

# 关于人工智能的最新假说

## ——知识的阈值理论

Douglas B. Lenat 和 Edward A. Feigenbaum

两位作者根据人工智能研究的已有成果，在1987年8月的国际人工智能会议上提出了三个基本假说：(1)知识原理——一个能有效地完成某项复杂任务的程序，必须具备相应领域的大量知识。(2)宽度假设——智能系统(程序)还应具备两种应付意外情况的能力：首先，它可以退出专业领域，转而应用一般性知识；其次，系统能够应用类比法推理，以便借用其它领域的知识。(3)实验验证假设——人们关于人工智能的思想和方法必须在大型的问题求解系统中得到检验。这三个假说统称为知识的阈值假说。

这一假说的提出，引起了学术界的广泛注意。有人认为它从更高层次上抓住了当前人工智能发展的关键问题，将给已经陷入困境的人工智能研究带来新的生机，使其产生质的飞跃，从而推动人工智能研究进入一个新的发展时期。有的则认为它只不过是一篇空泛的思辨性文章。

是非曲直让读者评说，现将其精华部分译出。

### 一、 知识原理

过去三十年，AI工作者一直在追求一个梦想，就是让计算机完成各种困难的认知任务。事实表明，要实现这一目标，计算机必须具备专业知识并能利用它。

智能，是一种在巨大搜索空间中迅速找到问题解的能力。“知识就是力量”这一假设已被公认，它可以进一步表述成：

**知识原理 (KP)：**一个系统之所以能体现智能理解和智能行为，主要是由于它能够利用专业领域的概念、事实、表示、方法、模型、隐喻以及领域的启发性知识，等等。

“专业领域”一词是非常重要的。知识常被看做是被编辑过的搜索。由于搜索的代价昂贵，而知识的存贮相对低廉。因此，知识原理主张在不牺牲已有能力的情况下进行概括，以便获得可以用于更广领域的知识。当

然，仅有一般性知识是不够的，因为一般性方法往往不能使问题求解空间的搜索有效地收敛。

在利用知识进行问题求解和通过搜索进行问题求解之间存在着连续性。知识和搜索在问题求解过程中常常是交替使用的，在遇到已知的事实或完备的方法后，这一过程才会终止。知识与搜索的交替使用经常被作为强调搜索的佐证。实际上，它同样也可作为强调知识重要性的佐证。

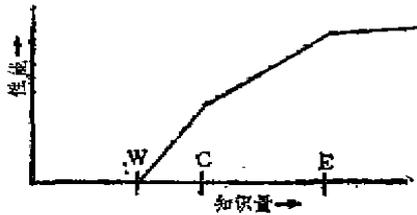
**形式化阈值：**要求解一个问题，至少得有形式化此问题的知识。有人说：找到了一个很好的知识表示方法，问题就解决一大半了。

目前的专家系统都已超过了形式化阈值。但要具备专家水平，还需要更多的知识。

**胜任阈值：**难题的求解非线性地依赖于知识。在达到胜任阈值以前，每增加一份知识，每能使系统性能得到一分改进。超过胜任阈值后，增加的知识是有用的，但利用率较低。

超过胜任阈值，就进入了专家圈子。这样的专家能协调地运用知识和技索。第一代专家系统含有的知识量一般都超过了胜任阈值，但都少于现有的人类专家的知识量。

**专家阈值：**随着系统知识量的不断增加，最终会导致所有异常情况都能得到处理，这时系统的知识量就达到了专家阈值。此后增加的知识对系统性能的改进将不会有实质性影响。



W—形式化阈值

C—胜任阈值

E—专家阈值

图1 知识与系统性能的关系

图1表明了知识与系统性能的函数关系。

人类专家一般都处在胜任水平与专家水平之间。这并不是说其它知识不重要，只是目前我们对那些知识还没有充分认识。例如，类比就可以把一个领域的知识应用到另一个领域。

现实世界的许多问题（或任务）所面向的领域比较窄，具有50~1000条if—then规则就能超过胜任阈值，再增加同样多的规则就几乎接近专家阈值了。由于专家的经验性知识目前还不能完全用规则显式地表示出来；而且一次性程序设计也不能完成规定的任务。因此，当专家系统发生错误的时候，只有通过人类专家来纠正。当然，这同时也给系统增加了新的规则，使得系统增量式地接近胜任水平或专家水平。

