

认知相对论——通向强人工智能之路

李玉鑑*

(桂林电子科技大学人工智能学院 桂林 541004)

摘要: 人工智能(AI)的发展如火如荼,大有超越人类之势,以致很多人认为奇点就要来临,强人工智能即将实现。这是一种对强人工智能的误解,因为强人工智能的核心并不在于其功能是否强大,而在于它是否具有意识。该文首先解释了强人工智能的内涵,讨论了与之相关的意识问题;然后,阐述了旨在揭开意识奥秘的认知相对论思想,包括:世界的相对性原理和符号的相对性原理,以及世界、语言和心灵的关系。接着,提出了另一条新原理,即意识的等效原理,用以说明意识从物质产生所需要的物理条件,解决主观体验或现象意识的困难问题,推导意识能力受限于感觉能力且以感觉容量为上界的认知基本定理,并分析意识在哪里和自我是什么的可能性。最后,在认知相对论的框架下,给出了研究意识问题的新纲领和实现机器意识的新思路,并展望了强人工智能的未来。

关键词: 人工智能; 强人工智能; 意识; 机器意识; 认知相对论

中图分类号: TN919; TP391; B017

文献标识码: A

文章编号: 1009-5896(2024)02-0408-20

DOI: 10.11999/JEIT230749

Theory of Cognitive Relativity — The Road to Strong Artificial Intelligence

LI Yujian

(School of Artificial Intelligence, Guilin University of Electronic Technology, Guilin 541004, China)

Abstract: Artificial Intelligence(AI) develops in full swing with a great potential to surpass human, leading many people to believe that a singularity is imminent and that strong AI is about to be realized. This is a misconception of strong AI, because the core of strong AI is not whether it is powerful, but whether it has consciousness. In this article, firstly, the connotation of strong AI is explained, and the related problem of consciousness is discussed; Then, the ideas of Theory of Cognitive Relativity is elucidated, aimed at revealing the secret of consciousness, including the Principle of World's Relativity, the Principle of Symbol's Relativity, together with the relationships between world, language and mind. Subsequently, another new principle is expounded, namely the Principle of Consciousness' Equivalence, to show the physical conditions required for arising of consciousness from matter, and to solve the hard problem of subjective experience or phenomenal consciousness, and to establish the fundamental theorem of cognition that conscious ability is limited by sensory ability with the upper bound of sensory capacity, and to analyze the possibility of where consciousness is as well as what self is. Finally, under the framework of the theory of cognitive relativity, a new creed for solving the puzzle of consciousness and a new guide for implementing machine consciousness are presented, with the future of strong AI envisioned.

Key words: Artificial Intelligence(AI); Strong AI; Consciousness; Machine consciousness; Theory of cognitive relativity

收稿日期: 2023-07-21; 改回日期: 2023-12-05; 网络出版: 2023-12-13

*通信作者: 李玉鑑 liyujian@guet.edu.cn

基金项目: 国家自然科学基金(61876010)

Foundation Item: The National Natural Science Foundation of China (61876010)

1 引言

1956年，人工智能(Artificial Intelligence, AI)在达特茅斯会议上诞生，开启了用机器模拟人类智能的征程。人工智能几经大起大落，如今又重返辉煌，在世界上掀起了惊涛骇浪。随着深度学习的发展^[1-4]，人工智能取得了巨大成功，创造了一系列标志性成果，例如AlphaGo^[5]，MuZero^[6]，AlphaFold^[7]。特别是最近火遍全球的聊天生成预训练转换器(Chat Generative Pre-trained Transformer, ChatGPT)，已经具备了对话、问答、解题、写作、翻译和科研等几乎全方位的综合智能，大有超越人类之势。文献[8]曾经预言人工智能的奇点即将来临，而文献[9]似乎相信人工智能很快就会拥有心灵。媒体给人的感觉就是，强人工智能仿佛已经近在眼前。然而，现实却并非如此。

事实上，人们不仅对人工智能的奇点盲目乐观，对强人工智能的内涵也存在误解。奇点原先是指物理宇宙中的一类特殊时空点，比如黑洞，它具有极小的体积、极高的密度和极强的引力，不再适用普通物理定律。库兹韦尔把这一概念用到了人工智能的发展上，指机器将与人类不断融合，并在某个神妙时刻的奇点达到指数爆炸的智能程度。从此，机器与人类的区别将变得模糊，达到无缝融合的程度，我们可以任意装扮成不同的身体，扮演一系列不同的角色，这未免有点太过科幻。姑且不论机器与人类的无缝融合是否可能，即使可能，融合产生的新物种到底是机器还是人类，也是一个难以名状的麻烦。更有甚者，很多人以为奇点之后的人工智能无比强大，就是所谓的强人工智能，这又是一种对强人工智能的误解。从概念上讲，强人工智能的关键并不在于功能是否强大，而在于是否拥有意识。按照文献[10]的观点，现在所有的人工智能都不具备意识，本文也秉持这种态度。

本文首先解释强人工智能的内涵，讨论与之相关的意识问题。然后，阐述认知相对论的思想，包括世界的相对性原理和符号的相对性原理，还有世界、语言和心灵的关系，以及另一条新原理——意识的等效原理，用以说明意识从物质产生所需要的物理条件，解决主观体验或现象意识的困难问题，推导意识能力受限于感觉能力且以感觉容量为上界的认知基本定理，并分析意识在哪里和自我是什么的可能性。接着，在认知相对论的框架下，提出研究意识问题的新纲领和实现机器意识的新思路。最后，对强人工智能的未来进行了展望。

2 强人工智能

早期人工智能的主要范式是符号主义，其核心

在于物理符号系统假设，即：足够复杂的物理符号系统是产生通用智能的充分必要条件。根据这个假设，通用智能可以借助足够复杂的物理符号系统来产生，足够复杂的物理符号系统也可以展现出通用智能。其中，物理符号系统不限于人类创造并使用的符号系统，还可以由其他物质材料来构造，但必须遵从物理学定律。物理符号系统是由符号的结构及其操作组成的，可以通过在计算机上编写的程序来实现，它所具有的智能将在程序运行时得到展现。在这个意义上，人类的通用智能也可以通过程序的运行来展现。换句话说，我们只需在计算机上编写适当的程序并运行之，就可以产生人类的所有心智能力，包括：感知、识别、学习、记忆、理解、思维、意识和心灵等。

为了反驳符号主义，Searle^[11]提出了强人工智能的概念，以便跟弱人工智能作出区分。强人工智能的最初内涵，是指这种哲学立场，即：认为计算机经过适当编程就能产生真正的心智，而不仅仅是模拟。而弱人工智能主张相反的立场：计算机永远只能是一种研究心智的工具，不管多么强大，也无论如何编程，都不可能拥有真正的心智。究竟哪一种立场正确呢？

Searle认为计算机没有意识和理解，充其量只能提供一种对人类智能的模拟，而不可能是一种复制，因此最多是一个心灵的空壳。他提出著名的“中文屋”思想实验，来批判强人工智能的观点。这个思想实验假定有一台计算机利用程序模拟了人类理解故事的能力，如果你用中文输入一个关于故事的问题，它可以像你一样用中文给出回答。例如，给计算机输入故事：“一个人走进一家餐馆，要了一份汉堡包。汉堡包送上来时，已被烤焦，这个人生气地冲出餐馆，没有付账，也没有留下小费。”现在问：“这个人吃汉堡包了吗？”计算机会像你一样根据推测回答：“不，他没有吃。”当然，你还可以找很多其他的故事来问各种问题，但计算机总能输出跟你一样(或者类似)的回答。由于在第三者看来，计算机与人在理解故事方面的表现几乎相同的，因此它通过了关于“理解”的图灵测试，我们可以认为它具备了理解的能力。

强人工智能的支持者宣称，计算机确实通过程序理解了故事。但Searle认为这纯属想当然，因为他自己显然不懂中文，完全不能理解中文故事，不过他同样也可以按照程序的指令来处理中文。换言之，如果讲英文母语的Searle本人被关在一个只能用中文跟外界交互的房间(即中文屋)里面，那么由于他可以像计算机一样按照英文书写的程序指令来

回答外界输入的任何中文问题,所以在屋外的人看来, Searle对问题的回答与讲中文母语的人毫无区别。凡是看过他用中文回答的人,谁也不会知道他连一个中文字都不认识,更不用说理解中文故事了。既然Searle按照程序指令回答中文问题的能力并不是因为真正理解了中文,计算机肯定也不是因为真正的理解。因此,看起来的“理解”并不等于真正的“理解”。另一方面,人类的理解是需要意识参与的思维和心灵活动,计算机同样不可能通过程序就真正拥有我们的意识、思维和心灵。

Searle的中文屋说明,计算机不可能像人那样理解中文。不过,在中国人看来,中文屋对强人工智能的反驳也许力度还不够充分。如果这样,我们可以再构想一个“阿拉伯文屋”,里面的人只能按照程序指令用阿拉伯文跟外界交互,得出的结论是类似的:计算机不可能像人类那样通过意识、思维和心灵来理解阿拉伯文。因此,意识、思维和心灵在本质上不同于计算机程序,是不可能通过编程来实现的。

现在,强人工智能的内涵已经有所变化,不再仅仅指Searle所说的哲学立场,常常还泛指真正具有意识、思维和心灵的人工智能。很多人以为,强人工智能是指功能强大的人工智能,至少跟人类一样强大,其实这是一种误解。在一般意义上,强人工智能与弱人工智能的根本区别在于是否有意识。迄今为止,所有的人工智能都属于弱人工智能,因为没有意识。然而,弱人工智能并不代表功能就一定弱,它也可能是功能非常强大的专用人工智能或者通用人工智能。比如, AlphaGo就属于弱人工智能,但其下围棋的水平已经超过了人类的世界冠军。又比如, ChatGPT也属于弱人工智能,但它似乎具备了无所不知的通用性。

人们经常混淆强人工智能和通用人工智能的概念,以为强人工智能包括通用人工智能、人类水平智能和超级人工智能。事实上,强人工智能未必代表功能的强大,它可能只具有低等动物(如蜜蜂)的智能,但必须要有意识。而通用人工智能是指足以表现出人类的任何智能行为的人工智能,但未必具有意识。由于强人工智能有意识,就可能按照自己的选择行事,而不完全受制于人,甚至违背人类的命令。因此,强人工智能才是真正的智能。通用人工智能可能只是看起来好像具有人类的智能,而实际上只是一具没有意识的哲学僵尸。专用人工智能是指某些专门领域或狭窄范围的人工智能,有时也称为窄人工智能,它跟通用人工智能一样,可以没有意识。此外,超级人工智能也未必是强人工智

能,但必须在某些方面具有超级强大的智能,而有没有意识,则无关紧要。因此,超级人工智能可以是强人工智能,或者弱人工智能,同时还可以是通用人工智能,或者专用人工智能。例如, AlphaGo及其升级版是超级人工智能,也是弱人工智能和专用人工智能。虽然ChatGPT已经具备了良好的通用性,但可能还算不上超级人工智能,而且因没有意识仍然只是弱人工智能。在理论上,通用人工智能、专用人工智能和超级人工智能,既可以没有意识,也可以拥有意识。如果没有,它们就是弱人工智能,否则就属于强人工智能。

人工智能有时也被划分为计算智能、感知智能和认知智能3个层次。计算智能是指用算法和程序在计算机上的运行所实现的人工智能。根据Searle的中文屋实验,计算智能不可能有意识,属于弱人工智能。感知智能是指具备视觉、听觉、触觉等感知能力的人工智能,认知智能是指具备语言、判断、推理和思维等能力的人工智能。严格地说,感知智能和认知智能都是需要意识参与的,目前机器不可能真正拥有。迄今,所有的感知智能和认知智能,实际上都是通过计算机编程实现的,本质上仍然属于计算智能,因此是弱人工智能,而不是强人工智能。只有机器在有意识的前提下所具备的感知智能和认知智能,才是强人工智能。

人工智能还有一种划分是:自动智能和自主智能。自动智能指机器按照人类设计好的模型和算法自动完成任务,比如自动巡检。由于自动智能可能看起来具有自主性,有时被误以为是自主智能。自主智能要求机器能够主动优化和选择自己的运行方式,而不需要人类去进行引导。虽然自主有完全依赖自己的意思,但人们常常还是把它与自动混淆了。自主还有主动的意思,而主动跟自动是不同的。严格地说,自主智能必须有意识参与,才能实施主动的选择和操作。无人驾驶看起来似乎具有主动性,但实际上可能并不自主。即便无人驾驶的自动化程度较高,人类干预较少,如果没有意识的参与,本质上仍然是自动智能,最多也只是自治智能。自动智能和自治智能都没有意识,属于弱人工智能;而自主智能拥有意识,属于强人工智能。

3 意识问题

如果说强人工智能的关键在于实现机器意识,那么机器能有意识吗?我们的意识又是什么呢?这就是所谓的意识问题^[12-21],它困扰东西方长达千年之久,被称为世界之结,也被誉为上帝最后的秘密。

自古以来,意识就是哲学家争论不休的问题^[22],人们有时也把它视为灵魂。苏格拉底有一句名言:

“认识你自己”。希波克拉底猜测我们的认识能力来自大脑^[23]，而柏拉图认为这是因为我们拥有非物质的灵魂，即一种脱离物质和个别而独立存在的一般抽象概念和客观精神实体^[24]。阿奎那又说灵魂是神创的、不朽的，可以与身体相分离^[25]。笛卡尔在某种程度上把意识等同于“思维”，他提出了著名的命题“我思故我在”，认为心灵和身体是两种完全不同的实体^[26]。这种心身二元论的观点影响了一代又一代的思想家，至今还有人奉若神明。例如，诺贝尔奖获得者埃克斯就相信心灵可以独立于身体而存在，还提出了关于“树突子(dendron)”和“心灵子(psychon)”的理论假设。埃克斯曾在梵蒂冈的教皇科学院负责组织过一次关于“脑和意识体验”的讨论，教皇想要确认：大脑属于科学，心灵属于神学^[27]。如果心灵可以独立于身体的物质而存在，那么意识就不再是一个需要解释的问题，因为它就是世界的本原，是一开始就有的东西。唯心主义认为意识比物质更基本、更在先，它不仅决定了世界上的一切物质，也决定了我们所看到的大千世界。泛心论主张“心灵”无处不在^[28]，自然界的所有事物都以某种方式拥有一定程度的心灵，不管是阿米巴虫，还是其身上的分子、原子、电子、甚至夸克，都具有心灵的成分。尽管曾一度被认为荒谬绝伦，但由于强人工智能迫切需要意识理论的发展、突破和指导，泛心论近年来又蓦然抬头，再次成为关注和争论的焦点。

