金观涛：反思“人工智能革命”丨文化纵横

原创 2017-08-02 金观涛 文化纵横



点击文末“阅读原文”，可便捷购买《文化纵横》2017年8月号纸刊。

✪金观涛 / 中国美术学院；台湾政治大学；香港中文大学中国文化研究所

《文化纵横》微信：whzh_21bcr

[导读] 近期，柯洁和AlphaGo的对战，再次掀起人工智能问题的讨论——人工智能对人类的认识论构成了根本性冲击。人如果不再是万物灵长，如果远离真正的生产劳动和创造，成为附着于机器社会的寄生群体，势必会给涉及宗教、经济、政治、伦理的各个方面带来革命性的冲击。

金观涛先生从“科学与人文”的视角，强调当前这场人工智能革命实质是对“智能”认识的某种退步，AlphaGo的工作原理是仿生学，而人类智能的内核——创造并使用符号的能力却遭到忽视。究其原因，这是现代社会中人文精神丧失、与科学被技术异化的恶果。

柯洁和AlphaGo的对决，再次掀起人工智能的讨论。伴随大数据时代的来临，互联网、人工智能、虚拟现实互相融合进而有可能重塑社会；特别是人工智能的蓬勃发展，似乎正在颠覆人类社会现有的组织、生产和生活形态。

本文力图从人类智能和社会长时段发展的视角，审视当下这场人工智能革命的实质，及其对现代社会的冲击。进一步来说，今天有必要在认识论层面上，理解“什么是人类智能”，重新界定人类智能的进步与社会形成、演变的关系。当前对这些问题的认识空前混乱，恰恰反映出科学被技术异化和人文精神的丧失。

一场退回到原点的革命

当前正在发生的这场人工智能革命，实质是一场“退回到原点的运动”，由此所引发的讨论，在某种意义上也是对人类“智能”认识的大倒退。所有这一切都和人文精神的沦丧有关，特别是对人工智能的发展历史的忽略。

其实，早在“人工智能”这一观念出现以前，AlphaGo的工作原理已经被数学家和生理学家发现了。1948年，美国应用数学家诺伯特·维纳提出了“控制论”，促使智能研究的行为主义大转向。与今日人工智能研究直接相关的，是当时的自动机理论，麦克卡洛和匹茨的神经网络模型就是这类自动机之一，其被视作人工智能研究的连接主义学派（其对智能的认识也是行为主义的）的开创者。神经网络模型的基本思想极为简单--大脑神经元的工作原理和电脑工程师所说的逻辑门一样，一个CPU 就是一个逻辑门的集合体。[1]在此基础上可以得到神经元的计算模型：系统的输入为I1, I2, ... ,IN，接受这些输入的神经元输出用如下法则规定：将这些信号经过加权（W1, W2, ... ,WN），如果其强度超过了一定的阈值（T），则该神经元就会发放一个信号y给其他神经元或者直接输出到外界。[2]这样，阈值的改变意味着输入和输出关系的变化。当阈值的改变由输出引起时，我们可以说这部自动机在“学习”。

简言之，复杂的智能行为被简化成机器对外界刺激的反应和反馈。低层次的反馈是达到目的的行动；高层次的反馈，因存在代表网络连接方式之参数随输出结果变化而不断地调整，其可以解释神经网络结构和相应行为模式的变化，也就是学习机制。

作为神经网络自动机的AlphaGo，其工作原理早已一清二楚，在今天没有新的内容。20世纪40-50年代，因电脑刚刚出现，加上数据量的限制，AlphaGo不可能被制造出来。近年来，伴随计算机硬件的发展，以及海量数据的积累，使得神经网络自动机的研究和制造突破了物质和技术条件的限制，这就是今天的人工智能“革命”。

仿生学和AlphaGo

既然早在20世纪40-50年代，AlphaGo的工作原理已得到清晰认知，为什么当时神经网络自动机的功能没有被冠以人工智能之名呢？因为神经网络模型对应的是生物本能，而不是人类智能。



