

归纳逻辑及其在归纳学习中的应用问题

王雨田*

(中央民族学院物理系, 北京 100081)

INDUCTIVE LOGIC AND THE PROBLEMS OF IT'S APPLICATIONS IN THE INDUCTIVE LEARNING

Wang Yutian

(Department of Physics, Central Institute for Nationalities, Beijing 100081)

Abstract Logicism has it's great significance for AI, but also has it's limit especially in the field of inductive simulation. Modern inductive logic as a kind of non—standard logic should be an important component part in the logical foundations of inductive learning. This is worth to make further study.

摘要 逻辑主义对人工智能具有重大的意义,但也有其局限性(特别是对于归纳的模拟).现代归纳逻辑作为一类非标准逻辑应成为归纳学习逻辑基础的重要组成部分.对此值得进一步加以探讨.

§ 0. 引言

机器学习、不确定推理的研究已取得重大进展,神经网络的研究也提出过一些学习模型.在人工智能的历史过程中,对归纳的模拟研究已提到了前所未有的重要地位,不少学者认为归纳学习已成为机器学习的核心.

知识工程是基于知识的.知识必须经过认识过程(包括学习)才可获取.认识过程只能开源于经验世界.人类通过对现实世界的感知,从中进行概括与归纳,再经过选择才能形成公理,进行演绎推理.可见归纳先于演绎,较演绎更为基本,演绎与归纳不可分. H. A. Simon 曾指出,在求证一个几何时,“得到正确结论的过程是演绎推理过程,而实际的心理过程则是归纳过程”,^[1]这在原则上应为计算机程序所能处理,为逻辑所能陈述.

从逻辑的观点看,演绎逻辑的形式化已甚完善,与物理符号主义是能适应的.人工智能在早期主要从事定理机器证明、问题求解与博弈等,主要是运用演绎逻辑工具.只有在专家

* 本文1991年6月收到.本文是国家高技术863—306—506—4课题资助项目.作者王雨田,教授,主要研究领域为现代逻辑与人工智能、系统科学与哲学.

系统或知识工程提出后,由于知识获取不能从公理出发,只能面向经验世界进行认知与学习,而学习一般是指“人(和计算机)增长知识和改善技能所经由的历程”,^[2]其中必需用到归纳,所以归纳学习才不能不提到机器学习中的核心地位.

但是,我们不能不承认,对归纳的研究要远比演绎更困难.归纳属于复杂性研究的范围,即使在逻辑学领域,不少逻辑学家尚且对其敬而远之,更何况对归纳的模拟,那就势必更为困难.所以,早在 1984 年,R. S. Michalski 在论及归纳学习的基础理论研究时指出:“建立它们在理论上所用的原理还未得到很好的解释,由于缺少通用的术语和适当的形式理论,使得比较不同的学习方法还很困难”,^[3]至今并未完全解决.

本文要着重指出:归纳逻辑特别是现代归纳逻辑应该而且有可能应用于归纳学习,成为其逻辑基础研究中的重要组成部分.

§ 1. 逻辑主义与归纳

近些年来,围绕着人工智能基础问题展开了热烈的讨论,其中的一个重要问题就是逻辑主义的评价与作用问题.无论是逻辑学派或其反对派,对于归纳这类不可靠推理都是感到困惑和棘手的.

无论如何,逻辑主义在人工智能的研究中已经发挥了重大的作用.逻辑主义者 J. McCarthy 首先提出陈述性知识表达的适用性问题.在 Nils J. Nilsson 看来,人工智能科学家必须对世界的直觉与常识的认识给出一个恰当的概念化形式,即由个体、函数和关系所构成的期望模型,并要给出相应的模型论“语义”.这种理论与方法可简称为概念化(Conceptualization).^[4]逻辑学派使用的这种方法所强调的是基于演绎逻辑的可靠性推理和模型论“语义”.一旦涉及大量存在的不可靠推理,尽管他们试图用极小模型、限制(circumscription)、概率等技术来逼近,但毕竟是很不理想的,是困难重重的.由此不难理解,对逻辑主义持反对态度的 L. Birnbaum 据此对概念化、模型论“语义”这类形式化方法采取强烈的批评态度,强调不可靠推理中合理的不一致性,并提出要以功能语义代替模型论“语义”,批评逻辑学派过分偏爱演绎与模型论“语义”,视之为“宝贝”,实际却“带来了反作用,忽视了其它同样重要的方法”.^[5]可见我们既要肯定逻辑主义的功绩与作用,尤需正视其局限性.