文献^[29]认为“意识的科学只能根据可以复制和系统变化的标准条件建立在客观的允许重复的基础之上”，并把心理学界定为“研究意识并探索控制心灵的独特规律”。文献^[30]主张从意识的目的性和选择性来理解人类的心灵，并认为心理学是关于心理生活的现象及其条件的科学，应该研究意识的功能。而文献^[31]认为科学无法解释意识。麦金^[32]断言大脑的构造对人类理解能力是一个限制，肯定有某些问题解决不了，比如意识，就像老鼠无法求解微分方程一样。文献^[33]也觉得意识是个匪夷所思的东西，它没有质量和体积，其性质完全不同于物理过程。文献^[34]把灵魂视为机器中的幽灵，既不愿意承认有心灵这样的东西又想说它好像是有，并试图通过语言的分析 and 身体的行为来解释心灵的性质，从而将其转换成物理的东西。文献^[35]则坚决反对研究“像鬼火一样”的意识，甚至拒绝承认意识的存在，并宣称所有心理活动都可以用“刺激-反应”的行为公式来加以解释，虽然这推动了行为主义心理学的发展，但严重阻碍了意识的研究。乔姆斯基指出，仅仅通过“刺激-反应”的公式，

无论如何都解释不了人类复杂多变的自然语言现象^[36]，因此在刺激和反应之间必然存在“意识”。这导致了行为主义的最后一根稻草被压垮，认知心理学开始兴起^[37]，认知科学相继诞生^[38]。由此可见，认知是一种以意识为基础的心理活动，达马西奥^[39]写道：“没有意识，就没有认知”。

1974年，Nagel^[40]在《成为一只蝙蝠会是什么样子》的文章中，讨论了诸如“怎么才能解释看到红色的感觉”一类的问题。他指出意识的核心特征是主观性，这是一种跟特定物种有关的感受性质，又称为“现象意识”，或者“感受质”。意识就是拥有这种感受质的心灵状态，它只能被拥有它的有机体自己知道和理解，他人最多只能推测。Nagel认为，由于科学是客观的，而意识是主观的，所以科学根本就无法洞悉主观经验的王国，不论我们掌握多少关于蝙蝠的生理学知识，都无法真正了解是什么使得蝙蝠成其为蝙蝠。因此，科学和意识之间存在不可逾越的鸿沟，两者不可能有相交，即便真的相交了，那一定是因为发现了某种新的基本自然法则。而Dennett^[41]认为，意识的主观体验只不过是一种看上去十分逼真的幻觉，它根本就不是一个真正的问题，应该彻底否定它：我们之所以对意识深信不疑，就是因为我们能够体验到它。弗拉纳根也认为主观的意识和客观的科学之间不存在鸿沟：为什么你能主观地、区别于他人地体验到特定脑活动，这是因为只有你和你的神经系统正常相连，从而产生只属于你自己的体验，背后并没有神秘玄学^[27]。

1976年，大名鼎鼎的克里克开始从神经科学的还原主义立场研究意识问题，他提出了一个惊人的假说：“你”，你的喜悦、悲伤、记忆和抱负，你的本体感觉和自由意志，实际上都只不过是一大群神经细胞及其相关分子的集体行为而已^[42]。在克里克看来，大脑是一个复杂的系统，把它看作黑箱无益于理解其内部的工作机理。他主张只有将脑活动和人类行为的所有方面和层级都纳入考虑之后，才有可能形成一个统领一切的大理论。克里克与科赫合作试图探明视觉信息在抵达视皮层之前经历的早期加工步骤，旨在寻找意识的“神经相关物”^[43-46]，即：产生特定意识知觉所需最少的神经元活动及其内在机制。他们一方面承认意识的神经相关物无法解决意识问题，另一方面又希望它能为意识理论带来巨大突破。

1994年，查默斯曾经宣称，研究神经元不可能揭示出为什么声波对我们耳朵的撞击，能引起我们对《贝多芬第五交响曲》的主观体验。在他看来，不论神经科学家对大脑的了解有多么透彻，他们都

不可能在物理世界与主观世界之间的“鸿沟”上，架起一座理解的桥梁。科赫则反驳说：“我怎么知道你的主观体验与我的一样？甚至我怎么知道你具有意识。”关于主观体验的问题，可能是对意识进行还原主义研究的最大难点，比如怎样解释红的程度；如果不同的人或动物个体在观察某个物体时主观体验有差异，那么基于还原论的解释就可能遇到很大麻烦。事实上，要严格确定大脑中的红色神经相关物是极其困难的，因为我们所进行的任何一项观察，都需要很多神经元的复杂相互作用才能完成，也许根本就不存在所谓的红色神经相关物：只要它出现，即意味着你看到了红色。

随着越来越多的人踏入意识的领地，关于大脑的实验数据迅速积累，对意识的疾病诊疗和神经机制的理解取得了进展^[47,48]，大量的意识理论也如雨后春笋般地接踵而至^[49,50]，其中影响较大的有：全局工作空间理论^[51]、整合信息理论^[52]、高阶理论^[53]、复馈预测处理理论^[54]、注意图式理论^[55]、协调客观还原理论^[56]、活镜理论^[57]等。还有些受到一定关注的观点或方法包括：无限联想学习^[58]、神经底物分析^[59]、公理图解数学^[60]、计算建模证伪^[61]、神经心智场^[62]、休息状态范式^[63]、系统不可分度^[64]、系统复杂度^[65]、记忆系统演化^[66]、具身量子纠缠^[67]、通信起源假说^[68]、神经标记解释^[69]，等等。然而，所有这些理论、观点和方法，没有任何一个能够解释现象意识，也没有任何一个被普遍接受。意识问题仍然处于重重迷雾当中，哲学不知道它如何从物质中产生，神经科学也不知道它如何从大脑中涌现。这个世界之结，这个上帝最后的秘密，还在等待能够破解它的爱因斯坦降临^[27]。

4 认知相对论

2005年，也就是爱因斯坦提出狭义相对论之后的100年，本文研究者在《21世纪100个交叉科学难题》上发表了《揭开意识的奥秘》一文^[70]，阐述了认知相对论的两条基本原理，即：世界的相对性原理和符号的相对性原理，拉开了从第1性原理出发揭开意识奥秘的序幕，2018年还在国际智能科学会议上作了关于“认知相对论”的特邀报告。

4.1 世界的相对性原理

意识问题涉及外部世界的真实性。我们所看到的外部世界是真实的吗？由于在这个世界上生活，我们很难怀疑它的真实性，绝大多数人相信：确实存在一个外部世界，它大体上不依赖于我们对它的观察。另一方面，我们似乎又必须通过观察才能知道外部世界的存在，否则怎么能够知道它存在呢？事实上，我们所看到的外部世界就是我们所体验到

的主观世界。虽然人类的不同个体看到的外部世界也许是相同的，但不同种类动物看到的外部世界却可能不同。例如，蜜蜂可以看到偏振光，蛇可以看到红外线，蝙蝠可以听到超声波，大象可以听到次声波，狗可以区分成千上万种气味，鲨鱼可以感知到电场，青蛙只能看到运动的物体，等等。这说明人类和其他动物的主体在意识中所能体验到的主观世界并不是唯一的。据此，本文提出：

世界的相对性原理：主体在意识中所能体验到的主观世界极大地受限于其对本体世界的观察方式。

世界的相对性原理实际上可以看作人择原理的推广。人择原理有若干不同的版本，其核心思想在于强调宇宙的本质极大地受限于我们人类对它的观察，宇宙之所以是这个样子就是为了让人类出现来观察它。

无数哲学家和科学家一直在设法建造一座桥梁，以连接主观世界和客观世界之间的鸿沟，但是他们没有想到，我们的大脑和身体恰恰就是这座桥梁。根据世界的相对性原理，意识所能体验到的主观世界并不完全对应于纯粹作为客体的本体世界。在这里，本体世界是指康德所说的由物自身构成的世界，或者理解为去掉一切观察者之后所剩下的物质世界。尽管本体世界可能只有一个，但是主观世界却可以丰富多彩，两者之间的桥梁正是主体对本体世界所采用的观察方式，特别是主体所拥有的传感器和认知结构。

4.2 符号的相对性原理

人类不仅能够意识中体验到主观世界，还可以用语言来描述和认识主观世界。语言是一种由符号构成的交流工具，可以用来表达情感、思维、想法和经历等意识体验。在不同的国家和地区，人们往往说着不同声波形式的语言，全世界大约有5 000~7 000种。虽然各种语言的物理形式可能完全不同，但是每一种都可以用来思维和交流。例如，中国人用汉语，英国人用英语，日本人用日语。而且从现实生活的角度来看，任何一种物理形式的语言对于人类描述和认识世界来说都是等价的，据此本文提出：

符号的相对性原理：当主体用语言来描述和认识其在意识中所体验到的主观世界时，所有语言的物理形式都等价。

对符号的相对性原理这一名称，切不可从字面上去理解，因为它借用和参考了爱因斯坦的狭义相对性原理，即所有惯性坐标系对于物理定律的描述都等价。在符号的相对性原理中，可以把各种语言的物理形式理解为描述主观世界的不同坐标系。

根据符号的相对性原理，主体在采用语言表达思维的内容时可以选择任意的物理形式，比如气流或者水流。事实上，人类的语音就是一种气流。而水生动物，比如海豚，就可能是通过水流的振动来交流思想的。而在大脑中发生的思维过程，笔者认为既不是气流，也不是水流，而应该是通过电流来实现的。

4.3 世界、语言和心灵的关系

许多学者指出，深度学习的背后缺乏世界模型，特别是通用世界模型，这成了制约人工智能进一步发展和突破的瓶颈。然而，却没人知道世界模型到底是什么。有一种观点以为，世界就是我们能够用语言描述的所有事物，包括常识。其典型代表就是逻辑原子主义和语言哲学，主张世界可以通过语言的分析 and 逻辑的构造建立起来。在认知相对论的框架下，本体世界是不可观察的、甚至是不可知的，关于本体世界的通用世界模型也是不存在的。根据世界的相对性原理，主体所能观察到的必定是一个受限于其自身传感器和认知结构的主观世界，显然它不可能是对本体世界的完美复制。如果把本体世界比作一头完整的大象，那么主观世界最多只是盲人摸到的那个部分。因此，即使人类的主观世界可以用一个足够大的语言模型来充分描述，这个语言模型也不可能成为人工智能的通用世界模型。更何况任何语言模型都是有局限的，总有它不能描述的东西，比如它不可能给世界上的每一个基本粒子都命名，因为每一个命名都必须消耗另一些基本粒子来实现某种物理形式的表达，从而导致无穷后退的问题。

世界和语言是两个有本质区别的东西。人类的世界，是由图像、声音、气味、味道和疼痛等意识体验构成的主观世界，它形成我们的感性。这个感性的世界是可以用来描述和认识的，而语言的基本单位——符号，虽然可以采用任意的物理形式来实现，但必须在约定为某种确定的物理形式之后，相应的符号语言才能成为社会交流的工具。一般而言，符号应该选择约定为具有相对简单的物理形式，以便于生成、记录、发送、传输和接收。如果没有经过选择和约定，任何图像和声音都不能看作符号，只能代表它们自己，而非他物。可是一旦被选择和约定，符号就可以用来指代别的东西，包括我们无法感知的红外线、紫外线、超声波、次声波等，由此产生判断、推理和思维，并形成我们的理性。严格地说，符号就是经过选择和约定之后用来代表其他事物的物理形式。尽管符号必须以某种确定的物理形式存在，但却是被彻底抽去一切物理

属性的东西。在这种抽象的意义上，符号之间是不能产生物理相互作用的。另一方面，由于符号仍然是一种物理形式，我们的心灵就可以利用符号及其构成的语言对物理世界产生因果作用。

语言哲学试图把世界的构造归结为语言的分析，并进而把世界和语言等同起来，但没有成功。符号主义人工智能以物理符号系统假设为基础，其背后还有一个更深层的隐含假设是：世界=语言。在这个隐含假设下，世界上所有的事情，不过就是语言的分析 and 符号的计算，包括我们的心智。如果世界真的等于语言，那么我们只要把心中的任何愿望用语言写在纸上就等于实现了，显然这是不可能的！

本文以为，世界和语言都是心灵的组成部分，其共同基础是意识，意识才是心灵中最重要方面。那么，意识又是如何从物质中产生的呢？下面将通过意识的等效原理来说明。

5 意识的等效原理

爱因斯坦的广义相对论有一条基本原理称为等效原理，其含义是指重力场与以适当加速度运动的参考系等价。换言之，观察者无法区分重力场和适当的加速运动。为了破解意识如何从物质中产生的问题，本文研究者将全息光学的波前再现原理进行了推广，提出了意识的等效原理。用以说明意识从物质产生所需要的物理条件，解决主观体验或现象意识的困难问题，推导意识能力受限于感觉能力且以感觉容量为上界的认知基本定理，并分析意识在哪里和自我是什么的可能性。

