

序区间值决策系统中最优可信规则获取与约简*

王天擎¹, 谢 军²

(1. 五邑大学 经济管理学院, 广东 江门 529020; 2. 江苏大学 计算机科学与通信工程学院, 江苏 镇江 212013)

摘要: 基于描述子的规则获取可导出序值决策系统中的所有可信规则,但对包含区间值序决策系统却不能有效支持。因此,首先根据每个属性值域的范围,提出了一个区间段值的概念,用以将序区间值决策系统转换为序区间段值决策系统;然后,在序区间段值决策系统中提出了基于区间段值的优势和弱势描述子概念,用以导出序区间值决策系统中的所有可信规则;最后,研究了两种新的描述子的约简以及相对约简问题,给出了相应的判定定理与区分函数。以上为从序区间值决策系统中获取有效的最优可信决策规则提供了一种新理论基础与操作手段。

关键词: 粗糙集; 区间值; 优势(弱势)描述子; 约简; 可信规则

中图分类号: TP18

文献标志码: A

文章编号: 1001-3695(2012)12-4482-04

doi:10.3969/j.issn.1001-3695.2012.12.020

Acquisition of all credible rules and reductions in ordered interval-valued decision systems

WANG Tian-qing¹, XIE Jun²

(1. School of Economics & Management, Wuyi University, Jiangmen Guangdong 529020, China; 2. School of Computer Science & Telecommunication Engineering, Jiangsu University, Zhenjiang Jiangsu 212013, China)

Abstract: In the ordered decision system, it can induce all of the credible rules based on the descriptors, but which cannot direct support for interval value. To solve this problem, this paper proposed a new concept of segmented interval value based on the range of domain of each condition attributes, then the ordered interval-valued decision system could be changed to an ordered segmented interval-valued decision system. Secondly, based on the segmented interval value, it proposed the concepts of the increasing and decreasing descriptors. Finally, it investigated the reductions and relative reductions in terms of the decision classes of these two new descriptors, and gave the judgment theorems and discernibility functions with respect to these reductions. These methods provide a theory and practical approach, which can generate all compact and effective optimal credible decision rules from the old ordered decision system including interval data.

Key words: rough sets; interval data; increasing(decreasing) descriptors; reduction; credible rules

0 引言

波兰科学家 Pawlak 等人^[1,2]开创了 Rough 集理论的先河,为处理不精确、不确定、不完备信息提供了新的数学工具。但经典的粗糙集理论是基于等价关系的,而在多准则决策问题中,描述对象的属性常常具有顺序性。为了解决这个问题, Greco 等人^[3]在序值信息系统中提出了基于优势关系的粗糙集模型。灰色系统理论最先由我国数学家 Deng^[4]提出,其中包括了灰色分类、灰色决策和灰色关系分析等内容。类似于粗糙近似空间和决策规则生成对区间值的处理正处于研究起步阶段^[5-8],因此,对于含区间值信息系统的研究有着重要的理论意义与实用价值。

同时,针对集值信息系统, Leung 等人^[9]提出了基于描述子的粗糙集模型。Wu 等人^[10]证明了使用基于描述子的粗糙集模型可以比基于容差关系的粗糙集模型获得更高的粗糙近似精度。Yang 等人^[11,12]证明在不完备序值决策系统中基于描述子可以导出所有可信规则,但对含区间值却还不能有效支持。

本文以包含区间值的序值决策系统为研究对象。首先根据描述子对区间值不能有效支持以及每个属性取值区间,提出了新的区间段值概念,用以将包含区间值的序值决策系统转换为序区间段值决策系统,并给出了其上的优势关系;然后,在序区间段值决策系统中提出了基于区间段值的优势和弱势描述子,用以获取其所有可信规则;最后,详细讨论了优势和弱势描述子的约简,更进一步讨论两种描述子关于决策类的相对约简问题,并给出了相应的判定定理与区分函数。以上工作最终为从原复杂序区间值决策系统中获取所有简洁、有效的最优可信规则提供了新的理论方法。

1 概述

1.1 序值决策系统

一个序值决策系统是四元组 $ODS = \langle U, AT \cup D, V, f \rangle$ 。其中: U 是一个被称为论域的非空有限对象集合; AT 为序条件属性集合, D 为序决策属性集合; V_{AT} 是所有条件属性的值域集合, V_D 是所有决策属性的值域集合,并且 V_{AT} 和 V_D 都是全序

收稿日期: 2012-04-16; 修回日期: 2012-05-25 基金项目: 江苏省高校自然科学基金资助项目(10KJB520004)

作者简介: 王天擎(1974-),男,湖南长沙人,副教授,硕士,主要研究方向为粗糙集、智能信息处理及数据挖掘(xinfuwtq@163.com); 谢军(1973-),男,湖南衡阳人,副教授,硕士,主要研究方向为粗糙集理论及其应用。

集, $V = V_{AT} \cup V_D$ 表示全体属性的值域集合; 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^[4] 在序值决策系统 ODS 中, 对于 $\forall A \subseteq AT$, 由 A 决定的优势关系记为 $DM(A)$ 且

$$DM(A) = \{(x, y) \in U^2 : \forall a \in A, f(x, a) \geq f(y, a)\} \quad (1)$$

在序值决策系统中, 决策属性集合 D 构成了论域上的划分 $CL = \{CL_t : t \in T\}$, 其中 $T = \{1, 2, \dots, n\}$ 。对于 $\forall r, s \in T$, 若 $r > s$, 则认为 CL_r 中的对象要优于 CL_s 中的对象。令 $CL_t^{\geq} = \bigcup_{s \geq t} CL_s, CL_t^{\leq} = \bigcup_{s \leq t} CL_s, t = 1, \dots, n; x \in CL_t^{\geq}$ 表示 x 至少属于决策类 CL_t ; $x \in CL_t^{\leq}$ 表示 x 至多属于决策类 CL_t 。

