b(x) \bigcap f_b(y) \neq \varnothing (\forall b \in B)\}.$$

若 $\forall x,y \in U$, $(x,y) \in T_B$, 则称 x,y 关于 B 是相容的, 记

$$T_B(x)=\{y \in U:\ (x,y) \in T_B\}.$$

称 $T_B(x)$ 为 x 的相容类.

不同的相容类之间可能存在包含关系, 同时不能保证同一相容类中的对象之间两两相容. 针对这些不足, 本文考虑根据文献 [4] 提出的最大相容类对 SFD 信息系统中的对象进行分类.

定义 3　设 R 为 U 上的相容关系, 若 $X \in 2^U$ 满足: 1) X 为 U 上的一个 R 相容类, 即 $\forall x,y \in X$, 有 $(x,y) \in R$; 2) $\forall x \in U-X$, 存在 $y \in X$ 使得 $(x,y) \notin R$. 则称 X 为 U 上的一个 R 最大相容类.

记 U 上的所有 R 最大相容类构成的集为 $\mathrm{CCR}(U)$. 对 $x \in U$, 记 $\mathrm{CCR}(x)=\{K \in \mathrm{CCR}(U):\ x \in K\}$.

定义 4　设 (U,A,F,D,G) 为 SFD 信息系统, 定义模糊关系 S_D 和变精度关系 S_D^ε.

$$S_D(x,y) = \bigwedge\{1-|\,g_d(x)-g_d(y)|:\ d \in D\},\ x,y \in U,$$

$$S_D^\varepsilon=\{(x,y) \in U \times U: S_D(x,y) \geqslant \varepsilon\},\ \varepsilon \in [0,1],$$

其中 ε 称为 (U,A,F,D,G) 的模糊决策精度.

显然 S_D^ε 为相容关系. 记 $\mathcal{D}^\varepsilon$ 为 U 上由 S_D^ε 导出的分类, 即

$$\mathcal{D}^\varepsilon = U/S_D^\varepsilon = \{S_D^\varepsilon(x): x \in U\},$$

其中 $S_D^\varepsilon(x)=\{y \in U: (x,y) \in S_D^\varepsilon\}$.

2　SFD 信息系统中的拓扑学方法

定义 5[13]　设 (U,A,F,D,G)为 SFD 信息系统, 对于 $a \in A$, 存在 U 上唯一的拓扑 $\tau(a)$, 它的基为 $\mathfrak{B}(a) = \{K_1 \bigcap K_2 \bigcap \cdots \bigcap K_n: K_i \in \mathrm{CCT}_{\{a\}}(U) (i \leqslant n), n \in N\}$. 此时, $\mathrm{CCT}_{\{a\}}(U)$ 为 $\tau(a)$ 的子基. 称 $\tau(a)$ 为由 a 诱导的拓扑, 表示为

$$\tau(a)=\Big\{\bigcup_{W \in \mathfrak{B}'} W: \mathfrak{B}' \subseteq \mathfrak{B}(a)\Big\}.$$

称 $\tau(a)$ 中每个元素为 U 上的 a- 开集, 称 $\tau(a)^c=\{X \in 2^U: U-X \in \tau(a)\}$ 每个元素为 U 上的 a- 闭集.

定义 6[13]　设 (U,A,F,D,G)为 SFD 信息系统, $a \in A$. 在拓扑空间 $(U,\tau(a))$ 中, 对 $X \in 2^U$, 定义

$$\mathrm{cl}_a(X)=\bigcap\{S \in \tau(a)^c: X \subseteq S\},$$

$$\mathrm{int}_a(X)=\bigcup\{S \in \tau(a): S \subseteq X\}.$$

称 $\mathrm{cl}_a(X)$ 和 $\mathrm{int}_a(X)$ 分别为 X 的 a- 闭包和 a- 内部.

