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


拓扑学是一个重要的数学分支，是研究拓扑学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 (如拓补基，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，设为轴的幂集，为闭区间，表示自然数集。为方便讨论，记为所有由集到集的映射构成的集。

定义 1[3] 令为集值模糊决策信息系统的近似集。其中：为对象集，为条件属性集；为决策属性集；与为条件属性集，对条件属性集和集；和分别表示为的信息值和值域；为决策属性集；
定义 2 设 $\langle U, A, F, D, G \rangle$ 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:

$$T_B = \{(x, y) \in U \times U : f_B(x) \cap f_B(y) \neq \emptyset (\forall b \in B)\}.$$  

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 $U$ 上的相关系，若 $x \in U^2$ 满足：(1) $X$ 为 $U$ 上的一个 $R$ 相容类，即 $\forall x, y \in X, (x, y) \in R; 2)$ $\forall x \in X - \{x\}$ 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 $CCR(U)$. 对 $x \in U$, 记 $CCR(x) = \{K \in CCR(U) : x \in K\}$.

定义 4 设 $\langle U, A, F, D, G \rangle$ 为 SFD 信息系统，定义模糊关系 $S_D$ 和变精度关系 $S_D^e$.

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

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

其中 $\epsilon$ 称为 $(U, A, F, D, G)$ 的模糊决策精度。

显然 $S_D^e$ 为相容关系。记 $D^e = U$ 上由 $S_D^e$ 导出的分类，即

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

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

2 SFD 信息系统中的拓扑学方法

定义 5[3] 设 $\langle U, A, F, D, G \rangle$ 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 $\mathfrak{B}(a) = \{K_1 \cap K_2 \cap \cdots \cap K_n : K_i \in CCT_{\{a\}}(U) (i \leq n), n \in N\}$. 此时，$CCT_{\{a\}}(U)$ 为 $\tau(a)$ 的子基。称 $\tau(a)$ 为由 $a$ 诱导的拓扑，表示为

$$\tau(a) = \left\{ \bigcup_{W \in \mathfrak{B}^e} W : \mathfrak{B}^e \subseteq \mathfrak{B}(a) \right\}.$$  

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

定义 6[3] 设 $\langle U, A, F, D, G \rangle$ 为 SFD 信息系统，$a \in A$. 在拓扑空间 $(U, \tau(a))$ 中，对于 $x \in U^2$，定义

$$c_{\tau}(X) = \bigcap\{S : S \in \tau(a) : X \subseteq S\},$$

$$\tau(a) = \bigcup\{S : \tau(a) \subseteq S \}.$$

称 $c_{\tau}(X)$ 和 $\tau(a)$ 分别为 $X$ 的 $a$-闭包和 $a$-内部。

定义 7[3] 设 $\langle U, A, F, D, G \rangle$ 为 SFD 信息系统，$a \in A$. 在拓扑空间 $(U, \tau(a))$ 中，对于 $x \in U^2$, 定义

$$\tau(a) = \bigcap\{S : S \in \tau(a) : X \subseteq S\},$$

$$\tau(a) = \bigcup\{S : \tau(a) \subseteq S \}.$$  

称 $c_{\tau}(X)$ 和 $\tau(a)$ 分别为 $X$ 的 $a$-闭包和 $a$-内部。
定义 11 设 \((U, A, F, D, G)\) 为 \(\varepsilon\)-精度不协调 SFD 信息系统，令

\[ P^*_2 = \{(x_i, x_j) \in U \times U : \exists \delta^2 \in \mathcal{D}^2 \text{ 使得 } x_j \notin \delta^2 \land [x_i]_{\delta} \subseteq \delta^2 \}, \]

\[ P^*_1 = \{(x_i, x_j) \in U \times U : \exists \delta^2 \in \mathcal{D}^2 \text{ 使得 } x_j \notin \delta^2 \land \delta^2 \cap [x_i]_{\delta} = \emptyset \}. \]

定义

\[ H^*_2(x_i, x_j) = \begin{cases} \{ \alpha \in A : x_j \notin [x_i]_{\alpha} \} \cup (x_i, x_j) & \text{在 } P^*_2; \\ \emptyset, (x_i, x_j) & \text{在 } P^*_1; \end{cases} \]

\[ H^*_1(x_i, x_j) = \begin{cases} \{ \alpha \in A : x_j \notin [x_i]_{\alpha} \} \cup (x_i, x_j) & \text{在 } P^*_2; \\ \emptyset, (x_i, x_j) & \text{在 } P^*_1; \end{cases} \]

称 \( H^*_2(x_i, x_j) \) 和 \( H^*_1(x_i, x_j) \) 分别为 \( x_i \) 和 \( x_j \) 的连通下近似 \( \varepsilon \)-精度辨识集和连通上近似 \( \varepsilon \)-精度辨识集，记

\[ H^*_2 = \{ H^*_2(x_i, x_j) : \forall x_i, x_j \in U \}, \]

\[ H^*_1 = \{ H^*_1(x_i, x_j) : \forall x_i, x_j \in U \}. \]

