

处理不确定性的外延方法和内涵方法的比较^{*}

The Comparison of Extensional Approach and Intensional Approach When Tackling Uncertainty

王 飞 刘大有 王松昕

(吉林大学计算机系 长春130023)

Abstract Extensional approach vs. intensional approach is taxonomy of tackling uncertainty. In this paper, we compare these two approaches. Extensional systems are computational convenient but semantically sloppy, while intensional systems are semantically clear but computational clumsy. The trade-off between semantic clarity and computational efficiency has been the main issue of concern.

Keywords Extensional approach, Intensional approach, Bayesian networks

1. 引 言

从语义的角度分类,研究不确定性的方法分为两种:外延方法和内涵方法。前者也称为产生式系统、基于规则的系统、基于过程的系统。MYCIN, PROSPECTOR 是典型的外延方法。外延方法把不确定性作为一般化(generalized)的真值附着在公式上,依照传统的经典逻辑,任何公式的不确定性都是其子公式不确定性的函数,因此公式间的联系就是选择合适的组合函数。例如,合取式 $A \wedge B$ 的不确定性值是 A 的不确定性值和 B 的不确定性值的函数,如 $\min$ 函数。

内涵方法也称为陈述系统(declarative system)或基于模型的系统。不确定性附着于事件的状态集或可能世界的子集,子集间的联系通过集合上的操作完成。例如, $p(A \wedge B)$ 是包含 A 为真和 B 为真的可能世界子集的概率值,不能从个体概率 $p(A)$ 和 $p(B)$ 中得出。

外延方法和内涵方法的优缺点针锋相对,外延方法计算简便,但语义不系统(sloppy);内涵方法语义清晰,但计算复杂。

2. 处理不确定性的外延方法

2.1 外延方法的优点

外延方法具有计算模块性,具体表现在局域性(locality)和隔离性(detachment)。局域性是指不考虑知识库任何其它的事实;隔离性是指不考虑某个事实是如何得出的。规则 $A \xrightarrow{x} B$ 可以理解为:一旦 A 的不确定性经历了变化 ΔA ,则不管事实库中包含有什么其它事实,也不管 ΔA 是如何得到的,以 ΔB 更新 B

的不确定性, ΔB 是 $x, \Delta A$ 及 B 当前不确定性的函数。

在内涵方法中,即便有条件概率表 $p(B|A)=p$ 及 A 为真,也不能做出任何关于 B 或者 $p(B)$ 的断言。因为 $p(B|A)=p$ 的含义是:如果 A 为真,并且 A 是事实库中唯一已知事实,那么 B 发生的概率为 p 。一旦事实库中出现事实 K ,那么断言 $p(B|A)=p$ 的许可被取消,替代以寻找 $p(B|A, K)$ 的概率。我们对这个条件概率一无所知,不能开始任何运算,除非能够证实知识库中除 A 外的所有其它事实均与 B 无关,这就是为什么无关性的证实内涵系统中非常重要的原因。

事实上,无关性的证实外延系统中也非常重要,但外延系统过于追求计算上的简便忽略了无关性的证实,因此带来语义上的缺陷。

2.2 外延方法的缺点

外延系统的计算简便性是以经常产生不一致为代价的,即会产生一些违背常识的结论。主要体现在以下几个方面:①双向推理的不恰当处理;②结论缩回的困难(retracting conclusion);③有相关来源的证据的不恰当使用。下面一一加以说明。

2.2.1 双向推理的作用 这里的双向推理不同于 MYCIN 中的双向推理,不是指数据流驱动推理和目标驱动推理,而是指使用诊断知识和预言知识进行的推理,简称诊断推理和预言推理。一条规则“如果发生火灾,则会有烟雾”,从“火灾”到“烟雾”是诊断推理;有“烟雾”则可能发生了“火灾”是预言推理。

使用预言性知识和诊断性知识综合进行推理的能力在不确定性推理中是非常重要的,这些知识的不正确使用会产生非常奇怪的结果。具有一般性的论述模

^{*} 本项研究得到高等学校博士学科点专项科研基金资助课题“处理不确定性信息的内涵方法之研究”和教育部符号计算与识工程重点实验室的资助。王 飞 博士生,刘大有 教授,博士导师,王松昕 博士生。

式溯因推理(abductive)——如果 A 蕴含 B,那么发现 B 为真使得 A 更可信。这个模式涉及两个方向的推理:从 A 到 B 的诊断推理和从 B 到 A 的预言推理。一般说来,人们不需要以两条独立的规则来进行溯因推理,即第一条规则(例如,“起火”意味着“有烟”)激发第二条规则(“有烟”使得“起火”更可信)。但是在外延方法中,第二条规则必须显式地表达出来,否则无法执行溯因推理;更糟糕的是,一旦引入第二条规则,第一条规则必须被移走,否则会产生循环,任何证据 A 的轻微支持经过 B 的放大反馈回 A,会很快产生关于 A 和 B 的强有力断言。一种比较流行的解决方法(例如 MYCIN)切断这种类型的循环,只允许诊断推理,没有预言推理。