定义 2^[4] 在序值决策系统 ODS 中, 有 $A \subseteq AT$, 对于 $\forall CL_t^{\geq} (1 \leq t \leq n)$, $CL_t^{\geq}$ 的下、上近似集记为 $A_*(CL_t^{\geq}), A^*(CL_t^{\geq})$ 且

$$\begin{aligned} A_*(CL_t^{\geq}) &= \{x \in U : D_A^+(x) \subseteq CL_t^{\geq}\} \\ A^*(CL_t^{\geq}) &= \{x \in U : D_A^-(x) \cap CL_t^{\geq} \neq \emptyset\} \end{aligned} \quad (2)$$

对于 $\forall CL_t^{\leq} (1 \leq t \leq n)$, $CL_t^{\leq}$ 的下、上近似集记为 $A_*(CL_t^{\leq}), A^*(CL_t^{\leq})$ 且

$$\begin{aligned} A_*(CL_t^{\leq}) &= \{x \in U : D_A^-(x) \subseteq CL_t^{\leq}\} \\ A^*(CL_t^{\leq}) &= \{x \in U : D_A^+(x) \cap CL_t^{\leq} \neq \emptyset\} \end{aligned} \quad (3)$$

其中: $D_A^+(x) = \{y \in U : (y, x) \in DM(A)\}, D_A^-(x) = \{y \in U : (x, y) \in DM(A)\}$ 。

1.2 描述子

定义 3 令 ODS 为一序值决策系统:

a) 对于 $\forall a \in AT, \forall v \in V_a$, 称 $(\geq_{a_i}, v)$ 为一个 AT —优势原子公式;

b) 任意一个 AT —优势原子公式或几个 AT —优势原子公式的合取称为一个 AT —优势描述子;

c) 若 $\bigwedge_{a_i \in A} (\geq_{a_i}, v_i)$ 为一个 AT —优势描述子, 令

$$\|\bigwedge_{a_i \in A} (\geq_{a_i}, v_i)\| = \{x \in U : \forall a_i \in A, f(x, a_i) \geq v_i\}$$

则称 $\|\bigwedge_{a_i \in A} (\geq_{a_i}, v_i)\|$ 为优势描述子 $\bigwedge_{a_i \in A} (\geq_{a_i}, v_i)$ 的支持集;

d) 若 s 是一个 AT —优势描述子, 则 s 中出现的所有属性的集合记为 $AT(s)$, 其中:

$$A = \{a_1, a_2, \dots, a_m\} \subseteq AT, \{v_1, v_2, \dots, v_m\} \in V_{a_1} \times V_{a_2} \times \dots \times V_{a_m}, 1 \leq i \leq m$$

定义 4 令 ODS 为一序值决策系统, 记

$$DES(\geq_{AT}) = \{s : s \text{ 是 } AT\text{—优势描述子}, \|s\| \neq \emptyset\}$$

$$DES(\leq_{AT}) = \{o : o \text{ 是 } AT\text{—弱势描述子}, \|o\| \neq \emptyset\}$$

对于 $\forall s \in DES(\geq_{AT})$, 若 $AT(s) = AT$, 则称 s 为一个完全 AT —优势描述子; 对于 $\forall o \in DES(\leq_{AT})$, 若 $AT(o) = AT$, 则称 o 为一个完全 AT —弱势描述子, 记

$$FDES(\geq_{AT}) = \{s : s \in DES(\geq_{AT}), s \text{ 是完全 } AT\text{—优势描述子}\}$$

$$FDES(\leq_{AT}) = \{o : o \in DES(\leq_{AT}), o \text{ 是完全 } AT\text{—弱势描述子}\}$$

2 基于区间段值的序值决策系统

定义 5 设 R 表示实数集, 对 $\forall x$ 的两个值 $\underline{x}, \bar{x} \in R (\underline{x} = \inf X, \bar{x} = \sup X)$, 如果 $\underline{x} \leq \bar{x}$, 则 x 可以定义为

$$\otimes x = [\underline{x}, \bar{x}] = \{y : y \in R, \underline{x} \leq y \leq \bar{x}\}$$

则称闭区间 $\otimes x$ 为一个区间值 (灰数)。全体区间值组成的集合记为 $I(R)$ 。

定义 6 一个含区间值的序值决策系统被称为序区间值

决策系统, 其为一个四元组: $OIDS = \langle U, AT \cup d, V, f_{\otimes} \cup f_d \rangle$ 。对于 $\forall a \in AT$, 有 $U \times AT \rightarrow V_a, f_{\otimes}(x, a) = [\underline{f}_{\otimes}(x, a), \bar{f}_{\otimes}(x, a)]$, 其中 $(x, a) \in \otimes AT, \underline{f}_{\otimes}(x, a) = \inf(V_a)$ 且 $\bar{f}_{\otimes}(x, a) = \sup(V_a)$ 。

