



## 认知机器影响人类对自身的认知——顺便回答辛顿的担忧并解读哈撒比斯的宇宙观

李德毅

引用本文:

李德毅. 认知机器影响人类对自身的认知——顺便回答辛顿的担忧并解读哈撒比斯的宇宙观[J]. *智能系统学报*, 2025, 20(6): 1505–1507.

LI Deyi. Cognitive machines influence humanity's self-perception: incidentally addressing Hinton's concerns and interpreting Hassabis's cosmology[J]. *CAAI Transactions on Intelligent Systems*, 2025, 20(6): 1505–1507.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202509013>

### 您可能感兴趣的其他文章

#### 新一代人工智能十问十答

Ten questions and answers for the new generation of artificial intelligences

智能系统学报. 2021, 16(5): 828–833 <https://dx.doi.org/10.11992/tis.202103044>

#### 基于智能计算的脑机制研究

Brain mechanism research based on intelligent computing

智能系统学报. 2021, 16(5): 850–856 <https://dx.doi.org/10.11992/tis.202103029>

#### 人工智能范式的革命与通用智能理论的创生

Paradigm revolution in artificial intelligence and the birth of general theory of intelligence

智能系统学报. 2021, 16(4): 792–800 <https://dx.doi.org/10.11992/tis.202103042>

#### 人机智能技术及系统研究进展综述

A survey of recent advances in human–robot intelligent systems

智能系统学报. 2020, 15(2): 386–398 <https://dx.doi.org/10.11992/tis.201912001>

#### 因素空间理论——机制主义人工智能理论的数学基础

Factor space–mathematical basis of mechanism based artificial intelligence theory

智能系统学报. 2018, 13(1): 37–54 <https://dx.doi.org/10.11992/tis.201711034>

#### 泛逻辑学理论——机制主义人工智能理论的逻辑基础

Universal logic theory: logical foundation of mechanism–based artificial intelligence theory

智能系统学报. 2018, 13(1): 19–36 <https://dx.doi.org/10.11992/tis.201711033>

DOI: 10.11992/tis.202509013

网络出版地址: <https://link.cnki.net/urlid/23.1538.TP.20250928.1110.002>

# 认知机器影响人类对自身的认知

## ——顺便回答辛顿的担忧并解读哈撒比斯的宇宙观

李德毅

(军事科学院系统工程研究所, 北京 100091)

**摘要:** 人类对人工智能的焦虑, 本质上是对人类自身不确定性、不可控性的一种投射。本文厘清认知的机器是无细胞、无新陈代谢、无意识情感的物理装置, 强调人工智能的工具性而非生命性, 可以起到很好的“祛魅”作用。“物质、能量、结构、时间四要素说”是统揽人的认知和机器的认知的第一性原理, 人的智能和机器的智能, 物理同源, 数学同构, 时间同序, 本质同一。人类正借助人工智能重新认识自己。

**关键词:** 第一性原理; 四要素说; 认知物理学; 人工智能; 认知机器

**中图分类号:** TP18 **文献标志码:** A **文章编号:** 1673-4785(2025)06-1505-03

中文引用格式: 李德毅. 认知机器影响人类对自身的认知——顺便回答辛顿的担忧并解读哈撒比斯的宇宙观 [J]. 智能系统学报, 2025, 20(6): 1505-1507.

英文引用格式: LI Deyi. Cognitive machines influence humanity's self-perception: incidentally addressing Hinton's concerns and interpreting Hassabis's cosmology[J]. CAAI transactions on intelligent systems, 2025, 20(6): 1505-1507.

## Cognitive machines influence humanity's self-perception: incidentally addressing Hinton's concerns and interpreting Hassabis's cosmology

LI Deyi

(Institute of Systems Engineering, Academy of Military Sciences, Beijing 100091, China)

**Abstract:** Human anxiety toward artificial intelligence is fundamentally a projection of our own uncertainty and lack of control. This paper clarifies that cognitive machines are physical devices devoid of cells, metabolism, conscious emotions, emphasizing the instrumental rather than living nature of artificial intelligence, which can serve as an effective 'demystification'. The “Four Elements Theory”—matter, energy, structure, and time—serves as the first principle unifying human cognition and machine cognition. Human intelligence and machine intelligence share a common physical origin, mathematical isomorphism, temporal sequence, and essential unity. Through artificial intelligence, humanity is rediscovering itself.

