一种基于连续不确定决策表的概念近似方法

花文健，　刘作良，　韩　兵
(空军工程大学 电讯工程学院，陕西 西安　710077)

摘　要:复杂决策表通常具有连续定量的属性,存在部分未知值或存在描述一个对象的属性有多个可能值的情况,或者3种情况并存,称为连续不确定决策表。通过分析发现,连续不确定决策表可视为一种多值表元决策表。利用Fuzzy集理论可将多值表元决策表转化为带有隶属度的单一表元决策表;基于此,给出了扩展信息表和决策表的定义,提出了对多值表元决策表中决策概念下近似及边界的计算方法。

关键词: 连续不确定决策系统; 扩展Rough集; 概念近似
中图分类号:TP311　　**文献标识码**:A　　**文章编号**:1009-3516(2005)01-0038-06

在运用CRST进行实际数据分析时,往往要对信息表(决策表)进行预处理,即补全表中缺失的信息、对连续值域的条件属性进行离散化,对多值表元进行单值转换。补全不完全信息(决策)表元的方法目前主要有Mean Completer算法、Combinatorial Completer算法及Roustida算法。这些算法的基本思想是选择可能的值用以填补未知的表元。另一方面,也有人尝试通过改造RST中的不可分辨关系,如利用扩展容差(相似)关系,在计算中直接处理未知表元。连续属性的离散化不是一个新课题,早在RST提出之前,就已经有了很多成果。针对Rough集特点,Nguyen H. S.和Skowron提出了布尔逻辑和Rough集理论相结合的离散化算法,并用于Rosetta系统中。

指挥决策问题中存在着一种很普遍的决策表[1],即决策表中既存在取多值的表元又有未知的表元,且表中有连续定量条件属性存在,称为连续不确定决策表。本文提出一种利用Fuzzy集理论将连续不确定决策表转化为表元单一的决策表,并用扩展Rough集理论对其进行了深入的分析。说明了连续不确定决策表可归结为多值表元决策表的原因,得出多值表元决策表的统一描述模型通过扩展Rough集理论中信息表、决策表、不可分辨关系及上下近似计算等重要概念及多值表元决策表的分析方法。

1　表元的统一描述模型

对连续不确定决策表预处理面临着离散化连续定量属性、补齐未知表元和多值表元的单值转化3个问题。本文利用Fuzzy集理论分析这3个问题中的表元得出其统一描述模型,为叙述方便,从实际问题中导出的决策表,称为原始决策表。

利用Fuzzy集理论将连续定量属性离散化,实质是通过在属性值域上建立若干个Fuzzy子集,每个Fuzzy子集用一个语义符号来标示,将原始属性值域内的每个值映射到语义符号集中的每个元素的过程。假定已给定属性a_i的语义符号集为$LS^i=(ls_1^i,ls_2^i,\cdots,ls_m^i)$,则对于$\forall ls_k^i\in LS^i(1\leqslant k\leqslant m)$,存在着$\mu_k^i:V_{ai}\to(0,1]$,隶属度$\mu_k^i(a_i(u_j))$表明$[u_j,a_i]$离散为$ls_k^i$的可能程度。但在Fuzzy离散化过程中,由于语义符号对应的Fuzzy子集其支集往往有交叠,那么在支集重叠的部分,1个属性值就可能对应着至少两个语义符号,即1个属性值以不同的或相同的隶属度对应着不止1个语义符号。一般情况下,不能确定究竟选择哪个。如果所有

收稿日期:2004-04-23
基金项目:国家"863"计划基金资助项目(86330604021)
作者简介:花文健(1976-),男,北京人,博士生,主要从事C^3I智能决策支持系统研究;
刘作良(1938-),男,四川成都人,教授,博士生导师,主要从事C^3I系统决策理论及技术研究.