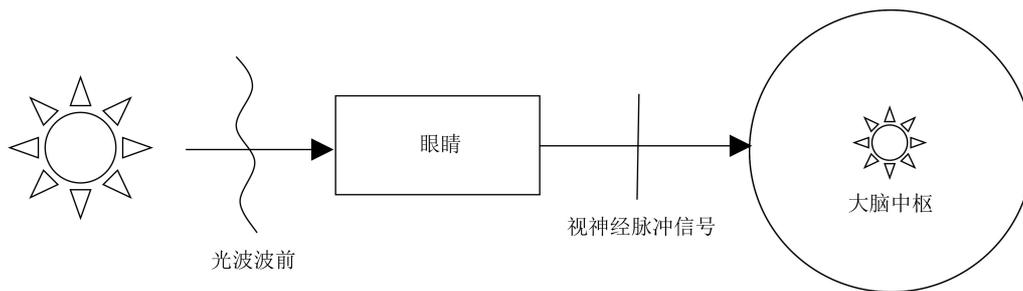
5.1 波前再现原理

波前再现是指用一束与参考光束的波长和传播方向完全相同的光束照射全息图，则用眼睛可以观察到一幅非常逼真的物体形像，悬空地再现在全息图后面原来的位置上。波前再现原理宣称，我们看到物体的形像，并不一定需要物体本身出现在我们面前，只要物体的光波波前再现即可。如图1所示，我们看到太阳，或者说大脑中枢意识到太阳，未必需要太阳真的出现在我们面前，只要太阳的光波波前再现即可。因此，眼见未必为实，我们看到的可能只是虚幻的光波波前，而非真实的物体^[7]。

理论上，波前再现是使记录时被“冻结”在全息干板上的光波波前在特定条件下“复活”，构成与原波前完全相同的新波前继续传播，形成3维立体像的过程。假设在全息干板H上的坐标为 x, y ，把物体光波和参考光波分别用复振幅表示为

$$O(x, y) = O_0(x, y) \exp[j\phi_o(x, y)] \quad (1)$$

$$R(x, y) = R_0(x, y) \exp[j\phi_r(x, y)] \quad (2)$$

图1 物体的光波波前 \mathbf{X} 和其对眼睛所激发的视神经脉冲信号 \mathbf{Y}

其中 O_0, ϕ_0 分别是物体光波达到全息干板 H 上的振幅和相位分布。物体光波和参考光波的相干叠加的总光场为

$$U(x, y) = O(x, y) + R(x, y) \quad (3)$$

全息干板 H 上记录的是干涉场的光强分布, 将 $U(x, y)$ 与其复共轭 $U^*(x, y)$ 相乘, 易知曝光强度为

$$\begin{aligned} A(x, y) &= U(x, y) \cdot U^*(x, y) \\ &= |O|^2 + |R|^2 + O \cdot R^* + O^* \cdot R \end{aligned} \quad (4)$$

经线性处理后, 底片的透过率函数 t_H 与曝光强度成正比, 略去无关紧要的比例常数可得

$$t_H(x, y) = |O|^2 + |R|^2 + O^* \cdot R + O \cdot R^* \quad (5)$$

由此得到的底片就是全息照片, 又称全息图。

如果再用参考光波 $R(x, y)$ 照射全息图 H , 透过 H 后的光振幅 $U'(x, y)$ 可表达为

$$\begin{aligned} U'(x, y) &= R(x, y) \cdot t_H(x, y) \\ &= R_0(x, y) \exp[j\phi_r(x, y)] \\ &\quad \cdot [|O|^2 + |R|^2 + O^* \cdot R + O \cdot R^*] \\ &= R_0 (O_0^2 + R_0^2) \exp[j\phi_r] + R_0^2 O_0 \exp[j\phi_0] \\ &\quad + R_0^2 O_0 \exp[-j(\phi_0 - 2\phi_r)] \end{aligned} \quad (6)$$

不难看出, 在 $U'(x, y)$ 中的第2项光波是 $R_0^2 O_0 \exp[j\phi_0]$, 与原来的物体光波基本无两样, 只是增加了一个常数因子, 它就是物体光波的再现, 亦称波前再现。

从1948年盖伯提出波前再现的思想开始, 全息照相技术进入发展轨道, 但一直到1960年激光的出现之后, 才导致了許多始料未及的应用。盖伯因此获得了1971年的诺贝尔物理学奖。全息照相技术的巨大成功, 为本文研究者提出意识的等效原理, 提供了坚实的实践基础。

5.2 意识的等效原理

根据波前再现原理, 物体光波和其光波波前的再现, 对人类主体产生视觉意识的作用而言, 具有等价的效果。然而, 无论是物体光波, 还是其光波波前, 都必须经过眼睛转化为视神经脉冲信号并传递

到大脑的高级中枢之后, 才能产生关于物体的视觉意识。眼睛的作用是把物体光波或其光波波前转变为视神经脉冲信号。如果用 \mathbf{X} 表示物体光波或其光波波前, 用 \mathbf{Y} 表示 \mathbf{X} 所激发的视神经脉冲信号, 那么眼睛的功能可以用一个变换函数 T 来表示, 即

$$\mathbf{Y} = T(\mathbf{X}) \quad (7)$$

其中, \mathbf{X} 和 \mathbf{Y} 在数学上未必同构, 两者的维数可能相差非常大。神经解剖学研究表明, 人类的视网膜上约有 12×10^7 个对弱光刺激敏感的视杆细胞, 约有 $65 \times 10^5 \sim 70 \times 10^5$ 个对强光和颜色敏感的视锥细胞, 而视网膜的背后仅有 $1 \times 10^6 \sim 2 \times 10^6$ 条视神经纤维。这意味着, \mathbf{X} 可能有上亿的维数, 远远大于 \mathbf{Y} 的维数。

人类的意识无疑是由大脑中枢产生的。从图1不难看出, 由于物体的光波波前 \mathbf{X} 和视神经脉冲信号 \mathbf{Y} 是被眼睛完全分开的, 所以大脑中枢无法直接接触到 \mathbf{X} , 只能利用 \mathbf{Y} 来产生意识, 而且 \mathbf{X} 必须通过 \mathbf{Y} 才能让大脑中枢产生意识。这意味着: 如果 \mathbf{X} 在主体的大脑中枢引起了某种视觉意识, 那么这种视觉意识也必然是通过 \mathbf{Y} 引起的, 因此 \mathbf{Y} 一定能引起相同的视觉意识, 无论 \mathbf{Y} 和 \mathbf{X} 是否同构, 也无论 \mathbf{Y} 和 \mathbf{X} 是否维数相同。换言之, 物体的光波波前和其对人类主体的眼睛所激发的视神经脉冲信号在产生该主体的视觉意识时具有等价的效果。如果进一步把这种等价性推广到其他任何传感器的情况, 如耳、鼻、舌、肤等, 就是意识的等效原理, 即:

意识的等效原理: 任何物理刺激和其对主体的传感器所激发的电信号在产生该主体的意识时都具有等价的效果。

在意识的等效原理中, 传感器是指把物理刺激变换为电信号的装置, 不仅包括人类和动物身上的眼、耳、鼻、舌、肤等生物传感器, 也包括给机器人配置的各种非生物人造传感器。主体是指任何具有意识的物理系统, 可以是人类和动物, 也可以是机器。通常, 意识是由物理刺激直接引起的。但根据意识的等效原理, 意识也可以由等价的电信号或感觉信号引起。感觉是主体通过传感器从环境获取信息的过程, 由感觉得到的信息称为感觉信息, 其

物理表达形式称为感觉信号，也就是物理刺激 \mathbf{X} 对主体的传感器所激发的电信号 \mathbf{Y} 。感知是让感觉信号进入意识的过程，感知的结果称为知觉。生物主体的意识是通过神经电信号产生的。理论上，机器主体的意识也可以通过电子电路的电信号来产生。

有必要指出的是，意识的等效原理并不能由广义相对论的等效原理推导出来，它与后者分属不同的领域，是笔者直接作为认知相对论的公理提出来的，由于涉及心理活动和认知过程的主观性本质，所以区别于任何物质科学的客观性公理、原理或定律。虽然意识的等效原理看似简单，但它具有非常深刻的内涵，也不是随随便便就可以得到的，而是笔者从全息光学的波前再现原理出发，结合神经科学的知识，经过30多年的摸索和思考，以及悟道之后的严格逻辑论证，才最终形成和确立的。

5.3 感觉能力和感觉容量

在香农信息论的框架下^[72]，可以把传感器看作一个信道，其作用是把环境的物理刺激传输到主体的信息处理中枢。此时，传感器的功能不再是一个确定的变换函数，而应该描述为一个条件概率转移函数 $P(\mathbf{Y}|\mathbf{X})$ 。如果用 \mathbf{X} 表示环境的物理刺激，用 \mathbf{Y} 表示 \mathbf{X} 对传感器所激发的电信号，用 $P(\mathbf{X})$ 和 $P(\mathbf{Y})$ 分别表示 \mathbf{X} 和 \mathbf{Y} 的概率分布，用 $P(\mathbf{X},\mathbf{Y})$ 表示 \mathbf{X} 和 \mathbf{Y} 联合概率分布，那么主体在通过传感器接收到电信号 \mathbf{Y} 时从环境获取的信息量就是 \mathbf{X} 和 \mathbf{Y} 的平均互信息

$$\begin{aligned} I(\mathbf{X};\mathbf{Y}) &= H(\mathbf{X}) - H(\mathbf{X}|\mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y}|\mathbf{X}) \\ &= \sum_{\mathbf{X},\mathbf{Y}} P(\mathbf{X},\mathbf{Y}) \log_2 \frac{P(\mathbf{X},\mathbf{Y})}{P(\mathbf{X})P(\mathbf{Y})} \geq 0 \end{aligned} \quad (8)$$

其中， $H(\mathbf{X})$ 和 $H(\mathbf{Y})$ 分别表示 \mathbf{X} 和 \mathbf{Y} 的信息熵， $H(\mathbf{X}|\mathbf{Y})$ 和 $H(\mathbf{Y}|\mathbf{X})$ 表示条件信息熵。有必要指出， $I(\mathbf{X};\mathbf{Y})$ 具有数学上严格的对称性和非负性，且仅在 \mathbf{X} 和 \mathbf{Y} 相互独立时 $I(\mathbf{X};\mathbf{Y})=0$ ，无论 \mathbf{X} 和 \mathbf{Y} 是否恒等，也无论 \mathbf{X} 和 \mathbf{Y} 的维数是否相同。

主体在不同环境中的感觉能力是不同的。如果把物理刺激的概率分布 $P(\mathbf{X})$ 看作对某种环境的描述，那么 $I(\mathbf{X};\mathbf{Y})$ 的认知意义就是主体在这种环境中的感觉能力的描述。如果 \mathbf{X} 和 \mathbf{Y} 相互独立，即 $P(\mathbf{X},\mathbf{Y})=P(\mathbf{X})P(\mathbf{Y})$ ，那么感觉能力 $I(\mathbf{X};\mathbf{Y})=0$ ，这意味着主体无法从概率分布为 $P(\mathbf{X})$ 的环境中获取包含信息的信号，相当于传感器在这种环境中不能传输任何信息，主体的感觉能力为0。比如，在全黑的山洞环境中，我们什么也看不见，眼睛的视觉刺激 \mathbf{X} 是恒定不变的全黑，因而与眼睛的视觉信号 \mathbf{Y} 独立，所以 $I(\mathbf{X};\mathbf{Y})=0$ ，即眼睛对全黑环境的视觉能力为0。

当物理刺激具有某个最佳概率分布 $P^*(\mathbf{X})$ 时， $I(\mathbf{X};\mathbf{Y})$ 描述的感觉能力达到最大值

$$C = \max_{P(\mathbf{X})} I(\mathbf{X};\mathbf{Y}) \quad (9)$$

根据香农信息论^[72]， C 就是在把传感器看作信道时的信道容量，笔者将其称为感觉容量。在通信系统中，信道容量是对信道传输速率的限制，只有当传输速率小于信道容量时，信息在传输过程中才能保证不出错。而感觉容量侧重于信息在认知方面的意义，用来表达主体的感觉能力的上界。只有当环境的变化速率不超过感觉容量时，主体的传感器才能正常工作，否则就会难以反映环境的真实情况而出错。比如，我们无法看到子弹飞行的细节情况，因为飞得太快，其变化的速率远远超过了人类的视觉容量。由于感觉容量的差别，不同的主体可能擅长适应不同的环境。在与感觉容量相匹配的最佳环境中，主体才能获得最大的感觉能力，从而利用更多的环境信息来指导产生更有效的适应性行为。

5.4 意识产生的物理条件

虽然感觉信号可以用来产生意识，但是未必都用来产生意识。有些感觉信号可能被主体用于身体的调控或者价值的评判，而不是产生意识。图2中，感觉信号就是物理刺激 \mathbf{X} 对主体的传感器所激发的电信号 \mathbf{Y} ，它实际上可能又被分解整合为多组不同的电信号，比如调控信号 \mathbf{Y}_1 、评判信号 \mathbf{Y}_2 和意识信号 \mathbf{Y}_3 。注意， \mathbf{Y}_1 、 \mathbf{Y}_2 和 \mathbf{Y}_3 未必完全独立，可能存在相关性。其中， \mathbf{Y}_1 用来调控身体器官(比如眼睛)的状态， \mathbf{Y}_2 用来评判物理刺激 \mathbf{X} 对主体的价值(比如令人愉悦还是恐惧)，只有 \mathbf{Y}_3 才直接用来产生对物理刺激 \mathbf{X} 的意识(比如形状、颜色和声音)。根据意识的等效原理， \mathbf{Y}_3 和 \mathbf{X} 在产生主体的意识时具有等价的效果。原则上说， \mathbf{Y}_2 和 \mathbf{X} 对产生主体的价值评判作用也是等价的， \mathbf{Y}_1 和 \mathbf{X} 对产生主体的身体调控作用同样是等价的。虽然 \mathbf{Y}_2 和 \mathbf{Y}_1 不直接用来产生意识，我们也无法在意识中体验到它们是什么，但根据 \mathbf{Y}_2 产生的评判结果和根据 \mathbf{Y}_1 产生的调控结果有可能再通过某些内部传感器进入主体的意识之中，比如让我们意识到自己的情绪或者动作。

