

徒量智不足以为人：从人禽之辨到人机之辨

郭海鹏

摘要：在儒家思想中，人之为人的本质是“仁”，其内容包含亲爱人的真情实感、反身克己的修养功夫、先天内在的良知本心、万物一体的感通性以及生生不息的创造性。相比而言，AI是“不仁之物”，是由“数据+模型+算法”构成的抽象数学对象和机械程序，无法与人类智能、心灵和意识相提并论。徒量智不能为人，人的性智或内在仁德如果不能彰显，或将堕落乃至麻木不仁之物也未可知。因此真正的危险并非来自AI，而是来自人自身的失衡畸形发展，即工具理性无约束的单向度发展导致人的异化、物化的危险。只有遵循“正德、利用、厚生、惟和”的思想，使性智和量智齐头并举，以达思修交尽，二智圆融的境界，才能发展出一个安全、良善的AI技术，以及人与AI各尽其性的新时代。

关键词：儒家思想；人工智能；人机之辨

中国分类号：G122 **文献标识码：**A **文章编号：**1673-5706（2019）02-0060-08

一、AlphaGo带来的无谓恐慌：机器真的战胜了人类吗？

2015年10月，名为“阿尔法围棋”（AlphaGo）的人工智能（AI）机器人，在没有任何让子的情况下，以5：0完胜欧洲围棋冠军、职业二段选手樊麾。2016年3月，又以4：1的总比分战胜围棋世界冠军李世石九段。随后，在中国棋类网站上以“大师”（Master）为注册帐号与中日韩数十位围棋高手进行快棋对决，连续60局无一败绩。2017年5月，在中国乌镇围棋峰会上，又以3：0的总比分碾压排名世界第一的世界围棋冠军柯洁，正式登顶成为“世界围棋第一高手”。在之后的闭幕式上，中国围棋协会宣布，授予AlphaGo围棋九段称号，成为中国棋院的第43名九段棋手。

AlphaGo的一系列胜利被视为是AI研究的一个突破性的里程碑，令AI领域的专家们也颇为吃惊，因为之前估计AI至少要花5到10年才能攻下围棋这一高复杂度的智力游戏。随着AlphaGo的胜利，现在基本上可以说所有的桌面棋类游戏都被AI征服了。相比之前的IBM的国际象棋AI程序“深蓝”，AlphaGo的深度学习算法更具普遍应用性，可能是向“通用AI”方向迈向的一步。对于社会大众来说，因为围棋一直以来被认为是人类智能的专属领地，AlphaGo的胜利令世界震惊，给人们带来了巨大的心理冲击，纷纷惊呼AI的威胁来了，甚至开始担心人类未来是否要被AI取代甚至奴役。作为一种革新性技术，AI对人类社会和经济发展的重要性毋庸置疑。与以往任何一项技术革新一样，AI发展也可能带来

响应的负面后果，如失业问题和加剧贫富分化等等，确实应该给予一定的重视，但是过度的恐慌实在是不必要的，大可不必过分渲染人工智能的威胁。

围棋是一种信息透明、结构明确且可用规则穷举的对弈游戏，AlphaGo 能够解决围棋问题并不代表它也能解决其他更加复杂和具不确定性的应用。事实上，AlphaGo 只是一个计算机程序，只能在人类给定的目标下进行模拟人类智能某一具体活动的优化计算，并不具有自主思考、设定目标的能力，更谈不上自我意识。单从 AlphaGo 的胜利就得出“机器战胜人类”的结论是无谓的恐慌。恰恰相反，AlphaGo 的胜利是人类智能的胜利，是计算机科学技术计算能力和软件算法设计方面进步的体现。同时，AlphaGo 的胜利也从侧面说明了人类智能和学习能力的威力。以柯洁为例，六岁开始学棋，十八岁获得第一个世界冠军，假设十二年间每天训练下四盘棋，加起来也不过是一万七千多盘，而 AlphaGo 进行学习训练动辄都是数百万甚至上千万局，比有史以来所有人类下过的棋局都要多得多。相比之下，人类棋手的学习效率则要高得多，围棋大师吴清源从 13 岁到 59 岁间下过的精彩棋局也不过 800 多局。^①最根本的是，AlphaGo 只能按照设定的程序进行机械冰冷的计算，永远也体会不到下棋的快乐。

