人工智能迷途：计算机的高技能等于高智能吗？

原创 2016-02-17 王培 赛先生



王培 （美国天普大学计算机与信息科学系）

人工智能(Artificial Intelligence，简称AI)开始于一个很简单的想法。当计算机被发明之后，不少人很快认识到它的能力远不限于数值计算，而是可以通过对信息编码和对指令编程完成很多其它以往只有人脑才能完成的任务。这自然就导出了是否人脑的功能都能在计算机中实现的问题。这类研究开始有不同的名称，最后“人工智能”这个词被大多数人接受了，尽管对其解释始终没有统一(见《当你谈论人工智能时，到底在谈论什么?》)。



计算机科学之父图灵（Alan Turing）

开始，AI研究者们以为可以在不远的将来实现他们的目标。这个领域的先驱图灵在1950年

预言在五十年内可以造出言语行为和人难以分辨的“思维机器”，而这个领域的奠基人明斯基和西蒙等人的估计甚至还要更乐观些。在AI历史上，曾经有过一些雄心勃勃的研发项目，如纽厄尔和西蒙的“通用问题求解器”和日本政府牵头的“第五代计算机系统”等，但都没有达到预期目的。这些失败让AI的名誉受到巨大的负面影响，导致经费断流、人员出走、岗位削减，史称“AI之冬”。

人工智能“脑死亡”了?

痛定思痛，AI的大部分研究者觉得失败的原因是前辈们太好高骛远了。用一本权威AI教科书[3]的说法，在八十年代末“人工智能接受了科学方法”，其特征是从热衷于发明新理论转向基于经典理论，从期望发现一揽子解决方案转向解决具体问题。这使得AI着力于一个一个地解决各种以往只有人能解决的实际问题，或一项一项地再现人的认知功能。由于各种“专家系统”在很多领域达到甚至超过了人类的水平，AI终于摆脱了“现代炼金术”的恶名。广为人知的例子包括战胜了国际象棋世界冠军的“深蓝”系统和近年来风头正劲的“深度学习”技术在图像、语音识别等领域的应用等。

但这个从通用系统向专用系统的转向是有代价的。由于“只有人能解决的问题”在不同的时期有不同的内容(如曾经包括四则运算、排序等)，以此定义“人工智能”实际上把它看成了计算机应用的前沿。由于这个前沿在不断扩展，所面对的问题千差万别，解法也因问题而异，所以不可能全被囊括于一个理论之中，因此也就说不清AI和普通计算机应用到底有什么区别。甚至有人说“人工智能”专指那些尚未被计算机解决的问题，因为一旦一个问题被计算机解决了，它也就不再是“只有人能解决的问题”了。以这种标准来界定一个研究领域自然会造成大量的思想混乱。

即使我们完全不在乎“人工智能”到底是什么意思，只是以“能抓耗子”来定义“好猫”，这种“分而治之”的战略仍有致命问题存在。人类的各种认知功能显然是有紧密的相互依存关系的，而人在各个领域的问题求解能力也不乏相通性和可转移性。在目前AI各子领域中，都有不少核心问题是依赖于其它子领域的。以自然语言理解为例，有什么证据说明语言理解能力是和推理、学习、记忆、感知、规划、决策等能力无关，而可以“就语言论语言”的?

对上面的诘问的常见答复是：智能之中的诸功能当然是互相依赖的，但由于它们太复杂，只能一个一个研究，将来再把它们集成一个完整的智能系统。或者说当具体问题解决多了，普适的办法也可能就找到了。

这种战略在很多领域的确是恰当的，但在AI研究中似乎并非如此。解决具体问题的最有效手段常常是通过深入分析该问题的特点得到的，因此未必能推广到其它问题。比如说“深蓝”的技术对计算机下围棋就不适用。类似地，当某种认知功能被孤立定义成一个计算过程时，其结果和该功能在人脑中的“原生态”常常已经是大不相同了。大略翻翻一本AI教科书(如 [3]) 就不难发现大量针对各种问题的解决方案。把它们都编成程序装到同一个计算机系统里并不难，难的是让它们协调工作。由于这些方案是基于不同的理论和预设的，它们完全没有协作的基础。这就解释了为什么整合各种AI功能的呼吁(如[1])收效甚微。