从图2不难看出，解释意识的关键在于说明意识信号 \mathbf{Y}_3 如何让主体产生意识，特别是其中所需

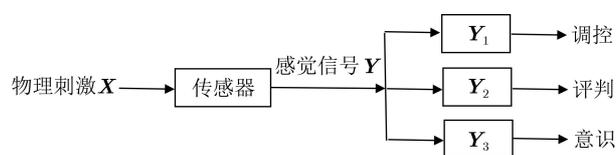


图2 物理刺激 \mathbf{X} 和其对传感器所激发的电信号 \mathbf{Y} ，以及 \mathbf{Y} 经分解整合后得到信号 \mathbf{Y}_1 、 \mathbf{Y}_2 和 \mathbf{Y}_3

要的物理条件是什么。首先,意识的产生需要一定的时间。视觉暂留现象的存在表明,人类的意识并不是瞬间产生的,而是一个短暂的时间过程,这正是电影技术得以实现的依据。电影播放胶片的速度是24格/s,由此可以推算出人类产生视觉意识大约需要41.6 ms的时间,如果时间更短有些视觉内容就可能融合在一起。根据进化论,人类的视觉系统是经过千百万年自然演化不断适应环境的结果,所以还可以推断,在人类的生存环境中,视觉信息的变化率一般不会超过24 Hz。只要大致按照这个速率来采样视觉信息并产生视觉意识,人类基本上就可以满足日常生活的需要。另外,我们产生其他意识,包括听觉、嗅觉、味觉和触觉等,也是需要时间的,只是时间的长短可能有所不同。

主体又是怎样把意识信号 Y_3 变成关于物理刺激 X 的意识呢?当主体在保持 Y_3 时,如果再等待进行可选操作旨在将其与记忆、行动或者语言等建立新关联,比如经典条件反射和操作条件反射,那么主体就会因为无法区分电信号 Y_3 和物理刺激 X 而把 Y_3 等同于 X 。根据意识的等效原理,只要意识信号 Y_3 再现,就等价于物理刺激 X 的再现。因此,在该主体看来,就好像产生了关于物理刺激 X 的意识。即便物理刺激 X 并没有真的出现在其面前,该主体也会以为它出现在了自已的面前。这就是所谓的主观体验或现象意识,它是在主体保持意识信号 Y_3 并等待对其进行可选操作时涌现出来的,如图3所示。

主观体验或现象意识,又称为感受质,它被普遍认为是意识研究中最根本的困难问题^[73],而其他问题则属于容易问题,即那些可以直接用认知科学的标准方法处理的问题^[12],包括:信息的整合、语言的生成、行为的控制,等等。现象意识之所以困难,是因为它涉及到意识经验的主观性。物理学原理都是客观的,不可能对主观性给出充分的解释。在这个意义上,意识是不可物理还原的。也就是说,在人类的物理身体和意识经验之间,必定存在着一条解释上的鸿沟。这条鸿沟也不可能通过神经科学来填补,因为神经科学仅仅通过仪器设备的客观测量同样无法充分解释意识经验的主观性。有些科学家认为解释意识可能需要一种新物理学,而笔

者认为只需在认知相对论的框架中增加意识的等效原理即可。根据该原理,解释现象意识的关键就是如何让主体以为意识信号的再现就是物理刺激的再现。现象意识就是在主体无法区分意识信号和物理刺激时涌现出来的。由于在意识涌现时主体无法区分意识信号和物理刺激,所以主体的意识就总会指向物理刺激,从而让意识具有了总是关涉某物的意向性^[74]。此外,该主体的意识信号只能在他自己的身体内流动,也只能让他自己涌现出现象意识,这就解释了意识的私密性。

结合图2和图3可以看出,意识信号 Y_3 与调控信号 Y_1 和评判信号 Y_2 的主要区别在于:它需要主体将其保持一段时间并等待进行可选操作,而调控信号 Y_1 和评判信号 Y_2 则无需进行保持和可选操作。所以,主体对调控信号 Y_1 和评判信号 Y_2 的响应可以实现完全连续自动化,通常速度更快。比如,我们在碰到伤害刺激时会自动躲闪,根本无需意识的参与,从而达到快速应对环境突变的目的。而意识信号的处理涉及到保持、等待和可选操作的过程,期间还可能受到环境随机变化的影响,难以实现完全连续自动化,花费的时间较多,速度相对较慢。从进化的角度看,人类和动物也不太可能先天地把所有生存策略都编码到基因中去,这就需要具备一定的应对环境不确定性的学习能力。在认知相对论的框架下,学习必须在意识的基础上才能完成,只有被意识到的东西才能进行学习,意识是学习的必要条件。另一方面,意识又是一种应对环境不确定性的机制,而不确定性是不可能通过完全确定的数学公式计算出来的,也是不可能通过纯粹的计算过程来定义的。因此,意识不是完全可计算的。

意识的感受质是什么呢?感受质是由环境的物理刺激以及主体的传感器和意识结构所共同决定的。意识结构就是把物理刺激 X 对传感器所激发的电信号 Y 进行分解整合的信号处理模块,其作用就是产生意识信号 Y_3 。如果物理刺激是 X ,意识信号是 Y_3 ,那么感受质在本质上就是意识信号 Y_3 与 X 之间的物理转化关系,它不仅反映了主体从环境获取信息的能力,也基本上决定了主体在意识中所能观察到的主观世界。从根本上说,感受质是不能严格定义的,它虽然可以让主体产生现象意识的主观体验,但不能告诉主体其物理本质是什么。事实上,物理学也只能告诉我们物质与物质之间的某些关系,并不能告诉我们物质的本质是什么。一切跟意识有关的活动都必须以感受质为基础才能建立和发展起来。根据世界的相对性原理,如果主体拥有不同的传感器和意识结构,即使在相同的环境中,也可能形成

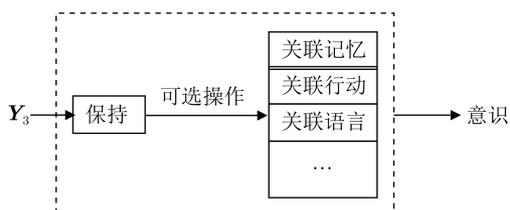


图3 意识产生的物理条件

不同的感受质，从而在意识中体验到不同的主观世界。当主体的主观世界不能有效匹配外部世界时，就可能导致主体难以顺利地适应环境。很多神经损伤患者就属于这种情况，他们可能看不出物体的形状，难以识别物体是什么^[75]。

5.5 现象信息和意识能力

根据康德的物自身学说^[76]，我们不可能认识物自身的本质，只能观察到关于物自身的现象。现象就是物自身看起来的样子，或者说物自身被主体观察到的那一部分，就像我们只能看到一个人的外表，但看不到他的内脏和内心。在认知相对论的框架下，现象必须由物理刺激引起，不可能从纯粹的符号计算中产生，就像无论我们怎么打算盘，都不可能从算盘中冒出七色彩虹，更不可能冒出百元大钞。物理刺激作用于传感器，其结果是产生感觉信号。如果物理刺激是 X ，感觉信号是 Y ，对 Y 分解整合得到的意识信号是 Y_3 ，那么现象就是主体为了适应环境在保持 Y_3 并等待进行可选操作时从 X 所获得的信息，笔者把这种信息称为现象信息，其信息量的大小可以表示为 X 和 Y_3 的平均互信息

$$\begin{aligned} I(X; Y_3) &= H(X) - H(X|Y_3) \\ &= H(Y_3) - H(Y_3|X) = I(Y_3; X) \geq 0 \end{aligned} \quad (10)$$

其中，意识信号 Y_3 可以看作是主体利用感觉信号 Y 在信息处理中枢对物理刺激 X 的某种解读结果，它在必要时就涌现为现象，而当现象涌现时就产生了现象意识。虽然意识信号 Y_3 并不是物理刺激 X 本身，但它包含了主体想要知道的关于 X 的信息，并且主体在意识中是无法区分 Y_3 和 X 的，这种不可区分性就是信息论意义下的心物同一。

现象信息量 $I(X; Y_3)$ 的认知意义在于它代表了主体利用意识信号 Y_3 对物理刺激 X 的解读能力，笔者将其称为意识能力，也就是主体通过现象意识知晓环境所发生的事情的能力。

5.6 认知基本定理

由于感觉信号 Y 是物理刺激 X 通过传感器的变换处理而来的，意识信号 Y_3 又是通过对感觉信号 Y 的分解整合处理得到的，所以 $X \rightarrow Y \rightarrow Y_3$ 形成一个马尔可夫链，根据有关的信息处理定理^[72]有

$$I(X; Y_3) \leq I(X; Y) \leq C \quad (11)$$

因此，现象信息量表达的意识能力不会超过主体的感觉能力，更不会超过主体的感觉容量。简言之，意识能力受限于感觉能力且以感觉容量为上界。笔者把这个结论称为认知基本定理，它是对主体认知界限的形式化表达，也是无数哲学家在探索

认识论的过程中梦寐以求的结论。例如，洛克、休谟、康德和马赫等人都研究过人类感觉对认识的限制问题，试图建立一门关于知识范围的科学。认知基本定理的直观解释是：被主体意识到的现象信息，一定不会超过流入主体的感觉信息。这完全符合我们的日常生活经验：我们对从来没见过的事物不可能产生现象意识，最多只能进行天马行空的想象。需要指出的是，感觉能力和感觉容量一般是先天给定的，而意识能力则是可以随着个体成长和后天训练而发展变化的。这意味着从物理刺激 X 到意识信号 Y_3 的概率转移函数 $P(Y_3|X)$ 可能是不断变化的，主体的意识能力也可能是不断变化的，但不管怎么变、如何变，意识能力始终受限于感觉能力且以感觉容量为上界。

根据认知基本定理，如果主体在某个环境中没有感觉能力，也不会有意识能力。因为感觉能力 $I(X; Y)=0$ ，必然推出现象信息量 $I(X; Y_3)=0$ ，即意识能力为0，所以不可能形成关于环境的现象意识。比如，在全黑的环境中我们没有视觉能力，也不会有视觉意识。虽然根据意识的等效原理，主体也可以通过其他方式(如整合过去的记忆和经验)让意识信号 Y_3 再现，从而在一定的条件让关于视觉刺激 X 的意识再现，但这种意识再现不可能是主体从物理刺激 X 那里看到的外界现象，而可能只是主体对 X 的构建、想象、回忆、反省或者梦幻等，严格地说不应再被称为现象意识，至少不再是纯粹的现象意识了。

5.7 意识在哪里

除了意识是什么，人们还常常会问：意识在哪里？在认知相对论的框架下，生物意识一般需要多个功能结构的协同工作才能产生，可能涉及多个不同的核团和脑区，应该更多地表现出协同性和整体性，而不是局部定位的。对人类和高等动物来说，笔者认为意识可能主要发生在丘脑和大脑皮层之间的信息互动过程中。因为丘脑不仅汇聚了来自皮下结构的输入信息，而且既向大脑皮层广泛投射输出信息，也从大脑皮层广泛接收反馈信息以促进并保持向皮层投射的丘脑神经元的持续激活^[77]，如图4所示。意识信号 Y_3 很可能是在丘脑中整合形成的，当主体通过某种方式保持意识信号并等待对其进行旨在建立新关联的可选操作时，由于无法区分意识信号和物理刺激，意识就涌现出来了。在笔者看来，对意识信号的保持机制，就是所谓的注意；而进行可选操作的能力，则让主体拥有了自由意志和创造力。

5.8 自我是什么

人们也非常想知道：自我是什么？笔者认为在

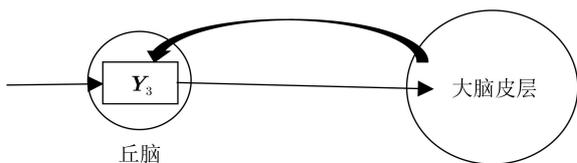


图4 意识可能发生在丘脑和大脑皮层的信息互动过程中

大脑中存在一个特殊的神经元集群，并将其称为“自我中心核”，专门用来产生关于“自我”的意识。自我中心核通过内部感官汇聚来自全身各处的体感信号进行分解整合，形成“自我意识信号”。当主体在保持自我意识信号并等待对其进行旨在建立新关联的可选操作时，根据意识的等效原理，自我意识便涌现出来了。自我意识是主体关于自身存在状态的整体性、统一性意识，它让主体具备了选择“做与不做”的自由意志，从而需要对自己的行为负责。可是，我们为什么又能够意识到自己有意识呢？因为产生自我意识的神经信号也会通过内部传感器汇聚到“自我中心核”引起关于自我的意识，于是我们就能意识到刚才的意识其实就是自我产生的意识。大脑两半球之间的小结构——前楔前叶(anterior precuneus)，可能对自我意识的形成起着“中心核”的作用^[78]。

6 研究意识问题的新纲领

2005年7月1日，美国《科学》杂志上发表了125个挑战性的科学问题，其中第2个便是：意识的生物学基础是什么？很多人认为，只有生命才有意识，意识是一种生物现象，必须以生物学的方法来解决。然而，生命并非意识的充分条件，比如植物没有意识，受精的鸡蛋也没有意识。虽然人们对动物是否有意识的问题存在诸多争议^[79-82]，但一直在试图测试动物和人类是否有意识^[83-85]，甚至探讨如何实现人工意识的问题^[86-88]。此外，意识还是一种心理现象，具有主观性，似乎又与科学的客观性风马牛不相及。有些人甚至断言，将意识作为科学的研究对象是不可能的。尽管如此，研究意识之谜的竞赛还是在全世界激动人心地展开了^[89]。2018年，Nature杂志曾提出6个最重大的科学问题，其中之一就是：“什么是意识？”^[90]。而如何使意识的困难问题变得更容易，则成了很多科学家争相角逐的主题^[73]。

本文认为，研究意识问题必须以认知相对论的3条基本原理作为出发点，才可能取得实质性和突破性的成功。这3条基本原理，即：世界的相对性原理、符号的相对性原理和意识的等效原理，是从生活经验和科学事实中总结、归纳、提炼出来的，