定义 7 设 OIDS 为一个序区间值决策系统, 对 $\forall a \in AT, \forall x \in U$, 设 x 在 a 上的取值范围在一个固定区间范围内, 即最小值为 $(\underline{SD}_a)$ 且最大值为 $(\overline{SD}_a)$ 之间时, 则称区间值 $\otimes SD_a$ 为所有对象在属性 a 上的固定取值区间值, 记

$$\otimes SD_a = [\underline{SD}_a, \overline{SD}_a] \quad (4)$$

定义 8 设 OIDS 为一序区间值决策系统, $\forall a \in AT, \forall x \in U$, 对属性 a 上的取值范围为一设定的固定区间值 $\otimes SD = [\underline{SD}_a, \overline{SD}_a]$, 则可以选取一合适正整数 $N_a (N_a > 0)$ 对 $\otimes SD$ 进行等长分段, 本文将每一个小分段区间称为一个区间段值。记 SD_a 为 $\otimes SD$ 所有区间段值的集合, SD_a^i 为第 i 个分段的区间段值, 则有

$$\begin{aligned} SD_a &= \{SD_a^1, SD_a^2, \dots, SD_a^{N_a}\} = \{[\underline{SD}_a, \underline{SD}_a + \text{Len}_a], \\ &[\underline{SD}_a + \text{Len}_a, \underline{SD}_a + 2 \times \text{Len}_a], \dots, [\underline{SD}_a + (N_a - 1) \times \\ &\text{Len}_a, \underline{SD}_a + N_a \times \text{Len}_a]\} \end{aligned} \quad (5)$$

其中: $\text{Len}_a = (\overline{SD}_a - \underline{SD}_a) / N_a$ 。

因此, 根据定义 7, 本文将完全由区间段值所组成的序区间值决策系统称为序区间段值决策系统—OISDS。同时, 为了将最初的序区间值决策系统—OIDS 有效地转换成序区间段值决策系统—OISDS, 下面引入灰白化原理。

定义 9 对于任一区间灰数值 $\otimes x \in [a, b]$, 将其白化值 $\tilde{\otimes} x$ 取为

$$\tilde{\otimes} x = \beta a + (1 - \beta) b, \beta \in [0, 1] \quad (6)$$

定义 10 形如 $\tilde{\otimes} x = \beta a + (1 - \beta) b, \beta \in [0, 1]$ 的白化称为等权白化。在等权白化中, 若取 $\beta = 1/2$ 而得到的白化值称为等权均值白化。

在现实生活中, 许多区间值都在某个基本值附近变动, 特别在区间灰数取值的分布信息缺乏时, 常采用等权均值白化, 而这也符合人们对区间值自然直观的解释。本文中对区间值的白化均采用比较直观、合理的等权均值白化。

因此, 根据定义 10, 将原序区间值决策系统—OIDS 中所有的区间值转换为等权均值白化值后, 即可得到一个序白化值决策系统—OWDS = $\langle U, AT \cup d, V, f_{\otimes}^{1/2} \cup f_d \rangle$ 。

定义 11 设 OWDS 为一序白化值决策系统, 对 $\forall a \in AT, \forall x \in U$, 若有

$$\begin{aligned} f_{\otimes}^{1/2}(x, a) \in SD_a^i &= [\underline{SD}_a + (i - 1) \times \text{Len}_a, \\ &\underline{SD}_a + i \times \text{Len}_a] \quad i = 1, 2, 3, \dots, n \end{aligned} \quad (7)$$

则记

$$\begin{aligned} f_{\otimes}(x, a) = SD_a^i &= [\underline{SD}_a + (i - 1) \times \text{Len}_a, \\ &\underline{SD}_a + i \times \text{Len}_a] \quad i = 1, 2, 3, \dots, n \end{aligned} \quad (8)$$

其中, 当同时出现 $f_{\otimes}^{1/2}(x, a) \in SD_a^i$ 且 $f_{\otimes}^{1/2}(x, a) \in SD_a^{i+1}$ 时, 本文规定 $f_{\otimes}^{1/2}(x, a) \in SD_a^{i+1}$ 。

根据定义 11, 可将序白化值决策系统—OWDS 中所有的白化值都转换为区间段值, 从而构建了一个序区间段值决策系统—OISDS = $\langle U, AT \cup d, V, f_{\otimes}^{SD} \cup f_d \rangle$ 。

定义 12 在序区间段值决策系统 OISDS 中, 对于 $\forall A \subseteq AT$, 由 A 决定的优势关系记为 $M(A)$ 且

$$M(A) = \{(x, y) \in U^2 : \forall a \in A, \overline{f_{\otimes}^{SD}(x, a)} \geq \overline{f_{\otimes}^{SD}(y, a)}\} \quad (9)$$

显然, M 是自反的和传递的, 但未必是对称的。

例1 表1为一序区间值决策系统, 其中 $AT = \{a_1, a_2, a_3\}$ 为序条件属性集合, d 为序决策属性, 则有 $CL = \{CL_1, CL_2\}$, 其中, $CL_1 = \{x_2, x_4, x_5, x_{10}\}$, $CL_2 = \{x_1, x_3, x_6, x_7, x_8, x_9\}$, 即 $CL_1^{\geq} = U, CL_2^{\geq} = CL_2, CL_1^{\leq} = CL_1, CL_2^{\leq} = U$ 。

表1 序区间值决策系统

U	a_1	a_2	a_3	d
x_1	[3.3, 3.9]	[1.6, 3.0]	[1.1, 1.9]	2
x_2	[1.2, 2.1]	[2.3, 2.8]	[1.3, 2.2]	1
x_3	[2.1, 2.8]	[1.2, 1.8]	[1.8, 3.0]	2
x_4	[1.2, 2.1]	[1.1, 1.6]	[1.1, 1.8]	1
x_5	[1.3, 1.7]	[2.2, 2.9]	[1.2, 1.9]	1
x_6	[2.9, 3.7]	[1.4, 2.1]	[1.7, 1.9]	2
x_7	[2.1, 2.9]	[1.1, 1.2]	[1.3, 1.9]	2
x_8	[2.7, 3.7]	[2.2, 2.9]	[1.8, 2.9]	2
x_9	[2.2, 2.7]	[1.2, 2.2]	[2.1, 2.7]	2
x_{10}	[1.0, 1.2]	[1.4, 1.9]	[1.3, 1.8]	1