新规则应与系统已有规则完全拟合。也就是说，你将分辨不出系统运行过程中什么时候运用了新规则，什么时候运用了旧规则。

**显式知识原理：**智能系统的知识必须以显式的形式表示出来（尽管经过编辑）。

把知识表示成显式的形式，就可以在其上应用元规则，进行知识获取、检查等等。这种显式的形式化知识将有利于进行类比和抽象。

有了足够的知识，还需不需要那种精巧的推理方法呢？

知识就是一切，对于智能行为来说，并不一定需要复杂、精巧的控制结构。一个标准的估值函数所使用的控制策略也是知识。

一方面，归纳、演绎、类比、特化、概括都要以知识为基础，因此知识成为我们的瓶颈，另一方面，策略和方法本身也是由一条条知识组成的。因此，我们说：知识就是一切。

## 二、知识原理的佐证

人的大部分思维活动，都是某种符号处理过程。五十年前，图灵机和图灵理论就暗示了计算机可以做为这一符号处理过程的形式化模型的基本思想。

1956年Dartmouth夏季AI会议以后，人们开始了最早的符号处理过程的研究。Newell和Simon首先提出了人工智能研究模式的基本原理假设：

**物理符号系统假设：**把现实世界的对象、活动和相互关系表示成符号间的互连结构，并在这种结构上应用符号处理过程，就可以模拟智能行为。因此，数字计算机完全可以表现出智能。

早期，人们费尽心机地研究能产生胜任水平的智能系统。图灵研究人-机会话、设计下棋程序。Newell和Simon也编过下象棋程序。从事这一方面研究的还有Gelernter、Moses、Samuel等。

在斯坦福大学，Lederberg与Feigenbaum研究了精确的推理方法，并同Buchanan和Djerassi合作设计了Dendral。通过对Dendral的实验研究，我们发现：这个系统之所以获得了成功，是由于它具备了大量的化学基本知

识和光谱化学知识。

今天，我们回头看，才发现知识在加速搜索过程中所起的作用。可是在六十年代，人们却认为人工智能研究的最终突破将是推理机的研制成功。

七十年代，人们开始设计医疗程序和科学问题求解程序，从此拉开了应用知识原理的序幕。MYCIN形成了一类咨询程序的范例，它的推理系统是简单的逆向推理链。DEC从1981年以来一直在使用R1程序，并进行了不断的修改，它的控制结构就是单纯的正向链。这些专家系统能用专业术语与专家们对话、显示推理链、并具有一定的知识获取能力，比如让专家在一定的规定下对知识库的规则进行修改、增添。

在过去十几年里，专家系统已渗透到工程、制造、地质、分子生物学、财政、机器故障诊断、信号处理等领域。但是其中大部分都是“分立”系统，而且这些系统使用的技术不外乎是启发信息技术，如利用经验性规则和定性规则等。此外，这些系统采用的推理技术都是一些弱方法，系统性能的好坏完全取决于它的知识库中的知识。

进入八十年代，很多AI工作者开始转向基于知识的观点。例如，常听搞机器学习的人说：一个机器学习程序在运行以前，必须在知识库中存有一定的知识，而不应该从头开始学习。搞视觉程序的人也说：一个视觉程序系统必须对其要分析的环境有很好的认识和理解。

### 三、 宽度假设

第一代专家系统是脆弱的。在没有超过其知识的有效范围以前，它具有很好的胜任能力。一旦超过其知识的有效范围，它便手足无措。我们人也会遇到这种情况，不过我们人能够胜任的问题很多，而且在遇到没有见过的问题时，并不是束手无策，例如我们可以运用常识性知识，也可以通过类比等方

法借鉴其它领域的知识或解决其它问题时积累的有关经验。

**宽度假设 (BH)：对于智能系统，它的问题求解器必须能利用逐渐增长的一般性知识，并且（或者）能进行类比推理。**

专业知识是一个领域内经验的总结，在遇到相应的问题时，它能使搜索很快收敛；但面临新问题时，就不得不转向一般性知识或进行类比推理。

每个人都掌握着一定的常识性知识。例如，我们说“做功需要能”、“水往低处流”等等。对于一个人来说，他所掌握的这种知识是具有一致性的（不矛盾的）。专家系统中没有这种简单的常识，因此就有可能出现笑话。如，对一个车体附有红斑点的旧车，一个皮肤疾病诊断程序可能会得出结论说：它得了伤寒病。