正是逻辑主义的这种局限性,使得人工智能科学家至今对机器学习、归纳学习、不确定推理的逻辑基础研究仍很薄弱.例如,R. S. Michalski 虽提到过与归纳有关的“弱蕴涵”(以 $::>$ 表示)和算法的逻辑基础;^[3]M. R. Genasareth 与 N. J. Nilsson 在对归纳的描述时也指出过,^[6]背景理论 Γ 不应逻辑蕴涵数据集 Δ ,即 $\Gamma \not\vdash \Delta$,句子集 Φ 可定义为归纳结论,即:

$$\Gamma \cup \Delta \not\vdash \emptyset \quad \text{iff} \quad (1) \quad \Gamma \cup \Delta \not\vdash \rightarrow \emptyset$$

$$(2) \quad \Gamma \cup \{\emptyset\} \not\vdash \Delta$$

但他们无论是用“ $::>$ ”或“ \vdash ”来表示与归纳有关的蕴涵,从逻辑的角度看,却均未给出尽可能严格的逻辑解释(包括模型论“语义”的解释),也就不足以将这类有关不可靠推理的蕴涵关系建立在坚实的逻辑基础之上.这是一个无法回避的难题.由此可知,在归纳学习逻辑基础的研究中引入归纳逻辑是一项必要的迫切任务.

§ 2. 现代归纳逻辑是一类非标准逻辑

传统的归纳逻辑可追溯到德谟克里特与亚里斯多德的归纳研究,亚氏将归纳推理看作

一种特殊的三段论,在方法上开创了以演绎研究归纳的技术. 培根与穆勒分别以三表法与求因果联系的四法奠定了归纳逻辑的基础. 他们与包括牛顿在内的一些科学家探讨过归纳的理论问题,强调了归纳的发现作用,并明确了枚举(enumeration)、排除(elimination)、类比(analogy)等基本的归纳方法.

由于当时将归纳与因果相联系而涉及到本体论问题,18世纪英国哲学家休谟(D. Hume)提出严谨的逻辑论证:凡企图以过去实例的概括去证明将来的归纳论断,则势必陷入循环论证,故归纳推理的合理性在逻辑上是不可解的. 这就是著名的“休谟问题”.

为了回避“休谟问题”而与科学理论相协调,考虑到归纳推理不保真,莱布尼兹、布尔、贝叶斯等先后提出以概率值作为归纳结论或然性度量的思想. 其后又在数理逻辑与形式系统方法的直接影响下,在本世纪20—50年代前后形成了以概率逻辑为主的现代归纳逻辑,其主要代表人物为 J. M. Keynes, H. Reichenbach 与 R. Carnap 等人,给出过多种不同的半形式化的、形式化的归纳概率逻辑系统. 由此可见,现代归纳逻辑是一类引入概率度量、运用演绎逻辑与形式化方法、以归纳推理为研究对象的非标准(或非经典)逻辑系统,是当代哲理逻辑(Philosophical Logic)的一个重要组成部分. 逻辑主义者承认,“二值逻辑对于表达不确定知识不太适用,可以用概率信息表示. 我们希望逻辑和概率论的适当组合对于解决不确定推理有所帮助”.^[4]现代归纳逻辑与此正是一致的.

根据对概率的不同解释与其度量的不同处理,现代归纳逻辑区分为贝叶斯派和非贝叶斯派. 在贝叶斯派中又分为逻辑贝叶斯派与主观贝叶斯派,由此划分出现代归纳逻辑的三个主要派别,即属于经验主义的频率派、逻辑派与私人主义派. 本文不可能予以评介,请参阅有关专著.^[7]以下仅从人工智能的逻辑主义角度进行一些对比的讨论.

