

浅论人工智能与哲学之间的关系

汪安平 3220180857 计算机学院 软件工程

摘要：哲学作为联结和沟通各个学科的枢纽，在整个科学领域扮演着极为重要的作用。人工智能作为近些年来爆炸式发展的学科，在解决一些类人脑的活动中取得了十分不错的效果。但同时人工智能也会涉及数学，历史，推理等众多学科，模拟人类的推理、感知、学习等能力，当人工智能向着越来越高的度前进的时候，人们不禁会联想到当计算机具有人类的感知，推理能力，人工智能会对人类带来哪些影响。这时，人工智能是什么？发源于哪里？人工智能的未来是哪里？这些层面很大问题的解答，势必会将哲学和这门极具焦点的学科紧紧的联系在一起。本论文主要从解释人工智能的发展前景方面详细讨论了人工智能与哲学之间紧密的关系。

关键词：哲学；人工智能；神经网络；问题；领域

在讨论人工智能和哲学之间的关系前，我更加的想简明扼要的向大家介绍什么是人工智能，以及什么是哲学，当然本文中的关于哲学是什么的概述仅仅从较为浅层的方面来引入的。

1 人工智能是什么

人工智能，也称作机器智能，指由人创造出来的机器所表现出来的智能。通常人工智能是指通过普通计算机程序的手段实现的人类智能技术。同时，人类的无数职业也正在被这样的人工智能技术所取代。人工智能根据智能的程度可以分为强人工智能和弱人工智能。强人工智观点认为有可能制造出真正的能推理和解决问题的智能机器，而且这样的机器将被认为是具有知觉、自我意识的。弱人工智能认为不可能制造出能真正地推理和解决问题的智能机器，这些机器只不过是看起来像是智能的，但是并不是真正的拥有智能，也不会有自主意识。弱人工智能是对比于强人工智能出现的，因为人工智能的研究一度处于停滞不前的状态下，直到类神经网络上有了强大的运算能力加以模拟后，才开始改变并大幅度超前。这些所谓的弱人工智能在神经网络的发展下已经有巨大的进步，但对于要如何集成成为强人工智能，现在还没有明确定论。

人工智能的研究是高度技术性和专业的，各分支领域都是深入且各不相通的，因而涉及的范围很广。人工智能的研究可以大致的分为几个技术问题。其分支领域主要集中在解决具体问题，其中之一是，如何使用不同的工具来完成特定的应

用程序。人工智能的核心问题包括构建能够跟人类相似甚至超越的推理、知识、学习、交流、感知、移动和操作物体的能力等。在当前阶段，强人工智能已经取得了初步的成果，甚至在视频识别，语言分析、棋类游戏等单方面的能力达到了超越人类的水平，而且人工智能还需要时间研究，比较流行的方法包括统计方法、计算智能和传统意义的人工智能。当前有大量的工具使用了人工智能技术，其中包括搜索和数学优化、逻辑推演等。而基于仿生学、认知心理学，以及基于概率论和经济学的算法等也在逐步的探索之中。思维来源于大脑，而思维控制行为，行为需要意志去实现，而思维又是对所有数据采集的整理，相当于数据库，所以人工智能最后会演变为机器替换人类。这是对人工智能比较官方的定义了。最近几年，人工智能的发展迅速，和我们的生活也越来越贴近了。越来越多的人工智能应用投放到生活生产之中，从出入公司、宿舍等的门禁人脸识别功能，到智能的无人驾驶技术。人工智能的相关学科也同时成为了广大学习者们聚集的“胜地”，譬如大数据，深度学习，机器学习，模式识别等。人工智能的应用领域也在当今社会生活中不断的扩展，其中在自动驾驶，图像处理，语音处理和自然语言处理的过程中尤为突出。