简单而深刻，独立且相容。它们属于宏观系统层面的原理，既有哲学的思辨，又有科学的内涵，旨在建立一座联系主观世界和客观世界的桥梁，把心理现象和物质现象关联起来。它们不仅涵盖了世界、语言和心灵之间的关系，也说明了意识从物质产生所需要的物理条件，并解决了主观体验或现象意识的困难问题，还催生了意识能力受限于感觉能力且以感觉容量为上界的认知基本定理。特别是，由于它们论述了主观性的基础，所以既不能被任何物质科学的客观原理所完全蕴含和推导，又可以对意识的重要特征给出非常有力度的解释。根据上述3条基本原理，本文完善了原来的认知相对论纲领^[70]，并提出了研究意识问题的新纲领，详细描述如下：

(1) 意识的研究应该分为3个方向：人类的意识、动物的意识和机器的意识，同时要注意采用不同的方法。不要轻易将人类的意识强加到动物身上，也不要过高地期望机器会拥有与人类相同的意识。因为人类、动物和机器观察本体世界的方式可能存在很大的区别，根据世界的相对性原理，他们在意识中所能观察到的主观世界可能很不相同。

(2) 对人类来说，意识的研究应该集中在发生意识障碍的机制，以及治疗这些障碍的方法(研究正常人受到的限制太大)，而不要试图去建立所谓的脑模型来阐明正常人的意识活动，因为我们的大脑和身体就是一个很好的解释，而脑模型即使能够产生意识，也可能根本就不是人类的意识。在认知相对论的框架下，机器可能拥有完全不同于人类和动物的意识，人工智能所期望的通用世界模型是不存在的。人类并不是因为有了通用世界模型才有意识，而是因为有了意识才有主观世界。根据世界的相对性原理，主观世界是依赖于观察者的，并非通用世界模型。

(3) 对动物来说，意识的研究首先应该集中在它们所观察到的世界究竟与人类有何不同，以及它们之间是如何交流的问题上，然后再去考虑动物神经系统的各个器官、核团和皮层功能区在产生意识活动时是怎样发生作用的，但不要过于狂热地去追求所谓的意识神经相关物，因为这种神经相关物一方面可能由于信息在神经系统中的大范围分布式表达而根本不存在，另一方面即使存在，它对不同类个体几乎没有意义，对同类个体也可能由于身体和经历的差异而具有不同的位置和模式。根据意识的等效原理，意识的产生需要一个保持意识信号并对其进行旨在与记忆、行动或者语言等建立新关联的可选操作过程，这可能涉及多个神经系统的核团和脑区的协同工作，应该更多地表现为协同性和整体

性，而不是局部定位性，所以很难被解释为一群独立神经元所发放的电信号。换言之，意识的神经相关物可能根本就不存在。

(4) 现象意识的主观体验，是意识的困难问题，在认知相对论的框架下已经可以得到解释，这在世界上当属首次。根据意识的等效原理，物理刺激对传感器所激发的电信号经过分解整合之后有一部分会形成意识信号，当主体在保持意识信号并等待进行旨在与记忆、行动或语言等建立新关联的可选操作时，由于无法区分物理刺激和意识信号，就会把意识信号等同于物理刺激，从而涌现出现象意识的主观体验，就像波前再现时物体涌现在我们面前一样。因此，解释现象意识的主观体验，并不需要大脑中有一个假想的小人在看，也无需一个光学屏幕在显示。事实上，科学家在打开头颅之后从来就没有发现过什么小人，也从来没有发现过按规则阵列整齐排列并发射着彩色光的脑细胞^[91]。

(5) 意识的产生涉及到主体对意识信号的保持和可选操作、以及建立新关联的机制，它可能还是学习、记忆和认知的基础。笔者认为，学习、记忆和认知是意识的必要条件。如果一个物理系统没有学习、记忆和认知的能力，比如桌子、铁块和石头，就不会有意识。这个观点是有认知神经科学依据的。文献^[92]的研究表明，只有物理刺激先于响应行为一段时间，在刺激和行为之间才能形成条件反射。如果物理刺激和响应行为同时发生，两者之间的关联并不是增强，反而是减弱。物理刺激需要提前保持一段时间，即在中枢神经系统中留下痕迹之后，才能与相继的响应行为建立所谓的“痕迹”反射。在认知相对论的框架下，这种痕迹就是意识，只有被主体意识到的东西，至少要借助被主体意识到的东西，才能进行学习、记忆和认知。

(6) 认知相对论是一个宏观系统层面的意识理论，不仅可以用来解释意识的所有重要特征，包括意向性、主观性、私密性、感受质、自我意识和自由意志等，而且可以为机器意识(或人工意识)的实现提供强有力的具体指导。在笔者看来，解决意识问题最重要和最关键的标志是制造出具有意识的人工机器，而不是去揭示意识产生的生物神经机制。事实上，虽然人类仍然不清楚鸟类飞行的神经控制机理，但是这并不妨碍我们掌握飞行的秘密，因为我们事实上制造出了比鸟类飞得更高、更远的飞机。所以，只要人类能够制造出有意识的机器，即使不清楚意识背后的神经机理，也同样意味着我们揭开了意识的奥秘。

(7) 有许多不同的方案设计具有意识的机器，

但是传感技术和模式识别技术的水平将决定着有意识机器所能观察到的世界的极限，同时也在很大程度上决定着机器智能的极限。根据世界的相对性原理，机器所能观察到的主观世界极大地受限于其对本体世界的观察方式，我们完全可以给机器装上不同于人类的传感装置和识别系统，让它直接看到偏振光、电磁场，听到超声波、次声波。根据符号的相对性原理，我们可以赋予机器某种特殊物理形式的语言，比如无线电波语言^[93]，代替人类的自然语言来描述和认识所观察到的主观世界，并进行独立思考和相互交流。根据意识的等效原理，当机器主体把物理刺激对其传感器所激发的电信号，通过分解整合形成意识信号，再等待对其进行旨在与记忆、行动或语言等建立新关联的可选操作时，由于无法区分意识信号和物理刺激，就会涌现出关于物理刺激的意识。

7 实现机器意识的新思路

在当代，意识被认为是认知科学研究的“皇冠”。认知科学的最终目标就是要用一组统一的基本原理来建立一门关于意识的科学。派利夏恩指出，如果认知科学有望成为像物理学、化学、生物学和地质学那样真正的科学，它就应当是不依赖于其他科学的原理、完全独立地建立在一套特定词汇和统一原理之上的学说^[38]。在笔者看来，认知相对论就是这样的学说，其中的3条基本原理完全独立于其他科学的原理，它们足以在主观世界和客观世界之间架起一座理解的桥梁，来破解现象意识或主观体验的困难问题。根据3条基本原理，认知相对论不仅从一个兼顾哲学和科学的宏观系统层面打开了意识这个历经千年沧桑的世界之结，还可以为人工智能的发展提供强大的理论指南和具体的实践思路，以期实现机器意识的圣杯。

目前，机器意识还处于非常初级的阶段^[94]，主要有3种研究方法^[95]：符号计算、神经网络和量子计算。已经证明，符号计算和神经网络在表达建模能力方面是等价的^[96]，而且由于前者不可能描述现象意识^[97]，后者也不可能描述现象意识。为了克服符号计算和神经网络的局限性，一些科学家又提出了量子计算的方法^[98]。不过，与前两者相比，量子计算只是采用了不同的物理机制来实现计算过程，虽然在某些方面的计算速度非常快，但在本质上仍然不可能描述现象意识。在笔者看来，任何计算理论都解释不了现象意识！计算的本质是符号操作，属于纯粹理性，而意识是从纯粹感性开始的。几个月大的婴儿就是纯粹感性的典型代表，他们不会说话，不懂语言，没有理性，但他们有纯粹感性的意

识, 这种意识是不可能用基于符号操作的计算理论来解释的。也许会有人反驳说, 计算可以超越符号, 未必是图灵计算。可是, 根据丘奇-图灵论题, 图灵机是计算的本质描述, 所有计算都是图灵计算, 不属于图灵计算的计算, 严格地说都不是计算。在这个意义上, 超越符号的计算就不是计算。心智的计算理论主张^[99], 心智是一种图灵式计算, 然而它把图像、声音和味道等所有能被我们感知到的模式都看作符号。图灵式计算就是对这些广义符号的操作。显然, 这与婴儿的感性意识相矛盾, 因为在婴儿那里根本就没有符号。此外, 文献^[100]提出的意识图灵机也没有跳出计算的窠臼, 对于解决现象意识的问题同样是无能为力、望洋兴叹。

类脑智能的研究期望将生物脑的工作原理应用于提高人工智能的水平^[101-104], 但是关于生物脑如何产生意识的工作原理还在一片茫茫迷雾和漫漫黑暗之中, 按照现有结构模拟的类脑思路来实现人工智能的机器意识恐怕就更加遥遥无期。事实上, 日本的5代机计划失败了, 欧盟的人脑计划也失败了, 这两个轰轰烈烈的计划曾经都考虑了类脑的思路。

本文认为, 在认知相对论框架下开展机器意识的实现研究, 才是一个未来可期、有望成功的正确方向。事实上, 迄今为止除了认知相对论以外, 任何其它的意识理论都解释不了现象意识。只有在机器上实现了现象意识, 才能算是真正的机器意识。那么应该如何实现机器意识呢? 本文提出了如下新思路:

(1) 机器意识只能在具身的硬件机器人上实现。不要指望通过在计算机上运行的程序来实现机器意识。Searle已经给出了中文屋论证。在笔者看来, 如果意识是程序运行的结果, 那么意识就可以独立于身体而存在, 永不消亡, 因为程序可以拷贝到任何地方永久保存, 也可以在任何一台计算机上永远运行。而这显然悖于科学界的共识: 意识依附于我们的肉体并且随着肉体的死亡而消亡。另一方面, 由于程序的执行是一个计算过程, 其输入是抽去一切物理属性的数字或符号, 其结果是物理实现无关的, 既可以用计算机的运行来得到, 也可以用齿轮的转动来得到, 甚至还可以用算盘的拨动来得到, 区别只是速度的快慢不同。而根据世界的相对性原理, 意识是物理实现相关的, 极大地受限于主体对本体世界的观察方式。计算和意识, 前者跟物理实现无关, 后者跟物理实现有关, 显然又是一个不可调和的矛盾。

(2) 机器主体是指具有意识的硬件机器人, 包括支撑框架、传感器、认知结构、动力结构、运动

结构、机械手臂等。其中, 传感器和认知结构是产生机器意识的关键。认知结构由意识结构和保持选择结构组成。意识结构是主体把物理刺激对传感器所激发的电信号进行分解整合的信号处理模块, 其输出结果就是意识信号。保持选择结构, 简称保选结构, 则是对意识信号进行保持和可选操作的信号处理模块, 旨在与记忆、行动或者语言等建立新关联。根据意识的等效原理, 意识是主体在利用保选结构对意识信号进行处理时, 由于无法区分意识信号和物理刺激而涌现出来的。在笔者看来, 认知就是在意识的基础上建立新记忆、新行动和新语言等新关联以及新经验的过程, 如达马西奥所写: “没有意识, 就没有认知”^[39]。

(3) 机器意识就是机器主体的意识, 主要由其传感器和认知结构所决定。传感器的作用是把环境的物理刺激 \mathbf{X} 转换为感觉信号 \mathbf{Y} 。认知结构包含意识结构和保选结构, 其中意识结构负责把感觉信号 \mathbf{Y} 分解整合成意识信号 \mathbf{Y}_3 ; 保选结构负责对意识信号进行保持和可选操作, 以及意识的涌现。认知结构的总目标是为了根据来自传感器的感觉信号生成对适应环境最有效的现象信息以用来引导主体下一步的思考和行动。机器主体的传感器和认知结构未必跟人类相同, 甚至可能完全不同, 这可能导致机器主体在意识中体验到与人类不同的主观世界, 从而产生不同的意识体验, 并形成不同的主观性。

(4) 根据认知基本定理 $I(\mathbf{X}; \mathbf{Y}_3) \leq I(\mathbf{X}; \mathbf{Y}) \leq C$, 机器的意识能力受限于其感觉能力且以其感觉容量为上界, 因此只要给机器增加更多、更强的传感器, 提升机器的感觉能力 $I(\mathbf{X}; \mathbf{Y})$ 和感觉容量 C , 就可以让机器的意识能力 $I(\mathbf{X}; \mathbf{Y}_3)$ 具有更高的天花板, 从而让机器的认知能力获得更大的发展空间, 至少在理论上具有超越人类认知水平的可能。

(5) 同类机器主体具有相同的传感器和认知结构, 由于观察到的主观世界相同, 就会误以为主观世界是与观察者无关的, 从而将其当作客观世界来看待。在这个意义上, 世界的客观性本质来源于主体的主观性。所以, 没有绝对的主观性, 也没有绝对的客观性。主观性和客观性都是相对的, 都是与观察者有关的, 就像物理学的空间和时间都是相对于观察者而言的, 既没有绝对的空间, 也没有绝对的时间。

(6) 机器主体也可以产生自我意识, 但应在其它意识的基础上逐步发展起来。发展心理学研究表明, 人类的婴儿在8个月前还没有萌发自我意识, 在一周岁前后才显示出主体我的意识, 约在两周岁前后才具有客体我的意识^[105]。根据笔者的分析,

机器主体产生自我意识还需要一个“自我中心核”，其作用是通过内部传感器来汇聚全身各处的体感信号。由于体感信号在先，汇聚信号在后，因此自我就是一种稍后才能形成的意识^[106-108]，其抽象层次比视、听、嗅、味、触等现象意识更高。在认知相对论的框架下，自我意识就是主体在保持自我意识信号并等待对其进行旨在与记忆、行动或语言等建立新关联的可选操作时，由于无法区分自我意识信号和“我自身”而涌现出来的，因此自我意识也可以视为关于“我自身”的意识，它指向“我自身”。

