

纯粹归纳逻辑框架下的类比推理

梁贤华

摘要: 纯粹归纳逻辑是卡尔纳普进路归纳逻辑的一种复兴, 其基本的目标是把归纳当作数理逻辑的一个分支来研究。在纯粹归纳逻辑的框架下, 类比推理被建立在一阶逻辑的基础之上, 其理论的核心是用距离函数来刻画相似性, 并且以相关性函数作为推理基础。因此, 在某种意义上, 可以把它看作是实例相关函数的一种延伸。而在实践的意义上, 则反映了人工智能对类比推理形式化的需求。

关键词: 纯粹归纳逻辑; 类比; 卡尔纳普

中图分类号: B81

文献标识码: A

在 2015 年出版的《纯粹归纳逻辑》一书中, 作者帕里斯 (J. Paris) 和温科斯卡 (A. Vencovska) 对卡尔纳普致以崇高的敬意, 把他追认为“纯粹归纳逻辑”之父。([17], 第 x 页)

卡尔纳普 (R. Carnap) 被认为是 20 世纪最重要、最具有影响力的归纳逻辑思想家, 他的关于归纳逻辑的文献至今依然是归纳逻辑研究的活水源头, 恰如普特南 (H. Putnam) 所指出的:

为数理归纳逻辑建立一个基础的最为重要并且最具雄心的尝试归之于卡尔纳普……他激励大批年轻的逻辑学家对归纳产生兴趣。可以说, 今天卡尔纳普的工作对于整个归纳逻辑领域的重要性, 好比弗雷格的工作在 20 世纪初时对演绎逻辑的重要性: 显然, 按照现在的情况来说, 它是不能令人满意的, 但是没有其它真正的可替代的方法了。要么攻克卡尔纳普进路的难题, 要么抛弃整个计划。([19], 第 302 页)

帕里斯和温科斯卡继承了卡尔纳普的基本思想, 并且从逻辑技术方面大大地推进了卡尔纳普的构想。

因为类比推理不仅是卡尔纳普归纳逻辑的一个重要研究对象, 而且它还构成了《纯粹归纳逻辑》的一个关键的推理原则, 所以, 从类比推理入手, 不仅可以揭示纯粹归纳逻辑的思想渊源, 而且可以进一步分析归纳逻辑的最新研究成果, 最终将有助于我们更深入地把握归纳逻辑的发展趋势。

收稿日期: 2019-06-09

作者信息: 梁贤华 南开大学哲学院
hallolxh@163.com

基金项目: 国家社科基金重大项目“现代归纳逻辑的新发展、理论前沿与应用研究”(15ZDB018)。

总的来说,在纯粹归纳逻辑的框架下进行类比推理,关键是用距离函数来刻画相似性,并且以相关性函数作为推理基础。由于卡尔纳普的身影在其中随处可见,所以我们可以把它看作是卡尔纳普进路归纳逻辑的一个重要的复兴。

1 卡尔纳普的类比推理研究进路

不少学者认为,以卡尔纳普为代表的概率逻辑进路是类比推理的一个理想出路。例如,赫茜(M. Hesse)断言:“如果一个强有力的论断声称,我们是通过一个更加熟悉的现象领域来为新现象提供理性的预测,那么这种论断的逻辑涉及一种被称为‘类比论证’的东西。似乎只有卡尔纳普的确证理论能够为类比论证的解释难题提供一种充分地普遍并且详尽的理论工具。”([7],第319页)

可以把卡尔纳普对类比推理的研究划分为两个阶段。第一个阶段以他早期的论文《论归纳逻辑》和著作《概率的逻辑基础》为代表,其核心思想是明确地把类比推理划入到归纳逻辑的范畴中,确立了类比推理的理论地位。

具体地,在《论归纳逻辑》中,卡尔纳普说:“类比推理适用如下的情况。我们已经知道的证据是这样的事实,个体事物 b 和 c 在某种特征上是一致的,并且, b 另外具有更多的特征;因此我们考虑假说: c 也具有这样的特征。”([1],第87页)这是一种常见的符合直觉的类比推理模式。

在《概率的逻辑基础》中,卡尔纳普列出了五种“最重要的”归纳逻辑:([2],第207页)

1. 直接推理:从全体到个体的推理。
2. 预测推理(又称外延推理):从一个样本到另一个样本的推理。特别地,当第二个样本只含有一个个体时,预测推理又被称为单一预测推理。
3. 类比推理:从一个个体到另一个个体的推理(基于它们的相似性)。
4. 逆向推理:从个体到全体的推理。
5. 普遍性推理:从一个样本到一个具有普遍形式的假说推理。

卡尔纳普强调,这五种类型的归纳推理既不是互相排斥的,也没有穷尽所有类型的归纳推理。

总的来说,在第一阶段中,卡尔纳普确立了类比推理在归纳逻辑中的地位,并且认为类比推理的核心基础是相似性。但在这个阶段他主要的研究方法是把类比纳入到他的确证逻辑当中,现在学界一般认为确证的方法是不成功的,所以这个研究类比的方法也被认为是失败的。特别是在古德曼发现绿蓝悖论之后,人们意识到即便是根据同样的证据,通过归纳推理可能会得到两个互相矛盾的结论。这使得整个卡尔纳普的归纳逻辑都受到了挑战。为了应对古德曼的挑战,卡尔纳普

将归纳逻辑划分应用归纳逻辑和理论归纳逻辑两个部分。于是，卡尔纳普的归纳思想发展到第二个阶段。

卡尔纳普类比思想的第二个阶段以他去世后发表的著作《归纳逻辑的基本系统》为代表，其核心策略是用量化的“距离”来测度“相似性”。这个方法鲜明、直观，被认为是一个很有理论价值的思想。特别是在人工智能的量化发展要求的背景下，该方法尤其具有现实的意义。而纯粹归纳逻辑所要发展的类比推理正是后期卡尔纳普的类比思想。

具体地，卡尔纳普所提出的研究类比推理的新方法如下：([3], 第 8–11 页)

首先，确定一个谓词簇 $S = \{\text{红, 橙, 黄, 绿, 蓝, 紫}\}$ ，然后将谓词簇 S 表征为六个相等区域的颜色空间，如图：

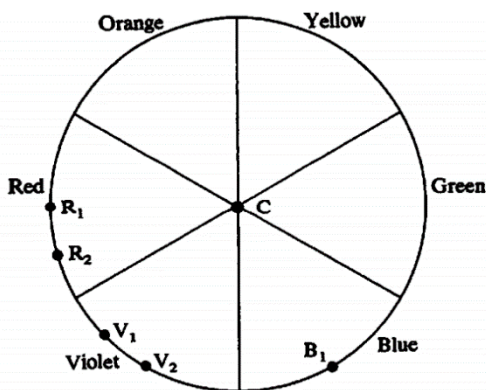


图 1: 卡尔纳普的颜色空间

其中圆心 C 代表自然灰色，圆周上依次被划分为红、橙、黄、绿、蓝、紫六个扇形区域，每个圆弧上的中点代表该种颜色的饱和点，或纯粹的色调，例如， R_1 表示饱和的红色，并且自 R_1 向圆心 C 逐渐靠近，意味着饱和的红色将逐渐变淡，最后将变成中性的灰色。

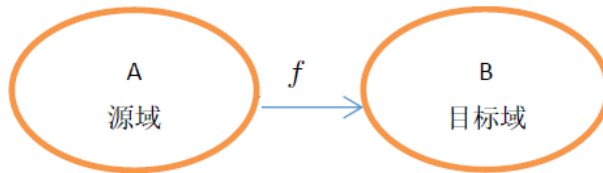
于是，当问及 R_2 （淡红色）和 V_1 （淡紫色）哪一个与 R_1 （饱和的红色）更相似时，人们可以毫不犹豫地說出答案： R_2 与 R_1 更相似。并且我们可以通过几何的距离来刻画这样的相似性比较： $d(R_1, R_2) < d(R_1, V_1)$ ，即距离越小，相似性就越大。卡尔纳普认为我们可以得到一个普遍的测度相似性的方法。

