

不同知识粒度下粗糙集的不确定性研究

王国胤^{1),2)} 张清华^{1),2)}

¹⁾(西南交通大学信息科学与技术学院 成都 610031)

²⁾(重庆邮电大学计算机科学与技术研究所 重庆 400065)

摘 要 粗糙集的不确定性度量方法,目前主要包括粗糙集的粗糙度、粗糙熵、模糊度和模糊熵.在不同知识粒度下,从属性的角度,给出了分层递阶的知识空间链,发现在分层递阶的知识粒度下部分文献中定义的粗糙集的粗糙熵和模糊度随知识粒度的变化规律不一定符合人们的认识规律.从信息熵的角度提出了一种粗糙集不确定性的模糊度量方法,证明了这种模糊度随知识粒度的减小而单调递减,弥补了现有粗糙熵和模糊度量粗糙集不确定性的不足.最后,分析了在不同知识粒度下粗糙度和模糊度的变化关系.

关键词 粗糙度;粗糙熵;模糊度;知识粒度;商空间
中图法分类号 TP18

Uncertainty of Rough Sets in Different Knowledge Granularities

WANG Guo-Yin^{1),2)} ZHANG Qing-Hua^{1),2)}

¹⁾(School of Information Science & Technology, Southwest Jiaotong University, Chengdu 610031)

²⁾(Institute of Computer Science & Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065)

Abstract Rougness, rough entropy, fuzziness, and fuzzy entropy are major methods for measuring the uncertainty of rough sets. In different knowledge granularity levels, a hierarchical knowledge space chain is proposed based on the attributes in information systems. Some regularities of the changing of rough entropy and fuzziness of a rough set with the knowledge granularity are found to be inconsistent with human cognition. A new method for measuring the fuzziness of rough sets is proposed based on information entropy. The fuzziness measured by the new method is monotonously decreasing with the refining of knowledge granularity in apporiximation spaces. It overcomes the problem of roughness and rough entropy. Finally, the relations of the changing of roughness and fuzziness are analyzed in different knowledge granularities.

Keywords roughness; rough entropy; fuzziness; knowledge granularity; quotient space

1 引 言

进入21世纪以来,不确定性问题的研究工作受到越来越多的关注^[1].如何对不确定性信息和数据进行更加有效的处理,从而发现不确定性信息中蕴

涵的知识和规律,是一个重要的研究课题^[2].Zadeh在1965年提出的模糊集(Fuzzy sets)理论^[3],Pawlak在1982年提出的粗糙集(Rough sets)理论^[4]和张钹、张铃在1990年提出的商空间理论^[5]是粒计算(granular computing)的三大基础数学理论,是处理不确定性问题的有效方法,已广泛应用于模

式识别、知识发现、问题求解以及不确定推理等领域. 模糊集作为经典康托集的推广, 利用隶属函数来表示对象关于集合的隶属程度, 重在区分属于同一集合的不同对象间的隶属程度, 其不足之处在于其隶属函数往往需由专家给出, 带有一定的主观性; 粗糙集理论是处理不完全和不精确信息的一种有效数学工具^[6], 建立在对论域分类的基础上, 将不确定知识用已知知识库中的知识来刻画, 对不确定问题的描述和处理比较客观, 但粗糙集理论是研究在给定的空间(知识基)上不同概念的表示、转换和相互依存问题的, 其论域是点集, 元素之间没有拓扑关系; 商空间理论基于复杂问题粒化的思想, 建立了一种商结构的形式化问题求解理论体系, 利用保真、保假原理来高效地获得问题的解或近似解, 它不仅针对给定的商空间(知识基)来讨论知识的表达问题, 而且利用对象之间的结构(偏序结构或拓扑结构), 在所有可能的商空间中找出最合适的商空间, 从不同商空间(不同角度)观察同一问题, 以便得到对问题不同角度的理解, 最终合成对原问题总的解(近似解)^[5]. 可以说, 模糊集理论是一种“软”计算方法, 粗糙集理论是“硬”计算方法, 而商空间理论是介于模糊集和粗糙集之间的一种问题求解(近似解)的计算方法, 模糊商空间可以利用分层递阶结构“廉价”地描述问题的不确定性^[7]. 另外, Gau 和 Buehrer 提出的 Vague 集理论, 通过对模糊对象赋予真、假隶属函数来处理模糊性, 是模糊集理论的扩充^[8]. 依靠各自的特点和优势, 这些方法已经广泛应用于对不确定、不精确、不完整信息的处理以及对大规模海量数据的挖掘和对复杂问题的求解^[9]. 李德毅认为^[1]: 在主、客观世界普遍存在的不确定性中, 随机性和模糊性是最重要的两种形式, 不确定性和确定性并非完全对立, 在一定程度上可以相互转化. 例如, 某一层次的不确定性可能是更高层次上的确定性, 种种不确定性中还可能隐藏着某些确定的规律等. 人工智能研究人员的任务, 就是寻找并且能够形式化地表示不确定性中的规律性, 至少是在某种程度上的规律性, 从而使机器能够模拟人类认识客观世界、认识人类本身的认知过程. 当前, 对于粗糙集的不确定性度量的方法主要有粗糙度、粗糙熵、模糊度和模糊熵.

在同一知识粒度的近似空间下, Chakrabarty^[10]等人较为详细地讨论了粗糙集的模糊性度量问题; Banerjee^[11]和 Huynh^[12]对模糊集的粗糙度进行了研究; 王国胤^[13-15]等人从信息观的角度分析了决策

信息系统的确定性, 并讨论了代数观和信息观意义下粗糙集的不确定性的异同; 梁吉业^[16-18]等人从信息熵、条件熵、互信息和知识粒度的角度分析了粗糙集的不确定性, 并给出了一种新的粗糙集的粗糙熵; 苗夺谦^[19-21]等人从粒计算和信息表示等角度研究了知识的粒度、知识的粗糙度与信息熵之间的关系.

然而, 随着属性个数的变化, 论域空间形成一个分层递阶结构(金字塔结构, 即商空间). 当知识空间中的知识粒度严格递减时, 一个粗糙集的粗糙度、粗糙熵、模糊度和模糊熵将怎样变化? 它们之间的关系又是如何? 关于这方面的研究工作, 已有一定的研究基础, 特别是研究粗糙精度、粗糙度、分类精度、粗糙熵和条件熵在不同知识粒度的近似空间下的变化已经比较详尽^[11-16, 18-21]. 综合分析上述研究工作可以发现, 粗糙集的粗糙度随着知识粒度的减小而单调递减, 这符合人们的认知直觉. 但是, 很多实际例子表明, 当属于一个集合的正域或负域中的知识颗粒被细分时, 粗糙集的粗糙度将不发生变化; 而且当属于集合边界域中的知识颗粒被细分时, 它的粗糙度可能也不发生变化, 这与人们的认知直觉不吻合. 为了克服这个问题, 有的研究者提出了粗糙熵, 如 Liang^[18]等人定义了一种粗糙熵, 它是集合 X 的粗糙度与近似空间中的知识粒度之积, 并得到结论: 这种粗糙熵随着知识颗粒的细分严格单调递减. 这个结论在一定程度上弥补了用粗糙度度量粗糙集不确定性的不足. 但是, 我们分析发现, 如果对集合 X 负域的知识颗粒(与 X 无关)进行细分, 粗糙度将不变(符合人们的认知规律), 但粗糙熵却严格递减(不符合人们的认知规律). 这说明与集合 X 无关的知识颗粒的变化也会导致 X 的粗糙熵的变化, 这与人们对不确定性问题的认知不符.

为此, 需要进一步研究粗糙集不确定性的另一度量方式——粗糙集的模糊度. 虽然在同一知识粒度的近似空间下粗糙集的模糊性得到研究者的关注^[10, 16, 18-21], 但是关于粗糙集的模糊度在不同知识粒度的近似空间(分层递阶的近似空间)下将如何变化的研究工作甚少. 粗糙集的模糊度随着近似空间中知识颗粒的细分将如何变化? 对这个问题的探索, 有利于发现不确定性问题中隐藏的某些确定规律. 从认知角度来讲, 集合 X 随着与它有关的知识颗粒的细分, 它的不确定性要降低, 模糊度也应该降低. 但是, 文献[10]给出的粗糙集模糊度在知识粒度细化的过程中可能反而会逐渐增加, 这与人们认知

不确定性问题的直觉相悖.