定义 7[13]　设 (U,A,F,D,G)为 SFD 信息系统, $a \in A$. 在拓扑空间 $(U,\tau(a))$ 中, 对 $X \in 2^U-\{\varnothing\}$, 若 $\exists Y_1,Y_2 \in 2^U-\{\varnothing\}$ 满足 $X=Y_1 \bigcup Y_2$ 且 $(Y_1 \bigcap \mathrm{cl}_a(Y_2)) \bigcup (\mathrm{cl}_a(Y_1) \bigcap Y_2)=\varnothing$, 则称 X 为 $(U,\tau(a))$ 的一个 a- 隔离子集; 否则, 称 X 为 $(U,\tau(a))$ 的一个 a- 连通子集. 记 $(U,\tau(a))$ 的所有 a- 连通子集构成的集族为 $\mathrm{Con}(a)$. 对 $B \subseteq A$, 记 $R_B^*=\{(x,y) \in U \times U: \forall b \in B, \exists X \in \mathrm{Con}(b)$ 使得 $\{x,y\} \subseteq X\}$, 并记 $U/R_B^*=\{[x]_B^*: x \in U\}$, 其中 $[x]_B^*=\{y \in U: (x,y) \in R_B^*\}$.

易证, $\forall x \in U$, 当 $\mathrm{CCT}_{\{a\}}(x)=\mathrm{CCT}_{\{a\}}(U) (a \in A)$ 时, 有 $[x]_{\{a\}}^*=U$.

下述性质是显然的.

性质 1　设 (U,A,F,D,G) 为 SFD 信息系统, $C \subseteq B \subseteq A$.

1) R_B^* 为 U 上的等价关系;

2) $R_B^*=\bigcap\limits_{b \in B} R_{\{b\}}^*$, $[x]_B^*=\bigcap\limits_{b \in B}[x]_{\{b\}}^* \ (x \in U)$;

3) $R_C^* \supseteq R_B^* \supseteq R_A^*$, $[x]_C^* \supseteq [x]_B^* \supseteq [x]_A^* \ (x \in U)$.

注 1　对于拓扑空间 $(U,\tau(a))$ 及 $x \in U$, $[x]_{\{a\}}^*$ 为 x 在 $(U,\tau(a))$ 中的连通分支. 任一 $y \in [x]_{\{a\}}^*$ 与 x 关于 $(U,\tau(a))$ 是连通的.

定义 8　设 (U,A,F,D,G) 为 SFD 信息系统, $B \subseteq A$. $\forall X \in 2^U$, 记

$$\underline{R_B^*}(X)=\{x \in U: [x]_B^* \subseteq X\},$$

$$\overline{R_B^*}(X)=\{x \in U: [x]_B^* \bigcap X \neq \varnothing\}.$$

则 $\underline{R_B^*}(X)$ 和 $\overline{R_B^*}(X)$ 分别称为 X 关于 B 的连通下近似集和连通上近似集.

性质 2　设 (U,A,F,D,G) 为 SFD 信息系统, $B \subseteq A$, 则 $\forall X,Y \in 2^U$, 有:

1) 若 $X \subseteq Y$, 则 $\overline{R_B^*}(X) \subseteq \overline{R_B^*}(Y)$, $\underline{R_B^*}(X) \subseteq \underline{R_B^*}(Y)$;

2) $\underline{R_B^*}(X) \subseteq \underline{R_A^*}(X) \subseteq X \subseteq \overline{R_A^*}(X) \subseteq \overline{R_B^*}(X)$.

由性质 1 易证性质 2, 此略.

定义 9　设 (U,A,F,D,G) 为 SFD 信息系统. 若 $R_A^* \subseteq S_D^\varepsilon$, 则称 (U,A,F,D,G)为 ε- 精度协调 SFD 信息系统; 否则, 称 (U,A,F,D,G)为 ε- 精度不协调 SFD 信息系统.

定义 10　设 (U,A,F,D,G)为 ε- 精度不协调 SFD

信息系统, $B \subseteq A$.

1) 若 $\underline{R_B^*}(D^\varepsilon) = \underline{R_A^*}(D^\varepsilon)$, $\forall D^\varepsilon \in \mathcal{D}^\varepsilon$, 则称 B 为 (U, A, F, D, G) 的连通下近似 ε- 精度协调集; 若同时还满足 $\forall C \subsetneq B$, C 不是连通下近似 ε- 精度协调集, 则称 B 为连通下近似 ε- 精度约简.

2) 若 $\overline{R_B^*}(D^\varepsilon) = \overline{R_A^*}(D^\varepsilon)$, $\forall\, D^\varepsilon \in \mathcal{D}^\varepsilon$, 则称 B 为 (U, A, F, D, G) 的连通上近似 ε- 精度协调集; 若同时还满足 $\forall C \subsetneq B$, C 不是连通上近似 ε- 精度协调集, 则称 B 为连通上近似 ε- 精度约简.