1. 频率派:以 H. Reichenbach 为代表,对概率以相对频率加以解释. 在其系统中给出个体、属性与关系,这相当于给出概念化形式:即以 $A: x_1, x_2, \dots, x_i, \dots$ 表示参考类(reference class),以 $B: y_1, y_2, \dots, y_i, \dots$ 表示属性类. 序列对 x_i, y_i 间为 1-1 对应的有序关系 R_{xy} , 以 $e(x, y)$ 表示 1-1 对应函项,则概率语句表示为:

$$(i) (x_i \in A \xrightarrow[p]{\ } y_i \in B)$$

其中 $\xrightarrow[p]{\ }$ 为概率蕴涵符号, p 为数值算子,即概率的取值,上述概率语句表示对于所有属于参考类 A 的第 i 个个体 x_i 以概率值 p 概率蕴涵所有属于属性集 B 的第 i 个属性 y_i , 即一般地表示对于任意的某个体以一定的概率蕴涵某属性. 在此,概率蕴涵是一类归纳关系. 同时可知,一概率语句是由三元组 $\langle A, B, e(x, y) \rangle$ 所决定,其约束条件很强,故只能是对于现实世界某种归纳过程的近似性期望模型.

现定义命题函项为: $f(x) =_{df} x_i \in A$

$$g(y) =_{df} y_i \in B$$

则可较严格地给出概率语句的一般式:

$$\begin{aligned} (x)(y) [f(x) \cdot e(x, y) \xrightarrow[p]{\ } g(y)] &\Leftrightarrow (i) (x_i \in A \xrightarrow[p]{\ } y_i \in B) \\ &\Leftrightarrow (A \xrightarrow[p]{\ } B) \\ &\Leftrightarrow [P(A, B) = p] \end{aligned}$$

显然,上式表示 A 类个体以概率 p 概率蕴涵 B 类属性. 进一步必需对概率蕴涵给出恰

当的解释.

H. Reichenbach 认为概率是无穷序列的相对频率极限值, 以此给出对概率蕴涵的经验主义解释. 以 $\sum_{i=1}^n (x_i \in A)$ 表示 x_i 当 $i=1, 2, \dots, n$ 时满足 $x_i \in A$ 的数目, 并给出如下定义:

$$N^n(A) = \sum_{i=1}^n (x_i \in A)$$

$$N^n(A, B) = \sum_{i=1}^n (x_i \in A) \cdot (y_i \in B)$$

则相对频率 $F^n(A, B) = N^n(A, B) / N^n(A)$,

显然可得: $[P(A, B) = p] = \lim_{n \rightarrow \infty} F^n(A, B)$.

我们可以用逻辑主义观点将此看作一种语义的解释. 它是建立在枚举法的基础之上的. 用机器归纳的术语来表示, 相对频率在实质上是对所覆盖的正例数的一种度量, 不过在机器归纳中是通过启发式搜索机制实现的, 在频率派归纳逻辑中则是以可重复性的无穷实验序列来描述的.

2. 逻辑派: 这派站在纯逻辑的立场引入贝叶斯型概率, 用演绎逻辑方法处理归纳推理, 很典型地表现出相当于逻辑主义关于概念化与模型论“语义”的特点, 其代表人物为 Carnap, 对归纳逻辑的研究具有深远的影响.

面对现实世界中复杂的归纳过程, Carnap 从纯逻辑、纯演绎的立场上对频率派的经验主义持批评的否定态度, 将频率解释的概率称为概率 2, 而将基于先验主义、具有分析性质的、进行逻辑解释的概率称为概率 1, 亦称确证函数或确证度 (degree of confirmation), 即 C 函数, 表示为: $C(h, e) = p$, 即证据 e 对假设 h 的确证程度, 以概率值 p 表示. 不难看出 C 函数在形式上与 $P[A, B]$ 的相似, 但其语义解释不同, 不仅 h 与 e 均为分析命题, 而且 C 函数作为对归纳蕴涵的刻划在逻辑派中是给出严格的模型论“语义”的.