定义 1 (距离函数). 对于任意的集合 X 和 $x, y, z \in X$ ，映射 $d: X \times X \rightarrow \mathbb{R}^+$ 是一个距离函数，当且仅当以下 4 个条件成立：

1. $d(x, x) = 0$;
2. $d(x, y) > 0$ ，对于任意的 $x \neq y$;

3. $d(x, y) = d(y, x)$ (对称性条件);
4. $d(x, y) + d(x, z) \geq d(y, z)$ (三角形法则)。

卡尔纳普的类比推理新构想和二十世纪末发展起来的结构主义理论不谋而合。根据结构主义的观点, 类比是一个从源域到目标域的映射。也就是说, 类比是根据已知的信息 (A) 来推断未知的可能情况 (B), 而在这个推理的过程中, 最重要的根据是 A 和 B 具有相似的逻辑结构 f 。其基本的推理图式如下:



从结构主义的视角来看, 卡尔纳普的几何距离测度方法有两个优点。第一个优点是它对源域的规定非常明确, 即规定了它们是有限的谓词语言集合 (红、橙、黄、绿、蓝、紫), 这个方法被帕里斯纳入了 PIL 的体系中。事实上, 马里奥 (M. C. Di Maio) 也采用了同样的方法来研究类比推理: “我将会聚焦于这样的情形: 所有的属性都属于同一个种类 (例如, 颜色类), 并且只涉及不同的属性之间的相似性 (例如, 红与橙之间的相似性, 要比红与绿之间的相似性要大)。” ([5], 第 375 页) 第二个优点是, 可用距离函数来量化地刻画相似性。从实用价值的角度来看, 这是一个不可忽略的优点。

希尔 (A. Hill) 把卡尔纳普的方法看作是使用纯粹归纳逻辑的技术来刻画谓词之间的相似性的一个范例。他把相似性划分为三个类型 ([8], 第 21 页):

- (1) 谓词之间的相似性。
- (2) 原子之间的相似性。
- (3) 句子之间的相似性。

例如, “灰色”和“黑色”是谓词之间的相似性的一个例子, 而“行星绕太阳转动”和“电子绕原子转动”则是句子之间的相似性的一个例子。

希尔认为纯粹归纳逻辑研究的是 (1) 和 (2) 这两种类型的相似性。

我们认为, 更准确地说, 纯粹归纳逻辑实际上研究的是原子之间的相似性, 因为谓词本身即是由原子构成的。

事实上, 卡尔纳普自己也是通过原子来定义 Q -谓词的。他在《概率的逻辑基础》中给出了一个例子 ([2], 第 125 页)。即对于给定的语言 $\mathcal{L} = \{P_1, P_2, P_3\}$, 由于每一个 P_i 都有两种不同的取值情况 (P_i 和 $\neg P_i$), 这样便有 $2^3 = 8$ 个 Q -谓词, 例如 $Q_1 = P_1 \wedge P_2 \wedge P_3$, $Q_2 = P_1 \wedge P_2 \wedge \neg P_3$ 。在“状态描述”一节中, 我们将看到, Q_1 和 Q_2 实际上是 PIL 体系中的两个不同的原子。

卡尔纳普的类比研究新思路启发了帕里斯等学者，并且人们发现了更多的测度相似性的方法，而最重要的一种测度方法是下面将要讨论的“汉明距离”，它被直接地用来刻画原子之间的相似性。

2 一元纯粹归纳逻辑框架下的类比推理

“纯粹归纳逻辑”的理念来源于卡尔纳普，其目的是把归纳当作数理逻辑的一个分支来研究，正如杰弗里 (R. C. Jeffrey) 所指出的：“卡尔纳普认为纯粹归纳逻辑的任务是为 c 函数提供一个数学的定义。” ([10], 第 300 页) 在《纯粹归纳逻辑》中，帕里斯和温科斯卡更进一步说：“纯粹归纳逻辑是把理性的概率当作数理逻辑的一个分支来处理的研究。” ([17], 第 i 页) 这是因为他们相信，正如可以把数学划分为应用数学和理论数学一样，同样可以把归纳划分应用归纳和理论归纳两个部分。直觉上，这种研究策略是可取的。

具体地，在纯粹归纳逻辑的背景下，帕里斯和温科斯卡把类比推理建立在一阶逻辑的基础之上，其理论基础是概率函数，并且相关性函数尤其重要，他们的基本目标是把类比当作是数理逻辑的一个特殊领域来研究。

系统地表述 PIL 体系下的类比推理的著作是《纯粹归纳逻辑》([17])。此外，帕里斯指导的博士论文《在归纳逻辑中通过类比进行推理》([8])，和硕士论文《一元归纳逻辑中的类比推理》([4])，都是在 PIL 体系下展开的类比推理研究，因此也是重要的参考文献。

由于纯粹归纳逻辑包括了一元归纳和多元纯归纳两大部分，并且后者比前者要复杂得多，为了简化起见，我们将把精力集中在一元归纳的部分。下面我们首先论述纯粹归纳逻辑的理论基础。

2.1 纯粹归纳逻辑常用符号的说明

纯粹归纳逻辑建立在一阶逻辑语言的基础之上，但是在一阶逻辑语言中不含有函数和等式，而且关系符号是有限的，常项是可数无穷的，也就是说，预设常项穷尽了论域中的所有个体（元素）。

下面是纯粹归纳逻辑中常用符号的说明：

- $\mathcal{L}$: 一阶语言；
- 常项: $a_1, a_2, \dots, a_n$ (或 $a_{i_1}, a_{i_2}, \dots, a_{i_m}$)；
- 变元: $x_1, x_2, \dots$
- SL : 一阶语言 $\mathcal{L}$ 的句子的集合；
- $QFSL$: 没有量词的一阶语言的句子的集合；

- 关系符号: $R_1, R_2, \dots, R_q$, 并且这些关系符号的参数数量分别用 $r_1, r_2, \dots, r_q$ 来表示。当关系符号的参数数量只有一个的时候, 也用 $P_1, P_2, \dots, P_q$ 来表示一元关系谓词。

特别地, 还常用缩略式“PIL”来表示“纯粹归纳逻辑”。

2.2 纯粹归纳逻辑的总问题

在某种意义上, 我们可以把纯粹归纳逻辑看作是理性决策的问题。也就是说, 纯粹归纳逻辑所要研究的总问题是: 对于一个理性的主体来说, 如何对进行概率赋值才是合乎理性的?