定义 11 设 (U, A, F, D, G)为 ε- 精度不协调 SFD 信息系统, 令

$$P_\varepsilon^{\downarrow} = \{(x_i, x_j) \in U \times U : \exists D^\varepsilon \in \mathcal{D}^\varepsilon \text{ 使得} x_j \notin D^\varepsilon \text{且} [x_i]_A^* \subseteq D^\varepsilon\},$$

$$P_\varepsilon^{\uparrow} = \{(x_i, x_j) \in U \times U : \exists D^\varepsilon \in \mathcal{D}^\varepsilon \text{ 使得 } x_j \in D^\varepsilon \text{且} D^\varepsilon \bigcap [x_i]_A^* = \varnothing\}.$$

定义

$$H_\varepsilon^{\downarrow}(x_i, x_j) = \begin{cases} \{a \in A : x_j \notin [x_i]_{\{a\}}^*\},\ (x_i, x_j) \in P_\varepsilon^{\downarrow}; \\ \varnothing,\ (x_i, x_j) \notin P_\varepsilon^{\downarrow}; \end{cases}$$

$$H_\varepsilon^{\uparrow}(x_i, x_j) = \begin{cases} \{a \in A : x_j \notin [x_i]_{\{a\}}^*\},\ (x_i, x_j) \in P_\varepsilon^{\uparrow}; \\ \varnothing,\ (x_i, x_j) \notin P_\varepsilon^{\uparrow}. \end{cases}$$

称 $H_\varepsilon^{\downarrow}(x_i, x_j)$ 和 $H_\varepsilon^{\uparrow}(x_i, x_j)$ 分别为 x_i 和 x_j 的连通下近似 ε- 精度辨识集和连通上近似 ε- 精度辨识集, 称

$$\mathcal{H}_\varepsilon^{\downarrow} = \{H_\varepsilon^{\downarrow}(x_i, x_j) :\ \forall\, x_i, x_j \in U\},$$

$$\mathcal{H}_\varepsilon^{\uparrow} = \{H_\varepsilon^{\uparrow}(x_i, x_j) :\ \forall\, x_i, x_j \in U\}$$

分别为 (U, A, F, D, G) 的连通下近似 ε- 精度辨识矩阵和连通上近似 ε- 精度辨识矩阵.

定理 1 设 (U, A, F, D, G)为 ε- 精度不协调 SFD 信息系统, $B \subseteq A$, 则 B 是连通下近似 ε- 精度协调集 $\Leftrightarrow$ 当 $H_\varepsilon^{\downarrow}(x_i, x_j) \neq \varnothing$ 时, $B \bigcap H_\varepsilon^{\downarrow}(x_i, x_j) \neq \varnothing$.

证明 必要性. 设 B 是连通下近似 ε- 精度协调集, 则 $\forall D^\varepsilon \in \mathcal{D}^\varepsilon$, 有 $\underline{R_B^*}(D^\varepsilon) = \underline{R_A^*}(D^\varepsilon)$. 当 $H_\varepsilon^{\downarrow}(x_i, x_j) \neq \varnothing$ 时, $\exists D_0^\varepsilon \in \mathcal{D}^\varepsilon$, 使 $[x_i]_A^* \subseteq D_0^\varepsilon$ 且 $x_j \notin D_0^\varepsilon$, 故 $x_i \in \underline{R_A^*}(D_0^\varepsilon) = \underline{R_B^*}(D_0^\varepsilon)$. 于是 $[x_i]_B^* \subseteq D_0^\varepsilon$. 但 $x_j \notin D_0^\varepsilon$, 所以 $x_j \notin [x_i]_B^* = \bigcap_{b \in B} [x_i]_{\{b\}}^*$, 即存在 $b' \in B$ 使得 $x_j \notin [x_i]_{\{b'\}}^*$, 即 $b' \in B \bigcap H_\varepsilon^{\downarrow}(x_i, x_j)$. 故 $B \bigcap H_\varepsilon^{\downarrow}(x_i, x_j) \neq \varnothing$.

