

Philosophical Thinking of Artificial Intelligence

顾 骏 主 编
郭毅可 副主编

人与机器：

思想人工智能

上海大学出版社

专题课题“全员育人：‘同向同行’的平台
系列课为例”，项目批准号：17JDSZ1013
“顾骏团队”

顾 骏 主 编
郭毅可 副主编

• 人文与智能丛书 •



人与机器： 思想人工智能

Philosophical Thinking of Artificial Intelligence

上海大学出版社

· 上海 ·

图书在版编目(CIP)数据

人与机器：思想人工智能/顾骏主编. —上海：
上海大学出版社, 2018. 8

ISBN 978 - 7 - 5671 - 3208 - 5

I. ①人… II. ①顾… III. ①人工智能—科学哲学
IV. ①TP18 - 02

—中国版本图书馆 CIP 数据核字(2018)第 173878 号

责任编辑 傅玉芳 庄际虹

徐雁华 陈 强

封面设计 柯国富

人与机器：思想人工智能

顾 骏 主 编

郭毅可 副主编

上海大学出版社出版发行

(上海市上大路 99 号 邮政编码 200444)

(<http://www.press.shu.edu.cn> 发行热线 021 - 66135112)

出版人 戴骏豪

*

南京展望文化发展有限公司排版

江阴金马印刷有限公司印刷 各地新华书店经销

开本 710 mm × 1000 mm 1/16 印张 20.75 字数 231 千

2018 年 8 月第 1 版 2018 年 8 月第 1 次印刷

ISBN 978 - 7 - 5671 - 3208 - 5/TP · 068 定价 38.00 元

目 录

前 言 人工智能之“天问” / 001

- 一、人工智能的形而上意义何在? / 004
- 二、被颠覆者有能力前瞻颠覆者吗? / 008
- 三、人工智能仅止于模拟人类智能吗? / 011
- 四、“单项型”能战胜“全能型”吗? / 012
- 五、不具有非理性成分是人工智能的劣势吗? / 016
- 六、人工智能需要自我意识吗? / 019
- 七、人工智能还是无机智能? / 021
- 八、算法是“道成肉身”吗? / 024
- 九、“零的增长”是可能的吗? / 028

第一章 图灵测试到底灵不灵? / 031

- 一、人工智能为何言必称图灵? / 033
- 二、图灵机与图灵测试有何共同之处? / 035
- 三、图灵测试有没有破绽? / 040
- 四、人类逻辑体系的起点在哪里? / 042
- 五、预设前提：图灵测试的问题在哪里? / 045
- 六、人类不该为自己的智能而骄傲吗? / 047

七、走出人类中心观：平行智能是可能的吗？ / 051

八、如何为图灵测试“打补丁”？ / 053

九、为什么有破绽的图灵仍然很灵？ / 054

十、思想超越：科学是如何发展的？ / 055

第二章 机器智能是如何生成的？ / 057

一、机器智能是怎么发展而来的？ / 059

二、ABC 时代，机器智能会超过人类吗？ / 065

三、机器智能主要有哪些内涵模块？ / 068

四、机器如何获得智能？ / 072

五、人需要向智能机器人学习吗？ / 081

六、机器智能到底模拟人类什么？ / 083

第三章 赢了围棋，机器就能战胜人类吗？ / 087

一、为什么是围棋？ / 088

二、围棋到底有多难？ / 091

三、五代 AlphaGo 如何横扫围棋高手？ / 092

四、AlphaGo 的原理何在？ / 094

五、机器智能与人类智能孰优孰劣？ / 097

六、人机分野：确定还是不确定？ / 100

七、世界是确定的吗？ / 102

八、机器智能的发展有边界吗？ / 105

九、人机交流是可能的吗？ / 108

第四章 “小冰”作品的诗意哪里来? / 111

- 一、诗歌何以人文? / 112
- 二、机器诗人真的技胜一筹? / 114
- 三、“小冰”诗作有何高妙? / 117
- 四、机器诗人如何创作? / 120
- 五、机器诗人如何看图作诗? / 126
- 六、机器诗人真能创作? / 128
- 七、机器诗人作品的诗情画意何来? / 132
- 八、人-机共融能促进艺术创作吗? / 134
- 九、科学家让“小冰”写诗意欲何为? / 136

