



形式背景上近似推理生成决策蕴涵研究

张家录 吴霞

Study on the Approximate Reasoning Models of Decision Implication in Formal Decision Context

ZHANG Jia-Lu, WU Xia

在线阅读 View online: <https://doi.org/10.16383/j.aas.c220705>

您可能感兴趣的其他文章

大数据智能决策

Big Data for Intelligent Decision Making

自动化学报. 2020, 46(5): 878–896 <https://doi.org/10.16383/j.aas.c180861>

基于近似子图的规则空间压缩算法

Rule-space Compression Algorithm Using Approximate Subgraph

自动化学报. 2019, 45(8): 1586–1598 <https://doi.org/10.16383/j.aas.2017.c170272>

污水处理决策优化控制

Advanced Decision and Optimization Control for Wastewater Treatment Plants

自动化学报. 2018, 44(12): 2198–2209 <https://doi.org/10.16383/j.aas.2018.c170257>

面向分布式数据流大数据分类的多变量决策树

A Multivariate Decision Tree for Big Data Classification of Distributed Data Streams

自动化学报. 2018, 44(6): 1115–1127 <https://doi.org/10.16383/j.aas.2017.c160809>

基于决策空间变换最近邻方法的Pareto支配性预测

Prediction of Pareto Dominance Using Nearest Neighbor Method Based on Decision Space Transformation

自动化学报. 2017, 43(2): 294–301 <https://doi.org/10.16383/j.aas.2017.c150877>

基于形式概念分析和语义关联规则的目标图像标注

Object Image Annotation Based on Formal Concept Analysis and Semantic Association Rules

自动化学报. 2020, 46(4): 767–781 <https://doi.org/10.16383/j.aas.c180523>

形式背景上近似推理生成决策蕴涵研究

张家录¹ 吴霞¹

摘要 决策蕴涵分析是形式概念分析研究的重要方面, 基于形式背景获取决策蕴涵、概念规则等知识是数据分析、机器学习的重要研究内容之一. 首先, 利用属性逻辑语义对决策蕴涵的特性进行刻画. 其次, 在经典二值逻辑框架下分析决策蕴涵、概念规则的基于全蕴涵三 I 推理思想及分离规则 (Modus ponens, MP) 和逆分离规则 (Modus tonens, MT) 的近似推理模式的特征, 证明决策蕴涵的 MP、MT 近似推理结论是决策蕴涵, 概念规则的 MP、MT 近似推理结论是概念规则等结论. 引进属性逻辑公式的伪距离, 在属性逻辑伪距离空间中分析推理对象范围参数变化对决策蕴涵 MP、MT 近似推理结论的影响. 最后, 提出若干通过 MP、MT 近似推理生成决策蕴涵、概念规则及拟决策蕴涵的模式和方法, 数值实验验证了所提方法的有效性.

关键词 决策形式背景, 属性逻辑, 决策蕴涵, 概念规则, 拟决策蕴涵, MP 近似推理, MT 近似推理

引用格式 张家录, 吴霞. 形式背景上近似推理生成决策蕴涵研究. 自动化学报, 2024, 50(11): 1-15

DOI 10.16383/j.aas.c220705 **CSTR** 32138.14.j.aas.c220705

Study on the Approximate Reasoning Models of Decision Implication in Formal Decision Context

ZHANG Jia-Lu¹ WU Xia¹

Abstract Decision implication analysis is an important aspect of formal concept analysis, and it is one of the contents of data analysis and machine learning to obtain the knowledge of decision implication and conceptual rule from formal decision context. Firstly, attribute logic semantics is used to analyze the characteristics of decision implication. Secondly, in the classical binary logic framework, the characteristics of modus ponens (MP) and modus tonens (MT) approximate reasoning models based on the all implication three- I inference idea is analyzed. It is proved that the MP and MT approximate reasoning conclusions of decision implications are decisions implications, the MP and MT approximate reasoning conclusions of concept rules are concept rules. The pseudo-metric of attribute logic formula is introduced, and the approximate reasoning of decision implication and its properties are analyzed in attribute logical pseudo-metric space in machine learning environment. Finally, some models and methods of generating decision implication, concept rules and quasi-decision implication through MP and MT approximate reasoning based on existing decision implication, concept rules and quasi-decision implication are proposed, numerical experiments show that the proposed methods are effective.

Key words Formal decision context, attribute logic, decision implication, concept rule, quasi-decision implication, MP approximate reasoning, MT approximate reasoning

Citation Zhang Jia-Lu, Wu Xia. Study on the approximate reasoning models of decision implication in formal decision context. *Acta Automatica Sinica*, 2024, 50(11): 1-15

基于信息系统的知识获取是机器学习、人工智能研究的重要方面, 而规则提取又是知识获取的主

要内容之一^[1]. 形式背景 (G, M, I) 是一个由对象集 G , 属性集 M 及 G 与 M 间的二元关系 I 组成的三元组. 在基于形式背景进行规则提取时, 规则是通过内涵属性集合之间的关系描述且通过对应的外延对象集合之间的包含 (或近似包含) 关系来反映, 其形式是一个如 $A \rightarrow B$ 的表达式, 其中 A 和 B 是属性集合, 其含义是形式背景中具有属性 A 的对象也具有属性 B ^[2-3].

为了从逻辑角度刻画蕴涵规则, 通常在更广泛意义下讨论蕴涵规则 $A \rightarrow B$, A 和 B 分别是条件属性集和决策属性集, 此时的蕴涵规则被称为决策蕴

收稿日期 2022-09-06 录用日期 2023-01-13

Manuscript received September 6, 2022; accepted January 13, 2023

湖南省自然科学基金 (2020JJ4561, 2020JJ4381), 湖南省信息和计算科学专业校企合作创新创业教育基地 (2020-301) 资助

Supported by Natural Science Foundation of Hunan Province (2020JJ4561, 2020JJ4381) and Information and Computing Science University-enterprise Cooperation Innovation and Entrepreneurship Education Base of Hunan Province (2020-301)

本文责任编辑 胡清华

Recommended by Associate Editor HU Qing-Hua

1. 湘南学院数学与信息科学学院 郴州 423000

1. College of Mathematics and Information Science, Xiangnan University, Chenzhou 423000

涵规则, 简称决策蕴涵^[3-4]. Zhai 等^[5-6]在抽象的条件属性和决策属性集上引进决策蕴涵, 从逻辑语义和逻辑语构 2 个方面对决策蕴涵进行了探讨^[6], 他们还在 Pavelka 模糊逻辑意义下提出并讨论了模糊决策蕴涵. 贾楠等^[7]分析了这种一般决策蕴涵的推理规则, 并讨论了推理规则集的完备性和无冗余性. 为适应智能决策的需要, 王琪等^[8]研究了含参模糊决策蕴涵. 张少霞等^[9]讨论了基于逻辑公式的逻辑型决策蕴涵, 使得决策知识的表示更加丰富. 抽象集合上的决策蕴涵讨论固然可以丰富决策蕴涵的理论, 但实际数据分析中在讨论决策蕴涵时更加关注信息系统上、特别是形式背景中的决策蕴涵 $A \rightarrow B$, 此时 A 和 B 分别是决策形式背景的条件属性集和决策属性集. 形式背景上的决策蕴涵还包括概念规则和粒规则 2 个特殊类型. Zhang 等^[10]对决策形式背景中的决策蕴涵、概念规则、粒规则的特征和彼此间的关系进行了研究, 借助决策蕴涵逻辑分析了概念规则和粒规则中的信息损失及其信息丢失背后的原因. 为有效进行多粒度意义下的数据分析, 李金海等^[11-12]讨论了多粒度意义下的决策蕴涵性质, 实现了跨粒度重组数据的决策蕴涵生成.

传统上, 人们习惯于应用二值逻辑及其推理以探究事物之间的确定因果关系和获取确定性新知识^[13]. 但是, 正如亚马逊的推荐系统^[14]和谷歌的流感预测^[15]等现实案例所展示, 对于大数据环境下的数据分析和应用, 人们更加关注事物之间的相关关系分析, 寻找数据之间的近似因果关系一样受到人们的关注和重视^[16]. 也正是因为这个原因, 人们不断尝试创新各种方法以探索事物之间的近似因果关系, 并依托这种近似因果关系从现有知识获取新知识^[17]. 推理是寻找事物之间因果关系, 从已有知识获取新知识的重要方法, 是数理逻辑的主要研究内容. 人们一般认为模糊逻辑及其推理应用是探究事物之间近似因果关系及获取不确定性知识的手段和方法^[18-19]. 然而, 正如文献 [18] 所指出, 借助于全蕴涵三 I 推理思想^[19], 经典二值逻辑系统中存在着与模糊推理相同的求解机制, 从而基于经典二值逻辑也可以进行近似推理的探讨.

决策蕴涵、概念规则都是信息系统上知识的表现形式, 决策蕴涵、概念规则推理是知识推理的一个重要方面. 因此, 将经典二值逻辑框架下的近似推理模式应用于决策蕴涵、概念规则, 讨论决策蕴涵、概念规则的推理模式及性质, 依据形式背景上的现有决策蕴涵、概念规则, 通过近似推理方法寻找新的决策蕴涵、概念规则是有意义的. 从知识发现的角度看, 基于现有数据集上的决策蕴涵、概念

规则等知识, 采用近似推理的方法获取更大数据集上的决策蕴涵、概念规则等新的知识是非常有价值的.

本文应用经典二值逻辑中的分离规则 (Modus ponens, MP)、逆分离规则 (Modus tonens, MT) 近似推理来探讨形式背景上的决策蕴涵和概念规则的近似推理的特征, 并提出实现方法. 首先, 讨论决策蕴涵的属性逻辑语义的性质. 其次, 在经典二值逻辑框架下分析基于全蕴涵三 I 推理思想的决策蕴涵和概念规则的 MP、MT 近似推理模式的特征, 证明决策蕴涵的 MP、MT 近似推理结论是决策蕴涵, 概念规则的 MP、MT 近似推理的结论是概念规则. 引进属性逻辑公式间的伪距离, 并在逻辑伪距离空间中分析对象参数变化对决策蕴涵的 MP、MT 近似推理结论的影响. 最后, 提出 2 个利用 MP、MT 近似推理来生成拟决策蕴涵的算法, 对 2 个算法的复杂性进行分析, 并通过数值实验对本文提出方法的有效性进行验证.

1 形式概念和属性逻辑概述

本节概述下文要用到的形式概念分析和属性逻辑的知识.

定义 1^[3]. 一个形式背景是一个三元组 $K = (G, M, I)$, 其中 G 为对象集合, M 为属性集合, I 是 G 到 M 的二元关系, gIm 表示对象 g 具有属性 m .

通常称不包含空行、空列、满行和满列的形式背景为正则形式背景^[20]. 本文只讨论正则形式背景.

定义 2^[3]. 在形式背景 $K = (G, M, I)$ 中, 对 $A \subseteq G$, 记 $\alpha^I(A) = \{m \in M | \forall g \in A, gIm\}$ (A 中对象共同具有属性的集合); 对 $B \subseteq M$, 记 $\beta^I(B) = \{g \in G | \forall m \in B, gIm\}$ (具有 B 中所有属性的对象集合); α^I, β^I 分别称为 K 的属性映射和对象映射.

定义 3^[3]. 在形式背景 $K = (G, M, I)$ 中, 对 $A \subseteq G, B \subseteq M$, 如果 $\alpha^I(A) = B, \beta^I(B) = A$, 则称 $C = (A, B)$ 是一个概念, 并称 A 是概念 C 的外延, B 是概念 C 的内涵. 形式背景 K 的全体概念记为 $C(K)$.

命题 1^[3]. 对 $A, A_1, A_2 \subseteq G, B, B_1, B_2 \subseteq M$.

1) $\alpha^I(A_1) \subseteq \alpha^I(A_2) \Rightarrow \alpha^I(A_2) \subseteq \alpha^I(A_1), \alpha^I(B_1) \subseteq \alpha^I(B_2) \Rightarrow \beta^I(B_2) \subseteq \beta^I(B_1)$;

2) $A \subseteq \beta^I(\alpha^I(A)), B \subseteq \alpha^I(\beta^I(B))$;

3) $A \subseteq \beta^I(B)$ 当且仅当 $B \subseteq \alpha^I(A)$;

4) $\alpha^I(A_1 \cup A_2) = \alpha^I(A_1) \cap \alpha^I(A_2), \beta^I(A_1 \cup A_2) = \beta^I(A_1) \cap \beta^I(A_2)$;

5) $\alpha^I(A_1 \cap A_2) \supseteq \alpha^I(A_1) \cup \alpha^I(A_2), \beta^I(A_1 \cap A_2) \supseteq \beta^I(A_1) \cup \beta^I(A_2)$;

6) $(\beta^I(\alpha^I(A)), \alpha^I(A)), (\beta^I(B), \alpha^I(\beta^I(B)))$ 都是概念.

定义 4^[21]. 设 $K = (G, M, I)$ 是一个形式背景, \neg, \wedge 和 \rightarrow 是逻辑联结词. 称 $a \in M$ 为一个原子属性公式, 全部原子属性公式生成的 $(\neg, \wedge, \rightarrow)$ 自由代数记为 $\Gamma(K)$, 称 $\Gamma(K)$ 中的元素为基于形式背景 K 的属性逻辑公式 (简称属性公式).

$\Gamma(K)$ 中的元素可以通过以下方式^[21] 得到:

- 1) $M \subseteq \Gamma(K)$;
- 2) 若 $\varphi \in \Gamma(K)$, 则 $\neg \varphi \in \Gamma(K)$;
- 3) 若 $\varphi, \psi \in \Gamma(K)$, 则 $\varphi \wedge \psi, \varphi \rightarrow \psi \in \Gamma(K)$;
- 4) $\gamma \in \Gamma(K)$ 当且仅当 γ 可经过以上 3 步有限次所产生.

本文讨论的决策蕴涵的前件和后件都是属性合取公式, 对于 $B \subseteq M$, 由 B 中的属性合取公式为 $\varphi = \bigwedge_{b \in B} b$. 在不会引起混淆时, 时常不区分属性集 B 及 B 中属性组成的合取公式. 同时, 为了表述简便, 也经常使用逻辑联结词 $\varphi \vee \psi = \neg(\neg\varphi \wedge \neg\psi)$.