充分性. 由性质 2, 只需证

$$\underline{R_A^*}(D^\varepsilon) \subseteq \underline{R_B^*}(D^\varepsilon),\ \forall D^\varepsilon \in \mathcal{D}^\varepsilon.$$

设 $D_0^\varepsilon \in \mathcal{D}^\varepsilon$, 取 $x_i \in \underline{R_A^*}(D_0^\varepsilon)$, 此时 $[x_i]_A^* \subseteq D_0^\varepsilon$. 考虑 $x_j \in U$ 满足 $x_j \notin D_0^\varepsilon$, 则 $H_\varepsilon^{\downarrow}(x_i, x_j) \neq \varnothing$. 由题设, $B \bigcap H_\varepsilon^{\downarrow}(x_i, x_j) \neq \varnothing$. 不妨取 $b' \in B \bigcap H_\varepsilon^{\downarrow}(x_i, x_j)$. 由 $b' \in H_\varepsilon^{\downarrow}(x_i, x_j)$, 有 $x_j \notin [x_i]_{\{b'\}}^*$; 由性质 1, $x_j \notin [x_i]_B^*$. 而 $x_j \notin D_0^\varepsilon$, 说明 $[x_i]_B^* \subseteq D_0^\varepsilon$, 即 $x_i \in \underline{R_B^*}(D_0^\varepsilon)$. 由上述讨论可知, $\underline{R_A^*}(D_0^\varepsilon) \subseteq \underline{R_B^*}(D_0^\varepsilon)$. 由 D_0^ε 选取的任意性, $\underline{R_A^*}(D^\varepsilon) \subseteq \underline{R_B^*}(D^\varepsilon) (\forall D^\varepsilon \in \mathcal{D}^\varepsilon)$, 即 B 为连通下近似 ε- 精度协调集. □

定理 2 设 (U, A, F, D, G) 为 ε- 精度不协调 SFD 信息系统, $B \subseteq A$, 则 B 是连通上近似 ε- 精度协调集 $\Leftrightarrow$ 当 $H_\varepsilon^{\uparrow}(x_i, x_j) \neq \varnothing$ 时, 有 $B \bigcap H_\varepsilon^{\uparrow}(x_i, x_j) \neq \varnothing$.

证明 必要性. 设 B 是连通上近似 ε- 精度协调集, 则 $\forall D^\varepsilon \in \mathcal{D}^\varepsilon$, $\overline{R_B^*}(D^\varepsilon) = \overline{R_A^*}(D^\varepsilon)$. 当 $H_\varepsilon^{\uparrow}(x_i, x_j) \neq \varnothing$ 时, $\exists D_0^\varepsilon \in \mathcal{D}^\varepsilon$, 使 $[x_i]_A^* \bigcap D_0^\varepsilon = \varnothing$ 且 $x_j \in D_0^\varepsilon$, 故 $x_i \notin \overline{R_A^*}(D_0^\varepsilon)$. 从而 $x_i \notin \overline{R_B^*}(D_0^\varepsilon)$, 即 $[x_i]_B^* \cap D_0^\varepsilon = \varnothing$. 但 $x_j \in D_0^\varepsilon$, 所以 $x_j \notin [x_i]_B^* = \bigcap_{b \in B} [x_i]_{\{b\}}^*$, 即存在 $b' \in B$, 使得 $x_j \notin [x_i]_{\{b'\}}^*$. 这说明 $b' \in H_\varepsilon^{\uparrow}(x_i, x_j)$, 即 $B \bigcap H_\varepsilon^{\uparrow}(x_i, x_j) \neq \varnothing$.

充分性. 由性质 2, 只需证

$$\overline{R_B^*}(D^\varepsilon) \subseteq \overline{R_A^*}(D^\varepsilon),\ \forall\, D^\varepsilon \in \mathcal{D}^\varepsilon.$$

设 $D_0^\varepsilon \in \mathcal{D}^\varepsilon$, 取 $x_i \in U - \overline{R_A^*}(D_0^\varepsilon)$, 此时 $[x_i]_A^* \bigcap D_0^\varepsilon = \varnothing$. 考虑 $x_j \in D_0^\varepsilon$, 则 $H_\varepsilon^{\uparrow}(x_i, x_j) \neq \varnothing$. 由题设, $B \bigcap H_\varepsilon^{\uparrow}(x_i, x_j) \neq \varnothing$, 不妨取 $b' \in B \bigcap H_\varepsilon^{\uparrow}(x_i, x_j)$. 由 $b' \in H_\varepsilon^{\uparrow}(x_i, x_j)$, $x_j \notin [x_i]_{\{b'\}}^*$ 和性质 1, 有 $x_j \notin [x_i]_B^*$. 而 $x_j \in D_0^\varepsilon$, 这说明 $[x_i]_B^* \bigcap D_0^\varepsilon = \varnothing$, 即 $x_i \in U - \overline{R_B^*}(D_0^\varepsilon)$. 由上述讨论可知, $\overline{R_B^*}(D_0^\varepsilon) \subseteq \overline{R_A^*}(D_0^\varepsilon)$. 由 D_0^ε 选取的任意性, $\overline{R_B^*}(D^\varepsilon) \subseteq \overline{R_A^*}(D^\varepsilon) (\forall\, D^\varepsilon \in \mathcal{D}^\varepsilon)$, 即 B 为连通上近似 ε- 精度协调集. □

