

类比推理综述^{*}(上)

徐冬溶¹ 潘云鹤² 张 畅² 王 选¹

(北京大学计算机研究所 北京 100871)¹ (浙江大学人工智能研究所 杭州 310027)²

摘 要 类比推理是一种重要的思维手段,可以解决很多难题。本文详细论述了类比的基础,古今中外各种类比推理及其理论。全文分成上、下两部分发表。上部分论及类比在日常生活中的重大作用,类比的基础和类比推理概述;下部分论及类比能力的发育,类比推理理论及几个例子。

关键词 类比, 类比推理, 相似论, 类比推理理论。

0242

1 类比在日常生活中的重大作用

类比是人类的重要能力,几乎一切智能活动都涉及到类比。人们通过类比进行理解、推理、学习和概括,尤其是小孩子。当我们面对一个新问题时往往回想起一个以前的类似问题作参考来求解。美国数学家 G·波利亚就十分强调类比推理的作用[Pol 1984],在他看来,类比是引导问题求解的有效方法。研究表明,人的学习过程大部分是通过模仿形成的,尤其是在认识初期,而这种模仿,实质上就是一种行为的类比。美国著名学习理论专家班拉图,将学习分为内模仿和外模仿。他认为外在的宏观模仿行为必须建立在内在的微观基础之上,没有这样的生理基础的类比,单纯依靠人在几年或十几年的外在模仿是不可想象的。

科学发展的历史表明,往往是一个学科的发展,通过类比,带动和发展了大批相关学科,并建立起一系列新的学科,从而极大地推动社会的进步。电器时代的文明就是发源于对磁、电子器件、爱迪生效应与半导体整流效应的研究与类比。

人类语言中最具特色和最为生动的比喻正是类比的一个特例。由比喻衍生出来的隐喻、明喻、通感拟人等修辞手法是强有力的表达手段。根据 Gentner [1980, 1982, 1983] 对儿童心理的研究,比喻是一种注重欣赏性和高雅艺术性的类比,单独的一个比喻可以在不同的对象之间同时传递数种相似性。

哲学上长期讨论假设的来源,而类比推理正是假设的主要来源之一。实际上,许多代表人类智能最高成就的重大发现就起源于类比,例如威尔逊云室、莫尔斯的电报信号中继站的设想。

有意识地使用类比推理这种重要思维手段,可以解决很多难题;法官的类比断案,律师引证辩论,以及教学实践的举例讲解和程序设计自动化。

关于类比推理的研究还很不足,一方面与其固有特性有关,另一方面,是由于它们也依赖于其他相关方面的发展,例如自然语言理解,知识在记忆中的组织方式等。

2 类比的基础

2.1 相似论

类比作为一种重要的思维、学习方式和一种重要的推理手段,无论在客观世界还是在科学领域的哲学、心理学以及其他各个学科中,都具有深厚的基础——相似性。

在人自然中,大至宇宙星云,小至原子、质子,远至上古洪荒,近至当代卫星上天、提取恐龙基因片段,无不存在着大量的相似之处。生命是一切动植物、菌株的共同现象,都有遗传物质 DNA。1961 年英国的彼得·米切尔的实验和现代化学进一步证明,人和植物拥有一个共同的祖先,都是由核前生物体进化而来的[Zha 1992]。在人类发展历史上,几乎所有的民族都经历相同的社会变革过程,都进食为数不多的几种粮食维持生命,并发展了语言、文字记载先民的事迹和传说,并在此基础上产生了音乐。自然界中的光电磁,从直观上看不存在什么共同点,但光却具有波粒二象性,也是一种电磁波,电具有磁性,磁具有电本质,也可用来产生光,光电磁是统一的。进而从能量角度看,电磁与机械能、化学能也存在着联系,本质上具有相似的内容。

客观世界是普遍联系的,依据这种联系的紧密

^{*} 本文受到国家自然科学基金、国家“八五”攻关项目、国家“八六三”计划及攀登计划资助。

程度,人们通常用相似、有联系以至无关来刻画它们,实际上,即使无关的事物也存在着联系即相似,只是依据即时的观点,这种相似被忽略了,基于这样一种广泛的相似性联系,类比得以生存并占据了日常生活中的极为重要的地位。发明家依据相似进行类比,发明了蒸汽机、飞机;文学家的比喻,使得作品如泣如诉;即使普通人也会依据相似对经验进行类比来判定某天是否会下雨。