第五章 机器人坐堂会让医生失业吗? / 139

- 一、医生为何要读片? / 141
- 二、医生读片需要有什么准备? / 143
- 三、医生如何读片? / 145
- 四、医生读片失误能避免吗? / 147
- 五、读片的认知意义是什么? / 150
- 六、机器是如何读片的? / 154
- 七、机器读片水平超过医生了吗? / 159
- 八、机器读片有什么优势和弱点? / 161
- 九、机器读片还将如何发展? / 163
- 十、未来医生和计算机如何联手造福人类? / 166

第六章 机器人独霸股市会颠覆市场秩序吗？ / 169

- 一、作为风险游戏，股市如何一路走来？ / 171
- 二、在风险集聚的股市上，人如何操作？ / 175
- 三、机器人如何做股票？ / 177
- 四、自动交易是如何构成的？ / 180
- 五、与人相比，机器人做股票有何优势？ / 182
- 六、机器人操作股票战绩如何？ / 185
- 七、机器人独占股市，会造成“齐步走”吗？ / 187
- 八、股市作为价格发现机制，会被算法取代吗？ / 189
- 九、机器交易做预测依据何在？ / 191
- 十、机器能让计划经济重新成为主流吗？ / 200
- 十一、有了机器智能，人类经济生活中的集体性智慧会被颠覆吗？ / 204

第七章 人工智能与机器人如何结伴而行？ / 207

- 一、机器人如何一路走来？ / 209
- 二、体力劳动和脑力劳动，哪个更容易模仿？ / 211
- 三、机器人为何难以拥有人类的“灵活性”？ / 214
- 四、机械智能与机器智能，区别在哪里？ / 217
- 五、机器人双足行走难在何处？ / 219
- 六、除了运动，人工智能还能给机器人带来什么？ / 223
- 七、机器人是如何训练成功的？ / 226
- 八、智能狗和狗智能也是平行关系吗？ / 227

第八章 机器人之间也有伦理关系? / 231

- 一、智能只存在于个体身上吗? / 233
- 二、世界上有哪些群体智能形态? / 234
- 三、为什么要让机器人具有群体智能? / 236
- 四、智能无人系统已具有群体智能了吗? / 240
- 五、机器人也会遭遇“囚徒困境”? / 242
- 六、伦理要求如何让群体走出“囚徒困境”? / 243
- 七、低智生物如何实现伦理要求? / 244
- 八、人类社会如何落实伦理要求? / 248
- 九、智能无人系统如何植入伦理安排? / 250
- 十、在机器人个体身上如何落实伦理要求? / 254
- 十一、智能无人系统的群体智能有边界吗? / 256

第九章 中国机器人何时成为机器中国人? / 259

- 一、仿生的本质是什么? / 262
- 二、模拟人类智能如何从自然人走向文化人? / 263
- 三、文化在哪里? / 264
- 四、当下人工智能研究的基本方法是什么? / 267
- 五、人工智能的算法是如何运演的? / 269
- 六、科技发展中西方思维优势何在? / 274
- 七、西方思维方式独领风骚,风险何在? / 276
- 八、未来人工智能研究能从中国文化汲取什么? / 278
- 九、二进位制的算法能与八卦对话吗? / 279