具体地, 即在“零知识”的情况下, 逻辑地, 或理性地, 主体 X 应该对一个句子 $\theta \in SL$ 在某个结构 M 中为真的可能性赋予怎样的信念? ([17], 第10页)

帕里斯已经意识到至少需要对三个地方作出进一步的解释。首先, “零知识”是指除了一阶逻辑的基本原理之外, 主体 X 没有任何其它的背景知识。这类似于一种“无知的状态”。在现实的世界中, 可能根本就不存在这样的主体。不过, 如果我们把具有“零知识”的主体 X 看作是一个理想的“机器人”主体, 则这样的一种认识状态还是可以设想的。正如帕里斯所指出的: “这样的一种初期的无知状态的更为具体的例子, 我们仅只需要向人工智能中的主体寻求。” ([15], 第315页)

其次, 在帕里斯看来, “逻辑”和“理性”是一对同义词, 但是他没有定义什么是“理性”, 倒是把这个问题交给我们的常识和直觉去处理。根据帕里斯的学生阿萨罗 (F. A. D'Asaro) 的理解: “在这种语境下, 我们应当把‘理性’仅只看作是非形式的、直觉的概念。用卡尔纳普的话来说, 可以把理性这个术语看作我们正在寻求说明的一个说明物, 因此我们不应给出它的定义。” ([4], 第11页) 对于帕里斯来说, 在最低限度的意义上, 如果我们把纯粹归纳逻辑当作是行动的指南, 则不会被认为是非理性的。

再次, “信念”一词表明了 PIL 体系所采用的概率解释是主观解释。强调主观解释在现代人工智能的背景下具有重要的意义, 因为该解释能够很好地刻画信息传递的作用, 并且具有很强的操作价值。

2.3 概率函数

定义 2. 令 w 为从 SL 到 $[0, 1]$ 之间的一个映射, 即, $w: SL \rightarrow [0, 1]$, 那么 w 是一个概率函数, 仅当对任意的 $\theta, \phi, \psi \in SL$, 以下的三个条件得到满足:

(P1) 如果 $\models \theta$, 那么 $w(\theta) = 1$;

(P2) 如果 $\models \neg(\theta \wedge \phi)$, 那么 $w(\theta \wedge \phi) = w(\theta) + w(\phi)$;

$$(P3) \quad w(\exists x\psi(x)) = \lim_{n \rightarrow \infty} w(\psi(a_1) \vee \psi(a_2) \vee \dots \vee \psi(a_n)).$$

P3 又被称为盖弗曼 (Gaifman) 公理 (条件), 在 PIL 中具有重要的特殊意义, 因为 “它意在捕捉这样的直觉—— a_i 穷尽论域, 并且因此 $\exists x\psi(x)$ 应当等于无穷析取 $\bigvee_{i=1}^{\infty} \psi(a_i)$ 。” ([12], 第 800 页) 也就是说, 最终我们可以得到一个收敛的极限值。对于归纳推理来说, 这当然是一个理想的性质。

另一方面, 条件概率是贝叶斯主义的核心支柱, 常常被誉为一种反映了 “向经验学习” 的精髓的逻辑工具。用主观贝叶斯主义的话语来说, 它刻画了新信息对主体 X 的信念更新所产生的作用。

定义 3. 假设 w 是满足定义 2 的概率函数, 并且 $\theta, \phi \in SL$, $w(\phi) > 0$ 。那么, 在已知 ϕ 的条件下发生 θ 的概率, 定义为映射函数 $w(\cdot | \phi) : SL \rightarrow [0, 1]$, 即

$$w(\theta | \phi) = \frac{w(\theta \wedge \phi)}{w(\phi)}$$

注意, 在上述定义中, 已经规定 $w(\phi) > 0$ 。考虑到可能存在 $w(\phi) = 0$ 的情况, 有时候条件概率公式也被写成 $w(\theta \wedge \phi) = w(\theta | \phi)w(\phi)$, 或对 $w(\phi) = 0$ 的情况重新下一个定义。

2.4 理性原理

理解 PIL 的一个比较好的方法是把它看作一种决策论。根据帕里斯和温科斯卡, 这是因为在许多情况下, 例如对一个空缺的岗位, 主管单位在决定应当雇佣谁的时候, 所寻求的并不是客观的真理, 而是力图使管理层内部达成一致的意见。所以, 如果我们希望将来的某一天能够把雇佣工人的工作交给人工智能主体 X 去做的话, 就要对施加于我们的偏好之上的理性观察的约束条件进行研究。([16], 第 358 页)

而 “约束条件” 可被理解为需要遵循的 “理性原理”。

为了便于读者理解, 我们首先介绍 “置换” 这个概念。

定义 4. 令 $S = \{x_1, x_2, \dots, x_n\}$, 那么 S 的置换 σ 是一个从 S 到 S 的一一映射, 即 $\sigma(x_i) = x_j$, 其中 $i, j \in \{1, 2, \dots, n\}$ 。

特别地, 当 $S = \{1, 2, 3, \dots\}$, 即 S 是一个自然数集合时, 可以把 S 的置换记作 $\sigma(1), \sigma(2), \sigma(3), \dots$, 或 $\sigma \in S_{\mathbb{N}+}$ 。

例 1. 假设 $S = \{1, 2, 3, 4\}$, 则 S 的一个的置换是:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

即 $\sigma(1) = 3$, $\sigma(2) = 1$, $\sigma(3) = 2$, $\sigma(4) = 4$ 。

有了置换的概念,我们便可以进一步得到 PIL 中的三个重要的理性原理 ([17], 第 33–35 页):

常项可交换性原理 (Ex) 令 σ 为一个从自然数到自然数的置换, 即 $\sigma: \mathbb{N} \rightarrow \mathbb{N}$, 则对 $\theta(a_1, a_2, \dots, a_n) \in SL$,

$$w(\theta(a_1, a_2, \dots, a_n)) = w(\theta(a_{\sigma(1)}, a_{\sigma(2)}, \dots, a_{\sigma(n)}))$$

对 Ex 的一个合理性证明是, 如果主体 X 对 $a_1, a_2, \dots, a_n$ 一无所知, 那么在进行概率赋值的时候, 他/她就不应该有区别地对待它们, 否则将会被认为是非理性的。([14], 第 6 页) 或者, 从另一个角度来看, Ex 是建立在这样的一种思想基础之上的: 理性主体的信念应当反映这样的一个事实——在他的主观世界中, 常项 a_i 完全是对称的 ([13], 第 5 页)。在这里, “对称性” 意味着它们具有同等的可能性。

例如, 假设从一个盒子中有放回地取球, 结果取得 2 个黑球, 1 个白球。那么 Ex 刻画的是这样的一种情况: 所得的结果与顺序无关。换句话说, 顺序“白、黑、黑”和顺序“黑、白、黑”的概率是一样的。

Ex 是 PIL 的一个标准预设, 即如果没有特别的说明的话, 就默认该条件得到满足。 Ex 考察的是常项的性质, 自然地, 我们可以设想, 对于谓词, 是否也有相似的性质呢? 这便是下面的原理所要回答的问题。

谓词可交换性原理 (Px) 假设关系谓词 $R_1, R_2, \dots, R_q$ 的参数数量都是 r , 则对于 $\theta, \theta' \in QFSL$, 如果 θ' 中的谓词是 θ 中的谓词的一个任意的置换, 那么 $w(\theta) = w(\theta')$, 即

$$w(\theta(R_1, R_2, \dots, R_q)) = w(\theta(R_{\sigma(1)}, R_{\sigma(2)}, \dots, R_{\sigma(n)}))$$

例如, $w(R_1(a_1) \wedge \neg R_2(a_2)) = w(R_2(a_1) \wedge \neg R_1(a_2))$ 。

另一个重要的原理是, 否定词 $\neg$ 也应具有对称性的性质。例如, 一枚均匀的硬币出现正面向上的概率和出现反面向上的概率应该是一样的。下面的原理所刻画的便是这样的一种理念。

强否定原理 (SN) 对于 $\theta, \theta^\neg \in QFSL$, 如果 $\theta^\neg$ 是将 θ 中的每一个 $\pm R$ 都用 $\mp R$ 来替换而得到的结果, 那么

$$w(\theta) = w(\theta^\neg)$$

例如, $w(R_1(a_1) \wedge \neg R_2(a_2)) = w(R_1(a_1) \wedge R_2(a_2))$ 。

SN 要求的是将 θ 中出现的每一个 R 都换成 $\neg R$, 但它并不要求同时地将有所的 R_i 都置换成 $\neg R_i$ ($i = 1, 2, \dots, q$)。因此, 作为 SN 的一种特殊情况, 我们还可以得到一个“弱否定原理”。

弱否定原理 (WN) 对 $\theta \in SL$, 假设 $\theta^\neg$ 是同时地将 θ 中的每一个 R_i 都置换为它的否定 $\neg R_i$ 所得的结果, 则 $w(\theta) = w(\theta^\neg)$ 。