相似是自然界的一种特性,而类比则正是充分利用这种特性进行的有意识的思维活动,各个学科从不同侧面对类比作了研究。

哲学家研究假设的来源,已开始把目光转向了类比。科学哲学曾经作为一种科学成就的评价方法和选择方法而存在,但现在把科学的发现过程的类比原理视为其重要研究对象。

皮亚杰在他的发生认识论中,为了说明认识的发生及知识的形成与发展,提出了四个基本概念,即图式、同化、顺应及平衡,他认为,图式的同化、平衡、顺应的相互作用,实质上就是一个类比的过程,包括类比源的激活、类比对应的建立、类比迁移及类比经验扩充。

值得注意的是,由于自然的类比推理不存在演绎推理形式逻辑的严格限制和条件,而是一种宽松的推理方式,有关增加类比可靠性的条件变得十分深刻,并使得对类比的最终结果进行验证成为必不可少的步骤。

2.2 相似的特征

相似性是人们对事物进行分类,抽取概念和进行推广时使用的组织性原理,在人类的心理中,相似性无所不在,表现形式多样,例如匹配的几率、对象的分类团体之间的集体观念、替换的偏差或发生的协同,分析这些数据,可为观察到的相似性求得理解(解释)并获得研究对象的基本结构。

关于相似性关系的理论分析通常使用几何模型,并将对象表达为坐标空间的点,由此,对象之间的不相似性就对应为空间两个点之间的距离。虽然有些对象表现为层次性的树状结构,而不是代数空间,但关于接近程度数据的分析本质上是可变量的,因此,关于相似性的研究通常总是基于空间观念的。然而,存在许多对象并不适于使用空间维度表达。在这些对象之间考察相似性,比较特征常常要比将对象看作空间点来计算特征更加合适。

一个关于相似性的距离度量函数 δ ,通常被定义为一个非负的值,并满足以下三个条件:

$$\textcircled{1}\text{最小性:}\delta(a,b)\geq\delta(a,a)=0$$

$$\textcircled{2}\text{对称性:}\delta(a,b)=\delta(b,a)$$

$$\textcircled{3}\text{三角不等性:}\delta(a,b)+\delta(b,c)\geq\delta(a,c)$$

使用几何方法衡量相似性,这些要求必须被满足。基于这三条假设,使用几何方法描述相似性做了很多富有成果的工作,如[Cor 1974][She 1974],然而相似性的特征表明,这些约束通常并不总是合适的。

相似性的最小性原理要求一个对象是自相似的,但这并不总是成立。例如,在信号分析过程中,两个相同信号被认为“相同”或“不相同”的情况,是以一个概率形式出现的,并且这个概率也并不保持为一个常数,是一个变量。因此,一个对象很容易被误认,而不是作为它本身被接纳。如果将概率的因素考虑在变量过程以内,就会与最小性原则相违背,从而与距离模型不相容。

相似性的对称性原则要求相似关系是对称和可逆的,但,说“a象b”不一定可以说成“b象a”,相似性有时带有明显的方向性。a称为主体,b称为参照物,通常b具有较突出的特征,是一个较为显著的刺激信号,而a通常是一个较不显著的对象,它以b为原型。因此,在相似的判定中,a、b的位置是由其相对刺激强度决定的,这一位置关系的确定,实际上与相似判断的不对称有关,这种方向性在直喻和隐喻中表现最为明显。因此,有时两个方向的相似性都可以引用,但所表达的意义完全不同,Rosch[Ros 1975]认为这种不对称性源于信息过程中与众不同的原型的特殊作用。在这里,如果将“原型”理解为一种相对的含义,则Rosch与Tversky的观点就是一致的。

相似性的三角不等性表明,当a与b相似并且b与c相似时,a与c不会非常不相似,这与直观是不一致的,因为相似是不能传递的,例如:A与B有共性因为他们同班,B与C有共性因为他们同性,但A与C却既不同班也不同性,一点共性都没有。此虽不足以驳倒三角不等性原则,但以此作为相似性的基础是不恰当的。

此外,这三条原则牵涉到对与函数的定义以及对 $\geq$ 、+、-运算的规定。不同的规定对上述三条的适用程度也不完全一样。

这三条规则本身是很弱的。关于它们的进一步分析和讨论,有专门论述[Bea 1968][Kra 1975][Tve 1970]。

总的来讲,相似的特征是:①部分的自反性(最

小性),②非对称性,③非传递性。此外,相似性还具有上下文相关特性,依据不同的上下文,相似性的表现也不同[Tve77]。

2.3 相似与相异

一种公认的假设是,关于相似与相异的判断是互补的,即它们之间的关系是一条斜率为-1的直线。有一些研究证实了这个假设[Hos 1972][Tve 1970]。然而,这种相似与相异的反比关系并不总是成立。