十、遵循形式逻辑的算法与中国顿悟式思维能对接吗？ / 282

十一、专注知识的人工智能与中国式道理能对接吗？ / 284

十二、专注于处理现象的人工智能与致力于透视表征的中国思维能对接吗？ / 288

十三、作为人工智能基石的图灵机与中国文化的“发现逻辑”能等价吗？ / 291

结 语 人工智能何时迎来“量子称霸”？ / 297

一、量子计算机从何而来？ / 299

二、量子力学何以既令人兴奋，又让人挠头？ / 301

三、量子计算机的历史意义何在？ / 304

四、量子计算机会带来人工智能的革命性改变吗？ / 306

五、量子人工智能的首要目标是什么？ / 309

六、量子称霸，谁该称王、称后？ / 310

后 记 / 313

前言

人工智能之“天问”

“人无远虑，必有近忧。”

——孔子《论语·卫灵公》

“先事虑事，先患虑患。先事虑事谓之接，接则事优成。先患虑患谓之豫，豫则祸不生。事至而后虑者谓之后，后则事不举。患至而后虑者谓之困，困则祸不可御。”

——荀子《荀子·大略》

人工智能是当下最受关注的技术创新和产业趋势。从积极方面说，人类能够利用人工智能提高生产和生活效率，这已成共识，可以讨论的只是在多大程度上实现这一点。从消极方面说，“第四次革命”“奇点”“颠覆性影响”等说法，形形色色，传递出人类对人工智能既爱又怕的复杂心态。为了利用人工智能，应对可能的冲击，前瞻人工智能可能给人类带来负面影响的研究越来越多，这同样是必要的，

只有把各种可能性都做了考虑,才能未雨绸缪,确保人工智能开发不至于走得过远,集聚失控风险。

问题是,目前国内有关人工智能风险的研究大多未涉及人工智能本身的属性,即使有,多半也只是一些碎片化言论,缺乏完整论述,更未对所取立场、视角和方法作方法论反思。因此,此类研究虽则在具体问题上见仁见智,提出了不少有价值的看法,但基本上停留在形而下的层面,将人工智能等同于人类已有的创造,并以人类过往对科学技术的有效管控为参照,对人工智能与人类智能的关系,采取简单乐观的态度,至多针对就业岗位减少等表层效应,提出预测和应对方案,而就人工智能开发整体上表示审慎态度和长远忧虑的论述极少。客观地说,如此态势对人工智能的发展来说,构成了相当友好的环境,但就正视人工智能可能带来的颠覆性影响,充分评估潜在的负面效应,提前做好风险应对的准备而论,则是明显不够的。一旦出现未曾预料的新现象、新进展、新证据或新后果,现有的前瞻研究很可能被证伪,无法起到科学预测应有的作用。事实上,从“深蓝”到 AlphaGo,遵循“摩尔定律”的人工智能,以其神速发展和已达到的智慧水平,一再令人大感意外。按照这样的轨迹和速率,在未来的日子里,人工智能震惊人类的频率和烈度,只会趋于密集和增高,盲目乐观将被证明是不明智甚至愚蠢的。作为对人类共同命运的关注,中国的人工智能风险研究尤须走出表层思考,努力提升思维层次,形成与发达国家学者对等交流的态势。从确保前瞻研究的方法论严谨和预测可靠出发,保持适度超脱,站在哲学高度,严肃提出问题,破除遮蔽,擦亮思维,激发洞察,才能站上前瞻研究的制高点。

一、人工智能的形而上意义何在？

关于人工智能最简洁也最直击要害的定性描述是，人工智能有可能颠覆人类智能，所谓“奇点”就是这一天到来的时间节点，据说从现在算起，还有30年左右的时间。英国物理学家、哲学家霍金生前多次提出明确警告，人工智能的崛起可能是人类文明的终结，“我们不确定我们会从人工智能得到无限帮助，还是被无限边缘化，甚至毁灭”。