本文从属性空间的角度,主要讨论不同知识粒度的近似空间下(即不同层次的商空间)粗糙集的模糊度的变化问题,提出一种基于信息熵的粗糙集的模糊度量方法,证明这种模糊度随着知识粒度的减小而单调递减,弥补粗糙度和粗糙熵对粗糙集不确定性度量的不足.这种模糊度的物理背景非常清楚,它既刻画出集合 X 的边界域中属于 X 的那部分元素“贡献”的不确定性,也刻画出不属于 X 的那部分元素“贡献”的不确定性,更精确地描述了粗糙集的不确定性.通过分析发现,如果集合 X 的边界域中的知识颗粒被“成比例”地细分,这种粗糙集的模糊度不会发生变化;如果集合 X 的边界域中的知识颗粒被“不成比例”地细分,这种粗糙集的模糊度将严格递减.这个结论克服了现有部分度量粗糙集不确定性方法的不足,与人们对不确定性问题的认知规律非常吻合.

本文第 2 节介绍相关基本概念;第 3 节讨论不同知识粒度下粗糙集的不确定性度量问题;第 4 节提出一种基于信息熵的粗糙集模糊度量方法;第 5 节讨论不同知识粒度下粗糙集的粗糙度和模糊度的变化关系;第 6 节是结束语.

2 相关基本概念

2.1 知识的粒度

定义 1^[6] 设一个信息系统是四元组 $S=(U, A, V, f)$, 其中 $U=\{x_1, x_2, \dots, x_n\}$ 是非空有限对象集, 称为“论域”, $A=\{a_1, a_2, \dots, a_m\}$ 是属性集, $V=\bigcup_{a \in C} V_a$, V_a 称为属性 a 的“值域”, $f_a: U \rightarrow V_a$ 是信息函数. 不可分辨关系: $IND(B)=\{(x, y) \in U \times U \mid \forall a \in B (f_a(x)=f_a(y))\}$ 是 U 上的等价关系, 所有等价类的集合记为 $U/IND(B)$, 简写为 U/B .

一个论域的划分构成粗糙集的一个近似空间, 划分中的每一个分块称为一个知识颗粒, 度量知识粒度的方法很多, 这里我们采用 Liang 等人给出的知识粒度的度量方法^[18]. 设 $U=\{x_1, x_2, \dots, x_n\}$, 属性集 $B(B \subseteq A)$ 对论域的划分 $U/B=\{X_1, X_2, \dots, X_m\}$, 则 U/B 的知识粒度定义为

$$G(U/B)=\frac{1}{|U|^2} \sum_{i=1}^m |X_i|^2 \quad (1)$$

容易证明: $\frac{1}{n} \leq G(U/B) \leq 1$ ($|\cdot|$ 表示集合的元素个数, 下同).

2.2 分层递阶的近似空间

任给一个信息系统 $S=(U, A, V, f)$, $A=\{a_1, a_2, \dots, a_m\}$ 是属性集, 任给一个属性子集 $B(B \subseteq A)$, 我们可以得到论域 U 的一个划分 U/B . U/B 中的每个元素 $[x]_B$ ($[x]_B$ 表示元素 x ($x \in U$) 的等价类) 表示近似空间的一个知识颗粒. 设 $P(A)$ 表示集合 $A=\{a_1, a_2, \dots, a_m\}$ 的幂集. 不难看出: 代数系统 $\langle P(A), \subseteq \rangle$ 构成一个完备的偏序格. 其中, \emptyset 是这个偏序格的最小元, A 是最大元.

定义 2. 在格 $\langle P(A), \subseteq \rangle$ 对应的 Hasse 图中, 从 \emptyset 到 A 的一条路径称为属性链.

例 1. $A=\{a_1, a_2, a_3\}$, $\langle P(A), \subseteq \rangle$ 对应的 Hasse 图如图 1 所示.

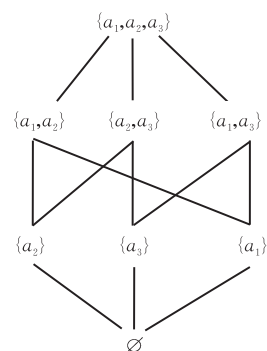


图 1 格 $\langle P(A), \subseteq \rangle$

如 $\emptyset \subseteq \{a_1\} \subseteq \{a_1, a_2\} \subseteq \{a_1, a_2, a_3\}$, $\emptyset \subseteq \{a_1\} \subseteq \{a_1, a_3\} \subseteq \{a_1, a_2, a_3\}$ 和 $\emptyset \subseteq \{a_3\} \subseteq \{a_2, a_3\} \subseteq \{a_1, a_2, a_3\}$ 等都是属性链.

定义 3^[21]. 设 $U=\{x_1, x_2, \dots, x_n\}$ 为非空有限论域, $P'=\{P'_1, P'_2, \dots, P'_l\}$ 和 $P''=\{P''_1, P''_2, \dots, P''_m\}$ 为 U 上的两个划分空间, 如果 $\forall P'_i \in P' (\exists P''_j \in P'' (P'_i \subseteq P''_j))$, 则称 P' 是 P'' 的细分空间, 记为 $P' \leq P''$.

定义 4^[22]. 设 $U=\{x_1, x_2, \dots, x_n\}$ 为非空有限论域, $P'=\{P'_1, P'_2, \dots, P'_l\}$ 和 $P''=\{P''_1, P''_2, \dots, P''_m\}$ 为 U 上的两个划分空间, 如果 $P' \leq P''$, 且 $\exists P'_i \in P' (\exists P''_j \in P'' (P'_i \subset P''_j))$, 则称 P' 是 P'' 的严格细分空间, 记为 $P' < P''$.

定理 1. 设格 $\langle P(A), \subseteq \rangle$ 中的任意一条属性链为 $\emptyset = B_0 \subset B_1 \subset B_2 \subset \dots \subset B_m = A$, 则 $U/B_m \leq U/B_{m-1} \leq \dots \leq U/B_1 \leq U/B_0 = \{U\}$.

在任何一条属性链下, 对象集 U 被分成不同的划分, 这些划分在“ \leq ”关系下构成一个金字塔结构, 称为分层递阶的近似空间.

例 2. 一个信息系统 $U=\{x_1, x_2, \dots, x_{10}\}$, $A=\{a_1, a_2, a_3\}$, 如表 1 所示.

表 1 一个信息系统

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}
a_1	1	1	1	1	1	1	1	2	2	3	3	3
a_2	0	1	1	1	1	1	1	2	2	3	3	3
a_3	0	0	0	0	0	0	1	2	3	4	4	5

如果取属性链 $\emptyset \subseteq \{a_1\} \subseteq \{a_1, a_2\} \subseteq \{a_1, a_2, a_3\}$, 可得到如下的分层递阶近似空间:

$$U/\emptyset = \{\{x_1, x_2, \dots, x_{10}\}\};$$
$$U/\{a_1\} = \{\{x_1, x_2, x_3, x_4, x_5, x_6, x_7\}, \{x_8, x_9\}, \{x_{10}, x_{11}, x_{12}\}\};$$
$$U/\{a_1, a_2\} = \{\{x_1\}, \{x_2, x_3, x_4, x_5, x_6, x_7\}, \{x_8, x_9\}, \{x_{10}, x_{11}, x_{12}\}\};$$
$$U/\{a_1, a_2, a_3\} = \{\{x_1\}, \{x_2, x_3, x_4, x_5, x_6\}, \{x_7\}, \{x_8\}, \{x_9\}, \{x_{10}, x_{11}\}, \{x_{12}\}\}.$$

在这个分层递阶的近似空间中,随着属性个数的增加,知识颗粒逐渐“细化”.