当 $CL_1^{\geq} = U$ 和 $CL_2^{\leq} = U$ 时, 有 $AT_*(CL_1^{\geq}) = AT_*(CL_2^{\leq}) = U$, 此时考虑任意对象在部分或者所有属性上的取值, 它的决策属性值必然是大于等于1或者小于等于2的, 因而根据这样的下近似集所导出的决策规则对于决策者来说没有意义。所以这两种情况笔者不予讨论。

若在表1的序区间值决策系统中, 设属性的固定取值区间值 $\otimes SD_{a_1} = [1.0, 4.0]$ 且 $\otimes SD_{a_2} = SD_{a_3} = [1.0, 3.0]$, 则设 $N_{a_1} = 3$ 且 $N_{a_2} = N_{a_3} = 2$, 根据定义8计算可得

$$\begin{aligned} SD_{a_1} &= \{SD_{a_1}^1, SD_{a_1}^2, SD_{a_1}^3\} = \{[1.0, 2.0], [2.0, 3.0], [3.0, 4.0]\} \\ SD_{a_2} &= \{SD_{a_2}^1, SD_{a_2}^2\} = \{[1.0, 2.0], [2.0, 3.0]\} \\ SD_{a_3} &= \{SD_{a_3}^1, SD_{a_3}^2\} = \{[1.0, 2.0], [2.0, 3.0]\} \end{aligned}$$

最后, 根据定义10、11, 及由以上计算的各属性区间段值即可构建表1的序区间段值决策系统, 如表2所示。

表2 序区间段值决策系统

U	a_1	a_2	a_3	d
x_1	[3.0, 4.0]	[2.0, 3.0]	[1.0, 2.0]	2
x_2	[1.0, 2.0]	[2.0, 3.0]	[1.0, 2.0]	1
x_3	[2.0, 3.0]	[1.0, 2.0]	[2.0, 3.0]	2
x_4	[1.0, 2.0]	[1.0, 2.0]	[1.0, 2.0]	1
x_5	[1.0, 2.0]	[2.0, 3.0]	[1.0, 2.0]	1
x_6	[3.0, 4.0]	[1.0, 2.0]	[1.0, 2.0]	2
x_7	[2.0, 3.0]	[1.0, 2.0]	[1.0, 2.0]	2
x_8	[3.0, 4.0]	[2.0, 3.0]	[2.0, 3.0]	2
x_9	[2.0, 3.0]	[1.0, 2.0]	[2.0, 3.0]	2
x_{10}	[1.0, 2.0]	[1.0, 2.0]	[1.0, 2.0]	1

3 序区间段值决策系统中基于描述子的可信规则获取

定义13 设 OISDS 为一序区间段值决策系统, 若 $\bigwedge_{a_i \in AT} (\geq_{a_i}, v_i)$ 为一个基于区间段值的完全 AT —优势描述子, 则若 $\| \bigwedge_{a_i \in AT} (\geq_{a_i}, v_i) \| \neq \emptyset$, 且有 $\| \bigwedge_{a_i \in AT} (\geq_{a_i}, v_i) \| \subseteq CL_i^{\geq} (2 \leq i \leq n)$, 则称

$$f(x, a_1) \geq v_1 \wedge f(x, a_2) \geq v_2 \wedge \cdots \wedge f(x, a_m) \geq v_m \rightarrow x \in CL_i^{\geq}$$

为一条优势描述子 $\bigwedge_{a_i \in AT} (\geq_{a_i}, v_i)$ 的可信决策规则, 记为 $\bigwedge_{a_i \in AT} (\geq_{a_i}, v_i) \rightarrow x \in CL_i^{\geq}$, 对于 $\forall x \in \bigwedge_{a_i \in AT} (\geq_{a_i}, v_i)$, 称 x 支持规则 $\bigwedge_{a_i \in AT} (\geq_{a_i}, v_i) \rightarrow x \in CL_i^{\geq}$ 。其中 $1 \leq i \leq m$ 。

例2 在表2所示的序区间段值决策系统中, 有 $V_{a_1} =$

$\{[1.0, 2.0], [2.0, 3.0], [3.0, 4.0]\}$, $V_{a_2} = V_{a_3} = \{[1.0, 2.0], [2.0, 3.0]\}$, $V_d = \{1, 2\}$ 。根据定义13, 使用基于区间段值的优势描述子方法可计算所有完全 AT —优势描述子, 得

