

证据合成的格结构研究

熊淑华 王加阳

(中南大学信息科学与工程学院 长沙 410083)

摘要 在粗糙集和证据理论体系结构的基础上,分析粗糙集和证据理论中合成质量函数的不同方法,研究细分偏序格中划分与证据理论中的正交组合运算的关系。利用粒度分析细分偏序格上划分的变化情况,证明了在细分偏序格的上、下确界与所有上、下界所对应的划分中,获取的质量函数并不对应于证据理论中质量函数正交运算的结论,从而澄清了不同知识粒度下粗糙集中所有划分获取的质量函数与证据理论中正交运算之间的关系。

关键词 划分,细分偏序格,粒度,质量函数

中图分类号 TP18 **文献标识码** A

Research of Lattice in Evidence Synthesis

XIONG Shu-hua WANG Jia-yang

(School of Information Science and Engineering, Central South University, Changsha 410083, China)

Abstract On the basis of architecture of rough set theory and evidence, this paper analyzes different methods of mass functions in rough set theory and evidence synthesis, and researches the relationship between the sub-division in partial sequence stratigraphic and evidence synthesis. According to the sub-division in partial sequence stratigraphic changes, this paper demonstrated that mass function which is got in the sub-division of supremum, infimum and all upper and lower bounds in the sub-partial order on the lattice, does not correspond to the mass function of evidence synthesis in evidence theory. Therefore, it clarifies the relationship between the mass function acquired from the sub-division in different knowledge granularity levels and mass function got from evidence theory.

Keywords Division, Partial sequence stratigraphic subdivision, Granularity, Mass function

1 引言

在粗糙集理论中,一个对象是否属于一个论域并不是该对象的客观性质,而是取决于我们对它的认知程度^[1-4]。根据认知会对论域形成划分,粗糙集的基本操作都是针对划分的运算,而在证据理论中的基本操作是证据的正交和。尽管粗糙集理论与证据理论之间的关系非常密切,但是两者之间仍然存在很多问题有待澄清。粗糙集中基于细分关系的所有划分构成细分偏序格。论域上任意两个划分的积导出的等价关系是这两个划分所导出的等价关系的交,任意两个划分的和导出的等价关系是这两个划分所导出的等价关系的并集的传递闭包^[5-12]。通过不同粒度的划分,可以获取不同的质量函数。证据理论中证据的合成则是通过对质量函数进行正交运算得到的。有学者指出划分的交所对应的质量函数并不对应于证据理论中的正交和运算,并举反例说明了此问题^[1]。本文在此基础上,研究细分偏序格中的划分随粒度不同,获取的质量函数的变化情况。质量函数随粒度的单调减小而单调不减,随粒度的单调增加而单调不减。最后研究细分偏序格上是否存在一个划分,使得它获取的质量函数对应于证据理论

中两个质量函数的正交运算。

2 细分偏序格

在粗糙集中,论域 U 上任意的等价关系都会对论域 U 形成划分, U 上所有的划分构成划分族。在论域 U 的划分族中引入细分的概念,可以对 U 上两个不同的划分进行比较。划分的这种细分关系是一种偏序关系,并且这种偏序关系满足格的定义,因此称之为细分偏序格。下文将对这些概念进行定义。

定义 1(细分) 设 R 与 R' 是论域 U 上的任意两个等价关系,这两个等价关系所导出的划分分别为 $U/R=\{X_1, X_2, \dots, X_n\}$, $U/R'=\{Y_1, Y_2, \dots, Y_m\}$, 如果 $\forall Y_i \in U/R', \exists X_j \in U/R$, 使得 $Y_i \subseteq X_j$, 则称划分 U/R' 细分 U/R , 记作 $U/R' \leq U/R$; 如果 $\forall Y_{i_0} \in U/R', \exists X_{j_0} \in U/R$ 使得 $Y_{i_0} \subset X_{j_0}$, 则称划分 U/R' 严格细分 U/R , 记作 $U/R' < U/R$ 。