2.3 粗糙集不确定性的几种度量方法

2.3.1 粗糙集的粗糙度

定义 5^[6]. 在一个信息系统中, $IND(B)$ 是 U 上的一个不可分辨关系, $[x]_B$ 表示对象 x 的等价类, 对象子集 $X \subseteq U$, X 的下近似集($\underline{B}X$)、上近似集($\overline{B}X$)和边界域($BN_B(X)$)分别定义如下:

$$\underline{B}X = \{x \in U | [x]_B \subseteq X\},$$
$$\overline{B}X = \{x \in U | [x]_B \cap X \neq \emptyset\},$$
$$BN_B(X) = \overline{B}(X) - \underline{B}(X).$$

定义 6^[22]. 在一个信息系统中, $IND(B)$ 是 U 上的一个不可分辨关系, $[x]_B$ 表示对象 x 的等价类, 对象子集 $X \subseteq U$, X 的粗糙精度和粗糙度为

粗糙精度: $\alpha_B(X) = \frac{R(X)}{R(X)}$;

粗糙度:

$$\rho_B(X) = 1 - \alpha_B(X) = 1 - \frac{R(X)}{R(X)} = \frac{BN_B(X)}{R(X)}.$$

显然,对于任意的 $X \subseteq U$, 都有 $0 \leq \alpha_B(X) \leq 1$ 且 $0 \leq \rho_B(X) \leq 1$.

如果 $\overline{B}(X) = \underline{B}(X) = X$, 即 $\rho_B(X) = 0$ (或 $\alpha_B(X) = 1$), 称 X 关于 B 是精确的;

如果 $\underline{B}(X) \subset \overline{B}(X)$, 即 $0 < \rho_B(X) \leq 1$ (或 $0 \leq \alpha_B(X) < 1$), 称 X 关于 B 是粗糙的.

2.3.2 粗糙集的粗糙熵

关于粗糙集的粗糙熵的定义形式很多, 这里我们采用 Liang 提出的粗糙熵.

定义 7^[18]. 设 $U = \{x_1, x_2, \dots, x_n\}$, 属性子集 $B(B \subseteq A)$ 对论域的划分 $U/B = \{X_1, X_2, \dots, X_m\}$,

$X \subseteq U$, 则属性集合 B 的熵定义为

$$E(B) = - \sum_{i=1}^m \frac{|X_i|}{|U|} \log_2 \frac{1}{|X_i|} \quad (2)$$

X 在划分 U/B 上的粗糙熵定义为

$$E_B(X) = \rho_B(X) E(B) \quad (3)$$

集合 X 的粗糙熵是粗糙度与属性集合 B 的熵之积.

2.3.3 粗糙集的模糊度

设 $U = \{x_1, x_2, \dots, x_n\}$ 是非空有限集, A 是 U 上的模糊集, $A(x_i)$ 是模糊集的隶属函数. 用 $P(U)$ 表示集合 U 上的所有经典集合, $F(U)$ 表示集合 U 上的所有模糊集合. 显然, $P(U) \subseteq F(U)$.

定义 8^[23]. $\forall A \in F(U)$, 若映射 $d: F(U) \rightarrow [0, 1]$ 满足条件:

- (1) $d(A) = 0$ 当且仅当 $A \in P(U)$;
- (2) $d(A) = 1$ 当且仅当 $\forall x_i \in U A(x_i) = \frac{1}{2}$;
- (3) $\forall x_i \in U \left(B(x_i) \leq A(x_i) \leq \frac{1}{2} \vee B(x_i) \geq A(x_i) \geq \frac{1}{2} \right) \rightarrow d(B) \leq d(A)$;

(4) $d(A) = d(A^c)$, 这里 A^c 是 A 的补集, 则称映射 d 是 $F(U)$ 上的一个模糊度, 记为 $d(\cdot)$.

设 U 是非空对象集, 对象子集 $X \subseteq U$, 则对于任意的 $x(x \in U)$, x 属于集合 X 的隶属函数为

$$\mu_X^B(x) = \frac{|X \cap [x]_B|}{|[x]_B|} \quad (4)$$

显然, $0 \leq \mu_X^B(x) \leq 1$, 它表示任意一个元素属于集合 X 的程度. 令 $F_X^B = \{\mu_X^B(x_1), \mu_X^B(x_2), \dots, \mu_X^B(x_n)\}$, 则 F_X^B 是集合 U 上的一个模糊集 (即 $F_X^B \in F(U)$). 由粗糙集上、下近似和边界的概念, 不难得出:

$$\underline{B}X = \{x \in U | \mu_X^B(x) = 1\};$$
$$\overline{B}X = \{x \in U | 0 < \mu_X^B(x) \leq 1\}.$$

模糊度是度量不确定问题的有力工具, 很多研究者对粗糙集的模糊度进行了分析, Chakrabarty^[10] 等人提出粗隶属函数可以导出模糊集, 并利用模糊集与它的最邻近清晰集间的距离来度量粗糙集的模糊性.

定义 9^[10]. 设 A 是 U 上的模糊集, 与 A 有关的最邻近的清晰集记为 Δ , 其定义为

$$\Delta(x_i) = \begin{cases} 0, & A(x_i) < 0.5 \\ 1, & A(x_i) > 0.5 \\ 0 \text{ 或 } 1, & A(x_i) = 0.5 \end{cases}$$

一般地,当 $A(x_i)=0.5$ 时,取 $\underline{A}(x_i)=1$,这时 $\underline{A}=A_{0.5}$,这里 $A_{0.5}$ 表示 A 的 0.5 截集.

Chakrabarty^[10] 等人利用模糊集 F_X^B 和它的最邻近清晰集 \underline{F}_X^B 之间的距离给出了粗糙集的两种模糊性度量.

(1) 线性模糊度:

$$d_{KI}(F_X^B)=\frac{2}{n}\sum_{i=1}^n|\mu_X^B(x_i)-\underline{\mu}_X^B(x_i)| \tag{5}$$

(2) 二次模糊度:

$$d_{Kq}(F_X^B)=\frac{2}{\sqrt{n}}\sqrt{\sum_{i=1}^n(\mu_X^B(x_i)-\underline{\mu}_X^B(x_i))^2} \tag{6}$$

其中, $\underline{\mu}_X^B(x_i)$ 表示 x_i 在模糊集 \underline{F}_X^B 中的隶属函数.

2.3.4 粗糙集的模糊熵

定义 10^[24]. $\forall A \in F(U)$, 若映射 $e: F(U) \rightarrow [0, +\infty)$ 满足条件:

- (1) $e(A)=0$ 当且仅当 $A \in P(U)$;
- (2) $e(A)$ 取得最大值当且仅当 $\forall x_i \in U, A(x_i)=\frac{1}{2}$;
- (3) $\forall x_i \in U \left(B(x_i) \leq A(x_i) \leq \frac{1}{2} \vee B(x_i) \geq A(x_i) \geq \frac{1}{2} \right) \rightarrow e(B) \leq e(A)$;

(4) $e(A)=e(A^c)$, 这里 A^c 是 A 的补集, 则称映射 e 是 $F(U)$ 上的一个模糊熵, 记为 $e(\cdot)$.

梁吉业^[16-18] 等人建立了粗糙集的一种模糊熵:

$$E_L(F_X^B)=\sum_{i=1}^n\mu_X^B(x_i)(1-\mu_X^B(x_i)) \tag{7}$$

并得出了相应的结论: 一个精确集的模糊熵等于 0, 一个粗糙集合与它的补集具有相同的模糊性.

2.3.5 信息熵

信息熵是一个非常广泛的概念, 1948 年 Shannon 信息熵^[25] 的提出为信息的不确定度量奠定了理论基础, Klir 基于 Shannon 熵提出了一种度量不确定性的信息熵^[26]:

$$H(F_X^B)=-\frac{2}{n}\sum_{i=1}^n\mu_X^B(x_i)\log_2^{\mu_X^B(x_i)} \tag{8}$$

容易验证, $H(\cdot)$ 不满足模糊度的定义(定义 8), 不是模糊度.

3 不同知识粒度下粗糙集的不确定性度量

目前, 度量粗糙集不确定性的方法主要有粗糙度、粗糙熵、模糊度和模糊熵. 在分层递阶的近似空

间下, 随着知识颗粒的细分, 不同层次上的知识粒度有何变化规律? 定理 2 和定理 3 揭示了这个变化规律.