对逻辑公式进行赋值是研究二值逻辑、模糊逻辑等逻辑系统语义的基本手段^[18], 但现有文献对决策逻辑^[1] 及决策蕴涵的逻辑研究较少有从赋值出发来进行^[6]. 作者在基于软集的决策分析^[22] 和基于形式背景的属性逻辑^[21, 23] 的研究中, 已经尝试从属性公式赋值的角度来对相关内容进行讨论.

设 $K = (G, M, I)$ 是形式背景. 对每个 $u \in G$, u 自然地决定了一个映射 $v_u^{(0)} : M \rightarrow \{0, 1\}$, $v_u^{(0)}(a) = I(u, a)$. 由于 $\Gamma(K)$ 是由原子属性公式集 M 生成的 $(\neg, \wedge, \rightarrow)$ 型自由代数, 故 $v_u^{(0)}$ 可唯一延拓成 $\Gamma(K)$ 上的一个映射 v_u , v_u 在 M 上的限制正是 $v_u^{(0)}$, 称映射 v_u 为属性公式集的一个 K -赋值, 全体 K -赋值之集记为 $\Omega_K = \{v_u : u \in G\}$. 显然, G 到 Ω_K 之间有一一对应 $\pi : G \rightarrow \Omega_K$, $\pi(u) = v_u$. 对每个 $a \in M$, a 决定了一个定义在 G 上的函数 $\bar{a} : G \rightarrow \{0, 1\}$, $\bar{a}(u) = I(a, u)$. 同样, 一个由 t 个原子属性公式 a_{i_1}, \dots, a_{i_t} 组成的属性公式 $\varphi = \varphi(a_{i_1}, \dots, a_{i_t})$ 也决定了 G 上的函数 $\bar{\varphi}(u) = \bar{\varphi}(\bar{a}_{i_1}(u), \dots, \bar{a}_{i_t}(u))$, $\bar{\varphi}$ 作用于 $\bar{a}_{i_1}, \dots, \bar{a}_{i_t}$ 的方式如同 φ 作用于 a_{i_1}, \dots, a_{i_t} 的方式一样. 例如, $\varphi = \neg a_1 \vee a_2 \rightarrow a_3$, 则 $\bar{\varphi}(u) = (1 - \bar{a}_1(u)) \vee \bar{a}_2(u) \rightarrow \bar{a}_3(u)$.

另一方面, 按照经典二值逻辑系统的赋值方法, 由映射 $v_u^M : M \rightarrow \{0, 1\}$ 所决定的公式集 $\Gamma(K)$ 上的映射 $v : \Gamma(K) \rightarrow \{0, 1\}$ 为 $\Gamma(K)$ 的赋值, 对 $\varphi \in \Gamma(K)$, $v(\varphi)$ 称为逻辑公式 φ 的值. $\Gamma(K)$ 的全部赋值之集记为 Ω . 一般地, Ω_K 为 Ω 的子集.

例 1. 一个形式背景 $K = (G, M, I)$ 见表 1, 其中对象集 $G = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$, 属性集 $M = \{a_1, a_2, a_3, a_4, a_5, a_6\}$.

根据 K -赋值的产生方法, Ω_K 有 7 个赋值. 而

表 1 形式背景 $K = (G, M, I)$
Table 1 A formal context $K = (G, M, I)$

G	a_1	a_2	a_3	a_4	a_5
u_1	1	1	0	0	1
u_2	0	0	1	1	0
u_3	1	0	1	0	0
u_4	0	1	0	1	0
u_5	0	1	0	1	1
u_6	1	0	1	0	1

按照二值逻辑赋值方法, Ω 有 $2^5 = 32$ 个赋值. 一般地, Ω_K 为 Ω 的真子集, 比如赋值 $v_0 : v_0(a_1) = 0, v_0(a_2) = 0, v_0(a_3) = 0, v_0(a_4) = 0, v_0(a_5) = 0$; $v_1 : v_1(a_1) = 1, v_1(a_2) = 1, v_1(a_3) = 1, v_1(a_4) = 0, v_1(a_5) = 0$; v_0, v_1 在 Ω 中, 而不在 Ω_K 中.

由于全体 K -赋值集 Ω_K 为全体赋值集 Ω 的真子集, 因此, 本文的 K -赋值也可作为文献^[24] 所讨论的数理逻辑部分赋值理论的具体实例.

在形式背景中, $I(u, a) = 1$ 表明一个对象 u 具有属性 a , 此时有 $\bar{a}(u) = 1$. 一般地, 如果 $\bar{\varphi}(u) = 1$, 则表明对象 u 具有属性公式 φ 中的全部属性, 此时也称对象 u 满足公式 φ . 对于 $\varphi \in \Gamma(K)$, 称满足 φ 的对象之集 $\{u \in G | \bar{\varphi}(u) = 1\}$ 为 φ 的 K -语义, 本文记 φ 的 K -语义为 $\|\varphi\|_K$, 在不会引起混乱时, 也可将下标 K 省略. 属性公式的 K -语义有以下命题 2.

命题 2^[21]. 设 $K = (G, M, I)$ 是一个形式背景, 则对任何 $\varphi, \psi \in \Gamma(K)$, 下式成立:

- 1) $\|\neg\varphi\| = G - \|\varphi\|$;
- 2) $\|\varphi \vee \psi\| = \|\varphi\| \cup \|\psi\|$;
- 3) $\|\varphi \wedge \psi\| = \|\varphi\| \cap \|\psi\|$;
- 4) $\|\varphi \rightarrow \psi\| = (G - \|\varphi\|) \cup \|\psi\|$.

命题 3^[21]. 设 $\Sigma = \{v_{u_1}, \dots, v_{u_l}\} \subset \Omega_K$, $\{r_1, \dots, r_l\}$ 是一个 $0 \sim 1$ 序列, 则存在属性公式 $\varphi \in \Gamma(K)$ 满足 $v_{u_i}(\varphi) = r_i, i = 1, \dots, l$.

在经典二值逻辑的确定性推理中, 有 2 个基本的推理规则: MP 和 MT, 这 2 个规则同样适应属性逻辑的推理. 但在人工智能的研究中同样也关注近似推理, 即对知识的近似获取同样感兴趣. 因此, 人们也经常需要如下适应于近似推理的广义 MP^[17-18]:

$$\left\{ \begin{array}{l} \text{已知} \quad \varphi \rightarrow \psi \\ \text{且给定} \quad \varphi^* \\ \hline \text{求} \quad \psi^* \end{array} \right. \quad (1)$$

式中, $\varphi, \psi, \varphi^*, \psi^* \in \Gamma(K)$.

在确定推理中, 上述广义 MP 式 (1) 是无解的. 在实际应用中, 虽然不能找到其精确解, 但可以寻求上述广义 MP 式 (1) 的近似解. 根据全蕴涵三 I 推理

思想, 所求的近似解 ψ^* 既要得到小前提 φ^* 的最大支持, 也还要得到大前提 $\varphi \rightarrow \psi$ 的最大支持, 即 ψ^* 应该使得对于所讨论范围内的对象, 蕴涵公式 $(\varphi \rightarrow \psi) \rightarrow (\varphi^* \rightarrow \psi^*)$ 能取得最大值 1. 因此, 给出如下广义 MP 解的定义.

定义 5^[17]. 设 $\varphi, \psi, \varphi^* \in \Gamma(K)$, Σ, Δ 为 G 的 2 个子集. 基于形式背景的近似推理的广义 MP 式 (1) 的 (Σ, Δ) 解是满足以下条件的属性公式 ψ^* :

- 1) $\forall x \in \Delta, \forall y \in \Sigma,$
 $(\overline{\varphi}(y) \rightarrow \overline{\psi}(x)) \rightarrow (\overline{\varphi^*}(y) \rightarrow \overline{\psi^*}(x)) = 1.$
- 2) 如果有 $\gamma \in \Gamma(K)$ 满足 $\forall x \in \Delta, \forall y \in \Sigma,$
 $(\overline{\varphi}(y) \rightarrow \overline{\psi}(x)) \rightarrow (\overline{\varphi^*}(y) \rightarrow \overline{\gamma}(x)) = 1,$ 则 $\forall x \in \Delta,$
 $\overline{\psi^*}(x) \leq \overline{\gamma}(x).$

定理 1^[17]. 设 $\varphi, \psi, \varphi^* \in \Gamma(K)$, Σ 和 Δ 是 G 的 2 个子集. 基于形式背景的广义 MP 式 (1) 的解 ψ^* 满足:

$$\overline{\psi^*}(x) = [\overline{\varphi^*}(y) \wedge (\overline{\varphi}(y) \rightarrow \overline{\psi}(x))], x \in \Delta \quad (2)$$

式中, \wedge 表示函数取小运算.

注 1. 由于 $\overline{\varphi}(y) \rightarrow \overline{\psi}(x) = (1 - \overline{\varphi}(y)) \vee \overline{\psi}(x)$, 因此式 (2) 可改写为:

$$\overline{\psi^*}(x) = \bigvee_{y \in \Sigma} \vee [\overline{\psi}(x) \wedge \bigvee_{y \in \Sigma} \overline{\varphi^*}(y)]$$

式中, \vee 表示函数取大运算.

从而有以下 4 种特殊情形:

- 1) 如果有 $y_0 \in \Sigma$, 使得 $\overline{\varphi^*}(y_0) \wedge (1 - \overline{\varphi}(y_0)) = 1$, 则 $\forall x \in \Delta, \overline{\psi^*}(x) = 1.$
- 2) 如果对任何 $y \in \Sigma$, 有 $\overline{\varphi^*}(y) = 0$, 则 $\forall x \in \Delta,$
 $\overline{\psi^*}(x) = 0.$
- 3) 如果有 $y_0 \in \Sigma$, 使得 $\overline{\varphi^*}(y_0) = 1$, 且 $\|\varphi^*\| \subseteq \|\varphi\|$, 则 $\forall x \in \Delta, \overline{\psi^*}(x) = \overline{\psi}(x).$
- 4) 如果 $\varphi = \varphi^*$, 则由式 (2) 有 $\psi = \psi^*$, 从而广义 MP 的确是经典 MP 的推广.

对于如下多规则 MP 近似推理问题^[17]:

$$\left\{ \begin{array}{l} \text{已知} \quad \varphi_1 \rightarrow \psi_1 \\ \quad \quad \quad \dots \\ \quad \quad \quad \varphi_n \rightarrow \psi_n \\ \hline \text{且给定} \quad \varphi^* \\ \text{求} \quad \quad \quad \psi^* \end{array} \right. \quad (3)$$

如果采用先推理后聚合的推理过程, 聚合算子取合取算子, 则其推理结论的属性公式函数为^[17]:

$$\overline{\psi^*}(x) = \bigwedge_{i=1}^n \bigvee_{y \in \Sigma} [\overline{\varphi^*}(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in \Delta \quad (4)$$

类似于二值逻辑推理的 MT, 考虑如下基于形式背景的近似推理的广义 MT^[17]:

$$\left\{ \begin{array}{l} \text{已知} \quad \varphi \rightarrow \psi \\ \hline \text{且给定} \quad \psi^* \\ \text{求} \quad \quad \varphi^* \end{array} \right. \quad (5)$$

式中, $\varphi, \psi, \varphi^*, \psi^* \in \Gamma(K).$

定义 6^[17]. 设 $\varphi, \psi, \psi^* \in \Gamma(K)$, Σ, Δ 是 G 的 2 个子集. 基于形式背景的近似推理的广义 MT 式 (3) 的 (Σ, Δ) 解是满足以下条件的属性公式 φ^* :

- 1) $\forall x \in \Delta, \forall y \in \Sigma,$
 $(\overline{\varphi}(y) \rightarrow \overline{\psi}(x)) \rightarrow (\overline{\varphi^*}(y) \rightarrow \overline{\psi^*}(x)) = 1.$
- 2) 如果有 $\chi \in \Gamma(K)$ 满足 $\forall x \in \Delta, \forall y \in \Sigma,$
 $(\overline{\varphi}(y) \rightarrow \overline{\psi}(x)) \rightarrow (\overline{\varphi^*}(y) \rightarrow \overline{\chi}(x)) = 1$ 则 $\forall y \in \Sigma,$
 $\overline{\psi^*}(y) \geq \overline{\chi}(y).$

定理 2^[17]. 设 $\varphi, \psi, \psi^* \in \Gamma(K)$, Σ, Δ 是 G 的 2 个子集. 基于形式背景的广义 MT 式 (3) 的解 φ^* 满足:

$$\overline{\varphi^*}(y) = \bigwedge_{x \in \Delta} [(\overline{\varphi}(y) \rightarrow \overline{\psi}(x)) \rightarrow \overline{\psi^*}(x)], y \in \Sigma \quad (6)$$

注 2. 由于 $\overline{\varphi}(y) \rightarrow \overline{\psi}(x) = (1 - \overline{\varphi}(y)) \vee \overline{\psi}(x)$, 因此式 (6) 可改写为:

$$\overline{\varphi^*}(y) = \bigwedge_{x \in \Delta} [\overline{\psi^*}(x) \vee (1 - \overline{\psi}(x))] \wedge [\overline{\varphi}(x) \vee \bigwedge_{x \in \Delta} \overline{\psi^*}(x)]$$

从而有以下 3 种特殊的情形:

- 1) 如果有 $x_0 \in \Delta$, 使得 $\overline{\psi^*}(x_0) \vee (1 - \overline{\psi}(x_0)) = 0$, 则 $\forall y \in \Sigma, \overline{\varphi^*}(x) = 0.$
- 2) 如果对任何 $x \in \Delta$, 有 $\overline{\psi^*}(y) = 1$, 则 $\forall y \in \Sigma,$
 $\overline{\varphi^*}(x) = 1.$
- 3) 如果有 $x_0 \in \Delta$, 使得 $\overline{\psi^*}(y) = 0$, 且 $\|\psi\| \subseteq \|\psi^*\|$, 则 $\forall y \in \Sigma, \overline{\varphi^*}(y) = \overline{\varphi}(y).$

对于如下多规则 MT 近似推理问题^[17]:

$$\left\{ \begin{array}{l} \text{已知} \quad \varphi_1 \rightarrow \psi_1 \\ \quad \quad \quad \dots \\ \quad \quad \quad \varphi_n \rightarrow \psi_n \\ \hline \text{且给定} \quad \psi^* \\ \text{求} \quad \quad \quad \psi^* \end{array} \right. \quad (7)$$

如果采用先推理后聚合的推理过程, 聚合算子取合取算子, 则其推理结论的属性公式函数为^[17]:

$$\overline{\psi^*}(y) = \bigwedge_{i=1}^n \bigwedge_{x \in \Delta} [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x)) \rightarrow \overline{\psi^*}(x)], y \in \Sigma \quad (8)$$

2 决策蕴涵的属性逻辑语义分析

为了方便展开决策蕴涵的逻辑推理, 本节利用属性逻辑的 K -赋值, 对决策蕴涵进行属性逻辑的语义分析.