可见, 在某种意义上说, 我们可以把 WN 看作是 SN 条件的弱化, 因此, 在 PIL 的体系中, 有时候, 当我们需要削弱前提条件的强度时, 一个重要的策略便是用 WN 来替换 SN 。 PIL 体系下的类比推理便使用了这个方法。

至此我们已经知道, 在缺乏关于现实世界的其它的参考信息的“零知识状态”预设下, 基于对“对称性”的考察, 得到了三个重要的理性原理 Ex 、 Px 和 SN 。相似地, 我们还可得到另外一个重要的原理:

原子可交换性原理 (Ax) 对 $\{1, 2, \dots, 2^q\}$ 的任意置换 σ , 和常项 $a_{i_1}, a_{i_2}, \dots, a_{i_m}$,

$$w\left(\bigwedge_{k=1}^m \alpha_{h_k}(a_{i_k})\right) = w\left(\bigwedge_{k=1}^m \alpha_{\sigma(h_k)}(a_{i_k})\right)$$

我们已经知道, Ex 刻画的是, 从一个盒子中有放回地取球, 所得结果的概率与取球的顺序无关。例如, 假设取得 2 个红球、1 个黄球和 2 个白球, 则顺序“红、白、红、白、黄”和顺序“黄、红、红、白、白”的概率是一样的。相似地, Ax 所刻画的直觉信念是: 在 Ex 被满足的前提下, Ax 等于说, 我们赋予这些颜色的名称是无关的。例如, 取得 2 个红球、1 个黄球和 2 个白球, 和取得 2 个绿球、1 个黄球和 2 个黑球的概率是一样的 ([15], 第 321 页)。

因为涉及到对“原子”的定义, 本应在“状态描述”一节内容之后给出理性原理 Ax 的, 但是, 为了突出 PIL 体系中的“原理”的重要意义, 我们决定将它放在这里。

2.5 状态描述

“状态描述”是纯粹归纳逻辑的一个核心概念, 该概念是卡尔纳普归纳逻辑的一个重要成果, 它的重要性见之于凯梅尼 (J. G. Kemeny) 的评论: “可以说, 卡尔纳普的整个归纳逻辑的方法建立在一个概念——‘状态描述’的基础之上。所有的其它概念都可以根据这个概念进行定义。” ([11], 第 39 页)

定义 5. 令 $R_1, R_2, \dots, R_q$ 为一阶语言 $\mathcal{L}$ 的关系谓词, 并且它们的参数数量分别是 $r_1, r_2, \dots, r_q$ 。那么, 对于来自 $a_1, a_2, \dots, a_m$ 的不同的常项 $a_{i_1}, a_{i_2}, \dots, a_{i_m}$, 状态描述 $\Theta(a_{i_1}, a_{i_2}, \dots, a_{i_m})$ 是一个具有如下形式的一阶语言 $\mathcal{L}$ 的句子:

$$\bigwedge_{j=1}^q \bigwedge_{c_1, c_2, \dots, c_{r_j}} \pm R_j(c_1, c_2, \dots, c_{r_j})$$

其中 $+R_j$ 表示 R_j , $-R_j$ 代表 $\neg R_j$, $c_1, c_2, \dots, c_{r_j}$ 来自于 $a_{i_1}, a_{i_2}, \dots, a_{i_m}$ (不必是不同的)。 ([17], 第 41–42 页)

例 2. 令 $\mathcal{L} = \{R_1, R_2\}$, 其中 R_1 是一元关系符号, R_2 是二元关系符号。那么对于常项 a_1, a_2 , 一个可能的状态描述是

$$\Theta(a_1, a_2) = \neg R_1(a_1) \wedge R_1(a_2) \wedge R_2(a_1, a_2) \wedge R_2(a_2, a_1) \wedge R_2(a_1, a_1) \wedge \neg R_2(a_2, a_2)$$

在 PIL 的体系中, 除了用 Θ 符号来表示状态描述之外, 有时候还用 Ψ 和 Φ 来表示, 以作出不同的区分。

特别地, 当关系谓词 $R_1, R_2, \dots, R_q$ 的参数数量都只有一个的时候, 我们便得到一元纯粹归纳逻辑下的状态描述。这时候, 我们便可以把 $a_{i_1}, a_{i_2}, \dots, a_{i_m}$ 的状态描述

$$\bigwedge_{j=1}^q \bigwedge_{c_1, c_2, \dots, c_{r_j}} \pm R_j(c_1, c_2, \dots, c_{r_j})$$

简化为

$$\begin{aligned} \Theta(a_{i_1}, a_{i_2}, \dots, a_{i_m}) &= \bigwedge_{j=1}^q \bigwedge_{k=1}^m \pm R_j(a_{i_k}) \\ &= \bigwedge_{k=1}^m \alpha_{h_k}(a_{i_k}) \end{aligned}$$

其中 $\alpha_1(x), \alpha_2(x), \dots, \alpha_{2^q}(x)$ 都是 $\mathcal{L}$ 的原子, 它们的一般形式是:

$$a_i(x) = \pm R_1(x) \wedge \pm R_2(x) \wedge \dots \wedge \pm R_q(x)$$

有时候, 也把原子 $\alpha_i(x)$ 记作 $R_1^{\epsilon_1} \wedge R_2^{\epsilon_2} \wedge \dots \wedge R_q^{\epsilon_q}$, $\epsilon_i \in \{0, 1\}$, $i \in \{1, 2, \dots, q\}$, $R_i^0 = \neg R_i$, $R_i^1 = R_i$ 。

也就是说, 对于 $\mathcal{L} = \{R_1, R_2, \dots, R_q\}$, 每个 R_j 都有两种可能的取值情况, 即 $\pm R_j$ (R_j 和 $\neg R_j$), 因此, 对语言 $\mathcal{L}$ 中的关系谓词按照字典式序列取合取公式, 便可以得到 2^q 个原子公式。([17], 第 50 页)

例 3. 令 $\mathcal{L}_2 = \{R_1, R_2\}$, 其中 R_1, R_2 是一元关系谓词。则 $\mathcal{L}_2$ 的原子是:

$$\begin{aligned} \alpha_1(x) &= R_1(x) \wedge R_2(x) & \alpha_2(x) &= \neg R_1(x) \wedge R_2(x) \\ \alpha_3(x) &= R_1(x) \wedge \neg R_2(x) & \alpha_4(x) &= \neg R_1(x) \wedge \neg R_2(x) \end{aligned}$$

如果 R_1 表示“红色”的性质, R_2 表示“黄色”的性质, 那么原子 $\alpha_1(x)$ 表示 x 混合了红、黄两种颜色。原子 $\alpha_4(x)$ 则表示 x 既没有红色也没有黄色。显然, $\alpha_1(x)$ 和 $\alpha_4(x)$ 完全是互相排斥的。事实上, 原子 $\alpha_1(x), \alpha_2(x), \alpha_3(x), \alpha_4(x)$ 都是互相排斥的。

在例 3 中, 我们利用颜色来说明原子的性质, 我们在往后的论述中, 还会用相似的例子作理论的说明。事实上, 这是 PIL 的一种常用的方法。正如帕里斯所指出的: 思考原子的一个有用的方法是把它看作是颜色——假设 a_i 是球体, 并且这些球体的唯一区别就是颜色。因此, 认识一个球体的颜色就认识了 a_i 的一切。([14], 第 7 页) 正如例 3 所揭示的, 对于给定的语言 $\mathcal{L}_2$, 原子 $\alpha_1(x), \dots, \alpha_4(x)$ 刻画了 x 可能具有的所有属性。