二、机器会思考吗？

1950 年，图灵在《计算机与智能》一文中开篇就问到：“我提议考虑这样一个问题：机器能思考吗？”并提出了著名的用以检验人工智能的方法“图灵测试”。在图灵测试中，把测试者和被测试者（一个人和一台机器）分隔开，由测试者向两者随意发问，如果测试者无法分辨被测试的哪个是机器，哪个是人，就可以说被测试的机器通过了图灵测试，并被认为具备人的智能。^②

人工智能也被称为机器智能（Machine

Intelligence），其主要目的是为了构造能象人类那样思考的机器。1956 年达特茅斯会议标志着 AI 的诞生，到会的专家们确定了用“人工智能（Artificial Intelligence）”一词作为本研究领域的名称，并确定了研究的前提是“学习或智能的任何其他特性的每一个方面都能被精确地加以描述，使得机器可以对其进行模拟。”^③

《人工智能哲学》一书中将其定义为“研究怎样制造计算机，并（或）为其编程，使其能做心灵所能做的那些事情”。^④关于 AI 与人类心灵和思维的关系有两个基本立场：“强 AI”和“弱 AI”。^⑤前者认为有可能制造出真正具有自我意识的 AI，只要运行适当的程序，计算机本身就是有思维的；后者则认为 AI 只是对人类智能的模仿，仅是用来研究人的思维的工具，不可能制造出真正具有自我意识的 AI。不难看出，图灵测试是弱 AI 的立场。

机器真的会思考吗？这取决于我们如何界定“机器”和“思考”。如果认为人也是一种机器，那么因为人会思考，因而当然可以说机器会思考，可是问题在于人不是机器。如何定义机器？在中文中，“机”指的是关键部件，即发动部件。“械”指的是一般器械或器具，即其它随动构件。“器”指的是一般的器件。19 世纪 60 年代以来，“机器”主要指来自西方的“machine”，“机械”用于描述西方的“machinery”和“machine”。20 世纪初，“机械”成为各种机械装置的泛称，是“机构（mechanism）”和“机器（machine）”的总称，其含义包括“机器”。^⑥英国机械学家威利斯（Robert Willis 1800-1875）在其《机构学原理》（The Principle of Mechanism, 1841 年）中所给的定义是：“任何机器（machine）都是由用各种不同方式连接起来的一组构件组成，使其一个构件运动，其余构件将发生一定的运动，这些构件与最初运动之构件的相对运动关系取决于它们之间连

① 吴清原著：《吴清源对局全集》，孔祥明译，成都：蜀蓉棋艺出版社，2010 年。

② A. M. Turing. “Computing Machinery and Intelligence” [J], Mind, Vol. 59, No. 236, pp. 433-460, 1950.

③ McCarthy, John; Minsky, Marvin; Rochester, Nathan; Shannon, Claude (1955). “A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence”. (<http://www-formal.stanford.edu/jmc/history/dartmouth/dartmouth.html>)

④ [英] 玛格丽特·博登编：《人工智能哲学》，刘西瑞等译，上海：上海译文出版社，2006 年，第 1 页。

⑤ Searle, John. R. (1980) Minds, brains, and programs. Behavioral and Brain Sciences 3 (3): 417-457.

⑥ 张柏春：《汉语术语“机器”与“机械”初探》，载自《机械技术史（2）——第二届中日机械技术史国际学术会议论文集》，2000 年。

接的性质。”^①中国机械学家和机械工程教育家刘仙洲给出了“机械”的中文定义：“机械者，两个以上之物体之组合体，动其一部分则其余各部各发生一定之相对运动或限制运动，吾人得利用之使一种天然能力或机械能力发生一定之效果工作者也。”^②总而言之，机械和机器都是由零部件构成的具有一定功能的装置，其运转在一定误差范围内可以用数学语言进行精确描述和控制。

传统意义上的机器可以代替人的体力劳动，计算机则用来代替人的脑力劳动——数学运算和符号操作。英国数学家查尔斯·巴贝奇早在1837年就发明了电子计算机的前身——分析机（Analytical Engine）。^③1936年，阿兰·图灵发明了一个抽象的机器“图灵机”，来模拟人们用纸笔进行数学运算的过程。图灵的伟大洞见是，这样的一台机器，尽管所能进行的操作都是局部且有限的，却可以模拟人类所能进行的任何计算过程。^④著名“丘奇——图灵论题”指出，所有可计算问题都可以由一台图灵机来计算。实际上，数学家们构造了许多可以表示“有效计算性”的形式系统，最终都证明其计算能力与图灵机等价。由于传统机械装置各构件的运动和控制作用都可以用数学加以清晰描述，即都是可计算的，因此图灵机可以模拟一切机器的行为。