一般地,共性衡量的增加会增加相似性的认定而削弱相异的判定,而相异性衡量权的加重则会削弱相似的认定而加强相异的判定。因此,在判定对象之间的相似和相异时,共性的作用是不一样的,反比关系不总是成立。换句话说,只有在判定相似和相异时对共性一视同仁,相似与相异的判定才是互补的。

用 $s(a,b)$ 表示 a,b 的相似判定值, $d(a,b)$ 表示 a,b 的相异判定值,则存在一个非负的函数 f 和常数 λ 和 q , 对于 a,b,c,e 有:

$$s(a,b) > s(c,e) \text{ iff}$$

$$\theta f(A \cap B) - f(A-B) - f(B-A) > \theta f(C \cap E) - f(C-E) - f(E-C) \text{ 和}$$

$$d(a,b) > d(c,e) \text{ iff}$$

$$f(A-B) + f(B-A) - \lambda f(A \cap B) > f(C-E) + f(E-C) - \lambda f(C \cap E)$$

其中 A,B,C,E 是与 a,b,c,e 分别对应的特性集,于是,一般必须要有 $\lambda=\theta$, 才有反比关系的成立。

研究表明,人在判定相似关系时,往往更多地关注共性,而在判定相异时,则往往忽略共性。于是在这两个过程中,共性的作用并不平等,在 $s(a,b)$ 中, θ 很大;在 $d(a,b)$ 中, λ 很小,于是得到:

$$\text{只要 } f(A \cap B) > f(C \cap E) \text{ 和 } f(A-B) + f(B-A) > f(C-E) + f(E-C)$$

就有 $s(a,b) > s(c,e)$ 和 $d(a,b) > d(c,e)$

这说明,一组存在较多共性同时具有较多相异性的对象比起一组存在较少共性和较少相异性的对象,在相似判定时会认为更相似,同时会在相异判定时被认为更相异(参见 Tversky 的实验)。

Tversky 的这个实验尚不足以体现著名程度对最终决策的作用,这是 Tversky 的一个失误,在我们看来,一个对象之所以著名,因为它被人熟知,即人们更多地了解其细节和属性,因而著名对象与非著名对象的属性框架内 slot 中值的数量是不一样的。其中,著名对象的 slot 值较多,而非著名对象的值较少。在 Tversky 的理论中,他默认了两种对象的值空

间是相同的。这是一个错误前提,因此,我们认为:

$$s(a,b) > s(c,e) \text{ iff}$$

$$F(a,b) \cdot (\theta f(A \cap B) - \alpha F(A-B) - \beta F(B-A)) > F(c,e) \cdot (\theta f(C \cap E) - \gamma f(C-E) - \epsilon f(E-C))$$

$$d(a,b) > d(c,e) \text{ iff}$$

$$F(a,b) \cdot (\alpha f(A-B) + \beta f(B-A) - \lambda f(A \cap B)) > F(c,e) \cdot (\gamma f(C-E) + \epsilon f(E-C) - \lambda f(C \cap E))$$

在 θ 很大, λ 很小时,只要

$$F(a,b)f(A \cap B) > F(c,e)f(C \cap E) \text{ 和}$$

$$F(a,b)(\alpha f(A-B) + \beta f(B-A)) > F(c,e)(\gamma f(C-E) + \epsilon f(E-C))$$

就有 $s(a,b) > s(c,e)$ 和 $d(a,b) > d(c,e)$

其中 $F(X,Y)$ 是对 X,Y 著名度的衡量, $\alpha, \beta, \gamma, \epsilon$ 体现了相似关系的非对称性。

当 $\alpha=\beta=\gamma=\epsilon$ 时,并不失一般地取为 1.0, $F(a,b)=F(c,e)$ 时,上式还原为 Tversky 的表达。

相似性与类比推理的关系不言而喻,而值得注意的是,对象间存在的差异对于类比也很重要,不相似性为类比的结果增添新的特色提供了可能。著名物理学家汤川秀树教授指出:“当两个事物之间的不同性和相似性都被认识得很清楚时,类比思维就会变得更加富有成果。”[Tan 1987]

2.4 相似性度量

关于相似性度量,可有几何性度量和特征集合方法。如前所述,几何方法常常与现实的相似性评判相违背,本节着重讨论后一类方法。Tversky 提出的相似度的属性计算理论具有代表性。

$\Delta = \{a, b, c, \dots\}$ 表示研究对象域,假定 Δ 中的个体都是用其属性来表征的, A, B, C 等分别对应于 a, b, c 等,表示其特征集合。这些特征对应各个成分,表示具体内容,从而间接表达出质量、复杂度等一些抽象信息。使用属性作为研究对象的手段是常用的 [Jak 1961], [Nei 1967], [Gib 1969], [Smi 1974]。