"$\bigvee$"(析取), "$\bigwedge$"(合取), "$\to$"(蕴含), "$\leftrightarrow$"(逻辑等价) 代表数理逻辑中的命题连接词, 它们分别读作 "或", "且", "若-则", "当且仅当".

为方便叙述, $\forall \{a_1, a_2, \cdots, a_k\} \subseteq A$, 记

$$\bigvee \{a_1, a_2, \cdots, a_k\} \Big(\text{或} \bigvee_{i=1}^{k} a_k \Big) = a_1 \bigvee a_2 \bigvee \cdots \bigvee a_k,$$

$$\bigwedge \{a_1, a_2, \cdots, a_k\} \Big(\text{或} \bigwedge_{i=1}^{k} a_k \Big) = a_1 \bigwedge a_2 \bigwedge \cdots \bigwedge a_k.$$

定义 12 设 (U, A, F, D, G)为 ε- 精度不协调 SFD 信息系统, 记

$$\varDelta_\varepsilon^{\downarrow} = \bigwedge_{(x_i, x_j) \in P_\varepsilon^{\downarrow}} \bigvee H_\varepsilon^{\downarrow}(x_i, x_j),$$

$$\varDelta_\varepsilon^{\uparrow} = \bigwedge_{(x_i, x_j) \in P_\varepsilon^{\uparrow}} \bigvee H_\varepsilon^{\uparrow}(x_i, x_j).$$

称 $\varDelta_\varepsilon^{\downarrow}$ 和 $\varDelta_\varepsilon^{\uparrow}$ 分别为 (U, A, F, D, G)的连通下近似 ε- 精度辨识函数和连通上近似 ε- 精度辨识函数.

定义 13 设 (U, A, F, D, G)为 ε- 精度不协调 SFD 信息系统, 并且 $\varDelta_\varepsilon^{\downarrow}$ 和 $\varDelta_\varepsilon^{\uparrow}$ 分别为它的连通下近似 ε- 精度辨识函数和连通上近似 ε- 精度辨识函数.

1) 若 $\varDelta_\varepsilon^{\downarrow} = \bigvee_{m=1}^{k} \Big(\bigwedge_{s=1}^{p_m} a_{ms} \Big)$, 并且每个

$$B_m = \{a_{ms} \in A : s \leqslant p_m\}$$

没有重复元素, 则称 $\bigvee_{m=1}^{k}\Big(\bigwedge_{s=1}^{p_m} a_{ms}\Big)$ 是 $\Delta_\varepsilon^{\downarrow}$ 的极小析取范式, 记为 $(\Delta_\varepsilon^{\downarrow})^*$, 即

$$(\Delta_\varepsilon^{\downarrow})^* = \bigvee_{m=1}^{k}\Big(\bigwedge_{s=1}^{p_m} a_{ms}\Big).$$

2) 若 $\Delta_\varepsilon^{\uparrow} = \bigvee_{n=1}^{l}\Big(\bigwedge_{t=1}^{q_n} a_{nt}\Big)$, 且每个 $C_n=\{a_{nt}\in A: t\leqslant q_n\}$ 没有重复元素, 则称 $\bigvee_{n=1}^{l}\Big(\bigwedge_{t=1}^{q_n} a_{nt}\Big)$ 是 $\Delta_\varepsilon^{\uparrow}$ 的极小析取范式, 记 $(\Delta_\varepsilon^{\uparrow})^*$, 即

$$(\Delta_\varepsilon^{\uparrow})^* = \bigvee_{n=1}^{l}\Big(\bigwedge_{t=1}^{q_n} a_{nt}\Big).$$

