

基于非单调推理的领域专家知识库的研究^{*}

The Research of Field Expert Knowledge Base Based on Nonmonotony Reasoning

夏幼明¹ 夏幼安² 徐天伟¹ 赵景秀¹

(云南师范大学计算机科学系 昆明650092)¹(郑州电缆厂职工大学 郑州450006)²

Abstract At first, this paper introduces the last results of knowledge presenting and reasoning, basic conception and research methods of nonmonotony. Then, we enlarge ability to express knowledge in semantic network language, and put forward knowledge presenting way of modal logic. Finally, we provide arithmetic of modal logic reason.

Keywords Semantic network, Presenting knowledge, Nonmonotony reason, Modal logic

1. 知识表示及推理概况

知识库在人工智能研究中占有重要地位。Ginsberg 等人介绍了多种知识库更新方法, 其中一类是基于模型的方法, 一类是基于公式的方法, 但所有这些方法在通常情况下都是难解的 (intractable)^[1]. 对基于模型的方法, Eiter 指出, 当知识库是 Horn 公式的集合, 新知识是 Horn 公式且长度有一常数上界时, 存在多项式时间算法^[2]. 马绍汉提出一类树型知识库的更新算法, 在树型约束图智能方式表示下, 该算法通过一个自底向上的过程, 得到更新后的知识库^[3].

确定推理的研究是基于经典的二值逻辑, 因而, 具有相当牢固的基础. 而不确定推理的研究却是另一种情形. 不确定推理的基础比确定推理的基础要弱得多. 到目前为止, 甚至还没有获得人们普遍接受的基础^[4]. 不确定推理面对下面几个问题: (1) 如何正确地、定性地刻画一个命题的不确定性? (2) 如何适当地度量一个命题的不确定性?

Silberschatz 指出, 对知识的合法、新奇、有用和可理解性的综合度量是采用对规则的关注的程度^[5]. 对规则的关注的程度包含对规则的客观的关注的程度和对规则的主观的关注的程度两个方面. 对规则的客观的关注的程度称为客观关注程度, 对规则主观关注的程度称为主观关注程度. 目前, 对规则的关注的程度的研究主要是对规则的客观关注程度的研究. 主要包括对于规则 $A \Rightarrow B$, Agrawal^[6] 提出了置信度 $P(B/A)$, Piatetsky-Shapiro^[7] 提出了事件独立性 $P(B|A)/P(a)P(B)$, Symth^[8] 提出了 J-Measure 函数等. 程继华^[9] 提出一种关注规则的多策略方法以解决对规则主观关注程度的度量.

Roberts, Goldstrein 以及 Fahlman 在对多语义网络的继承问题中提出了一种研究方法^[9], 而马殿富提出了一种非单调继承方法研究多语义网络继承问题^[10], 这很相似于以前的 Thouratzky^[11], Horty^[12], Etherington^[13] 等提出的各种方法. Horty 的方法虽然能够产生唯一扩张, 但是存在不稳定性问题. Etherington 方法通过将语义网络的链翻译为中间语言——缺省逻辑, 对语义网络的链 (Link) 给出统一的定义. 这种研究语义网络的方法针对同一语义网络能产生多个扩张, 其主要思想为: 提出的应用于非循环语义网络图的可废除继承的分析方法, 能够产生适合于继承直觉的非多义性结论, 基于信任理论的继承推理与以前的继承推理相比较不总是一致. 这个不同的推理策略自然产生不同的结果. 此方法是采用语义网络的术语直接建立继承理论. 这个不同于非单调继承逻辑.

2. 非单调推理

在经典的逻辑系统中, 以一个无矛盾的公理系统为基础, 每当加入新的事实, 往往能推出新的结论, 而且总是保持原有的结论不变. 这种推理称为单调的, 当新的事实被推翻时, 原有推理称为非单调的.