在定义了一元谓词状态描述后, 便可以进一步探讨状态描述的数学表征问题。

首先, 定义一个论域 D_{2^q} , 使得 $D_{2^q} \subseteq R^{2^q}$, 即

$$D_{2^q} = \{ \langle x_1, x_2, \dots, x_{2^q} \rangle \mid x_i \geq 0, \sum_{i=1}^{2^q} x_i = 1, i = 1, 2, \dots, 2^q \}$$

其中 $x_i = w(\alpha_i(x))$ 。

其次, 定义状态描述上的函数。

令 $\vec{x} = \langle x_1, x_2, \dots, x_{2^q} \rangle \in D_{2^q}$, 则 $w_{\vec{x}}$ 是定义在状态描述 $\Theta(a_{i_1}, a_{i_2}, \dots, a_{i_m})$ 上的一个概率函数, 仅当

$$\begin{aligned} w_{\vec{x}}(\Theta(a_{i_1}, a_{i_2}, \dots, a_{i_m})) &= w_{\vec{x}} \left(\bigwedge_{k=1}^m \alpha_{h_k}(a_{i_k}) \right) \\ &= \prod_{k=1}^{2^q} x_k^{m_k} \end{aligned}$$

其中 $m_1, m_2, \dots, m_{2^q}$ 分别是原子 $\alpha_1, \alpha_2, \dots, \alpha_{2^q}$ 的数量。

在前述两个定义的基础上, 可以进一步地表述著名的德菲涅表征定理 (de Finetti's Representation Theorem) ([17], 第 55 页):

定理 1. 令 $\mathcal{L} = \{R_1, R_2, \dots, R_q\}$, 其中 $R_1, R_2, \dots, R_q$ 是一元谓词。如果 $\mathcal{L}$ 上的概率函数 w 满足 Ex , 那么存在一个测度 μ , 使得 μ 在 D_{2^q} 的布尔子集上, 并且

$$\begin{aligned} w \left(\bigwedge_{k=1}^m \alpha_{h_k}(a_{i_k}) \right) &= w(\Theta(\alpha_1^{m_1}, \alpha_2^{m_2}, \dots, \alpha_{2^q}^{m_{2^q}})) \\ &= \int_{D_{2^q}} x_1^{m_1} x_2^{m_2} \dots x_{2^q}^{m_{2^q}} d\mu \end{aligned}$$

相反地, 给定在 D_{2^q} 的布尔子集上的测度 μ , 通过上述定义得到的概率函数 w 唯一地扩展到 SL 上的一个满足 Ex 的概率函数。

其中, μ 又被称为 w 的“德菲涅先验” (de Finetti prior)。

德菲涅表征定理是纯粹归纳逻辑的一个关键的数理逻辑工具, 许多定理需要通过它才能得到证明。但是该定理很复杂, 因此很容易成为人们诟病 PIL 体系的根据。

2.6 相关性

对于一个理性的主体 X 来说, 如果他/她观察到某个事件在过去发生的次数越多, 那么他/她就越有理由相信该事件在将来发生的可能性越大。这又被比喻为“向经验学习”。例如, 假设我们来到了一个新的地方, 不妨称之为 A 地。如果在第一天我们连续地看到了 9 只黑色的乌鸦, 那么我们就会有较大的信心预期: 我们在 A 地看到的第 10 只乌鸦将也是黑色的。而刻画这种向“经验学习”直觉信念的一个方法是:

实例相关原则 (PIR) 对于 $\mathcal{L}$ 的原子 $\alpha(x)$ 和 $\theta(a_1, a_2, \dots, a_n)$,

$$w(\alpha(a_{n+2}) \mid \alpha(a_{n+1}) \wedge \theta(a_1, a_2, \dots, a_n)) \geq w(\alpha(a_{n+2}) \mid \theta(a_1, a_2, \dots, a_n))$$

也就是说, 在已有证据 $\theta(a_1, a_2, \dots, a_n)$ 的基础上获悉 a_{n+1} 具有性质 α , 将会增强我们断言下一个常项 a_{n+2} 也具有同样的性质的信心。([17], 第 70 页)

定理 2. Ex 蕴涵 PIR 。

也就是说, 如果我们相信 Ex 是合理的, 那么根据逻辑的结果, 我们便应该同意 PIR 也是合理的。这个定理意味着卡尔纳普归纳逻辑蓝图的一个重大的胜利。不过, 该定理的证明过程十分复杂, 并且严重地依赖德菲涅表征定理, 即使是对有一定数学背景知识的人来说, 该定理也是不容易把握的。帕里斯对此深有忧虑, 他认为, 可能存在这样的一种情况: 即人们可能接受 Ex , 但是并不接受 PIR 。因此, 他寄希望于将来, 但愿将来的某一天有人能够提出更加容易理解的证明方法。([14], 第 9 页)

2.7 卡尔纳普的连续统方法

我们在前面已经知道, 可以把 PIL 看作是卡尔纳普归纳逻辑的最新研究成果。在这里我们便可以看到一个重要的根据: 卡尔纳普的连续统概率函数 c_λ^C , 其独特的记号被保留下来了——需知道, 帕里斯特别地定义了概率函数, 在 PIL 的体系中, 除了卡尔纳普的连续统函数之外, 都是用 w 来表示概率函数的。

定义 6. 令 $\mathcal{L} = \{R_1, R_2, \dots, R_q\}$, 其中 $R_1, R_2, \dots, R_q$ 均为一元谓词。定义概率函数 c_λ^C 如下 ([15], 第 316 页):

$$c_\lambda^C \left(\alpha_j(a_{n+1}) \mid \bigwedge_{k=1}^n \alpha_{h_k}(a_k) \right) = \frac{m_j + \lambda 2^{-q}}{n + \lambda}$$

其中, m_j 是原子 α_j 出现的次数, $\sum_{j=1}^{2^q} m_j = n$, $0 \leq \lambda \leq \infty$ 。

特别地, 如果 n 是一个固定的自然数, 那么,

- 当 $\lambda \rightarrow \infty$ 时,

$$\begin{aligned} \lim_{\lambda \rightarrow \infty} \frac{m_j + \lambda 2^{-q}}{n + \lambda} &= \lim_{\lambda \rightarrow \infty} \frac{\frac{m_j}{\lambda} + 2^{-q}}{\frac{n}{\lambda} + 1} \\ &= 2^{-q} \end{aligned}$$

- 当 $\lambda = 0$ 时,

$$c_0^{\mathcal{L}} = \left(\bigwedge_{k=1}^n \alpha_{h_k}(a_k) \right) = \begin{cases} 2^{-q} & \text{如果 } h_1 = h_2 = \dots = h_n, \\ 0 & \text{其他情况。} \end{cases}$$

- 当 $\lambda = \infty$ 时,

$$c_{\infty} = \left(\bigwedge_{j=1}^n \alpha_{i_j}(a_{t_j}) \right) = 2^{-nk}$$

下面是一个源自于卡尔纳普的例子，阿萨罗进一步把它整合到 PIL 体系中来。（[4]，第 35–36 页）这个例子不仅能够帮助我们更好地把握卡尔纳普连续统函数，而且它还有助于我们理解 PIL 体系中的类比推理的重要意义。

例 4. 令 $\mathcal{L} = \{P_1, P_2, P_3\}$ 。假设某地的昆虫的种类是有限的，即可用语言 $\mathcal{L}$ 的原子来充分地刻画，不妨设昆虫的颜色类型正好等于语言 $\mathcal{L}$ 所构成的原子的个数，即 $2^3 = 8$ （种）。假设已经观察到了 20 只昆虫，分布如下：