定理 2^[21]. 设格 $\langle P(A), \subseteq \rangle$ 中的任意一条链为 $\emptyset=B_0 \subset B_1 \subset B_2 \subset \dots \subset B_m=A$, 则 $G(U/B_{i+1}) \leq G(U/B_i)$ ($i=0, 1, 2, \dots, m-1$; 下同).

定理 3^[21]. 设格 $\langle P(A), \subseteq \rangle$ 中的任意一条链为 $\emptyset=B_0 \subset B_1 \subset B_2 \subset \dots \subset B_m=A$, 如果 $U/B_{i+1} < U/B_i$, 则 $G(U/B_{i+1}) < G(U/B_i)$.

在分层递阶的近似空间上, 随着知识粒度的减小, 粗糙集的粗糙度将如何变化? 定理 4 回答了这个问题.

定理 4^[6]. 设格 $\langle P(A), \subseteq \rangle$ 中的任意一条链为 $\emptyset=B_0 \subset B_1 \subset B_2 \subset \dots \subset B_m=A$, 对于任意的 $X \subseteq U$, 有 $\rho_{B_{i+1}}(X) \leq \rho_{B_i}(X)$.

定理 4 揭示了集合 X 的粗糙度随知识粒度减小而单调递减. 注意: 如果 $U/B_{i+1} < U/B_i$ (严格的细分关系), 不一定有 $\rho_{B_{i+1}}(X) < \rho_{B_i}(X)$ (严格单调递减). 如例 2 中, $X=\{x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9\}$, 取属性链 $\emptyset \subset \{a_1\} \subset \{a_1, a_2\} \subset \{a_1, a_2, a_3\}$, 则 $U/\{a_1, a_2, a_3\} < U/\{a_1, a_2\}$, 而 $\rho_{\{a_1, a_2, a_3\}}(X) = \rho_{\{a_1, a_2\}}(X)$. 这表明集合 X 在不同知识粒度的近似空间中可能得到相同的粗糙度. 为了克服这个问题, Liang^[18] 给出一种粗糙熵 $E_B(X) = \rho_B(X)E(B)$, 该粗糙熵随着近似空间中知识粒度减小会有何变化规律呢?

定理 5^[21]. 设格 $\langle P(A), \subseteq \rangle$ 中的任意一条链为 $\emptyset=B_0 \subset B_1 \subset B_2 \subset \dots \subset B_m=A$, 对于任意的 $X \subseteq U$, 若 $G(U/B_{i+1}) < G(U/B_i)$, 则 $E_{B_{i+1}}(X) < E_{B_i}(X)$.

定理 5 表明, 随着分层递阶的近似空间中知识粒度的减小, $E_B(X)$ 严格单调递减. 这个结论在一定程度上弥补了粗糙度的缺陷. 但是, 我们分析发现, 当近似空间中知识粒度的减小是由于集合 X 负域中的知识颗粒(与 X 无关)被细分时, 粗糙度不会改变(符合认知规律), 但粗糙熵 $E_B(X)$ 却严格递减(不符合认知规律). 这表明与集合 X 无关的知识颗粒(X 的负域中的知识颗粒)细分时, 粗糙集的粗糙熵会减小, 与人们的认知规律不吻合.

例 3. 设 $U=\{x_1, x_2, \dots, x_{10}\}$, $U/B_i=\{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6, x_7\}, \{x_8, x_9, x_{10}\}\}$, $U/B_{i+1}=\{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6, x_7\}, \{x_8, x_9\}, \{x_{10}\}\}$, $X=\{x_3, x_4, x_5, x_6, x_7\}$, 则 $\rho_{B_i}(X) = \rho_{B_{i+1}}(X) = \frac{3}{7}$;

$E_{B_i}(X) = \frac{3}{70}(8 + \log_2^{729}), E_{B_{i+1}}(X) = \frac{3}{70}(8 + \log_2^{108}),$
所以 $E_{B_i}(X) > E_{B_{i+1}}(X).$

因此,用粗糙熵度量粗糙集的不确定性还是存在一定的局限性.根据商空间理论中解释“模糊”和“清晰”之间粒度变化的关系“模糊在一定粒度下会变得清晰,而清晰在一定粒度下会变得模糊”和李德毅指出的^[1]“不确定性和确定性并非完全对立,在一定程度上可以相互转化”,本文接下来重点讨论,在分层递阶的近似空间中,粗糙集模糊度随着知识粒度的变化而变化的情况.

设格 $\langle P(A), \subseteq \rangle$ 中的任意一条链为 $\emptyset = B_0 \subset B_1 \subset B_2 \subset \dots \subset B_m = A$, 对于任意的 $X \subseteq U$, 模糊集 $F_X^{B_i}$ 与 $F_X^{B_{i+1}}$ 的模糊度的大小关系如何呢? 对这个问题的讨论要比粗糙度和粗糙熵复杂得多.

(1) 如果 $U/B_i = U/B_{i+1}$, 对任意的模糊性度量方法, $F_X^{B_i}$ 与 $F_X^{B_{i+1}}$ 的模糊度都相等;

(2) 如果 $U/B_{i+1} < U/B_i$, 容易证明: $d_{Kl}(F_X^{B_{i+1}}) \leq d_{Kl}(F_X^{B_i})$. 但 $d_{Kq}(F_X^{B_{i+1}})$ 和 $d_{Kq}(F_X^{B_i})$ 的大小关系不确定.

如例 2 中取 $X = \{x_6, x_7, x_8, x_9\}, U/\{a_1, a_2\} < U/\{a_1\}$, 则

$$\begin{aligned} F_X^{(a_1)} &= \left\{ \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, 1, 1, 0, 0, 0 \right\}; \\ F_X^{(a_1, a_2)} &= \left\{ 0, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, 1, 1, 0, 0, 0 \right\}; \\ d_{Kl}(F_X^{(a_1, a_2)}) &= d_{Kl}(F_X^{(a_1)}) = \frac{2}{12} \left(\frac{2}{7} \times 7 \right) = \frac{1}{3}; \\ d_{Kq}(F_X^{(a_1)}) &= \frac{2}{\sqrt{12}} \left(\left(\frac{2}{7} \right)^2 \times 7 \right) = \frac{4}{7\sqrt{3}}; \\ d_{Kq}(F_X^{(a_1, a_2)}) &= \frac{2}{\sqrt{12}} \left(\left(\frac{2}{6} \right)^2 \times 6 \right) = \frac{4}{6\sqrt{3}}. \end{aligned}$$

这个例子说明,如果用 $d_{Kl}(\cdot)$ 和 $d_{Kq}(\cdot)$ 来测量粗糙集的模糊度有以下缺陷:用 $d_{Kl}(\cdot)$ 来测量 $X = \{x_6, x_7, x_8, x_9\}$ 的模糊度, $U/\{a_1, a_2\} < U/\{a_1\}$, $G(U/\{a_1, a_2\}) < G(U/\{a_1\})$, 且 $\frac{6}{8} = \rho_{\{a_1, a_2\}}(X) < \rho_{\{a_1\}}(X) = \frac{7}{9}$, 这表明随着知识粒度的减小,粗糙度在降低,然而 X 的线性模糊度却不变 $d_{Kl}(F_X^{(a_1, a_2)}) = d_{Kl}(F_X^{(a_1)})$, 二次模糊度反而增加 $d_{Kq}(F_X^{(a_1, a_2)}) > d_{Kq}(F_X^{(a_1)})$, 这与人们的直觉相悖.

4 基于信息熵的粗糙集模糊度

为了能够将信息熵应用来测量粗糙集的模糊

度,我们进一步分析发现:粗糙集的模糊性来自边界的两个部分,一部分是边界域中属于集合 X 的元素,一部分是边界域中不属于集合 X 的元素,而式(8)的信息熵只考虑了前面一部分,没有涉及第二部分.为此,我们提出一种新的基于信息熵的粗糙集的模糊度量方法:

$$\begin{aligned} d_Z(F_X^B) &= -\frac{1}{n \ln 2} \sum_{i=1}^n [\mu_X^B(x_i) \ln \mu_X^B(x_i) + \\ &\quad (1 - \mu_X^B(x_i)) \ln (1 - \mu_X^B(x_i))] \end{aligned} \quad (9)$$

直观上讲,式(9)由 $\mu_X^B(x_i) \ln \mu_X^B(x_i)$ 和 $(1 - \mu_X^B(x_i)) \ln (1 - \mu_X^B(x_i))$ 两部分信息熵构成,前者主要反映属于集合 X 的元素“贡献”的不确定性,后者主要反映不属于集合 X 的元素“贡献”的不确定性,这两部分同时考虑才能更精确地刻画粗糙集的不确定性.接下来,我们验证 $d_Z(\cdot)$ 满足定义 8.