定义 7^[5]. 称形式背景 $K_1 = (G, C, I)$ 和 $K_2 =$

(G, D, J) 的并置 $K = (G, C, D, I, J)$ 为决策形式背景, 并分别称 C 和 D 为条件属性集和决策属性集. 若 $A \subseteq C, B \subseteq D$, 且有 $\|\wedge_{a \in A} a\| \subseteq \|\wedge_{b \in B} b\|$, 称属性公式 $\wedge_{a \in A} a \rightarrow \wedge_{b \in B} b$ 为决策蕴涵, 称 $\wedge_{a \in A} a$ 为决策蕴涵前件 (或前提), $\wedge_{b \in B} b$ 为决策蕴涵后件 (或结论). 如果 A, B 还分别是条件概念和决策概念的内涵属性集, 称决策蕴涵 $\wedge_{a \in A} a \rightarrow \wedge_{b \in B} b$ 为概念规则. 记决策形式背景 K 的全体决策蕴涵之集为 K_D , 全体概念规则之集为 K_C .

为了不引起混淆, 也经常简记 $\|\wedge_{a \in A} a\|$ 为 $\|A\|$, $\wedge_{a \in A} a \rightarrow \wedge_{b \in B} b$ 为 $A \rightarrow B$.

从形式上看, 文献 [5-6] 讨论的抽象集合上的决策蕴涵 $A \rightarrow B$ 是一个由逻辑联结词 \rightarrow 将 2 个属性合取公式联结起来的属性公式. 但是, 由于数据分析讨论决策蕴涵时常常是在信息系统或决策形式背景上进行的, 而且还要将概念规则和粒规则纳入决策蕴涵的范畴来统一讨论, 因此基于决策形式背景上的决策蕴涵 $A \rightarrow B$ 比一般抽象属性集上的决策蕴涵多了一个限制条件 $\|\wedge_{a \in A} a\| \subseteq \|\wedge_{b \in B} b\|$ [4, 10-11].

定义 8^[5]. 设 (G, C, D, I, J) 是决策形式背景.

1) 对于属性集 $T \subseteq C \cup D$ 和决策蕴涵 $A \rightarrow B$, 若 $A \not\subseteq T \cap C$ 或 $B \subseteq T \cap D$, 称 T 满足 $A \rightarrow B$ (或称 T 是 $A \rightarrow B$ 的模型), 记为 $T \models A \rightarrow B$. 设 $L \subseteq K_D$, 若 T 满足 L 中的任意一个决策蕴涵, 称 T 是 L 的一个模型, 记为 $T \models L$.

2) 若对任意 $T \subseteq C \cup D$, $T \models L$ 蕴涵 $T \models A \rightarrow B$, 称 $A \rightarrow B$ 为 L 的语义结论, 记为 $L \models A \rightarrow B$.

从逻辑观点看, 一个决策蕴涵是通过蕴涵联结词将条件属性合取公式和决策属性合取公式连接起来的属性逻辑公式, 定义 8 和语义结论等术语是命题逻辑中的相应术语在决策蕴涵中的本质特征刻画. 因此, 在引进属性逻辑语言后, 可以借用命题逻辑中的赋值方法对决策蕴涵、概念规则等进行讨论.

利用前面引入的 K -赋值, 以下给出决策蕴涵的 K -语义结论定义.

定义 9. 设 $K = (G, C, D, I, J)$ 是决策形式背景, $A \rightarrow B \in K_D$, $L \subseteq K_D$, v 是 K -赋值. 记 $v(L) = \wedge_{A_1 \rightarrow B_1 \in L} v(A_1 \rightarrow B_1)$, 如果对任何 K -赋值 $v \in \Omega_K$, 都有 $v(L) = 1$ 蕴涵 $v(A \rightarrow B) = 1$, 称 $A \rightarrow B$ 为 L 的 K -语义结论, 记为 $L \models_K A \rightarrow B$.

从逻辑语义角度看, 决策蕴涵集的语义结论是从全部模型 L 的角度来讨论, 而 K -语义结论只从形式背景对象集所决定的 K -赋值来考虑. 之间的关系正是计量逻辑学中基于全部赋值和部分赋值推理的关系在决策蕴涵的具体体现 [13, 24].

命题 4. 设 $K = (G, C, D, I, J)$ 是决策形式背

景, $A \rightarrow B \in K_D$.

1) $T \subseteq C \cup D$ 是 $A \rightarrow B$ 的模型当且仅当对赋值 $v_T \in \Omega$, 有 $v_T(A \rightarrow B) = 1$, v_T 是由 T 决定的赋值.

2) 如果 $v \in \Omega_K$, 使得 $v(A \rightarrow B) = 1$, 则 $T_0 = \{m | m \in C \cup D, v(m) = 1\}$ 是 $A \rightarrow B$ 的模型.

证明. 1) 设 $T \subseteq C \cup D$ 是 $A \rightarrow B$ 的模型, v_T 是由 T 所决定的赋值, 即 T 中的原子属性公式赋值为 1, $C \cup D - T$ 中的原子属性公式赋值为 0. 则当 $A \not\subseteq T \cap C$ 时, 有 $a \in A$ 使 $v_T(a) = 0$, 于是 $v_T(A \rightarrow B) = v_T(A) \rightarrow v_T(B) = 1$. 而当 $B \subseteq T \cap D$ 时, 对任何 $b \in B$ 有 $v_T(b) = 1$, 此时同样有 $v_T(A \rightarrow B) = v_T(A) \rightarrow v_T(B) = 1$. 因此总有 $v_T(A \rightarrow B) = 1$.

反之, 假设 T 不是 $A \rightarrow B$ 的模型. 若令 $T = \{m | m \in C \cup D, v(m) = 1\}$, 则可断定 T 是 $A \rightarrow B$ 的模型, 即 $A \not\subseteq T \cap C$ 或 $B \subseteq T \cap D$. 否则, 如果 $A \subseteq T \cap C$ 且 $B \not\subseteq T \cap D$, 则有 $v_T(A) = 1$ 且 $v_T(B) = 0$, 从而 $v_T(A \rightarrow B) = 0$. 与已知矛盾.

2) 如果 T_0 不是 $A \rightarrow B$ 的模型, 则 $A \subseteq T_0 \cap C$ 且 $B \not\subseteq T_0 \cap D$. 从而 $\forall a \in A, a \in T_0 \cap C$, 且 $\exists b \in B, b \notin T_0 \cap D$, 于是 $v(A) = 1, v(B) = 0$. 因此 $v(A \rightarrow B) = 0$, 与已知 $v(A \rightarrow B) = 1$ 矛盾. 所以 T_0 是 $A \rightarrow B$ 的模型. \square

由命题 4 的 1) 可知, T 是 $L \subseteq K_D$ 的模型当且仅当对任何的 $A_1 \rightarrow B_1 \in K_D$, 有 $v_T(A_1 \rightarrow B_1) = 1$. $L \models A \rightarrow B$ 当且仅当如果对任何的 $A_1 \rightarrow B_1$ 和 $v \in \Omega$, 有 $v(A_1 \rightarrow B_1) = 1$, 则一定也有 $v(A \rightarrow B) = 1$. 由于 K -赋值 Ω_K 只是全部模型 $T \subseteq C \cup D$ 所决定赋值集 Ω 的真子集, 因此如果 $L \models A \rightarrow B$, 则 $L \models_K A \rightarrow B$. 反之, 即使对任何 K -赋值 $v \in \Omega_K$, 当对任何 $A_1 \rightarrow B_1 \in L$, $v(A_1 \rightarrow B_1) = 1$ 时, 都有 $v(A \rightarrow B) = 1$, 也不一定断定 $L \models A \rightarrow B$. 因此, 命题 4 的 2) 只能用来判定 $A \rightarrow B$ 不是 L 的语义结论. 即如果有 K -赋值 $v_0 \in \Omega_K$, 使得 $\forall A_1 \rightarrow B_1 \in L, v_0(A_1 \rightarrow B_1) = 1$, 但 $v_0(A \rightarrow B) = 0$, 这时就可以断言 $A \rightarrow B$ 不是 L 的语义结论. 本文例 2 也说明了这一点.

例 2. 一个决策形式背景 $K = (G, C, D, I, J)$ 见表 2, 对象集 $G = \{u_1, u_2, u_3, u_4\}$, 条件属性集 $C = \{a_1, a_2, a_3, a_4\}$, 决策属性集 $D = \{d_1, d_2, d_3\}$.

$L = \{a_1 \wedge a_2 \wedge a_3 \rightarrow d_1 \wedge d_3\}$, 显然对任何 K -赋值 $v \in \Omega_K$, 当 $v(a_1 \wedge a_2 \wedge a_3 \rightarrow d_1 \wedge d_3) = 1$ 时, 有 $v(a_1 \wedge a_2 \rightarrow d_1 \wedge d_3) = 1$, 即 $L \models_K a_1 \wedge a_2 \rightarrow d_1 \wedge d_3$. 若取 $T = \{a_1, a_2, a_4, d_1, d_2\}$, 显然有 $v_T(a_1 \wedge a_2 \wedge a_3 \rightarrow d_1 \wedge d_3) = 1$ 的模型, 但 T 却不是 $a_1 \wedge a_2 \rightarrow d_1 \wedge d_3$ 的模型, 于是 $a_1 \wedge a_2 \rightarrow d_1 \wedge d_3$ 不是 L 的语义结论.

命题 5. 设 $K = (G, C, D, I, J)$ 是决策形式背景. 对 $A \subseteq C, B \subseteq D$, 有:

表 2 决策形式背景 $K = (G, C, D, I, J)$ Table 2 A formal decision context $K = (G, C, D, I, J)$

G	a_1	a_2	a_3	a_4	d_1	d_2	d_3
u_1	1	1	1	1	1	0	1
u_2	0	0	1	0	0	1	0
u_3	1	0	0	0	0	1	1
u_4	0	1	1	0	1	0	0

1) $A \rightarrow \alpha^J(\beta^I(A))$ 是决策蕴涵;

2) $A \rightarrow B$ 是决策蕴涵当且仅当 $B \subseteq \alpha^J(\beta^I(A))$.

证明. 1) 显然有 $\alpha^J(\beta^I(A)) \subseteq D$, 且 $\|A\| = \beta^I(A) \subseteq \beta^I(\alpha^J(\beta^I(A))) = \|\alpha^J(\beta^I(A))\|$. 所以 $A \rightarrow \alpha^J(\beta^I(A))$ 是决策蕴涵.

2) 若 $A \rightarrow B$ 是决策蕴涵, 则 $\|A\| = \beta^I(A) \subseteq \|B\| = \beta^J(B)$. 因此 $B \subseteq \alpha^J(\beta^J(B)) \subseteq \alpha^J(\beta^I(A))$.

反之, 如果 $B \subseteq \alpha^J(\beta^I(A))$, 则 $\|A\| = \beta^I(A) \subseteq \beta^J(\alpha^J(\beta^I(A))) \subseteq \beta^J(B) = \|B\|$. 所以 $A \rightarrow B$ 是决策蕴涵. \square

命题 6. 设 $K = (G, C, D, I, J)$ 是决策形式背景, $A \rightarrow B \in K_D$, $L \subseteq K_D$. 则 $A \rightarrow B$ 不是 L 的 K -语义结论当且仅当存在 $d \in B$ 使得 $A \rightarrow d$ 不是 L 的 K -语义结论.

证明. 如果 $A \rightarrow B$ 不是 L 的 K -语义结论, 则存在 K -赋值 $v_0 \in \Omega_K$, 虽然 $\forall A_1 \rightarrow B_1 \in L$, $v_0(A_1 \rightarrow B_1) = 1$, 但 $v_0(A \rightarrow B) = 0$, 即 $v_0(A) > v_0(B)$. 于是存在 $d \in B$, 使得 $v_0(A) = 1$, 而 $v_0(d) = 0$. 所以 $A \rightarrow d$ 不是 L 的 K -语义结论.

反之, 若存在 $d \in B$ 使得 $A \rightarrow d$ 不是 L 的 K -语义结论, 则有 K -赋值 $v_0 \in \Omega_K$, 使得 $\forall A_1 \rightarrow B_1 \in L$, $v_0(A_1 \rightarrow B_1) = 1$, 但 $v_0(A \rightarrow d) = 0$ 意味着 $v_0(A) = 1$, $v_0(d) = 0$. 于是 $v_0(A) = 1 > 0 = v_0(d) \geq v_0(B)$, 即 $v_0(A \rightarrow B) = 0$. 所以 $A \rightarrow B$ 不是 L 的 K -语义结论. \square

3 决策蕴涵和概念规则的近似推理模式

第 1 节给出的形式背景上的广义 MP、广义 MT 的前提和结论所涉及的公式是一般的属性公式, 且通常是在给定的对象范围内 (Σ, Δ) 探讨推理规则的解. 由于本文是在决策形式背景上来讨论决策蕴涵推理, 而决策蕴涵中的前件和后件分别是条件原子属性公式和决策原子属性合取公式, 且通常是通过全部的对象集合之间的包含关系来反映. 因此, 本节在分析决策蕴涵的 MP 近似推理、MT 近似推理的特征时, 推理的对象范围参数 (Σ, Δ) 是全部对象集 G .

3.1 决策蕴涵和概念规则的 MP 近似推理模式

本节讨论决策形式背景上基于广义 MP 的决

策蕴涵的近似推理 (以下简称决策蕴涵的 MP 近似推理) 的特征.