Fuzzy 子集的支集完全重叠,表元$[u_j,a_i]$可表示为

$$\{(ls_1^i,\mu_1^i(a_i(u_j))),(ls_2^i,\mu_2^i(a_i(u_j))),\cdots,(ls_m^i,\mu_m^i(a_i(u_j)))\} \tag{1}$$

其中$\mu_k^i>0$,否则,μ_k^i不全大于0,$k=1,\cdots,m$。

在实际中,还存在一种更复杂的情况,即表元可能为属性值域上的 Fuzzy 值,如表元为属性值域上的三角 Fuzzy 数。同理,表元可表示为

$$\{(ls_1^i,\mu_1^i(\tilde{a}_i(u_j))),(ls_2^i,\mu_2^i(\tilde{a}_i(u_j))),\cdots,(ls_m^i,\mu_m^i(\tilde{a}_i(u_j)))\} \tag{2}$$

利用 Fuzzy 集一致性测度[2],下面给出表元分别为清晰和 Fuzzy 时的隶属度计算方法。

设U为连续不确定决策表的论域,$x\in U$;A为属性集,$a_i\in A$,$1\leqslant i\leqslant |A|=n$;$V_i$为属性$a_i$的值域,$V_i\subset \mathbf{R}$,$\mathbf{R}$是实数集;$\tilde{v}_{ih}$为$V_i$上的第$h$个 Fuzzy 子集,其隶属函数为$\mu_{\tilde{v}_{ih}}(a_i(x))$,$h=1,\cdots,m_i$,$m_i$是$V_i$上 Fuzzy 子集的数量,每个$\tilde{v}_{ih}$的语义符号相应地表示为$v_{ih}$。令$\pi(\cdot,\tilde{v}_{ih})$表示“·”隶属于语义符号$v_{ih}$的隶属度,则有:

1)$a_i(x)$为清晰值时:

$$\pi(a_i(x),\tilde{v}_{ih})=\sup_{a_i(x)\in V_i}\{\min[1,\mu_{\tilde{v}_{ih}}a_i(x)]\}=\mu_{\tilde{v}_{ih}}(a_i(x)) \tag{3}$$

2)$\tilde{a}_i(x)$为V_i上的 Fuzzy 值时:

$$\pi(\tilde{a}_i(x),\tilde{v}_{ih})=\sup_{a_i(x)\in V_i}\{\min[\mu_{\tilde{a}_i(x)}(a_i(x)),\mu_{\tilde{v}_{ih}}(a_i(x))]\}=\mu_{\tilde{v}_{ih}}(\tilde{a}_i(x)) \tag{4}$$

对于未知表元来说,补全方法实际是根据相应属性下其他已知表元确定未知表元,用使用频率最高或最有可能的表元补充未知表元,或使用全部语义符号来逐个替换计算,根据计算结果的评判准则确定其中一个语义符号的过程。补全计算的核心是从可能的语义符号集合中找出最有可能的一个语义符号。假定可以给出未知表元可能的取值集合,这个集合可以是未知表元相应属性离散语义符号集合,也可以是语义符号集合的一个子集。如:设表元$[u_j,a_i]$未知,其相应属性a_i的离散语义符号集合为$LS^i=(ls_1^i,ls_2^i,\cdots,ls_m^i)$,$[u_j,a_i]$可能取值集合为$\{ls_2^i,ls_3^i,ls_4^i\}$或$LS^i$。需要注意的是:如果属性$a_i$是连续定量的且又存在未知表元,那么对属性作 Fuzzy 离散化和作补全的先后顺序会导致LS^i本身的改变,若先做 Fuzzy 离散化,则LS^i是属性值域上的一个 Fuzzy 子集族。假定对原始决策表先进行了 Fuzzy 离散,那么可将LS^i视为一个清晰论域。用 Fuzzy 集理论,补全算法思想可被描述为:LS^i上存在一个关于未知表元$[u_j,a_i]$的 Fuzzy 集$\mu_i:LS^i\to(0,1]$。对于每个ls_k^i都存在着隶属度$\mu_i(ls_k^i)$表明用ls_k^i填补$[u_j,a_i]$的可能程度,这时表元$[u_j,a_i]$可表示为

$$\{(ls_1^i,\mu_1^i(ls_1^i)),(ls_2^i,\mu_2^i(ls_2^i)),\cdots,(ls_m^i,\mu_m^i(ls_m^i))\} \tag{5}$$