为了给出这一严格的模型论“语义”, Carnap 不得不用一种极其简单的语言附以严格的约束条件来表达现实世界中复杂的归纳过程. 在逻辑主义看来, Carnap 选择的是一种极为简单的概念化.

Carnap 首先给出演绎的 \mathcal{L} 语言系统, 这是一个十分简单的一阶逻辑系统. 为了简化, 将个体域由无穷降为有穷, 即由 \mathcal{L}_∞ 简化为 \mathcal{L}_n , 再将一元谓词取为 π 个, 则此一阶语言为 \mathcal{L}_n^π . 进一步, 他对谓词的描述范围再加限制, 对关系不加考虑, 仅考虑我们在系统中所需考虑的某性质, 称为 Q 性质, 则用以表示 Q 性质的谓词称为 Q 谓词.

由此可见, \mathcal{L}_n^π 语言所描述的世界是很简单的, 其中只有 n 个个体 a_1, a_2, \dots, a_n , π 个一元谓词 P_1, P_2, \dots, P_π , 且将每个谓词限于描述一种 Q 性质, 而有无此性质又取 2 值, 即要么具有此性质, 要么不具有此性质, 故 Q 谓词只能有 2^π 个. Q 谓词可以用 P 谓词的合取式表示为:

$$\#p_1 \wedge \#p_2 \wedge \dots \wedge \#p_\pi, \#p_i \text{ 取 } p_i \text{ 或 } \neg p_i, i=1, 2, \dots, \pi.$$

例如: $(\neg p_1 \wedge p_2 \wedge \neg p_3 \wedge \dots \wedge p_\pi)(a_1)$ 表示个体 a_1 不具有性质 1 与 3 而只具有 π 个性质中的其他性质.

可见任意不同的 Q 谓词均互斥. 由上述种种规定, 才有可能运用 \mathcal{L}_n^π 语言对所描述的世界给出充分的刻划. 为此, Carnap 对莱布尼兹提出的可能世界 (possible world) 给出一类严

格的形式化描述.

定义:状态描述(state description):对给定个体域的一切可能状态予以完全描述的语句集,亦即对所规定的世界的一种充分描述.以 S_i 表示状态描述,则可定义为:

$$S_i =_{df} Q_{i_1}(a_1) \wedge Q_{i_2}(a_2) \wedge \dots \wedge Q_{i_n}(a_n)$$

其中 Q_{i_j} 为任一 Q 谓词. 状态描述共有 2^n 个. 显然, S_i 也可用二值矩阵表示如下:

	p1	p2	...	pn
a1	}			
a2		1/0		
...				
an				

显然,共有 2^n 个二值矩阵,任一 S_i 均对应其中的一个取值. 且每一状态描述也必然是互斥的.

在上述概念化工作的基础上, Carnap 进一步将可能世界的形式化建构与语句的语形逻辑相联系起来,给出其归纳蕴涵的模型论“语义”. \mathcal{L}^n 系统中的语句(合式公式),其逻辑真值由该语句在其中成立的可能世界数目来确定. 可以证明:

定理:对任一语句 h , 存在 r 个状态描述,使得 $h \Leftrightarrow S_1 \vee S_2 \vee \dots \vee S_r$.

由此可知:确证函数 C 可以由 h 与 e 所共同占有的状态描述数与 e 所占有的状态描述数的比率来确定其值. 为此,要给出测度函数.

定义:正则测度函数 $m(j)$ 是使语句 j 成立的那些状态描述的总和,这是一个数值函数. 当其值取正实数值且其所有状态描述值的总和为 1 时为正则的.

定义:确证函数由正则测度函数定义如下:

$$c(h, e) = m(e \wedge h) / m(e) \quad \text{当 } m(e) \neq 0$$

这就是 Carnap 通过对可能世界的形式化、数量化的处理对确证函数(亦即确证度)所给出的模型论“语义”. 这是纯逻辑地严格给出的,在理论上是很漂亮的.