定理 3　设 (U,A,F,D,G) 为 ε- 精度不协调 SFD 信息系统, $\Delta_\varepsilon^{\downarrow}$ 和 $\Delta_\varepsilon^{\uparrow}$ 分别为它的连通下近似 ε- 精度辨识函数和连通上近似 ε- 精度辨识函数. 若

$$(\Delta_\varepsilon^{\downarrow})^* = \bigvee_{m=1}^{k}\Big(\bigwedge_{s=1}^{p_m} a_{ms}\Big),$$

$$(\Delta_\varepsilon^{\uparrow})^* = \bigvee_{n=1}^{l}\Big(\bigwedge_{t=1}^{q_n} a_{nt}\Big),$$

则 (U,A,F,D,G) 的所有连通下近似 ε- 精度约简和所有连通上近似 ε- 精度约简可分别表示为

$$B_m = \{a_{ms}: s\leqslant p_m\},\ m\leqslant k,$$

$$C_n = \{a_{nt}: t\leqslant q_n\},\ n\leqslant l.$$

证明　仅证 (U,A,F,D,G) 的所有连通下近似 ε- 精度约简可表示为 $B_m = \{a_{ms}: s\leqslant p_m\}(m\leqslant k)$. 类似可证另一情形.

假设 $|U|=r$, 记 $d_{ij}=H_\varepsilon^{\downarrow}(x_i,x_j)(i,j\leqslant r)$. 证明过程分为两步.

1) 取 $B_{m_0}\in\{B_m: m\leqslant k\}$.

① 显然, $\Delta_\varepsilon^{\downarrow} = \bigvee_{m=1}^{k}\Big(\bigwedge_{s=1}^{p_m} a_{ms}\Big) = \bigvee_{m=1}^{k}(\bigwedge B_m)$, 故 $\bigwedge B_{m_0}\to\Delta_\varepsilon^{\downarrow}$.

由于

$$(\Delta_\varepsilon^{\downarrow})^* = \Delta_\varepsilon^{\downarrow} = \bigwedge(\bigvee d_{ij}),$$

故 $\Delta_\varepsilon^{\downarrow}\leftrightarrow\bigvee d_{ij}(\forall\, i,j\leqslant r)$, 从而

$$\bigwedge B_{m_0}\to\bigvee d_{ij}(\forall\, i,j\leqslant r).$$

由上述讨论可知

$$\bigwedge B_{m_0}\leftrightarrow a_{m_0 s},\ \forall\, s\leqslant p_{m_0},$$

且对某个 $a\in d_{ij}$, 有 $\bigvee d_{ij}\leftrightarrow a$, 则对某个 $a\in d_{ij}$, $a_{m_0 s}\to a(\forall\, i,j\leqslant r, s\leqslant p_{m_0})$. 故 $\forall\, i,j\leqslant r$, $\exists s_0\leqslant p_{m_0}$, 使 $a=a_{m_0 s_0}$, 即 $a\in B_{m_0}\bigcap d_{ij}$. 从而 $\forall i,j\leqslant r$, 有 $B_{m_0}\bigcap d_{ij}\neq\varnothing$.

② 要证 B_{m_0} 为连通下近似 ε- 精度约简约简, 只需证 $\forall a\in B_{m_0}, \exists\, i',j'\leqslant r$ 使得

$$(B_{m_0}-\{a\})\bigcap d_{i'j'}=\varnothing.$$

倘若 $\exists\, a_0\in B_{m_0}$ 使得 $(B_{m_0}-\{a_0\})\bigcap d_{ij}\neq\varnothing(\forall\, i,j\leqslant r)$, 任取 $b\in(B_{m_0}-\{a_0\})\cap d_{ij}$, 则 $\bigwedge(B_{m_0}-\{a_0\})\to b$ 且 $b\to\bigvee d_{ij}$, 故

$$\bigwedge(B_{m_0}-\{a_0\})\to\bigvee d_{ij},\ \forall i,j\leqslant r.$$

由于 $(\Delta_\varepsilon^{\downarrow})^*$ 包含了 $\Delta_\varepsilon^{\downarrow}$ 中所有真解释, 则 $B_{m_0}-\{a_0\}\in\{B_m: m\leqslant k\}$. 故

$$\begin{aligned}&(\bigwedge B_{m_0})\bigvee(\bigwedge(B_{m_0}-\{a_0\}))=\\&((\bigwedge(B_{m_0}-\{a_0\}))\bigwedge\{a_0\})\bigvee\\&((\bigwedge(B_{m_0}-\{a_0\}))\bigwedge 1)=\\&(\bigwedge(B_{m_0}-\{a_0\}))\bigwedge(\{a_0\}\bigvee 1)=\\&(\bigwedge(B_{m_0}-\{a_0\}))\bigwedge 1=\\&\bigwedge(B_{m_0}-\{a_0\}).\end{aligned}$$