缺乏预言推理会使系统不能进行另外一种非常重要的不确定性推理,我们称之为“解释排除”(explain away):即如果已知 A 蕴含 B、C 蕴含 B 和 B 为真,那么一旦发现 C 为真将使 A 更不可信。换句话说,发现对一个事实的第二种解释使得第一种解释更不可信。这样的多原因之间的交互作用出现于很多应用中。例如,盗窃和地震都会引起警报,发现警报是由地震引起的将使盗窃发生的可信度下降;再比如,灯泡发出红光使得看起来是红颜色的手套真的是红颜色的可能性减小。

为了实现“解释排除”推理,系统必须能够进行从证据到假说的诊断推理和从假说到证据的预言推理。有时可能有一种强制手段(如,枚举所有例外情况)在没有循环推理的情况下表述“解释排除”,但任何这样做的系统都是以牺牲模块性为代价的,即牺牲了局域性和隔离性。更确切地说,任何在自然规则基础上进行模块化推理并且平等地看待所有规则的系统必然不能实现某些重要的不确定性推理模式。

2.2.2 结论撤回或撤消的困难 局域性原理被充分应用于经典逻辑的推理规则中。规则“如果 P 则 Q”的含义是:如果 P 为真,即使事实库中含有其它事实,我们也可以在不做任何进一步分析的基础上作出断言 Q。然而,在不确定性推理中,这种忽略事实库中的其它事实的做法是不允许的。例如,假设我们有规则 $R = \text{“地面湿,则天下雨了(规则强度为 C)”}$ 。事实库中存在“地面湿”的事实不见得能使我们增加“天下雨了”的确定性,因为事实库中可能包含有“刚刚有洒水车经过”这样的证据,这种起到反对作用的证据有时是很容易发现的,例如“邻居家的草坪是干的”,这个事实直接反对“天下雨了”;有时这种证据隐匿在语法描述能力不足的背后,比如,尽管描述“洒水车经过”的证据不直接支持或反对“天下雨了”的可能性,但事实上它削弱了规则 R。这种削弱在基于规则的系统不能体现,一

旦规则 R 被激发,关于“天下雨了”的可信度的增加不能被撤回,因为“洒水车经过”和“天下雨了”之间不存在直接联系。假如规则 R 声明为“地面湿 $\wedge$ 洒水车没有经过 $\rightarrow$ 天下雨了”,把“洒水车”和“天下雨了”联系起来,这样做,规则的设计者必须考虑所有事实间的所有关系,收集整理大量的相互联系但联系不紧密的事实,一方面难以完备收集不紧密的联系,另一方面损坏了外延系统的模块性,而贝叶斯网建造时只需考虑事实间的直接因果关系,不考虑繁琐的事实间不紧密联系的同时利用网络结构体现出所有事实间的所有关系。

下面举例说明隔离性的缺陷。在演绎推理中,如果 K 蕴含 P 且 P 蕴含 Q,从推理链 $K \rightarrow P \rightarrow Q$,可以由 K 为真得到 Q 为真。推理链不适用于不确定性推理。系统含有两条分别看上去正确的规则“洒水车经过则地面是湿的”,“地面是湿的说明天下雨了”。显然,发现“洒水车经过”的事实不但不能得出“天下雨了”的结论,相反却削弱了对“天下雨了”的支持。

考虑另外一个例子,如图1所示。正常情况下,“警报”使我们考虑发生“盗窃”的可能性。如果你正在上班,邻居打电话告诉你家的警报器响了,你可能马上冲回家,尽管知道可能有其它原因使得警报器响了。但是如果听到电台报告说附近发生了地震,你会认为家里的警报是地震引起的,这时认为家中发生盗窃的可能性消失了。这样的推理过程是按下面的流程进行的:先从结果到原因(电台预报 $\rightarrow$ 地震),然后从原因到结果(地震 $\rightarrow$ 警报),再从结果到原因(警报 $\rightarrow$ 盗窃)。注意,这个推理模式的流程是:我们有规则“如果警报响则发生了盗窃”,听到广播,警报响变得更可信,而盗窃更不可信。简而言之,我们有规则“ $A \rightarrow B$ ”,并且 A 更可信时, B 更不可信。这样的情况和我们局域性度更新的结果相反。

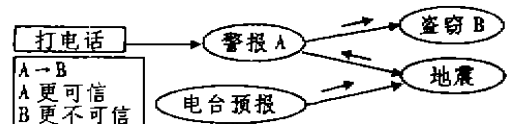


图1

2.2.3 有相关来源的证据(非独立证据) 外延系统具有局域性和隔离性,只对不确定性值(信度)的大小感兴趣,不关心引起其改变的来源。结果,相同或相关的信息可能不止一次地影响同一结论(在一些研究中,已注意避免出现这种情况)。Henrion 一个关于传染性疾病的例子说明在基于规则的系统的局域策略中会遇到的问题。图2说明在一般情况下,多个相互独立的证据分别影响假说的信度,但发现这些证据的来源是相同的应该减少假说的信度。