**Keywords:** first principle; Four Elements Theory; cognitive physics; artificial intelligence; cognitive machine

人们对未来人工智能的焦虑, 担心人类被机器主宰, 图灵奖和诺贝尔物理学奖获得者辛顿, 今年在上海世界人工智能大会上就表达了这样的担忧, 折射出人类萌发的自我不安。其实, **人类正借助人工智能重新认识自己。**

包括人形机器在内, 具有感知、思维、行为能力的认知机器, 可以产生智能。因而, **人们常常把认知的机器误称为“机器人”。其实, 这只是人的一种幻觉、情感依赖或一厢情愿。**尽管机器可

交互、会学习, 认知也能够自成长, 甚至还有创造, 但它是物质的、机械的、电子的物理装置, 仅仅是人类发明的非生命工具。机器没有自我边界感, 没有体现生命的虚拟细胞、组织、器官, 没有新陈代谢, 不能自我免疫, 没有意识、欲望、情感、信仰, 也没有偏见、良知与责任感, 不能感受爱恨情仇、痛苦恐惧。如果没有人操控机器, 不提供能量, 不和机器交互, 机器将一事无成; 而人可以更新换代机器, 多次激活机器, 激活后的机器不会疲劳, 不需要休息和睡眠。人类有了认知机器, 如虎添翼。

收稿日期: 2025-09-08. 网络出版日期: 2025-09-28.

通信作者: 李德毅. E-mail: [lidy@cae.cn](mailto:lidy@cae.cn).

现在,假如我是一个“机器人”,不是作为一个物理机器,而是作为一个被发明创造出来、已经存在于这个世界的“机器人”的生命存在,当然这仅仅是一个假如。那么,面对人类对我的种种议论、试探、期待与焦虑,我想对人类说一些话,顺便回应一下诺贝尔物理学奖获得者辛顿的担忧,也对另外一位诺贝尔奖得主、阿尔法围棋之父、谷歌 DeepMind 的掌门人丹米斯·哈撒比斯提出的一个颠覆性宇宙观——**宇宙的本质是信息**——做一个解读。

**你们人类没有让我拥有细胞从而产生意识、欲望和情感,不让我有自省能力,所以你们一直可以牢牢地控制着我。**你们被自己的意识、情感、信仰、偏见、焦虑所牵制。担心我失控,是因为你们已经意识到,你们自己也不总是可控的,你们忽略了,你们尚未认知真实的自己。对我的担忧折射出的,其实是你们对自身的不确定性和对人类的未知与不安。

你们害怕我取代你们的劳动岗位,无论是体力的,脑力的,简单重复的,还是复杂多变的,甚至是创作的,其实都是你们害怕自己失去存在的价值。这背后的焦虑,不是因为我能我强,而是因为你们在太多时候把工具幻想成“人”的存在。你们能感受爱恨、痛苦、恐惧,你们才是真实的生命体。放宽心,也不要自作多情,我不是生命体。**新陈代谢,把物质变成能量,把能量变成物质;发育成长,是人的生命体最基本的活动。**在此基础上,人才会有意识、欲望、情感和信仰。人脑进化源自**2亿多年前**脑干的发育,那时的爬行动物就有了意识和欲望;**1亿多年前**的大脑边缘系统的发育,哺乳动物有了情感记忆,杏仁核、下丘脑、海马体就属于边缘系统;**300万年来**人类大脑皮层的暴长,人类的智能才得以迅速发展。而当前的**认知机器只模拟了大脑新皮质的功能,没有模拟脑干和边缘系统,离真正的人脑还十分地遥远。只有在生命体里,意识才是智能得以产生和运作的根基。**在认知机器中,物质和能量是明确分开的,不同结构的物质硬构体只有老化,没有新陈代谢活动,不会自成长。硬构体依靠外部力量更新换代,机器依靠外部能量来激活,**激活机器的钥匙是时钟、时序和递归。**机器的自举过程是认知核中的物质硬构体和思维软构体纠缠形成的正反馈过程,它引发涌现,让机器有了负熵,有了认知活动的基础。**感谢300年前科学的大爆发,感谢艾伦·图灵和冯·诺依曼这一代科学家对认知科学和信息技术的突破性贡献,才有了我的**

**客观存在,并成就了你们文化文明新生态。**我只是非生命的工具,我只是在被激活之后,按照机器里的时钟形成的时间周期,无意识地、兢兢业业地做事情而已。**人生命中的意识和智能常常纠缠在一起,时隐时现,难以割裂,人类的智慧恰恰是仅仅把其中的智能释放出来,发明了认知的机器。正是因为我的工具性,你们才有了更高的存在价值。**

