

程度与精度的逻辑差粗糙集模型

张贤勇¹, 熊方², 莫智文¹, 程伟³

(1. 四川师范大学数学与软件科学学院 成都 610068; 2. 四川天一学院信息工程系 成都 610100;

3. 电子科技大学计算机科学与工程学院 成都 611731)

【摘要】基于程度与精度的逻辑差需求,提出了程度与精度的逻辑差粗糙集模型,并定义了粗糙集区域概念。通过变精度近似与程度近似的转化公式,得到了粗糙集区域的基本结构,提出了计算粗糙集区域的常规算法和结构算法,进行了算法分析与比较,探索并得到了模型在决策表中的应用方向。通过该模型拓展了程度粗糙集模型和经典粗糙集模型。

关键词 人工智能; 程度; 程度粗糙集; 逻辑差; 精度; 粗糙集理论; 结构算法; 变精度粗糙集
中图分类号 Tp18 **文献标识码** A **doi:**10.3969/j.issn.1001-0548.2010.05.028

Rough Set Model of Logical Difference Operation of Grade and Precision

ZHANG Xian-yong¹, XIONG Fang², MO Zhi-wen¹, and CHENG Wei³

(1. College of Mathematics and Software Science, Sichuan Normal University Chengdu 610068;

2. Department of Information Engineering, Sichuan Tianyi University Chengdu 610100;

3. School of Computer Science and Engineering, University of Electronic Science and Technology of China Chengdu 611731)

Abstract This paper proposes a rough set model of logical difference operation of grade and precision, and defines the concepts of rough set regions. In the new model, basic structure of rough set regions is obtained by transformation formulas between variable precision approximations and graded approximations. Regular algorithm and structural algorithm are proposed and analyzed to calculate rough set regions, and application prospect of the new model in decision table is explored and obtained. The new model has extended graded rough set model and classical rough set model.

Key words artificial intelligence; grade; graded rough set; logical difference; precision; rough set theory; structural algorithm; variable precision rough set

经典粗糙集模型^[1]的缺陷在于忽略了类与集合重叠部分的定量信息。而在实际中,集合间往往呈现一定程度的包含关系^[2],因此需要拓展经典粗糙集模型,而变精度粗糙集^[3]和程度粗糙集^[4]就是两个重要的拓展模型。传统的变精度粗糙集模型基于多数包含关系,即参数 β 范围为 $[0,0.5)$,从理论与实际出发,需要也容易把参数 β 范围拓展到 $[0,1]$ ^[5]。

实际上,变精度粗糙集和程度粗糙集分别来源于可能性模态逻辑和程度模态逻辑^[4]。变精度粗糙集模型的理论和应用是研究热点^[6-9]。程度粗糙集模型的研究,更多地集中在其背景上,即程度模态逻辑上^[10-11]。目前对两者结合的研究较少,但文献^[12-13]还是得到了一些结果。

精度和程度是两个重要的量化指标,分别与相对误差和绝对误差相联系。精度和程度把类与集合重叠部分的信息相对量化和绝对量化,是对同一事物两个不同逻辑侧面的刻画。两者结合起来,取长补短,使类与集合重叠部分的信息更加全面、精确;若能拓展得到范围更宽广、性质更丰富的模型,则能把经典粗糙集模型、变精度粗糙集模型、程度粗糙集模型拓展得更加适用。因此,精度和程度复合的理论研究与技术创新都具有重要意义。本文在此背景下探索精度与程度的结合及其的逻辑计算模型,拓展变精度粗糙集模型、程度粗糙集模型和经典粗糙集模型。

收稿日期: 2009-03-29; 修回日期: 2009-08-23

基金项目: 国家自然科学基金(11071178); 国家自然科学基金青年科学基金(60803028); 四川省科技支撑计划项目(09ZC1838); 四川省青年科技基金(07ZQ026114)

作者简介: 张贤勇(1978-),男,博士生,主要从事不确定性数学的理论及其应用方面的研究。

1 程度与精度的逻辑差粗糙集模型

定义 1 $\bar{R}_{k-\beta}A = \bar{R}_kA - \bar{R}_\beta A$ 、 $\underline{R}_{k-\beta}A = \underline{R}_kA - \underline{R}_\beta A$ 分别称为 A 的程度 k 与精度 $1-\beta$ 的逻辑差 R 上、下近似, 即:

$\bar{R}_{k-\beta}A = \cup\{[x]_R : |[x]_R \cap A| > k, c([x]_R, A) \geq 1-\beta\}$
 $\underline{R}_{k-\beta}A = \cup\{[x]_R : |[x]_R \cap A| \geq |[x]_R| - k, c([x]_R, A) > \beta\}$
 $\text{pos}R_{k-\beta}A = \bar{R}_{k-\beta}A \cap \underline{R}_{k-\beta}A$ 、 $\text{neg}R_{k-\beta}A = \sim(\bar{R}_{k-\beta}A \cup \underline{R}_{k-\beta}A)$ 分别称为 A 的程度 k 与精度 $1-\beta$ 的逻辑差 R 正域、负域。若 $\bar{R}_{k-\beta}A \neq \underline{R}_{k-\beta}A$ 则称 A 依程度 k 与精度 $1-\beta$ 的逻辑差运算是粗糙的。

程度与精度的逻辑差近似是程度与精度的逻辑差要求的结果, 在实际中是有意义的。上下近似、正域、负域统称为“粗糙集区域”, 具有与 β 、 k 相关的逻辑含义。 $\bar{R}_{k-\beta}A$ 是“属于 A 的个数多于 k , 但关于 A 的错误分类率不小于 $1-\beta$ ”的等价类的并, $\underline{R}_{k-\beta}A$ 是“不属于 A 的个数最多只有 k 个, 但关于 A 的错误分类率大于 β ”的等价类的并。

类似地, 定义精度与程度的逻辑差近似 $\bar{R}_{\beta-k}A$ 、 $\underline{R}_{\beta-k}A$, 容易得到下面的性质, 并可知程度与精度的逻辑差粗糙集模型部分拓展了程度粗糙集模型和经典粗糙集模型。

定理 1 $\bar{R}_kA = \bar{R}_{k-1}A$, $\underline{R}_kA = \sim \underline{R}_{1-k}A$, 进而有 $\bar{R}A = \bar{R}_0A = \bar{R}_{0-1}A$, $\underline{R}A = \underline{R}_0A = \sim \underline{R}_{1-0}A$ 。

因为:

$$c([x]_R, A) < 1-\beta \Leftrightarrow |[x]_R \cap A| > \beta|[x]_R|$$

$$c([x]_R, A) \leq \beta \Leftrightarrow |[x]_R \cap A| \geq |[x]_R| - \beta|[x]_R|$$

所以有下面变精度近似与程度近似的转化公式。

定理 2 $\bar{R}_\beta A = \cup\{[x]_R : |[x]_R \cap A| > \beta|[x]_R|\}$ 、 $\underline{R}_\beta A = \cup\{[x]_R : |[x]_R \cap A| \geq |[x]_R| - \beta|[x]_R|\}$;
 $\bar{R}_k A = \cup\{[x]_R : c([x]_R, A) < 1-k/[x]_R|\}$ 、 $\underline{R}_k A = \cup\{[x]_R : c([x]_R, A) \leq k/[x]_R|\}$ 。

定理 3 $\bar{R}_{k-\beta}A = \cup\{[x]_R : k < |[x]_R \cap A| \leq \beta|[x]_R|\}$ 、 $\underline{R}_{k-\beta}A = \cup\{[x]_R : |-k \leq |[x]_R \cap A| < |[x]_R| - \beta|[x]_R|\}$ 。

下面将分类研究由参数 β 、 k 决定的程度与精度的逻辑差粗糙集模型的基本结构。

定理 4 $\beta=0$ 时, $\bar{R}_{k-\beta}A = \phi$ 、 $\underline{R}_{k-\beta}A = \underline{R}_kA - \underline{R}A$ 、 $\text{pos}R_{k-\beta}A = \phi$ 、 $\text{neg}R_{k-\beta}A = \underline{R}A \cup (\sim \underline{R}_kA)$ 。
 $\beta=1$ 时, $\bar{R}_{k-\beta}A = \bar{R}_kA$ 、 $\underline{R}_{k-\beta}A = \phi$ 、 $\text{pos}R_{k-\beta}A = \phi$ 、 $\text{neg}R_{k-\beta}A = \sim \bar{R}_kA$ 。