定理 3. 设 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 是决策形式背景 $K = (G, C, D, I, J)$ 的决策蕴涵集, φ^* 是条件属性合取公式, f 是由 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 的 MP 近似推理得到的解函数:

$$f(x) = \bigwedge_{i=1}^n \bigvee_{y \in G} [\overline{\varphi^*}(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in G$$

如果记 $W = \{x \in G | f(x) = 1\}$, $\alpha^J(W) = \{d \in D | \forall x \in W, xJd\}$. 当 $\alpha^J(W)$ 非空时, 令 $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$, 则 $\varphi^* \rightarrow \psi^*$ 是决策蕴涵.

证明. 如果 $\|\varphi^*\|$ 是空集, $\varphi^* \rightarrow \psi^*$ 当然是决策蕴涵. 因此, 不妨设 $\|\varphi^*\|$ 非空, 并设 $x_0 \in \|\varphi^*\|$, 有 $\overline{\varphi^*}(x_0) = 1$. 设:

$$f_i(x) = \bigvee_{y \in G} [\overline{\varphi^*}(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], i = 1, \dots, n$$

根据注 1, 在 $\overline{\varphi^*}(x_0) = 1$ 的条件下, 只有当 $\|\varphi^*\| \subseteq \|\varphi_i\|$ 时, 才有 $f_i(x) = \overline{\psi}_i(x)$; 否则, $f_i(x)$ 恒等于 1. 不妨设在决策蕴涵集 L 的前件属性公式集 $\{\varphi_1, \dots, \varphi_n\}$ 中有 i_k 个属性公式 $\varphi_{i_1}, \dots, \varphi_{i_k}$ 满足条件 $\|\varphi^*\| \subseteq \|\varphi_{i_1}\|, \dots, \|\varphi^*\| \subseteq \|\varphi_{i_k}\|$. 从而:

$$f(x) = \bigwedge_{i=1}^n f_i(x) = \overline{\psi}_{i_1}(x) \wedge \dots \wedge \overline{\psi}_{i_k}(x), x \in G$$

如果记 $W = \{x \in G | f(x) = 1\}$, $\alpha^J(W) = \{d \in D | \forall x \in W, xJd\}$, 当 $\alpha^J(W)$ 非空时, 令 $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$, 则对 $x \in G$:

$$\overline{\psi^*}(x) = \bigwedge_{d \in \alpha^J(W)} \overline{d}(x) \geq f(x) = \overline{\psi}_{i_1}(x) \wedge \dots \wedge \overline{\psi}_{i_k}(x)$$

上式说明, 当 $\overline{\varphi^*}(x_0) = 1$ 时, 有 $\overline{\varphi}_{i_1}(x_0) = 1, \dots, \overline{\varphi}_{i_k}(x_0) = 1$. 又由于 $\varphi_{i_1} \rightarrow \psi_{i_1}, \dots, \varphi_{i_k} \rightarrow \psi_{i_k}$ 是决策蕴涵, 故有 $\|\varphi_{i_1}\| \subseteq \|\psi_{i_1}\|, \dots, \|\varphi_{i_k}\| \subseteq \|\psi_{i_k}\|$. 从而 $\overline{\psi}_{i_1}(x_0) = 1, \dots, \overline{\psi}_{i_k}(x_0) = 1$. 因此:

$$\overline{\psi^*}(x_0) \geq f(x_0) = \overline{\psi}_{i_1}(x_0) \wedge \dots \wedge \overline{\psi}_{i_k}(x_0) = 1$$

即 $\|\varphi^*\| \subseteq \|\psi^*\|$. 所以 $\varphi^* \rightarrow \psi^*$ 是决策蕴涵. \square

下面探讨基于现有决策蕴涵经过 MP 近似推理生成新的决策蕴涵的模式 1.

模式 1. 决策蕴涵的 MP 近似推理生成模式

前提. 决策蕴涵集 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$, 条件属性合取公式 φ^* .

结论. 决策蕴涵 $\varphi^* \rightarrow \psi^*$.

1) 计算属性公式函数 $\overline{\varphi^*}, \overline{\varphi}_i, \overline{\psi}_i, i = 1, \dots, n$.

2) 计算 MP 近似推理解函数:

$$f(x) = \bigwedge_{i=1}^n \bigvee_{y \in G} [\overline{\varphi^*}(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in G$$

3) 计算 $W = \{x \in G | f(x) = 1\}$, $\alpha^J(W) = \{d \in D | \forall x \in W, xJd\}$.

4) 生成决策蕴涵 $\varphi^* \rightarrow \bigwedge_{d \in \alpha^J(W)} d$.

5) 对任何 $B \subseteq \alpha^J(W)$, 生成决策蕴涵 $\varphi^* \rightarrow \bigwedge_{d \in B} d$.

例 3. 一个决策形式背景 $K = (G, C, D, I, J)$ 见表 3, 对象集 $G = \{u_1, u_2, u_3, u_4, u_5, u_6, u_7\}$, 条件属性集 $C = \{a_1, a_2, a_3, a_4, a_5, a_6\}$, 决策属性集 $D = \{d_1, d_2, d_3\}$.

表 3 决策形式背景 $K = (G, C, D, I, J)$

Table 3 A formal decision context $K = (G, C, D, I, J)$

G	a_1	a_2	a_3	a_4	a_5	a_6	d_1	d_2	d_3
u_1	1	1	0	0	1	1	1	0	1
u_2	0	0	1	1	1	1	1	1	0
u_3	1	0	1	0	0	1	0	0	1
u_4	0	1	0	1	0	1	1	0	0
u_5	0	1	0	1	1	1	1	1	0
u_6	1	0	1	0	1	1	0	1	1
u_7	0	1	1	1	0	0	1	1	1

给定 2 个决策蕴涵 $a_1 \wedge a_2 \rightarrow d_1 \wedge d_3$, $a_1 \wedge a_6 \rightarrow d_3$ 和条件属性公式 $a_1 \wedge a_2 \wedge a_6$, 根据决策蕴涵生成模式 1, MP 近似推理得到如下非平凡决策蕴涵 $a_1 \wedge a_2 \wedge a_6 \rightarrow d_1 \wedge d_3$, $a_1 \wedge a_2 \wedge a_6 \rightarrow d_1$, $a_1 \wedge a_2 \wedge a_6 \rightarrow d_3$.

按照决策蕴涵的定义, 对于由原子属性公式构成的蕴涵公式, 当 $\|\varphi\|$ 是空集时, $\varphi \rightarrow \psi$ 一定是决策蕴涵, 但这类决策蕴涵在实际知识获取中作用不是很大, 且极大地增大了决策蕴涵的个数. 因此, 决策蕴涵的 MP 近似推理生成模式 1 没有考虑生成这类平凡的决策蕴涵. 定理 3 和决策蕴涵生成模式 1 给出的决策蕴涵的 MP 近似推理是基于广义 MP 的, 是从逻辑角度来考虑的. 注意到, 广义 MP 的前提是条件属性公式, 前提是决策蕴涵, 而决策蕴涵的前件和后件分别是由条件属性和决策属性的合取构成的属性公式. 从知识发现的角度看, 属性合取公式本质上描述了知识的内涵, 因此决策蕴涵的 MP 近似推理的实质是一种从条件形式背景上的知识来获取决策形式背景上的知识的方法. 如果将条件属性理解为原因, 决策属性理解为结果, 则决策蕴涵的 MP 近似推理遵循的是一种“由因及果”的知识推理模式.

当 $\varphi^* \rightarrow \psi^*$ 是决策蕴涵时, 对任何由 ψ^* 中部分原子属性公式组成的合取公式 γ , $\varphi^* \rightarrow \gamma$ 都是决策蕴涵. 显然 $\varphi^* \rightarrow \bigwedge_{d \in \alpha^J(\|\varphi^*\|)} d$ 是决策蕴涵且含有最多的决策属性. 由定理 3 的证明过程可知, $f(x)$ 和 $\varphi^*(x)$ 不一定总相等, 从而 $\alpha^J(W)$ 和 $\alpha^J(\|\varphi^*\|)$ 不

一定相同. 因此, 对给定的 φ^* , 通过 MP 近似推理得到的决策蕴涵 $\varphi^* \rightarrow \bigwedge_{d \in \alpha^J(W)} d$ 不一定是多规则 MP 近似推理问题式 (3) 含有最多决策蕴涵 $\varphi^* \rightarrow \bigwedge_{d \in \alpha^J(\|\varphi^*\|)} d$. 这是因为由 MP 近似推理得到的决策蕴涵 $\varphi^* \rightarrow \bigwedge_{d \in \alpha^J(W)} d$ 还受到了大前提 $\varphi^* \rightarrow \gamma$ 的约束. 定理 3 的证明过程表明了 $\varphi^* \rightarrow \bigwedge_{d \in \alpha^J(W)} d$ 是得到给定前提 φ^* 和 $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 最大支持的含有最多决策属性的决策蕴涵.

定理 4. 在定理 3 条件下, 有 $\{\varphi_1 \rightarrow \psi_1, \dots, \varphi_n \rightarrow \psi_n\} \models_K \varphi^* \rightarrow \psi^*$.

证明. 由定理 3 可知, $\varphi^* \rightarrow \psi^*$ 是决策蕴涵. 由于 $G \sim \Omega_K$ 间存在一一对应 $\pi: G \rightarrow \Omega_K$, $\pi(x) = v_x$. 对 $x \in G$ 决定的 $v_x \in \Omega_K$, 如果 $v_x(\varphi_1 \rightarrow \psi_1) = 1, \dots, v_x(\varphi_n \rightarrow \psi_n) = 1$, 则 $v_x(\varphi_1) \leq v_x(\psi_1), \dots, v_x(\varphi_n) \leq v_x(\psi_n)$. 即 $\bar{\varphi}_1(x) \leq \bar{\psi}_1(x), \dots, \bar{\varphi}_n(x) \leq \bar{\psi}_n(x)$. 因此:

$$f(x) = \bigwedge_{i=1}^n \bigvee_{y \in G} [\bar{\varphi}^*(y) \wedge (\bar{\varphi}_i(y) \rightarrow \bar{\psi}_i(x))] \geq \bigwedge_{i=1}^n [\bar{\varphi}^*(x) \wedge (\bar{\varphi}_i(x) \rightarrow \bar{\psi}_i(x))]$$

$\forall x \in G$, 如果记 $W = \{x \in G | f(x) = 1\}$, $\alpha^J(W) = \{d \in D | \forall x \in W, xJd\}$, $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$. 则 $\bar{\psi}^*(x) = \bigwedge_{d \in \alpha^J(W)} \bar{d}(x) \geq f(x) \geq \bar{\varphi}^*(x)$. 从而对任何 $v_x \in \Omega_K$, $v_x(\varphi^*) \leq v_x(\psi^*)$, $v_x(\varphi^* \rightarrow \psi^*) = 1$. 所以 $\{\varphi_1 \rightarrow \psi_1, \dots, \varphi_n \rightarrow \psi_n\} \models_K \varphi^* \rightarrow \psi^*$. \square

定义 10. 对于给定的属性公式 $\varphi_1, \dots, \varphi_k$, 称属性公式集 $\varphi_1, \dots, \varphi_k$ 不相容是指 $\forall x \in G, \bar{\varphi}_1(x) \wedge \dots \wedge \bar{\varphi}_k(x) = 0$; 否则, 称为相容.

命题 7. 设 $\varphi \rightarrow \psi$ 是决策形式背景 $K = (G, C, D, I, J)$ 的决策蕴涵, φ^* 是与 φ 不相容的条件属性公式, 则由 $\varphi \rightarrow \psi$ 和 φ^* 经过 MP 近似推理可以得到前件为 φ^* 的全部决策蕴涵.

证明. $\varphi \rightarrow \psi$ 和 φ^* 的 MP 近似推理的解函数为:

$$f(x) = \bigvee_{y \in G} [\bar{\varphi}^*(y) \wedge (1 - \bar{\varphi}(y))] \vee [\bar{\psi}(x) \wedge \bigvee_{y \in G} \bar{\varphi}^*(y)]$$

由于 φ^* 与 φ 不相容, 故 $\forall x \in G, \bar{\varphi}^*(x) \wedge \bar{\varphi}(x) = 0$. 于是 $\forall x \in G, f(x) = 1$. 由定理 3 的证明可知, 对任何 $W \subseteq G = \{x \in G | f(x) = 1\}$, $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$, $\varphi^* \rightarrow \psi^*$ 是决策蕴涵. \square

命题 7 表明, 一个与决策蕴涵前提不相容的条件属性公式 φ^* , 与决策蕴涵一起可以 MP 近似推理得出前件为 φ^* 的全部决策蕴涵.

命题 8. 在定理 3 的条件下, 有 $\{\varphi_1 \rightarrow \psi_1, \varphi_2 \rightarrow \psi_2\} \models_K \varphi_1 \wedge \varphi_2 \rightarrow \psi_1 \wedge \psi_2$.

证明. 由定理 3, $\{\varphi_1 \rightarrow \psi_1, \varphi_2 \rightarrow \psi_2\} \models_K \varphi_1 \wedge \varphi_2 \rightarrow \psi^*$, ψ^* 是由 $\{\varphi_1 \rightarrow \psi_1, \varphi_2 \rightarrow \psi_2\}$ 和 $\varphi_1 \wedge \varphi_2$

的 MP 近似推理的结论. 因此:

$$f(x) = \bigwedge_{i=1}^2 \bigvee_{y \in G} [\varphi_1 \wedge \varphi_2(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))] = \overline{\psi}_1(x) \wedge \overline{\psi}_2(x)$$

$W = \{x \in G | \overline{\psi}_1(x) \wedge \overline{\psi}_2(x) = 1\}$, 又由于 $\alpha^J(W) = \{d \in D | \forall x \in W, xJd\}$, 故结论 $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$ 包含 $\psi_1 \wedge \psi_2$ 的全部原子属性公式. 所以 $\{\varphi_1 \rightarrow \psi_1, \varphi_2 \rightarrow \psi_2\} \models_K \varphi_1 \wedge \varphi_2 \rightarrow \psi_1 \wedge \psi_2$. \square

定理 5. 设 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 是决策形式背景 $K = (G, C, D, I, J)$ 的概念规则集, φ^* 是条件概念的内涵属性公式, f 是由 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 的 MP 近似推理得到的解函数:

$$f(x) = \bigwedge_{i=1}^n \bigvee_{y \in G} [\overline{\varphi}^*(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in G$$

若记 $W = \{x \in G | f(x) = 1\}$, $\alpha^J(W) = \{d \in D | \forall x \in W, xJd\}$. 当 $\alpha^J(W)$ 非空时, 令 $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$, 则 $\varphi^* \rightarrow \psi^*$ 是概念规则.