证明. $d_Z(F_X^B) = 0$ 当且仅当 $\forall x_i \in U (\mu_X^B(x_i) = 0 \vee \mu_X^B(x_i) = 1)$, 即 F_X^B 是普通的康托集, $F_X^B \in P(U)$. 定义 8 的条件(1)满足.

对于任意的 $x_i (x_i \in U)$, 令 $\mu_X^B(x_i) = t_i (0 \leq t_i \leq 1)$, 令 $f(t_i) = t_i \ln t_i + (1 - t_i) \ln (1 - t_i)$, 易证,函数 $f(t_i)$ 在唯一的极值点 $t_i = \frac{1}{2}$ 处取得最小值 $-\ln 2$.

所以, $d_Z(F_X^B)$ 在点 $\mu_X^B(x_i) = \frac{1}{2}$ 处取得最大值 1. 定义 8 的条件(2)满足.

对于任意的 $x_i (x_i \in U)$, 由于 $f(t_i) = t_i \ln t_i + (1 - t_i) \ln (1 - t_i)$ 在区间 $[0, \frac{1}{2}]$ 单调递减, 在 $[\frac{1}{2}, 1]$ 单调递增, 在 $t_i = \frac{1}{2}$ 处取得最小值, 所以, $d_Z(F_X^B) = -\frac{1}{n \ln 2} \sum_{i=1}^n f(t_i)$ 在区间 $[0, \frac{1}{2}]$ 单调递增, 在 $[\frac{1}{2}, 1]$ 单调递减, 在 $t_i = \frac{1}{2}$ 处取得最大值. 因此, 当 $\mu_X^B(x_i) = t_i \leq t'_i = \mu_X^{B'}(x_i) \leq \frac{1}{2}$ 或 $\mu_X^B(x_i) = t_i \geq t'_i = \mu_X^{B'}(x_i) \geq \frac{1}{2}$ 时, 有 $d_Z(F_X^B) \leq d_Z(F_X^{B'})$. 定义 8 的条件(3)满足.

$d_Z((F_X^B)^c) = d_Z(F_X^B)$ 显然成立, 定义 8 的条件(4)满足.

综上所述, $d_Z(\cdot)$ 是粗糙集的一种模糊度. 下面, 我们讨论模糊度 $d_Z(\cdot)$ 随近似空间中知识粒度的减小的变化趋势.

定理 6. 设格 $\langle P(A), \subseteq \rangle$ 中的任意一条链为

$\emptyset = B_0 \subset B_1 \subset B_2 \subset \cdots \subset B_m = A$, 如果 $U/B_{i+1} < U/B_i$, 则对于任意的 $X \subseteq U$, 都有 $d_Z(F_X^{B_{i+1}}) \leq d_Z(F_X^{B_i})$.

证明. 设 $U/B_i = \{P_1, P_2, \dots, P_r\}$, $U/B_{i+1} = \{Q_1, Q_2, \dots, Q_t\}$ ($r < t$). 因为, $U/B_{i+1} < U/B_i$, 令 $\Delta B_i = B_i - B_{i+1}$ 表示属性增量. 则属性增量 ΔB_i 一定对 $U/B_i = \{P_1, P_2, \dots, P_r\}$ 中的至少一个元素进行细分. 为简化证明, 我们不妨设 U/B_i 中只有 P_1 被 ΔB_i 分为两个部分 (分为多个部分的证明情况类似), $P_1 = Q_i \cup Q_j$ ($Q_i, Q_j \in U/B_{i+1}$), U/B_i 的其它元素不变 (其它情况可以根据这种情况进行证明). 下面分情况讨论:

(1) 当 $P_1 \cap X = \emptyset$ 时, 对于任意的 $x (x \in P_1)$, $\mu_X^{B_i}(x) = \frac{|P_1 \cap X|}{|P_1|} = 0$. 因为 $P_1 = Q_i \cup Q_j$ ($Q_i \cap Q_j = \emptyset$), 所以, 对于任意 $x (x \in Q_i$ 或者 $x \in Q_j)$, $\mu_X^{B_{i+1}}(x) = \frac{|Q_i \cap X|}{|Q_i|} = \frac{|Q_j \cap X|}{|Q_j|} = 0$. 因此, 属性增量 ΔB_i 对 $U/B_i = \{P_1, P_2, \dots, P_r\}$ 的细分不改变模糊集 $F_X^{B_i}$ 的隶属函数值, 即 $F_X^{B_i} = F_X^{B_{i+1}}$, 所以 $d_Z(F_X^{B_{i+1}}) = d_Z(F_X^{B_i})$.

(2) 当 $P_1 \subseteq X$ 时, 对于任意的 $x (x \in P_1)$, $\mu_X^{B_i}(x) = \frac{|P_1 \cap X|}{|P_1|} = 1$. 由于 $P_1 = Q_i \cup Q_j$ ($Q_i \cap Q_j = \emptyset$), 所以, 对于任意的 $x (x \in Q_i$ 或者 $x \in Q_j)$, $\mu_X^{B_{i+1}}(x) = \frac{|Q_i \cap X|}{|Q_i|} = \frac{|Q_j \cap X|}{|Q_j|} = 1$. 因此, 属性增量 ΔB_i 对 $U/B_i = \{P_1, P_2, \dots, P_r\}$ 的细分不改变模糊集 $F_X^{B_i}$ 的隶属函数值, 即 $F_X^{B_i} = F_X^{B_{i+1}}$, 即 $d_Z(F_X^{B_{i+1}}) = d_Z(F_X^{B_i})$.

(3) 当 $P_1 \cap X \neq \emptyset$, 且 $P_1 \cap X \neq P_1$ 时, 因为 $P_1 = Q_i \cup Q_j$, 则 $|P_1| = |Q_i| + |Q_j|$ ($Q_i \cap Q_j = \emptyset$),

$$\begin{aligned} d_Z(F_X^{B_i}) &= \frac{1}{n \ln 2} \sum_{i=1}^n [-\mu_X^{B_i}(x_i) \ln \mu_X^{B_i}(x_i) - \\ &\quad (1 - \mu_X^{B_i}(x_i)) \ln (1 - \mu_X^{B_i}(x_i))] \\ &= \frac{1}{n \ln 2} \left\{ \sum_{x_i \in P_1} [-\mu_X^{B_i}(x_i) \ln \mu_X^{B_i}(x_i) - \right. \\ &\quad (1 - \mu_X^{B_i}(x_i)) \ln (1 - \mu_X^{B_i}(x_i))] + \\ &\quad \sum_{x_j \notin P_1} [-\mu_X^{B_i}(x_j) \ln \mu_X^{B_i}(x_j) - \\ &\quad (1 - \mu_X^{B_i}(x_j)) \ln (1 - \mu_X^{B_i}(x_j))] \left. \right\}, \end{aligned}$$

下面分类讨论:

① 如果 $Q_i \cap X = \emptyset$, 设 $|P_1 \cap X| = a$ 且 $|P_1| - |P_1 \cap X| = b$, 则公式

$$\begin{aligned} &\sum_{x_i \in P_1} [-\mu_X^{B_i}(x_i) \ln \mu_X^{B_i}(x_i) - (1 - \mu_X^{B_i}(x_i)) \ln (1 - \mu_X^{B_i}(x_i))] \\ &= -a \ln \frac{a}{a+b} - b \ln \frac{b}{a+b}, \end{aligned}$$

$$\begin{aligned} &\sum_{x_i \in Q_i \cup Q_j} [-\mu_X^{B_{i+1}}(x_i) \ln \mu_X^{B_{i+1}}(x_i) - \\ &\quad (1 - \mu_X^{B_{i+1}}(x_i)) \ln (1 - \mu_X^{B_{i+1}}(x_i))] \\ &= \sum_{x_i \in Q_j} [-\mu_X^{B_{i+1}}(x_i) \ln \mu_X^{B_{i+1}}(x_i) - \\ &\quad (1 - \mu_X^{B_{i+1}}(x_i)) \ln (1 - \mu_X^{B_{i+1}}(x_i))] \end{aligned}$$

$$= -a \ln \frac{a}{a+b_1} - b_1 \ln \frac{b_1}{a+b_1},$$

这里 $|Q_j \cap X| = a_1 = a$, $|Q_j| - |Q_j \cap X| = b_1 < b$. 令函数 $f(a, b) = -a \ln \frac{a}{a+b} - b \ln \frac{b}{a+b}$, 因为 $\frac{\partial f}{\partial b} = \ln \frac{a+b}{b} > 0$, 所以 $f(a, b)$ 关于 b 是增函数. 因为 $b_1 < b$, 所以

$$-a \ln \frac{a}{a+b} - b \ln \frac{b}{a+b} \geq -a \ln \frac{a}{a+b_1} - b_1 \ln \frac{b_1}{a+b_1}.$$

② 如果 $Q_j \subseteq X$, 则 $|Q_i \cap X| = a_1 < a$, $|Q_i| - |Q_i \cap X| = b_1 = b$,

$$\begin{aligned} &\sum_{x_i \in Q_i \cup Q_j} [-\mu_X^{B_{i+1}}(x_i) \ln \mu_X^{B_{i+1}}(x_i) - \\ &\quad (1 - \mu_X^{B_{i+1}}(x_i)) \ln (1 - \mu_X^{B_{i+1}}(x_i))] \\ &= \sum_{x_i \in Q_i} [-\mu_X^{B_{i+1}}(x_i) \ln \mu_X^{B_{i+1}}(x_i) - \\ &\quad (1 - \mu_X^{B_{i+1}}(x_i)) \ln (1 - \mu_X^{B_{i+1}}(x_i))] \\ &= -a_1 \ln \frac{a_1}{a_1+b} - b \ln \frac{b}{a_1+b}. \end{aligned}$$

因为 $\frac{\partial f}{\partial a} = \ln \frac{a+b}{a} > 0$, 所以 $f(a, b)$ 关于 a 是增函数. 又因为 $a_1 < a$, 所以 $-a \ln \frac{a}{a+b} - b \ln \frac{b}{a+b} \geq -a_1 \ln \frac{a_1}{a_1+b} - b \ln \frac{b}{a_1+b}$.

③ 如果 $Q_i \cap X \neq \emptyset$ 且 $Q_i \cap X \neq Q_i$, $Q_j \cap X \neq \emptyset$ 且 $Q_j \cap X \neq Q_j$, 令 $|X \cap Q_i| = a_1 > 0$, $|X \cap Q_j| = a_2 > 0$, $|Q_i| - |X \cap Q_i| = b_1 > 0$, $|Q_j| - |X \cap Q_j| = b_2 > 0$, 此时, $a_1 + a_2 = a$, $b_1 + b_2 = b$.

$$\begin{aligned} &\sum_{x_i \in P_1} [-\mu_X^{B_i}(x_i) \ln \mu_X^{B_i}(x_i) - (1 - \mu_X^{B_i}(x_i)) \cdot \\ &\quad \ln (1 - \mu_X^{B_i}(x_i))] = -a \ln \frac{a}{a+b} - b \ln \frac{b}{a+b}; \end{aligned}$$

$$\begin{aligned} &\sum_{x_i \in Q_i \cup Q_j} [-\mu_X^{B_{i+1}}(x_i) \ln \mu_X^{B_{i+1}}(x_i) - \\ &\quad (1 - \mu_X^{B_{i+1}}(x_i)) \ln (1 - \mu_X^{B_{i+1}}(x_i))] \\ &= -a_1 \ln \frac{a_1}{a_1+b_1} - b_1 \ln \frac{b_1}{a_1+b_1} - \end{aligned}$$

$$(a-a_1)\ln \frac{a-a_1}{a-a_1+b-b_1}-$$
$$(b-b_1)\ln \frac{b-b_1}{a-a_1+b-b_1}.$$

又令 $F(a_1, b_1) = -a_1 \ln \frac{a_1}{a_1+b_1} - b_1 \ln \frac{b_1}{a_1+b_1} -$

$$(a-a_1)\ln \frac{a-a_1}{a-a_1+b-b_1} - (b-b_1)\ln \frac{b-b_1}{a-a_1+b-b_1},$$

求解 $F(a_1, b_1)$ 的最大值. 对 $F(a_1, b_1)$ 求偏导数, 得方程组:

$$\begin{cases} \frac{\partial F}{\partial a_1} = 0 \\ \frac{\partial F}{\partial b_1} = 0 \end{cases}.$$

解该方程组得: $\frac{a_1}{b_1} = \frac{a}{b}$. 此时, $\frac{a_2}{b_2} = \frac{a}{b}$, 这表明函数

$F(a_1, b_1)$ 在 $\frac{a_1}{b_1} = \frac{a_2}{b_2} = \frac{a}{b}$ 处取得最大值 $-a \ln \frac{a}{a+b} -$

$b \ln \frac{b}{a+b}$. 所以, $F(a_1, b_1) \leq -a \ln \frac{a}{a+b} - b \ln \frac{b}{a+b}$.

根据以上①, ②和③, 有

$$\sum_{x_i \in P_1} [-\mu_X^{B_i}(x_i) \ln \mu_X^{B_i}(x_i) - (1-\mu_X^{B_i}(x_i)) \ln (1-\mu_X^{B_i}(x_i))] \geq \sum_{x_i \in Q_i \cup Q_j} [-\mu_X^{B_i}(x_i) \ln \mu_X^{B_i}(x_i) - (1-\mu_X^{B_i}(x_i)) \cdot \ln (1-\mu_X^{B_i}(x_i))].$$

所以, $d_Z(F_{X^{B_i+1}}^{B_i}) \leq d_Z(F_{X^{B_i}}^{B_i})$.

综上所述, 定理 6 得证. 证毕.

当属性增量 ΔB_i 将 P_1 划分为 Q_i, Q_j ($Q_i \neq \emptyset, Q_j \neq \emptyset, Q_i \cap Q_j \neq \emptyset$) 两个细的知识颗粒时, 即 $P_1 = Q_i \cup Q_j, X \subseteq U$, 如果

$$\frac{|P_1 \cap X|}{|P_1| - |P_1 \cap X|} = \frac{|Q_i \cap X|}{|Q_i| - |Q_i \cap X|} = \frac{|Q_j \cap X|}{|Q_j| - |Q_j \cap X|},$$

则称 P_1 被属性增量 ΔB_i “成比例”细分. 特别地, 当 $Q_i = \emptyset$ 或者 $Q_j = \emptyset$ (即 P_1 没有被分解) 时, 我们视为一种特殊的“成比例”细分.

推论 1. 当且仅当属性增量 ΔB_i 将 $U/B_i = \{P_1, P_2, \dots, P_r\}$ 中的每个知识颗粒进行“成比例”细分时, 有 $d_Z(F_X^{B_i}) = d_Z(F_X^{B_i+1})$.

如果属性增量 ΔB_i 将 $U/B_i = \{P_1, P_2, \dots, P_r\}$ 中的某个知识颗粒进行“不成比例”细分后, 粗糙集的模糊性将严格递减.

续例 2. 设 $X = \{x_6, x_7, x_8, x_9\}, U/\{a_1, a_2\} \prec U/\{a_1\}$, 则

$$F_X^{(a_1)} = \left\{ \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, \frac{2}{7}, 1, 1, 0, 0, 0 \right\};$$
$$F_X^{(a_1, a_2)} = \left\{ 0, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, \frac{2}{6}, 1, 1, 0, 0, 0 \right\}.$$

由第 3 节可知, $d_{K_q}(F_X^{(a_1, a_2)}) = \frac{4}{6\sqrt{3}} > d_{K_q}(F_X^{(a_1)}) =$

$\frac{4}{7\sqrt{3}}$, 这与人们的认知相悖. 用 $d_Z(\cdot)$ 计算得

$$d_Z(F_X^{(a_1)}) = \frac{-1}{12\ln 2} \left(\left(\frac{2}{7} \ln \frac{2}{7} + \frac{5}{7} \ln \frac{5}{7} \right) \times 7 + 0 \right) = \frac{1}{12\ln 2} \times 65.88;$$

$$d_Z(F_X^{(a_1, a_2)}) = \frac{-1}{12\ln 2} \left(\left(\frac{2}{6} \ln \frac{2}{6} + \frac{4}{6} \ln \frac{4}{6} \right) \times 6 + 0 \right) = \frac{1}{12\ln 2} \times 45.56.$$

可见, $d_Z(F_X^{(a_1)}) > d_Z(F_X^{(a_1, a_2)})$. 这说明二次模糊度随着知识粒度的减小反而增加, $d_Z(\cdot)$ 随着知识粒度的减小而单调递减.