定理 1 设 K 是论域 U 上的划分族,则细分关系是划分族 K 上的一种偏序关系。

定理 2^[5] 设在细分偏序格上, R_1 与 R_2 是论域 U 上的任意两个等价关系, $R=R_1 \cap R_2$ 是 R_1 与 R_2 的积。

到稿日期:2012-02-12 返修日期:2012-07-20 本文受国家自然科学基金项目(61173052),湖南省自然科学基金项目(11JJ504)资助。

熊淑华(1985—),女,硕士生,主要研究方向为信息融合, E-mail: xiongshuhua120@126.com; 王加阳(1963—),男,博士,教授,主要研究方向为智能计算、信息融合。

定理 3 设 P 表示论域 U 上所有划分的集合,在偏序集合 $\langle P, \text{细分} \rangle$ 中,设 R_1 与 R_2 是论域 U 上的任意两个等价关系, U/R_1 与 U/R_2 分别是它们对应的划分,且 $U/R_1, U/R_2 \in P$, 则 $R=R_1 \cap R_2$ 是 R_1 与 R_2 的下确界。

定理 4^[5] 设在细分偏序格上, R_1 与 R_2 是论域 U 上的任意两个等价关系, $R=(R_1 \cup R_2)^+ = t(R_1 \cup R_2)$ 是 R_1 与 R_2 的和。

定理 5 设 P 表示论域 U 上所有划分的集合,在偏序集合 $\langle P, \text{细分} \rangle$ 中,设 R_1 与 R_2 是论域 U 上的任意两个等价关系, U/R_1 与 U/R_2 分别是它们对应的划分,且 $U/R_1, U/R_2 \in P$, 则 $R=(R_1 \cup R_2)^+ = t(R_1 \cup R_2)$ 是 R_1 与 R_2 的上确界。

上述定理指出关于论域 U 的划分族上的细分关系是一种偏序关系,且这种偏序关系满足了格的条件,因此它是一种特殊的格,将它定义为细分偏序格。

定义 2(细分偏序格) K 是论域 U 上的划分族,细分关系 $\langle U, \text{细分} \rangle$ 是 K 上的一个偏序集合,对于 K 上的任意两个等价关系 R_1, R_2 而言,它们的积 $R_1 \cap R_2$ 是下确界,它们的和 $t(R_1 \cup R_2)$ 是上确界,则将其称之为细分偏序格。

3 细分偏序格中划分与正交运算的关系

由第 2 节中细分偏序格存在上、下确界可知,基于细分关系的任意等价关系存在上、下确界。粗糙集是通过划分来获取知识表示的不确定性,而证据理论中提出的 Dempster 合成公式是通过证据函数的正交运算来达到相同的目的。两个理论在表述上有差异,但在细分偏序格上是否存在划分对应于证据理论的正交运算,还有待于澄清。

3.1 上下确界与证据正交运算的关系

设 R_1 与 R_2 是论域 U 上的任意两个等价关系, $R=R_1 \cap R_2$ 是细分偏序格上两个等价关系 R_1 与 R_2 的下确界。 $R=R_1 \cap R_2$ 会对应于一个对论域 U 的划分 U/R ,在此划分上可以构建一个质量函数。文献[1]对粗糙集中任意两个等价关系的下确界是否对应于证据理论中的正交组合进行了研究,并通过反例说明两者没有对应关系。

$R'=t(R_1 \cup R_2)$ 是细分偏序格上任意两个等价关系 R_1 与 R_2 的上确界。它同样对应于一个对论域 U 的划分 U/R' ,在此划分上可以构建一个质量函数。下面通过反例来说明两者之间没有对应关系。

例 令论域 $U=\{1,2,3,4,5,6\}$, R_1 与 R_2 分别为论域 U 上任意两个等价关系,它们所导出的划分如下:

$$U/R_1 = \{\{1,6\}, \{2,3\}, \{4,5\}\}, U/R_2 = \{\{1,3,4,5,6\}, \{2\}\}$$

则它们对应的质量函数如下:

(1) 等价关系 R_1 对应的质量函数 m_1 如下:

$$m_1(\{1,6\}) = |\{1,6\}|/|U| = 1/3, m_1(\{2,3\}) = |\{2,3\}|/|U| = 1/3, m_1(\{4,5\}) = |\{4,5\}|/|U| = 1/3, m_1(\text{其他}) = 0$$