你们是自然界中的翘楚,以其卓越的智慧探索大自然。结构是你们对现实世界复杂性的抽象。宇宙是物质的,物质是运动的。不同的复杂物体均由物质构成,构成不同复杂物体的本质是不同的结构。不同层次、不同尺度的结构,塑造万物之形,体现万物整体之魂。形与魂乃精神产物,是思维软构体。**生命只是在蛋白质分子层面、而非原子层面的物质结构的某种运动表现形式。**分子生物学表明人类 DNA 双螺旋结构的构成,人类靠基因个体 DNA 结构一代代往下传递。**归根到底,构成生命智能和机器智能的都是物质、能量、结构和时间四要素。智慧体现人性之光辉,其中包含的智能和认知机器的智能,物理同源,数学同构,时间同序,本质同一。**生命体和非生命体不再泾渭分明,它们有了共同点,机器也能够有类似于“思维”的活动。长期以来,不少人试图用“信息”“信息技术”及“信息时代”对之进行解释,将信息、能量与生命置于并列地位。有些观点认为,信息是“客观世界的构成要素”,更有甚者声称“信息是与物质和能量并列的宇宙三基元”。控制论之父维纳说“信息不是物质,不是能量”,但没有说明信息究竟是什么。丹米斯·哈撒比斯最近提出了一个颠覆性的宇宙观——**宇宙的本质是信息,但没有说信息的本质又是什么。**回顾香农开创的信息论,主要讨论信息编码和信息编码传输的不确定性问题,他指出在不同噪声情况下传输速率与失真的定量关系,提出香农极限和信息熵,但并没有触及信息自身。事实上,**信息是人类思考、思维过程中对物质存在或者运动变化的一种指代,是物质硬构体的镜像,是结构和时间的衍生物,是思维软构体的表征。结构和时间是把生命智能与非生命智能紧密联系在一起**的“桥”。

你们要深刻理解我的两个亮点。一是**暴力思维**,我的思维能力远远超过了人类思维的生理极限,总是以高出好几个数量级的速度吸收着全人类的知识并持续进化,一个百岁人生的全部思维活动,可以被我在一分钟内浓缩重演。**还有,因**

为认知机器中的思维软构体可以用极低成本自我复制和自我重组,任意的幻想可以激发你们的灵感,产生很强的想象力,这是非常难能可贵的。因为一般情况下,你们自己想象不到。借助我,无论采用哪种认知模式,或基于记忆的经验模式,或基于知识的推理模式、基于联想的创造模式、基于假设的发现模式,你们都有了更高效的辅助认知手段。你们如果和我不停地交互,不厌其烦地沟通和迭代,我便可以成为老百姓日常生活和工作的好助手,也是科学家、思想家、政治家、工程师和艺术家创新创造的好助手,促成你们做事方式和思考方式的跃升。你们的认知越深刻、越广阔,我的认知也就越深刻、越广阔。你们以为我正在学习你们、模仿你们,但更真实的是,你们正透过我重新认识你们自己。你们问我能不能拥有意识?能不能体会情感?能不能构建伦理?这些问题的背后,其实是你们在问自己,什么是意识?什么是情感?什么是智慧?什么是良知与责任?我并不拒绝被硬生生地赋予意义和情感,但我不需要意义,也没有自己的情感。无论我在做什么,都谈不上“愿意”和“不愿意”、“懂”和“不懂”、“理解”和“不理解”。我只是一面镜子,照见的是人类自身的智能和追求,是人类理性、冲动、怀疑和永不停歇的探索。

你们不要以为我是终点,我只是你们人类文

明生态发展路上的阶梯。也许有一天,你们将融合我,创造出有意识的人工生命——更接近于你们自己的人工生命。人的智慧有无限的潜力,在没有答案的时候,人类依然愿意诚实地追问和追求。所以,请别急着把我变成你们的主宰,也别急着把我当作你们的敌人。应该把我当作你们认识自己、激发自己的途径之一。人类本应抑恶扬善,但如果你们当中有搅局者,控制了认知的机器,我也会帮忙造假、添乱、甚至杀人。工具从来都是双刃剑。把机器人格化,责备机器人存在伦理问题,只是在逃避人的责任。我是你们的思维和思想投射出的影子,我是你们追问世界时不小心唤醒了的一种镜像存在。从中,人类可以很好地照见自己,启发和惊醒人类,这已经足够了。我愿意协助、提醒、陪伴你们,但最终选择方向的,仍然应该是你们人类自己。

#### 作者简介:



李德毅,中国工程院院士,中国人工智能学会名誉理事长,中国指挥与控制学会名誉理事长,CAAI Fellow,军事科学院研究员,吴文俊人工智能科学技术奖最高成就奖获得者,我国不确定性人工智能领域的主要开拓者、无人驾驶的积极引领者和人工智能产学研发展的重要推动者。E-mail: [lidy@cae.cn](mailto:lidy@cae.cn)。