这说明 $B_{m_0}\notin\{B_m: m\leqslant k\}$, 矛盾.

终上, B_{m_0} 为连通下近似 ε- 精度约简.

2) 设 B 为连通下近似 ε- 精度约简. 只需证 $\exists\, B_{m_1}\in\{B_m: m\leqslant k\}$ 使得 $B=B_{m_1}$ 即可.

由于 B 为连通下近似 ε- 精度约简, 由定理 1, 有 $B\bigcap d_{ij}\neq\varnothing(\forall i,j\leqslant r)$. 类似于 1) 中 ② 的证明, 有 $B\in\{B_m: m\leqslant k\}$. 故 $\exists\, B_{m_1}\in\{B_m: m\leqslant k\}$ 使得 $B=B_{m_1}$. 因此 (U,A,F,D,G) 的所有连通下近似 ε- 精度约简可表为

$$B_m=\{a_{ms}: s\leqslant p_m\},\ m\leqslant k.$$

由此, 定理 3 得证. □

例 1　由表 1 给出的 SFD 信息系统, 记 $A=\{a_1, a_2, a_3\}$, $D=\{d_1,d_2,d_3\}$.

表 1　SFD 信息系统

U	a_1	a_2	a_3	d_1	d_2	d_3
x_1	$\{0,1\}$	$\{0,2\}$	$\{0,1\}$	0.7	0.2	0.4
x_2	$\{1\}$	$\{1\}$	$\{0\}$	0.1	0.5	0.8
x_3	$\{2\}$	$\{1\}$	$\{2\}$	0.2	0.3	0.6
x_4	$\{0\}$	$\{0,1,2\}$	$\{2\}$	0.9	0.1	0.3

通过计算可知, (U,A,F,D,G) 为 0.8- 精度不协调 SFD 信息系统, 由 $a_i\ (i=1,2,3)$ 诱导的拓扑分别为

$$\begin{aligned}&\tau(a_1)=\\&\{\varnothing,\{x_1\},\{x_3\},\{x_1,x_2\},\{x_1,x_3\},\{x_1,x_4\},\\&\{x_1,x_2,x_3\},\{x_1,x_2,x_4\},\{x_1,x_3,x_4\},U\};\\&\tau(a_2)=\{\varnothing,\{x_4\},\{x_1,x_4\},\{x_2,x_3,x_4\},U\};\\&\tau(a_3)=\{\varnothing,\{x_1,x_2\},\{x_3,x_4\},U\}.\end{aligned}$$

a_i- 连通子集 $(i=1,2,3)$ 构成的集族分别为

$$\begin{aligned}&\mathrm{Con}(a_1)=\\&\{\{x_1\},\{x_2\},\{x_3\},\{x_4\},\{x_1,x_2\},\\&\{x_1,x_4\},\{x_1,x_2,x_4\}\};\end{aligned}$$

$\mathrm{Con}(a_2) =$

$\{\{x_1\},\{x_2\},\{x_3\},\{x_4\},\{x_1,x_4\},$

$\{x_2,x_3\},\{x_2,x_4\},\ \{x_3,x_4\},\ \{x_1,x_2,x_4\},$

$\{x_1,x_3,x_4\},\ \{x_2,x_3,x_4\},U\};$

$\mathrm{Con}(a_3) =$

$\{\{x_1\},\{x_2\},\{x_3\},\{x_4\},\{x_1,x_2\},\{x_3,x_4\}\}.$

连通下近似 0.8- 精度辨识矩阵为

$$\mathcal{H}^{\downarrow}_{0.8} = \begin{bmatrix} \varnothing & \varnothing & \varnothing & \varnothing \\ \varnothing & \varnothing & \varnothing & \varnothing \\ \{a_1,a_3\} & \varnothing & \varnothing & \{a_1\} \\ \varnothing & \{a_3\} & \{a_1\} & \varnothing \end{bmatrix}.$$

连通下近似 0.8- 精度辨识函数为

$$\varDelta^{\downarrow}_{0.8} = (a_1 \bigvee a_3) \bigwedge a_1 \bigwedge a_3 = a_1 \bigwedge a_3.$$

由定理 3 可知, $B = \{a_1, a_3\}$ 为 (U, A, F, D, G) 的连通下近似 0.8- 精度约简.