指挥决策问题的原始决策表中,多值表元现象主要发生在决策表元,决策属性值域通常为离散符号集合。设第i个决策属性为d_i,相应的符号集合为$S^i=\{S_1^i,S_2^i,\cdots,S_m^i\}$,$[u_j,d_i]$为多值表元,表现为$[u_j,d_i]=S_1^i,S_2^i,\cdots,S_m^i,\cdots$。多值表元单值化是指在已确定的多个属性值中选定可能性大的属性值的过程。由此可将多值表元扩展描述为:在符号集S^i上存在 Fuzzy 集$\mu_i:S^i\to(0,1]$,对于每个$s_k^i\in S^i(1\leqslant k\leqslant m)$都存在着隶属度$\mu_i(s_k^i)$表明$[u_j,d_i]$取值为$s_k^i$的可能程度,即表元$[u_j,d_i]$的取值可表示为

$$\{(s_1^i,\mu_i(s_1^i)),(s_2^i,\mu_i(s_2^i)),\cdots,(s_m^i,\mu_i(ls_m^i))\} \tag{6}$$

在单值表元的情况下,式(6)中仅有一个隶属度为1,其余全为0;在多值情况下,为使问题有意义,式(6)中存在至少两个大于0的隶属度。

从式(1)、(2)、(5)、(6)可看出,表元的描述是“属性值所属符号、隶属度”的组合。隶属度表明了属性值所属符号的可能程度。

定义 1　多值表元 MVE(Multi - Valued Entry)。令W_i为a_i的语义符号集。如果a_i是连续属性,那么将其转化为取值为 Fuzzy 子集$\tilde{w}_{ih}$的语义变量,相应的语义符号为$w_{ih}\in W_i$;如果a_i是离散属性,其取值亦为$w_{ih}\in W_i$,$h=1,\cdots,m_i$,$m_i=|W_i|$;$x\in U$,$a_i\in A$称

$$[x,a_i]=\{(w_{ih},\pi_{ih}(x))\mid w_{ih}\in W_i \text{ 且 } \pi_{ih}(x)>0\} \tag{7}$$

为多值表元,记为 MVE(x,a_i)。其中,

$$\pi_{ih}(x)=\begin{cases}\pi(\tilde{a}_i(x),\tilde{w}_{ih}) & a_i\text{ 值域连续 }\tilde{a}_i(x)\text{是其上的 Fuzzy 子集}\\ \pi(a_i(x),\tilde{w}_{ih}) & a_i\text{ 值域连续},a_i(x)\text{清晰}\\ \pi(a_i(x),w_{ih}) & a_i\text{ 值域离散}\end{cases} \tag{8}$$

$\pi_{ih}(x)$ 表示对象 x 用第 i 个属性的第 h 个语义符号 w_{ih} 描述时的可能程度值。

2 MVE 信息系统与决策系统的分解

对决策分类(概念)近似是利用条件属性的分划元素去近似表示决策分类的过程。因此应首先考虑 MVE 信息表及其上的分划。

2.1 信息系统的扩展

从定义 1 易见,连续不确定型决策表的表元可以统一描述为 MVE。然而,若使用经典 RST 近似原理分析还需对决策表做分解。

根据定义 1 易知,一个 MVE 信息系统 $\tilde{I}_{MVE}=\langle U,A,W,\tilde{\zeta}_{MVE}\rangle$,且应满足 $x\in U$,对于 $\forall a_i\in A$ 有:$\tilde{\zeta}_{MVE}(x,a_i)=\mathrm{MVE}(x,a_i)$,即每个表元都可视为属性 a_i的语义符号与其上可能程度值组成的组合表示,即