$$\begin{aligned} s_1 &= (\geq_{a_1}, [1.0, 2.0]) \wedge (\geq_{a_2}, [1.0, 2.0]) \wedge (\geq_{a_3}, [1.0, 2.0]), \|s_1\| = \{x_1, x_2, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}\} \\ s_2 &= (\geq_{a_1}, [1.0, 2.0]) \wedge (\geq_{a_2}, [1.0, 2.0]) \wedge (\geq_{a_3}, [2.0, 3.0]), \|s_2\| = \{x_3, x_8, x_9\} \\ s_3 &= (\geq_{a_1}, [1.0, 2.0]) \wedge (\geq_{a_2}, [2.0, 3.0]) \wedge (\geq_{a_3}, [1.0, 2.0]), \|s_3\| = \{x_1, x_2, x_5, x_8\} \\ s_4 &= (\geq_{a_1}, [1.0, 2.0]) \wedge (\geq_{a_2}, [2.0, 3.0]) \wedge (\geq_{a_3}, [2.0, 3.0]), \|s_4\| = \{x_8\} \\ s_5 &= (\geq_{a_1}, [2.0, 3.0]) \wedge (\geq_{a_2}, [1.0, 2.0]) \wedge (\geq_{a_3}, [1.0, 2.0]), \|s_5\| = \{x_1, x_3, x_6, x_7, x_8, x_9\} \\ s_6 &= (\geq_{a_1}, [2.0, 3.0]) \wedge (\geq_{a_2}, [1.0, 2.0]) \wedge (\geq_{a_3}, [2.0, 3.0]), \|s_6\| = \{x_3, x_8, x_9\} \\ s_7 &= (\geq_{a_1}, [2.0, 3.0]) \wedge (\geq_{a_2}, [2.0, 3.0]) \wedge (\geq_{a_3}, [1.0, 2.0]), \|s_7\| = \{x_1, x_8\} \\ s_8 &= (\geq_{a_1}, [2.0, 3.0]) \wedge (\geq_{a_2}, [2.0, 3.0]) \wedge (\geq_{a_3}, [2.0, 3.0]), \|s_8\| = \{x_8\} \\ s_9 &= (\geq_{a_1}, [3.0, 4.0]) \wedge (\geq_{a_2}, [1.0, 2.0]) \wedge (\geq_{a_3}, [1.0, 2.0]), \|s_9\| = \{x_1, x_6, x_8\} \\ s_{10} &= (\geq_{a_1}, [3.0, 4.0]) \wedge (\geq_{a_2}, [1.0, 2.0]) \wedge (\geq_{a_3}, [2.0, 3.0]), \|s_{10}\| = \{x_8\} \\ s_{11} &= (\geq_{a_1}, [3.0, 4.0]) \wedge (\geq_{a_2}, [2.0, 3.0]) \wedge (\geq_{a_3}, [1.0, 2.0]), \|s_{11}\| = \{x_1, x_8\} \\ s_{12} &= (\geq_{a_1}, [3.0, 4.0]) \wedge (\geq_{a_2}, [2.0, 3.0]) \wedge (\geq_{a_3}, [2.0, 3.0]), \|s_{12}\| = \{x_8\} \end{aligned}$$

根据定义13基于优势描述子的可信决策规则描述, 可导出如下可信决策规则:

$$\begin{aligned} R_1: f(x, a_1) &\geq [1.0, 2.0] \wedge f(x, a_2) \geq [1.0, 2.0] \wedge f(x, a_3) \geq [2.0, 3.0] \rightarrow f(x, d) \geq 2 \\ R_2: f(x, a_1) &\geq [2.0, 3.0] \wedge f(x, a_2) \geq [1.0, 2.0] \wedge f(x, a_3) \geq [1.0, 2.0] \rightarrow f(x, d) \geq 2 \\ R_3: f(x, a_1) &\geq [3.0, 4.0] \wedge f(x, a_2) \geq [1.0, 2.0] \wedge f(x, a_3) \geq [1.0, 2.0] \rightarrow f(x, d) \geq 2 \end{aligned}$$

4 序区间段值决策系统中基于描述子的最优可信规则

4.1 基于描述子的约简

定义14 设 OISDS 为一序区间段值决策系统, $s \in \text{FDES}(\geq_{AT})$, 对于 $\forall s' \in \text{DES}(\geq_{AT})$, 若 s' 满足

$$s' \succeq s, \|s'\| = \|s\|, \text{对于 } \forall s'' \succ s', \|s''\| \neq \|s\|$$

则称 s' 是 s 的一个约简优势描述子。

对于一个基于区间段值的 AT —优势描述子 s 来说, 它的约简优势描述子是保持 $\|s\|$ 中的元素不发生变化的最小数量原子公式的合取, $\|s\|$ 的所有约简优势描述子记为 $md(s)$ 。

若 $s \in \text{FDES}(\geq_{AT})$, 不妨假设 $s = \bigwedge_{a_i \in AT} (\geq_{a_i}, v_i)$, 对于 $\forall x \in U$, 令

$$D_{AT}(s, x) = \begin{cases} \{a_i \in AT: f(x, a_i) < v_i\} & x \notin \|s\| \\ AT: x \in \|s\| \end{cases}$$

则称 $D_{AT}(s, x)$ 为优势描述子 s 与对象 x 的分辨属性集, 称 $D_{AT}(s) = \{D_{AT}(s, x): s \in \text{FDES}(\geq_{AT}), x \in U\}$ 为优势描述子 s 的分辨矩阵。

定理 1 设 OISDS 为一序区间段值决策系统,若 $s \in \text{FDES}(\geq_{AT})$, $s' \in \text{DES}(\geq_{AT})$ 且 $s' \geq s$, 那么就有

$$\|s'\| = \|s\| \Leftrightarrow D_{AT}(s, x) \cap AT(s') \neq \emptyset (\forall x \in U, x \notin \|s\|)$$

证明 必要性。对于 $\forall x \in U$, 假设 $x \notin \|s\|$, 那么根据 $\|s'\| = \|s\|$ 有 $x \notin \|s'\|$ 。因此, 必存在 $a_j \in AT(s') \subseteq AT$ 使得 $f(x, a_j) < v_j$, 这里 $v_j \in V_{a_j}$ 且 $(\geq_{a_j}, v_j) \in s'$ 。由已知条件 $s' \geq s$, 可得 $a_j \in D_{AT}(s, x)$, 从而推出 $D_{AT}(s, x) \cap AT(s') \neq \emptyset$ 。