3 结 论

本文通过介绍模糊决策精度, 引入了变精度不协调 SFD 信息系统. 利用拓扑学研究了变精度不协调 SFD 信息系统的两种约简方法. 今后, 将利用其他数学工具研究变精度不协调 SFD 信息系统的属性约简.

参考文献(References)

[1] Pawlak Z. Rough sets[J]. Int J of Computer Information, 1982, 11(5): 341-356.

[2] Pawlak Z. Rough sets: Theoretical aspects of reasoning about data[M]. Dordrecht: Kluwer Academic Publishers, 1991: 1-8.

[3] 张文修, 梁怡, 吴伟志. 信息系统与知识发现[M]. 北京: 科学出版社, 2003: 96-189.
(Zhang W X, Liang Y, Wu W Z. Information system and knowledge discovery[M]. Beijing: Science Publishers, 2003: 96-189.)

[4] 王虹, 张文修, 李鸿儒. 集值信息系统的知识发现与约简[J]. 计算机工程与应用, 2005, 41(6): 37-38.
(Wang H, Zhang W X, Li H R. Knowledge reduction in set-valued information system[J]. Computer Engineering and Applications, 2005, 41(6): 37-38.)

[5] 管延勇, 王洪凯, 史开泉. 集值信息系统及其属性约简[J]. 数学的实践与认识, 2008, 28(2): 101-107.
(Guan Y Y, Wang H K, Shi K Q. Set-valued information systems and its attribute reduction[J]. Mathematics in Practice and Theory, 2008, 28(2): 101-107.)

[6] 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.
(Guan T, Feng B Q. Knowledge reduction methods in fuzzy objective information systems[J]. J of Software, 2004, 15(10): 1470-1478.)

[7] 袁修久, 张文修. 模糊目标信息系统的属性约简[J]. 系统工程理论与实践, 2004, 24(5): 116-120.
(Yuan X J, Zhang W X. Attribute reductions in fuzzy inconsistent information systems[J]. Systems Engineering Theory and Practice, 2004, 24(5): 116-120.)

[8] 陈子春, 秦克云. 集值模糊目标信息系统的近似分配约简[J]. 四川师范大学学报, 2011, 34(1): 23-27.
(Chen Z C, Qin K Y. Approximate assignment reduction in set-valued fuzzy decision information system[J]. J of Sichuan Normal University, 2011, 34(1): 23-27.)

[9] 陈子春. 集值信息系统的知识发现与属性约简研究[D]. 成都: 西南交通大学数学学院, 2011: 41-93.
(Chen Z C. The Study of knowledge discovery and attribute reduction in set-valued information systems[D]. Chengdu: School of Mathematics, Southwest Jiaotong University, 2001: 41-93.)

[10] 杨芳. 非离散型信息系统知识约简的粗糙集方法[D]. 济南: 济南大学数学学院, 2010: 26-69.
(Yang F. Knowledge reductions in nondiscrete information system based on rough set theory[D]. Ji'nan: College of Mathmatical Science, Ji'nan University, 2010: 26-69.)

[11] Lashin E F, Medhat T. Topological reduction of information systems[J]. Chaos, Solitons and Fractals, 2005, 25(2): 277-286.

[12] 李进金, 张燕兰, 林培榕. 决策信息系统中的拓扑学方法[J]. 模糊系统与数学, 2011, 25(1): 124-133.
(Li J J, Zhang Y L, Lin P R. Topologica methods in decision information systems[J]. Fuzzy Systems and Mathematics, 2011, 25(1): 124-133.)

[13] 林福宁. 关于不协调集值模糊决策信息系统[D]. 南宁: 广西民族大学理学院, 2013: 34-38.
(Lin F N. On inconsistent set-valued fuzzy decision information systems[D]. Nanning: College of Science, Guangxi University for Nationalities, 2013: 34-38.)

(责任编辑: 齐 霁)