$$[x,a_i]=\{(w_{ih},\pi_{ih}(x))\} \tag{9}$$

为使问题有意义,式(9)满足 $\pi_{ih}(x)>0,1\leqslant h\leqslant m_i,m_i=|W_i|$。

文献[3,4]提供了信息系统单值化分解策略,即把属性 a_i 替换为属性 a_i及其语义符号的组合(a_i,w_{ih}) ,$1\leqslant h\leqslant m_i,m_i=|W_i|$ 为语义符号集合元素数,实际上就是用 m_i列(a_i,w_{ih}) 替换原信息表中一列 a_i。显然,这样做会使属性数量增加。根据 LEM2 规则[5]推导算法,属性数量增加会直接导致算法中"for 循环"个数增加,从而算法的时间复杂性呈幂指数律增加。另一方面,如果增加原信息表中对象个数则不会导致规则推导算法时间复杂性以幂指数律的速度增加(但会使空间复杂性呈线性增长)。因此,我们采用将 MVE 信息表中 MVE 对象分解为子对象来替代原对象的策略。将每个 MVE 对象 x 视为其子对象 x^j的组合,每个子对象的表元为单值表元$(w_{ih},\pi_{ih}(x^j))$。

定理 1 设 $P\subseteq A$ 为描述 x 的属性子集。$\forall q_i\in P$,q_i的语义符号集 W_i 有 $m_i>1$ 个元素;令 x^j表示 $x\in U$ 在 q_i下的子对象,$[x,q_i]$ 如式(9)所示,则表元$[x^j,q_i]=(w_{ih},\pi_{ih}(x^j))$,其中 $\pi_{ih}(x^j)=\pi_{ih}(x)$;且 x^j的个数为

1)当 P 中单个属性有多值现象时:

$$j=1,2,\cdots,p_i \qquad p_i=|\{h\mid \pi_{ih}(x)>0\}|\leqslant m_i \tag{10}$$

2)当 P 中多个属性有多值现象时:

$$j=1,2,\cdots,\prod_{i:p_i>0}p_i, \qquad p_i=|\{h\mid \pi_{ih}(x)>0\}|\leqslant m_i \tag{11}$$

定理 1 给出了子对象的表示及其数量。由此可看出,在一个属性集合 P 下,存在一个描述 x^j的向量:

$$((w_{1h},\pi_{1h}(x^j)),\cdots,((w_{ih},\pi_{ih}(x^j)),\cdots,((w_{nh},\pi_{nh}(x^j))) \tag{12}$$

利用 Yager 提出的含参数 l 的 T - 模算子,属性集 P 描述 x^j的 P - 可能程度 $\pi_p(x^j)$ 为

$$\pi_p(x^j)=1-\min\{1,[\sum_{i:q_i\in P}(1-\pi_{ih}(x^j))^l]^{1/l}\} \tag{13}$$

当 $l\to\infty$ 时,式(13) 为

$$\pi_p(x^j)=\min\{\pi_{1h}(x^j),\pi_{2h}(x^j),\cdots,\pi_{nh}(x^j)\} \tag{14}$$

其中,$n=|P|$,n 为非负整数。

根据定理 1,子对象为所属对象的一部分,换句话说,每个子对象是以一定程度表现所属对象。为了度量各个子对象在属性集 P 下对所属对象的表现程度,引人相对系数。

定义 2 设 $P\subseteq A$ 为属性集,x^j 为 x 的子对象,p 为 x^j的个数,P 描述 x^j的 P - 可能程度 $\pi_p(x^j)$,称

$$\eta_p(x^j)=\frac{\pi_p(x^j)}{\sum_{t=1}^{p}\pi_p(x^t)} \tag{15}$$

为 x^j在属性集 P 下相对于 x 的相对系数。

由此,可导出表元唯一、单值且带有隶属度的扩展信息表 $\tilde{I}$。

定义 3 扩展信息表 $\tilde{I}=\langle U',\eta,A,W,\tilde{\zeta}\rangle$ 是一个 5 元组,其中U'为子对象的全体,记为子对象论域;η 为相对系数集,即 $\eta=\{\eta_A(x')\}$,$x'\in U'$;A 为非空有限属性集合;$W=\bigcup_{i:a_i\in A}W_i$;$W_i$是属性 a_i的值域,是 a_i上