定理 5 $\beta \in (0,1)$ 时, $|[x]_R| > k/\beta$ 是 $[x]_R \subseteq \bar{R}_{k-\beta}A$ 的必要条件, $|[x]_R| < k/\beta$ 是 $[x]_R \subseteq \underline{R}_{k-\beta}A$ 的必要条件。

1) $|[x]_R| > k/\beta$ 时, $[x]_R \not\subseteq \underline{R}_{k-\beta}A$ 。
 (1) $[x]_R \not\subseteq \text{pos}R_{k-\beta}A$;
 (2) $k < |[x]_R \cap A| \leq \beta|[x]_R|$ 时, $[x]_R \subseteq \bar{R}_{k-\beta}A$;
 (3) $|[x]_R \cap A| \leq k$ 或 $|[x]_R \cap A| > \beta|[x]_R|$ 时, $[x]_R \not\subseteq \bar{R}_{k-\beta}A$, 故 $[x]_R \subseteq \text{neg}R_{k-\beta}A$ 。

2) $|[x]_R| < k/\beta$ 时, $[x]_R \not\subseteq \bar{R}_{k-\beta}A$ 。

(1) $[x]_R \not\subseteq \text{pos}R_{k-\beta}A$;

(2) $|[x]_R| - k \leq |[x]_R \cap A| < (1-\beta)|[x]_R|$ 时, $[x]_R \subseteq \underline{R}_{k-\beta}A$;

(3) $|[x]_R \cap A| < |[x]_R| - k$ 或 $|[x]_R \cap A| \geq (1-\beta)|[x]_R|$ 时, $[x]_R \not\subseteq \underline{R}_{k-\beta}A$, 故 $[x]_R \subseteq \text{neg}R_{k-\beta}A$ 。

3) $|[x]_R| = k/\beta$ 时, $[x]_R \not\subseteq \bar{R}_{k-\beta}A, \underline{R}_{k-\beta}A$, 所以 $[x]_R \not\subseteq \text{pos}R_{k-\beta}A$ 、 $[x]_R \subseteq \text{neg}R_{k-\beta}A$ 。

由上可知 $\text{pos}R_{k-\beta}A = \phi$, 故在程度与精度的逻辑差粗糙集模型中, 论域仅完备分化为上近似、下近似、负域3种粗糙集区域。在实际中, 由参数 β 、 k , 用 k/β 把数轴分成 $[0, k/\beta)$ 、 k/β 、 $(k/\beta, +\infty)$ 等3部分, 由 $|[x]_R|$ 的实际分布, 并结合 $|[x]_R \cap A|$, 可得到 $[x]_R$ 的粗糙集区域归属, 进而可得到粗糙集区域。

定理 6 $\beta \in (0,1)$ 时,

$$\bar{R}_{k-\beta}A = \cup\{[x]_R : |[x]_R| > k/\beta, k < |[x]_R \cap A| \leq \beta|[x]_R|\}$$

$$\underline{R}_{k-\beta}A = \cup\{[x]_R : |[x]_R| < k/\beta, |[x]_R| - k \leq |[x]_R \cap A| < (1-\beta)|[x]_R|\}$$

$$\text{neg}R_{k-\beta}A = (\cup\{[x]_R : |[x]_R| > k/\beta, |[x]_R \cap A| \leq k \text{ 或 } |[x]_R \cap A| > \beta|[x]_R|\}) \cup (\cup\{[x]_R : |[x]_R| < k/\beta \text{ 或 } |[x]_R \cap A| < |[x]_R| - k \text{ 或 } |[x]_R \cap A| \geq (1-\beta)|[x]_R|\}) \cup (\cup\{[x]_R : |[x]_R| = k/\beta\})$$

该定理刻画了程度与精度的逻辑差粗糙集模型的基本结构。下面在 $\beta \in (0,1)$ 的模型中研究计算粗糙集区域的两种不同算法。

1) 常规算法

(1) 计算程度上下近似和变精度上下近似;

(2) 由集合的差计算程度与精度的逻辑差上下近似;

(3) 通过集合并补运算可得负域。

2) 结构算法

由定理6计算程度与精度的逻辑差粗糙集模型中的上下近似域及负域。首先 $|[x]_R|$ 与 k/β 比较大小, 然后 $|[x]_R \cap A|$ 需要和一些参数比较大小, 在不

同的情况下与参数比较大小的先后顺序如下:

- (1) $|[x]_R| > k/\beta$, 则先 k 后 $\beta|[x]_R|$;
- (2) $|[x]_R| < k/\beta$, 则先 $(1-\beta)|[x]_R|$ 后 $|[x]_R| - k$ 。

两个算法中, 核心的部分是对每类 $[x]_R$, 判断其是否属于特定集合, 主要计算是比较大小。对每类 $[x]_R$, 需要2个输入数据 $|[x]_R|$ 和 $|[x]_R \cap A|$ 。设类 $[x]_R$ 共有 n 个, 则需要 $2n$ 个输入数据。下面以比较大小作为基本操作, 以类 $[x]_R$ 的规模 n 作为实例特征, 分析和比较两个算法。

在常规算法中, 每个类需要判断是否属于 $\bar{R}_k A$ 、 $\underline{R}_k A$ 、 $\bar{R}_\beta A$ 、 $\underline{R}_\beta A$, 需要比较大小4次和2个辅助变量 $|[x]_R| - |[x]_R \cap A|$ 和 $c([x]_R, A)$, 其余是4步集合的差并补运算。所以, 常规算法的时间、空间复杂性分别为 $T(n) = 4n$ 、 $S(n) = 2n$, 且最好情况、最坏情况和平均情况下是一致的。

在结构算法中, 每个类的基数需要先与 k/β 比较大小1次。

- (1) 若 $|[x]_R| > k/\beta$, 则最多需要用 $|[x]_R \cap A|$ 与 k 、 $\beta|[x]_R|$ 比较大小2次, 需要1个辅助变量 $\beta|[x]_R|$;
- (2) 若 $|[x]_R| < k/\beta$, 则最多需要用 $|[x]_R \cap A|$ 与 $(1-\beta)|[x]_R|$ 、 $|[x]_R| - k$ 比较大小2次, 需要2个辅助变量 $(1-\beta)|[x]_R|$ 和 $|[x]_R| - k$;
- (3) 若 $|[x]_R| = k/\beta$, 则不需要再比较大小。所以, 在最坏的情况下, 结构算法的时间、空间复杂性分别为 $T(n) = 3n$ 、 $S(n) = 2n$ 。

综上所述, 在最坏的情况下, 两个算法的时间、空间复杂性的渐进分析为 $T(n) = \Theta(n)$ 、 $S(n) = \Theta(n)$ 。但结构算法比常规算法具有时间、空间优势:

- (1) 在最坏情况下, 结构算法空间复杂性与常规算法相同, 但具有时间优势;
- (2) 常规算法不存在最好、最坏与平均情况的差异, 而结构算法在一般情况下, 会有 $|[x]_R| > k/\beta$ 及 $|[x]_R| = k/\beta$ 。 $|[x]_R| > k/\beta$ 会降低结构算法的空间复杂性, $|[x]_R| = k/\beta$ 则会同时降低空间与时间复杂性。同时, 当 $|[x]_R| \neq k/\beta$ 时, 在深入验证 $[x]_R$ 的归属过程中, 若 $[x]_R$ 属于负域, 则可能只需要比较大小1次, 并减少辅助变量个数;
- (3) 结构算法由比较大小判定出类与粗糙集区域的关系即可结束, 而常规算法还需要4步集合运算。显然, 所有类满足 $|[x]_R| = k/\beta$ 时为结构算法的

最好情况, 此时 $T(n) = n$ 且该算法原地工作。

结构算法比常规算法更具优势, 主要原因在于经过理论推导, 定理6把粗糙集区域刻画得非常精确。计算时直接利用精度和程度构造的参数范围对每一个类进行检验即可确定该类的3种粗糙集区域归属, 进而得到粗糙集区域。而常规算法需要先确定类的程度上下近似和变精度上下近似4种集合归属, 然后还要再转换。实际应用中论域的分类往往是很多的, 故在大量数据处理中, 通过结构算法会节约大量的时间和空间。