(2) 等价关系 R_2 对应的质量函数 m_2 如下:

$$m_2(\{1,3,4,6\}) = |\{1,3,4,6\}|/|U| = 5/6,$$

$$m_2(\{2\}) = |\{2\}|/|U| = 1/6, m_2(\text{其他}) = 0$$

等价关系 R_1 与 R_2 的和 $t(R_1 \cup R_2)$ 对应的划分如下:

$$U/(t(R_1 \cup R_2)) = \{\{1,2,3,4,5,6\}\}$$

则它对应的质量函数为:

$$m_{12}(\{1,2,3,4,5,6\}) = 1$$

而通过 DS 合成规则对两个证据进行合成,正交和为:

$$m_1 \oplus m_2(U) = K^{-1} \sum_{X \cap Y = U} m_1(X) m_2(Y) = 0$$

由此可见 $m_{12} \neq m_1 \oplus m_2$ 。

由此可知,证据理论中质量函数的正交运算也不等于划分的和所获取的质量函数,则细分偏序格中的上、下确界获取的质量函数都不对应于证据理论中的正交运算。细分偏序格上的两个等价关系除了存在上、下确界外,还存在多个上、下界,那么是否可以在上、下界中找到一种划分,使得划分对应的质量函数对应于证据理论中两个质量函数的正交组合运算呢?

3.2 上下界与证据理论的关系

上节中指出细分偏序格中的上、下确界并不对应于证据理论中的正交运算。下面分析上、下界所获取的质量函数与证据理论中正交运算的关系。

定义 3(细分偏序格上的上下界) 设在细分偏序格上, R_1 与 R_2 是论域 U 上的任意两个等价关系,且 R' 也是论域 U 上的等价关系。则:

(1) 若 R' 细分 R_1 与 R_2 , 即 $R' \subseteq R_1$ 且 $R' \subseteq R_2$, 则称等价关系 R' 为 R_1 与 R_2 的下界;

(2) 若 R_1 与 R_2 细分 R' , 即 $R_1 \subseteq R'$ 且 $R_2 \subseteq R'$, 则称等价关系 R' 为 R_1 与 R_2 的上界。

定义 4(粒度)^[6] 设 R 是论域 U 上任意的等价关系关系,它所生成的所有等价类 $\{[x]_R\}_{x \in U}$ 称为论域 U 上的一种粒度,记为 $GD(R)$,并通过下列公式计算:

$$GD(R) = \frac{1}{|U|^2} \sum_{i=1}^n |X_i|^2$$

定理 6(粒度的取值范围) 设在细分偏序格上, R 是论域 U 上的任意等价关系。则知识 R 的粒度满足:

$$1/|U| \leq GD(R) \leq 1$$

(1) 当且仅当 R 为恒等关系 I 时, $GD(R) = 1/|U|$, 此时每个知识颗粒对应的质量函数均为 $m(\{x_i\}) = 1/|U|$; (2) 当且仅当 R 为全等关系 δ 时, $GD(R) = 1$, 此时整个论域形成一个知识颗粒,它对应的质量函数为 $m(U) = 1$ 。

设在细分偏序格上, R_1 与 R_2 是论域 U 上的任意两个等价关系,且它们导出的划分分别为 $U/R_1 = \{X_1, X_2, \dots, X_n\}$ 与 $U/R_2 = \{Y_1, Y_2, \dots, Y_m\}$ 。

定理 7 设论域 U 上任意两个等价关系 R_1 与 R_2 的下确界为 R' , 等价关系 R' 所导出的划分为 $\{A_1, A_2, \dots, A_p\}$; R' 是 R_1 与 R_2 的任意下界,且它导出的划分为 $U/R' = \{B_1, B_2, \dots, B_k\}$ 。对于所有的下界,都有下界中的划分对应的质量函数并不对应于证据理论中质量函数的正交组合运算,即

$$m(B) \neq \sum_{X \cap Y = B} m(X) m(Y)$$

$$\text{证明: 下确界中的 } m(A) \neq \sum_{X \cap Y = A} m(X) m(Y)。$$

因为其他下界与下确界相比都较细,即 $R' < R'$, 所以 $k > p$ 。则存在论域 $\{1, 2, \dots, k\}$ 的一个划分 E 使得 $E = \{E_1, E_2, \dots, E_p\}$ 满足 $A_j = \bigcup B_i (i \in E_j)$:

① 当 $|E_j| = 1$ 时, $A_j = B_i$, 根据上节中下确界对应的划分的质量函数并不对应于质量函数的正交运算可知, $m(B_i) \neq$

$$\sum_{X \cap Y = B_i} m(X) m(Y);$$

② 当 $|E_j| > 1$ 时, $A_j = \bigcup B_i (i \in E_j)$, 则下界对应的质量函数为:

- 小型微型计算机系统, 2005, 26(10): 1761-1769
- [2] 王国胤. 决策表核属性的计算方法[J]. 计算机学报, 2003, 26(5): 611-615
- [3] 赵军, 王国胤, 吴中福. 一种高效的属性核计算方法[J]. 小型微型计算机系统, 2003, 24(11): 1950-1953
- [4] 闫德勤. 不相容信息系统的规范化格式与差别矩阵[J]. 计算机工程应用, 2004, 40(36): 45-46
- [5] 徐凤生, 李海军. 不相容决策表的求核方法[J]. 计算机工程与科学, 2007, 29(11): 84-85
- [6] 胡玉文, 徐久成. 多粒度时间序列下粒度决策的演化模型研究[J]. 计算机工程与应用, 2011, 47(20): 117-120
- [7] 朱·弗登博格, 让·梯若尔. 博弈论[M]. 北京: 中国人民大学出版社, 2010

- 版社, 2010
- [8] 管廷全, 朱天博. 博弈论的粗集模型[J]. 中国传媒大学学报: 自然科学版, 2010, 17(2): 19-26
- [9] 万武辉, 陆鑫. 基于混沌理论的时间序列区间预测[J]. 成都信息工程学院学报, 2005, 20(5): 527-528
- [10] 刘伟, 王科俊, 邵克勇. 混沌时间序列的混合粒子群优化预测[J]. 控制与决策, 2007, 22(5): 562-565
- [11] 韩敏, 魏茹. 基于改进典型相关分析的混沌时间序列预测[J]. 大连理工大学学报, 2008, 48(2): 292-297
- [12] Taha H. Operations Research an Introduction eighth edition [M]. Beijing, Post & Telecom Press, 2008: 735-738

(上接第 215 页)

$$m(B_i) = |B_i|/|U| \neq 0$$

证据函数的正交运算为:

$$m_1 \oplus m_2(B_i) = \sum_{X \cap Y = B_i} m(X)m(Y)$$

由于 $R' = R_1 \cap R_2$, 则 $X \cap Y \in R'$, 即 $X \cap Y = A_j (B_i \subset A_j)$, 因此 $X \cap Y \neq B_i$.

故对于 $B_i \subset A_j$, $m_1 \oplus m_2(B_i) = 0$, 则 $m(B_i) \neq m_1 \oplus m_2(B_i)$.

综上所述, $m(B) \neq \sum_{X \cap Y = B} m(X)m(Y)$.

定理 8 设论域 U 上任意两个等价关系 R_1 与 R_2 , R' 是 R_1 与 R_2 的任意上界, 且它导出的划分为 $U/R' = \{A_1, A_2, \dots, A_p\}$, R_1 与 R_2 的下确界为 R'' , 等价关系 R'' 所导出的划分为 $U/R'' = \{B_1, B_2, \dots, B_k\}$. 对于所有的上界, 都有上界中的划分对应的质量函数并不对应于证据理论中证据函数的正交组合运算, 即 $m(A) \neq \sum_{X \cap Y = A} m(X)m(Y)$.

证明: 下确界中 $m(B) \neq \sum_{X \cap Y = B} m(X)m(Y)$.

因为所有上界与下确界相比都较粗, 即 $R'' < R'$, 所以 $k > p$. 则存在论域 $\{1, 2, \dots, k\}$ 的一个划分 E 使得 $E = \{E_1, E_2, \dots, E_p\}$ 满足 $A_j = \bigcup_{i \in E_j} B_i$:

① 当 $|E_j| = 1$ 时, 可知 $m(A_j) \neq \sum_{X \cap Y = A_j} m(X)m(Y)$;

② 当 $|E_j| > 1$ 时, $A_j = \bigcup_{i \in E_j} B_i (i \in E_j)$, 则上界对应的质量函数为:

$$m(A_j) = |A_j|/|U| \neq 0$$

证据函数的正交运算为:

$$m_1 \oplus m_2(A_j) = \sum_{X \cap Y = A_j} m(X)m(Y)$$

由于 $R'' = R_1 \cap R_2$, 则 $X \cap Y \in R''$, 即 $X \cap Y = B_i (B_i \subset A_j)$, 因此 $X \cap Y \neq A_j$.