非单调推理的三种方式: 设 S 为命题, (1) 限定理论: 当且仅当没有事实证明 S 在更大范围内存在时, S 只在指定范围内成立. (2) 缺省理论: S 在缺省条件下成立是指当且仅当没有事实证明 S 不成立时 S 是成立的, 产生这一假设的背景是: 在非单调推理中, 知识库不够丰富, 难以支持人或系统所需的推理, 因此需要想当然地对知识库进行扩张, 添加新的内容, 扩张的内容即是缺省知识. (3) 自认识理论: 如果我知道 S 并且我不知道任何事实与 S 矛盾, 则 S 是成立的. 换句话说,

^{*} 本课题得到云南省自然科学基金资助 (2000F0049M). 夏幼明 教授, 工学硕士, 现从事人工智能、数据库等领域研究.

如果推理者并不相信命题 S 的非,则 S 和推理者的当前知识是一致的。

3. 扩张语义网络的知识表示能力

在人工智能的研究领域中,知识表示是研究的主要内容之一;其中,语义网络作为知识表示的主要方法引起人们广泛的重视^[4]。本文是在文[15]提供的 SnetL 语言形式化描述的基础上给出新的研究结果,即,对 SnetL 语言进行扩充,用于构建领域专家知识库以及 Agent 常识库。在文[16]中,给出了将 SnetL 语言描述的知识转换成逻辑谓词形式的表示以及转换的规则和方法。

定义1 将形如 $(L_1; L_2; \dots; L_n; N_1, T_1; T_2; \dots; T_l; R, M_1; M_2; \dots; M_t; N_2)$ 的表示称为三元组。其中, $L_1, L_2, \dots, L_n, T_1, T_2, \dots, T_l, M_1, M_2, \dots, M_t$ 称为标号, N_1, N_2 为结点, R 为关系。当 $n \neq 0$ 时,称 $L_i (i=1, 2, \dots, n)$ 为对结点 N_1 的修饰; $t \neq 0$ 时,称 $M_j (j=1, 2, \dots, t)$ 为对结点 N_2 的修饰; $l \neq 0$ 时,称 $T_s (s=1, 2, \dots, l)$ 为对结点 R 的修饰。

定义2 无前缀标号的三元组称为主三元组;含前缀标号的三元组称为标号三元组。

定义3 标号作为前缀出现在三元组前面时,成为标号定义。当标号在三元组内部出现时,称为标号使用。

定义4 一个标号三元组如果仅有标号定义而没有标号使用时则称这个三元组为无嵌套标号三元组;否则,称为嵌套标号三元组。

定义5 由若干三元组表示的每一句话,构成一个语义子网。

例如:用语义网络语言表示:小王拿了朵红花。

(小王,拿, L_1 :花)
 L_1 :(颜色(花),is,红)

我们将 L_1 所确定的标号三元组中的结点:颜色(花)称为函数结点;该标号三元组的语义为:颜色(花)是红的,即红花。

我们对语义网络语言 SnetL 扩展内容如下:

·语义说明表示

〈函数节点〉::= $\langle$ 函数名 $\rangle\langle$ 节点 $\rangle$

〈函数名〉::= $\langle$ 标识符 $\rangle$

〈固有关系〉::= imply (蕴涵)| compare (比较)| similar_to (相似)| relation (关系)| belong (属于)| object (对象)| more_than (多于)| less_than (少于)| logic (逻辑)| action (行为)| be (是)| have (有)| position (位置)| quality (质量)| attribute (属性)| property (性质)| degree (程度)| number (数量)| Timply (时间蕴涵)

为了增加语义描述的功能,增加了关系类型句,其语法定义为:

〈关系类型句〉::= $\langle$ 关系名 $\rangle, \text{TYPE}$ |类型,〈关系类型名 $\rangle$

〈关系类型名〉::= imply (蕴涵)| compare (比较)| similar_tp (相似)| relation (关系)| belong (属于)| object (对象)| logic (逻辑)| action (行为)| be (是)| have (有)| position (位置)| quality (质量)| attribute (属性)| property (性质)| degree (程度)| exception (除外)

exception 类型的关系用于模态逻辑的表示。例如,除了鸵鸟之外,鸟会飞。

$(L_1, \text{exception}, L_2);$
 $L_1: (L_3, \text{imply}, L_4);$
 $L_3: (?x, \text{one_of}, \text{鸟});$
 $L_4: (?x, \text{fly},);$
 $L_2: (\text{鸵鸟}, \text{subclass}, \text{鸟});$
其中, $?x$ 表示变量。