图灵机或任意一个AI程序只能按其程序进行机械操作，无法与人脑所具备的心灵和意识相提并论。首先图灵机的计算能力是有限的，计算理论的研究表明，存在“不可计算的问题”。图灵在1936年的论文中就提出了第一个不可计算的问题，即图灵机“停机问题”：不存在这样一个程序（算法），它能够计算任何图灵机在给定输入上是否会停机（结

束计算）。^⑤实际上，存在无穷多“不可计算的”数学问题。因为每个图灵机都可以对应一个自然数，因而全部图灵机的集合可一一对应于自然数集合，而全部问题则对应于更高级别的无穷集合，问题永远比算法多得多。哥德尔不完全性定理也指出，任何足够强的形式系统都是不完备的，即在系统内存在一个为真但无法在系统内推导出的命题。这些为真、但又不可证明的命题，电脑不能生成或推导出来，人脑却能看出它是真的，因而人脑和心灵具有某种电脑（形式系统）无法实现的直觉和洞察力，所以电脑或人工只能不可能象人脑那样“思考”，不可能拥有象人脑那样的心灵和意识。^⑥

电脑和人脑的另一个分别是电脑只能遵照给定语法处理符号，无法理解符号的含义，人脑却可以理解语义。1980年，美国哲学家西尔勒（John Searle）提出了名为“中文屋子”的假想实验，来反驳强AI的观点。^⑦西尔勒假设一个人呆在一个封闭的房子里，有输入和输出缝隙与外部相通。输入的是中文问题，而他却对中文一窍不通。房子内有一本英语的指令手册（相当于计算机程序），从中可以找到相应的规则，产生中文答案。他可以把作为答案的中文符号写在纸上，并输出到屋子外面。这样，看起来他能处理输入的中文问题，并给出正确答案（如同一台计算机通过了图灵测试）。但是实际上，他对那些中文问题毫无感觉，根本不理解其中的任何一个词。

人工智能程序是“数据+模型+算法”，其本质是真而不实的抽象的数学对象，可以用不同的编程语言实现，也可以在不同的硬件平台上运行，是“理事两分”。人的心灵和思考过程则是具体而实

① 冯立升：《中国机械工程史研究的若干问题》，载自张柏春、李成智主编：《技术史研究十二讲》，北京：北京理工大学出版社，2006年，第89页。

② 刘仙洲：《中国机械工程发明史（第一编）》，北京：科学出版社，1962年，第5—6页。

③ Bromley, Allan G. The Evolution of Babbage's Calculating Engines. Annals of the History of Computing, Vol 9 (1987): 113-136.

④ Alan Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. Proceedings of the London Mathematical Society, Series 2, 42 (1936-7), pp 230-265.

⑤ Alan Turing. On Computable Numbers, with an Application to the Entscheidungsproblem. Proceedings of the London Mathematical Society, Series 2, 42 (1936-7), pp 230-265.

⑥ [英]罗杰·彭罗斯著：《皇帝新脑》，许明贤等译，长沙：湖南科技出版社，2007年，第122页。

⑦ Searle, John. R. (1980) Minds, brains, and programs. Behavioral and Brain Sciences 3 (3): 417-457.

在的，与大脑、身体乃至身心所处的文化和社会环境密不可分，是“理在事中”。借用冯友兰的新理学本体论思想的概念，一个是“真际”，一个是“实际”。冯友兰认为，哲学研究的对象是真际，科学研究的对象是实际。^①因此某种意义上，强AI的立场是“反科学的”，认为人的心灵和思考活动可以从实际的生理基础和物理环境中抽离出来，企图以抽象的虚理来取代具体的实在，犯了“以理限事”的错误^②。怀特海在《科学与现代世界》一书中将现代科学中此类把抽象的东西当成实在的东西的错误称为“错置具体性之谬误（Fallacy of Misplaced Concreteness）”。^③

三、仁者人也

在儒家思想中，关于人的本质的讨论最早是通过“人禽之辨”展开和深入的，最终确立了“仁”为人之所以为人的本质，而且“人”与“仁”可以互为定义。《礼记·中庸》：“仁者，人也。”《孟子·尽心篇》：“仁也者，人也；合而言之，道也。”朱子《孟子集注》释曰：“仁者，人之所以为人之理也。然仁，理也；人，物也。以仁之理，合于人之身而言之，乃所谓道者也。”在儒家思想中，作为人之本质的“仁”，主要具有以下涵义。