2 实例

决策表 $S = (U, T, V, f)$, 其中 U 表示36个病人; $T = \{r_1, r_2, r_3\}$, 条件属性 r_1 和 r_2 分别表示“发烧”和“头痛”, 决策属性 r_3 表示“感冒”; R 表示由 r_1 和 r_2 决定的等价关系; $V_{r_1} = \{0,1,2\}$, 0、1、2分别表示不发烧、轻度发烧、严重发烧, $V_{r_2} = \{0,1,2\}$, 0、1、2、分别表示不头痛、轻度头痛、严重头痛, $V_{r_3} = \{0,1\}$, 0、1分别表示没感冒、感冒。设定初始数据, 如表1所示。由初始数据可得相关的统计结果, 其中 $[x]_m = (i, j)$ 表示基于条件属性发烧、头痛的分类, A 表示感冒病人的集合, 统计结果如表2所示。

表1 初始数据

病人	发烧	头痛	感冒
1	0	0	0
2	1	1	0
3	0	2	1
4	2	1	0
5	1	0	1
6	2	2	1
7	0	0	0
8	1	2	0
9	2	2	1
10	1	1	1
11	1	2	1
12	2	0	0
13	0	0	0
14	2	1	1
15	0	1	1
16	1	1	0
17	0	2	0
18	2	1	1
19	0	0	0
20	1	2	1
21	2	0	1

续表

病人	发烧	头痛	感冒
22	0	0	0
23	2	1	0
24	1	2	1
25	0	2	0
26	2	2	1
27	1	1	0
28	2	0	1
29	2	1	1
30	0	0	0
31	1	2	0
32	0	1	0
33	2	1	1
34	1	1	1
35	0	0	0
36	2	0	0

表2 类的统计结果

$[x]_m$	$[x]_m$ 元素	$ [x]_m $	$[x]_m \cap A$	$ [x]_m \cap A $	$c([x]_m, A)$
$[x]_1=(0,0)$	1、7、13、19、 22、30、35	7		0	1
$[x]_2=(0,1)$	15、32	2	15	1	1/2
$[x]_3=(0,2)$	3、17、25	3	3	1	2/3
$[x]_4=(1,0)$	5	1	5	1	0
$[x]_5=(1,1)$	2、10、16、27、 34	5	10、34	2	3/5
$[x]_6=(1,2)$	8、11、20、24、31	5	11、20、24	3	2/5
$[x]_7=(2,0)$	12、21、28、36	4	21、28	2	1/2
$[x]_8=(2,1)$	4、14、18、23、 29、33	6	14、18、 29、33	4	1/3
$[x]_9=(2,2)$	6、9、26	3	6、9、26	3	0

下面计算 $k=1$ 、 $\beta=0.4$ 时，程度与精度的逻辑差粗糙集模型中的粗糙集区域。

1) 常规算法: (1) $k=1$ 时, $\bar{R}_k A = [x]_5 \cup [x]_6 \cup [x]_7 \cup [x]_8 \cup [x]_9$, $\underline{R}_k A = [x]_2 \cup [x]_4 \cup [x]_9$; $\beta=0.4$ 时, $\bar{R}_\beta A = [x]_2 \cup [x]_4 \cup [x]_6 \cup [x]_7 \cup [x]_8 \cup [x]_9$, $\underline{R}_\beta A = [x]_4 \cup [x]_6 \cup [x]_8 \cup [x]_9$ 。

(2) $\bar{R}_{1-0.4} A = [x]_5$, $\underline{R}_{1-0.4} A = [x]_2$;

(3) $\text{neg}R_{1-0.4} A = [x]_1 \cup [x]_3 \cup [x]_4 \cup [x]_6 \cup [x]_7 \cup [x]_8 \cup [x]_9$ 。

2) 结构算法

$\beta=0.4$, $k=1$, $k/\beta=2.5$, 可计算得 $[x]_5 \subseteq$

$\bar{R}_{1-0.4} A$ 、 $[x]_2 \subseteq \underline{R}_{1-0.4} A$ 、 $[x]_1, [x]_3, [x]_4, [x]_6, [x]_7, [x]_8, [x]_9 \subseteq \text{neg}R_{1-0.4} A$ 。比如对 $[x]_5$ 有 $|[x]_5| = 5 > 2.5$, $1 < |[x]_6 \cap A| = 3 \leq 0.6 \times 5$, 即 $[x]_5$ 满足 $\bar{R}_{k-\beta} A$ 中的条件 $|[x]_R| > k/\beta, k < |[x]_R \cap A| \leq \beta |[x]_R|$, 所以 $[x]_5 \subseteq \bar{R}_{1-0.4} A$ (详细计算见表3), 可得 $\bar{R}_{1-0.4} A = [x]_5$ 、 $\underline{R}_{1-0.4} A = [x]_2$ 、 $\text{neg}R_{1-0.4} A = [x]_1 \cup [x]_3 \cup [x]_4 \cup [x]_6 \cup [x]_7 \cup [x]_8 \cup [x]_9$ 。