昆虫颜色：	α_1	α_2	α_3	α_4	α_5	α_6	α_7	α_8
$m_i =$ 观察到的相应昆虫数量：	2	6	4	2	3	0	1	2

不妨令

$$\begin{aligned} \alpha_1 &= P_1(x) \wedge P_2(x) \wedge P_3(x) & \alpha_2 &= P_1(x) \wedge P_2(x) \wedge \neg P_3(x) \\ \alpha_3 &= P_1(x) \wedge \neg P_2(x) \wedge P_3(x) & \alpha_4 &= \neg P_1(x) \wedge P_2(x) \wedge P_3(x) \\ \alpha_5 &= \neg P_1(x) \wedge \neg P_2(x) \wedge P_3(x) & \alpha_6 &= \neg P_1(x) \wedge P_2(x) \wedge \neg P_3(x) \\ \alpha_7 &= P_1(x) \wedge \neg P_2(x) \wedge \neg P_3(x) & \alpha_8 &= \neg P_1(x) \wedge \neg P_2(x) \wedge \neg P_3(x) \end{aligned}$$

根据卡尔纳普的 $c_{\lambda}^{\mathcal{L}}$ 函数，“第 21 只昆虫的颜色是 α_1 ”的概率为

$$c_{\lambda}^{\mathcal{L}} = \left(\alpha_1(a_{21}) \mid \bigwedge_{k=1}^{20} \alpha_{h_k}(a_k) \right) = \frac{2 + 2^{-q}}{20 + \lambda}$$

由于颜色为 α_8 的昆虫数量也是 2，所以“第 21 只昆虫的颜色是 α_8 ”的概率为

$$c_{\lambda}^{\mathcal{L}} = \left(\alpha_8(a_{21}) \mid \bigwedge_{k=1}^{20} \alpha_{h_k}(a_k) \right) = \frac{2 + 2^{-q}}{20 + \lambda}$$

显然,根据卡尔纳普连续统函数的计算结果,“第21只昆虫的颜色是 α_1 ”的概率,与“第21只昆虫的颜色是 α_8 ”的概率,是相等的。

但是,两者的预测概率真的是完全一样吗?答案是:不一样。

因为当我们观察这8个原子的谓词(P_1, P_2, P_3)前面的符号($\pm$)时,便会发现 $\alpha_2, \alpha_3, \alpha_4$ 和 α_1 更相相似,其数量是 $m_1 + m_2 + m_3 + m_4 = 2 + 6 + 4 + 2 = 14$ 。而 $\alpha_5, \alpha_6, \alpha_7$ 则和 α_8 更加相似,但其数量只有 $m_5 + m_6 + m_7 + m_8 = 3 + 0 + 1 + 2 = 6$ 。显然, $14 > 6$, 因此,基于现有的观察数据,第21只昆虫的颜色是 α_1 的概率要大于其颜色为 α_8 的概率。这正是PIL体系下的类比推理研究所要解释的问题。而类比推理第一个需要解决的问题就是如何刻画相似性的问题。

2.8 汉明距离和相似性的刻画

相似性是类比推理的核心概念,如果我们能够把握它,也就是掌握了类比推理的制高点。

而我们已经知道,卡尔纳普提出了用几何距离函数来测度相似性的构想,并且他的 Q -谓词实质上是由原子所构成的。因此,一个很自然的推断就是,我们能否直接地考察原子之间的相似性呢?答案是肯定的。事实上,直接对原子进行考察还有一个优点:可以对给定的语言进行逻辑分析而不至于避免陷入本体论预设的泥潭——卡尔纳普的颜色空间是有争议的。

下面我们通过一个例子来说明原子之间的相似性。我们知道,在例3中,对给定的语言 $\mathcal{L} = \{R_1, R_2\}$, $\mathcal{L}$ 有 $2^2 = 4$ 个原子,它们分别是:

$$\begin{aligned}\alpha_1(x) &= R_1(x) \wedge R_2(x) & \alpha_2(x) &= \neg R_1(x) \wedge R_2(x) \\ \alpha_3(x) &= R_1(x) \wedge \neg R_2(x) & \alpha_4(x) &= \neg R_1(x) \wedge \neg R_2(x)\end{aligned}$$

那么,首先,我们可以把同一性看作是相似性的一种特殊情况,例如,原子 α_1 与其自身是相似的,当获知 a_n 满足 α_1 时,自然增加了 a_{n+1} 也满足 α_1 的预测概率,即有

$$w(\alpha_1(a_{n+1}) \mid \alpha_1(a_n) \wedge \psi(\vec{a})) \geq w(\alpha_1(a_{n+1}) \mid \psi(\vec{a}))$$

而这实际上就是实例相关原则。帕里斯说:“实例相关原则,可以被解释成‘由类比所支持’的一个例子。”([17], 第165页)因此,在某种意义上,我们可以把PIL体系中的类比看作相关性函数的一种扩展。其次,当我们比较原子 $\alpha_1, \alpha_2, \alpha_4$ 时,便于会发现, α_1 和 α_2 只有关系符号 R_1 前面的标志是不同的,而 α_1 和 α_4 的两个关系符号 R_1 和 R_2 的标志都是不同的。因此, α_1 和 α_2 更加相似。所以,获知 a_n 满足 α_2 ,要比获知它满足 α_4 更有利于预测 a_{n+1} 满足 α_1 。而刻画这种原子之间的相似性的一个重要工具是“汉明距离”(Hamming distance) ([8], 第36页):

定义 7. 两个原子之间的汉明距离, 例如, α_j 和 α_k 之间的汉明距离, 是满足 $\alpha_j(x) \models R(x) \Leftrightarrow \alpha_k(x) \models \neg R(x)$ 的谓词 R 的数量, 记为 $\|\alpha_j - \alpha_k\|$ 。

即是说, 原子 α_j 和 α_k 的汉明距离, 体现在 $\pm R_i$ 上的差别的具体数量。

具体地, 在例 3 中, $\alpha_1 \models R_1(x)$, $\alpha_2 \models \neg R_1(x)$, 即原子 α_1 和 α_2 的区别只是关系符号 R_1 的标志不同, 所以 $\|\alpha_1 - \alpha_2\| = 1$ 。另一方面, 由于 $\alpha_1 \models R_1(x)$, $\alpha_4 \models \neg R_1(x)$, 并且 $\alpha_1 \models R_2(x)$, $\alpha_4 \models \neg R_2(x)$, 因此 $\|\alpha_1 - \alpha_4\| = 2$ 。

我们还可以用图示来刻画汉明距离。例如, 令正方形的边长是单位 1, 则我们可以分别地用正方形的四个顶点来表示例 3 中的四个原子。这样, 两个不同的原子之间的汉明距离, 便是沿着正方形的周长运动的两个顶点之间的最短距离。例如, 从原子 $R_1(x) \wedge R_2(x)$ 到原子 $\neg R_1(x) \wedge R_2(x)$ 和 $\neg R_1(x) \wedge \neg R_2(x)$ 的汉明距离分别是 1 和 2。

$$\begin{array}{ccc}
 R_1(x) \wedge R_2(x) & \xrightarrow{1} & R_1(x) \wedge \neg R_2(x) \\
 \downarrow 1 & \text{---} & \downarrow 1 \\
 \neg R_1(x) \wedge R_2(x) & \xrightarrow{1} & \neg R_1(x) \wedge \neg R_2(x)
 \end{array}$$

(Note: The diagram shows a square with vertices representing atomic states. The top-left vertex is $R_1(x) \wedge R_2(x)$, the top-right is $R_1(x) \wedge \neg R_2(x)$, the bottom-left is $\neg R_1(x) \wedge R_2(x)$, and the bottom-right is $\neg R_1(x) \wedge \neg R_2(x)$. The edges are labeled with the number of predicate differences (Hamming distance). The distance between top-left and top-right is 1. The distance between top-left and bottom-left is 1. The distance between top-left and bottom-right is 2 (via the top-right or bottom-left vertex).)