首先,特定对象涉及的数据总体具有丰富的内容和复杂的形式,需要抽取、收集有关知识、特征。因而,一个特征属性集合可以看作是这个先期抽取和收集工作的产品。其次,“特征”常用作二值性变量,例如有值或无值,有声或无声,但常需作为一种连续性或至少是维度性的(有序的数列,基数性的)变量。这样一种表达,与一个非空间性的单维定向结构是相似的。这样结构可以被表达为一种互相重叠的链,

或者是树形结构。Restle 对于集合理论表达有过专门论述[Res 1959]。

用 $s(a, b)$ 表示“a 与 b 相似的程度”，其中，a, b 表示 Δ 集中的两个元素，则 $s(a, b) > s(c, d)$ 表示“a 比 c 象 d 更象 b”。相似性度量应满足以下三个基本假设。

①匹配性(matching): $s(a, b) = F(A \cap B, A - B, B - A)$

②单调性: 如果 $A \cap B \supset A \cap C, A - B \subset A - C$ 和 $B - A \subset C - A$, 则 $s(a, b) \geq s(a, c)$

函数 F 被称为匹配函数, 它衡量两个被看作特征集合体的对象的匹配程度, 因此, 衡量相似不再采用几何的空间距离方法, 而以匹配函数的集合理论计算。

为了确定匹配函数形式, 需引入一些新原则和前提。

Φ 表示与 Δ 中元素有关的一切特征, 用 X, Y, Z ... 表示 Φ 的子集, 是一些特征集合。

$F(X, Y, Z)$ 是有定义的, 如果 $\exists a, b \in \Delta, a \cap B = X, A - B = Y, B - A = Z$, 由此 $s(a, b) = F(A \cap B, A - B, B - A) = F(X, Y, Z)$

定义 $V \cong W$, 对 $\forall V, W \subset \Phi$, 如果 $\exists X, Y, Z \subset \Phi$, 并且下式之一成立: (I) $F(V, Y, Z) = F(W, Y, Z)$; (II) $F(X, V, Z) = F(X, W, Z)$; (III) $F(X, Y, V) = F(X, Y, W)$ 。

如果 $(A \cap B) \cong (C \cap D), (A - B) \cong (C - D), (B - A) \cong (D - C)$ 这三式中, 有一式或二式或三式成立, 则称对偶 (a, b) 和 (c, d) 分别是一次一致, 二次一致或三次一致的。

③独立性: 假设 (a, b) 与 (c, d), (a', b') 与 (c', d') 都是二次一致的, 并且, (a, b) 与 (a', b'), (c, d) 与 (c', d') 在第三个成分上是一次一致的, 则:

$$s(a, b) \geq s(a', b') \text{ 与 } s(c, d) \geq s(c', d')$$

这条公理的含义是, 用 S 作度量, F 有三个变量, 其中任意两个变量的联合作用是独立于第三个变量的。因此, (a, b) 与 (a', b'), (c, d) 与 (c', d') 虽是一次一致的, 其间的关系存在一定的相似, 但它们之间的相似程度比较是保证统一步调的。这一点, 由 (a, b) 与 (c, d), (a', b') 与 (c', d') 之间的横向的二次一致性维系。因此在两个方向上的一致性互相独立的, 其中的一次一致性于此间的二次一致性不发挥作用。例如图 1 中:

$$s(a, b) = F(\text{圆下巴, 笑, 直眉}) \geq F(\text{尖下巴, 生气, 直眉}) = s(a', b') \text{ iff}$$

$$s(c, d) = F(\text{圆下巴, 笑, 曲眉}) \geq F(\text{尖下巴, 生气, 曲眉}) = s(c', d')$$

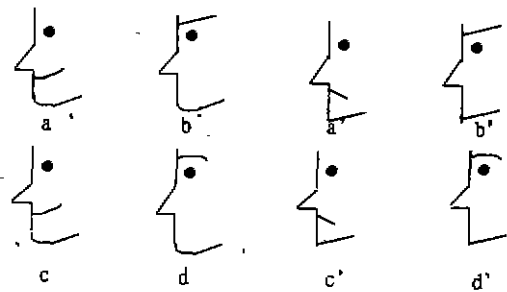


图 1 一个例子

值得指出的是, 每当对公理进行一个检验, 关于特征的一个预定的解释就给定了, 因而完全可能在不同的解释下公理 3 成立与否不同。因此公理 3 是给出一个合理解释的约束, 一般选择符合这一约束的解释作为相似度量标准。