续例 3. 设 $U = \{x_1, x_2, \dots, x_{10}\}, U/B_i = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6, x_7\}, \{x_8, x_9, x_{10}\}\}, U/B_{i+1} = \{\{x_1, x_2, x_3\}, \{x_4, x_5, x_6, x_7\}, \{x_8, x_9\}, \{x_{10}\}\}, X = \{x_3, x_4, x_5, x_6, x_7\}$, 则 $\rho_{B_i}(X) = \rho_{B_{i+1}}(X) = \frac{3}{7};$

$$E_{B_i}(X) = \frac{3}{70} (8 + \log_2^{729}), E_{B_{i+1}}(X) = \frac{3}{70} (8 + \log_2^{108}),$$

所以, $E_{B_i}(X) > E_{B_{i+1}}(X)$; 而模糊集 $F_X^{B_i}$ 和 $F_X^{B_{i+1}}$ 相等 ($F_X^{B_i} = F_X^{B_{i+1}} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}, 1, 1, 1, 1, 0, 0, 0)$), 所以, $d_Z(F_X^{B_i}) = d_Z(F_X^{B_{i+1}})$. 这说明与集合 X 无关的知识颗粒的细化导致粗糙熵严格递减, 但模糊度 $d_Z(\cdot)$ 不变.

式(9)依赖两部分信息熵, 既利用了度量不确定性的 Shannon 熵, 也结合粗糙集的特点, 同时构造集合 X 的边界域中属于 X 的那部分元素“贡献”的不确定性和不属于 X 的那部分元素“贡献”的不确定性, 非常直观. 随着集合 X 的边界域上的知识颗粒的“不成比例”的细分, 粗糙集的模糊度将严格递减; 而集合 X 边界域上的知识颗粒被“成比例”细分时, 粗糙集的模糊度不变. 这更加准确地刻画出人们对不确定性问题的认知规律.

5 不同知识粒度下粗糙集的粗糙度和模糊度的变化关系

这里主要讨论粗糙度 $\rho_B(\cdot)$ 和模糊度 $d_Z(\cdot)$ 在分层递阶的近似空间中随知识粒度的变化而变化的关系.

性质 1. 设信息系统 $S = (U, A, V, f)$ 中, $B \subseteq$

$A, X \subseteq U$, 如果 X 是关于 B 精确的, 则 $d_{Kl}(F_X^B) = d_{Kq}(F_X^B) = d_Z(F_X^B) = 0$.

这个性质表明, 任何关于 B 的精确集的模糊度都等于 0. 除了用数值来表示粗糙集的不确定特征外, 也可以用拓扑特征^[22]来刻画.

- (1) 如果 $\underline{B}(X) \neq \emptyset, \bar{B}(X) \neq U$, 称 X 是粗糙可定义的;
- (2) 如果 $\underline{B}(X) = \emptyset, \bar{B}(X) \neq U$, 称 X 是内不可定义的;
- (3) 如果 $\underline{B}(X) \neq \emptyset, \bar{B}(X) = U$, 称 X 是外不可定义的;
- (4) 如果 $\underline{B}(X) = \emptyset, \bar{B}(X) = U$, 称 X 是全不可定义的.

性质 2. 设信息系统 $S = (U, A, V, f)$ 中, $B \subseteq A, X \subseteq U$, 如果 $d_Z(F_X^B) = 1$, 则 $\rho_B(X) = 1$, 且 X 关于 B 是全不可定义的.

这个性质的逆不一定成立. 即一个集合 X 关于 B 是全不可定义时, 它的粗糙度等于 1, 但是模糊度不一定等于 1.

性质 3. 设信息系统 $S = (U, A, V, f)$ 中, $B_1 \subset B_2 \subseteq A, X \subseteq U$, 且 $\rho_{B_1}(X) > \rho_{B_2}(X)$, 则 $d_Z(F_X^{B_1}) > d_Z(F_X^{B_2})$.

该性质说明, 在一条属性链上, 如果粗糙度降低, 必然导致粗糙集的模糊度降低. 该性质的逆不一定成立, 即如果 $d_Z(F_X^{B_1}) > d_Z(F_X^{B_2})$ 时, $\rho_{B_1}(X) > \rho_{B_2}(X)$ 不一定成立. 这说明粗糙集的模糊度降低时, 粗糙度未必降低.

性质 4. 设信息系统 $S = (U, A, V, f)$ 中, $B_1 \subset B_2 \subseteq A, X \subseteq U$, 如果 $d_Z(F_X^{B_1}) = d_Z(F_X^{B_2})$, 则 $\rho_{B_1}(X) = \rho_{B_2}(X)$.

随着分层递阶的近似空间中知识粒度的减小, 如果粗糙集的模糊度不变, 则粗糙集的粗糙度也不变. 性质 3 和性质 4 表明模糊度比粗糙度对知识粒度的变化更“灵敏”.

性质 5. 设信息系统 $S = (U, A, V, f)$ 中, $B_1 \subset B_2 \subseteq A$, 且 $G(U/B_1) > G(U/B_2)$, 则 $\rho_{B_1}(X) \geq \rho_{B_2}(X)$, $d_Z(F_X^{B_1}) \geq d_Z(F_X^{B_2})$.

随着近似空间中的知识粒度的减小, 粗糙集的粗糙度、模糊度不一定严格递减.

性质 6. 设信息系统 $S = (U, A, V, f)$ 中, $B_1 \subset B_2 \subseteq A, X \subseteq U$, 如果 $\rho_{B_1}(X) > \rho_{B_2}(X)$ 或者 $d_Z(F_X^{B_1}) > d_Z(F_X^{B_2})$, 则 $G(U/B_1) > G(U/B_2)$.

在分层递阶的近似空间中, 知识粒度随着粗糙集的粗糙度或模糊度的降低而必然降低.

性质 7. 设信息系统 $S = (U, A, V, f)$ 中, $B \subseteq A, X \subseteq U$, 若 $\rho_B(X) < 1$, 则 $d_Z(F_X^B) < 1$.

在分层递阶的近似空间中, 这些性质刻画了粗糙度、模糊度随知识粒度的变化而变化的规律. 本文给出的粗糙集模糊度 $d_Z(\cdot)$ 随知识粒度的变化规律更加符合人们的认知规律.

6 结束语

粗糙集的粗糙度、粗糙熵、模糊度和模糊熵虽然都是度量粗糙集的不确定性的, 但它们之间有一定的联系和区别. 粗糙性从集合的边界区域的角度来刻画粗糙集的不确定性, 随着知识粒度的减小, 如果集合的边界区域变小, 粗糙度降低, 粗糙集的不确定性下降, 具有很好的直观性; 而粗糙集的模糊性用元素属于某个集合的隶属函数的大小来刻画粗糙集的不确定性, 与集合的边界区域大小和知识粒度的大小有关. 粗糙集的粗糙性具有一定的“几何”特点, 而模糊性具有一定的“代数”特点, 它们从直观和抽象两个方面分别刻画出粗糙集的不确定性, 具有一定的互补性. 本文从信息熵的角度提出了一种新的粗糙集的模糊性度量方法 $d_Z(\cdot)$, 该方法用“两个”方面的信息熵来刻画粗糙集的不确定性, 具有非常形象和直观的特点. 这种模糊度既有信息熵度量不确定性的优势(是两部分信息熵构成), 又能克服粗糙度和 Liang^[18]定义的粗糙熵对粗糙集不确定性度量的不足, 也能弥补 Chakrabarty^[10]等人提出的粗糙集模糊度随着知识粒度减小反而可能增加的缺陷. 在研究中, 我们发现不是满足定义 8 的任意模糊度都会随着知识颗粒的细分而严格单调递减, 因此, 在构造测量粗糙集的模糊度的测量方法时, 除了满足定义 8 外, 还应该增加一个约束条件: 随着知识颗粒的细分, 粗糙集的模糊度单调递减.