深度学习的基础是神经网络研究

举个例子。1959年，麦克卡洛、匹茨与美国认知心理学家杰罗姆·莱特文等合作完成了一篇论文--《蛙眼告诉蛙脑什么》。他们发现，蛙眼能看到的是相对于背景运动着并具有特定形态的物体，而对背景视而不见。如一只昆虫或者天敌的“影子”只要从眼前掠过，青蛙会立即做出反应：扑向食物或者逃进水中。麦克卡洛等人用一个神经网络的数学模型来解释青蛙看到了什么。蛙眼神经网络的功能是青蛙的本能，不能被视为智能。但是从神经网络来讲，它和AlphaGo没有区别。正因如此，当时神经网络自动机的研究对应的是仿生学。

生物神经网络的本能源于遗传，而AlphaGo的能力是通过学习获得的，两者似乎不可同日而语。青蛙对运动物体的反应快过人，狗的嗅觉是人不可能有的。这都是自古就广为人知的事实，从来没有人会认为它们能超过人类，因为人具有智能，生物本能不是智能。智能是一种学习能力，这正是AlphaGo引起广泛注意的原因。其实，上述疑问之所以产生，是出于对智能和学习关系的混淆。生物的本能亦是通过“学习”获得的，只是不被称为学习罢了。

什么是深度学习？其本质是神经网络结构（规定神经元联接之阈值）因网络功能而改变，它与生物本能因自然选择而形成的过程相同。AlphaGo通过人喂给它的大数据或下棋的胜负来修改网络结构，青蛙则通过适者生存进化出对运动物体做出迅速反应的神经网络，从结果规定网络结构这一点上看，两者一模一样。换言之，今日所谓系统的深度学习，其理论基础是行为主义和连接主义。学习机制为“自下而上”的选择路径，这只是生物学意义上“自然选择”导致物种形成的过程。只是生物界不需要某一物种以下围棋来生存，故自然选择这一“深度学习机制”没有创造出人所不能匹敌的围棋高手。

任何一种达到目的的过程和学习都需要不同层次的反馈，但这些高层次的反馈绝不等同于智能。其实，神经网络自动机在60年前并未被视作有智能的，人工智能研究的出现正是立足于对智能和本能的区分。

如何定义“智能”？长期以来，关于人类智能的思考，基本上是思辨性的。人们从自明的意识出发，通过内省来认识“记忆”、“学习”甚至是自我的形态。行为主义开创了用行为模式来定义智能的全新思路。举个例子，一只狗被汽车撞了，下次见了汽车就会害怕，通常人们是用狗有记忆来说明这一点的（心理学称之为生物对外界刺激的敏感化）；行为主义认为，狗有记忆等价于其行为模式改变。因为行为模式是可以通过观察来研究的，甚至可以定量地加以测量，故行为主义开启了心理学科学化、定量化的转向。[3]如前所说，“控制论”的出现使得学习所依靠的反馈机制被发现，它促使智能研究的行为主义大转向。实际上，用图林实验判别机器会不会思考，其基础也是智能的行为主义定义。毫无疑问，这是智能研究的一大进步。借用维纳唯一的学生阿尔贝勃的说法来概括：这是在智能甚至意识解释中赶走鬼魂的工作。但是，智能的行为主义定义正确吗？

就记忆本身而言，AlphaGo通过人喂给它的大数据或下棋的胜负得到的记忆，只是智能所具有的各种记忆中的一种，即最初等的“默会知识”。心理学研究早就指出：记忆分“陈述性记忆”和“非陈述性记忆”。默会知识是非陈述性记忆，比如骑自行车、游泳等，人们可以通过实践（反馈学习）来获得，但却不能用语言来描述它，甚至主体都不知道自己“为何知道”。另一种是陈述性的，也就是人类能够以符号系统予以表述的知识，其可以被传递、重构、遗忘甚至歪曲。行为主义和连接主义对智能的认识很大程度上只停留在其非陈述性的一面。然而，如果没有陈述性记忆，知识不能经理性加工、传递，亦不能社会化。也就是说，仅仅用非陈述性记忆是无法理解最简单的智能的。事实上，20世纪50年代末，心理学的行为主义学派受到广泛地批评，控制论研究的第二阶段开始了。智能有了更为深入也较为准确的定义，人工智能的研究从此起步。在此意义上，今日忽略对行为主义的批评、回到控制论刚兴起时状态的“人工智能革命”，是某种对智能认识的倒退。

