你这是什么逻辑? | 王培AI专栏

原创 2016-08-11 王培 赛先生



尽管神经元网络对很多专门问题的确非常有效，但仍未解决通用智能系统的核心问题。因为 “思维” （或称其为 “智能” 、 “认知” 等等）不能被简单看成用数据拟合一个函数。本文试图说明，为什么仍有理由认为思维遵循某种 “逻辑” ，但这种逻辑不是 “经典逻辑” 。经典逻辑并没有错，但是在知识和资源不足的情况下不能用，而非公理化逻辑则正体现了在这种情况下的理性原则。

王培（美国天普大学计算机与信息科学系）

我在《人工智能迷途：计算机的高技能等于高智能吗？》中介绍了“通用人工智能”的概念，并在《计算机会有超人的智能吗？》中提到了我设计的“纳思”系统[1]。这个系统的全名是“非公理化推理系统”（Non-Axiomatic Reasoning System，NARS）。顾名思义，我试图构建一个遵从逻辑规则的计算机推理系统，以此实现通用人工智能。

有些对人工智能有了解的读者看到这里已经开始摇头了：“基于逻辑的人工智能不是条死胡同吗？现在最有希望实现人工智能的技术不是深度神经元网络吗？”我在《机器是如何被骗并骗人的？》中已经简单说明了我对神经元网络的看法：这项技术对很多专门问题的确非常有效，但仍未解决通用智能系统的核心问题。**“思维”（或称其为“智能”、“认知”等等）不能被简单看成用数据拟合一个函数。**

当然，神经元网络有局限性不说明逻辑就没有局限性。实际上，基于逻辑的人工智能所遭到的批评要多得多。下面我们得从头说起。

“逻辑”一般是指正确的推理规则，而也常被引申来指正确的思维规则或论证规则。“首先总结人的思维规律，然后据此构建计算机系统”听起来像是个实现“思维机器”的合理计划。这里的主要难点在第一步，因为对什么是“人类的逻辑”，学者们仍然是众说纷纭，甚至思维是否真有“逻辑”都是个有争议的话题。这个论题在逻辑学、心理学、哲学等领域已经讨论了两千多年了。一方面，有理由认为人类思维的确有某些普适规则，否则我们完全不可能互相理解或彼此说服。另一方面，对这种规则的严格刻画仅仅获得了有限的成功，而大量思维活动仍处于“无逻辑可言”的境地。

就其内容而言，一个“逻辑”系统主要包括：

1) 关于**语言**的规定，比如一套语法规则。只有满足要求的语句才可以被处理。

2) 关于**语义**的规定，即指出如何在原则上确定语词的意义和语句的真值。

3) 关于**规则**的规定，即指出如何正确地从前提语句组合中推出结论语句。

在历史上影响最大的逻辑系统是亚里士多德建立的**传统逻辑**和弗雷格、罗素等建立的**数理逻辑**。这些 “经典” 逻辑系统在很多领域中取得了伟大的成就，但在解释人类思维时仍有明显的局限性，包括下列广为人知的问题：

· **不确定性的表示与处理**：经典逻辑中的语句非真即假，而人类知识中往往不可能完全消除各种不确定性，尤其是预测、解释、概括、分类等推理活动似乎无法得到完全确定的结论。

· **非演绎推理的合法性**：人类思维在很大程度上依赖于归纳、归因、类比等推理类型，但它们都可能从真前提中推出假结论，即不具有传统意义下的 “保真性” ，所以在经典逻辑中均不合法。