（一）仁者爱人

儒家思想的核心是“仁”，其首义为“爱人”，指人与人之间相亲相爱之情。许慎《说文解字》曰：“仁，亲也，从人从二。”《论语·颜渊》中樊迟问仁，子曰：“爱人。”《中庸》曰：“仁者，人也，亲亲为大。”《论语·学而》：“孝悌也者，其为仁之本与！”儒家思想把“孝”“悌”“慈”等人伦亲情作为践仁之根本和起点，以恕道推己及人，老吾老以及人之老，幼吾幼以及人之幼，以至于天下一家，“四海之内，皆兄弟也”，“凡天下疲癃、残疾、惇独、鰥寡，皆吾兄弟之颠连而无告者也”。人际交往中体现的仁又被提炼为一以贯之的忠恕之

道：尽己之心为忠，推己及人为恕，己欲立而立人，己欲达而达人，己所不欲，勿施于人。

基于人与人相亲相爱的真实情感和推己及人的忠恕之道，儒家构造了五种基本的人伦关系和行为准则，即五伦：君臣、父子、兄弟、夫妇、朋友。在五伦中尽职尽责，成为儒家学习做人和教育的主要内容。《孟子·滕文公上》：“使契为司徒，教以人伦：父子有亲，君臣有义，夫妇有别，长幼有序，朋友有信。”依此义，人是能在具体的人伦关系中恰当地表达仁爱的人。

（二）反身克己

仁的表现，对外是爱人，对内则是克己，要求能反躬自省，克制和约束自己，凡事合乎“礼”的要求。《论语·颜渊》：“颜渊问仁。子曰：‘克己复礼为仁。一日克己复礼，天下归仁焉！为仁由己，而由人乎哉？’颜渊曰：请问其目。子曰：非礼勿视，非礼勿听，非礼勿言，非礼勿动。”曾子也说过：“吾日三省吾身。”能反身克己反映出人是身心合一的整体。1993年在湖北荆门出土的郭店楚简的67个“仁”字，都没有写作从人从二的“仁”，而是写作上身下心的“𠂔”字。^④这一字形也印证了仁字身心一体的含义。《大学》和《中庸》里强调的“慎独”，就是一种反身克己的自我修养功夫。依此义，人是能够反躬自省的身心合一的有机整体。

（三）良知本心

儒家思想认为，道不远人，仁乃天命之性，就存在我们的心中。《论语·述而》：“子曰：‘仁远乎哉？我欲仁，斯仁至矣。’”孟子曰：“仁，人心也；义，人路也。”儒家主张性善论，认为“人之初，性本善”，这里说的不是人性就是现成的良善，而是人具有与生俱来的潜在的善根和为善的可能，即根植于心的良知良能。《孟子·尽心上》：“人之所不学而能者，其良能也；所不虑而知也，良知也。”《诗·大雅·烝民》：“天生烝民，有物有则，民

① 冯友兰：《三松堂全集第四卷·新理学》，郑州：河南人民出版社，2001年，第19页。

② 语出王夫之《续春秋左氏传博议下·土文伯论日食》：“有即事以穷理，无立理以限事。故所恶于异端者，非恶其无能为理也，罔然仅有得于理，因立之以概天下也。”

③ Alfred North Whitehead, Science and the Modern World[M], pp52. THE MACMILLAN COMPANY, 1925.

④ 余兰兰：《郭店楚简中的“仁”字研究》，《湖北大学学报》（哲学社会科学版）2012年第1期。

之秉彝，好是懿德。”《大学》所讲的“明德”，《中庸》讲“天命之为性”，孟子讲“人皆有不忍人之心”，“恻隐之心，人皆有之；羞恶之心，人皆有之；恭敬之心，人皆有之；是非之心，人皆有之”，“人皆可以为尧舜”，荀子讲“涂之人可以为禹”，都是在揭示这个先天的良知本心。孟子认为这是人之所以区别于禽兽的本质。《孟子·离娄下》：“人之所以异于禽兽者几希；庶民去之，君子存之。舜明于庶物，察于人伦，由仁义行，非行仁义也。”王阳明更把“致良知”作为其心学思想的核心，将吾心之良知视为是天理。《传习录中·答顾东桥书》：“若鄙人所谓致知格物者，致吾心之良知於事事物物也。吾心之良知，即所谓天理也。致吾心良知之天理於事事物物，则事事物物皆得其理矣。”在《咏良知四首示诸生》中，王阳明写道：“个个人心有仲尼，自将闻见苦遮迷。而今指与真头面，只是良知更莫疑。”依此义，人是有良知或先天道德意识的人。