更重要的一点是，即使是这些打上“智能”标签的技术仍常常让我们觉得缺了些重要的东西。的确，它们在很多方面比人更强，因为它们速度更快、容量更大、功能更稳定，但是这本来就是计算机的长项，而它们仍然缺乏适应性、灵活性、创造性等直观上和智能密切相关的特征。即使是最新的机器学习算法也仍然受制于其中种种对应用领域的预设，因而和人的学习能力的普遍性不可同日而语。

明斯基曾在一次采访中对AI主流转向专用系统大表不满，斥之为“AI已经脑死亡了”。

通用人工智能的崛起

其实就像历史上类似情形中那样，总还是有些不改初衷的研究者，尽管他们只能在学术界的边缘沉默坚持。在2001年，他们中的一些人为了改变这种状况，开始把他们的研究成果编成一本书。最早草拟的书名是《Real AI》(真正的人工智能)，在2003年全书完稿时定为《Artificial General Intelligence》(通用人工智能，简称AGI)。这本书在出版过程中又经历了些周折，故2007年才得以面世[2]。在此期间，在2006年于华盛顿特区召开了为期两天，大约50人参与的AGI研讨会([4]就是这次会议文集的引言)。在这些准备活动的基础上，AGI年会(http://agi-conf.org/)从2008年开始举办，而AGI学报(http://www.degruyter.com/view/j/jagi)也在2009年创刊了。时至今日，AGI的概念已经广为流传，尽管人们对其有各种各样的理解和反应。

一个常常被当作AGI同义词的概念是“强人工智能”(Strong AI)。直观来说这是可以理解的，因为AGI的野心显然比目前的AI要强很多。但是按照这个词的发明人哲学家塞尔的说法，即使一个计算机系统的外部表现完全像人，那也只是个“弱人工智能”，而“强人工智能”必须真正拥有自我意识。这样一来一个AI系统是“强”还是“弱”就完全没有外部标准了,这是很多AGI研究者所不同意的。

与此相反，AI和AGI的差别是有明确外部标准的。如果要用一句话来概括，可以说AGI是把“智能”看作一个问题，而目前主流AI是将其看作一组问题。从AGI的观点看，“智能”当然是个很复杂的问题，必须一步一步地解决，但这里的“一步”不是“下围棋”、“识别照片中的事物”或“听懂中国话”，而是要根据一个智能理论来划分。这就像盖一栋房子应该从画图纸开始，然后据此一间间、一层层地建，而不是各人根据自己的想法分别去建客厅、书房、卧室、厨房、阁楼、卫生间等等，最后再设法把它们拼在一起。这听上去像常识，只是这样工程周期显然比只造一个房间会长很多，而且在全部完工之前不容易看出好坏。因此，在主流AI的工地上，绝大部分施工队都在忙着造各种房间，而不愿花功夫去捉摸那些成败难料的图纸。

尽管有“建造多功能通用系统”这个共识，目前AGI领域内对“智能到底是指什么”仍旧未达成一致意见，而在《当你谈论人工智能时，到底在谈论什么?》中介绍的几种观点在目前AGI研究中均不同程度地有所体现。既然对“做什么”都有不同看法，对“怎么做”的答复就更是百家争鸣了。尽管如此，已经有若干AGI项目离开了纸上谈兵的阶段，开始有具体系统可以被检验了，虽然离完成都还有很远的距离。