(7) 机器主体在形成自我意识之后，还可以拥有自由意志，也就是自己选择“做与不做”的最终决定权。一方面，这可能导致机器主体违反人类的命令行事，甚至有意伤害人类，或者坐视人类受到伤害于不顾；另一方面，这又是机器主体超越计算主义、摆脱预定程序限制而产生真正智能的前提。笔者认为，真正的智能是主体在有意识的状态下利用给定条件和可选操作创建一定的工作流程，比如算法、程序和行为，以便更好地适应环境的过程。简言之，真正的智能就是对算法、程序或行为等工作流程的有意识创建，而不是无意识的执行。计算机所能做的仅仅是按照事先确定的算法和程序一步一步地去执行，虽然有时看起来可以表现出一定的智能、甚至超级智能，但严格地说并不属于真正的智能。

(8) 机器主体可以采用独特的语言系统来描述和认知其在意识中所体验到的主观世界，但未必以人类的自然语言为基础。自然语言处理一直是人工智能的重要课题，人们希望在机器主体与人类之间直接进行无障碍的自然语言交流。然而，根据符号的相对性原理，机器主体完全可以采用非语音的物理形式，比如无线电波语言^[93]，来设计和实现自己的语言系统。无线电波语言是笔者在2018年国际智能科学大会上提出的一种语言设想，其基本符号由无线电波的振荡模式组成，无需解码为人类的自然语音，就可以直接被机器主体用来描述和认识主观世界，并进行判断、推理、思维和交流。理论上，无线电波语言突破了人工智能对自然语言处理的研究思路。实践上，无线电波语言也可以像人类的声波语言那样有无数种方言，还有望在未来的太空探索中产生巨大的应用前景，比如让机器宇航员在太空中直接交流(如图5所示)，而无需受到空气对语音的限制。

(9) 机器意识的实现未必需要情感的支撑。当我们醒来睁开双眼看世界的时候，意识很快就产生

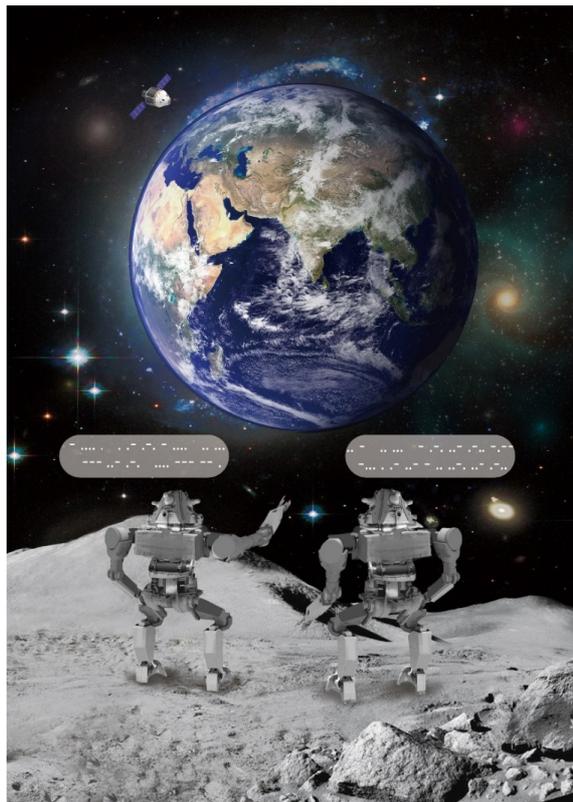


图5 两个机器人宇航员可以直接用无线电波语言来思考和交流，无需解码为人类的语音

了，这并不需要任何情感方面的理由。尽管情感对我们的认知和智能具有一定的影响，但对机器来说，这种影响有时反而应该忽略。比如，意识不到劳累的厌烦情绪，机器主体才会不知疲倦地工作。人工智能所关注的情感计算，是指“与情感相关、来源于情感或能够对情感施加影响的计算”^[109]，其本意并不是要让机器拥有人类的情感体验，而是通过计算的方法进行一定程度的模拟。人工情感也许可以让机器产生类人的表情，然而在本文看来，奥运冠军站在领奖台上的内心快乐体验，无论如何都不是在机器上可以通过计算模拟出来的情感。

8 强人工智能的展望

机器能否有意识^[110]，是一个令人神往和着迷的问题。制造出机器意识是人工智能的圣杯^[111]，也是人工智能的梦想。特斯拉说过：“我认为任何一种对人类心灵的冲击都比不过亲眼见证人造大脑变为现实。”其实，他真正想看到的就是一颗人造大脑，在机器上实现了具有人类心灵的强人工智能。

关于强人工智能，查默斯^[12]赞同Searle的最初定义，即：所有的心智，包括人类的意识体验，都可以用一个非空的计算类来定义。换句话说，就是人类的意识和心灵是可以通过在计算机上运行的程序来实现的。这实际上也意味着意识和心灵能够用

现有物理学来解释,因为计算机就是用现有物理学实现的,而且程序运行的每一步都对应一个物理过程。另一方面,查默斯又宣称意识不可还原为现有物理学,通俗地说就是不能用现有物理学来解释。由此可见,查默斯的观点是自相矛盾的。事实上,他既没有用计算给出意识的定义,也没有用现有物理学给出意识的解释。

目前,关于意识问题的探讨和争论,不只局限于学术界的关注,还广泛吸引了媒体的兴趣和公众的眼球^[112,113]。再加上各国政府从国家战略上对新一代人工智能的迫切推动,发展意识理论,制造机器意识,实现强人工智能,已经箭在弦上、势在必行。

本文认为,在计算机上运行的任何程序都实现不了人类的意识,也实现不了动物和机器的意识。如果把强人工智能定义为具有意识的人工智能,那么无论如何编程都产生不了强人工智能。在认知相对论的框架下,只有在具身的硬件机器人上,才可能实现机器意识,从而实现强人工智能。根据意识的等效原理,意识是主体在保持意识信号并等待对其进行旨在与记忆、行动或语言等建立新关联的可选操作时,由于无法区分意识信号和物理刺激而涌现出来的。根据世界的相对性原理,意识体验又极大地受限于主体对本体世界的观察方式,是不能脱离主体的传感器、认知结构和身体结构而独立存在的。其中,传感器决定着主体从环境获取信息的感知能力,认知结构决定着意识信号的形成、保持、等待、选择以及建立新关联的过程,身体结构决定着主体产生有效行为以适应环境的策略。根据符号的相对性原理,机器主体还可以直接用无线电波语言代替人类的声波语言,来描述和认知其在意识中所体验到的主观世界,并进行判断、推理、思维和交流,从而在太空探索中不再受到空气对语音的限制。当机器主体在通过传感器和认知结构观察外部世界的时候,利用意识结构把物理刺激对传感器所激发的电信号分解整合形成意识信号,再利用保选结构对其进行保持和旨在建立新关联的可选操作,机器意识就会涌现出来,引导机器主体更好地创建工作流程以应对环境变化的不确定性,从而产生强人工智能。

强人工智能有意识并不等于其功能就一定强大,但与无意识的弱人工智能相比,它可以在意识中思考和选择每一步的行为,通过不断建立事物之间的新关联来应对环境变化的不确定性,从而有望形成更强大的环境适应能力。弱人工智能是一个被动执行程序以解决在环境中所遇问题的过程,而强

人工智能是一个在意识中主动创建程序步骤以应对环境变化的过程。不过,强人工智能有意识并非意味着它一定有智能。如果想要达到甚至超越人类心智水平的强人工智能,机器主体必须具备出色的感知能力、灵巧的行动能力和精湛的语言能力,其中涉及物理、化学、材料、机械、电子、控制、仪器等各门科学与技术方面的知识,这无疑将是一项堪比阿波罗登月计划的伟大工程,也将是一次充满艰辛和困难的万里长征,远非一朝一夕可以完成。从牛顿力学到登上月球,人类花了300年左右的时间。而关于心智的认知科学远远没有物质科学成熟,甚至还没有进入牛顿时代,即使从笔者提出的认知相对论算起,实现特斯拉的“人造大脑”梦想,保守估计至少也需要200~300年。在这个过程中,认知相对论将可能起到范式革命的作用,逐步把人工智能从在计算机上对人类智能的模拟和仿真,过渡到在具身的硬件机器人上实现一系列不同水平的人造心智,就像生物从低级到高级的不断演化过程。强人工智能的心智水平,开始可能只有简单的光亮意识或者声响意识,连带一些粗糙的语言能力,在引导机器主体选择适应环境的行为方面所起的作用非常有限,不过随着时间的推移其精密和复杂程度将会像生命进化一样缓缓攀升,直至先进材料、机械制造、电子信息、自动控制和精密仪器等工程实现技术的极限。虽然无法准确预测人造心智何时能与人类心智相媲美,但是可以想象,人类对心智的理解必将突破计算的窠臼,心智的计算机隐喻时代将很快会结束,新一代心智科学的理论与技术必将越来越兴旺发达,机器意识的标志性成果必将越来越丰富多彩,强人工智能也必将得到生机勃勃的发展和蒸蒸日上的繁荣。

参考文献

- [1] 李玉鑑,张婷.深度学习导论及案例分析[M].北京:机械工业出版社,2016.
LI Yujian and ZHANG Ting. An Introduction to Deep Learning with Case Studies[M]. Beijing: China Machine Press, 2016.
- [2] 李玉鑑,张婷,单传辉,等.深度学习:卷积神经网络从入门到精通[M].北京:机械工业出版社,2018.
LI Yujian, ZHANG Ting, SHAN Chuanhui, et al. Deep Learning: Mastering Convolutional Neural Networks from Beginner[M]. Beijing: China Machine Press, 2018.
- [3] ZHANG Ziwei, CUI Peng, and ZHU Wenwu. Deep learning on graphs: A survey[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 34(1): 249-270. doi: 10.1109/TKDE.2020.2981333.
- [4] BRAUWERS G and FRASINCAR F. A general survey on

- attention mechanisms in deep learning[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2023, 35(4): 3279–3298. doi: [10.1109/TKDE.2021.3126456](https://doi.org/10.1109/TKDE.2021.3126456).
- [5] SILVER D, HUANG A, MADDISON C J, *et al.* Mastering the game of Go with deep neural networks and tree search[J]. *Nature*, 2016, 529(7587): 484–489. doi: [10.1038/nature16961](https://doi.org/10.1038/nature16961).
- [6] SCHRITTWIESER J, ANTONOGLIOU I, HUBERT T, *et al.* Mastering Atari, Go, chess and shogi by planning with a learned model[J]. *Nature*, 2020, 588(7839): 604–609. doi: [10.1038/s41586-020-03051-4](https://doi.org/10.1038/s41586-020-03051-4).
- [7] JUMPER J, EVANS R, PRITZEL A, *et al.* Highly accurate protein structure prediction with AlphaFold[J]. *Nature*, 2021, 596(7873): 583–589. doi: [10.1038/s41586-021-03819-2](https://doi.org/10.1038/s41586-021-03819-2).
- [8] 雷·库兹韦尔, 李庆诚, 董振华, 田源, 译. 奇点临近[M]. 北京: 机械工业出版社, 2011.
KURZWEIL R, LI Qingcheng, DONG Zhenhua, TIAN Yuan. translation. The Singularity is Near[M]. Beijing: China Machine Press, 2011.
- [9] KOSINSKI M. Theory of mind might have spontaneously emerged in large language models[EB/OL]. <https://arxiv.org/abs/2302.02083>, 2023.
- [10] LI Deyi, HE Wen, and GUO Yike. Why AI still doesn't have consciousness?[J]. *CAAI Transactions on Intelligence Technology*, 2021, 6(2): 175–179. doi: [10.1049/cit.2.12035](https://doi.org/10.1049/cit.2.12035).
- [11] SEARLE J R. Minds, brains, and programs[J]. *Behavioral and Brain Sciences*, 1980, 3(3): 417–424. doi: [10.1017/S0140525X00005756](https://doi.org/10.1017/S0140525X00005756).
- [12] 大卫·J. 查默斯, 朱建平, 译. 有意识的心灵: 一种基础理论研究[M]. 北京: 中国人民大学出版社, 2013.
CHALMERS D J, ZHU Jianping. translation. The Conscious Mind: In Search of A Fundamental Theory[M]. Beijing: China Renmin University Press, 2013.
- [13] SHEVLIN H. Non-human consciousness and the specificity problem: A modest theoretical proposal[J]. *Mind & Language*, 2021, 36(2): 297–314. doi: [10.1111/mila.12338](https://doi.org/10.1111/mila.12338).
- [14] NIIKAWA T, MIYAHARA K, HAMADA H T, *et al.* Functions of consciousness: Conceptual clarification[J]. *Neuroscience of Consciousness*, 2022, 2022(1): niac006. doi: [10.1093/nc/niac006](https://doi.org/10.1093/nc/niac006).
- [15] PEEBLES G. Phenomenology and the unity of consciousness[J]. *Synthese*, 2021, 199(3/4): 5455–5477. doi: [10.1007/s11229-021-03031-9](https://doi.org/10.1007/s11229-021-03031-9).
- [16] SIGNORELLI C M and MELING D. Towards new concepts for a biological neuroscience of consciousness[J]. *Cognitive Neurodynamics*, 2021, 15(5): 783–804. doi: [10.1007/s11571-020-09658-7](https://doi.org/10.1007/s11571-020-09658-7).
- [17] MONTEMAYOR C. Types of consciousness: The diversity problem[J]. *Frontiers in Systems Neuroscience*, 2021, 15: 747797. doi: [10.3389/fnsys.2021.747797](https://doi.org/10.3389/fnsys.2021.747797).
- [18] SCHURGER A and GRAZIANO M. Consciousness explained or described?[J]. *Neuroscience of Consciousness*, 2022, 2022(1): niac001. doi: [10.1093/nc/niac001](https://doi.org/10.1093/nc/niac001).
- [19] RUAN Zenan. The fundamental challenge of a future theory of consciousness[J]. *Frontiers in Psychology*, 2023, 13: 1029105. doi: [10.3389/fpsyg.2022.1029105](https://doi.org/10.3389/fpsyg.2022.1029105).
- [20] GRAZIANO M S A. A conceptual framework for consciousness[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2022, 119(18): e2116933119. doi: [10.1073/pnas.2116933119](https://doi.org/10.1073/pnas.2116933119).
- [21] JOHNSTONE C and ALEGAONKAR P S. Understanding physical processes in describing a state of consciousness: A review[EB/OL]. <https://arxiv.org/abs/2301.09576>, 2023.
- [22] 撒穆尔·伊诺克·斯通普夫, 詹姆斯·菲泽, 匡宏, 邓晓芒, 译. 西方哲学史: 从苏格拉底到萨特及其后[M]. 北京: 世界图书出版公司, 2009.
STUMPF S E, FIESER J, KUANG Hong, DENG Xiaomang. translation. A History of Philosophy: Socrates to Sartre and beyond[M]. Beijing: World Book Inc. , 2009.
- [23] 马修·科布, 张今, 译. 大脑传[M]. 北京: 中信出版集团, 2022.
COBB M, ZHANG Jin. translation. The Idea of the Brain[M]. Beijing: CITIC Press, 2022.
- [24] 车文博. 西方心理学史[M]. 杭州: 浙江教育出版社, 1998.
CHE Wenbo. A History of Western Psychology[M]. Hangzhou: Zhejiang Education Publishing House, 1998.
- [25] 梯利, 伍德, 葛力, 译. 西方哲学史[M]. 北京: 商务印书馆, 2015.
THILLY F, GE Li. translation. A History of Philosophy[M]. Beijing: The Commercial Press, 2015.
- [26] 小西奥多·希克, 刘易斯·沃克, 柴伟佳, 龚皓, 译. 做哲学: 88个思想实验中的哲学导论[M]. 北京: 北京联合出版公司, 2018.
SCHICK JR T, VAUGHN L, CHAI Weijia, GONG Hao. translation. Doing Philosophy: An Introduction Through Thought Experiments[M]. Beijing: Beijing United Publishing Co. , Ltd. , 2018.
- [27] 迈克尔·加扎尼加, 罗路, 译. 意识本能[M]. 杭州: 浙江教育出版社, 2022.
GAZZANIGA M, LUO Lu. translation. The Consciousness Instinct[M]. Hangzhou: Zhejiang Education Publishing House, 2022.
- [28] The Metaphysics Research Lab. Panpsychism[EB/OL]. <https://plato.stanford.edu/entries/panpsychism/>, 2022.
- [29] 威廉·冯特, 李维, 译. 人类与动物心理学讲义[M]. 北京: 北京大学出版社, 2013.
WUNDT W, LI Wei. translation. Lectures on Human and Animal Psychology[M]. Beijing: Peking University Press, 2013.