可见, 汉明距离越小, 则原子之间的相似性越大。于是, 有

类比原则 (AP_S) 对 $\mathcal{L}$ 的任意原子 $\alpha_h, \alpha_i, \alpha_j$, 如果

$$\|\alpha_h - \alpha_i\| < \|\alpha_h - \alpha_j\|$$

那么,

$$w(\alpha_h(a_{n+1}) \mid \alpha_i(a_n) \wedge \psi(\vec{a})) \geq w(\alpha_h(a_{n+1}) \mid \alpha_j(a_n) \wedge \psi(\vec{a}))$$

其中 $\psi \in QFSL$, 并且 ψ 不包含 a_n 和 a_{n+1} 。

当上述类比原则中的“ $\geq$ ”为更强的“ $>$ ”关系时, 称之为“严格类比原则”, 即有

严格类比原则 (SAP_S) 对 $\mathcal{L}$ 的任意原子 $\alpha_h, \alpha_i, \alpha_j$, 如果

$$\|\alpha_h - \alpha_i\| < \|\alpha_h - \alpha_j\|$$

那么,

$$w(\alpha_h(a_{n+1}) \mid \alpha_i(a_n) \wedge \psi(\vec{a})) > w(\alpha_h(a_{n+1}) \mid \alpha_j(a_n) \wedge \psi(\vec{a}))$$

综上所述, 在类比原则 AP_S 和 SAP_S 中, 关键的条件是汉明距离——它构成了判断相似性的基础。

2.9 类比原理 (S)AP_S 的基本性质及其局限性

为了更加深入地研究 PIL 中的类比推理的性质, 一个有用的方法便是考察它与 PIL 的体系中其它重要性质的逻辑关系。而这对于我们更好地把握 PIL 也是很有必要的。

我们先给出一个源自于卡尔纳普的理性原则的定义 ([17], 第 61 页):

规则性原则 (R_{EG}) 对任意的 $\theta \in QFSL$, $w(\theta) > 0$ 。

可以从两个方面来理解为什么要规定 R_{EG}: 一方面, 由于 PIL 的总问题预设了理性主体 X 没有任何关于 θ 的背景知识, 所以没有充足理由把 θ 看作是不可能事件; 另一方面, 更为现实的意义是, 规定 $w(\theta) > 0$ 可以避免条件概率出现分母为零的情况。

定义了 R_{EG} 之后, 便可以给出严格类比原则 SAP_S 的两个重要性质: 一、SAP_S 与 Ax 不可并存; 二、SAP_S 蕴涵 R_{EG}。可分别地用命题的方式来刻画 ([17], 第 166 页):

命题 1. SAP_S 与 Ax 不一致。

证明. 不妨设存在不同的原子 $\alpha_h(x), \alpha_i(x), \alpha_j(x)$, 使得 $\|\alpha_h - \alpha_i\| < \|\alpha_h - \alpha_j\|$, 并且令 $\psi(\vec{a})$ 为重言式。那么, 根据 SAP_S, 有

$$w(\alpha_h(a_{n+1}) \mid \alpha_i(a_n) \wedge \psi(\vec{a})) > w(\alpha_h(a_{n+1}) \mid \alpha_j(a_n) \wedge \psi(\vec{a}))$$

然而, 根据原子可交换性原理 Ax, $w(\bigwedge_{i=1}^n \alpha_{h_i}(a_i)) = w(\bigwedge_{i=1}^n \alpha_{\sigma(h_i)}(a_i))$, 故有

$$w(\alpha_h(a_{n+1}) \mid \alpha_i(a_n) \wedge \psi(\vec{a})) = w(\alpha_h(a_{n+1}) \mid \alpha_j(a_n) \wedge \psi(\vec{a}))$$

可见, 两者的结论是矛盾的。□

那么, 人们自然会问, 命题 1 是否意味着 Ax 和 SAP_S 之间至少有一个是错误的? 我们认为更恰当的理解方式是把 SAP_S 看作是 Ax 的一个矫正原则, 也就是说, SAP_S 涉及到对所收集到的证据的综合考虑, 而 Ax 主要刻画的是主体 X 在证据不足的时候对各种不同的情况赋予同等的概率。其实, 在例 4 中我们便已经看到了类比原则对卡尔纳普连续统函数的一个修正。

命题 2. SAP_S 蕴涵 R_{EG}。

证明. 根据 R_{EG}, 对于任意的 $\psi \in QFSL$, $w(\psi) > 0$ 。

假设 SAP_S 不蕴涵 R_{EG}, 不妨令 $w(\psi) = 0$ 。因此, 将有

$$w(\alpha_h(a_{n+1}) \mid \alpha_i(a_n) \wedge \psi(\vec{a})) = w(\alpha_h(a_{n+1}) \mid \alpha_j(a_n) \wedge \psi(\vec{a}))$$

而根据 SAP_S, 上式中的 “=” 应为 “>”, 矛盾。□

此外, 还可以用其它的命题和定理来进一步刻画 SAP_S 的性质。

命题 3. 当 $q = 1$, 即只有一个谓词的时候, $\mathcal{L} = \{P\}$ 。则如果 w 在语言 $\mathcal{L}$ 中满足 Ex , Px , 和 SN , 那么 w 也满足 AP_S 。

对于这个命题, 关键的目标是证明 ([9], 第 1301–1302 页):

$$w(P(a_{n+1}) \mid P(a_n) \wedge \psi(\vec{a})) \geq w(P(a_{n+1}) \mid \neg P(a_n) \wedge \psi(\vec{a}))$$

PIL 的另一个重要的关注点是, 在谓词语言的收缩或扩张的情况下, 已有的定理或命题是否还成立。命题 4 表明 AP_S 不受语言的收缩 (或扩张) 的影响 ([9], 第 1300–1301 页):

命题 4. 令 $\mathcal{L} = \{P_1, P_2, \dots, P_q\}$ 。假设 $\mathcal{L}$ 上的概率函数 w 满足 AP_S , 并且 $\mathcal{L}^- \subset \mathcal{L}$, 那么 w 被限制于 $\mathcal{L}^-$ 时仍然满足 AP_S 。

根据前述的分析可知, PIL 的体系中的类比推理通常是建立在 Ex , Px 和 SN 这几个预设之上的, 但是, 由于这些前提条件太强了, 因此使得 SAP_S 只能适用于非常有限的情况。这可从下面的定理得到说明 ([17], 第 166–167 页):

定理 3. 当 $q > 2$ 时, SL 上没有函数同时满足 Ex , Px , SN 和 SAP_S 。

可见, 对于语言 $\mathcal{L}$, 当它的关系谓词多于两个的时候, SAP_S 的局限性便突显出来了, 即它只能适用于非常有限的范围。其中一个重要原因是因为它的前提条件要求太苛刻了, 因此, 为了扩大 SAP_S 的应用范围, 一个基本的思路便是弱化前提条件, 例如, 用更弱的预设 WN 来取代 SN 。这是将来进一步研究的思路。

2.10 对等类比原则

一个可能大家都比较熟悉的科学类比是: 由于火星和地球的结构颇为相似, 而地球上生命存在, 所以火星上可能也有生命存在。更为具体地, 由于已经证实水、氧气、适宜的温度是生命存在的基本条件, 如今发现火星上也有水和空气, 并且其气候也在一定的范围内变化, 基于结构上的相似性, 人们很自然会推测火星上可能也有生命存在。也就是说, 当我们发现了某个结构 $P_1(a_1) \wedge P_1(a_2) \wedge P_1(a_3)$, 很自然会联想到结构 $P_2(a_4) \wedge P_2(a_5) \wedge P_2(a_6)$ 可能也是存在的。这样的推理是合乎理性的, 因为两者在结构上是相似的。而捕捉这种由“结构相似性所支持”的直觉信念的一种形式化方法是“对等类比原则”。([17], 第 168 页)