定义的定性语义符号集合；$\tilde{\zeta}$ 是扩展信息函数，满足：$x' \in U', a_i \in A, \tilde{\zeta}(x', a_i) = (w_{ih}, \pi_{ih}(x'))$。

显然，经典 RST 信息表是 $\tilde{I}$ 的一种特殊情况.

2.2 P－不可分辨关系

P－不可分辨关系使对象集合和属性集合之间建立起密切的联系，是进行近似计算的基础。

定义 4　设 $\tilde{I}$ 是扩展信息系统，对于 $P \subseteq A$，$\forall q_i \in P, x' \in U', y' \in U'$ 称

$$\tilde{R}_p = \{(x', y')\}: q_i(x') = q_i(y') = w_{ih}\} \tag{16}$$

为 U' 上的 P－可能不可分辨关系，其中，$q_i(x') = w_{ih}$ 表示对象 x' 的属性 q_i 具有属性值 w_{ih}，$w_{ih} \in W_i$。

定理 2　$\tilde{R}_p$ 是 U' 上的等价关系。

证明：可直接验证 $\tilde{R}_p$ 满足自反性，对称性和传递性。

由于 U' 上的等价关系必然决定着 U' 上的分划存在，所以，易得推论 1。

推论 1：若 $\tilde{R}_p$ 是 U' 上的 P－可能不可分辨关系，则必然产生 U' 上的一个分划：

$$A = U'/\tilde{R}_p = \{[x']_p : x' \in U'\} \tag{17}$$

其中，$[x']_p = \{y' : (x', y') \in \tilde{R}_p\} = \{y' : q_i(y') = q_i(x')(\forall q_i \in P)\}$。

推论 2：$\tilde{R}_p$ 是 U' 上的等价关系，则 $(U', \tilde{R}_p)$ 为 Pawlak 近似空间。

证明：可由近似空间定义[6]直接验证。

分划 A 的元素 $[x']_p$，称为一个 P－等价类。不难看出 P－等价类 $[x']_p$ 中的元素为具有相同语义符号 w_{ih} 的子对象，但每个子对象具有不同的可能程度值。可通过集成每个元素的可能程度值来度量用 w_{ih} 描述 $[x']_p$ 的可能程度值。

定义 5　设 $[x']_p$ 为 P－等价类，$x'_i \in [x']_p$ 被 w_{ih} 描述的可能程度值记为 $\pi_{ih}(x'_i)$，则：

$$\pi_{ih}([x']_p) = 1 - \min\{1, [\sum_{x'_i \in [x']_P} (1 - \pi_{ih}(x'_i))^l]^{1/l}\}$$

当 $l \to \infty$ 时，称：

$$\pi_{ih}([x']_p) = \min_{x'_i \in [x']_P} \pi_{ih}(x'_i) \tag{18}$$

为 $[x']_p$ 用 w_{ih} 描述的可能程度值。

定义 6　设 $[x']_p$ 为 P－等价类，子对象 $y' \in U'$ 的相对系数为 $\eta_p(y')$，称

$$\mathrm{Card}([x']_p) = \sum_{y' \in [x']_P} \eta_p(y') \tag{19}$$

为 $[x']_p$ 的基。

2.3 近似计算

定义 7　设 $(U', \tilde{R}_p)$ 为 Pawlak 近似空间，$[x']_P \in A$，对于任意 $Y \subseteq U'$，

称：

$$\underline{\tilde{R}}_p(Y) = \{x' \in U' : [x']_P \subseteq Y\} = \cup\{[x']_P : [x']_P \subseteq Y\} \tag{20}$$