高技能≠高智能

在AGI中的“G”(General)不意味着能解决一切问题，而是指在系统设计时不限定其应用范围。“通用”自然是个程度问题，“通用”和“专用”也是相对而言的，但这个差别仍然是明显存在的，比如“深蓝”就完全不能算个通用系统。一个常见的误解是把一个通用系统看作一批专用系统的组合。实际上它们是两类非常不同的系统，不能相互替代。不仅AGI不能通过堆砌现有AI技术来实现，而且即使在AGI实现后也不会抹杀专用技术的价值。如[4]中所指出的，对一个可以明确刻画的应用问题而言，专用系统往往比通用系统更有效和更可靠。通用系统的应用价值是在那些问题和解法都难以事先确定的领域。在理论价值上，这二者的差别就更显著了。各种现存AI理论对于揭示智能的一般原理贡献甚微，而AGI理论则往往是从一般原理入手的。

在《人工智能：何为“智”?》中，我介绍了我对“智能”的工作定义是“在知识和资源相对不足的条件下的适应能力”。按照这个定义构建的系统可以被置于很多不同的环境中，而其智能使得它能够逐渐通过学习获得在该领域解决问题的能力。照这种观点看，“智能”不是“解决具体问题的能力”，而是“获得解决具体问题的能力的能力”，因此是一种通用的“元能力”(meta-ability)。

我对于“能力”和“元能力”的区别的最早认识来自于小时候读到的一个寓言故事。故事的框架很老套：两兄弟遇到一个白胡子老神仙，他们选择了不同的法宝，因此走上了不同的人生道路。这个故事的特别之处是其中的法宝是两双手套，一双是白色的，另一双是黄色的。带上黄手套就可以直接拥有各种技艺，而带上白手套就可以很快学会各种技艺。后面的情节大家都能猜到：选了黄手套的马上变身为成功人士，而选了白手套的仍需四处学徒。有一天两个法宝同时失效了，这时前成功人士立刻被打回原形，而那位前学徒工却仍身怀长技。这个故事的出处我已经找不到了(如果哪位读者知道请告诉我)，两双手套的颜色也可能被我搞反了，但故事中的教益却对我影响深远，且不仅仅是“不要贪图急功近利”这么简单。

在我看来，“能力”可以被分解成“技能”和“智能”两个方面。一种“技能”指解决某类问题的专门化能力，而“智能”指那种获得和提高各种技能的能力。如果把系统的总技能看成时间的函数，那么智能就是这个函数的导数。系统在某一时刻的技能等于其“本能”（先天技能）部分和智能所贡献的部分的叠加。这两个部分的关系在人类和计算机系统中非常不同。因为人类的本能差别较小，同龄人的技能差别可以近似地反映智能差别，这就是“智商”定义的基础。相反，计算机系统的技能既可以来自本能，也可以来自智能。当一个计算机系统完全靠人编的程序解决问题，而不根据自己的经验对其行为进行任何调整时，它仍可能拥有很高的技能，但完全没有智能。而另一方面，一个高智能系统如果不依赖于预先编好的应用程序，在开始运行时可能并没有多高的技能。

根据上述分析，把人类智商测试直接用于人工智能是错误的，因为计算机系统在一个时刻的“智能值”和“技能值”之间没有相关性，而前者必须通过后者的变化来衡量。智能的具体衡量标准和方法仍是AGI领域中的一个热点课题，但无论如何，这种衡量不能以一套固定的考题来进行，因为那样一来，一定会有人去编制一些在这种考试中表现出色，但此外一无所能的系统。这绝不是正确的导向 — 图灵测试已经因此而遭到“鼓励欺骗”的指责了。

“机器学习”还远远不够

把智能理解为学习能力远非新想法。心理学家如皮亚杰已经明确地指出智能是一种高级适应能力，图灵认为思维机器要通过学习实现其功能，而“机器学习”从一开始就是AI的一个子领域。在这种情况下，为什么上述想法不能在主流AI中发展呢?这是因为像“智能”一样，对“学习”这个概念也有许多种不同的理解，而我的理解和主流AI的很不一样。