霍金是为数不多的敢作“盛世危言”的科学家。尽管他对人工智能和利用人工智能造福人类，总体上抱乐观态度，但一个“不确定”已清楚地表现出这位科学家、哲学家的思想独立和思维严谨。仅仅“颠覆性影响”这个说法本身，就足以让人不敢在尚未进行合理而系统的质疑前，就对人类智能与人工智能的关系，想当然地做过于乐观的遐想。

现在的问题不是霍金的担忧是否杞人忧天，而是我们能否在高于单纯技术的层次上思考人工智能的前景，进而思考人类智能乃至人类的前景。在此，借鉴中国传统文化的三对范畴——体与用、道与器和天与人，逐级提升，对人工智能的形而上性质及其与人类智能的关系，提出一些问题，并做方法论的阐释。

1. 人工智能是“体”还是“用”？

“体”与“用”是中国哲学的基本范畴。西学东渐之后，中国学者提出“中学为体，西学为用”，主张以中国文化为本体，西方文化为效用，利用西方文化来帮助中国实现文化更新。这种封闭态度随即遭到批判，但体用之别的视角是有价值的，中国坚持走自己的路，始终

强调“中国特色的社会主义”，就有类似策略隐含其中。

套用这一思路，着眼风险的研究首先必须确定人工智能到底应该被看作具有自身存在和固有逻辑的“本体”，还是仅仅服务于人类目的的某种“效用”，即为人类意愿所左右的技术。如果仅承认其效用，那只需像当年对待电能和石油、现在对待高铁一样就行，没有必要过多忧虑，什么对人类的颠覆性影响，完全无从谈起：不成其为“本体”，何来人工智能取代人类智能的问题？

反过来，如果把人工智能置于“本体”地位，则必须考虑，引入人工智能是否意味着在本来由人类智能一统天下的世界上，出现了竞争对手？一元的智能世界是否就此变成二元的智能世界？两种智能的关系又将如何？未来的趋势又会怎么演变？如此这般的问题就不再是借助人工智能来提升人类智能那么简单，也不只是人类智能和人工智能到底谁会最后统治地球的问题，还将拓展到诸如外星文明、宜居星球等更具科幻感的宏大话题。

在天文学领域中，人类曾经跨出勇敢的一步，从“地心说”进到了“日心说”，基督教宇宙秩序从此被颠覆。在人工智能领域，人类是否也会迈出同样勇敢的一步，在看待人工智能时从“用”进到“体”？这两项思想行动之所以称得上勇敢，因为其挑战的是同样的思维误区、同样的刻板模式和同样的自我遮蔽——人类中心主义。

2. 人工智能是“道”还是“器”？

“道”与“器”是中国哲学的另一对基本范畴。孔子说“君子谋道不谋食”，“朝闻道夕死可矣”，还说“君子不器”。“道”具有形而上的属性，是世界的本源，而“器”只有形而下的属性，用于满足实际生活

需要。今日所谓“人才”者，在孔子看来，差不多只是用于完成功能性事务的人形工具。简言之，“道”与“器”这一对范畴问的是：人到底是为了超越性的更高存在而存在，还是仅为生存而生存？

把这一思考方式运用于人工智能，虽不够严谨，却仍有价值。因为一旦承认人工智能具有本体属性，必然会提出进一步的问题，即人工智能究竟是作为一种更高存在，比如“道”的实现形式，还是仅止于其自身的工具性存在，即所谓的“器”？换言之，人工智能的发展是按照某个外在指令不断演化，还是仅仅依靠人类获得技术层面的逐步完善？

如果视之为“器”，那只需要讨论如何把人工智能设计得更完善，所谓负面影响即便纳入考虑，也仅限于实用层面，根本不用关心“颠覆性”问题：人工智能跳不出人类手掌，能带来多大毁灭性影响？充其量不过技术失误，就像核电站泄漏，不是核能挣脱了人类控制，只是人类在设计、制造或运行中出现了失误或疏忽而已。所以，前瞻也只需要从人类的角度，注意避免人工智能可能存在的缺陷和因此导致的风险。