充分性。因为 $s' \geq s$, 所以很明显有 $\|s'\| \supseteq \|s\|$, 故只需证明 $\|s'\| \subseteq \|s\|$ 即可。对于 $\forall x \in U, \forall x \notin \|s\|$, 根据假设 $D_{AT}(s, x) \cap AT(s') \neq \emptyset$, 所以必定存在 $a_j \in AT(s')$ 使得 $f(x, a_j) < v_j$, 这里 $v_j \in V_{a_j}$ 且 $(\geq_{a_j}, v_j) \in s'$, 故 $x \notin \|s'\|$, 即 $\|s'\| \subseteq \|s\|$ 。

定义 15 设 OISDS 为一序区间段值决策系统, $s \in \text{FDES}(\geq_{AT})$, 令

$$\Delta(s) = \bigwedge_{x \in U} (\vee D_{AT}(s, x))$$

称 $\Delta(s)$ 为优势描述子 s 的区分函数。

根据布尔推理理论, 由定义 14 及定理 1 可得如下结论。

定理 2 设 OISDS 为一序区间段值决策系统, $\Delta(s)$ 的极小析取范式中的每个合取项对应的属性子集即为 s 的约简优势描述子中出现的属性集合。

例 3 在表 2 所示的序区间段值决策系统中, 基于区间段值优势描述子的区分矩阵如表 3 所示。

表 3 表 2 中优势描述子的区分矩阵

U	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}
s_1	AT	AT	AT	AT	AT	AT	AT	AT	AT	AT
s_2	a_3	a_3	AT	a_3	a_3	a_3	a_3	AT	AT	a_3
s_3	AT	AT	a_2	a_2	AT	a_2	a_2	AT	a_2	a_2
s_4	a_3	a_3	a_2	a_2, a_3	a_3	a_2, a_3	a_2, a_3	AT	a_2	a_2, a_3
s_5	AT	a_1	AT	a_1	a_1	AT	AT	AT	AT	a_1
s_6	a_3	a_1, a_3	AT	a_1, a_3	a_1, a_3	a_3	a_3	AT	AT	a_1, a_3
s_7	AT	a_1	a_2	a_1, a_2	a_1	a_2	a_2	AT	a_2	a_1, a_2
s_8	a_3	a_1, a_3	a_2	a_1, a_2, a_3	a_1, a_3	a_2, a_3	a_2, a_3	AT	a_2	a_1, a_2, a_3
s_9	AT	a_1	a_1	a_1	a_1	AT	a_1	AT	a_1	a_1
s_{10}	a_3	a_1, a_3	a_1	a_1, c	a_1, a_3	a_3	a_1, a_3	AT	a_1	a_1, a_3
s_{11}	AT	a_1	a_1, a_2	a_1, a_2	a_1	a_2	a_1, a_2	AT	a_1, a_2	a_1, a_2
s_{12}	a_3	a_1, a_3	a_1, a_2	a_1, a_2, a_3	a_1, a_3	a_2, a_3	a_1, a_2, a_3	AT	a_1, a_2	a_1, a_2, a_3

4.2 描述子的相对约简与决策规则

定义 16 设 OISDS 为一序区间段值决策系统, 优势描述子 $s \in \text{FDES}(\geq_{AT})$, 若 $\exists CL_n^{\geq} (2 \leq n \leq m)$ 且 $\|s\| \subseteq CL_n^{\geq}$, 则称优势描述子 s 在 $CL_n^{\geq}$ 中是相容的。

定义 17 设 OISDS 为一序区间段值决策系统, 设优势描述子 $s \in \text{FDES}(\geq_{AT})$ 且 s 在 $CL_n^{\geq}$ 中是相容的, 若 $s' \geq s$ 且 s' 在 $CL_n^{\geq}$ 中是相容的, 称 s' 为 s 的相对相容优势描述子; 若 s' 是 s 的相对相容优势描述子且 $\forall s'' > s'$, s'' 不是 s 的相对相容优势描述子, 称 s' 为 s 的相对约简优势描述子。

显然, 若 s' 是 s 的相对相容优势描述子, 那么称 $s' \rightarrow x \in CL_n^{\geq}$ 为 $s \rightarrow x \in CL_n^{\geq}$ 的约简优势可信规则, 其中 $s \rightarrow x \in CL_n^{\geq}$ 为由优势描述子 s 导出的初等可信规则; 若 s' 是 s 的相对约简优势描述子, 那么称 $s' \rightarrow x \in CL_n^{\geq}$ 为最优优势可信规则。

定理 3 设 OISDS 为一序区间段值决策系统, 设 $s \in \text{FDES}(\geq_{AT})$ 且 s 在 $CL_n^{\geq} (2 \leq n \leq m)$ 中是相容的, 若 $\forall s' \in \text{DES}(\geq_{AT})$, 则 s' 是 s 的相对相容优势描述子 $\Leftrightarrow s' \geq s, \forall x \in U - CL_n^{\geq}, D_{AT}(s, x) \cap AT(s') \neq \emptyset$ 。

证明 必要性。因为 s' 是 s 的相对相容优势描述子, 所以很明显 $s' \geq s$ 。假设 $\exists x \in U - CL_n^{\geq}$ 且 $D_{AT}(s, x) \cap AT(s') = \emptyset$, 那么有 $x \in \|s'\|$ 。根据条件可知 s' 在 $CL_n^{\geq}$ 中是相容的, $\|s'\| \subseteq CL_n^{\geq}$, 从而可得 $x \in CL_n^{\geq}$ 。这与假设 $x \in U - CL_n^{\geq}$ 相矛盾。