什么是“智能”？

什么是陈述性记忆？这是主体使用符号来表达对象及形成的记忆，它比行为主义更接近智能之本质，这就进入到智能研究的另一个学派--符号主义。

艾什比是早期控制论的另一位代表人物，他对智能研究最大的贡献是，第一次给出了“智能”的较准确的定义。什么是智力？艾什比认为智力是进行正确（合适）选择的能力，选择能力的放大即为智力的放大，整个智能行为包含着符号（一种特殊的信息）之选择、编码、传输和解码，以及上述过程的自我叠代（和反馈相对应）。[4]符号（信息）选择的过程一旦实现自动化，可以说是有智能的。这就是符号主义学派创始人西蒙和纽厄尔所提出的“物理符号系统假说”（PSSH, Physical Symbol System Hypothesis）。它可以概括如下：“任一物理符号系统如果是有智能的，则必能执行对符号的输入、输出、存储、复制、条件转移和建立符号结构这样6种操作。”[5]这一自动化的符号选择则是“模仿数理科学的发展方式，将知识系统地整理成公理体系。这种方法将数学严格公理化，从公理出发，由逻辑推理得到引理、定理、推论。广义而言，将数学发现整理成一系列的逻辑代数运算，将直觉洞察替代为机械运算”。[6]

1956年夏天，数十名来自数学、心理学、神经学、电脑科学与电气工程等各种领域的学者，聚集在位于美国新罕布什尔州汉诺威市的达特茅斯学院，讨论如何用电脑的符号运算类比人的智慧，

并根据会议发起人之一的麦卡锡的建议，正式把这一学科领域命名为“人工智能”。符号主义学派的创始人西蒙和纽厄尔参加了这场具有历史意义的会议，而且他们带到会议上的“逻辑理论家”是当时唯一可以工作的人工智能软体，引起了与会代表的极大兴趣与关注。因此，西蒙、纽厄尔以及达特茅斯会议的发起人麦卡锡和明斯基被公认为人工智能的奠基人，被称为“人工智能之父”。

用符号运算类比人的智慧，比神经网络自动机更为高级，其形成的以及处理的符号系统，更接近由理性学习得到的陈述性记忆。但为什么随着电脑运算的进步和大数据的出现，这一对智能更准确定义的领域中没有发生革命？反而是在神经网络自动机的领域中出现了突破呢？关键在于，西蒙和纽厄尔在艾什比的基础上，虽然将人们对“智能”的理解向前推进了一大步，但仍然没有把握人类智能。人类智能不同于生物本能之处在于，人不仅具备选择的能力，还具备创造符号以及利用符号系统把握世界、并赋予世界意义的能力。

在数学符号系统的运行规则背后，起码还存在如下四个层面，它们才涉及智能的本质。第一，主体是自由的，可以给出符号并用符号系统指涉对象。语言就是这样的符号系统；第二，人用语言传递知识，组织社会，产生社会行动。这一切会反过来使主体意识复杂化，社会意识由此起源；第三，人会意识到自己有自由意志，让意志指向某种符号系统，创造一个应然世界；第四，应然世界的演变会进一步放大主体的自由，以创造出更为复杂的符号系统。数学只是人创造出来的众多符号系统中最不可思议的一种，数学真实具有和经验世界不同的真实性标准。

而人工智能的符号主义学派，其所着眼的智能只是第四个层面上的一小点（即数学推理）。当智能之冰山主体（即上述四个层面）还在水下，仅仅看到冰山之一角，我们能把握智能吗？当然不能！换言之，虽然对智能的符号主义认识比神经网络自动机深入得多，但仍然是瞎子摸象。用电脑科技和大数据实现的，只能是那些完全搞清楚的东西。今日人工智能讨论的巨大盲点在于，很多人把生物的本能当作人类智能。

智能研究中不可化约的两类变量

近年来，随着电脑技术的突飞猛进和大数据时代的到来，对人工智能的符号主义研究亦取得长足进步。20世纪70年代，地图四色问题在电脑的参与下得到证明，应该被视为人工智能研究所取得的第一个成果。四色问题证明之难点是数学家必须处理1700多种不同类型的基本图，在某种意义上超出了人处理无规则基本图的能力，必须借助运算机放大人的判断（选择）能力。此外，平面几何命题的机器证明当然也可以作为人工智能的成果。这些案例中，电脑只是在做人指派给它的工作，整个证明和解决问题的方案都由人给出，人机互动共同达到目标。它不像AlphaGo那样自动学习、自行给出表面上具有人类智能才能得到的方案。