分别对于 \((U, A, F, D, G)\) 的连通下近似 \( \varepsilon \)-精度辨识矩阵和连通上近似 \( \varepsilon \)-精度辨识矩阵。

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

\[ R^*_1(D^2) = R^*_1(D^2), \forall D^2 \in \mathcal{D}^2 \text{ 且 } R^*_1(D^2) \subseteq R^*_0(D^2). \]

\[ R^*_0(D^2) = R^*_0(D^2), \forall D^2 \in \mathcal{D}^2. \]

讨论可知，\( R^*_1(D^2) \subseteq R^*_0(D^2) \). 但 \( D^0 \) 选取的任意性，

\[ R^*_1(D^2) \subseteq R^*_0(D^2)(\forall D^2 \in \mathcal{D}^2), \text{ 即 } B \text{ 为连通下近似 } \varepsilon \text{-精度辨识集} \]
没有重复元，即有

\[ \bigvee_{m=1}^{k} \left( \bigwedge_{s=1}^{p_m} a_{ms} \right) \]

是 \( \Delta_t \) 的极小析取形式，记为 \( (\Delta_t)^* \)，即

\[ (\Delta_t)^* = \bigvee_{m=1}^{k} \left( \bigwedge_{s=1}^{p_m} a_{ms} \right). \]

2) 若 \( \Delta_t = \bigvee_{m=1}^{k} \left( \bigwedge_{s=1}^{q_m} a_{nt} \right) \) 且每个 \( C_n = \{ a_{nt} \in A : \]

\[ t \leq q_n \} \] 没有重复元，即有

\[ (\Delta_t)^* = \bigvee_{m=1}^{l} \left( \bigwedge_{s=1}^{q_n} a_{nt} \right). \]

定义 3 设 \( (U, A, F, D, G) \) 为 \( \varepsilon \)-精度不协调 SFD 信息系统，\( \Delta_t^2 \) 和 \( \Delta_t^3 \) 分别为它的连通下近似 \( \varepsilon \)-精度辨识函数和连通上近似 \( \varepsilon \)-精度辨识函数。若

\[ (\Delta_t^2)^* = \bigvee_{m=1}^{k} \left( \bigwedge_{s=1}^{p_m} a_{ms} \right), \]

\[ (\Delta_t^3)^* = \bigvee_{m=1}^{l} \left( \bigwedge_{s=1}^{q_n} a_{nt} \right), \]

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

\[ B_m = \{ a_{ms} : s \leq p_m \} \]

\( m \leq k \),

\[ C_n = \{ a_{nt} : t \leq q_n \} \]

\( n \leq l \).

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

假设 \( |U| = r \), 记 \( d_{ij} = H^2_{ij}(x_i, x_j) \) \( (i, j \leq r) \)。证明过程分为两步。

1) 取 \( B_m = \{ B_m : m \leq k \} \).

① 显然，\( \Delta_t = \bigvee_{m=1}^{k} \left( \bigwedge_{s=1}^{p_m} a_{ms} \right) = \bigvee_{m=1}^{k} (\bigwedge_{s=1}^{p_m} a_{ms}). \) 故

\[ \bigwedge B_m \rightarrow \Delta_t^2. \]

由于

\[ (\Delta_t)^* = \Delta_t^2 \]

\[ = \bigwedge \left( \bigvee d_{ij} \right) \]

故 \( \Delta_t^2 \iff \bigvee d_{ij} (\forall i, j \leq r) \)，从而

\[ \bigwedge B_m \rightarrow \bigvee d_{ij} (\forall i, j \leq r). \]

由上述讨论可得

\[ \bigwedge B_m \leftrightarrow a_{ms}, s \leq p_m. \]

且对某个 \( a \in d_{ij} \)，有 \( \bigvee d_{ij} \leftrightarrow a \)。则对某个 \( a \in d_{ij} \)，\( a_{ms} \rightarrow a (\forall i, j \leq r, s \leq p_m) \)，故 \( a \in d_{ij} \rightarrow d_{ij}, a_{ms} \rightarrow d_{ij} \)。从而 \( \forall i, j \leq r, x_{\emptyset} \leq p_m, a = a_{ms} \text{ 函数 } \)。即有

\[ \bigwedge B_m \bigcap d_{ij} \neq \emptyset. \]

② 要证 \( B_m \) 为连通下近似 \( \varepsilon \)-精度约简约简，只需证 \( \forall a \in B_m, \exists i', j' \leq r \) 使得

\[ (B_m - \{ a \}) \bigcap d_{ij} = \emptyset. \]