（四）万物一体

先秦儒学多侧重于切身的孝悌忠信之教和修齐治平之道，对于超越具体生活经验的宇宙人生的普遍道理较少深入探讨。《论语·公冶长》中子贡发出感叹：“夫子之文章，可得而闻也；夫子之言性与天道，不可得而闻也。”到了宋代，理学家们为了回应佛道思想的挑战，开始阐发一种贯通宇宙自然和人生命运的新儒学。

宋儒程颢着重强调仁的“感通”之义，以“与物同体”论仁。一个人如果不把宇宙万物看作是与自己息息相关的一个整体，而以为与自己相互隔绝不相干，则是一种麻木不仁的病态表现。《河南程氏遗书》卷四中程颢说：“若夫至仁，则天地为一身，而天地之间，品物万形为四肢百体。夫人岂有视四肢百体而不爱者哉？圣人，仁之至也，独能体是心而已，曷尝支离多端而求之自外乎？故‘能近取譬’者，仲尼所以示子贡以为仁之方也。医书有以手足风顽谓之四体不仁，为其疾痛不以累其心故也。夫手足在我，而疾痛不与知焉，非不仁而何？世之忍心无恩者，其自弃亦若是而已。”张载在《西铭》中提出了人与天地同体同性、“民胞物与”的思想：“乾称父，坤称母；予兹藐焉，乃混然中处。故天地之塞，吾其体；天地之帅，吾其性。民，吾同胞；物，

吾与也。”王阳明在《大学问》中也说：“大人者，以天地万物为一体者也。”依此义，人是有同理心、能与他人及物相感通之人。

（五）生生之德

在儒家思想中，仁不仅是内在于人的心性本体，还是生生不息的宇宙本体。《周易·系辞》曰：“日新之谓盛德，生生之谓易。”又曰：“天地之大德曰生。”还描述上天之道为：“一阴一阳之谓道……显诸仁，藏诸用。”这是说天道在万物的创生化育中，在仁中显现，天的本质就是创造性本身，也就是仁。在现代汉语中，仍保留着把植物的果核称为“仁”的用法，如花生仁、核桃仁，是生长的根源。

宋儒也认为“生生”也是仁。二程说，“万物之生意最可观，此元者善之长也，斯所谓仁也”（《二程集》）。朱子《语类》载：“问‘仁者天地生物之心’。曰：‘天地之心，只是个生。凡物皆是生，方有此物。如草木之萌芽，枝叶条干，皆是生方有之。人物所以生生不穷者，以其生也。才不生，便干枯杀了。’”宋儒又讲，人之心，得之于天，即是天地之心。吾人的本体，也是宇宙万物的本体。南宋王应麟说：“人者，天地之心也。仁，人心也。人而不仁，则天地之心不立矣。为天地立心，仁也。”为天地立心，即是参赞天地之化育。清代戴震认为：“仁者，生生之德也”，“气化流行，生生不息”是仁（《孟子字义疏证》）。以此而论，仁便是宇宙大化洪流生生不息的动力之源，是天地之心，是宇宙的本体。

人是万物之灵，天道下贯成为人的性命，人若能尽心尽性，便可以与天地参。《中庸》：“唯天下至诚，为能尽其性；能尽其性，则能尽之性；能尽人之性，则能尽物之性；能尽物之性，则可以赞天地之化育；可以赞天地之化育，则可以与天地参矣。”依此义，人是通过尽心尽性践仁来体现天道生生不息之创造性的人。

综上所述，在儒家思想中，人之为人的本质是仁，其内容包含亲爱人的真实情感、反身克己的修养功夫、先天内在的良知本心、万物一体的感通性以及生生不息的创造性。

四、徒量智不足以为人

从上面的讨论可知，人工智能的本质是抽象的数与机械的算法过程，人的本质是具体而真实的仁，

能爱人、会反省、可感通、有良知、知是非，是一个生生不已的有机生命，因此两者不能同提并论。从中西文化比较的角度来看，“仁”与“智”代表了中西文化的突出特征。自轴心时代开始，中西方文化就开始沿着不同的方向探索各自的轴心突破。中国开辟了一条以“仁”为指引的人文主义道路，西方则是一条以“智”为指引的理性主义道路。人工智能是现代科技文明发展的最新成果，是西方逻辑理性之理智主义的产物，可以一直追溯到古希腊毕达哥拉斯学派“万物皆数”的理念。