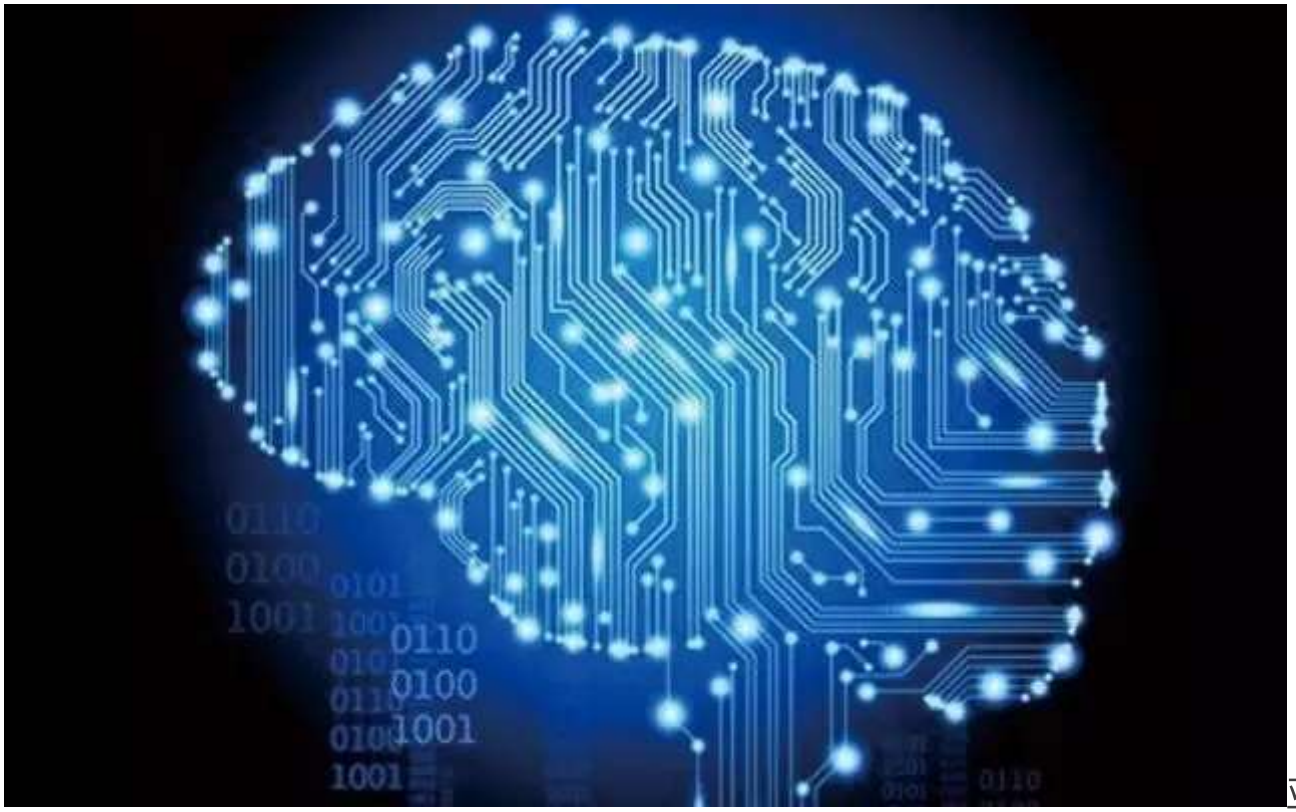
· **可修正的常识推理**：那些一般有效，但有例外的 “常识” 常常在推理中被作为真前提来使用。在认识到例外发生后，这种推理的一些结论需要修正。比如看到一只鸟，人们会依据常识认为它会飞，直到获得相反的证据。

研究者们对这些问题的态度大致可分为三类。在人工智能的背景下，其典型表现形式分别是：

1. 保守主义：“人的思维不符合经典逻辑那是因为人脑没进化好。人工智能应该遵循经典逻辑，而不是模仿人的坏榜样。”
2. 修正主义：“人的思维不符合经典逻辑那是因为经典逻辑没设计对。人工智能应该遵循一种接近人的思维现实的新逻辑。”
3. 激进主义：“人的思维不符合经典逻辑那是因为思维根本不符合任何逻辑。人工智能应该遵循一种“非逻辑”方案，比如神经元网络。”