进行类比时，常常要对两个问题进行匹配。现实世界有很多相似的因果关系、相似的结构，因此进行类比推理是可能的。

所谓类比，是指如果A和B具有某种相似性，那么它们就可能具有其它相似的性能。这是个一般法则，对于具体问题来说，通常还有不同的类比方法，这取决于所掌握的数据多少。假如类比空间是n维的，我们就可以设计一个知识体，使它在n阶矩阵中进行这种类比推理。

为什么要考虑因果关系呢？这是因为，如果原因A和B没有经过从具体到一般的概括，那么它们之间的相似性充其量是一种表面的巧合，而不具有启发性。

类比能够增加知识库中知识的相关性。它能对有关数据提出新解释，增添新知识，填充缺省值，建议可行方案，等等。现今，我们设计专家系统时，都是用手工方式把有关知识放进一定的数据结构中去，还要进行编码设计。将来，类比有可能不只是作为一种推理方法，甚至可以作为增加新知识的辅助手段。

## 四、 宽度假设的佐证

目前,能够说明宽度假设的事实还不多,因此,我们暂且只能把它称为假设;否则,就可以称之为宽度原理了。不过,研究一下AI各领域目前面临的难题,我们同样能得到说明宽度假设正确的事实。这里,我们看一下自然语言理解(NL)和机器学习(ML)。

**4.1 自然语言理解** 在自然语言理解中,需要很多知识密集型技巧,例如辨析词义、分析代词的指代以及省略掉的含义等。对一个特殊的文体,自然语言理解程序可以编入二十世纪关于美国的一些史料,一些常识作为语义分析、问题解答以及隐喻分析用。

从某种意义上说,NL工作者已经解决了语言理解问题。不过,要设计一个通用图灵可测试系统,还要给它增加更多的语义信息,而且程序的语义部分将越来越类似于受宽度假设支配的大型知识库。

我们是不是言过其辞了?实际上,你只要翻翻报纸,就会发现一个惊人的现实,几乎字里行间无不使用着隐喻之含意。NL工作者,还有词典,为了避免用法混淆,常给一个单词赋予几个含义。

另一方面,我们在读一本书时,一般事先都是具备一定知识。而且,作者在著书时,无形中就假定了读者已知道这些常识。如果语言理解程序不知道这些,那么用户与这个程序系统就很难进行交流。

今天,NL工作者大都认识到了设计一个大型语义分析构件的必要性。不过,只是近来才有人真正开始研制这样的系统(CYC [Lenat 86]和日本的电子词典研制工程(EDR))。因此,只有过几年,我们才能得到更有力的事实来证明宽度假设的正确性。

**4.2 机器学习** AM和Eurisko 是含有知识量较多的两个学习系统。我们在这两个系统中进行了一些机器学习的实验,尤其对启发性技术的本质和工作原理进行了研究[Lenat和Brown]。我们发现:尽管AM和Eurisko已存有一些知识,但它们的学习能力最终由知识库的大小决定。正象Porter最近同我们说的,没有一定的知识,也很难学习到新的东西,

**知识递进学习:**如果掌握的知识有限,就不要

指望能很快学习很多东西。

培根曾指出:“要想走在前,理论得当先”。当然一个新的理论在它产生时,人们还很难预测它是否很有用。这个问题反映在机器学习中,就是有人提出机器学习应该是基于解释的学习和基于目标的学习。

人类从“无知”开始学起,这是不是与上面的讨论相矛盾呢?这还很难说,因为我们至今还不知道婴儿的大脑活动是从哪儿开始的呢!这些问题的进一步探讨超出了本文,也许也超出了二十世纪的科学研究范围。

人生下来除了有个结构很好的大脑外,还需要花很长的时间进行学习。婴儿的大脑活动过程我们还不清楚,一旦孩子开始学说话了,也就进入了符号学习阶段。符号学习也是传统的AI研究的核心问题。

## 五、 实验验证假设

我们这些科学家们都自认为富有创造力。实际上,同大自然相比,我们的想象力太贫乏了。因此,对于我们来说,发现要比发明容易得多。经验表明,过早地数字化将掩盖世界的多样性、复杂性及其神秘性。这导致了我们的中心方法学假设—AI的研究范例:

**实验验证假设(EH):**我们对智能的了解还很贫乏。在我们面前,自然界把它最神奇的奥秘掩藏起来了。因此,研究AI的最有效办法是,把我们的假设编到程序中去,通过程序的运行来获得数据。意外结果将修正我们的假设(也促使我们重新提出新的假设),学习系统的进步就靠这些否定我们假设的实验结果来诱发。也就是说,AI程序必须具备实验者本身想不到的行为。

早期的AI程序曾使其编制者惊奇不已,如Newell和Simon的LT程序,Gelernter的几何定理证明程序。可是,由于沉溺于公理系统的定理证明,就愈加看不到现实世界的奇妙性了。

AM程序没有什么明确的目标,它只能说是个实验,把几百条数学启发式规则应用

到100个简单的概念上，看看会产生什么结果。AM产生了很多令人惊讶的结果，其中有很多已用到Eurisko的构造中。这里需要再指出的是，一个智能程序系统的最终能力极限并不是我们一般所想的诸如运行时间、新的表示技术等，而是机器当中需要很多一致性。沿着我们指出的途径，就是要进行大的实验：

**困难问题假设：解决简单问题有很多方法，要想提高胜任水平和胜任宽度（解决问题的有效范围），我们需要一个系统，它能使智能的测试和智能水平的提高变得容易。**

AI界目前的混乱也许是由于我们把AI目标与为了达到目标所采取的策略混为一谈了。各学科领域都有自己的范例，一系列值得考虑的问题、一系列可以尝试的方法、一系列评价结果的标准等。一个范例之所以被接受往往是其认知方面的合理性。不过，每一范例的应用范围是很有限的。AI的一些范例是随着各种目标和各种策略逐渐形成的。有种倾向认为：任何一个策略都可应用于任何目标的实现，这更引起了AI界目前的混乱。

知识原理、宽度假设、实验验证假设都是一种策略。采取什么策略进行研究，这是每个AI研究者都要考虑的问题。

## 六、不同的观点

**6.1 关于唯美主义的观点** 人们喜欢简单的一致性的假设，追求形式上的完美。然而，AI研究方法与唯美主义者的方法截然不同。按照困难问题假设，AI研究者必须集中主要精力来研究领域问题。

很多AI工作者从“尊重科学”的愿望出发，一直在寻求一个优美的解。而且认为，只有上升到数学形式描述的高度，AI才能被认为是严谨的科学。但实际上，形式化还从来没有促进过一个学科的早期发展。例如，在设计新的药物分子时，生物化学家都知道，应用Schroedinger波动方程来计算能量的最小值，其效力是很低的。所以尽管有这一理论存在，但生物化学家们还是喜欢用一些不很成熟的启发式方法。许多药物公司用这种方法赚了大钱。

唯美主义者的一个观点就是提倡高度的并行性或高速设备。问题是在知识丰富的领域，最复杂的任务往往不能被并行处理。假如我们让一百万人去发掘治愈癌症的药方，在0.2秒时间内我们是找不到办法的。每一项癌症实验要用几个月或几年的时间才能完成，每一次实验中只有不多的结果有研究价值，上次实验结果又决定下一次实验应该做的内容。

有些问题的处理时间会随着问题的增大而指数增长，采用微程序设计方法或并行方法只能去掉一个常数因子；采用基于知识的方法也许会增加一恒定的开销，但更重要的是，有可能缓解或消除时间的指数增长。随着所面临的问题变得愈加复杂，搜索时间的指数增长也将会变得愈加严重。

**6.2 关于专家系统的观点** 知识原理是目前专家系统开发的基础。下面是关于专家系统的其它一些观点。

专家系统的一个主要能力源泉，或者说它之所以那么快地被建造出来，是由于它具有多个规则的最佳合作迭加性。如，通过黑板或者分割规则集，就可以把一些规则集成一个规则群，从而形成更大的专家系统的知识源。

更进一步，就是利用类比方法把几百个大的专家系统连在一起，达成一个更大的最佳合作体系。这个梦想在具体实施中失败了。因为当拓宽每一个“元素”的领域时，就要考虑如何把它们有机地耦合在一起，这种“胶合剂”将会变得愈来愈复杂。在专家系统技术的真正潜力被认清以前，我们似乎需要一个由宽度假设支配的大系统。

把很多分立的专家系统结合起来形成一个大系统，就要找到一种方法，它能将我们一般人所具有的大量抽象概念、模型、事实、经验规则、表示技术等组织成一个接近完全一致性的体系。