对等类比原则 (CP) 令 $\theta, \theta' \in SL$, 并且使得 θ' 是通过某些在 θ 中并没有出现过的新常项或关系符号, 来取代 θ 中的某些常项或关系符号而到的结果。则有

$$w(\theta \mid \theta') \geq w(\theta)$$

根据 CP ，对于前述的火星上可能存在生命的类比，有

$$w(P_2(a_4) \wedge P_2(a_5) \wedge P_2(a_6) \mid P_1(a_1) \wedge P_1(a_2) \wedge P_1(a_3)) \geq w(P_2(a_4) \wedge P_2(a_5) \wedge P_2(a_6))$$

又例如，假设 R_1 是一元关系，而 R_2 和 R_3 是二元关系，令

$$\theta = \neg R_1(a_1) \wedge \exists x R_2(x, a_2) \quad \theta' = \neg R_1(a_3) \wedge \exists x R_3(x, a_2)$$

根据 CP ，则有

$$w(\neg R_1(a_1) \wedge \exists x R_2(x, a_2) \mid \neg R_1(a_3) \wedge \exists x R_3(x, a_2)) \geq w(\neg R_1(a_1) \wedge \exists x R_2(x, a_2))$$

第二个例子反应了帕里斯和温科斯卡在 2017 年所提出的多元谓词类比新构想。他们指出， CP 中的陈述 θ 和 θ' 所要捕捉的是结构相似性所蕴涵的一个重要意义：它们具有同样的句法形式，只是某些名称发生了改变（[18]，第 404 页）。

可见，多元归纳类比也是 PIL 体系下的类比推理的一个重要的发展方向。这种研究动态值得我们深入关注。因为前述的类比原则 SAP_S 刻画的是同一种类的事物之间（例如棕色和黄色）的相似性，而对等类比原则 CP 所刻画的是不同的事物之间的相似结构。相比而言，后者更接近日常生活意义上的“类比”。希尔更是把对等类比原则划入“句子之间的相似性”范畴。

3 形式化是归纳类比的发展趋势吗？

这首先要搞清楚什么是“形式化”，如果它是指全面符号化、逻辑化的话，那么，正如我们在前述的分析中所看作到的，类比原则 SAP_S 的适用范围是非常有限的——当语言 $\mathcal{L}$ 的关系谓词多于 3 个的时候，便没有任何函数能够同时满足 Ex ， Px ， SN 和 SAP_S 了。这对于类比的形式化计划来说无疑是一个重大的打击。

但是，如果我们把“形式化”的概念弱化，把它看作是一种以逻辑工具作为辅助手段的理性决策方法，那么，它的发展前景还是很广阔的。

第二种理解“形式化”的方式更符合帕里斯和温科斯卡的思想。正如我们在“纯粹归纳逻辑的总问题”一节内容中所揭示的，他们明确地把“逻辑地”弱化为“理性地”，即把逻辑问题还原为决策问题，大大削弱了对“真”的承诺。这是因为，决策问题往往并不涉及到“真”的问题，例如，某个公司招聘员工的时候所考虑就不是形式逻辑的“真”问题，而是新员工能否胜任工作的问题，或者说，能否为公司创造利润的问题。

类比推理的形式化发展和以简特纳 (D. Gentner) 为代表的结构主义类比理论所取得的成功密切不可分——后者在认知心理学、科学解释模型、人工智能类比推理等方面都取得了重要的进展，大大鼓舞了学界对类比推理研究的信心，并且推进了类比推理形式化的发展。

事实上,纯粹归纳逻辑框架下的类比推理和结构主义类比推理的理念基本上是一致的。帕里斯和温科斯卡在 2017 年发表的最新论文《整合类比支持到纯粹归纳逻辑中》的一个脚注中指出:“这个类比概念潜藏的基础类似于在人工智能(AI)中所利用的结构影射理论,参见 [6]。”([18], 第 403 页)

可见,在一定的意义上,我们可以说 PIL 体系下的类比体现了由结构主义类比所推动的人工智能类比的基本发展趋势。在《纯粹归纳逻辑》一书的前言中,帕里斯和温科斯卡更是明确地说:“我们希望这部著作能够吸引哲学家和人工智能共同体的兴趣。”([17], 第 6 页)从现实的观点来看,这是因为,在现代科技发展的背景下,人们的理性决策越来越多地涉及到人工智能的参与,而这对类比推理的理论发展提出了新的要求。正是在这个意义上,我们说,形式化是归纳类比的一种基本发展趋势。

参考文献

- [1] R. Carnap, 1945, “On inductive logic”, *Philosophy of Science*, **12(2)**: 72–97.
- [2] R. Carnap, 1950, *Logical Foundations of Probability*, Chicago: University of Chicago Press.
- [3] R. Carnap, 1980, “A basic system of inductive logic”, in R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, vol. II, pp. 7–155, Berkeley and Los Angeles: University of California Press.
- [4] F. A. D’Asaro, 2014, *Analogical Reasoning in Unary Inductive Logic*, MSc. dissertation, University of Manchester.
- [5] M. C. Di Maio, 1995, “Predictive probability and analogy by similarity in inductive logic”, *Erkenntnis*, **43(3)**: 369–394.
- [6] D. Gentner, 1983, “Structure-mapping: A theoretical framework for analogy”, *Cognitive Science*, **7(2)**: 155–170.
- [7] M. B. Hesse, 1964, “Analogy and confirmation theory”, *Philosophy of Science*, **31(4)**: 319–327.
- [8] A. Hill, 2013, *Reasoning by Analogy in Inductive Logic*, PhD dissertation, University of Manchester.
- [9] A. Hill and J. B. Paris, 2013, “An analogy principle in inductive logic”, *Annals of Pure and Applied Logic*, **164(12)**: 1293–1321.
- [10] R. C. Jeffrey, 1973, “Carnap’s inductive logic”, *Synthese*, **25(3–4)**: 299–306.
- [11] J. G. Kemeny, 1952, “Extension of the method of inductive logic”, *Philosophical Studies*, **3(3)**: 38–42.
- [12] J. Landes, J. B. Paris and A. Vencovska, 2010, “A characterization of the language invariant families satisfying spectrum exchangeability in polyadic inductive logic”, *Annals of Pure and Applied Logic*, **161(6)**: 800–811.

- [13] J. Landes, J. Paris and A. Vencovska, 2008, “Some aspects of polyadic inductive logic”, *Studia logica*, **90(1)**: 3–16.
- [14] J. Paris, 2012, “Pure inductive logic”, FEWS Tutorial, Munich 2012, http://fitelson.org/few/paris_notes.pdf.
- [15] J. B. Paris and P. Waterhouse, 2009, “Atom exchangeability and instantial relevance”, *Journal of Philosophical Logic*, **38(3)**: 313–332.
- [16] J. Paris and A. Vencovska, 2011, “A note on irrelevance in inductive logic”, *Journal of Philosophical Logic*, **40(3)**: 357–370.
- [17] J. Paris and A. Vencovska, 2015, *Pure Inductive Logic*, Cambridge: Cambridge University Press.
- [18] J. Paris and A. Vencovska, 2017, “Combining analogical support in pure inductive logic”, *Erkenntnis*, **82(2)**: 401–419.
- [19] H. Putnam, 1975, *Mathematics, Matter and Method*, Cambridge, MA: Harvard University Press.

(责任编辑: 罗心澄)

Analogical Reasoning in the Framework of Pure Induction Logic

Xianhua Liang

Abstract

Pure induction logic is a revival of the Carnap's inductive logic approach. Its basic goal is to study induction as a branch of mathematical logic. In the framework of pure induction logic, analogical reasoning is based on first-order logic, and its central idea is to use the distance function to measure similarity, and the relevance function is a tool of reasoning for it. Therefore, in a certain sense, analogical reasoning can be seen as an extension of relevance function. In practical, it reflects the need of artificial intelligence for the formalization of analogical reasoning.