相反，如果视之为“道”，那就必须承认发展人工智能必定服从某种固有逻辑，顺应某种规律。因此，关于人工智能的风险研究必定致力于揭示其逻辑指向，找到必然趋势，然后才能预测相关技术会产生什么影响，包括负面影响，否则，预测难免成为一堆碎片，而前瞻也将成为鼻子底下的探险之旅。

3. 人工智能是“天”还是“人”？

承认了人工智能是“体”，反映了“道”的要求，接踵而来的问题

是,这个“道”是“天道”还是“人道”?也就是说,人工智能作为模拟智能,模拟的到底是存在于宇宙中的智能一般还是人类的特殊智能?

中国传统文化信奉“天人合一”,所以,天道即人道。但既然存在天道和人造两种形态,那彼此的关系问题就不会是没有价值的。“人造服从天道”是中国文化的根本信念。老子说:“天地不仁,以万物为刍狗。”在这里,“仁”体现人造,属于人文价值,居于核心地位,但在“天道”运行范围内,却无足轻重。司马迁说的“究天人之际,通古今之变”,意思更加明白,天与人是有区别的,而边界在哪里,就是一个好问题。

把这一视角运用于人工智能研究,如果其模拟的是“人造”,那么模拟者要超越被模拟者,就相当困难,人类意志决定了人工智能往哪个方向发展,发展到什么限度,一旦人类发觉可能遭遇重大风险或威胁,可以直接中止人工智能的进一步研发。

如果人工智能模拟的是“天道”,那无论人类怎么干预,人工智能总会遵循固有逻辑,沿着不受人控制的方向发展下去。在这个过程中,自然指令才是人工智能发展的最终动力,超越人类智能因此是完全可以想象、绝对符合逻辑的。毕竟,至今没有人能证明,人类智能是宇宙最高智能形态,不可超越。“道可道,非常道”,人类智能也好,人工智能也好,都只是智能一般的实现形态,不是智能一般本身。所以,仅凭对人类智能的不完全认识,不足以断言智能一般是怎么回事,也就无法断言人类智能必定凌驾于人工智能之上。事实上,智能机器人有了深度学习能力之后,超越人类智能的现象已多次出现。警报早已响起,研究者不能选择性失聪。

最近有报道称,在 Facebook 进行的研究项目中,两个智能机器

人在对话中自发生成了人类无法理解但确有意义的语言，出于审慎的考虑，项目被暂时中止了。

对于这样一个案例，站在人工智能只是模拟“人道”的立场上，会相信人类的理性和控制力，只要关闭研发的主动权仍在人类手里，就不用担心人工智能会颠覆人类智能。而站在“天道”的立场上则会推测，无论人类看没看到威胁，愿不愿意接受风险，都将身不由己地继续研发，即便出现更为骇人的情形，也不得不任其发展，因为有超越人类的自然机制在控制着人类本身，驱使人类有违本意地放任人工智能发展，哪怕最后毁灭人类自身。

纵观世界，人类装备核武器、制造巨量垃圾、严重污染环境，且应对无方，早已准备好了毁灭自身的条件，不差人工智能这一条。如果自然界演化规律确实只是让人类做一回“天地之过客”，那在人工智能研发的关键时刻，人类的自由意志将只会被用于开启自己墓道的闸门！

这三个维度以及内含的两个不同方向，对于前瞻人工智能都不是没有意义的，但风格不同，层次不同，关注点不同，最后结论也将完全不同。考虑到目前关于人工智能的研究在“用”“器”和“人”的方向上比较多，基于补缺的目的，这里有意更多地往“体”“道”和“天”的方向上，来前瞻人工智能及其可能的风险。

二、被颠覆者有能力前瞻颠覆者吗？

理性是人类认识和改变世界的终极手段，但现代社会科学尤其是社会学恰恰产生于对于人类理性的悖论性认识：社会学在认识到