就人之身心性命而言，仁是其本体，智是本体一个方面的发用。宇宙万物一体而又有别，仁的本质可以通天地万物为一体，而智的运用则以与物有别、相对为二为前提。仁之发用为智，形成一知性主体，可以观察测量外物，并运用其逻辑数学的理性进行综合，揣测自然的量化规律和法则，进而形成知识的体系，开辟出一科学之途径，再以所产生的知识利用加工自然界之物质、能量和资源以满足人欲之需，从而造就了一个各种机器层出不穷的现代物质文明，其成就不可谓不大。然而它的弊病在于，若知性主体一直向外逐物而不知返，积习愈久愈深，则心陷于物而僵化、物化、机械化，使得德性主体不能彰显，人将堕落而不成为人。

关于东西方文化“仁”与“智”的分别，熊十力先生在《新唯识论》里面有一个“性智”和“量智”的说法。他说：“性智者即是自性的明解，此中自性即曰本体，在宇宙论上通万有而言其本原则曰本体，即此本体以为吾人所以生之理而言则亦曰自性……《易》所谓始万物之知即此性智是也。”关于“量智”，他说：“量智是思维和推度与简择等作用，能明辨事务之理则及于所行所历简择得失，故名为量智，亦名理智。此智原是性智的发用而卒别于性智者，因为性智作用依官能而发现，即官能得假之以自用……逐取乎物、分别乎物、运用乎物，由此积习日深日精而成功后起之一种明锐势用，所谓量智是已。”^①

儒家思想要求人彰显性智或德性主体，发明良

知，做一个真正的人。儒家自古以来一向严守人禽之辨，就是因为看到了人心有陷溺和堕落的危险。孔子以“孝”别人禽。《论语·为政》：“子游问孝。子曰：‘今之孝者，是谓能养。至于犬马，皆能有养；不敬，何以别乎？’”孟子以“心”别人禽。《孟子·离娄下》：“人之所以异于禽兽者几希；庶民去之，君子存之。”《孟子·告子上》：“由是观之，无恻隐之心，非人也；无羞恶之心，非人也；无辞让之心，非人也；无是非之心，非人也。”荀子以“义气”、“能群”和“能辩”别人禽。《荀子·王制》：“水火有气而无生，草木有生而无知，禽兽有知而无义；人有气，有生，有知，亦且有义气，故最为天下贵也。力不若牛，走不若马，而牛马为用，何也？曰：人能群，彼不能群也。”《荀子·非相》：“人之所以为人者何已也？曰：以其有辨也。饥而欲食，寒而欲暖，劳而欲息，好利而恶害，是人之所生而有也，是无待而然者也，是禹桀之所同也。然则人之所以为人者，非特以二足而无毛也，以其有辨也。今夫狴狴形状亦二足而无毛也，然而君子啜其羹，食其馘。故人之所以为人者，非特以其二足而无毛也，以其有辨也。夫禽兽有父子，而无父子之亲，有牝牡而无男女之别。故人道莫不有辨。”《礼记》以“礼”别人禽。《礼记·曲礼上》：“鹦鹉能言，不离飞鸟。猩猩能言，不离禽兽。今人而无礼，虽能言，不亦禽兽之心乎？夫唯禽兽无礼，故父子聚麀。是故圣人作，为礼以教人，使人以有礼，知自别于禽兽。”

AI则是量智运用的产物，其目的却是想要在人之外再造一个人，殊不知徒量智不能为人，性智如果不能彰显，人或将堕落为禽兽乃至麻木不仁之物也未可知。从存在的状态来比较，AI是仁体发用转为量智运作而产生一个抽象的数理对象和机械的算法过程，属于“理”的范畴，是纯粹的潜在，包括人和禽兽在内的生命体则是鲜活而具体的实在。比较AI、禽兽和人，AI程序为仁之偏倚所出和抽象投射，有理，但无气、无生、无知、无情、无义；禽兽得仁之小体，有理、有气、有生，在较低的级别上有知、有情，但是无义、无礼、无心；只有人得仁之全体，