本文的目的就是要说明为什么仍有理由认为思维遵循某种“逻辑”（所以激进主义是错的），但这种逻辑不是“经典逻辑”（所以保守主义也是错的）。其实持这种修正主义想法的人不少，也不限于人工智能领域。在逻辑学界已经有多种“非经典逻辑”被提出来了，包括多值逻辑、模糊逻辑、概率逻辑、归纳逻辑、非单调逻辑、相关逻辑、超协调逻辑等等。这些逻辑系统的一个共同特点是试图在某一方面改造经典逻辑以符合现实推理活动的要求，而同时在其它方面仍遵循经典逻辑的传统（见参考资料[2]）。这些“保守的修正主义”方案各自都取得了一定的成功，但如果我们的目的是给推理活动以至于思维活动提供一个统一的规范性理论，那么这些方案就是“头疼医头，脚疼医脚”了，即对“病源”的挖掘还不够深。如果前述问题都是同一个“疾病”的不同“症状”呢？那我们就该需要一种“激进的修正主义”了。这正是我的立场。

如我在《人工智能：何为“智”？》中说明的，有理由把智能看成一种在现实的知识和资源的约束下发挥的高级适应能力。这就要求智能系统必须是“开放”的，即新知识和新问题可以在任何时刻出现，且新知识可能和系统已有的知识相冲突，而新问题也可能超出系统已有的知识范围或时空资源限度。不难看出，**我们的日常思维活动常常是在这种情况下进行的，而我们也希望人工智能系统能在这种条件下工作。但问题是这种系统中的推理总体上无法遵循经典逻辑。**



亚里士多德逻辑仍是以一般人类推理活动为关注对象的，尽管它只是把握了一种特定的（但也是常用的）推理形式：三段论式的二值演绎推理。到了数理逻辑，关注对象已经从一般推理活动变为数学定理证明了。为了保证逻辑学的规范性不受思维中的偶然因素所影响，弗雷格等曾极力批判 “心理主义” 对逻辑学的影响。同样是研究推理活动，心理学的角色像是记者，其任务是忠实地描述这一活动的各种形式，而逻辑学的角色则像是法官，其任务是严格的评判这些形式是否正确。这两者当然是有本质不同的。

在数学中，推理的正确性（逻辑有效性）是有严格标准的，其最典型的形式就是定理的证明过程。这一过程从一组其真值不受怀疑的公理出发，靠推理规则的保真性来逐步证明各个定理。这一过程的可靠性保证了数学中定理的权威地位。但数学之外怎么办呢？传统的办法是把相对可靠的知识（ “事实” ）权当 “公理” 来用，然后靠推理规则的保真性获得大量具有同等可靠性的导出结论。这个 “公理化方法” 的另一个名称是 “假说演绎法” ，其中的前提被当作假说，而靠演绎得到结论。不管怎么称呼，这类方法使得数理逻辑在数学之外也得到了广泛的应用。

一旦推理系统的知识和资源是相对不足的，其中的推理就不再能被看作公理系统中的定理证明，因为系统中没有其真值完全免受未来经验挑战的 “公理” 作为起点。即使某些知识足够可靠，仅靠它们也往往不足以为所有问题提供可靠的答复，更不要说所有答复都还有时间要求。因为系统的加工和存储能力都是有限的，即使对已有的知识也会有 “没记住” 和 “没想到” 的情况。逻辑学界对这类挑战的传统回应是说解决这些问题超出了逻辑学的范围，因为逻辑学只关注封闭系统中的推理问题，且不考虑资源约束。但在现实情况下真的没有 “逻辑” 可言吗？在没有办法保证结论绝对正确的情况下，真的是随便怎么说都一样有理了吗？