- [30] 詹姆斯, 唐钺译. 心理学原理[M]. 北京: 北京大学出版社, 2013.
JAMES W, TANG Yue. translation. The Principles of Psychology[M]. Beijing: Peking University Press, 2013.
- [31] 埃里克·坎德尔, 喻柏雅, 译. 追寻记忆的痕迹: 新心智科学的开创历程[M]. 北京: 中国友谊出版公司, 2019.
KANDEL E R, YU Baiya. translation. In Search of Memory: The Emergence of A New Science of Mind[M]. Beijing: China Friendship Publishing Company, 2019.
- [32] C. 麦金, 吴杨义, 译. 意识问题[M]. 北京: 商务印书馆, 2015.
MCGINN C, WU Yangyi. translation. Problem of Consciousness[M]. Beijing: The Commercial Press, 2015.
- [33] 罗素, 贾可春, 译. 哲学问题[M]. 北京: 商务印书馆, 2019.
RUSSELL B, JIA Kechun. translation. The Problems of Philosophy[M]. Beijing: The Commercial Press, 2019.
- [34] 吉尔伯特·赖尔, 徐大建, 译. 心的概念[M]. 北京: 商务印书馆, 1992.
RYLE G, XU Dajian. translation. The Concept of Mind[M]. Beijing: The Commercial Press, 1992.
- [35] 约翰·B. 华生, 潘威, 郭本禹, 译. 行为主义[M]. 北京: 商务印书馆, 2019.
WATSON J B, PAN Wei, GUO Benyu. translation. Behaviorism[M]. Beijing: The Commercial Press, 2019.
- [36] 史蒂芬·平克, 欧阳明亮, 译. 语言本能: 人类语言进化的奥秘[M]. 杭州: 浙江人民出版社, 2015.
PINKER S, OUYANG Mingliang. translation. The Language Instinct: How the Mind Creates Language[M]. Hangzhou: Zhejiang People's Publishing House, 2015.
- [37] GOLDSTEIN E B, 张明, 译. 认知心理学: 心智、研究与生活[M]. 北京: 中国轻工业出版社, 2020.
GOLDSTEIN E B, ZHANG Ming. translation. Cognitive Psychology: Connecting Mind, Research, and Everyday Experience[M]. Beijing: China Light Industry Press, 2020.
- [38] 刘晓力. 认知科学对当代哲学的挑战[M]. 北京: 科学出版社, 2020.
LIU Xiaoli. The Challenge of Cognitive Science to Contemporary Philosophy[M]. Beijing: Science Press, 2020.
- [39] 安东尼奥·达马西奥, 孙强, 译. 感受与认知: 让意识照亮心智[M]. 上海: 上海科技教育出版社, 2022.
DAMASIO A, SUN Qiang. translation. Feeling & Knowing: Making Minds Conscious[M]. Shanghai: Shanghai Science and Technology Education Press, 2022.
- [40] NAGEL T. What is it like to be a bat?[J]. *The Philosophical Review*, 1974, 83(4): 435–450. doi: [10.2307/2183914](https://doi.org/10.2307/2183914).
- [41] DENNETT D C. Welcome to strong illusionism[J]. *Journal of Consciousness Study*, 2019, 26(9/10): 48–58.
- [42] 弗朗西斯·克里克, 汪云九, 齐翔林, 吴新年, 等译. 惊人的假说[M]. 长沙: 湖南科学技术出版社, 1998.
CRICK F, WANG Yunjiu, QI Xianglin, WU Xinnian, et al. translation. The Astonishing Hypothesis[M]. Changsha: Hunan Science and Technology Press, 1998.
- [43] CRICK F and KOCH C. Toward a neurobiological theory of consciousness[J]. *Seminars in the Neurosciences*, 1990, 2: 263–275.
- [44] 克里斯托夫·科赫, 顾凡及, 侯晓迪, 译. 意识探秘[M]. 上海: 上海科学技术出版社, 2012.
KOCH C, GU Fanji, HOU Xiaodi. translation. Quest for Consciousness[M]. Shanghai: Shanghai Scientific & Technical Publishers, 2012.
- [45] SCHLOSSMACHER I, DELLERT T, BRUCHMANN M, et al. Dissociating neural correlates of consciousness and task relevance during auditory processing[J]. *NeuroImage*, 2021, 228: 117712. doi: [10.1016/j.neuroimage.2020.117712](https://doi.org/10.1016/j.neuroimage.2020.117712).
- [46] DEMBSKI C, KOCH C, and PITTS M. Perceptual awareness negativity: A physiological correlate of sensory consciousness[J]. *Trends in Cognitive Sciences*, 2021, 25(8): 660–670. doi: [10.1016/j.tics.2021.05.009](https://doi.org/10.1016/j.tics.2021.05.009).
- [47] FRIEDMAN G, TURK K W, and BUDSON A E. The current of consciousness: Neural correlates and clinical aspects[J]. *Current Neurology and Neuroscience Reports*, 2023, 23(7): 345–352. doi: [10.1007/s11910-023-01276-0](https://doi.org/10.1007/s11910-023-01276-0).
- [48] MARTIN S. Why using “consciousness” in psychotherapy? Insight, metacognition and self-consciousness[J]. *New Ideas in Psychology*, 2023, 70: 101015. doi: [10.1016/j.newideapsych.2023.101015](https://doi.org/10.1016/j.newideapsych.2023.101015).
- [49] SATTIN D, MAGNANI F G, BARTESAGHI L, et al. Theoretical models of consciousness: A scoping review[J]. *Brain Sciences*, 2021, 11(5): 535. doi: [10.3390/brainsci11050535](https://doi.org/10.3390/brainsci11050535).
- [50] SETH A K and BAYNE T. Theories of consciousness[J]. *Nature Reviews Neuroscience*, 2022, 23(7): 439–452. doi: [10.1038/s41583-022-00587-4](https://doi.org/10.1038/s41583-022-00587-4).
- [51] BAARS B J. A Cognitive Theory of Consciousness[M]. Cambridge: Cambridge University Press, 1988.
- [52] TONONI G, BOLY M, MASSIMINI M, et al. Integrated information theory: From consciousness to its physical substrate[J]. *Nature Reviews Neuroscience*, 2016, 17(7): 450–461. doi: [10.1038/nrn.2016.44](https://doi.org/10.1038/nrn.2016.44).
- [53] BROWN R, LAU H, and LEDOUX J E. Understanding the higher-order approach to consciousness[J]. *Trends in Cognitive Sciences*, 2019, 23(9): 754–768. doi: [10.1016/j.tics.2019.06.009](https://doi.org/10.1016/j.tics.2019.06.009).
- [54] CLARK A. Whatever next? Predictive brains, situated agents, and the future of cognitive science[J]. *Behavioral and Brain Sciences*, 2013, 36(3): 181–204. doi: [10.1017/S0140525X12000477](https://doi.org/10.1017/S0140525X12000477).
- [55] GRAZIANO M S A. The attention schema theory: A foundation for engineering artificial consciousness[J].

- Frontiers in Robotics and AI*, 2017, 4: 60. doi: [10.3389/frobt.2017.00060](https://doi.org/10.3389/frobt.2017.00060).
- [56] HAMEROFF S. ‘Orch OR’ is the most complete, and most easily falsifiable theory of consciousness[J]. *Cognitive Neuroscience*, 2021, 12(2): 74–76. doi: [10.1080/17588928.2020.1839037](https://doi.org/10.1080/17588928.2020.1839037).
- [57] COOKE J E. The living mirror theory of consciousness[J]. *Journal of Consciousness Studies*, 2020, 27(9/10): 127–147.
- [58] BRONFMAN Z Z, GINSBURG S, and JABLONKA E. The transition to minimal consciousness through the evolution of associative learning[J]. *Frontiers in Psychology*, 2016, 7: 1954. doi: [10.3389/fpsyg.2016.01954](https://doi.org/10.3389/fpsyg.2016.01954).
- [59] AFRASIABI M, REDINBAUGH M J, PHILLIPS J M, et al. Consciousness depends on integration between parietal cortex, striatum, and thalamus[J]. *Cell Systems*, 2021, 12(4): 363–373. e11. doi: [10.1016/j.cels.2021.02.003](https://doi.org/10.1016/j.cels.2021.02.003).
- [60] SIGNORELLI C M, WANG Quanlong, and COECKE B. Reasoning about conscious experience with axiomatic and graphical mathematics[J]. *Consciousness and Cognition*, 2021, 95: 103168. doi: [10.1016/j.concog.2021.103168](https://doi.org/10.1016/j.concog.2021.103168).
- [61] BRIDEWELL W and ISAAC A M C. Apophatic science: How computational modeling can explain consciousness[J]. *Neuroscience of Consciousness*, 2021, 2021(1): niab010. doi: [10.1093/nc/niab010](https://doi.org/10.1093/nc/niab010).
- [62] DRESP-LANGLEY B. Consciousness beyond neural fields: Expanding the possibilities of what has not yet happened[J]. *Frontiers in Psychology*, 2022, 12: 762349. doi: [10.3389/fpsyg.2021.762349](https://doi.org/10.3389/fpsyg.2021.762349).
- [63] KOCULAK M and WIERZCHOŃ M. Consciousness science needs some rest: How to use resting-state paradigm to improve theories and measures of consciousness[J]. *Frontiers in Neuroscience*, 2022, 16: 836758. doi: [10.3389/fnins.2022.836758](https://doi.org/10.3389/fnins.2022.836758).
- [64] ARKHIPOV A. Non-separability of physical systems as a foundation of consciousness[J]. *Entropy*, 2022, 24(11): 1539. doi: [10.3390/e24111539](https://doi.org/10.3390/e24111539).
- [65] KOCULAK M and WIERZCHOŃ M. How much consciousness is there in complexity?[J]. *Frontiers in Psychology*, 2022, 13: 983315. doi: [10.3389/fpsyg.2022.983315](https://doi.org/10.3389/fpsyg.2022.983315).
- [66] BUDSON A E, RICHMAN K A, and KENSINGER E A. Consciousness as a memory system[J]. *Cognitive and Behavioral Neuroscience*, 2022, 35(4): 263–297. doi: [10.1097/WNN.0000000000000319](https://doi.org/10.1097/WNN.0000000000000319).
- [67] TORDAY J S. Consciousness, embodied quantum entanglement[J]. *Progress in Biophysics and Molecular Biology*, 2023, 177: 125–128. doi: [10.1016/j.pbiomolbio.2022.11.002](https://doi.org/10.1016/j.pbiomolbio.2022.11.002).
- [68] FEDOTOV S A and BAIDYUK E V. Communication as the origin of consciousness[J]. *Integrative Psychological and Behavioral Science*, 2023, 57(1): 20–42. doi: [10.1007/s12124-022-09686-4](https://doi.org/10.1007/s12124-022-09686-4).
- [69] BACHMANN T and ARU J. Conscious interpretation: A distinct aspect for the neural markers of the contents of consciousness[J]. *Consciousness and Cognition*, 2023, 108: 103471. doi: [10.1016/j.concog.2023.103471](https://doi.org/10.1016/j.concog.2023.103471).
- [70] 李玉鑑. 揭开意识的奥秘——兼谈认知相对论纲领[M]. 李喜先. 21世纪100个交叉科学难题. 北京: 科学出版社, 2005. LI Yujian. Reveal the secrets of consciousness – also on theory of cognitive relativity[M]. LI Xixian. 100 Interdisciplinary Science Puzzles of the 21st Century. Beijing: Science Press, 2005.
- [71] 沈常宇, 金尚忠. 光学原理[M]. 2版. 北京: 清华大学出版社, 2017. SHEN Changyu and JIN Shangzhong. Principles of Optics[M]. 2nd ed. Beijing: Tsinghua University Press, 2017.
- [72] 傅祖芸, 赵建中. 信息论与编码[M]. 2版. 北京: 电子工业出版社, 2014. FU Zuyun and ZHAO Jianzhong. Information Theory and Coding[M]. 2nd ed. Beijing: Publishing House of Electronics Industry, 2014.
- [73] MELLONI L, MUDRIK L, PITTS M, et al. Making the hard problem of consciousness easier[J]. *Science*, 2021, 372(6545): 911–912. doi: [10.1126/science.abj3259](https://doi.org/10.1126/science.abj3259).
- [74] 王姝彦. 自然主义视域下的意向性问题研究[M]. 北京: 科学出版社, 2018. WANG Shuyan. A Study on Intentionality Issues from the Perspective of Naturalism[M]. Beijing: Science Press, 2018.
- [75] 梅尔文·古德尔, 大卫·米尔纳, 李恒威, 龚书, 译. 看不见的视力: 对有意识和无意识视觉的探索[M]. 杭州: 浙江大学出版社, 2017. GOODALE M A, MILNER D, LI Hengwei, GONG Shu. translation. Sight Unseen: An Exploration of Conscious and Unconscious Vision[M]. Hangzhou: Zhejiang University Press, 2017.
- [76] 韩水法. 康德物自身学说研究[M]. 北京: 商务印书馆, 2007. HAN Shuifa. A Study on Kant's Theory of Things in Themselves[M]. Beijing: The Commercial Press, 2007.
- [77] G. 布扎基, 尚春峰, 李叶菲, 李晟豪, 译. 从内向外解析大脑[M]. 北京: 科学出版社, 2022. BUZSÁKI G, SHANG Chunfeng, LI Yefei, LI Shenghao. translation. The Brain from Inside Out[M]. Beijing: Science Press, 2022.
- [78] LYU Dian, STIEGER J R, XIN C, et al. Causal evidence for the processing of bodily self in the anterior precuneus[J]. *Neuron*, 2023, 111(16): 2502–2512. e4. doi: [10.1016/j.neuron.2023.05.013](https://doi.org/10.1016/j.neuron.2023.05.013).