把各领域的专家系统结合在一起来的过程中，各领域本身具有的简单性和特异性都会给这种最佳合作系统的建造带来困难。这也导致了目前的专家系统大都是分立的、简单的、知识是自成体系的。在着手建造专家系统时，首先就要同专家们对话，找到他们解决问题的最复杂的知识以及一些技巧。例如，欲问：专家是如何回避关于时间和可信度的一般性推理的？也许从专家们的回答中发现，他们从问题的信息中抓住的时间和可信度的数据结构是很特殊，也是很简单的。

有一点对于专家系统和逻辑学家都适用，那就是对于一个大的专家系统的知识库，没有必要（甚至不可能）实现全局一致性。大系统只需满足局部一致性，我们称之为有理的。例如，在物理学中，光的波动说和光的粒子说并存了几十年。知识空间是自我证明的。在一个前后一致的系统中，推导出矛盾只能使它比一般的没有任何进展的那类系统更严谨，因为此时系统可以回溯，达到新的一致性（或称之为平衡态）。

## 七、 结论：超过局部最大值

我们的观点是：

\*要想有效地解决难题，必须掌握专业领域的知识。

\*为了应付新情况，必须具备一般性知识以及广泛的其它领域的知识。

\*我们已有了很多关于智能机理的理论，下一步需要进行更多的实验，继续建造大的实验基地观测智能系统的行为、研究类比推理的机制、设计机器学习与自然语言理解系统……

\*尽管机器学习、自然语言理解及其它一些AI领域已取得了成绩，但在八十年代，通过手工把知识编到程序中仍旧是最快的方法。

知识原理要求人们对各领域的难题求解时所用的知识具体化、形式化。这种正式的知识教育要比那种师傅带徒弟的方式快得多。正是基于知识原理的观点，才促使人们研制了那么多有效益的专家系统。

宽度假设要求消耗必要的资源来建造大型知识库。而且要求系统掌握各门学科中的经验和思想。此外，系统还要对每一学科有一定深度的了解、掌握各领域的一般性理论。

作为宽度假设的一个应用，考虑建造一个基于知识的系统，使之覆盖大部分工程设计。这个课题已由日本的EDR工程和Bob Kahn国家研究所分别选中。两个小组都把此课题当做一个中期（1994年）目标。它肯定比单一的专家系统应用领域宽得多，但要比受宽度假设约束的通用知识库窄得多。

更狭窄的应用领域就是所谓的“律师工

作站”或“植物学家工作站”，它们同样也是很有前途的。它们将具有一般的常识性知识、各领域所独具的一般性结论、合适的表示技术以及领域中的具体细节的知识（如抽取知识的特殊方法、联机数据库中有关此领域的一些概念等）。这种工作站对于专家们将是很有诱惑力的，专家们也将会乐于在这种工作站上工作，并对知识库中的知识进行修改、增删。

几十年来，我们把自己的职业生涯都倾注在KP、BH和EH的研究工作上。也许有人会说：这是个骇人的想法。但在科学事业上，总得有点冒险精神。说它是骇人的不外乎有两点：（a）对于大多数AI研究者；这几个假设并非显然，（b）这些假设对于我们这些倡导者来说也并不是完全能接受的！

为什么这些假设不是显然的呢？大多数AI研究者侧重于一些非常小的问题。最后的结果是，只利用很少的知识就解决了问题。因此看上去，这些问题的求解与KP、BH和EH是无关系的。只有在处理现实世界的难题时，人们才会感受到这些假设的正确性。

为什么这三个假设是“难以接受”的呢？因为它们不是唯美的！

我们目前已清除了很大的障碍——已经解决了那些与知识无关的问题。换个说法——已经跨过了与知识无关问题的局部最大值。越过这一“山巅”，另一座“山峰”又耸立在我们面前，那就是专家系统。我们同样地把专家系统看做是局部最大的。AI将要越过这个局部最大值。本文表达了作者关于专家系统的未来发展的一孔之见，即超大型知识库的前景是光明的，类比推理、常识性知识也是具有很大潜力的。

按KP和EH建造的系统对未来究竟会产生什么影响还很难预测。公共教育目前还不能达到为每个学生提供一个机智的、胜任的老师。AI将会改变这种情况。我们现在的娱乐业采取的方式还是消极的，AI将会给这一行业带来转机，使得“旁观者”变成“实施者”。