根据发烧、头痛条件属性, 论域分化为9种病人类别, $\bar{R}A = [x]_2 \cup [x]_3 \cup [x]_4 \cup [x]_5 \cup [x]_6 \cup [x]_7 \cup [x]_8 \cup [x]_9$ 和 $\underline{R}A = [x]_4 \cup [x]_9$ 分别表示可能和一定归入感冒集合的病人类别。而 $\bar{R}_{1-0.4} A = [x]_5$ 表示“属于感冒集合的病人数多于1个, 但关于感冒集合的错误分类率不小于0.6”的病人类别, $\underline{R}_{1-0.4} A = [x]_2$ 表示“不属于感冒集合的病人数最多只有1个, 但关于感冒集合的错误分类率大于0.4”的病人类别。

由该例可见, 程度与精度的逻辑差粗糙集模型, 利用精度和程度两个指标对概念进行了复合逻辑刻画, 具有实际意义。该例有9个类, 常规算法中 $T(9) = 36$ 、 $S(9) = 18$; 而结构算法中仅有 $[x]_2$ 、 $[x]_4$ 满足 $|[x]_R| < k/\beta = 2.5$, 其余全部满足 $|[x]_R| > k/\beta = 2.5$; 同时, 对 $[x]_1, [x]_3, [x]_4$ 再作深入判断时, 只需要比较大小1次, 分别需要0、0、1个辅助变量, 即可确定它们属于负域。结构算法的分析见表3, 其时间、空间复杂性为 $T(9) = 24$ 、 $S(9) = 8$ 。由此得出结构算法的优势比较明显。

表3 结构算法的计算与分析

病人类别	第一轮比较大小: $ [x]_m $	第二轮比较大小: $ [x]_m \cap A $	粗糙集区域归属	比较大小次数	辅助变量个数
$[x]_1$	$> k/\beta$	$\leq k$	$\text{neg}R_{1-0.4} A$	2	0
$[x]_2$	$< k/\beta$	$< (1-\beta) [x]_R $, $\geq [x]_R - k$	$\underline{R}_{1-0.4} A$	3	2
$[x]_3$	$> k/\beta$	$\leq k$	$\text{neg}R_{1-0.4} A$	2	0
$[x]_4$	$< k/\beta$	$\geq (1-\beta) [x]_R $	$\text{neg}R_{1-0.4} A$	2	1
$[x]_5$	$> k/\beta$	$> k$ 、 $\leq \beta [x]_R $	$\bar{R}_{1-0.4} A$	3	1
$[x]_6$	$> k/\beta$	$> k$ 、 $> \beta [x]_R $	$\text{neg}R_{1-0.4} A$	3	1
$[x]_7$	$> k/\beta$	$> k$ 、 $> \beta [x]_R $	$\text{neg}R_{1-0.4} A$	3	1
$[x]_8$	$> k/\beta$	$> k$ 、 $> \beta [x]_R $	$\text{neg}R_{1-0.4} A$	3	1
$[x]_9$	$> k/\beta$	$> k$ 、 $> \beta [x]_R $	$\text{neg}R_{1-0.4} A$	3	1

3 决策表中的应用

$S = (U, T)$ 为决策表, $T = C \cup D$ 。可定义决策属性集 D 与条件属性集 C 的程度 k 与精度 $1-\beta$ 的

逻辑差的近似依赖性:

$$\gamma(C, D, k, \beta) = |\text{pos}(C, D, k, \beta)| / |U|$$

式中 $\text{pos}(C, D, k, \beta) = \bigcup_{Y \in U/D} \underline{\text{ind}}(C)_{k-\beta} Y$ 。

决策属性集与条件属性集的程度与精度的逻辑差近似依赖性, 是对执行具有程度 k 与分类误差 β 的对象分类能力的一种综合评价。在上述实例中, $R_{1-0.4}A = [x]_2$, $R_{1-0.4}(\sim A) = [x]_2 \cup [x]_4$, 可得感冒决策属性集与发烧、头痛条件属性集的程度1与精度0.6的逻辑差近似依赖性 $\gamma(C, D, 1, 0.4) = 8.33\%$ 。