充分性。假设 s' 在 $CL_n^{\geq}$ 中是不相容的, 那么必定有 $y \in U$ 且 $y \in \|s'\|$ 和 $y \notin CL_n^{\geq}$ 。根据条件知, 若 $y \in U - CL_n^{\geq}$, 则有 $D_{AT}(s, y) \cap AT(s') \neq \emptyset$ 。这与由 $y \in \|s'\|$ 得到的 $D_{AT}(s, y) \cap AT(s') = \emptyset$ 是相矛盾的。

定义 18 设 OISDS 为一序区间段值决策系统, 对于 $\forall s \in \text{FDES}(\geq_{AT})$, 若 s 在 $CL_n^{\geq} (2 \leq n \leq m)$ 中是相容的, 令

$$\Delta_n(s) = \bigwedge_{x \in U - CL_n^{\geq}} (\vee D_{AT}(s, x))$$

其中: $\Delta_n(s)$ 称为优势描述子 s 在 $CL_n^{\geq}$ 中的相对区分函数。

根据布尔推理理论, 由定义 17 及定理 3 可得如下结论。

定理 4 设 OISDS 为一序区间段值决策系统, $\Delta_n(s)$ 的极小析取范式中的每个合取项对应的属性子集即为 s 的相对约简优势描述子中出现的属性集合。

例 4 对于表 2 所示的序区间段值决策系统中:

a) 根据定义 18 和定理 4, 并最终通过规则的合并, 求得所有基于区间段描述子的最优优势可信决策规则为:

(a) $RR_1: f(x, a_1) \geq [2.0, 3.0] \rightarrow f(x, d) \geq 2$, 则得表 1 中的包含最优优势可信决策规则为:

$$\text{Rule}_1: f(x, a_1) \geq [2.1, 2.8] \rightarrow f(x, d) \geq 2;$$

$$\text{Rule}_2: f(x, a_1) \geq [2.1, 2.9] \rightarrow f(x, d) \geq 2;$$

$$\text{Rule}_3: f(x, a_1) \geq [2.2, 2.7] \rightarrow f(x, d) \geq 2.$$

(b) $RR_2: f(x, a_3) \geq [2.0, 3.0] \rightarrow f(x, d) \geq 2$, 则得表 1 中的包含最优优势可信决策规则为:

$$\text{Rule}_4: f(x, a_3) \geq [1.8, 3.0] \rightarrow f(x, d) \geq 2;$$

$$\text{Rule}_5: f(x, a_3) \geq [1.8, 2.9] \rightarrow f(x, d) \geq 2.$$

b) 同上, 也可求得表 2 中所有基于区间段描述子的最优弱势可信决策规则(略)。

经过研究对比以上导出的最优可信规则可以发现, 一条基于区间段值导出的最优可信规则往往都对应着多条基于原区间值导出的最优可信规则。因此, 采用本文提出的基于区间段的描述子用以获取全部可信规则的方法具有简洁性和实用性。

5 结束语

粗糙集理论在包含区间值决策系统中的扩充近年来是粗糙集理论研究的一个热点问题。本文主要研究了在包含区间值序决策系统中导出所有可信规则和知识约简方法。首先, 提出了区间段值的概念; 然后在序区间段值决策系统中提出了基于区间段值的优势和弱势描述子用以导出所有的可信规则; 接着给出了求得描述子所有约简的具体操作手段以及相应的判定定理与区分函数, 从而可以获得简化的决策规则; 最后, 用实例进行了验证。综上, 本文工作都为从序区间值决策系统中获取简洁和实用的知识提供了新的理论方法和技术手段。

在本文所做工作的基础上, 笔者下一步的研究方向将是对不完备序区间值决策系统中基于描述子的可信规则获取方法进行讨论与研究。

(下转第 4517 页)

解。在相同维数和群体数条件下,本文算法在寻优精度、寻优率方面均优于文献[7],表现出较强的全局寻优能力和较高的搜索精度。此外,本文算法方差较小,说明本文算法在全局搜索能力和稳定性方面优于文献[7]算法。其原因在于本文算法设计的各种免疫算子,增强了算法的局部搜索能力,增强了解的精度和稳定性。

为了进一步说明算法在求解全局最优解方面的能力,下面的多峰函数为例,进行说明。

$$\max f_1(x, y) = 1 + x \sin 4\pi x - y \sin(4\pi y + \pi) + \frac{\sin(6\sqrt{x^2 + y^2})}{6\sqrt{x^2 + y^2} + 10^{-15}} \quad x, y \in [-1, 1]$$

该函数为多峰函数,有四个全局最大值 2.118,对称分布于 $(+0.64, +0.64)$, $(-0.64, +0.64)$, $(+0.64, -0.64)$, $(-0.64, -0.64)$ 。该函数存在大量局部极大值,尤其是在中间区域有一取值与全局最大值很接近的局部极大值 2.077。

图1、2为两种算法运行结束后所搜索到的最优解分布。由算法结果可以看出:本文算法具有较优的全局搜索能力,同时还可搜索到部分次优解;而文献[7]算法容易陷入局部最优,且全局搜索能力差。说明本文设计的各种算子是有效的。