参 考 文 献

[1] Li De-Yi, Liu Chang-Yu, Du Yi, Han Xu. Artificial intelligence with uncertainty. Journal of Software, 2004, 15(11): 1583-1594(in Chinese)
(李德毅, 刘常昱, 杜钰, 韩旭. 不确定性人工智能. 软件学报, 2004, 15(11): 1583-1594)

- [2] Liu Jie, Chen Xiao-Ping, Cai Qing-Sheng, Fan Yan. Recognition structure of uncertainty: A unified framework for representation, reasoning and learning. *Journal of Software*, 2002, 13(4): 649-651(in Chinese)
(刘洁, 陈小平, 蔡庆生, 范焱. 不确定信息的认知结构表示、推理和学习. *软件学报*, 2002, 13(4): 649-651)
- [3] Zadeh L A. Fuzzy sets. *Information and Control*, 1965, 8: 338-353
- [4] Pawlak Z. Rough sets. *International Journal of Computer and Information Science*, 1982, 11: 341-356
- [5] Zhang Bo, Zhang Ling. *Theory of Problem Solving and Its Applications*. Beijing: Tsinghua University Press, 1990(in Chinese)
(张钹, 张铃. *问题求解理论及应用*. 北京: 清华大学出版社, 1990)
- [6] Wang Guo-Yin. *Rough Set Theory and Knowledge Discovery*. Xi'an: Xi'an Jiaotong University Press, 2001(in Chinese)
(王国胤. *粗糙集理论与知识获取*. 西安: 西安交通大学出版社, 2001)
- [7] Zhang Ling, Zhang Bo. *The Theory and Applications of Problem Solving-Quotient Space Based Granular Computing*. 2nd Version. Beijing: Tsinghua University Press, 2007(in Chinese)
(张铃, 张钹. *问题求解理论及应用-商空间粒度计算理论及其应用*. 第2版. 北京: 清华大学出版社, 2007)
- [8] Gau W L, Buehrer D J. Vague sets. *IEEE Transactions on Systems, Man and Cybernetics*, 1993, 23(2): 610-614
- [9] Wang Guo-Yin, Zhang Qing-Hua, Hu Jun. An overview of granular computing. *CAAI Transactions on Intelligent Systems*, 2007, 2(6): 8-26(in Chinese)
(王国胤, 张清华, 胡军. 粒计算研究综述. *智能系统学报*, 2007, 2(6): 8-26)
- [10] Chakrabarty K, Biswas R, Nanda S. Fuzziness in rough sets. *Fuzzy Sets and Systems*, 2000, 110: 247-251
- [11] Banerjee M, Pal S K. Roughness of a fuzzy set. *Information Sciences*, 1996, 93: 235-246
- [12] Huynh V H, Nakamori Y. A roughness measure for fuzzy sets. *Information Sciences*, 2005, 73: 255-275
- [13] Wang Guo-Yin, Yu Hong, Yang Da-Chun. Decision table reduction based on conditional information entropy. *Chinese Journal of Computers*, 2002, 25(7): 1-8(in Chinese)
(王国胤, 于洪, 杨大春. 基于条件信息熵的决策表约简. *计算机学报*, 2002, 25(7): 1-8)
- [14] Wang Guo-Yin, Zhao Jun, An Jiu-Jing, Wu Yu. A comparative study of algebra viewpoint and information viewpoint in attribute reduction. *Fundamenta Informaticae*, 2005, 68(3): 289-301
- [15] Zhao Jun, Wang Guo-Yin. Research on system uncertainty measures based on rough set theory//*Proceedings of the RSKT2006*. Chongqing, 2006: 227-232
- [16] Liang J Y, Chin K S, Dang C Y. A new method for measuring uncertainty and fuzziness in rough set theory. *International Journal of General Systems*, 2002, 31(4): 331-342
- [17] Liang Ji-Ye, Li De-Yu. *Uncertainty and Knowledge Acquisition in Information Systems*. Beijing: Science Press, 2005(in Chinese)
(梁吉业, 李德玉. *信息系统中的不确定性与知识获取*. 北京: 科学出版社, 2005)
- [18] Liang J Y, Shi Z Z. The information entropy, rough entropy and knowledge granulation in rough set theory. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2004, 12(1): 37-46
- [19] Miao Duo-Qian, Fan Shi-Dong. The calculation of knowledge granulation and its application. *Systems Engineering-theory & Practice*, 2002, 22(1): 48-56(in Chinese)
(苗夺谦, 范世栋. 知识的粒度计算及其应用. *系统工程理论与实践*, 2002, 22(1): 48-56)
- [20] Miao Duo-Qian, Wang Jue. An information representation of the concepts and operations in rough set theory. *Journal of Software*, 1999, 10(2): 113-116(in Chinese)
(苗夺谦, 王珏. 粗糙集理论中概念与运算的信息表示. *软件学报*, 1999, 10(2): 113-116)
- [21] Miao Duo-Qian, Wang Guo-Yin, Liu Qing et al. *Granular Computing: Past, Present and Future Prospects*. Beijing: Science Press, 2007(in Chinese)
(苗夺谦, 王国胤, 刘清等. *粒计算: 过去、现在与展望*. 北京: 科学出版社, 2007)
- [22] Zhang Wen-Xiu, Wu Wei-Zhi, Liang Ji-Ye, Li De-Yu. *Theory and Methods of Rough Sets*. Beijing: Science Press, 2001(in Chinese)
(张文修, 吴伟志, 梁吉业, 李德玉. *粗糙集理论与方法*. 北京: 科学出版社, 2001)
- [23] Yang Lun-Biao, Gao Yi-Ying. *Fuzzy Mathematics: Theory and Application*. Guangzhou: South China University of Technology Press, 2004(in Chinese)
(杨纶标, 高英仪. *模糊数学原理及应用*. 广州: 华南理工大学出版社, 2004)
- [24] Liu X C. Entropy, distance measure and similarity measure of fuzzy sets and their relations. *Fuzzy Sets Systems*, 1992, 52: 305-318
- [25] Shannon C E. A mathematical theory of communication. *The Bell System Technical Journal*, 1948, 27: 379- 423
- [26] Klir G J, Wierman M J. *Uncertainty Based Information*. New York: Physica-Verlag, 1998



WANG Guo-Yin, born in 1970, Ph. D. , professor, Ph. D. supervisor. His research interests include Rough set, granular computing, data mining and knowledge technology etc.

ZHANG Qing-Hua, born in 1974, Ph. D. candidate. His main research interests include intelligent information processing and granular computing etc.

Background

Uncertainty is an important property of uncertain set theories. Roughness, rough entropy, fuzziness, and fuzzy entropy are major methods for measuring the uncertainty of rough sets. However, we find that the regularities of the changing of rough entropy and fuzziness of a rough set with the knowledge granularity are inconsistent with human cognition.

Granular computing (GrC) is a new method for simulating human thinking and solving complicated problems. It could be regarded as an umbrella covering the theories, methodologies and techniques of granularity. From the view of granular computing, the uncertainty of rough set should vary in different knowledge granularity levels. We find there is some limitation for roughness and rough entropy to measure the uncertainty of a rough set in different knowledge

granularity levels.

A new method for measuring the fuzziness of rough sets is proposed based on information entropy in this paper. The fuzziness measured by the new method is monotonously decreasing with the refining of knowledge granularity in approximation spaces. It overcomes the problem of roughness and rough entropy. In addition, the relations of the changing of roughness and fuzziness in different knowledge granularities are analyzed.

This work is partially supported by the National Natural Science Foundation of China (No. 60573068 and No. 60773113), Science & Technology Research Program of the Municipal Education Committee of Chongqing of China (No. KJ060517) and Natural Science Foundation of Chongqing of China (No. 2008BA2017).