对于形象性的对象, 为判定其相似性, 还必须引入一些更加全局性、总体性的约束, 例如对称性, 连接性, 复杂度, 线型, 面积等等。Tversky 在三条公理之外, 还提出了一些原则, 它们是可求解性和不变性。可求解性要求研究对象的属性足够丰富, 相似性判断能够进行。它包含一个中值约束和一个负变元约束, 这样, 该域的相似性方程可解[Tve 1977]。不变性要求属性的值区间之间, 在 F 的三个变量位置上发挥等价性的作用, 即这些区间的值, 对 F 的权是一致的。这条性质建立起属性集合中不同属性取值区间的对应性, 进而要求即使是同一属性的取值在 F 的三个变量位置上也保持一样的稠密度。

在此基础上, 定义 $s(a, b) = \theta f(A \cap B) - \alpha f(A - B) - \beta f(B - A)$, 且 $s(a, b) \geq s(c, d)$ iff $s(a, b) \geq s(c, d)$, $\alpha, \beta, \theta > 0$, 这就是所谓的相似性衡量的“相反性模型”, 因为相似的度量同时考虑了相同性特征和相异性特征, 并且, 由 α, β, θ 来调节其作用强度。

当 $\theta = 1, \alpha = \beta = 0$, 则 $s(a, b) = f(A \cap B)$, 以共性作为衡量标准; 当 $\theta = 0, \alpha + \beta = 1$ 时, $-s(a, b) = f(A - B) + f(B - A)$, 对称的差异性作为衡量标准。

在这里, f 是一个反映特征显著性的函数, 衡量指定特征对相似性的贡献。f 综合考虑对象的显著性, 比如强度, 频率, 熟悉度, 形态好坏和信息内容。 θ, α, β 的取值决定于相似性衡量的上下文要求。相反性模型主要就是通过 θ, α, β 的相反性规则构建的。

Tversky 的这一理论在形式上十分简洁, 也很

直观,但并不是唯一可用的方法。事实上,由于对象的复杂特性,这一理论的可操作性有待加强,实用性也需加强。从本质上讲,这一理论与冯·诺依曼的冒险决策的经典理论[Neu 1947]是类似的。或许正是由该理论启发得到的,但它们也存在显著不同。

关于相似性度量的模型,还有一种比率模型[Gre 1975][Sjo 1972],定义 $s(a,b)$ 为

$$s(a,b) = \frac{f(A \cap B)}{f(A \cap B) + \alpha f(A - B) + \beta f(B - A)},$$

$$(\alpha, \beta \geq 0)$$

这样, $s(a,b)$ 被归一化为 0~1 之间的值,当 $\alpha = \beta = 1$ 时,

$$s(a,b) = \frac{f(A \cap B)}{f(A \cup B)}$$

当 $\alpha = \beta = 1/2$ 时,

$$s(a,b) = \frac{2f(A \cap B)}{f(A) + f(B)} \quad (\text{参见[Eis 1959]})$$

当 $\alpha = 1, \beta = 0$ 时,

$$s(a,b) = \frac{f(A \cap B)}{f(A)} \quad (\text{参见[Bus 1951]})$$

总的来说,Tversky 的理论扬弃了几何模型下相似性度量的优缺点,提出了一个更广泛的理论衡量方法。但几条公理太强,只适于那些特征明确的对象,对于广泛存在的不确定因素以及未知特性的情况尚缺乏考虑。

许多研究者使用自己定义的相似度量,与 Tversky 的工作没有必然联系。例如, Kling [Kli 1971], [Win 1978, 1980]。

3 类比推理概述

3.1 类比研究发展历史

有关类比的研究已经在多个领域展开,最早的研究工作见于心理学和哲学,并且是与学习有关的。在哲学领域,可以追溯到 Aristotle 和 Plato [Bur 1973]。时间近一些的,有 Black, Peirce, Hesse, Wylie, Burrell, McNery, 和 Boss 等人的工作 [Bur 1973] [Hes 1966] [Wyl 1985] [McI 1968] [Ros 1981] [Dun 1926] [Bla 1962] [Pei 1934], 他们对 Aristotle 等人的局限作了讨论并研究了类比与哲学语言、类比与隐喻 (metaphor) 的关系,给出类比的总体描述。在心理学领域的工作,有 Duncker, Dorolle, Herbert 等 [Dun 1926] [Dor 1949] [Her 1898], 着重研究类比在认识过程中的作用以及类比的心理过程。在计算机领域围绕人工智能进行的类比推理研究掀起了最近一次的类比研究热潮, Gentner, Holyoak, Carbonell, Gick, Keane, Rumelhart, Winston 等人的研究为类比推理的研究定下了基调,并在极大程度上