另一类具有代表性的例子是近年来数位人文研究的兴起。在观念史研究中，为了揭示历史上普遍观念的变化，必须对表达普遍观念的关键词进行统计分析。理解关键词的意义、将其放到历史语境中认识过去曾经盛行过的普遍观念，以及历史学家在心中重演普遍观念和社会组织的互动，上述每一步都涉及上述智能定义的四个方面，这当然是电脑无法做到的，其必须依靠具有主体性（能理解他人智力）的人，也就是人文学者，但是建立文本数据库，从数据库中提取不同年代代

表普遍观念的关键词，对其进行统计、分类，特别是找到关键词在文本中形成的关系网络，这些都离不开电脑和处理符号系统的人工智能方法。当所有历史文献都被输入电脑，成为历史学家拥有的大数据时，人文研究必定面临一个转折点：如果再不利用人工智能，人文学者将不能处理大数据中所包含的海量数据，以发现历史真实。这难道不是大数据时代带来的人工智能革命的一部分吗？

举一个我们研究工作中亲身经历的例子。20年前，我与青峰在香港中文大学工作时，开始建立数据库并用关键词的词频统计分析来研究中国现当代思想的形成，其中“主义”的出现和成熟是当代意识形态取代传统儒家思想的重要指标。1890年至1930年中国出现过多少种“主义”？其种类和频度怎样变化？我们可以从数据库中提出有关“主义”的所有例句，再将每一个句子放到相应文本中做意义分析，找出其所属的类型，然后统计新的主义在每年中出现的次数，看其如何转化，但由于数据庞大，多达5万条，这就大大超出个别人文学者的研究能力，单凭人文研究者很难解决这一问题。2008年，我到台湾政治大学任教，我们建立了和资讯科学家合作的数位人文研究团队，由一个人文硕士和IT硕生的合作，很快找到了统计分析方法，原来近现代中国共出现过1680多种“主义”。直到20世纪20年代，每年都有几十种新的主义被提出。[7]当时我和青峰就对学生讲，当人文学者遇到仅凭个人的能力没法解决或无法从总体上把握的问题时，可以借助电脑技术来解决。这就是当前方兴未艾的“数位人文学”或称“数字人文学”。

但必须指出，在数位人文研究中，存在两类不可缺少但不能互相化约的变量，一类是人文的，另一类是引进电脑和人工智能技术加以处理的。无论人工智能取得多大突破，在人文研究中，电脑技术都是人文学者的辅助工具，不可能替代研究者对关键词的意义分析以及在心中重演普遍观念和社会组织如何互动，使这项研究呈现出人文研究的主题。至今为止，人工智能的发展都没有可能涉及上述人类智能的四个基本层面。而人文学者的研究重心正是人类智能规定的东西，立足于这四个层面之外的人工智能，只能充当助手而已。因此，数位人文研究的核心在人文，所有科技手段的应用都必须围绕人文研究的需要来展开。

这一原则同样可运用到人类智能研究上，也就是说，我们用实验方法研究人的大脑神经网络的结构、认识大脑的各种功能以至于破解意识之谜时，都存在类似于数位人文研究中两类不可化约的变量。一类是因果性（互为因果的网络）的，即可以用神经网络自动机（今日人工智能方法）加以分析（或模拟）的，另一类是内省性的，它类似于人文研究中需要用研究者的智能去理解的内容。对人类智能的深入认识必须同时运用这两类变量，将其投射到用神经网络自动机（即目前所谓的人工智能）的研究上，可以直接得到的推论是，AlphaGo无论怎样通过深度学习进化都不会具有人类智能。