证明. 由定理 3 可知, $\varphi^* \rightarrow \psi^*$ 是决策蕴涵. 显然有 $(\|\varphi^*\|, \alpha^J(W))$ 是决策概念, 即 ψ^* 是决策概念的内涵属性公式. 与定理 3 的证明类似, 可以证明 $\|\varphi^*\| \subseteq \|\psi^*\|$, $\varphi^* \rightarrow \psi^*$ 是概念规则. \square

由定理 5, 在基于 MP 近似推理的决策蕴涵生成模式中, 用概念规则代替决策蕴涵, 条件概念属性公式代替条件属性公式, 可以得到相应的基于 MP 近似推理的概念规则生成方法.

例 4 (续例 3). $a_2 \wedge a_4 \rightarrow d_1$, $a_3 \wedge a_4 \rightarrow d_2$ 是两条概念规则, $a_2 \wedge a_3 \wedge a_4$ 是条件概念内涵属性公式, 利用决策蕴涵生成模式 1, 得到概念规则: $a_2 \wedge a_3 \wedge a_4 \rightarrow d_1$, $a_2 \wedge a_3 \wedge a_4 \rightarrow d_2$, $a_2 \wedge a_3 \wedge a_4 \rightarrow d_3$, $a_2 \wedge a_3 \wedge a_4 \rightarrow d_1 \wedge d_2$, $a_2 \wedge a_3 \wedge a_4 \rightarrow d_2 \wedge d_3$, $a_2 \wedge a_3 \wedge a_4 \rightarrow d_1 \wedge d_3$, $a_2 \wedge a_3 \wedge a_4 \rightarrow d_1 \wedge d_2 \wedge d_3$.

以上结论表明, 本文给出的 MP 近似推理方法是一种决策蕴涵和概念规则的推理模式, 即决策蕴涵和概念规则集在 MP 近似推理模式下是封闭的.

3.2 决策蕴涵和概念规则的 MT 近似推理模式

本节讨论决策形式背景上基于广义 MT 的近似推理 (以下简称 MT 近似推理) 的特征.

定理 6. 设 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 是决策形式背景 $K = (G, C, D, I, J)$ 的决策蕴涵集, ψ^* 是决策属性合取公式, 由 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 ψ^* 的 MT 近似推理得到的解函数为:

$$g(y) = \bigwedge_{i=1}^n \bigwedge_{x \in G} [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x)) \rightarrow \overline{\varphi}^*(x)], y \in G$$

若记 $V = \{y \in G | g(y) = 1\}$, $\alpha^I(V) = \{c \in C |$

$\forall y \in V, yIc\}$. 当 $\alpha^I(V)$ 非空时, 令 $\varphi^* = \bigwedge_{c \in \alpha^I(V)} c$, 则 $\varphi^* \rightarrow \psi^*$ 是决策蕴涵.

证明. 设 $y_0 \notin \|\psi^*\|$, 即 $\psi^*(y_0) = 0$. 如果记:

$$g_i(y) = \bigwedge_{x \in G} [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x)) \rightarrow \overline{\varphi}^*(x)], \\ i = 1, \dots, n$$

由注 2, 在 $\psi^*(y_0) = 0$ 的条件下, 只有满足 $\|\psi_i\| \subseteq \|\psi^*\|$ 时, 才会有 $g_i(y) = \overline{\varphi}_i(y)$; 否则, $g_i(y)$ 在 G 上恒为 0. 不妨设决策蕴涵集 L 的后件属性公式集 $\{\psi_1, \dots, \psi_n\}$ 中有 i_l 个属性公式 $\psi_{i_1}, \dots, \psi_{i_l}$ 满足条件 $\|\psi_{i_1}\| \subseteq \|\psi^*\|, \dots, \|\psi_{i_l}\| \subseteq \|\psi^*\|$. 从而有:

$$g(y) = \bigwedge_{i=1}^n g_i(y) = \overline{\varphi}_{i_1}(y) \wedge \dots \wedge \overline{\varphi}_{i_l}(y), y \in G$$

记 $V = \{y \in G | g(y) = 1\}$, $\alpha^I(V) = \{c \in C | \forall y \in V, yIc\}$. 当 $\alpha^I(V)$ 非空时, 令 $\varphi^* = \bigwedge_{c \in \alpha^I(V)} c$, 注意到都是原子属性公式的合取, 于是有:

$$\overline{\psi}^*(y) = \bigwedge_{c \in \alpha^I(V)} \overline{c}(y) = g(y) = \overline{\varphi}_{i_1}(y) \wedge \dots \wedge \overline{\varphi}_{i_l}(y)$$

$\forall y \in G$. 由于 $\varphi_{i_1} \rightarrow \psi_{i_1}, \dots, \varphi_{i_l} \rightarrow \psi_{i_l}$ 是决策蕴涵, 故有 $\|\varphi_{i_1}\| \subseteq \|\psi_{i_1}\|, \dots, \|\varphi_{i_l}\| \subseteq \|\psi_{i_l}\|$. 又由 $\overline{\psi}^*(y_0) = 0$, 知 $\overline{\psi}_1(y_0) = 0, \dots, \overline{\psi}_l(y_0) = 0$, 从而 $\overline{\varphi}_1(y_0) = 0, \dots, \overline{\varphi}_l(y_0) = 0$. 故 $\overline{\psi}^*(y_0) = \overline{\varphi}_{i_1}(y_0) \wedge \dots \wedge \overline{\varphi}_{i_l}(y_0) = 0$, 即 $y_0 \notin \|\varphi^*\|$. 因此 $\|\varphi^*\| \subseteq \|\psi^*\|$. 所以 $\varphi^* \rightarrow \psi^*$ 是决策蕴涵. \square

定理 6 给出的决策蕴涵的 MT 近似推理是基于广义 MT 的. 从知识发现的思维过程看, 广义 MT 有着与 MP 相反的思维模式. 因此, 决策蕴涵的 MT 近似推理本质上应属于“由果溯因”的知识推理模式.

命题 9. 在定理 6 的条件下, 有 $\{\varphi_1 \rightarrow \psi_1, \dots, \varphi_n \rightarrow \psi_n\} \models_K \varphi^* \rightarrow \psi^*$.

证明. 由定理 6 知 $\varphi^* \rightarrow \psi^*$ 是决策蕴涵. 由于 G 和 Ω_K 之间有一一对应 $\pi: G \rightarrow \Omega_K$, $\pi(y) = v_y$. 对 $y \in G$ 决定的 $v_y \in \Omega_K$, 如果 $v_y(\varphi_1 \rightarrow \psi_1) = 1, \dots, v_y(\varphi_n \rightarrow \psi_n) = 1$, 则 $v_y(\varphi_1) \leq v_y(\psi_1), \dots, v_y(\varphi_n) \leq v_y(\psi_n)$. 即 $\overline{\varphi}_1(y) \leq \overline{\psi}_1(y), \dots, \overline{\varphi}_n(y) \leq \overline{\psi}_n(y)$. 因此:

$$g(y) = \bigwedge_{i=1}^n \bigwedge_{x \in G} [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x)) \rightarrow \overline{\varphi}^*(x)] \leq \bigwedge_{i=1}^n [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(y)) \rightarrow \overline{\varphi}^*(y)] = \overline{\psi}^*(y)$$

$\forall y \in G$. 又 $V = \{y \in G | g(y) = 1\}$, $\alpha^I(V) = \{c \in C | \forall y \in V, yIc\}$, $\varphi^* = \bigwedge_{c \in \alpha^I(V)} c$, 于是 $\overline{\psi}^*(y) \geq g(y) = \bigwedge_{c \in \alpha^I(V)} \overline{c}(y) = \overline{\varphi}^*(y)$. 故 $\forall v_y \in \Omega_K$, $v_y(\varphi^*) \leq v_y(\psi^*)$, $v_y(\varphi^* \rightarrow \psi^*) = 1$. 所以 $\{\varphi_1 \rightarrow \psi_1, \dots, \varphi_n \rightarrow \psi_n\} \models_K$

$\varphi^* \rightarrow \psi^*$. \square

定理 7. 设 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 是决策形式背景 $K = (G, C, D, I, J)$ 的概念规则集, ψ^* 是决策概念内涵属性合取公式, g 是由 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 ψ^* 的 MT 近似推理得到的解函数:

$$g(y) = \bigwedge_{i=1}^n \bigwedge_{x \in G} [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x)) \rightarrow \overline{\varphi^*}(x)], y \in G$$

若记 $V = \{y \in G | g(y) = 1\}$, $\alpha^I(V) = \{c \in C | \forall y \in V, yIc\}$. 当 $\alpha^I(V)$ 非空时, 令 $\varphi^* = \bigwedge_{c \in \alpha^I(V)} c$, 则 $\varphi^* \rightarrow \psi^*$ 是概念规则.

证明. 显然有 $(\|\varphi^*\|, \alpha^I(V))$ 是条件概念, 即 φ^* 是条件概念的内涵属性公式. 与定理 6 的证明类似, 可证明 $\|\varphi^*\| \subseteq \|\psi^*\|$, $\varphi^* \rightarrow \psi^*$ 是概念规则. \square

由定理 5 和定理 7 可知, 决策蕴涵和概念规则集在 MT 近似推理模式下也是封闭的.

应用定理 6 的结论, 可以得到下述基于 MT 近似推理的决策蕴涵和概念规则生成模式 2.

模式 2. 决策蕴涵的 MT 近似推理生成模式

前提. 决策蕴涵集 $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$, 决策属性合取公式 ψ^* .

结论. 决策蕴涵 $\varphi^* \rightarrow \psi^*$.

- 1) 计算属性公式函数 $\overline{\psi^*}, \overline{\varphi}_i, \overline{\psi}_i, i = 1, \dots, n$;
- 2) 计算 MT 近似推理解函数:

$$g(y) = \bigwedge_{i=1}^n \bigwedge_{x \in G} [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x)) \rightarrow \overline{\varphi^*}(x)], y \in G.$$

- 3) 计算 $V = \{y \in G | g(y) = 1\}$, $\alpha^I(V) = \{c \in C | \forall y \in V, yIc\}$;
- 4) 生成决策蕴涵 $\bigwedge_{c \in \alpha^I(V)} c \rightarrow \psi^*$;
- 5) 对任何 $A \supseteq \alpha^I(A)$, 生成决策蕴涵 $\bigwedge_{c \in A} c \rightarrow \psi^*$.

由定理 7, 在基于 MT 近似推理的决策蕴涵生成模式中, 用概念规则代替决策蕴涵, 决策概念内涵属性公式代替决策属性公式, 可以得到相应的基于 MT 近似推理的概念规则生成方法.

例 5 (续例 3). 对于给定的 2 条概念规则 $(a_2 \wedge a_4 \wedge a_5 \wedge a_6 \rightarrow d_1 \wedge d_2, a_1 \wedge a_3 \wedge a_5 \wedge a_6 \rightarrow d_1 \wedge d_3)$ 和决策概念内涵属性公式 d_2 , 利用决策蕴涵生成模式 2, 经过 MT 近似推理得到概念规则 $a_1 \wedge a_2 \wedge a_3 \wedge a_4 \wedge a_5 \wedge a_6 \rightarrow d_2$.

4 近似推理生成决策蕴涵的有效性分析

第 3 节分析了决策形式背景全体对象集上的决策蕴涵的 MP、MT 近似推理的特征, 给出了利用 MP、MT 近似推理生成决策蕴涵的方法. 在人工智能环境下, 在应用 MP、MT 近似推理进行知识获取时, 经常需要基于部分数据集上的现有决策蕴涵经

过 MP、MT 近似推理动态地挖掘更大数据集上的决策蕴涵, 即遇到更多的是数据对象动态变化时的决策蕴涵挖掘. 另一方面, 利用少量的决策蕴涵借助 MP、MT 近似推理模式产生更多的决策蕴涵, 这正是知识推理的目的. 如何评价近似推理的好坏, 即如何评价近似推理结论的有效性, 目前还没有见到公认的标准. 本节引入属性逻辑伪距离, 对 MP、MT 近似推理结论的有效性进行刻画, 进而提出基础数据对象集动态变化时应用 MP、MT 近似推理来获取决策蕴涵的算法.

例 6 (续例 3). 给定决策蕴涵 $a_1 \wedge a_2 \rightarrow d_1 \wedge d_3, a_1 \wedge a_6 \rightarrow d_3$ 和条件属性公式 $a_2 \wedge a_6$.

如果近似推理的基础数据对象集是整个对象集, MP 近似推理得到的解函数为 $f(x) = 1, x \in G$, 这时 MP 近似推理的结论是空集. 如果推理范围限定为 $\Sigma = \{u_1, u_2, u_3, u_6, u_7\}$, MP 近似推理得到的解函数为 $f(x) = \overline{d_1} \wedge \overline{d_3}(x), x \in G$, 从而可得属性公式 $a_2 \wedge a_6 \rightarrow d_1 \wedge d_3$. 注意到, 在整个对象集 G 上, 它不是决策蕴涵, 但如果限定在对象集 Σ 上, 则它是决策蕴涵. 本文称限定在部分对象集上的是决策蕴涵的属性公式为拟决策蕴涵.

定义 11. 设 $K = (G, M, I)$ 是形式背景, 属性公式 $\varphi, \psi \in \Gamma(K)$, $\Sigma \subseteq G$. 定义 φ 在 Σ 上的真度为:

$$\tau_{\Sigma}(\varphi) = \int_{\Sigma} \overline{\varphi}(x) dP$$

式中, P 是 G 上的概率测度. 属性公式 φ, ψ 的 Σ -伪距离定义为:

$$\rho_{\Sigma}(\varphi, \gamma) = 1 - \tau_{\Sigma}((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi))$$

当 P 是 G 上的均匀概率测度 (计数测度) 时, 定义 11 中定义的真度可表示为:

$$\tau_{\Sigma}(\varphi) = \frac{1}{m} \sum_{i=1}^m \overline{\varphi}(x_i)$$

式中, m 为对象集 Σ 的元素个数.