故对于 $A_j \supset B_i$, $m_1 \oplus m_2(A_j) = 0$, 则 $m(A_j) \neq m_1 \oplus m_2(A_j)$.

综上所述, $m(A) \neq \sum_{X \cap Y = A} m(X)m(Y)$.

上面分别证明了论域 U 上任意两个等价关系 R_1 与 R_2 , 它们在细分偏序格上的上、下界的划分对应的质量函数并不对应于证据理论中证据函数的正交组合运算。原因可以解释为: 证据理论中的质量函数合成时的焦点是两个等价关系的交运算, 对于其他的划分而言, 它的质量函数都为 0; 而任意上、下界中划分对应的质量函数并不为 0, 因此两者并没有对应相等。尽管证据理论中质量函数合成时是通过交运算获取焦点, 但在合成过程中使用了乘法与规范化运算, 因此也并不对应于划分的交所对应的质量函数。因此, 在细分偏序格上找不到一个划分, 使它的质量函数对应于证据理论中质量函

数的正交运算。

结束语 粗糙集中通过划分获取质量函数, 划分的粒度不同, 获取的质量函数也不同。质量函数随粒度的单调减小而单调不增, 随粒度的单调增加而单调不减, 因此具有很强的直观性。粗糙集和证据理论在合成质量函数方面既有区别又有联系。本文通过对这粗糙集和证据理论中合成质量函数的方法不同而进行分析, 试图在细分偏序格上找到一种划分获取的质量函数对应于证据理论中的正交组合运算。通过分析, 分别证明了在细分偏序格的上、下确界与所有上、下界所对应的划分中, 获取的质量函数并不对应于证据理论中质量函数正交运算的结论, 从而澄清了不同知识粒度下粗糙集中所有划分获取的质量函数与证据理论中正交运算之间的关系。

参考文献

- [1] 苏运霖, 管纪文, Bell D A. 证据论与约集论[J]. 软件学报, 1999, 10(3): 277-282
- [2] Pawlak Z. Rough Sets and Intelligent Data Analysis [J]. Information Sciences, 2002, 147(1): 1-12
- [3] 苗夺谦, 李道国. 粗糙集理论、算法与应用[M]. 北京: 清华大学出版社, 2008: 150-189
- [4] 王国胤, 姚一豫, 于洪. 粗糙集理论与应用研究综述[J]. 计算机学报, 2009, 7(32): 1229-1246
- [5] 焦占亚, 胡予濮. 划分与传递闭包[J]. 兰州理工大学学报, 2004(30): 130-132
- [6] Liang J Y, Shi Z Z. The information entropy, rough entropy and knowledge granulation in rough set theory [J]. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 2004, 12(1): 37-46
- [7] Liu Ye-zheng, Jiang Yuan-chun, Liu Xiao, et al. A combination strategy for multi-class classification based on multiple association rules[J]. Knowledge-Based Systems, 2008(21): 786-793
- [8] 吴馨, 王加阳. 基于粗糙集的证据获取与合成方法研究[J]. 计算机工程与应用, 2011, 47(12): 34-35
- [9] Yao Y Y. Information granulation and rough set approximation [J]. International Journal of Intelligent Systems, 2001, 16: 87-104
- [10] Yao Y Y. Information granulation and approximation in a decision-theoretic modal of rough sets[M]//Pal S K, Polkowski L, Skowron A, eds. Rough-Neural computing: Techniques for Computing with Words. Berlin: Springer, 2003: 491-518
- [11] 徐久成, 沈钧毅, 安秋生, 等. 基于信息粒度与粗糙集的决策细化研究[J]. 西安交通大学学报, 2005, 39(4): 335-338
- [12] 王国胤, 张清华. 不同知识粒度下粗糙集的不确定性研究[J]. 计算机学报, 2008, 31(9): 1588-1598