同时,基于这种“语义”,他严格地将归纳蕴涵与演绎逻辑蕴涵(实质蕴涵)加以区别,将前者称为 \mathcal{L} -一部分蕴涵(partial \mathcal{L} -implication),是以正则测度函数,确证函数给出其“语义”的.

在归纳推理方面,他形式地定义简单定律为带有全称量词而其辖域不带量词的语句,如:

$$(x)(p_1x \rightarrow p_2x)$$

还定义了带有条件的简单定律. 他均用部分蕴涵研究其归纳推理的确证度问题.

3. 私人主义派:以上两派是对归纳的客观研究,而私人主义派则是对其主观因素作用的研究,强调主观信念对概率取值的影响. 早在 1926 年, F. P. Ramsey 已对与归纳推理有关的主观概率进行过研究^[9]. 他提出推理者对某假说的置信程度应与该假说的最高打赌商相等,后者是一种客观的度量. 其后 de Finetti 证明,若一置信函数是合理的(即不是选择必输的荷兰赌),当且仅当它满足概率演算的一般基本公理. 可见这是一种主观贝叶斯派. 可以因人而异,表明推理者主观因素的作用. 不过私人主义者仍力图为主观作用找到适当的客观度量. Carnap, L. J. Savage 与 R. C. Jeffrey 等人均从事过这方面的研究.

在人工智能中运用贝叶斯公式时,考虑到先验概率难以给出等缺点,Duda 等人提出了对主观贝叶斯方法加以改进,于 1978 年应用于专家系统 PROSPECTR 中.对于规则

IF E THEN H

在主观贝叶斯方法中则成为

IF E THEN (LS, LN) H

其中参数 LS、LN 分别为充分性度量与必要性度量.这些参数与先验概率均由领域专家主观给出.这表明在人工智能中对概率的评估也是必需考虑到主观因素作用的.现代归纳逻辑的私人主义,主观贝叶斯派也已有过类似的考虑与研究,是可加以借鉴的.

§ 3. 归纳逻辑应用于归纳学习问题

由上节对现代归纳逻辑的评介与对比分析,可提出如下论点:

1. 归纳逻辑.特别是现代归纳逻辑与归纳学习有着共同的研究对象,在所研究的问题和方法上也有共同或近似之处.将二者结合起来不仅是必要的,而且是可能的.

D. Kirsh 指出:逻辑学派的一个基本前提是可以不考虑智能的“载体”或“物质基础”,亦即可“在与载体无关的层次上”找出一套结构化的概念体系(即概念化),只要这套概念化对智能主体的智能(包括知识)能“具有足够强的解释力”就被证明是站得住的,是可使用的.因此,“概念化并不是指系统(机器)所具有的关于世界的概念,而是系统(机器)的设计者把他自己对世界的认识做到机器里面,从而创造出一个与系统将来真正工作环境(世界)最接近的数学结构,包括一系列对象、函数和关系”,^[10]这套方法建立在控制论的功能模拟的基础之上.现代归纳逻辑作为一类演绎化、形式化、数量化的非标准逻辑是满足上述条件的,它作为逻辑学家对归纳世界的一种认识与解释,并给出一套相应的数学逻辑结构,是能适当用于归纳学习概念化研究的.

机器归纳的研究进展在当前遇到一些难点.与归纳学习相比,归纳逻辑的研究早已取得不少成果,也有过不少困惑与教训,若能引入加以借鉴,是有好处的.在此仅举一例加以说明,在归纳逻辑中已研究过三种悖论,其中 Hemple 悖论又称确证悖论,是由于全称命题与其逆否命题等价而导致的悖论,在归纳学习中,对于正例集与反例集的关系,如果我们不注意到这个逻辑性质,又不加以妥善处理,就会给出某些合乎逻辑但不能使用的结果.因此,在国内外已有一些学者注意到这类问题,并进行结合性的研究,建立了基于归纳逻辑的归纳学习系统.