如上所述，我认为“学习”是一个通用的元能力，而且是智能的必要条件，但在目前的AI研究中，“学习”一般是被看做是诸多认知能力之一，是起锦上添花作用的，因而AI系统不一

定都有学习能力。即使在机器学习的研究中，“学习”也是一般被看做达到一个确定的输入输出关系的手段。这种学习系统的技能在“训练”阶段的确随时间增长，但达到某个水平后学习就停止了，而此后系统进入“工作”阶段。与此相反，在我的设计中，系统没有截然分开的“训练”阶段和“工作”阶段，学习是终生的，无处不在的，尽管在系统的“幼年期”和“成年期”学习和工作的比例有量的区别。最后，目前各种机器学习算法虽然有一定的通用性，但离AGI的要求还相差很远。

“错把技能当智能”是一个目前普遍存在的问题。对于只关心某个具体应用的人来说，到底一个计算机系统的某项功能是程序员设计好的还是系统自己学的，其间的差别大概并不重要，但如果关注点是系统的发展潜力和适用范围，这里的差别可就大了。

以我的研究为例，“知识相对不足”意味着系统必须不断学习新知识以应对新问题，而“资源相对不足”意味着系统必须不断调整策略以满足时间约束，而这种运行环境是不在现有AI研究(包括各种学习算法)的视野之内的。更大的差别体现在对阶段性结果的评价上。由于我的工作不是基于现有主流AI所接受的理论(主要是数理逻辑、概率统计、计算理论)之上的，而且尚未有具体应用成果，这种工作在他们看来就没什么价值。反过来，在我看来主流AI做的基本是“人工技能”，也不是我所感兴趣的。虽然其他AGI研究者和我的观点不完全一样，他们同样对主流AI的现有规范不满，而对AGI的研究目标和评价标准有更强的认同感。

目前介绍AI技术热点的报道很多，但其中的问题也不少，尤其是常常把这些技术做直观理解和外推，因而赋予它们很多它们本来不具有的性质。我在《赛先生》上写这个专栏的初衷是以我自己的研究为线索，逐步、系统、深入地分析AI、AGI和认知科学中的一些理论问题，澄清一些核心概念，并简单地介绍一个完整的智能理论。之所以牵扯到认知科学，是因为包括我在内的一些AGI研究者的目的不仅仅是要造思维机器，而且希望循此途径认清智能、认知、思维、心灵、意识等现象的本质和一般规律，这就不可避免地涉及到心理学、语言学、哲学等相关学科，而所谓“认知科学”正是指这些学科的交叉区域。我知道有很多读者的兴趣是在“人工智能是否会威胁人类”之类的话题上，但目前的这类讨论大量充斥着概念错误，所以在涉及这些话题之前，我们还是需要先把基本概念搞清楚。

参考资料

[1] Ronald J. Brachman, “AAAI Presidential Address: (AA)AI — more than the sum of its parts”, AI Magazine 27(4), 19-34, 2006

[2] Ben Goertzel and Cassio Pennachin (editors), Artificial General Intelligence, Springer, 2007

[3] Stuart Russell and Peter Norvig, Artificial Intelligence: A Modern Approach, 3rd edition, Pearson, 2010

[4] Pei Wang and Ben Goertzel, “Introduction: Aspects of artificial general intelligence”, in Advance of Artificial General Intelligence, 1-16, IOS Press, 2007

了解作者信息和他在赛先生上已发表的文章，请点击“阅读原文”访问他的主页。

投稿、提供新闻线索、转载授权请联系：iscientists@126.com

商务合作事宜请联系：dll2004@163.com

更多精彩文章：您可以回复"年份+月份"，如201510即可获取月度文章，或返回主页点击子菜单获取最新文章、往期文章或直达赛先生微博。谢谢！



微信号：iscientists



阅读原文



微信扫一扫
关注该公众号