·标准关系表示 用 one_of 、 subset 、 subclass 、 part_of 和 be 等系统保留字分别定义节点之间的成员、子集、子类、部分,是等标准概念关系以及关系之间的标准概念关系。

〈 one_of 关系式〉::= $\langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}, \text{one_of}, \langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}$

〈 subset 关系式〉::= $\langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}, \text{subset}, \langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}$

〈 subclass 关系式〉::= $\langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}, \text{subclass}, \langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}$

〈 part_of 关系式〉::= $\langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}, \text{part_of}, \langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}$

〈 be 关系式〉::= $\langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}, \text{be}, \langle$ 限定节点 $\rangle\{; \langle$ 标号 $\rangle\}$

上述五个关系式中,第一个节点和第二个节点必须同时为〈简单节点〉或〈关系名〉,换句话说,前后要一致。

·修饰关系表示

〈修饰关系〉::= time (时间)| before (以前)| start_time (开始时间)| end_time (结束时间)| way (方式)| place (地点)| reason (原因)| possible_reason (可能原因)| result (结果)| goal (目的)| degree (程度)

·量词关系表示 修饰关系通常用于对关系的进一步说明。

〈量词关系〉::= every (每个)| exist (存在)| some (某些)| only (唯一)| number (n个)| most (大多数)

most 用于模态逻辑知识的表示,在下面将给出其应用。

·关系性质表示 为使语义网络表示的一致性,对原有的标准关系库和关系性质库进行了修改,统一由三元组的形式表示出来。

〈关系性质关系式〉::= $\langle$ 关系名 $\rangle, \text{have}$ |具有,〈性质

1)) | (<关系名>, <性质关系>, <关系名>)
 <性质1> ::= translation(传递) | symmetry(对称)
 <性质关系> ::= inverse(逆关系) | equal(等价关系) |
 negation(否定关系)

• 模态逻辑的知识表示

① 多数蕴涵关系: (A, Mimply, B) , 其中, Mimply 是一种多数蕴涵的知识表示, 其语义为: A 成立, 多数的 B 成立, 那么, B 成立。

② 例外蕴涵关系: $(A, \text{Eimply}, (C, B))$, 其中, Eimply 是一种例外蕴涵的知识表示, 其语义为: A 成立, 除 C 之外 B 成立, 那么, B 成立。

③ 依赖时间蕴涵关系: $(A, L_1; L_2; R, B)$,

$L_1: (R, \text{start_time}, t_1);$

$L_2: (R, \text{end_time}, t_2);$

$(R, \text{TYPE}, \text{Timply})$

其中, Timply 是一种依赖时间蕴涵的知识表示, 其语义为: 由三元组 (A, R, B) 表示的命题, 在时间段 t_1 到 t_2 之间为真。

4. 模态逻辑推理的算法设计基本思想

在多数蕴涵关系表示的知识 (A, Mimply, B) 中, 进行多数蕴涵推理 M 由下列步骤进行:

M1 检查 A 在知识库中是否成立, 若无法得到 A 成立的事实, 则, 不能确认 B 成立, 即, 不做进一步推理;

M2 确定多数 B 成立的含义, 这里有两种方法解决此问题: 方法一: 在知识库中检查是否有 most B 的事实, 若有则 B 成立; 方法二: B 和非 B(B 的否定) 成立的比例数是否达到领域知识库中给定的数据, 若达到则 B 成立。

在对例外蕴涵的知识表示 $(A, \text{Eimply}, (C, B))$ 的例外蕴涵推理 E, 其方法为:

E1 检查 A 在知识库中是否成立, 若无法得到 A 成立的事实, 则, 不能确认 B 成立, 即, 不做进一步推理;

E2 在知识库中检查有关关系类型是 exception 的事实, 在这些事实中是否有形如 (B, R, C) 的事实。若有则 B 成立; 否则, B 不成立, 其中, R 的类型是 exception。