由于 φ 是一些属性的合取, 按照决策逻辑思想, $\overline{\varphi}(x) = 1$ 意味着支持“ x 支持 φ ”, 即对象 x 具有 φ 中的全部属性, 因此真度 τ_{Σ} 可以被理解为一种“投票模型”. 由于在实际数据处理中形式背景通常是用表格表示的, 这时对象 x 的概率 $P(x)$ 其实就是与 x 有相同属性值的对象数与全体对象个数的比值, 而公式 φ 的真度 $\tau_{\Sigma}(\varphi)$ 其实就是 Σ 上具有 φ 中全部属性的对象的概率和. 当 P 是 G 上的计数测度时, $\tau_{\Sigma}(\varphi)$ 是 Σ 上具有 φ 中全部属性的对象个数与 Σ 中全体对象个数的比值.

命题 10. 设 $\varphi, \psi \in \Gamma(K)$, $\Sigma \subseteq G$. 则:

$$\begin{aligned} \rho_{\Sigma}(\varphi) &= \int_{\Sigma} |\overline{\varphi}(x) - \overline{\psi}(x)| dP = \\ &= \frac{1}{m} \sum_{i=1}^m |\overline{\varphi}(x_i) - \overline{\psi}(x_i)| \end{aligned}$$

证明.

$$\begin{aligned} \rho_{\Sigma}(\varphi, \psi) &= 1 - \tau_{\Sigma}((\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)) = \\ &= 1 - \int_{\Sigma} (\overline{\varphi \rightarrow \psi}(x) \wedge \overline{\psi \rightarrow \varphi}(x)) dP = \\ &= 1 - \int_{\Sigma} (\overline{\varphi}(x) \rightarrow \overline{\psi}(x) \wedge \overline{\psi}(x) \rightarrow \overline{\varphi}(x)) dP = \\ &= 1 - \int_{\Sigma} (1 - \overline{\varphi}(x) + \overline{\psi}(x)) \wedge \\ &= (1 - \overline{\psi}(x) + \overline{\varphi}(x)) dP = \\ &= \int_{\Sigma} [1 - (1 - \overline{\varphi}(x) + \overline{\psi}(x)) \wedge \\ &= (1 - \overline{\psi}(x) + \overline{\varphi}(x))] dP = \\ &= \int_{\Sigma} |\overline{\varphi}(x) - \overline{\psi}(x)| dP \quad \square \end{aligned}$$

设 $F \subseteq \Gamma(K)$, $\Sigma \subseteq G$, 称 $\gamma \in \Gamma(K)$ 是 F 的 Σ -语义结论是指 Σ 对所决定的任何赋值 $v_x \in v_{\Sigma}$, 如果 $v_x(F) = \bigwedge \{v_x(\varphi) | \varphi \in F\} = 1$, 则有 $v_x(\gamma) = 1$. F 的全体 Σ -语义结论之集记为 $D_{\Sigma}(F)$.

定义 12. 设 $K = (G, M, I)$ 是形式背景, $F \subseteq \Gamma(K)$, $\varphi \in \Gamma(K)$, $\Sigma \subseteq G$, $\varepsilon > 0$. 如果 $\text{div}(F) = \inf\{\rho_{\Sigma}(\varphi, \gamma) | \gamma \in D_{\Sigma}(F)\}$, 则称 φ 为 F 的误差小于 ε 的结论, 记为 $\varphi \in D_{\varepsilon}(F)$.

定理 8. 设 $K = (G, M, I)$ 是形式背景, $F \subseteq \Gamma(K)$, $\varphi \in \Gamma(K)$, $\Sigma \subseteq G$, $\varepsilon > 0$. 则 $\varphi \in D_{\varepsilon}(F)$ 当且仅当 $\sup\{\tau_{\Sigma}(\gamma \rightarrow \varphi) | \gamma \in D_{\Sigma}(F)\} > 1 - \varepsilon$.

证明. 设 $\varphi \in D_{\varepsilon}(F)$, 则存在属性公式 $\gamma_0 \in D_{\Sigma}(F)$ 使得 $\rho_{\Sigma}(\varphi, \gamma_0) < \varepsilon$. 则:

$$\begin{aligned} \rho_{\Sigma}(\varphi, \gamma_0) &= 1 - \tau_{\Sigma}((\varphi \rightarrow \gamma_0) \wedge (\gamma_0 \rightarrow \varphi)) \geq \\ &= 1 - \tau_{\Sigma}(\gamma_0 \rightarrow \varphi) \end{aligned}$$

因此

$$\begin{aligned} 1 - \sup\{\tau_{\Sigma}(\gamma \rightarrow \varphi) | \gamma \in D_{\Sigma}(F)\} &\leq \\ 1 - \tau_{\Sigma}(\gamma_0 \rightarrow \varphi) | \gamma \in D_{\Sigma}(F) &< \varepsilon \end{aligned}$$

所以, $\sup\{\tau_{\Sigma}(\gamma \rightarrow \varphi) | \gamma \in D_{\Sigma}(F)\} > 1 - \varepsilon$.

反之, 如果 $\sup\{\tau_{\Sigma}(\gamma \rightarrow \varphi) | \gamma \in D_{\Sigma}(F)\} > 1 - \varepsilon$, 则 $\inf\{1 - \tau_{\Sigma}(\gamma \rightarrow \varphi) | \gamma \in D_{\Sigma}(F)\} < \varepsilon$. 易证:

$$(\varphi \rightarrow \varphi \vee \gamma) \wedge (\varphi \vee \gamma \rightarrow \varphi) \approx \gamma \rightarrow \varphi$$

则:

$$\begin{aligned} \rho_{\Sigma}(\varphi, \varphi \vee \gamma) &= 1 - \tau_{\Sigma}((\varphi \rightarrow \varphi \vee \gamma) \wedge (\varphi \vee \gamma \rightarrow \varphi)) = \\ &= 1 - \tau_{\Sigma}(\gamma \rightarrow \varphi) \end{aligned}$$

故 $\inf\{\rho_{\Sigma}(\varphi, \varphi \vee \gamma) | \gamma \in D_{\Sigma}(F)\} < \varepsilon$. 于是存在 $\gamma \in D_{\Sigma}(F)$ 使得 $\rho_{\Sigma}(\varphi, \varphi \vee \gamma) < \varepsilon$. 又由 $\gamma \in D_{\Sigma}(F)$ 有 $\varphi \vee \gamma \in D_{\Sigma}(F)$. 因此:

$$\inf\{\rho_{\Sigma}(\varphi, \gamma) | \gamma \in D_{\Sigma}(F)\} \leq \rho_{\Sigma}(\varphi, \varphi \vee \gamma) < \varepsilon$$

所以, $\varphi \in D_{\varepsilon}(F)$. \square

定理 9. 设 $\Sigma \subset \Delta \subseteq G$, ψ^* 是 $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 的 MP 近似推理结论, 则 ψ^* 是 L 的后件属性公式集 $F = \{\psi_1, \dots, \psi_n\}$ 的误差小于 $(|\Delta| - |\Sigma|)/|\Delta|$ 的结论, 其中 $|\Delta|$ 表示 Δ 的元素个数.

证明. $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 的 MP 近似推理的 (Σ, Δ) 解函数为:

$$f(x) = \bigwedge_{i=1}^n \bigvee_{y \in \Sigma} [\varphi_i^*(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in \Delta \quad (9)$$

设 $\|\varphi^*\|_{\Delta} \cap \Sigma \subseteq \|\varphi_{i_1}\|_{\Delta} \cap \Sigma, \dots, \|\varphi^*\|_{\Delta} \cap \Sigma \subseteq \|\varphi_{i_k}\|_{\Delta} \cap \Sigma$. 从而 $f(x) = \overline{\psi}_{i_1}(x) \wedge \dots \wedge \overline{\psi}_{i_k}(x)$, $x \in \Delta$. 令 $\gamma = \psi_{i_1} \wedge \dots \wedge \psi_{i_k}$, 显然有 $\gamma \in D_{\Sigma}(F)$. 又记 $W = \{x \in \Delta | f(x) = 1\}$, $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$. 由式 (9) 和决策蕴涵生成模式 1 可知, $\forall x \in \Sigma$,

$$\overline{\varphi^*}(x) = \bigwedge_{d \in \alpha^J(W)} \overline{d}(x) \geq f(x) = \overline{\psi}_{i_1}(x) \wedge \dots \wedge \overline{\psi}_{i_k}(x),$$

即 $\forall x \in \Sigma$, $\overline{\gamma \rightarrow \psi^*}(x) = 1$. 所以:

$$\begin{aligned} \tau_{\Delta}(\gamma \rightarrow \psi^*) &= \int_{\Delta} \overline{\gamma \rightarrow \psi^*}(x) dP \geq \\ \frac{|\Sigma|}{|\Delta|} &= 1 - \frac{|\Delta| - |\Sigma|}{|\Delta|} \end{aligned}$$

\square

命题 11. 设 $\Sigma \subset \Delta_1 \subset \Delta_2$, $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 的 MP 近似推理的 (Σ, Δ_1) , (Σ, Δ_2) 结论分别是 ψ_1^* 和 ψ_2^* . 则 ψ_2^* 比 ψ_1^* 包含有更多的原子属性公式, 且 ψ_2^* 和 $F = \{\psi_1, \dots, \psi_n\}$ 的距离小于 ψ_1^* 和 $F = \{\psi_1, \dots, \psi_n\}$ 的距离.

证明. $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 的 MP 近似推理的 (Σ, Δ_1) , (Σ, Δ_2) 解函数分别为 $f_1(x)$ 和 $f_2(x)$:

$$f_1(x) = \bigwedge_{i=1}^n \bigvee_{y \in \Sigma} [\varphi_i^*(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in \Delta_1$$

$$f_2(x) = \bigwedge_{i=1}^n \bigvee_{y \in \Sigma} [\varphi_i^*(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in \Delta_2$$

显然 $\forall x \in \Delta_1$, $f_1(x) = f_2(x)$. 不妨设在决策蕴涵集 L 的前件属性公式集 $\{\varphi_1, \dots, \varphi_n\}$ 中有 i_k 个前件属性公式 $\varphi_{i_1}, \dots, \varphi_{i_k}$ 满足条件 $\|\varphi^*\|_{\Delta} \cap \Sigma \subseteq \|\varphi_{i_1}\|_{\Delta} \cap \Sigma, \dots, \|\varphi^*\|_{\Delta} \cap \Sigma \subseteq \|\varphi_{i_k}\|_{\Delta} \cap \Sigma$. 因此, 类似于定

理 3 的证明, 不妨设:

$$f_1(x) = \overline{\psi}_{i_1}(x) \wedge \cdots \wedge \overline{\psi}_{i_k}(x), x \in \Delta_1$$

$$f_2(x) = \overline{\psi}_{i_1}(x) \wedge \cdots \wedge \overline{\psi}_{i_k}(x), x \in \Delta_2$$

记 $\gamma = \psi_{i_1} \wedge \cdots \wedge \psi_{i_k}$, $W_1 = \{x \in \Delta_1 | f(x) = 1\}$, $\psi_1^* = \bigwedge_{d \in \alpha^J(W_1)} d$, $W_2 = \{x \in \Delta_2 | f(x) = 1\}$, $\psi_2^* = \bigwedge_{d \in \alpha^J(W_2)} d$. 显然 $W_1 \subseteq W_2$, 从而 ψ_1^* 比 ψ_2^* 包含有更多的决策原子属性公式, 于是 $\forall x \in \Delta_2$, $\overline{\psi}_1^*(x) \leq \overline{\psi}_2^*(x)$. 因此由伪距离的定义知, ψ_2^* 与 γ 间的伪距离比 ψ_1^* 与 γ 间的伪距离要小. \square

由命题 11, 当 $\Sigma \subset \Delta_1 \subset \Delta_2$ 时, $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 的 MP 近似推理的 (Σ, Δ_1) 结论 ψ_1^* 比 (Σ, Δ_2) 结论 ψ_2^* 有更小的精确度.

定理 10. 设 $\Sigma \subset \Delta \subseteq G$, φ^* 是 $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 ψ^* 的 MT 近似推理结论. 则 φ^* 是 L 的前件属性公式集 $F = \{\varphi_1, \dots, \varphi_n\}$ 的误差小于 $(|\Delta| - |\Sigma|)/|\Delta|$ 的结论.

证明. $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 ψ^* 的多规则 MT 近似推理的 (Σ, Δ) 解函数为:

$$g(y) = \bigwedge_{i=1}^n \bigwedge_{x \in \Sigma} [(\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x)) \rightarrow \overline{\varphi}^*(x)], y \in \Delta \quad (10)$$

设 $\|\varphi_{i_1}\|_{\Delta} \cap \Sigma \subseteq \|\varphi^*\|_{\Delta} \cap \Sigma, \dots, \|\varphi_{i_n}\|_{\Delta} \cap \Sigma \subseteq \|\varphi^*\|_{\Delta} \cap \Sigma$. 从而 $f(x) = \overline{\varphi}_{i_1}(x) \wedge \cdots \wedge \overline{\varphi}_{i_n}(x)$, $x \in \Delta$. 令 $\gamma = \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_n}$, 显然有 $\gamma \in D_{\Sigma}(F)$. 又记 $V = \{y \in \Delta | g(y) = 1\}$, $\varphi^* = \bigwedge_{c \in \alpha^I(V)} c$. 由式 (9) 和决策蕴涵生成模式 2 可知:

$$\overline{\varphi}^*(y) = \bigwedge_{c \in \alpha^I(V)} \overline{c}(y) \geq g(y) = \overline{\varphi}_{i_1}(x) \wedge \cdots \wedge \overline{\varphi}_{i_n}(x)$$

$\forall y \in \Sigma$. 则 $\forall y \in \Sigma$, $\overline{\gamma} \rightarrow \overline{\varphi}^*(y) = 1$. 所以:

$$\tau_{\Delta}(\gamma \rightarrow \varphi^*) = \int_{\Delta} \overline{\gamma} \rightarrow \overline{\varphi}^*(y) dP \geq \frac{|\Sigma|}{|\Delta|} = 1 - \frac{|\Delta| - |\Sigma|}{|\Delta|}$$

\square

由定理 10、命题 11 的证明, 可证明如下结论成立.