智能研究伴随着对人类意识的探索

我在《关于意识的哲学思考》一文中，指出意识的解释必须将行为主义方法和内省方法结合起来。如果没有行为主义模型，该研究不能通过实验来证明。但意识决不是仅仅靠输入-输出的因果性模式能把握的，故必须把因果模型和内省变量结合起来。在此意义上，必须意识到意识的解释不能排除意识。正因如此，在人类智能研究中，必须发展出一种将科学与哲学结合在一起的新方法。我称之为意识解释的递归方案。

它可以简述如下：在意识（人类智能）的研究中，每一次都需要涉及两批不同性质的变量（参数）。一批是行为主义的变量（参数），我们记为B 集合中的元素，它是可以用实验测量并用来建构科学模型的。另一批是只有通过人的内省才能理解的变量（参数），我们记为C集合。在每一次解释中，必须用B集，因为只有这些参数形成的关系可以用实验证明；但单纯用B集合中的元素是不够的，我们必须附加C集合中的元素；附加之所以可能，这是因为我们有意识，可以用意识理解意识。事实上，正恰恰是基于可理解性，构成的整体模型才是合理的（虽然不是每一个细节都可以用实验检验）。

表面上看，因为解释中包含因素C，对意识的科学解释仍没有最后达成。但我认为，只要采用一种类似定义递归函数的方法，可以将解释中C类因素不断化约，使其为B集合中的元素和更原初的意识因素之组合。即随着每一次解释向前推进一步，研究者必须注意采用C集合中不同的元素（避免循环论证），并去发现更为基本的元素。我们可以期待内省的意识因素越来越少。也就是说，只要意识解释的递归方案不断顺利展开。意识之谜或许在有一天可以解开。

我在《关于意识的哲学思考》一文中还提出，伴随对意识研究的深入，实验测量与人文思辨的界限会越来越模糊。但是这两类变量仍是不可以化约的，其根本原因是人有自由意志，即人的主体性。它不可能被还原为因果性。[8]

被忽略的内核：主体的自由

如果我们承认人是动物进化来的（这一点似乎毋庸置疑），既然AlphaGo这样的智能只相当于动物本能，为什么不能想象如同物种演化那样，具有生物本能的神经网络在学习过程中突然涌现出人类智能呢？这正是当前人工智能讨论中另一个经常使人们进入误区的观点。

确实，能否使用符号是人和动物的本质区别。早在19世纪，德国哲学家恩斯特·卡西勒指出人是会使用符号的动物，人的“智能”可以更准确地定义为创造符号及其使用规则，并自由使用符号来与同类交流的能力。这一定义是当前科学界的共识。前面我在论述人类智能四个基本层面时，每一个层面都涉及创造和使用符号。正因如此，从动物向智人的飞跃亦被简化为发明符号和使用符号，它是用系统演化中的“涌现”来说明的。

最近有十分走红的《人类简史》正是这样概括智能和社会加速进步的机制：“智人主要是一种社会性的动物，社会合作是我们得以生存和繁衍的关键。”“只有智人能够表达关于从来没有看过、碰过、耳闻过的事物，而且讲得煞有其事”，“传说、神话、神以及宗教也应运而生”；而“通过文字创造出想象的现实，就能让大批互不相识的人有效合作，而且效果还不只如此。正由于大规模的人类合作是以虚构的故事作为基础，只要改变所讲的故事，就能改变人类合作的方式”。这使得“智人就能依据不断变化的需求迅速调整行为。这等于开启了一条采用‘文化演化’的快速道路，而不再停留在‘基因演化’这条总是堵车的道路上。走上这条快速道路之后，智人合作的能力一日千里，很快就远远甩掉了其他所有人类和动物物种”。[9]

该书勾勒出人类使用符号对社会形成的意义，以及文化如何加速人类智能的进化，这无疑是正确的。但《人类简史》却没有揭示人创造符号、使用符号沟通、建立社会，以及文化演变背后起关键作用的内核，容易使人忽略人类智能和社会之间的关系，从而导致对智能“涌现”和加速进步的错误概括：这些能力似乎可以仅仅从神经网络自动机的学习过程中自行产生出来。