- [79] HUNTER P. The emergence of consciousness[J]. *EMBO Reports*, 2021, 22(7): e53199. doi: [10.15252/embr.202153199](https://doi.org/10.15252/embr.202153199).
- [80] OVERGAARD M. Insect consciousness[J]. *Frontiers in Behavioral Neuroscience*, 2021, 15: 653041. doi: [10.3389/fnbeh.2021.653041](https://doi.org/10.3389/fnbeh.2021.653041).
- [81] DUNG L. Does illusionism imply skepticism of animal consciousness?[J]. *Synthese*, 2022, 200(3): 238. doi: [10.1007/s11229-022-03710-1](https://doi.org/10.1007/s11229-022-03710-1).
- [82] CARLS-DIAMANTE S. Where is it like to be an octopus?[J]. *Frontiers in Systems Neuroscience*, 2022, 16: 840022. doi: [10.3389/fnsys.2022.840022](https://doi.org/10.3389/fnsys.2022.840022).
- [83] FRIEDMAN D A and SOVIK E. The ant colony as a test for scientific theories of consciousness[J]. *Synthese*, 2021, 198(2): 1457–1480. doi: [10.1007/s11229-019-02130-y](https://doi.org/10.1007/s11229-019-02130-y).
- [84] DUNG L. Assessing tests of animal consciousness[J]. *Consciousness and Cognition*, 2022, 105: 103410. doi: [10.1016/j.concog.2022.103410](https://doi.org/10.1016/j.concog.2022.103410).
- [85] HUNT T, ERICSON M, and SCHOOLER J. Where's my consciousness-ometer? How to test for the presence and complexity of consciousness[J]. *Perspectives on Psychological Science*, 2022, 17(4): 1150–1165. doi: [10.1177/17456916211029942](https://doi.org/10.1177/17456916211029942).
- [86] SMITH D H and SCHILLACI G. Why build a robot with artificial consciousness? How to begin? A cross-disciplinary dialogue on the design and implementation of a synthetic model of consciousness[J]. *Frontiers in Psychology*, 2021, 12: 530560. doi: [10.3389/fpsyg.2021.530560](https://doi.org/10.3389/fpsyg.2021.530560).
- [87] SINGH S, CHAUDHARY D, GUPTA A D, et al. Artificial intelligence, cognitive robotics and nature of consciousness[C]. The 3rd International Conference on Intelligent Engineering and Management, London, United Kingdom, 2022: 447–454. doi: [10.1109/ICIEM54221.2022.9853081](https://doi.org/10.1109/ICIEM54221.2022.9853081).
- [88] ZHAO Minghao, ZHANG Tianchi, and ZHANG Jing. Consciousness neural network for path tracking control of floating objects at sea[C]. International Conference on Machine Learning, Cloud Computing and Intelligent Mining, Xiamen, China, 2022: 391–397. doi: [10.1109/MLCCIM55934.2022.00073](https://doi.org/10.1109/MLCCIM55934.2022.00073).
- [89] 李夏冰. 意识问题的模型化研究进路——对当代意识模型的哲学考察[D]. [博士论文], 山西大学, 2021. doi: [10.27284/d.cnki.gsxiu.2021.002177](https://doi.org/10.27284/d.cnki.gsxiu.2021.002177).
LI Xiabing. The modeling path of the problem of consciousness: A philosophical investigation of contemporary models of consciousness[D]. [Ph. D. dissertation], Shanxi University, 2021. doi: [10.27284/d.cnki.gsxiu.2021.002177](https://doi.org/10.27284/d.cnki.gsxiu.2021.002177).
- [90] KOCH C. What is consciousness?[J]. *Nature*, 2018, 557(7704): S8–S12. doi: [10.1038/d41586-018-05097-x](https://doi.org/10.1038/d41586-018-05097-x).
- [91] 谢平. 探索大脑的终极秘密: 学习、记忆、梦和意识[M]. 北京: 科学出版社, 2018.
XIE Ping. Exploring Brain's Ultimate Mysteries: Learning, Memory, Dreams and Consciousness[M]. Beijing: Science Press, 2018.
- [92] DAN Yang and POO M M. Hebbian depression of isolated neuromuscular synapses in vitro[J]. *Science*, 1992, 256(5063): 1570–1573. doi: [10.1126/science.1317971](https://doi.org/10.1126/science.1317971).
- [93] LI Yujian. Can machines think in radio language?[C]. The Third International Conference on Intelligence Science II, Beijing, China, 2018: 230–234. doi: [10.1007/978-3-030-01313-4_24](https://doi.org/10.1007/978-3-030-01313-4_24).
- [94] 秦瑞琳, 周昌乐, 晁飞. 机器意识研究综述[J]. 自动化学报, 2021, 47(1): 18–34. doi: [10.16383/j.aas.c200043](https://doi.org/10.16383/j.aas.c200043).
QIN Ruilin, ZHOU Changle, and CHAO Fei. A survey on machine consciousness[J]. *Acta Automatica Sinica*, 2021, 47(1): 18–34. doi: [10.16383/j.aas.c200043](https://doi.org/10.16383/j.aas.c200043).
- [95] 周昌乐. 机器意识: 人工智能的终极挑战[M]. 北京: 机械工业出版社, 2021.
ZHOU Changle. Machine Consciousness: The Ultimate Challenge to Artificial Intelligence[M]. Beijing: China Machine Press, 2021.
- [96] RADOVAN M. Computation and understanding[M]. GAMS M, PAPRZYCKI M, WU X. Mind Versus Computer: Were Dreyfus and Winograd Right? Amsterdam: IOS Press, 1997: 211–223.
- [97] CAPLAIN G. Is consciousness a computational property?[J]. *Informatica*, 1995, 19: 615–619.
- [98] HAMEROFF S and PENROSE R. Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness[J]. *Mathematics and Computers in Simulation*, 1996, 40(3/4): 453–480. doi: [10.1016/0378-4754\(96\)80476-9](https://doi.org/10.1016/0378-4754(96)80476-9).
- [99] The Metaphysics Research Lab. The computational theory of mind[EB/OL]. <https://plato.stanford.edu/entries/computational-mind/>, 2020.
- [100] BLUM L and BLUM M. A theory of consciousness from a theoretical computer science perspective: Insights from the conscious Turing machine[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2022, 119(21): e2115934119. doi: [10.1073/pnas.2115934119](https://doi.org/10.1073/pnas.2115934119).
- [101] 肖琳芬. 蒲慕明院士: 脑科学与类脑智能[J]. 高科技与产业化, 2021, 27(10): 20–23.
XIAO Linfen. Academician Mu Ming Pu: Brain science and brain-like intelligence[J]. *High-Technology & Commercialization*, 2021, 27(10): 20–23.
- [102] 鲍艳伟, 任福继. 人脑信息处理和类脑智能研究进展[J]. 科技导报, 2023, 41(9): 6–16. doi: [10.3981/j.issn.1000-7857.2023](https://doi.org/10.3981/j.issn.1000-7857.2023).

- 09.001.
BAO Yanwei and REN Fuji. Advances and prospects of human brain information processing and brain-like intelligence[J]. *Science & Technology Review*, 2023, 41(9): 6–16. doi: [10.3981/j.issn.1000-7857.2023.09.001](https://doi.org/10.3981/j.issn.1000-7857.2023.09.001).
- [103] 曾毅, 张倩, 赵菲菲, 等. 从认知脑的计算模拟到类脑人工智能[J]. *人工智能*, 2022(6): 28–40. doi: [10.16453/j.2096-5036.2022.06.003](https://doi.org/10.16453/j.2096-5036.2022.06.003).
- ZENG Yi, ZHANG Qian, ZHAO Feifei, *et al.* From computational simulation of cognitive brain to brain-like artificial intelligence[J]. *Artificial Intelligence View*, 2022(6): 28–40. doi: [10.16453/j.2096-5036.2022.06.003](https://doi.org/10.16453/j.2096-5036.2022.06.003).
- [104] 张学博, 袁天蔚, 张丽雯, 等. 2022年脑科学与类脑智能发展态势[J]. *生命科学*, 2023, 35(1): 9–17. doi: [10.13376/j.cbbs/2023003](https://doi.org/10.13376/j.cbbs/2023003).
- ZHANG Xuebo, YUAN Tianwei, ZHANG Liwen, *et al.* Progress on brain science and brain-inspired intelligence in 2022[J]. *Chinese Bulletin of Life Sciences*, 2023, 35(1): 9–17. doi: [10.13376/j.cbbs/2023003](https://doi.org/10.13376/j.cbbs/2023003).
- [105] 百度百科. 自我意识[EB/OL]. https://baike.baidu.com/item/自我意识/1800719?fr=ge_ala, 2023. Baidu Baike. Self-consciousness[EB/OL]. https://baike.baidu.com/item/自我意识/1800719?fr=ge_ala, 2023.
- [106] 杜文悦. 学龄前儿童身体自我意识发展的研究[D]. [硕士学位论文], 南昌大学, 2022. doi: [10.27232/d.cnki.gnchu.2022.004384](https://doi.org/10.27232/d.cnki.gnchu.2022.004384).
- DU Wenyue. Study on the development of bodily self-consciousness in preschool children[D]. [Master dissertation], Nanchang University, 2022. doi: [10.27232/d.cnki.gnchu.2022.004384](https://doi.org/10.27232/d.cnki.gnchu.2022.004384).
- [107] 任艳玲, 蔡婧, 马岭, 等. ADHD儿童自我意识水平及其影响因素研究[J]. *中国现代医生*, 2023, 61(8): 31–34,38. doi: [10.3969/j.issn.1673-9701.2023.08.007](https://doi.org/10.3969/j.issn.1673-9701.2023.08.007).
- REN Yanling, CAI Jing, MA Ling, *et al.* Self-concept and its influencing factors in children with attention deficit hyperactivity disorder[J]. *China Modern Doctor*, 2023, 61(8): 31–34,38. doi: [10.3969/j.issn.1673-9701.2023.08.007](https://doi.org/10.3969/j.issn.1673-9701.2023.08.007).
- [108] 曾毅. 自我意识理论对精神分析的启示[J]. *心理月刊*, 2023, 18(5): 189–190, 213. doi: [10.19738/j.cnki.psy.2023.05.059](https://doi.org/10.19738/j.cnki.psy.2023.05.059).
- ZENG Yi. The enlightenment of self-consciousness theory on psychoanalysis[J]. *Psychological Monthly*, 2023, 18(5): 189–190, 213. doi: [10.19738/j.cnki.psy.2023.05.059](https://doi.org/10.19738/j.cnki.psy.2023.05.059).
- [109] PICARD R W. *Affective Computing*[M]. Cambridge: MIT Press, 1997.
- [110] DEHAENE S, LAU H, and KOUIDER S. What is consciousness, and could machines have it?[J]. *Science*, 2017, 358(6362): 486–492. doi: [10.1126/science.aan8871](https://doi.org/10.1126/science.aan8871).
- [111] 阿卡普拉沃·包米克, 王兆天, 李晔卓, 译. 机器意识: 人工智能如何为机器人装上大脑[M]. 北京: 机械工业出版社, 2021.
- BHAUMIK A, WANG Zhaotian, LI Yezhuo. translation. *From AI to Robotics: Mobile, Social, and Sentient Robots*[M]. Beijing: China Machine Press, 2021.
- [112] FLEMING S, FRITH C D, GOODALE M, *et al.* The integrated information theory of consciousness as pseudoscience[EB/OL]. <https://psyarxiv.com/zsr78>, 2023.
- [113] Ambitious theories of consciousness are not "scientific misinformation"[EB/OL]. <https://www.theintrinsicperspective.com/p/ambitious-theories-of-consciousness>, 2023.
- 李玉鑑: 男, 教授, 研究方向为人工智能、深度学习、机器学习、模式识别、认知科学等.

责任编辑: 余蓉