为近二十年的类比研究制定了研究方向,他们中的一部分人是从心理学角度,例如 Gentner 等人,从研究儿童学习过程来看待类比推理的。所有这些研究成果是认知科学和心理学结合起来的产物 [Kea 1988], 使得用计算机模拟智能的类比行为成为可能,他们提出了一大批类比推理理论与模型,例如 Gentner 的 SME, Winston 的类比理论, Holyoak & Thagard 的 ACME 等,并实现了一些实验性系统。

1968 年 Evans 提出了第一个类比模型即 $A::B::C::?$, 比如在交通灯问题上, $Red::Stop::Green::?$, 用以解答一类几何图形变化问题的智力测验题,这个模型具有严格的形式并局限于特定的狭窄领域以内,与类比所进行的复杂而广阔的应用背景还具有很大距离,但这个模型对以后的类比推理研究起了启发性的作用。1971 年 Kling [Kli 1971] 提出了第一个类比推理样板,用以辅助代数定理证明。机器类比的真正确立,始于 Winston 的工作,他指出了因果关系在类比中的作用 [Win 1986], Carbonell 在 1981, 1983 年提出了类比的转换方法,将已知的问题转换成类似的问题解法 [Car 1981, 1983a]。从此,关于机器类比推理的研究迅速兴旺起来。

我国学者在这方面也进行了卓有成效的研究工作,取得了很大进展,南京大学计算机系从自动程序设计出发,系统地研究了类似关系和类比推理,分别从认识论和方法论角度给出了类似的两个形式定义并证明了其等价性,并进一步研究类似的序、归约树的性质 [YiB 1990]。北京航空航天大学计算机系的研究将相似性划分为语义相似、结构相似和目标相似,从知识表达角度在语义层、具体层和抽象层上研究了类比的性质,实现了一个自动程序设计系统 APA [LiB 1993]。中国科技大学计算机系的研究者提出了一种基于联想的框架类比方法,实现了一个从命题到图形逻辑的类比学习系统 GAP,“填补了图论中图的逻辑推理的空白” [Cai 1989]。浙江大学计算机系的研究从智能 CAD 出发,对类比推理作了发展,将相似性扩充为一种广义对应性,提出了类比生成的概念 [XuD 1993, 95a] 和综合推理的理论 [Pan 1994], 并在此基础上实现了一个广告设计 CAD 系统 ADS-2.0 [XuD 1995d]。

3.2 类比推理的分类

根据 Halle 的研究 [Hall 1989], 相似性观点可以分为三类,即语法相似 (或称结构相似), 语义相似和语境相似,依此类比对应也分为三类: 结构对应、语义 (类) 对应和语境对应。

结构对应要求保持基与靶之间的结构上的一致,如 Gentner 的结构映射(SM)理论[Gen 1983]。她将对应限制在关系同名及关系间的同名关系即高阶关系对应上。Haraguchi[Har 1986,87]用 Herband 模型[Emd 1976][Apt 1982]间的部分等价关系定义两个逻辑程序[Kow 1979]之间的对应,也是一种结构对应关系,实质上是对谓词的同名和函数的同名建立对应。Pötschke[Pöt 1986]用标号图的同构关系和最大同构子图用为类比对应,面向一类能用标号图表示的对象的结构类比问题,但由于标号图表达能力的欠缺和图同构算法复杂性[Aho 1976]的限制,应用前景狭窄。

语义(类)对应基于语义模板和语义类的预定义,其通用性有限,人工痕迹明显。例如 Evans[Eva 68]预定义了映射各部分的一一对应,Kling[Kli 1971]则定义了一组谓词的语义模板,还规定了各模板之间的序,作为相似性度量。Winston[Win 1978, 1980]预先将对应分成几类,规定匹配只能在同类中进行,采用简单的属性重叠法测量相似性,求取最佳匹配。

语境相关与上下文相关,因而弹性很大,是一个较为模糊的概念,在理解隐喻(Metaphor)[Hob 1983]和在交互环境中理解类比[Bur 1986]这样一些不适于形式化的领域应用较多。

与这三种类似观点相应,Keane 提出了一种二分法[Kea 1988]即语法性(Syntactic)类比和实用性类比(Fragmatic),他实际上将结构性相似对应归入语法性类比,将语境相关归入实用性类比,而将部分语义(类)对应看作是这两类的融合,因为语义类相似包含了预定的结构,但在相似匹配中又考虑了对象的实际内涵,如重要性。他将 Gentner 的 SM 归入语法类,将 Burstein[Bur 1973,86],Carbonell[Car 1981, 82, 83a],Holyoak[Hol 1985],Hollandt[Hol 1986]以及 Keane[Kea 1985,87,88],Kedar-Cabelli[Ked 1985a]的工作归入后一类,因为他们都一致强调规划、目标和类比目的的重要作用。首先提出“实用性途径”这一概念的是 Holyoak 和他的合作者[Hol 1985a],在 Keane 看来,分立的两类方法都不能很好地解决问题,必须将两者有机地融合起来, Gentner 对 SME 的修正[Gen 1989b]以及后来的 Holyoak 和 Thagard 的 ACME[Hol 1989]正代表了这样一种趋向。