具体而言，第一，创造和使用符号需要一种仅仅通过反馈学习的神经网络自动机不可能具有的能力，这就是主体的自由。两台AlphaGo可以互相沟通，或许人不知道它们在讲什么，但是它们有“对象”意识吗？这里的关键是，什么是用符号指涉对象？其前提是符号和对象关系的任意性，因为用符号指涉对象，只是主体间的约定。对同一事物，不同语言进行表达的词汇可以完全不一样。《人类简史》一书只从“讲故事”的角度来理解符号使用与人类社会的发展，而没有意识到主体的自由是语言发明的前提。如果没有自由的主体，人只是对外来刺激做出反应，并通过反应效果的反馈来改变神经网络结构的自动机，根本不会发明符号并用它来表达对象。什么是主体？什么是主体的自由？虽然哲学家一直在思考，但至今对其认识不甚明了。近年来，用DNA考古追溯人类起源，已搞清智人起源于20-30万年以前，即一种会使用语言、具有与现代人相同心智的人已经存在，从此，人类将自己和动物界完全区别开。确定无疑的是，在使用语言的背后正是主体的出现和主体的自由，但对于推动这一巨变的机制，可以说是一无所知。

第二，当人通过语言组织成社会后，自我意识开始复杂化。在这一过程中会发生突变，这就是人可以进一步意识到自己（主体）是自由的。“人是自由的”乃是创造符号之前提，而“人知道主体是自由的”是一种二阶思维，即对自由的自我意识。这是主体性又一巨大进步，它通常被称为人具有自由意志。今天经常看到一些研究，认为自由意志只是一种假象，人的行为是因果性的（被决定的）。这些研究之所以陷入误区而不自知，就是在智能研究中忽略内省变量（或以为它可以化约为行为主义变量）。其实，每一个人只要扪心自问，都知道自己是有自由意志的，存在对“善”、“恶”、“对”、“错”进行选择的自由。如果人没有自由意志，道德对人便毫无意义，高度复杂的社会组织以及可继承性的文明不可能存在。人在何种条件下意识到自己（主体）是自由的？对于这一革命性变化之原因，20世纪至今的人文研究已获得了巨大进展，这就是超越突破的发生和轴心文明的形成。也就是说，人的自由意志不仅是社会意识的属性，还是对社会意识反思的产物。

我在《轴心文明和现代社会》一书中将超越突破定义为人从社会中走出来，寻找不依赖于社会的终极价值，让其来组织社会。从此以后，人类社会告别古文明，进入一种有不死文化传统可继承的社会。高级宗教和道德起源了，不死的文化价值系统得以产生，成为智能进一步进化的载体。[10]可惜的是，这种基于人文和历史的研究，虽对于揭示智能的本质不可缺少，但因人文精神一直处于认知科学和人工智能研究之外，不被科学家理解。智力加速发展的机制被简单归为语言和古代社会本身，而不知道轴心文明以前的古文明的文化常在社会演化中灭绝。

智力进步第三种更重要的力量，来自于人终于认识到自己应该是自由的。这是什么意思？人知道自己是自由的，并不意味着个人自主成为社会组织原则。在传统社会，自主的个体生活在终极关怀或道德价值规定的有机体内，他们的合作和创造力的发挥一直受到有机体“容量”的限制。只有人认识到自己应该是自由的，自由的意义才能超越其他价值，成为“应然社会”的组织原则。现代科学的诞生是智能最伟大的革命，它和现代市场经济的出现同步，其背后是一种可以容纳自主个人无限制发挥自己创造性，并和其他主体合作的社会组织方式。这一切可以归为“自主性成为应然社会的组织原则”。人工智能的符号主义学派高度重视数学公理化思维模式，但他们没有想到，这种思维模式只有在现代社会中才能成熟。也就是说，从人类智能摆脱了种种天然桎梏并开始突飞猛进，到人工智能研究有可能出现，都是基于现代社会的起源并日益成为人类文明的组织框架。

结语

一旦理解智能（社会性）进化存在着上述三种推动力量，人工智能革命对现代社会带来的冲击也就一清二楚了。或许有一天神经网络自动机在学习中会发生不可思议的突变：涌现出主体的自由并发展出它们的语言，和其他神经网络自动机建立社会。但它们取代人类社会同样要经历类似于人类社会演变--从部落社会到古文明、轴心文明再到现代社会的产生。也许该过程正在宇宙某处发生，但在外星文明被发现前，这一切和人工智能冲击当代社会没有什么关系。