2 什么是哲学

要想对哲学是如何和人工智能之间产生关系有很好的理解，就要求我们对哲学是什么有一个初步的了解，每个人了解的哲学会不同，会有区别。我这里介绍一种比较浅显的易于理解哲学的解释，当然这也可能只是哲学的浅层理解。数学、物理、化学、经济、政治、外语甚至包括炒菜做饭、待人接物，这些都是有学问的，每种学问都能解决某方面的问题。这种学问很多，每种学问中有有一些小学问。无论你的学问是多还是少，你肯定有一套有意或者无意的办法，用于记忆这些学问，这就是哲学。哲学的作用就是用来记忆、关联各种学问。这么说，即使不学习哲学，每个人也都会哲学吗？明确的说是，不是因为不懂哲学，相反，是因为我们每个人都懂哲学，哲学就像空气一样无处不在，似乎不重要但又离不开。大家可能还有一个疑问，如果每个人都是哲学家，那么还要康德、黑格尔这样的哲学家干吗呢？因为每个人的哲学都在自己大脑中，没有把它写在纸上面、表达出来。而哲学家的工作就是将它们表达出来。不同的哲学家差别很大：老子的哲学是老子哲学，康德的哲学是康德哲学，没有对错之分。这跟物理、数学等有用的学问不同，它们是由对错的，如果不按照牛顿力学计算行星轨道，结果就不对；

如果不按川菜的做法炒菜，炒出来的就不是川菜。

每个人的哲学不同，但也有一些共同的规律，我虽然不能照抄你的哲学，但是能借鉴。哲学家将哲学表达出来后，别人就可以借鉴以提高自己的哲学。重要的是：表达出来的哲学可以使全人类支撑所有的知识。数学、物理、化学、经济、政治、外语这些学科就像是一个个陈放知识的盘子，总要有一个更大的盘子来陈放之前的这些各个领域所对应的小盘子，那么这个大盘子就是哲学。同时，在所有的学科中至少有两门基础学科：物理和数学。物理很强大，很基础，化学、生物学等都可以归于物理，但它无法支撑数学；数学也很强大很基础，但是它同样不能支撑物理。这就说明，物理和数学都不是最基础的，一定有一种比它们更基础的、可以支撑的、区分他们的学科，这就是哲学。当然，哲学可以支撑各个学科，并不是哲学能够取代各个学科。或许更多的人会说，哲学不是回答“我从哪来，我到哪去”这些终极问题的吗？其实这并不矛盾。哲学在回答终极问题时需要关联各个学科，而哲学把各种学科关联起来的目的之一也就是解答或者你挂机问题及单一学科不能解答的问题。所以，对于个人还是全人类来说，哲学都是起到支撑、区分、关联各个学科的作用。

3 哲学和人工智能之间的关系

今天是人工智能的时代。在当前的趋势下面，有一点可以肯定：人工智能至少要在一定程度上像人一样记忆各种知识。也就是说，人工智能需要有一种记忆知识的规则，它需要懂哲学。知识图谱是 google 等人工智能公司采用的知识存储技术，它是基于本体论的，而本体论就是哲学的一种理论。不难理解：人工智能所能够达到的高度，在一定程度上取决于它所采用的哲学。今天还是一个全球大变动的时代，新经济、新科技、新思潮汹涌澎湃。宇宙的本质是什么？人类的未来是什么？人工智能能否取代人类？解答这些终极问题终究是离不开哲学的。其实我之前的这些讨论都是有一个重要的前提才能够成立，就是：哲学有支撑、关联各个学科的作用，而不是和这些学科并列的学科。

人工智能是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。这样学术性的解释，恐怕很难于理解人工智能到底为何物。我们可以换个角度，从哲学的层面来说一说“人工智能是什么”、“人工智能从哪里来”以及“人工智能要到哪里去”这三个问题。