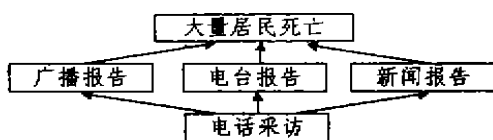


图2 规则不能局部地简单组合

外延系统的局域性太强,不能识别出证据的共同来源,因此,有时把并非是相互独立的证据当作相互独立的证据去更新假说的信度,从而进行一些不正确的推理。

3. 处理不确定性的内涵方法

内涵方法在语义方面非常清晰,因为它有全局一致的描述模型,不存在双向推理、相关证据处理方面的问题,但计算复杂性是其致命的弱点。在较复杂的环境下,内涵方法基本是不可用的。

举个例子形象说明外延方法和内涵方法这两种极端的情况。外延方法:一个人骑着一匹马穿过一片有地雷的区域。他可能选择一匹经过训练的、对地雷比较敏感的马,但是危险是实际存在的,有经验的人(知识工程师)会阻止马跑得太快以避免发生危险;

内涵方法:安全无虞的语义保证每一步前进都不会发生危险,但是前进非常困难,因为在每走一步之前都必须重新检测整个区域。

显然,在计算模块化和语义一致性之间存在矛盾,如何在计算的简便高效和语义的清晰之间选择一种较好的折衷方案是不确定性处理的重点研究内容。近年

来逐渐引起广大人工智能学者极大重视的贝叶斯网就是一种较好兼顾计算简单和语义清晰的方案。从本质上讲,贝叶斯网是一种内涵方法,是全概率分布的紧凑表示形式,由于它采用图形表示方式嵌入了大量的独立关系,使得计算复杂性大为下降,在很大程度上克服了内涵方法在计算方面的困难,使得贝叶斯网是可行的和实用的,逐渐成为不确定性处理的主流方法。

小结 本文比较不确定性处理的外延方法和内涵方法,分别指出了它们各自的优缺点,在实际处理不确定性的相关问题时,可以根据需要在计算简便性和语义清晰性间取舍,贝叶斯网是一种较好兼顾两者的解决方案。

参考文献

- 1 Pearl J. Fusion, propagation, and structuring in belief networks. *Artificial Intelligence*, 1986, 29: 241~288
- 2 Pearl J. Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kaufmann, San Mateo, CA
- 3 Pearl J. Belief networks revisited. *Artificial Intelligence*, 1993, 59: 49~56
- 4 Pearl J. Causal diagrams for empirical research. *Biometrika*, 1995, 82: 669~710
- 5 Cooper G. Computational complexity of probabilistic inference using Bayesian belief networks (Research note). *Artificial Intelligence*, 1990, 42: 393~405
- 6 Friedman N, Yakhini M. On the sample complexity of learning bayesian networks. *UAI'96*
- 7 Charniak E. Bayesian Networks without Tears. In *AAAI'91*
- 8 Neufeld E. Direction in uncertainty reasoning. *The Knowledge Engineering Review*, 1997, 12(4): 413~415

(上接第122页)

性的函数或过程对数。其它相似性度量类似

基于相似性度量我们可以构造出相似候选构件次序表,用户就能够选出匹配度最高的可复用构件,下面给出选取可复用构件的算法

```

Begin
  查找复用成分库
  if 完全相同的匹配 Then 结束
else
  集中相似的构件
  for 每个构件
    计算匹配程序
  end
  排列和挑选最好的
  修改构件
fi
END

```

其中的匹配程度可用相似性度量 $A(C_i, C_j)$ 的值来衡量。

选择相似的构件是一个分类问题,相似性程度依赖于如何组织该集合,可以谨慎地选择相关的属性,将密切相关的构件分成组,并按意义组织它们。

结束语 随着应用系统的日益复杂,软件构件化作为软件系统发展的重要方向,引起了人们的普遍关注,软构件技术将会受到越来越多的重视,我们提出的基于 OWE(Object With Event)软件构件描述方法,以及基于相似性度量的软件选择方法为开发复杂的应用系统提供了一个很好的解决途径,对提高软件生产率、改善软件产品质量、缩短产品交付时间有着十分重要的意义。

参考文献

- 1 郑明春,张家重. 关于软件复用. *计算机科学*, 1994, 21(4)
- 2 Marovac N. Embedded Documentation For Semi-automatic Program Construction and Software Reuse. *ACM Software Engineering Notes*, 1998, 23(3): 70~74
- 3 Whittleben B. Models and Languages for Component Description and Reuse. *ACM SIGSOFT Software Engineering Notes*, 1992, 20(2)
- 4 Castano S, Antonellis V D. Engineering a library of reusable conceptual components. *Information and Software Technology*, 1997, 39: 65~76