为 Y 关于近似空间 $(U', \tilde{R}_p)$ 的 P－下近似；

称：

$$\overline{\tilde{R}}_p(Y) = \{x' \in U' : [x']_P \cap Y \neq \emptyset\} = \cup\{[x']_P : [x']_P \cap Y \neq \emptyset\} \tag{21}$$

为 Y 关于近似空间 $(U', \tilde{R}_p)$ 的 P－上近似；

称：

$$\tilde{B}_{np}(Y) = \overline{\tilde{R}}_p(Y) - \underline{\tilde{R}}_p(Y) \tag{22}$$

为 Y 关于近似空间 $(U', \tilde{R}_p)$ 的 P－边界。

定理 3　集合 $Y \subseteq U'$，$Y^c \subseteq U'$，满足 $Y \cap Y^c = \emptyset$ 和 $Y \cup Y^c = U'$，则

1）Y 为 $\tilde{R}_p$ Rough 可定义（或 $\tilde{R}_p$ 全不可定义）当且仅当 Y^c 为 $\tilde{R}_p$ Rough 可定义（或 $\tilde{R}_p$ 全不可定义）；

2）Y 为 $\tilde{R}_p$ 外（内）Rough 不可定义当且仅当 Y^c 为 $\tilde{R}_p$ 内（外）Rough 不可定义。

证明（略）。

定义 8　设 $(U', \tilde{R}_p)$ 为 Pawlak 近似空间，$Y \subseteq U'$ 的 P－下近似和 P－上近似分别为 $\underline{\tilde{R}}_p(Y)$ 和 $\overline{\tilde{R}}_p(Y)$，则关于近似空间 $(U', \tilde{R}_p)$ 的精度为

$$\alpha_{\tilde{R}_p}(Y)=\frac{\mathrm{Card}(\underline{\tilde{R}}_p(Y))}{\mathrm{Card}(\overline{\tilde{R}}_p(Y))} \tag{23}$$

其中,

$$\mathrm{Card}(\underline{\tilde{R}}_p(Y)) = \sum_{[x']_p\subseteq\underline{\tilde{R}}_p}\mathrm{Card}([x']_p) = \sum_{x'\subseteq\underline{\tilde{R}}_p}\eta_p(x') \tag{24}$$

$$\mathrm{Card}(\overline{\tilde{R}}_p(Y)) = \sum_{[x']_p\subseteq\overline{\tilde{R}}_p}\mathrm{Card}([x']_p) = \sum_{x'\subseteq\overline{\tilde{R}}_p}\eta_p(x') \tag{25}$$

定义 9 设 U' 为 $\tilde{I}$ 的论域,Ω 为 U' 上的分类,$\Omega=\{\omega_i\}$,$\omega_i\subseteq U'$,满足 $\omega_i\cap\omega_j=\emptyset$,$1\leqslant i\neq j\leqslant n$。$\Omega$ 的上,下近似分别为

$$\underline{\tilde{R}}_p(\Omega) = \{\underline{\tilde{R}}_p(\omega_1),\underline{\tilde{R}}_p(\omega_2),\cdots,\underline{\tilde{R}}_p(\omega_n)\} \tag{26}$$

$$\overline{\tilde{R}}_p(\Omega) = \{\overline{\tilde{R}}_p(\omega_1),\overline{\tilde{R}}_p(\omega_2),\cdots,\overline{\tilde{R}}_p(\omega_n)\} \tag{27}$$

Ω 的 P - 近似质量为

$$\gamma_{\tilde{R}_p}(\Omega) = \frac{\sum_{t=1}^{n}\mathrm{Card}(\overline{\tilde{R}}_p(\omega_t))}{\mathrm{Card}(U')} \tag{28}$$