命题 12. 设 $\Sigma \subset \Delta_1 \subset \Delta_2$, $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 ψ^* 的 MT 近似推理的 (Σ, Δ_1) , (Σ, Δ_2) 结论分别是 φ_1^* 和 φ_2^* , 则 φ_2^* 比 φ_1^* 包含更多的原子属性公式, 且 φ_2^* 和 $F = \{\varphi_1, \dots, \varphi_n\}$ 的距离小于 φ_1^* 和 $F = \{\varphi_1, \dots, \varphi_n\}$ 的距离.

由例 6 可知, $\{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$ 和 φ^* 在部分对象集 $\Sigma \subset \Delta$ 的具有局部意义的 MP 近似推理结论 ψ^* 不一定使得 $\psi^* \rightarrow \varphi^*$ 是整个对象集 G 上的决策蕴涵. 由定理 9、定理 11、命题 11 和命题 12

可知, 利用 MP、MT 近似推理提取局部意义的决策蕴涵 (以下称为拟决策蕴涵) 生成算法 1 和算法 2.

算法 1. MP 近似推理生成拟决策蕴涵算法

输入. $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$, 条件属性合取公式 φ^* , 部分决策形式背景 $K = (\Delta, C, D, I, J)$, 对象集 $\Sigma \subseteq \Delta$.

输出. 拟决策蕴涵 $\varphi^* \rightarrow \psi^*$.

- 1) 计算 $\|\varphi^*\|_{\Sigma}$, $\|\varphi_i\|_{\Sigma}$, $i = 1, \dots, n$, 其中 $\|\varphi^*\|_{\Sigma} = \{u \in \Sigma | \forall a \in \varphi^*, I(u, a) = 1\}$;
- 2) 找出满足条件 $\|\varphi^*\|_{\Sigma} \subseteq \|\varphi_i\|_{\Sigma}$ 的前件属性公式 $\varphi_{i_1}, \dots, \varphi_{i_k}$, 相应的后件属性公式为 $\psi_{i_1}, \dots, \psi_{i_k}$;
- 3) 如果没有满足条件 2), 推理问题无解; 否则, 令 $\gamma = \psi_{i_1} \wedge \cdots \wedge \psi_{i_k}$;
- 4) 计算 $W = \{u \in \Delta | \forall d \in \Sigma, J(u, d) = 1\}$, $\alpha^J(W) = \{d \in D | \forall u \in W, uJd\}$.
- 5) 令 $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$, 计算属性公式 γ 和 ψ^* 的伪距离;
- 6) 获取拟决策蕴涵 $\varphi^* \rightarrow \psi^*$.

算法 1 的时间复杂度分析如下:

1) 计算 $\|\varphi^*\|_{\Sigma}$, $\|\varphi_i\|_{\Sigma}$, $i = 1, \dots, n$ 的时间复杂度为 $O((n+1)|C||\Sigma|)$, n 为 L 中的决策蕴涵个数, $|C|$ 是条件属性集 C 的元素个数, $|\Sigma|$ 是对象集 Σ 的元素个数.

2) 找出满足条件 $\|\varphi^*\|_{\Sigma} \subseteq \|\varphi_i\|_{\Sigma}$ 的前件属性公式 $\varphi_{i_1}, \dots, \varphi_{i_k}$ 的时间复杂度为 $O(n|\Sigma|)$.

3) 计算 W 的时间复杂度为 $O(|D||\Delta|)$; 计算 $\alpha^J(W)$ 的时间复杂度也是 $O(|D||\Delta|)$, $|D|$ 是决策属性集 D 的元素个数, $|\Delta|$ 是对象集 Δ 的元素个数.

4) 计算 $\psi^* = \bigwedge_{d \in \alpha^J(W)} d$ 的时间复杂度为 $O(|D|)$, 计算 γ 和 ψ^* 的伪距离需要先计算属性公式函数 $\overline{\gamma}$ 和 $\overline{\psi}^*$, 其时间复杂度为 $O(2|D||\Delta|)$, 计算 γ 和 ψ^* 的伪距离时间复杂度为 $O(|\Delta|)$, 所以算法 1 中步骤 5) 的时间复杂度为 $O(2|D||\Delta|)$.

由于 $\Sigma \subseteq \Delta$ 且 $|\Sigma|$ 与 $|\Delta|$ 相差不是很大, 所以 MP 近似推理生成一个拟决策蕴涵算法的时间复杂度为 $O((n+1)|C \cup D||\Delta|)$.

算法 2. MT 近似推理生成拟决策蕴涵算法

输入. $L = \{\varphi_i \rightarrow \psi_i | i = 1, \dots, n\}$, 决策属性合取公式 ψ^* , 部分决策形式背景 $K = (\Delta, C, D, I, J)$, 对象集 $\Sigma \subseteq \Delta$.

输出. 拟决策蕴涵 $\varphi^* \rightarrow \psi^*$.

- 1) 计算 $\|\psi^*\|_{\Sigma}$, $\|\psi_i\|_{\Sigma}$, $i = 1, \dots, n$, 其中 $\|\psi^*\|_{\Sigma} = \{u \in \Sigma | \forall a \in \psi^*, J(u, a) = 1\}$;
- 2) 找出满足条件 $\|\psi_i\|_{\Sigma} \subseteq \|\psi^*\|_{\Sigma}$ 的后件属性公式 $\psi_{i_1}, \dots, \psi_{i_k}$, 相应的前件属性公式为 $\varphi_{i_1}, \dots, \varphi_{i_k}$;
- 3) 如果没有满足式 (2) 中条件的公式, 推理问题无解; 否则, 令 $\lambda = \varphi_{i_1} \wedge \cdots \wedge \varphi_{i_k}$;
- 4) 计算 $V = \{u \in \Delta | \forall c \in \lambda, I(u, c) = 1\}$, $\alpha^I(V) = \{c \in$

$C|\forall u \in V, uIc\}$;

- 5) 令 $\varphi^* = \bigwedge_{c \in \alpha^I(V)} c$, 计算属性公式 λ 和 φ^* 的伪距离;
- 6) 获取拟决策蕴涵 $\varphi^* \rightarrow \psi^*$.

类似于 MP 近似推理生成拟决策蕴涵算法的时间复杂度分析, MT 近似推理生成拟决策蕴涵算法的时间复杂度也是 $O((n+1)|C \cup D||\Delta|)$.

5 数值实验与讨论

本节参考文献 [25], 通过 3 个数值实验验证本文算法的有效性. 由于 MT 近似推理生成拟决策蕴涵算法与 MP 近似推理生成拟决策蕴涵算法的推理机制、步骤和计算表达式类似, 因此实验只使用 MP 近似推理生成拟决策蕴涵算法进行. 基于定理 9 和命题 11 的证明过程, 算法的有效性通过 MP 近似推理结论 ψ_Δ^* 与相应大前提的后件合取公式 $\gamma = \psi_{i_1} \wedge \dots \wedge \psi_{i_k}$ 的伪距离 $\rho_\Delta(\gamma, \psi_\Delta^*)$ 、后件属性公式合取式对 MP 近似推理结论的平均支持度 $\tau_\Delta(\gamma, \psi_\Delta^*)$ 、MP 近似推理获取拟决策蕴涵所时间成本来评价.

实验环境为 Win7 64 位 Intel (R) 3.4 GHz, 4 GB 内存. 实验 1 用来测试 MP 近似推理生成拟决策蕴涵个数及与相应后件公式的伪距离. 实验数据设置为条件属性个数 20, 决策属性个数 10, 对象个数 2000, 分成 4 个数据组, 每组 500 个数据, 数据的属性值 (包括条件属性和决策属性) 随机生成, 数据填充比强度 $|I|/|G||M|$ 为 20%. 每组前 400 个数据为基础数据集 Σ , MP 近似推理的大前提是由 Σ 产生的决策蕴涵集 $L = \{\varphi_u \rightarrow \psi_u | u \in \Sigma\}$, $\varphi_u(c) = I(c, u)$, $c \in C$ 是对象 u 决定的条件属性公式, 决策属性公式 $\psi_u(d) = \bigwedge_{x \in \Sigma} [\beta^I(\varphi_u)(x) \rightarrow J(x, d)]$, $d \in D$, 后 100 个数据中的对象 u 决定的条件属性公式 φ_u^* 作为 MP 近似推理的小前提, 并用来测试算法 1 的效果.

各组后 100 个数据经过 MP 近似推理生成的拟决策蕴涵个数见表 4.

各组数据经过 MP 近似推理产生的拟决策蕴涵的后件 ψ_Δ^* 与相应后件合取公式 $\gamma = \psi_{i_1} \wedge \dots \wedge \psi_{i_k}$ 的最大伪距离和最小伪距离见表 5.

实验 1 基于 400 个实验基础数据, 通过近似推理获得新增加 100 个数据所确定的拟决策蕴涵. 实验结果显示经过 MP 近似推理产生的拟决策蕴涵的后件与相应后件合取公式的伪距离有点大. 但在实际知识推理中, 往往依据大量的数据通过近似推理发现少量新增数据所隐含的新知识. 由定理 9 可知, 当待获取新知识的新增数据个数与基础数据集的数据个数的比值 $(|\Delta| - |\Sigma|)/|\Delta|$ 较小时, MP 近似

表 4 生成的拟决策蕴涵个数

Table 4 The number of generated quasi-decision implications

数据组别	生成的拟决策蕴涵个数
数据组 1	62
数据组 2	58
数据组 3	74
数据组 4	71

表 5 生成的拟决策蕴涵后件与后件合取公式的伪距离

Table 5 The metric between the consequent of generated quasi-decision implications and the consequent conjunctive

数据组别	最大伪距离	最小伪距离
数据组 1	0.195	0.147
数据组 2	0.189	0.160
数据组 3	0.197	0.152
数据组 4	0.194	0.161

推理产生的拟决策蕴涵的后件与相应后件合取公式的伪距离也比较小. 实验 2 和实验 3 也都有效地支持了这个理论分析.

实验 2 用来测试对象集 Σ 、 Δ 的变化对 MP 近似推理生成拟决策蕴涵的个数及与相应后件合取公式的伪距离的影响. 实验数据设置为条件属性为 20 个, 决策属性为 10 个, 实验基础数据对象集 Σ 的元素个数 $|\Sigma|$ 为 1000, 近似推理新增对象以 10 个为 1 组的频率, 逐步增加到 100 个, 从而 Δ 的元素个数 $|\Delta|$ 动态地逐步增加到 1010, 1020, \dots , 1100, 相应的数据组分别被命名为数据组 1, 数据组 2, \dots , 数据组 10, 数据填充比强度与实验 1 相同, 为 20%.

基于基础数据对象集 Σ 产生的 MP 近似推理的决策蕴涵集 L 和逐次新增加数据中的对象 u 所决定的条件属性公式 φ_u 进行 MP 近似推理. MP 近似推理生成的拟决策蕴涵的个数以及生成的拟决策蕴涵的后件与相应后件合取公式 $\gamma = \psi_{i_1} \wedge \dots \wedge \psi_{i_k}$ 的最小伪距离见表 6.

由表 5 和表 6 可以看出, 在基础数据集确定前提下, 从新增加数据所生成的拟决策蕴涵个数以及生成的拟决策蕴涵的后件与相应的后件合取公式 $\gamma = \psi_{i_1} \wedge \dots \wedge \psi_{i_k}$ 的最小伪距离都随着测试数据的增加而增加, 增加趋势基本是线性的. 这与定理 9、命题 11 的理论分析结论相吻合.

实验 3 用来测试在同一个数据集上、采用不同方式获取拟决策蕴涵时的时间成本比较. 实验 3 的实验数据取值方式和数据填充比强度同实验 1. 基础数据对象集 Σ 的元素个数 $|\Sigma|$ 为 2000 个, MP 近

表 6 测试数据变化生成的拟决策蕴涵表

Table 6 A table of generated quasi-decision implication as test data changes

数据组别	拟决策蕴涵个数	最小伪距离
数据组 1	3	0.0082
数据组 2	5	0.0157
数据组 3	8	0.0256
数据组 4	12	0.0324
数据组 5	16	0.0418
数据组 6	21	0.0527
数据组 7	25	0.0619
数据组 8	29	0.0718
数据组 9	34	0.0821
数据组 10	39	0.0913

似推理的新增数据对象以 100 个为 1 组的频率, 逐步增加到 500 个, 从而 Δ 的元素个数 $|\Delta|$ 动态地逐次增加到 2 100, 2 200, \dots , 2 500, 相应的数据组分别被命名为数据组 1, 数据组 2, \dots , 数据组 5.

根据命题 5 中步骤 1), 由包含 2 000 个对象的基础数据集 Σ 中的每个对象 u 产生决策蕴涵 (对象 u 决策蕴涵产生的过程同实验 1), 以如此产生的全部决策蕴涵集 $L = \{\varphi_u \rightarrow \psi_u | u \in \Sigma\}$ 作为 MP 近似推理的大前提. MP 近似推理的小前提是各数据组中新增加对象 $u \in \Delta - \Sigma$ 所决定的条件属性公式 φ_u^* . MP 近似推理的对象范围参数 Σ 为 2 000 个对象组成的基础数据集, 参数 Δ 分别为在 Σ 上逐次加入新数据得到的数据组 1, 数据组 2, \dots , 数据组 5. 运用算法 1, 分别由 L 和 φ_u^* 进行 MP 近似推理生成以 φ_u^* 为前件的拟决策蕴涵. 采用 2 种方式获取 $\Delta - \Sigma$ 上的新增决策蕴涵: 方式 1) 是由各数据组的新增加对象 $u \in \Delta - \Sigma$ 所决定的条件属性公式 φ_u^* 和 L 直接在 Δ 上进行 MP 近似推理以获取 Δ 上的拟决策蕴涵, 即通过:

$$f_{\Delta u}(x) = \bigwedge_{i=1}^n \bigvee_{y \in \Delta} [\overline{\varphi^*}(y) \wedge (\overline{\varphi}_i(y) \rightarrow \overline{\psi}_i(x))], x \in \Delta$$

$$W_u = \{x \in \Delta | f_{\Delta u}(x) = 1\}, \psi_{\Delta u}^* = \bigwedge_{d \in \alpha^J(W_u)} d$$

获取 Δ 上的决策蕴涵 $\varphi_u^* \rightarrow \psi_{\Delta u}^*$.