在对依赖时间蕴涵的知识表示的推理中, 其困难是如何确定时间。因此, 要求给出对时间的判断。因此, 如果在知识库中已知有:

$(A, L_1; L_2; R, B)$,

$L_1: (R, \text{start_time}, t_1);$

$L_2: (R, \text{end_time}, t_2);$

$(R, \text{TYPE}, \text{Timply})$

依赖时间蕴涵推理 T 为:

• 108 •

T1 检查 A 在知识库中是否成立, 若无法得到 A 成立的事实, 则, 不能确认 B 成立, 即, 不做进一步推理;

T2 对 B 检查是否有对时间的说明, 如果没有则无法判断 B 成立, 结束; 否则, 设对 B 说明的时间为 t_3 , 如果 t_3 在 t_1 和 t_2 之间, 则 B 成立; 否则, 无法判断 B 成立, 结束。

在上述的三个算法中, 都要确定 A 是否在数据库中, 这时需要调用下面的查找算法 S。

S1 在数据库中直接检查是否有知识与 A 匹配, 若有, 则成功结束;

S2 检查蕴涵关系中的结论部分, 是否与 A 匹配, 如果有这样的三元组存在, 则设其形式为 $((A_1, \dots, A_n), \text{imply}, A)$, 然后, 分别将 $A_1, \dots, A_n$ 作为查找的子问题, 继续查找, 如果 $A_1, \dots, A_n$ 都查找成功, 则结束。

S3 检查标准关系表示部分, 如果存在这样三元组, 使得 B 作为第一节点出现, 即, (A, θ, C) , 其中, θ 是标准关系 one-of 或 subset 或 subclass 或 part-of 或 be, 那么, 将 C 作为继续查找的对象, 转到 S1。

参考文献

1. Eiter T, Gottlob G. On the complexity of propositional knowledge base revision, update, and counterfactuals. *Artificial Intelligence*, 1992, 57(2~3): 227~270
2. 马绍汉, 陶雪红. 一类树型知识库的更新算法. *软件学报*, 1999, 10(11): 1174~1178
3. Kyburg K. Why do we need foundation or modeling uncertainties. In: Ambrosio D, Smets, Bonissone, eds. *Proc of the Conf. on Uncertainty in AI*. San Mateo, California: Morgan Kaufmann Publishers, 1991. 438~442
4. Silberschatz A, Stonebraker M, Ullman J. What Makes patterns interesting in Knowledge discovery systems. *IEEE Trans on Knowledge and Data Engineering*, 1996, 8(6): 970~974
5. Agrawal R, Srikant R. Fast algorithms for mining association rules. [Tech Rep. RJ9839]. IBM Almaden Research Center, 1994
6. Piatetsky-Shapiro G. Discovery, analysis, and presentation of strong rules. *Advances in Knowledge Discovery and Data Mining*. Menlo Park, California: AAAI/MIT Press, 1991. 229~238
7. Smyth P, Goodman R M. An information theoretic approach to rule induction from database. *IEEE Trans on Knowledge and Data Engineering*, 1992, 4(4): 301~316
8. 程继华. 挖掘所关注规则的多策略方法研究
9. Fahlman S. Design sketch for a million element NETL machine. *AAAI-80*, 240~252
10. 马殿富. 信任理论——一种非单调继承推理方法. *98人工智能进展*. 1999. 45~51
11. Touretzky D. The mathematics of Inheritance System. [S. I.] Morgan kaufmann, 1986
12. Thomason H R, Touretzky D. S Skeptical Theory of inheritance in Nonmonotonic Semantic Network. *AI42*, 1990. 311~348
13. Etherington D W. Formalizing Nonmonotonic Reasoning System. *AI31*, 1987. 41~85
14. 陆汝铃. 人工智能. 科学出版社出版, 1996. 107~149
15. 陆汝铃. 专家系统开发环境. 科学出版社, 1994. 247
16. 夏幼明, 张春霞, 徐天伟, 夏幼安. 语义网络的知识获取及转换的研究. *云南师范大学学报*, 1996, 6. 40~45