2.4 决策系统的扩展

分类是由决策属性确定。在原始决策表中,如果对象有多值决策表元,则不能被确定地指定到某个分类中。若将带有多值决策表元的对象分解为带有可能程度单值表元的子对象,则子对象可以某种程度被唯一地指定到一个分类中。

假定原始决策表条件属性集 A 对应的表经过 2.1 节所述方法分解产生 $\tilde{I}$,即 $x\in U$ 分解为子对象 $x^j\in U'$ ($j=1,2,\cdots,p$),那么在条件属性集 A 下,x^j 的相对系数为 $\eta_A(x^j)$。设 x^j 有多值决策表元如式(6)所示,可继续将 x^j 分解为 x^{jk},$k=1,\cdots,l$。即经两次分解 x 得到 x^{jk},称 x^{jk} 为二次子对象,x 共有 $p\times l$ 个 x^{jk}。实际上,任一 x^{jk} 被描述为一个既具有条件单值表元和决策单值表元的向量。

定义 10 设 D 为决策属性集,二次子对象 x^{jk} 的相对系数为

$$\eta_D=(x^{jk})=\frac{\pi_D(x^{jk})}{\sum_{g=1}^{l}\pi_D(x^{jg})} \tag{29}$$

其中,$\pi_D(x^{jk})=\min\{\pi_{1h}(x^{jk}),\pi_{2h}(x^{jk}),\cdots,\pi_{|D|h}(x^{jk})\}$。

定义 11 设 A 为条件属性集、D 为决策属性集、子对象 x^j 的相对系数为 $\eta_A(x^j)$、二次子对象 x^{jk} 的相对系数为 $\eta_D(x^j)$,称

$$\eta_{A,D}(x^{jk})=\eta_A(x^{jk})\cdot\eta_D(x^{jk}) \tag{30}$$

为二次子对象 x^{jk} 相对于 x 的综合相对系数。

定义 12 扩展决策表 $\tilde{DT}=\langle U'',\eta,Q,W,\tilde{\zeta}\rangle$ 是一个 5 元组,其中:U'' 为二次子对象的全体,记为二次子对象论域;η 为综合相对系数集,即 $\eta=\{\eta_{A,D}(x'')\}$,$x''\in U''$;$Q=A\cup D$,且 $A\cap D=\emptyset$,A 和 D 分别为有限非空的条件属性集和决策属性集;$W=\cup_{i:q_i\in Q}W_i$;W_i 是属性 q_i 的值域,是 q_i 上的定性语义符号集合;$\tilde{\zeta}$ 是扩展信息函数,即 $\tilde{\zeta}(x'',q_i)=(w_{ih},\pi_{ih}(x''))$,表示二次子对象 x'' 用第 i 个属性 q_i 的第 h 个语义符号描述的可能程度值为 $\pi_{ih}(x'')$。

3 结论

连续不确定决策表是利用 Rough 集理论设计和实现智能指挥决策支持系统时常见的决策表。本文将这种决策表预处理为扩展决策表,并给出了计算扩展决策表中决策概念的下近似和边界的方法,这样就为规则推导算法提取知识提供了输入数据。需要注意的是,基于经典 Rough 集决策表的规则推导算法,如 LEM 或 LEM2 算法,其涉及的基本概念,如覆盖、最小规则集等,都需在本文所述扩展 Rough 集方法上做进一步扩展后才能利用上述下近似和边界推导出决策规则。这是我们正在研究的问题。

参考文献:

[1] 花文健,刘作良,杨　凡. 一种新的知识表示模型及其规则推导[J]. 空军工程大学学报(自然科学版),2003,6(4):44 - 47.

[2] Zadeh L A. Fuzzy Sets as A Basis for Theory of Possibility [J]. Fuzzy Sets and Systems, 1978,(1):3 - 28.

[3] Fernandez J M, Salido, Murakami S. Rough Set Analysis of A General Type of Fuzzy Data Using Transitive Aggregation of Fuzzy Similarity Relations [J]. Fuzzy Sets and Systems, 2003, 139: 635 - 660.