既然如此，我们又如何看待人工智能对现代社会生活的巨大影响呢？就其本身而言，它和历史上有过的新科技带来的生活方式和生产方式大变革相似，并不会改变现代社会的基本结构。20世纪

初，英国文学大师福斯特在《大机器停转之日》中描绘了如下场景：有台大机器照顾着人类的衣食住行，安排着人类的一举一动。但有一天，这个精心设计的、全知全能的、无所不在的大机器停转了，社会很快陷入崩溃状态。其实，这本是对20世纪极权社会的想象。但20世纪的历史和这一预言相反，现代社会的危机如法西斯主义的兴起与其说是新科技和新生产方式带来，还不如说是民族主义和现代民族国家本身存在着缺陷。

今日对人工智能冲击的忧虑，重要的不是其本身，而是现代社会包容一日千里科技革命的能力急骤地退步，其背后是现代社会立足的两大支柱本身受到科技革命的侵蚀。现代性的前提是科学理性与人文信仰的二分，当科学理性等同于技术，意味着它已异化。“上帝已死”和道德被等同于追求利益，表明终极关怀开始退出社会。如何安顿个人生命的终极意义？一个由没有道德、信仰的人组成的社会是否可能容纳科技和经济的进一步发展？这都是人类文明正面临的巨大挑战。

也许，现代社会的衰落在短期内不可避免，但人类智能长时期后退是不可能的。青峰在《让科学的光芒照亮自己》的修订版序言中，曾提及阿西莫夫的著名科幻系列小说《基地》。这本书的主题是银河系文明面临不可避免的衰落之时，为了缩短黑暗过渡期、保留今后文明复兴的种子，人类在星系遥远的边陲分别建立科学和人文两个基地。阿西莫夫并没有谈及文明衰落的原因，但青峰从他把科学和人文作为两个互不往来的基地的设想中，猜想阿西莫夫很可能认为文明衰落是科技过分发达被专业化异化和人文精神衰落引起的。[11]不管银河帝国的衰落的原因是不是科学的异化和人文精神的丧失，现代社会的重建只能依靠科学和人文两大支柱。

注释：

[1]杰夫·霍金斯、桑德拉·布拉克斯莉：《人工智能的未来》，贺俊杰、李若子、杨倩译，陕西科学技术出版社2006年版，第9页

[2] 集智俱乐部：《科学的极致：漫谈人工智能》，人民邮电出版社2015年版，第一章。

[3][8] 金观涛：《关于意识的哲学思考》，载《科学文化评论》2009年第3期。

[4] 详见W.R.艾什比：《控制论导论》，张理京译，科学出版社1965年版，第281-282页；W. R. Ashby, “Design for an Intelligence-Amplifier”, In Claude E. Shannon and J. McCarthy ed., Automata Studies, Princeton University Press, 1956.

[5] 《赫伯特·西蒙和艾伦·纽厄尔--人工智能符号主义学派的创始人》，载科学网，http://blog.sciencenet.cn/blog-39714-253943.html

[6] 《机械蛮力和人类智能：符号主义和联接主义的魔咒》，载雷锋网，https://www.leiphone.com/news/201509/kON77d06sXwWYY1k.html

[7] 詹荃亦、王乃昕：《“主义”的数位人文研究》，载项洁等主编：《数位人文在历史学研究的应用》，台湾大学出版中心2011年版，第219-245页。

[9] 尤瓦尔·赫拉利：《人类简史：从动物到上帝》，林俊宏译，中信出版社2014年版。

[10] 金观涛：《轴心文明和现代社会》，法律出版社即将出版。

[11] 刘青峰：《让科学的光芒照亮自己：近代科学为什么没有在中国产生》，新星出版社2006年版。

本文原载《文化纵横》2017年08月号。图片来源于网络，欢迎个人分享，媒体转载请获得许可。



- 吴军：大数据能够战胜人的逻辑推理吗？
- 胡凌：“机器人”与“法律人”的“战争”
- 机器之心编译：人工智能革命

文化
纵横

文化重建 价值重生

转载须知

后台回复“转载”获取授权

进入微店

点击“阅读原文-进店”进入

投稿邮箱

wenhuazongheng@gmail.com

阅读原文