首先，人工智能是生产力发展的产物。哲学一直都不是追究“什么”是“什

么”，更多是追求为什么“是”和“如何是”的问题。那么人工智能为什么被称为人类智慧的容器呢？大家所达成的共识是，人工智能被研究出来的一个重要的目标就是是机器能够胜任一些通常需要人类智能才能够完成的复杂工作。在1954年，美国人戴沃尔设计了世界上第一台可编程机器人，在当时，生产力还没有那么高度发展的时代，它的发明创造可能对于“复杂工作”的意义不大。在1981年的时候，日本经济产业省拨款8.5亿用于研究人工智能计算机项目，随后，英美等国纷纷响应，开始向信息技术领域的研究提供资金。这时候，人工智能的意义逐渐的凸现出来。从实用主义的角度上来说，人类已经逐渐布局人工智能发展战略，已获得更大的经济效益。从哲学物质与意识的辩证关系上来说，人类已经逐步认识到，人工智能被生产出来是服务于人类智能的。

再来说，人工智能“如何是”人类智慧容器的体现。1984年，美国人莱纳特带头启动了Cyc项目，其主要的目的是是人工智能的应用能够完成类似人类推理的工作。人工智能技术中心的应用，我们认为是外界环境、机器和人的认知感知器共同作用的结果，三者缺一不可。在我们人类自己都很难解释清楚自身“智能”到底为何物的有限认识下，对定义什么是“人工”制造的“智能”就更加的困难了。

关于“人工智能从哪里来”，这里浅述个人愚见。人工智能可能来源于人类的匮乏感，因为匮乏，因为有需求，甚至可以说是因为人类对未知探索的兴趣，甚至是为了解决一些可以被解决的人类社会中的问题，人工智能就慢慢的出现了。起初他可能是一些类似于机器人，机器的定义，随着这些工具越来越高级，能完成人类才能完成的复杂工作，人工智能这个词也逐渐的进入人们的视野。它能够对人的意识和思维的信息过程进行模拟，解放人类部分的“思考力”。有了高准确率的计算人工智能，我们解读数学问题的部分思考能力被解放了，人类可以有更多精力更高效地做其他人工智能暂时不能涉及到的领域。

最后思考“人工智能要到哪里去”的问题，我想借用康德认为哲学需要回答的三个问题来说一说。即“我能知道什么”、“我应该做什么”、“我可以期待什么”这分别对应于认知、道德和信仰。目前来说，我们认为人工智能是没有意识的，它对事物的认知，也是工程师输入代码编程的产物，那么随着科学技术的发展，未来人工智能会不会有意识、道德和信仰呢？这是值得我们人类去深思的问题。未来的人工智能极有可能是集科学技术、人文艺术、哲学伦理为一体的“有机化合物”，是各种有限认知和无限想象相互叠加和相互碰撞的结果。对于人工智能

的哲学思考还需要不断强化，以得到更新的认知。

人工智能是需要哲学实时参与其中的。哲学要做的第一件事就是思考大问题，澄清基本关系。与哲学家相比较，一般的自然科学家往往只是在自己的研究领域中预设了相关问题的答案，却很少系统的反思这些答案的合法性。还有就是，哲学的特性，哲学在不同学科的研究成果之间虚招汇通点，而不受某一具体学科视野之局限。举一个例子，用军事上的彼方，哲学更像是战略性的思考。如果你是在一个炮兵学院里面，不同的研究炮兵战术的军官会讨论炮兵战术所牵扯到的具体的几何学问题。但是站在战略层面，它可能对于这些比较细微的问题忽略掉，更多地考虑炮兵在军事编制中所扮演的角色，站在更高的层面上看。这可能帮助大家理解哲学应该是干什么的。