[4] Slavka Bodjanova. Approximation of Fuzzy Concepts in Decision Making [J]. Fuzzy Sets and Systems, 1997, 85:23 - 29.

[5] Jerzy Stefanowski. On Rough Set Bsed Approaches to Induction of Decision Rules[A]. Lech Polkowski, Andrzej Skowron. Rough Sets in Knowledge Discovery[C]. Heidelberg: Physica - Verlag, 1998,500 - 529.

[6] 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.

[7] 刘普寅,吴孟达. 模糊理论及其应用[M]. 长沙:国防科技大学出版社,2000.

(编辑:门向生)

An Approach to Approximation of Decision Concepts of Continuous Uncertain Decision System

HUA Wen - jian, LIU Zuo - liang, HAN Bing

(The Telecommunication Engineering Institute, Air Force Engineering University, Xi'an, Shaanxi 710077, China)

Abstract: There are always various types of entries in DT from complex decision problem of command and control, that is, entries with a continuous quantitative attribute, unknown entries or multi - valued entries, and the three types of entries often occur in the same DT. The DT with these entries named Continuous Uncertain Decision Table (CUDT) can not be analyzed directly by CRST. Fortunately, by modeling those three types of entries in CUDT with Fuzzy Sets theory, CUDT can be transformed into a special DT called Extended Decision Table (EDT) in which each entry is associated with a membership degree. An extended CRST based on this idea is proposed to transform the CUDT into EDT, and the lower approximations and the boundaries of decision concepts in EDT are also calculated by the extended CRST.

Key words: continuous uncertain decision system; extended Rough sets; concept approximation

参考文献:

[1] 花文健,刘作良,杨　凡. 一种新的知识表示模型及其规则推导[J]. 空军工程大学学报(自然科学版),2003,6(4):44 - 47.

[2] Zadeh L A. Fuzzy Sets as A Basis for Theory of Possibility [J]. Fuzzy Sets and Systems, 1978,(1):3 - 28.

[3] Fernandez J M, Salido, Murakami S. Rough Set Analysis of A General Type of Fuzzy Data Using Transitive Aggregation of Fuzzy Similarity Relations [J]. Fuzzy Sets and Systems, 2003, 139: 635 - 660.

[4] Slavka Bodjanova. Approximation of Fuzzy Concepts in Decision Making [J]. Fuzzy Sets and Systems, 1997, 85:23 - 29.

[5] Jerzy Stefanowski. On Rough Set Bsed Approaches to Induction of Decision Rules[A]. Lech Polkowski, Andrzej Skowron. Rough Sets in Knowledge Discovery[C]. Heidelberg: Physica - Verlag, 1998,500 - 529.

[6] 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.

[7] 刘普寅,吴孟达. 模糊理论及其应用[M]. 长沙:国防科技大学出版社,2000.

(编辑:门向生)

An Approach to Approximation of Decision Concepts of Continuous Uncertain Decision System

HUA Wen - jian, LIU Zuo - liang, HAN Bing

(The Telecommunication Engineering Institute, Air Force Engineering University, Xi'an, Shaanxi 710077, China)

Abstract: There are always various types of entries in DT from complex decision problem of command and control, that is, entries with a continuous quantitative attribute, unknown entries or multi - valued entries, and the three types of entries often occur in the same DT. The DT with these entries named Continuous Uncertain Decision Table (CUDT) can not be analyzed directly by CRST. Fortunately, by modeling those three types of entries in CUDT with Fuzzy Sets theory, CUDT can be transformed into a special DT called Extended Decision Table (EDT) in which each entry is associated with a membership degree. An extended CRST based on this idea is proposed to transform the CUDT into EDT, and the lower approximations and the boundaries of decision concepts in EDT are also calculated by the extended CRST.

Key words: continuous uncertain decision system; extended Rough sets; concept approximation