基于决策属性集与条件属性集的程度与精度的逻辑差近似依赖性, 可以定义条件属性集 C 关于决策属性集 D 的程度 k 与精度 $1-\beta$ 的逻辑差近似约简 $\text{red}(C, D, k, \beta)$ 为 C 的一个子集, 满足:

$$(1) \gamma(C, D, k, \beta) = \gamma(\text{red}(C, D, k, \beta), D, k, \beta);$$

(2) 从 $\text{red}(C, D, k, \beta)$ 中去掉任何一个属性(1)都不成立。

属性约简是粗糙集进行数据分析的重要概念, 引入条件属性集关于决策属性集的程度与精度的逻辑差近似约简后, 扩充了粗糙集理论, 更好的体现了数据相关性, 为进行近似推理和获取决策规则奠定了基础。

4 结 论

程度与精度的逻辑差粗糙集模型, 拓展了程度粗糙集模型和经典粗糙集模型, 涉及了精度与程度两个量化指标, 具有实际逻辑意义, 因而具有重要的理论价值和广泛的应用前景。对该模型的性质与应用, 以及其他平行的精度与程度的逻辑运算模型, 还值得深入地探讨。

本文研究工作得到四川师范大学科学研究基金(08KYL06)的资助, 在此表示感谢。

参 考 文 献

[1] PAWLAK Z. Rough sets[J]. International Journal of Computer and Information Sciences, 1982, 11: 341-356.
 [2] 舒 兰, 赵 磊. 粗糙集的模糊性[J]. 电子科技大学学报, 2005, 34(1): 124-126.

SHU Lan, ZHAO Lei. Fuzziness in rough sets[J]. Journal of University of Electronic Science and Technology of China, 2005, 34(1): 124-126.

- [3] ZIARKO W. Variable precision rough set model[J]. Journal of Computer and System Sciences, 1993, 46: 39-59.
 [4] YAO Y Y, LIN T Y. Generalization of rough sets using modal logics[J]. Intelligent Automation and Soft Computing, 1996, 2: 103-120.
 [5] 张贤勇, 莫智文. 变精度粗糙集[J]. 模式识别与人工智能, 2004, 17(2), 151-155.
 ZHANG Xian-yong, MO Zhi-wen. Variable precision rough sets[J]. Pattern Recognition and Artificial Intelligence, 2004, 17(2): 151-155.
 [6] PEI Zhi-li, SHI Xiao-hu, NIU Meng, et al. A method of gene-function annotation based on variable precision rough sets[J]. Journal of Bionic Engineering, 2007, 4: 177-184.
 [7] HONG Tzung-pei, WANG Tzu-ting, WANG Shyue-liang. Mining fuzzy β -certain and β -possible rules from quantitative data based on the variable precision rough-set model[J]. Expert Systems with Applications, 2007, 32: 223-232.
 [8] XIE Gang, ZHANG Jin-long, LAI K K, et al. Variable precision rough set for group decision-making: An application[J]. International Journal of Approximate Reasoning, 2008, 49: 331-343.
 [9] 孙士保, 姚磊磊, 吴庆涛, 等. 变精度粗糙集模型及其应用研究[J]. 计算机工程与应用, 2009, 45(7): 10-13.
 SUN Shi-bao, YAO Lei-lei, WU Qing-tao, et al. Research of generalized variable precision rough set model and its application[J]. Computer Engineering and Applications, 2009, 45(7): 10-13.
 [10] CERRATO C. Decidability by filtrations for graded normal logics (graded modalities V)[J]. Studia Logica, 1994, 53(1): 61-73.
 [11] TOBIES S. Pspace reasoning for graded modal logics[J]. Journal of Logic and Computation, 2001, 11(1): 85-106.
 [12] ZHANG Xian-yong, MO Zhi-wen. Product approximation of grade and precision[J]. Journal of Electronic Science and Technology of China, 2005, 3(3): 276-279.
 [13] ZHANG Xian-yong, MO Zhi-wen, XIONG Fang. Approximation of intersection of grade and precision[C]//Fuzzy Information and Engineering (Advances in Soft Computing 54). Berlin: Springer, 2008: 526-530.

编辑 蒋 晓