AI（人工智能）的研究目的，即是在人造机器上通过模拟人类的智能行为，最终实现机器智能。很显然，要做到这一点，就必须对“何为智能”这个问题做出解答。如果你认为实现“智能”的实质就是去尽量模拟自然智能体的生物学硬件。你就会去努力人脑的结构，并用某种数学模型去重建一个简化的神经网络，这也是联结主义者所做的。现在我们都知道有一个类脑研究计划，这种研究有复杂版本和简单版本，复杂版本就是蓝脑计划一样，把大脑运作的信息流程尽可能逼真的模拟出来，比较简单的就是简化的神经网络。倘若站在专业的脑科学研究者的立场之上，神经网络很不神经，离真正的神经活动来说，它是高度简化，但是站在宏观的立场之上，至少你说神经呀 UN 网络也是受大脑的启发和影响。这个路线很多人认为是对的，我认为可以做出一些成果，但是不要指望它能在各项应用中达到我们人类的要求。如果你认为人工智能的实质仅仅在于智能体在行为层面和人类行为的相似。那么你就会用尽一切办法来填满你理想中的智能机器的“心智”。由此看来，正因为自身研究对象的不确定性，AI 研究者在哲学层面上对于“智能”的不同理解，才会在技术实施的层面上产生如此大的影响。很明显，人工智能学科内部的基本分歧，在相对成熟的自然科学那里是比较罕见的。其次，AI 科学资深的研究手段，缺乏删除不同理论假设的决定性判决力，这在很大程度上也就为哲学思辨的展开预留了空间。

人工智能领域的哲学都具体体现在哪些方面呢？现实是很多人工智能的工作者觉得哲学讨论是空谈，殊不知他们的很多基本观点都是属于哲学的。一般来说，任何一个科学技术领域都有哲学观念的问题，因为每个理论都有其逻辑起点，包括基本预设、覆盖范围、直观背景等。这些问题上的争论往往不能完全在该领

域之内得到解决，而需要在更大的尺度和范围之内进行考虑，也因此而进入了哲学的范围。由于人工智能的研究对象涉及智能、认知、四维、心灵、意识等在哲学中反复被讨论和使用的概念，这哲学的关系比其他领域更加地密切和复杂。人工智能和其他科学技术领域最大的不同是其中研究规范的复杂。一个领域内有多个相互竞争的规范并不奇怪，但在其他领域中，主要争论的是对现象的解释和对问题的解决方法，而在人工智能领域中争议首先是什么是要解释的现象和解决的问题。人工智能的直观意义很简单，让计算机像人脑那样工作。但是计算机毕竟不是人脑，也不可能在所有的方面都能够像人脑，因此一个人工智能系统只能是在某些方面像人脑，而在其他方面可能完全不像。在这一点上，主流观点是把智能理解为“能解决那些人脑可以解决的问题”的能力。这种把“智能”看作“解题能力”的观点实际上是一个关于智能和思维的哲学信念，尽管很多人工智能工作者是将其作为一个自明的前提来接受的。

一些人工智能的批评者感到了人工智能在这方面的问题。他们正确地指出人的思维不能再人工智能现有的理论框架中得到解释，但错误地以此作为“真正的人工智能是不可实现”的理由。这些人工智能的批评者和他们的批评对象一样，把计算机的一种特定的使用方式，即遵循算法来解决问题，当做了其唯一可能的使用方式，因此不能为这个领域的突破指引方向。同时把“智能”看做一组算法的另一个后果就是人工智能领域的身份危机。“只有人脑可以解决的问题”是一个随时间而变得概念。目前人工智能如此流行，以至于之前叫自动化和计算机应用的工作都归于此名下。其结果是无法说清人工智能和计算机科学技术其他领域的区别何在，更遑论建立一个统一的理论基础。同时还会造成人工智能领域的碎片化。把每个问题独立定义和解决的过程中，各个认知功能之间的内部联系被割断了。其结果之一就是这些功能在很多所谓人工智能的系统中的表现和它们在人的思维活动中的不同。对于那些主要以人工智能作为研究人类智能或一般智能的途径的研究者来说，这种“分而治之”的办法是有缺陷的。即使完全从应用的需求来看，各个认知功能的协调运用在解决问题过程中常常是不可或缺的。以自然语言为例，说“语言理解不需要推理”听上去明显有问题，但至今语言理解系统一般都没有多少推理能力。

正是上述考虑导致了近年来“通用人工智能”研究的发展，其主要特征就是对智能的通用性和整体性的强调，以及对主流人工智能基本预设的挑战。要理解这一类研究工作和传统人工智能的区别，我们需要重新审视“智能到底是什么”