Carbonell 从推理方法入手,将类比推理分为两大类[Car 1986]:转换型(Transformational)类比推理和推导型(Derivational)类比推理。

转换型方法简单地根据类比对应将基域(Base Domain)中的一些性质、关系等直接对应到靶域(Target Domain)中作为类推结果,并在靶域中验证推论正确性。这个过程通常缺乏足够充分的理由,正确性不易保证,通常只能用于较简单的直觉推论,例如 Greiner[Gre 1988]预先将公式进行组合,使类比在一个公式上进行,从而降低对验证的需求,而这种预先公式化并组合的要求,对一般问题来讲是一种强条件。Gentner 的结构映射理论[Gen 1980,83]也归于这类转换型方法,为了较好地保证结论的正确性,她为其理论附加了一个“系统性原则”,规定只有与整个系统的总体特性不相违背的对象(关系、属性、对象)才能被映射到靶中。Winston 通过语义网的因果链将基中存在因果关系的结点映射到靶中,实现语义网的扩张和推理[Win 1978,80,82],Pötschke 根据标号图的同构关系进行其余部分的同构扩张,也都是这种直接的转换方法。Carbonell[Car 1983]也是将问题求解规划直接转移到靶中,再加上以必要的转换操作进行修改。

推导型类比与转换型方法不同,通过转换基中解的推导过程和决策点的方法,在靶中建立可行的推理途径,求得靶问题的解,因而目标解的合法性被分解为目标途径合法性的验证,从基础上对结果进行了维护。这些工作是分步进行的,其搜索空间减小,难度降低,开销较小。这一类型的方法主要包括 Kling[Kli 1971]的类比辅助定理证明系统 ZORBA-1, Ulrich 和 Moll[Ulrich 1977]的类比综合程序等。Carbonell 在其转换型工作的基础上,提出了推导的类比方法,指出要在靶中重构一个推演过程是相当复杂的,尤其是对于复杂系统,Carbonell 给出了一系列的工作步骤,但未能一致化和形式化。Scoff 和 Scherlis 的工作[Sch 1984][Die 1987],研究了程序开发过程的类比;提出程序开发的产品包括程序和积累经验观点,而这些知识和经验却只能进行非形式的描述性表示[Sch 1984][Die 1987]。

伊波考察了各种类比推理形式,认为根据不同的论域和着眼点以及不同的“相似型”,可以归结为四种形式,即属性类比、射类比、结构类比和扩张类比。具体细节和基本形式可参见[YiB 1989]和[XU 1995b]。

属性类比的推理理由十分有限,但存在某种进一步的限制时,可能例外。其相似识别是最简单的涵盖法,即 $|p(a) \cap p(b)|$ 。将所有属性一视同仁,没有重要性区分,但导出结论一般可信度较高,而效率较低,为了减小搜索空间,须采用额外的启发式规则。Winston 的重要性原则和预分类方法已经结合了这

样的规则,有效抑制了组合爆炸。

最早的类比程序 ANALOGY[Eva 1968]就是处理射类比的,用以解答一类智力测验图。射的相似度由图形状态中对应子图形的 Add, Delete 和 Trans 操作定义,以此求得靶的图形转换函数。Kling 的定理证明[Kli 1971],Ulrich 和 Moril 对类比过程和程序过程综合的推导的结合[Ulir 1977],以及 Carbonell 的问题求解系统[Car 1981, 1983a],Pötschte 的有向图类比都是射类比的例子[Pöt 1986]。

结构类比方法与前几种类比的结构化方法观点基本是一致的,Gentner 的结构映射理论是一个典型代表。此外还有张光鉴、张铁声的类比认知模型所提出的图式同构[Zha 1986, 92],以及 Haraguchi 的 Herbrand 域上的两个逻辑程序间的类比对[Har 1986, 87]。

扩张类比是由一些元素间的相似扩张为系统间的相似性,可以看成结构类比的一种逆向行为。这一类的研究尚且不多,Greiner 的“有用的类比”(Useful Analogy)部分地处理了这一问题[Gre 1988a]。