倘若 \( \exists a \in a \in B_m \) 使得 \( (B_m - \{ a \}) \bigcap d_{ij} \neq \emptyset \) \( (\forall i, j \leq r) \)，任取 \( b \in (B_m - \{ a \}) \bigcap d_{ij} \)，则 \( \bigwedge (B_m - \{ a \}) \rightarrow b \) 且 \( b \rightarrow d_{ij} \)。故

\[ \bigwedge (B_m - \{ a \}) \rightarrow \bigvee d_{ij}, \forall i, j \leq r. \]

由于 \( (\Delta_t^2)^* \) 包含了 \( \Delta_t^2 \) 中所有真解释，则 \( B_m - \{ a \} \in \{ B_m : m \leq k \} \)。故

\[ (B_m - \{ a \}) \bigvee (B_m - \{ a \}) = (B_m - \{ a \}) \bigvee (B_m - \{ a \}) \bigvee (B_m - \{ a \}) \bigvee (B_m - \{ a \}) \bigvee (B_m - \{ a \}) \bigvee (B_m - \{ a \}) \bigvee (B_m - \{ a \}) = (B_m - \{ a \}). \]

这说明 \( B_m \notin \{ B_m : m \leq k \} \)。矛盾。

终上，\( B_m \) 为连通下近似 \( \varepsilon \)-精度约简。

2) 设 \( B \) 为连通下近似 \( \varepsilon \)-精度约简。只需证 \( \exists B_m \in \{ B_m : m \leq k \} \) 使得 \( B = B_m \)，即可。

由于 \( B \) 为连通下近似 \( \varepsilon \)-精度约简，由定理 1，有 \( B \bigcap d_{ij} \neq \emptyset \) \( (\forall i, j \leq r) \)。类似于 1) 中②的证明，有 \( B \in \{ B_m : m \leq k \} \)。故 \( \exists B_m \in \{ B_m : m \leq k \} \) 使得 \( B = B_m \)，因此 \( (U, A, F, D, G) \) 的所有连通下近似 \( \varepsilon \)-精度约简可表示为

\[ B_m = \{ a_{ms} : s \leq p_m \} \]

\( m \leq k \).

由此，定理 3 得证。□

例 1 由表 1 给出的 SFD 信息系统，设 \( A = \{ a_1, a_2, a_3 \}, D = \{ d_1, d_2, d_3 \} \)。

<table>
<thead>
<tr>
<th>U</th>
<th>a_1</th>
<th>a_2</th>
<th>a_3</th>
<th>d_1</th>
<th>d_2</th>
<th>d_3</th>
</tr>
</thead>
<tbody>
<tr>
<td>x_1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.7</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>x_2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0.1</td>
<td>0.5</td>
<td>0.8</td>
</tr>
<tr>
<td>x_3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>0.2</td>
<td>0.3</td>
<td>0.6</td>
</tr>
<tr>
<td>x_4</td>
<td>0</td>
<td>0.1</td>
<td>2</td>
<td>0.9</td>
<td>0.1</td>
<td>0.3</td>
</tr>
</tbody>
</table>

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

\[ \tau(a_1) = \{ \emptyset, \{ x_1 \}, \{ x_1, x_2 \}, \{ x_1, x_2, x_3 \}, \{ x_1, x_2, x_3, x_4 \} \}; \]

\[ \tau(a_2) = \{ \emptyset, \{ x_1 \}, \{ x_1, x_2 \}, \{ x_2, x_3, x_4 \} \}; \]

\[ \tau(a_3) = \{ \emptyset, \{ x_1 \}, \{ x_2 \} \}; \]

\( a_i \)-连通子集 \( (i = 1, 2, 3) \) 构成的集族分别为

\[ \text{Con}(a_1) = \{ \{ x_1 \} \}; \]

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

\[ \{ x_1, x_3 \}, \{ x_1, x_2, x_3 \}, \{ x_1, x_2, x_4 \} \}. \]
Con(a_2) =
\{\{x_1\}, \{x_2\}, \{x_3\}, \{x_4\}, \{x_1, x_4\}, \{x_2, x_3\}, \{x_3, x_4\}, \{x_1, x_2, x_4\}, \{x_1, x_3, x_4\}, \{x_2, x_3, x_4\}\};

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
\[
H_{0.8}^i = \begin{bmatrix}
\emptyset & \emptyset & \emptyset & \emptyset \\
\emptyset & \emptyset & \emptyset & \emptyset \\
\{a_1, a_3\} & \emptyset & \emptyset & \{a_1\} \\
\emptyset & \{a_3\} & \{a_1\} & \emptyset
\end{bmatrix}.
\]

连通下近似 0.8-精度辨识函数为
\[
\Delta_{0.8}^i = (a_1 \lor a_3) \land a_1 \land a_3 = a_1 \land a_3.
\]

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)