和“怎样在计算机中实现智能”等有哲学意味的问题。通用人工智能和其他人工智能系统的不同很多都可以追溯到对智能的理解。在我看来，智能是“一个适应系统在知识和资源不足的条件下工作的能力”。在这种条件下工作意味着系统中的所有来自过去经验的知识都可能被未来经验挑战，而且系统在解决一个问题时一般没时间考虑到所有的相关知识，因此无法保证所有结论都是绝对正确或最优的。这种情况下的理性只能是一种“相对理性”，即适应性，就是说系统只能用过去经验来尽力应对目前的新情况，用有限的资源来尽量满足当下的要求。按照这种想法设计的系统和目前常见的主流人工智能系统有所不同。由于系统经验和资源需求随时间变化，系统对一个问题的解答往往不是固定的，因此“问题”和“解答”不能被看成经典意义下的“计算”或“函数”，并且解题过程也不遵循一个确定的“算法”，而是需要具体问题具体分析。即使是同一个问题实例，在不同的情形下所得到的处理也可能很不同。

这个意义下的人工智能不再是传统计算理论所能涵盖的。实际上这二者是互补的：前者研究知识和资源不足时的工作方式，而后者研究知识和资源充足时的工作方式。智能不是直接体现在系统解决问题的能力上，而是体现为一种“元能力”（即获得解决问题能力的能力）。在这个意义下的智能自然是通用的，而系统解决某一领域中的问题的能力完全来自于它自身的经验，包括通过感知运动界面得到的直接经验和通过语言界面得到的间接经验。因此，这种对智能的理解导致计算机在信息加工原则上向人类靠拢，而非复制人的个别解决问题能力。尽管让计算机解决各种问题有其重大价值，但这种工作并不一定加深我们对思维一般规律的理解。人工智能几十年的历史已经充分说明了这一点。当现有的理论不能满足问题的要求时，正确的办法是尝试建立新理论，而不是削足适履，把问题限制在已有理论所能覆盖的范围。

4 总结：哲学决定人工智能的高度

像科学史上其他重大课题一样，人工智能对人类既是个挑战又是个机会。因此，对其社会后果抱谨慎态度是绝对必要的，尤其是要尽量防止仓促地把理论成果投入实际应用。对人工智能安全问题的研究必须以对相关理论和技术问题的深入了解为前提。具体说来，由于真正的智能系统必须是有适应性的，其行为不仅是被其（先天）设计所决定的，更取决于其（后天）教育、训练、使用等环节，而后者往往在有关的讨论中被忽视。

由于人工智能的领域特征，其中有大量的哲学问题。我们甚至可以说只有在哲学上正确的方案才存在技术上成功的可能。但不幸的是，很多人工智能工作者缺乏在哲学层面上检讨其理论预设的兴趣和能力，而哲学工作者又常常缺乏深入了解技术问题的勇气和基础。在人工智能领域，人工智能工作者和哲学工作者相互理解的努力会必将促使双方受益，所以在以后很长一段时间内，哲学对人工智能的贡献将决定人工智能在这段时间所能够达到的高度。

参考文献

- [1] 维基百科-人工智能
- [2] 维基百科-人工智能史
- [3] 李开复，王咏刚. 《人工智能》. 文化发展出版社，2017
- [4] 吴军. 《智能时代：大数据与智能革命重新定义未来》. 中信出版社，2016
- [5] 尼克. 《人工智能简史》. 人民邮电出版社，2017
- [6] 徐英瑾. 《心智、语言和机器——维特根斯坦哲学与人工智能科学的对话》. 人民出版社，2013
- [7] 董军. 《人工智能哲学》. 科学出版社，2015
- [8] 毛华滨，高新民. 从意向性看人工智能的发展方向[J]. 科学技术与辩证法，2005，(5)：30.
- [9] McCorduck, Pamela. Machines Who Think[M]. Natick, MA: Alice and Klaus Peters, Ltd, 2004: 5.
- [10] 杰夫·霍金斯，桑德拉·布拉克斯莉. 人工智能的未来[M]. 贺俊杰，李若子，杨倩，译. 西安：陕西科学技术出版社，2006：215.