我们的研究认为,直到目前为止几乎所有的类比推理工作都是一种求解性的或求知性的工作。在给定的相似性条件约束下,依据这样的相似轨迹,为待映射部分搜寻一个至少看上去是有理由的映射象,因而已有的绝大部分工作都可以归为求解过程,它们缺少生成性的能力,即产生出原本不存在的内容的能力,而这种生成性,对于设计活动是极为重要的。另一方面,这种类比实质上可以理解为是对域间对象相似关系、域内对象映射关系以及基和靶内映射之间关系的一种关系类比过程,即高阶关系类比。从这两点出发,面向设计活动,我们提出了类比生成的概念。此时,传统的类比化为类比生成的一个特例[XuD 1995a]。

根据诸葛海等人的研究,类比推理可以划分为推定型和求解型两类[Zhu 1992]。推定型是指根据已知的前提,仿照逻辑演绎推理方式推出结论,求解型则是根据已知问题或已知问题的求解方法来求解新问题,其中推定型又分为保真型和非保真型。而求解型则又有两种方法即解映射和推导过程映射。

根据诸葛海的定义和对推定型的描述:

(A)保真型基本模式:

$$\begin{aligned} (1) O_1 \vdash O_2 \\ (2) \exists O_1', O_1' \in [O_1]_s \\ O_1' \vdash O_2', O_2' \in [O_2]_s \end{aligned}$$

其中 O_1 与 O_1' , O_2 与 O_2' 分别同构。

(B)非保真型基本模式:

$$(1) O_1 \vdash O_2$$

$$\begin{aligned} (2) \exists O_1' \text{ 且 } O_1' \in [O_1]_s \\ O_1' \vdash O_2', ACF, O_2' \in [O_2]_s \end{aligned}$$

其中,ACF 为类比结论确定程度, $ACF = f(SD_1(O_1, O_1'), SD_2(O_2, O_2'))$, $SD(O_1, O_2)$ 为 O_1 和 O_2 的相似程度。

可以发现,推定型实质上是在同构与整体推导之间进行类比推理,而求解型则是在问题描述与解之间进行类比推理,因而,抽象地看,都是在整个问题的一部分相似关系(如同构对应描述)和另一部分未知对应(如 $O_1 \vdash O_2'$, 问题解答)之间进行类比推理。所以本质上,这两种推理都是求解型的。其所谓的“求解型”类比,实质上是 Carbonell 二分类法中转换型和推导型的一个并集。但在诸葛海关于规则格和规则格同态一节讨论中给出了下面命题和推论:

命题: (1) 如果 $O_1 \vdash O_2, O_1' \vdash O_2'$, 则 $O_1 \wedge O_1' \vdash O_2 \wedge O_2'$; (2) 如果 $O_1 \vdash O_2, O_1' \vdash O_2'$, 则 $O_1 \vee O_1' \vdash O_2 \vee O_2'$ 。

推论: $\langle U', \leq \rangle$ 为格, 对于 $O_i \in U' (i=1, 2, 3)$, 以下各式成立:

$$(1) \text{ 如果 } O_1 \vdash O_2, O_1 \vdash O_3, \text{ 则 } O_1 \vdash O_2 \vee O_3;$$

$$(1) \text{ 如果 } O_1 \vdash O_2, O_3 \vdash O_2, \text{ 则 } O_1 \wedge O_3 \vdash O_2.$$

由此可见,他的类比推理实际上已包含一种基于符号的类比生成因素,由两个前提得出一条结论,这条结论不包含在任一条前提中, $A \vee B$ 可以看作是 A 和 B 的一种生成性的中间状态。Qian[Qia 1992]也认识到了类比过程中,两个源之间区域的重要性,提出了 Between-domain Analogy 的概念。这些,与我们的工作不谋而合。

3.3 类比推理的内容

整个类比推理的过程,通常可以划分为四个阶段:1)类比源的检索。2)建立对应关系及映射。3)推理和方案转换。4)结果学习。对应这四个阶段,类比推理的关键在于:

1)如何有效地组织类比源的存储和检索到一个合适的和有用的类比源,涉及知识表达、检索策略、相似度计算等一些问题。

2)如何分解对象并在其间建立对应,涉及避免组合爆炸、启发式匹配、相似计算、系统性保持、以及映射透明度等问题。

3)如何将解答转换到目标中,并保持其合理性。

4)如何更新、淘汰知识并维护系统相容性。

所有的类比研究工作都围绕以上这些问题展开,较多地集中于有关第二阶段的内容;第一阶段也有涉及;第三阶段的内容在一些实验系统中被局部地研究到;第四部分主要与知识工程的研究相关,专门著述不多。(未完待续)