以上分析表明,本文算法对于解决上述多模态函数优化问题是非常有效和可靠的,

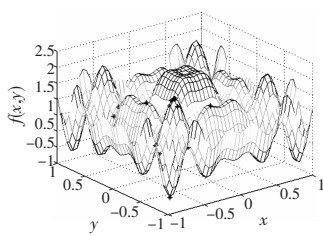


图1 文献[7]运行结束后的最优解分布

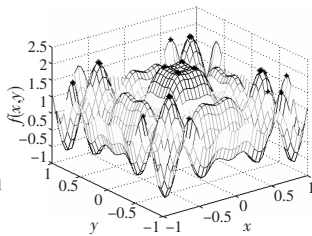


图2 本文算法运行结束后的最优解分布

3 结束语

本文针对多模函数优化问题,提出了一种求解多模函数优化的新型免疫文化基因算法。通过采用克隆算法进行全局搜索,并加入 Baldwin 学习机制局部搜索策略,帮助算法逃离局部最优解。仿真实验证明,算法寻找最优值能力更强,同时也

更加稳定可靠。

参考文献:

- [1] 毕晓君,王艳娇.用于多峰函数优化的多小生境人工蜂群算法[J].系统工程与电子技术,2011,54(11):11-15.
- [2] 朱光宇.模因内三角概率选择混合蛙跳算法[J].计算机集成制造系统,2009,56(3):10-15.
- [3] 傅清平.基于新型免疫算法的多峰函数优化[J].计算机应用研究,2011,43(10):10-15.
- [4] 彭利兵,黄辉先,阮挺,等.多峰函数优化的自适应小生境克隆选择算法[J].计算机工程与应用,2011,34(2):18-21.
- [5] 毕晓君,王义新.多峰值函数优化的拥挤差分进化算法[J].哈尔滨工程大学学报,2011,43(2):34-38.
- [6] 陆青,梁昌勇,杨善林,等.面向多峰值函数优化的自适应小生境遗传算法[J].模式识别与人工智能,2009,43(2):86-91.
- [7] 叶文,欧阳中辉,朱爱红,等.求解多峰函数优化的多小生境克隆选择算法[J].系统工程与电子技术,2010,23(5):210-214.
- [8] 薛文涛,吴晓蓓,徐志良.用于多峰函数优化的免疫粒子群网络算法[J].系统工程与电子技术,2009,34(5):213-217.
- [9] 戚玉涛,刘芳,焦李成.基于信息素模因的免疫克隆选择函数优化[J].计算机研究与发展,2008,45(6):991-997.
- [10] GONG Mao-guo, JIAO Li-cheng, LIU Fang, et al. Memetic computation based on regulation between neural and immune systems: the framework and a case study [J]. Science China (Information Sciences), 2010, 45(11): 2131-2138.
- [11] 陈杰,陈晨,张娟,等.基于 Memetic 算法的要地防空优化部署方法[J].自动化学报,2010,32(2):234-238.
- [12] 张明明,赵曙光,王旭.基于 Baldwin 效应的自适应有性繁殖遗传算法及其仿真研究[J].系统仿真学报,2010,22(10):2229-2332.
- [13] MATZINGER P. The danger model: a renewed sense of self [J]. Science, 2002, 296(5566): 301-305.
- [14] GONG Mao-guo, JIAO Li-cheng, ZHANG Li-ning, et al. Immune secondary response and clonal selection inspired optimizers [J]. Progress in Natural Science, 2009, 19(2): 237-253.
- [15] CHEN Jie, XIN Bin, PENG Zhi-hong. Statistical learning makes the hybridization of particle swarm and differential evolution more efficient: a novel hybrid optimizer [J]. IEEE Trans on Evolutionary Computation, 2008, 6(3): 239-251.

(上接第4488页)

参考文献:

- [1] PAWLAK Z, SKOWRON A. Rudiments of rough sets [J]. Information Sciences, 2007, 177(1): 3-27.
- [2] PAWLAK Z, SKOWRON A. Rough sets: some extensions [J]. Information Sciences, 2007, 177(1): 28-40.
- [3] GRECO S, MATARAZZO B, SLOWINSKI R. Rough approximation by dominance relations [J]. International Journal of Intelligent Systems, 2002, 17(2): 153-171.
- [4] DENG Ju-long. The primary methods of grey system theory [M]. Wuhan: Huazhong University of Science and Technology Press, 2005.
- [5] QIAN Yu-hua, LIANG Ji-ye, DANG Chuang-yin. Interval ordered information systems [J]. Computers and Mathematics with Applications, 2008, 56(8): 1994-2009.
- [6] CHENG Yi, MIAO Duo-qian, FENG Qin-rong. Positive approximation and converse approximation in interval-valued fuzzy rough sets [J]. Information Sciences, 2011, 181(11): 2086-2110.
- [7] WANG Zhou-jing, LI K W, XU Jian-hui. A mathematical programming

approach to multi-attribute decision making with interval-valued intuitionistic fuzzy assessment information [J]. Expert Systems with Applications, 2011, 38(10): 12462-12469.

- [8] CHEN Ting-yu, WANG Hsiao-pin, LU Yen-yu. A multicriteria group decision-making approach based on interval-valued intuitionistic fuzzy sets: a comparative perspective [J]. Expert Systems with Applications, 2011, 38(6): 7647-7658.
- [9] LEUNG Y, WU W Z, ZHANG W X. Knowledge acquisition in incomplete information systems: a rough set approach [J]. European Journal of Operational Research, 2006, 168(1): 164-180.
- [10] WU W Z, XU Y H. On two types of generalized rough set approximations in incomplete information systems [C] // Proc of IEEE International Conference on Granular Computing. 2005: 303-306.
- [11] YANG Xi-bei, XIE Jun, SONG Xiao-ning, et al. Credible rules in incomplete decision system based on descriptors [J]. Knowledge-Based Systems, 2009, 22(1): 8-17.
- [12] YANG Xi-bei, ZHANG Ming, DOU Hui-li, et al. Neighborhood systems-based rough sets in incomplete information system [J]. Knowledge-Based Systems, 2011, 24(6): 858-867.