① 熊十力：《新唯识论》（壬辰删定本），北京：中国人民大学出版社，2006年，第23—24页。

有理、有气、有生、有知、有义、有礼，而且最重要的是有心。

对 AI 的追求和推崇是人类倚仗量智妄夺天工的企图，是以部分僭越整体的偶像崇拜。在此过程中，心沉溺于物，迷于途而不知返，德性主体亦无暇彰显，从而可能形成人的物化、机械化、算法化，造成主体道德意识之沦丧。关于人的物化的严重后果，《礼记·乐记》曾警告过：“人生而静，天之性也。感于物而动，性之欲也。物至知知，然后好恶形焉。好恶无节于内，知诱于外，不能反躬，天理灭矣。夫物之感人无穷，而人之好恶无节，则是物至而人化物也。人化物也者，灭天理而穷人欲者也。于是有悖逆诈伪之心，有淫佚作乱之事，是故强者胁弱，众者暴寡，知者诈愚，勇者苦怯，疾病不养，老幼孤独不得其所。此大乱之道也。”20 世纪发生的两次世界大战和冷战等都印证了这一点。

五、正德利用厚生惟和

虽然对 AI 不必有无谓的恐慌，不用担心一个超智能 AI 有一天会统治和奴役人类，但对 AI 带来的现实问题还是要引起高度的重视。首先，AI 很可能会带来一定程度的失业和贫富分化等社会问题。一些能够被 AI 取代的工作岗位将会消失，如同汽车的发明导致马车夫失业一样。掌握了 AI 技术的人群可能因此拉大与未能掌握 AI 技术的人之间的贫富差距，如果数字鸿沟带来的后果一样。其次，要重视 AI 应用中的安全问题和有关法律问题。AI 在运行中可能脱离原先设计的目标而对人带来安全隐患，例如无人驾驶汽车。对于 AI 系统的责任认定也需要及早开始研究和立法规范引领。第三，要特别警惕 AI 可能被恶意应用。例如，AI 被设计执行反人道的毁灭性任务如 AI 武器系统。跟核武器和生化武器一样，AI 武器系统如果落入反人类的恐怖分子手中，后果不堪设想。未来也可能爆发高度智能化的 AI 战争，面对 AI 参与的军事攻击，如何有效地防御和反击是每一个国家都要面对的现实军事问题。

然而，真正的危险并非来自 AI，而是来自人自

身，即工具理性毫无约束的发展导致人不能为人的危险，这已经成为全球各文明都正在面对的普遍问题。面对良知沉溺和物化的危险，儒家思想提供的解决途径是向内反躬自省，体认良知本心，以彰显性智，挺立道德主体，以之驾驭知性主体和工具理性。人类在发展工具理性的同时，道德理性和人文价值必须同时得到均衡的发展。《尚书·大禹谟》提出：“正德、利用、厚生、惟和。”包括 AI 在内的科技都是利用厚生之事，其发展无碍，但要以正德为首，以和谐为目的，即首先要端正生命，最终要达到天地人已普遍的和谐。

雅斯贝尔斯在《历史的起源与目标》一书中提出了“轴心时期”这一著名概念，指出在公元前 8 世纪至公元前 2 世纪，中国、印度、希腊和希伯来文化都不约而同出现了一次巨大变迁，即“轴心突破”。^①这些轴心文化沿着不同的方向探索和发展了人类的潜力：中国是道德践履，印度是沉思冥想，希腊是理性思辨，希伯来是宗教信仰。在过去几百年间，西欧经由文艺复兴、地理大发现、宗教改革、科技革命、工业革命，逐步发展出一个工具理性高度发达的、世俗的、充斥各种机器的现代科技物质文明。电子计算机、互联网和 AI 是这一文明发展的最高级形式。现代文明的弊病在过去百年全面爆发，在全球范围内人与人、人与自然以及人与自我三大关系全面恶化。

雅斯贝尔斯在《历史的起源与目标》一书中还说过：“直至今日，人类一直靠轴心期所产生、思考和创造的一切而生存。每一次新的飞跃都回顾这一时期，并被它重燃火焰。自那以后，情况就是这样。轴心期潜力的苏醒和对轴心期潜力的回忆，或曰复兴，总是提供了精神动力。”^②自 20 世纪后期以来，随着全球一体化的发展，以“全球意识”、“生态意识”和“对话意识”为标志的“第二轴心时代”呼之欲出。人们期望，第二轴心时代的展开能够超越现代科技文明的弊病，达成人与人、人与自然以及人与自我三大关系的全面和谐。而要达成这一伟