方式 2) 是先由基础数据集 Σ 产生决策蕴涵集 $L = \{\varphi_u \rightarrow \psi_u | u \in \Sigma\}$, 再运用 MP 近似推理算法 1 获取 Δ 上的拟决策蕴涵 (以下称这种获取方式为增量获取方式). 算法效率是通过 L 中的决策蕴涵后件对 MP 近似推理结论的平均支持度 $\tau_{\Delta}(\psi_1 \wedge \dots \wedge \psi_n \rightarrow \psi_{\Delta}^*)$ 以及获取 Δ 上拟决策蕴涵所用时间成本来刻画. 采用 2 种不同方式获取拟决策蕴涵时, 拟决策蕴涵集 L 的后件对 MP 近似推理结论的平均支持度 $\tau_{\Delta}(\psi_1 \wedge \dots \wedge \psi_n \rightarrow \psi_{\Delta}^*)$ 和时间成本对比如表 7 所示.

由表 7 可以看出, 对给定的基础数据集 Σ , 随着数据集 Δ 的数据个数的增加, 决策蕴涵集 L 的后件公式集对 MP 近似推理结论的支持度会减少, 其变化幅度不是很显著, 当 MP 近似推理的参数数据集 Σ 与 Δ 的数据个数比 $|\Sigma|/|\Delta|$ 较大时, 实际支持度 τ_{Δ} 比 $|\Sigma|/|\Delta|$ 要小很多, 定理 9 中的比值 $|\Sigma|/|\Delta|$ 只是支持度 τ_{Δ} 的下界. 2 种方式总时间成本随着实验数据的增加而增加, 增加的趋势基本上是线性的. 与方式 1) 相比, 方式 2) 获取决策蕴涵的总时间成本可能要多些. 但如果是在已有 Σ 上的决策蕴涵的基础上来获取 Δ 上的新增决策蕴涵, 方式 2) 充分利用了 Σ 上的已有决策蕴涵, 减少了获取决策蕴涵 $L = \{\varphi_u \rightarrow \psi_u | u \in \Sigma\}$ 的时间成本, 扣除获取 Σ 上的决策蕴涵的时间成本后, 增量获取决策蕴涵所用时间成本就要少很多.

方式 2) 是基于已有决策蕴涵借助近似推理模式产生新的更多决策蕴涵, 体现了人工智能时代通过学习来获取新知识的思想方法, 是知识推理的有效手段. 但如何评价近似推理好坏, 目前还没有公认的标准. 本文引入的“决策蕴涵后件对 MP 近似推理结论的平均支持度”和“决策蕴涵前件对 MT 近似推理结论的伪距离”都是通过公式的真度来定义的. 按照文献 [18] 的支持度理论, 一个由条件联结词连接的蕴涵公式的真度本质上刻画了前件公式对后件公式的支持度, 上述评价标准实际上刻画了推理大前提中的知识对近似推理得到的结论知识的

表 7 后件集对结论的支持度和获取拟决策蕴涵时间成本对比

Table 7 Comparison of the support degree of consequent set to the conclusion and time consumption of obtaining quasi-decision implications

数据组别	L 的后件集对结论支持度 τ_{Δ}	方式 1) 总时间成本 (s)	方式 2) 总时间成本 (s)	增量获取所用时间成本 (s)
数据组 1	0.933	0.204	0.223	0.103
数据组 2	0.923	0.283	0.304	0.147
数据组 3	0.914	0.361	0.389	0.198
数据组 4	0.903	0.450	0.484	0.258
数据组 5	0.893	0.539	0.258	0.326

支持程度, 因此, 它们应是一个比较有效的知识推理评价标准. 由实验 1 ~ 实验 3 可知, 本文提出的应用 MP 和 MT 近似推理生成拟决策蕴涵的方法具有较好效果.

本文提出的近似推理方法的一个应用是针对有缺失值的不完全决策形式背景的属性值补齐. 这时对象范围参数选择为包含缺失值的对象集, 对象范围选择不包含缺失值的对象集, 进而根据缺失值是条件属性值还是决策属性值, 选择应用 MP 或 MT 近似推理来补齐有缺失值的不完全决策形式背景, 进而实现不完全决策形式背景上的决策蕴涵和概念规则等知识的获取.

6 结束语

本文从数据分析需求的角度讨论决策蕴涵, 利用属性逻辑的 K -赋值对决策蕴涵进行属性逻辑语义刻画, 在此基础上研究决策蕴涵、概念规则的近似推理. 分析经典二值逻辑框架下基于全蕴涵三 I 推理思想的决策蕴涵和概念规则的 MP、MT 近似推理模式的特征, 证明决策蕴涵的 MP、MT 近似推理结论是决策蕴涵, 概念规则的 MP、MT 近似推理结论是概念规则, 引进属性逻辑伪距离, 分析人工智能环境下通过 MP、MT 近似推理生成拟决策蕴涵的模式和效果. 提出若干基于现有决策蕴涵、概念规则通过 MP、MT 近似推理产生新的决策蕴涵、概念规则、拟决策蕴涵的模式和方法, 并通过实际例子和数值实验来说明这些模式的应用和效果. 人工智能环境下的机器学习是一种基于现有知识通过一定的规则获取新知识的过程. 本文提出的基于现有决策蕴涵、概念规则通过近似推理来产生新的决策蕴涵及概念规则的方法正是依据这个思想来进行的. 本文提出的方法是一种获取决策蕴涵、概念规则的新方法, 为人工智能、大数据环境下的知识发现、知识推理提供了新的途径. 未来工作包括以下 2 个方面:

1) 决策蕴涵和概念规则的近似推理机制的逻辑语构方面也是值得探讨的.

2) 将本文提出的形式背景上的近似推理模式推广到模糊形式背景, 研究模糊形式背景中模糊决策蕴涵和模糊概念规则的近似推理机制及方法.

References

- Zhang Wen-Xiu, Liang Yi, Wu Wei-Zhi. *Information Systems and Knowledge Discoveries*. Beijing: Science Press, 2003. (张文修, 梁怡, 吴伟志. 信息系统与知识发现. 北京: 科学出版社, 2003.)
- Li Jin-Hai, Wei Ling, Zhang Zhuo, Zhai Yan-Hui, Zhang Tao, Zhi Hui-Lai, et al. Concept lattice theory and method and their research prospect. *Pattern Recognition and Artificial Intelligence*, 2020, **33**(7): 619–642
- (李金海, 魏玲, 张卓, 翟岩慧, 张涛, 智慧来, 等. 概念格理论与方法及其研究展望. 模式识别与人工智能, 2020, **33**(7): 619–642)
- Ganter B, Wille R. *Formal Concept Analysis Mathematical Foundations 3*. Berlin: Germany Springer-Verlag, 1999.
- Li J H, Mei C L, Kumar C A, Zhang X. On rule acquisition in decision formal contexts. *International Journal of Machine Learning and Cybernetics*, 2013, **4**(6): 721–731
- Zhai Y H, Li D Y, Qu K S. Decision implications: A logical point of view. *International Journal of Machine Learning and Cybernetics*, 2014, **5**: 509–516
- Zhai Y H, Li D Y, Qu K S. Fuzzy decision implications. *Knowledge-Based Systems*, 2013, **37**: 230–236
- Jia Nan, Zhang Shao-Xia, Zhai Yan-Hui, Li De-Yu. Study of inference rules and deduction process on decision implications. *Journal of Frontiers of Computer Science and Technology*, 2020, **14**(2): 344–352
- (贾楠, 张少霞, 翟岩慧, 李德玉. 决策蕴涵上的推理规则和推理过程研究. 计算机科学与探索, 2020, **14**(2): 344–352)
- Wang Qi, Li De-Yu, Zhai Yan-Hui, Zhang Shao-Xia. Parameterized fuzzy decision implication. *Journal of Computer Research and Development*, 2022, **59**(9): 2066–2074
- (王琪, 李德玉, 翟岩慧, 张少霞. 含参决策蕴涵. 计算机研究与发展, 2022, **59**(9): 2066–2074)
- Zhang Shao-Xia, Li De-Yu, Zhai Yan-Hui. Logic-type decision implication. *Journal of Zhengzhou University (Natural Science)*, 2020, **52**(1): 1–7
- (张少霞, 李德玉, 翟岩慧. 逻辑型决策蕴涵. 郑州大学学报(理学版), 2020, **52**(1): 1–7)
- Zhang S X, Li D Y, Zhai Y H, Kang X P. A comparative study of decision implication, concept rule and granule. *Information Sciences*, 2020, **508**: 33–49
- Li Jin-Hai, Li Yu-Fei, Mi Yun-Long, Wu Wei-Zhi. Meso-granularity labeled method for multi-granularity formal concept analysis. *Journal of Computer Research and Development*, 2020, **57**(2): 447–458
- (李金海, 李玉斐, 米允龙, 吴伟志. 多粒度形式概念分析的介粒度标记方法. 计算机研究与发展, 2020, **57**(2): 447–458)
- Li Jin-Hai, Wang Fei, Wu Wei-Zhi, Xu Wei-Hua, Yang Xi-Bei, She Yan-Hong. Review of multi-granularity data analysis methods based on granular computing. *Journal of Data Acquisition and Processing*, 2021, **36**(3): 418–435
- (李金海, 王飞, 吴伟志, 徐伟华, 杨习贝, 折延宏. 基于粒计算的多粒度数据分析方法综述. 数据采集与处理, 2021, **36**(3): 418–435)
- Wang G J, Zhou H J. Quantitative logic. *Information Science*, 2009, **179**: 226–247
- Linden G, Smith B, York J. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 2003, **7**(1): 76–80
- Freyer D A, Hsieh Y H, Levin S R, Pines J M, Mareiniss D P, Mohareb A, et al. Google flu trends: Correlation with emergency department influenza rates and crowding metrics. *Clinical Infectious Diseases*, 2012, **54**(4): 463–469
- Liang Ji-Ye, Feng Chen-Jiao, Song Peng. A survey on correlation analysis of big data. *Chinese Journal of Computers*, 2016, **39**(1): 1–18
- (梁吉业, 冯晨娇, 宋鹏. 大数据相关分析综述. 计算机学报, 2016, **39**(1): 1–18)
- Zhang Jia-Lu, Wu Xia. Study on the approximate reasoning models based on the attribute logic. *Journal of Frontiers of Computer Science and Technology*, 2023, **17**(4): 848–856
- (张家录, 吴霞. 属性逻辑的近似推理模式研究. 计算机科学与探索, 2023, **17**(4): 848–856)
- Wang Guo-Jun. *Non-classical Mathematical Logic and Approximate Reasoning*. Beijing: Science Press, 2000.
- (王国俊. 非经典数理逻辑与近似推理. 北京: 科学出版社, 2000.)
- Wang Guo-Jun. The triple I algorithm for all implication of fuzzy reasoning. *Science in China (Series E)*, 1999, **29**(1): 43–53

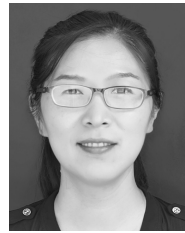
- (王国俊. 模糊推理的全蕴涵三 I 算法. 中国科学 (E 辑), 1999, **29**(1): 43–53)
- 20 Zhang Wen-Xiu, Wei Ling, Qi Jian-Jun. Attribute reduction theory and approach to concept lattice. *Science in China (Series E)*, 2005, **35**(6): 628–639
(张文修, 魏玲, 祁建军. 概念格的属性约简理论与方法. 中国科学 (E 辑), 2005, **35**(6): 628–639)
- 21 Luo Jia-Qi, Wu Xia, Zhang Jia-Lu, Du Jia-Tian. The description of object granules by using attribute logic formulas based on formal concept. *Fuzzy Systems and Mathematics*, 2020, **34**(1): 41–48
(罗佳琪, 吴霞, 张家录, 杜佳甜. 基于形式概念的对象粒的属性公式描述. 模糊系统与数学, 2020, **34**(1): 41–48)
- 22 Zhang J L, Wu X, Lu R H. Decision analysis methods combining quantitative logic and fuzzy soft sets. *International Journal of Fuzzy Systems*, 2020, **22**(6): 1801–1814
- 23 Wu Xia, Zhang Jia-Lu. Attribute logical formulas description of object granules based on the property-oriented (object-oriented) concepts. *Pattern Recognition and Artificial Intelligence*, 2020, **33**(9): 767–775
(吴霞, 张家录. 基于面向属性 (对象) 概念的对象粒的属性公式描述. 模式识别与人工智能, 2020, **33**(9): 767–775)
- 24 Zhang Jia-Lu, Chen Xue-Gang, Zhao Xiao-Dong. Theory of probability semantics of classical propositional logic and its application. *Chinese Journal of Computers*, 2014, **37**(8): 1775–1785
(张家录, 陈雪刚, 赵晓东. 经典命题逻辑的概率语义及其应用. 计算机学报, 2014, **37**(8): 1775–1785)
- 25 Mi Y L, Liu W Q, Shi Y, Li J H. Semi-supervised concept learning by concept-cognitive learning and concept space. *IEEE Transactions on Knowledge and Data Engineering*, 2022, **34**(5): 2429–2442



张家录 湘南学院数学与信息科学学院教授. 主要研究方向为智能信息处理, 知识发现和非经典数理逻辑与近似推理. E-mail: zjl0735@163.com

(ZHANG Jia-Lu Professor at the College of Mathematics and Information Science, Xiangnan University.

His research interest covers intelligent information processing, knowledge discovery, and nonclassical mathematical logic and approximate reasoning.)



吴霞 湘南学院数学与信息科学学院教授. 主要研究方向为智能信息处理, 非经典数理逻辑与近似推理. 本文通信作者.

E-mail: wuxia351@163.com

(WU Xia Professor at the College of Mathematics and Information

Science, Xiangnan University. Her research interest covers intelligent information processing, nonclassical mathematical logic and approximate reasoning. Corresponding author of this paper.)