我在参考资料[1] 中描述了一个 “非公理化逻辑” NAL（译成 “娜拉” 怎样？），其主要目标就是为智能系统（包括人类和计算机）中的推理活动建立一个统一的规范性模型。我试图以适应性原则为基础定义一种 “相对理性” ，即在绝对意义下的 “正确” 不可得的情况下， “合理的” 的答案就是系统在现实资源约束下所能找到的与以往经验一致程度最高的答案。在这里取代 “公理” 或 “事实” 成为真理标准的是系统的 “经验” ，而经验是随时间积累的，因此不是一个恒定不变的基准。如我在《证实、证伪、证明、证据：何以为 “证” ？》中所介绍的，在非公理推理系统 “纳思” 中，一个陈述的真值是对已有证据的度量。在知识不足的情况下，真值中的不确定性是不可避免的，而推理规则的 “保真性” 不再是保证推理结论与公理或事实完全一致，而是保证推理结论和前提的关系符合这种基于证据的真值定义。在《 “意思” 是什么意思？》中，我说明了在纳思中一个语词或概念的意义不是其定义或指称的事物，而是和其它语词或概念在经验基础上所建立的联系，因此会随着经验积累和情境变化而改变。

和上述语义学原则相配合，NAL的语法规则和推理规则属于“词项逻辑”，即更接近亚里士多德逻辑而非数理逻辑。具体说来，词项逻辑中的典型陈述是 “主谓句” 形式的，如 “甲是乙” ；其中的典型推理规则是 “三段论” 形式的，如从 “甲是乙” 和 “乙是丙” 中推出 “甲是丙” 。在这种逻辑中，一个典型陈述表达了两个概念间的可替代关系（ “是” ），而一个的典型推理规则表达了一种从已有的概念替代关系中建立新的替代关系的方式。作为一种多值逻辑，NAL中的每个陈述依照《证实、证伪、证明、证据：何以为 “证” ?》中的定义得到一个真值，以刻画经验对其支持（或反对）的程度。每个推理规则根据前提（及其真值）确定结论（及其真值）。

这篇短文自然无法描述NAL的细节（见参考资料[1]），而只是指出有这么一种 “逻辑” 存在，并例举其一些有趣的特征。



首先，说NAL是个逻辑，是因为其目的是将 “合理的推理” 规范化、形式化和计算机化。**这里以 “推理” 涵盖思维活动的各种基本步骤。**在目前的纳思系统中，NAL不仅统一实现了多种传统意义下的 “推理” （如演绎、归纳、归因、比较、类推、修正、选择、抽象、预测、解释等），而且统一解释了很多传统上不被看作 “推理” 的认知功能或现象（如学习、规划、决策、感知、行动、通讯、记忆、注意、灵感、直觉、想象、情感、游戏、审美等）。上述功能和现象以后会逐步介绍，而这里只是讨论和“逻辑”这个概念关系最密切的问题。

说NAL是 “规范性” 的，是因为上述功能都是建立在一个统一的 “相对理性” 的基础之上的。因为这个理论预设比其它规范性推理模型（如数理逻辑和概率论）的预设更接近于人的现实，所以纳思的表现和人的思维活动更加接近，尽管其设计目标并不是在所有细节上模拟人类行为（见《图灵测试是人工智能的标准吗？》中的讨论）。比如说，很多以往被看作 “谬误” 和 “偏见” 的现象在纳思中成为不可避免的（尽管并不一定是有益的）。

虽然有上述理由和优点，NAL还是被很多人（尤其是逻辑学家和人工智能学家）看成个怪物，因为它和他们所熟悉的 “逻辑” 太不一样了。从理论基础到技术细节，NAL挑战了大量逻辑学界广为接受的观念和预设。比如说，从一组正确的观察中可以 “合乎逻辑地” 导出一个可能被后面的观察推翻的预测吗？