① [德] 卡尔·雅斯贝尔斯：《历史的起源与目的》，魏楚雄、俞新天译，北京：华夏出版社，1989 年，第 7—8 页。

② [德] 卡尔·雅斯贝尔斯：《历史的起源与目的》，魏楚雄、俞新天译，北京：华夏出版社，1989 年，第 14 页。

大目标，各文明一方面要批判现代科技文明的弊端，同时也保留其优点，并且从各自前现代的传统思想资源中汲取智慧，包括儒家的“仁爱”、佛教的“慈悲”、基督教的“博爱”和伊斯兰教的“兄弟之爱”等，以建设一个和谐的后现代生态文明。

对于中国，21世纪要实现中华民族的伟大复兴，离不开中华文化的伟大复兴，即要返本开新，回到中华文化生命仁智合一、以仁摄智的状态，实现传统文化的创造性转化；同时要扩体充用，首先挺立文化主体，然后大胆吸取西方文化的优点和长处，为我所用，实现传统文化的创新性发展。在这样一个既融合前现代文明的智慧，又有现代文明发达科技的后现代社会形态中，儒家的内圣外王之道怎么行？熊十力先生曾提出一个观点：“思修交尽，上达天德。”他说：“专尚思辨者，可以睿理智，而以缺乏修为故，则理智终离其本，无可语上达。专重修为者，可以养性智，而以不务思辨故，则性智将遗其用，无可成全德。是故思修交尽，二智圆融，而后为至人之学。”^①现代西方一味追求“量智”的物质文明缺乏修为功夫，故没有上达之路，可以用中国的“性智”之学补之；但如果只讲“性智”，不务思辨，则“性智”将不能发挥其全体大用。因此最理想的状态是中西结合，性智和量智齐头并举，思修交尽，二智圆融。只有这样，才可能导出一个安全、良善的AI技术以及人与AI各尽其性的未来新世界。

参考文献：

- [1] 吴清源著. 吴清源对局全集[M]. 孔祥明, 译. 成都: 蜀蓉棋艺出版社, 2010.
- [2] A. M. Turing. Computing Machinery and Intelligence[J]. Mind, Vol. 59, No. 236, pp. 433-460, 1950.
- [3] [英] 玛格丽特·博登编. 人工智能哲学[M]. 刘西瑞等, 译. 上海: 上海译文出版社, 2006.
- [4] Searle, John. R. Minds, brains, and programs[J]. Behavioral and Brain Sciences, 3 (3): 417-457, 1980.

[5] 张柏春. 汉语术语“机器”与“机械”初探[C]//机械技术史(2)——第二届中日机械技术史国际学术会议论文集, 2000.

[6] 冯立升. 中国机械工程史研究的若干问题[M]//张柏春, 李成智主编. 技术史研究十二讲. 北京: 北京理工大学出版社, 2006.

[7] 刘仙洲. 中国机械工程发明史(第一编)[M]. 北京: 科学出版社, 1962.

[8] Bromley, Allan G. The Evolution of Babbage's Calculating Engines[J]. Annals of the History of Computing, Vol 9:113-136, 1987.

[9] Alan Turing. On Computable Numbers, with an Application to the Entscheidungsproblem[J]. Proceedings of the London Mathematical Society, Series 2, 42, pp 230-265, 1936.

[10] [英] 罗杰·彭罗斯著. 皇帝新脑[M]. 许明贤等译. 长沙: 湖南科技出版社, 2007.

[11] 冯友兰. 三松堂全集第四卷·新理学[M]. 郑州: 河南人民出版社, 2001.

[12] Alfred North Whitehead, Science and the Modern World[M]. The Macmillan Company, 1925, pp52.

[13] 余兰兰. 郭店楚简中的“仁”字研究[J]. 湖北大学学报(哲学社会科学版), 2012, 39(1).

[14] [德] 卡尔·雅斯贝斯. 历史的起源与目的[M]. 魏楚雄, 俞新天译. 北京: 华夏出版社, 1989.

[15] 熊十力. 新唯识论(壬辰删定本)[M]. 北京: 中国人民大学出版社, 2006.

作者: 郭海鹏, 北京师范大学—香港浸会大学联合国际学院(UIC)教授、全人教育办公室主任

责任编辑: 周修琦

^① 熊十力:《新唯识论》(壬辰删定本), 北京: 中国人民大学出版社, 2006年, 第28页。