2. 归纳逻辑至少可以在以下三方面可改进与拓宽归纳学习系统的研究与开发:

- (1)对归纳过程及其机理的理论分析.对归纳推理类型与特征的分析.
- (2)处理归纳过程的形式化方法与工具,尤其是对归纳蕴涵及其语义的处理.
- (3)运用上述方法与工具处理归纳推理的效用评估及所存在的局限性与困难.

以上三方面是由现代归纳逻辑的研究程序所决定的.与研究演绎不同,现实的归纳过程与机理是多种多样的,归纳的原型是复杂的.在建立归纳的形式系统或半形式系统之前,必须对现实世界的归纳原型进行理论分析,找出特点,加以分类.然后才有可能根据某些类型的归纳特点进行形式化描述,采用适当的逻辑与数学工具给出相应的形式化系统.对于这种系统可以适用于哪些类型的归纳,会出现什么问题(包括可能出现的归纳悖论)要在应用过

程中加以评估.原则上,建立通用的归纳机器(归纳系统)是不可能的,建立局部归纳(local induction)的有限归纳机器(适用于某些类型的归纳系统)则是可能的.

现代归纳逻辑的研究程序与逻辑主义方法和程序是有相似之处的.具有归纳智能的“载体”及其机理一般是难以掌握的,因此逻辑主义者一般只是在与载体无关的层次上去找具有足够解释力的概念化系统及其语义,以便进行对归纳的模拟,就算达到了目标.事实上,所谓与载体无关,并不是不对载体的归纳机理与功能进行可能的理论分析,否则是不可能进行概念化的.在这方面归纳逻辑的研究对于归纳学习具有启发与供鉴的意义.当然,这些相似并不排除又有重大的区别,在此不作讨论.

3. 要充分重视现代归纳逻辑的局限性与困难,也要看到逻辑主义对于归纳研究的局限性.它们都有待进一步加以改进与突破.

现代归纳逻辑的诸派在 50—60 年代均暴露出各自的缺点与局限性.事实已经证明,用演绎的、形式系统的、古典概率来处理归纳是存在严重困难的.上述的 Carnap 的 \mathcal{L}_N 系统及其模型论的“语义”,体现了演绎逻辑与概率的结合,在逻辑上虽很漂亮,却没有实用价值,最终以其不可克服的困难而失败.这一教训使我们不能不正视逻辑主义运用于归纳的重大局限性与困难.

从 60—70 年代以来,当代归纳逻辑的研究力图摆脱三派的困境,另辟新径.这对于归纳学习系统的改进与开发,对于逻辑主义的进一步发展也是有意义的.从这些当代的研究进展来看,对于研究具有复杂性特点的归纳与科学发现来看,逻辑是必要而不充分的,现有的逻辑工具是远远不能适应的.只有将逻辑工具与数学工具的改进进一步和认知科学、心理学、系统科学、哲学、科学哲学的研究结合起来,才可能取得重大进展.

参考文献

- 1 H. A. Simon, 司马贺译,人类的认知,科学出版社,1986,171.
- 2 P. R. Cohen and E. A. Feigenbaum, The Handbook of Artificial Intelligence, Vol. III, 1982.
- 3 R. S. Michalski, A Theory and Methodology of Inductive Learning, in Machine Learning, Vol. I, 1984.
- 4 N. J. Nilsson, Logic and Artificial Intelligence, in Artificial Intelligence 47, 1990, 中译文见“人工智能基础”.
- 5 I. Birnbaum, Rigor Mortis: A Response to Nilsson's 'Logic and Artificial Intelligence', in Artificial Intelligence 47, 1991, 中译文见“人工智能基础”.
- 6 M. R. Genesereth and N. J. Nilsson, Logical Foundations of Artificial Intelligence, 1987.
- 7 王雨田,现代逻辑学导引(下册),中国人民大学出版社,1988.
- 8 R. Carnap, Logical Foundations of Probability, 1950.
- 9 F. P. Ramsey, Truth and Probability, 1926.
- 10 H. Kirsh, Foundations of A. I.: the Big Issues, 中译文见“人工智能基础”, 9.