对这类疑问，我的回答是：在不同的情境中适用的 “逻辑” 是不同的。**经典逻辑并没有错，但是在知识和资源不足的情况下不能用，而非公理化逻辑则正体现了在这种情况下的理性原则。**在不能假定未来和过去完全一样的情况下，没有哪种预测方法能保证不出错，因此按照经典逻辑什么结论也证明不了。NAL在此情况下会提供和已有证据一致性最高的结论。这种结论在和**未来**经验比较时可能会**错**，但在和**过去**经验比较时的确是**对**的。像我在《人工智能：何为 “智” ？》中所指出的，对具有适应性的系统而言（不论是人还是机器），经验性信念的评价标准应当是过去经验，而非未来经验。因此，“错误的预测” 仍可能是 “合乎逻辑的” 。

但另一方面，并不是每种认知功能都对应一种特定的 “逻辑” 。比如说，国内一种流行的说法是 “思维可分为抽象思维、形象思维、灵感思维三种形式” 。但在纳思中，这三种 “思维” 只不过是系统在不同条件下的不同表现，其差别是相对的，而其中的逻辑是统一于NAL的。这里细节以后再说，但起码我们不能仅由于它们在外部表象上的差异就断定它们是不同的思维形式。这就好比说思维的结果可能对也可能错，但这不能被解释为我们是在交替使用 “正确思维” 和 “错误思维” 两种思维机制。

恰当地区分不同的逻辑还有助于澄清很多错误观念。比如说，在对人工智能的能力和局限的讨论中，不断有人以哥德尔不完全性定理为依据断言人工智能不可能到达人类智能的水平。既然这个定理说当一个推理系统足够强有力时，我们总可以找到一个其不

能证明的真命题，那岂不是说明我们总比这种系统聪明？这个断言的问题之一是哥德尔不完全性定理是关于公理化推理系统的，因此对于纳思这样的非公理化系统完全不适用，因此用这个定理不能说明我们永远比纳思高明。当然，这不是说纳思是具有“完全性”的（即可以证明所有真理），因为所有开放系统（包括我们自己）根据“开放性”的定义就都是不完全的，因此在这一点上“人”和“机”并无根本差别。**这里的教训是不能想当然地把人工智能系统看成公理化推理系统。实际上哥德尔不完全性定理尽管很伟大，但和人工智能并无直接关系。**

尽管目前“经典逻辑”常被当成“逻辑”的同义语，定理证明的逻辑和适应性行为的逻辑实际上是根本不同的。二者的确有共同点，但差异足够大，以至于应当被当成不同的系统。各种非经典逻辑都是看到了这一点，但大都从经典逻辑走得不够远。为满足通用人工智能的需求，在“逻辑”上很可能必须以“激进修正主义”的态度重起炉灶不可，这也正是为什么包括我在内的人工智能研究者干起了逻辑学家的活计，因为人工智能的任务、工具和制约条件为逻辑研究提供了一个独特的视角。这一实验的结果又会反过来深化我们对人类思维的认识，这就是我为什么在《人工智能危险吗？》中说**人工智能的研究在很大程度上也体现了人类自我认识的努力。**

参考资料

[1] Pei Wang, Non-Axiomatic Logic: A Model of Intelligent Reasoning, World Scientific Publishing, 2013

[2] Susan Haack, Deviant Logic, Fuzzy Logic: Beyond the Formalism, University of Chicago Press, 1996

了解作者信息和他在赛先生上已发表的文章，请点击页面左下角的“阅读原文”访问他的主页。本文图片来源于网络。

投稿、授权等请联系：iscientists@126.com

您可回复"年份+月份"（如201510），获取指定年月文章，或返回主页点击子菜单获取或搜索往期文章。

赛先生由百人传媒投资和创办，文小刚、刘克峰、颜宁三位国际著名科学家担任主编，告诉你正在发生的科学。上帝忘了给我们翅膀，于是，科学家带领我们飞翔。

微信号：iscientists

▲

长按图片识别二维码关注我们

Read more