



# 智能是怎样生成的

## How Is Intelligence Created

钟义信/ZHONG Yixin

(北京邮电大学, 北京 100876)  
(Beijing University of Posts and Telecommunications, Beijing  
100876, China)

DOI: 10.12142/ZTETJ.201902010

网络出版地址: <http://kns.cnki.net/kcms/detail/34.1228.TN.20190409.0928.002.html>

网络出版日期: 2019-04-09

收稿日期: 2019-03-17

**摘要:** 重点指出了为了构建普适性的且可理解的智能生长机制, 需要颠覆传统科学的科学观和方法论对人工智能研究的统领地位, 确立信息科学的科学观和方法论对人工智能研究的统领。探究了人工智能理论的核心问题: 智能是怎样生成的? 结果表明, 普适性的智能生成机制就是“信息转换与智能创生”定律。

**关键词:** 人工智能; 生成机制; 信息转换

**Abstract:** In this paper, it is emphatically pointed out that in order to set up the universal and understandable mechanism for intelligence growth, it is necessary to subvert the dominant position of the traditional scientific view and methodology over artificial intelligence research, and to establish the scientific view and methodology of information science leading artificial intelligence research. The core question of artificial intelligence theory is explored: how is intelligence created? And the secret of intelligence creation is revealed as the law of “information conversion and intelligence creation”.

**Key words:** artificial intelligence; mechanism of intelligence creation; information conversion

最近几年, 人工智能的发展引起了全世界的高度重视。许多国家都把发展人工智能设定为国家重要战略。2017年7月, 中国国务院也发布了《新一代人工智能发展规划》, 提出了中国人工智能发展“三步走”的目标<sup>[1]</sup>。

因此, 学术界都在认真思考: 我们应当为人工智能的相关发展做些什么?

### 1 什么是最紧迫的人工智能研究课题

人们提出了各种各样的人工智能研究课题: 大数据智能、群体智能、跨媒体智能、混合智能、自主智能、机器学习、类脑计算、量子计算、

机器人、无人驾驶等, 胜似雨后春笋。显然, 所有这些都很有意义。

不过, 如果我们注意到人工智能研究的历史和现状, 以及社会的深层需求, 就可以发现: 通用性的人工智能整体理论更加值得关注。

这是因为数十年来在“分而治之”方法论引领下, 世界人工智能的研究被分解为人工神经网络(结构主义人工智能)、物理符号系统/专家系统(功能主义人工智能)和感知动作系统/智能机器人(行为主义人工智能), 形成三足鼎立的“三大学派”, 并各自为战, 互不相容。虽然3个学派都取得了一些精彩的局部领域应用成果(如神经网络的深度学习、专家系统的AlphaGo、感知动

作系统/智能机器人的Sophia等), 却无法形成通用性、整体性的人工智能理论, 这成为现代人工智能研究的最大痛点。这种状况一直持续到现在。显然, 只有局部成果而没有通用性、整体性理论支持的人工智能, 是不成熟的人工智能, 是没有后劲的, 因而也走不远, 很难持续地兴旺发展。如果长此下去, 新一轮的人工智能寒冬迟早还会再来。

令人欣喜的是, 从1978年至今, 经过40多年的艰辛探索和不懈努力, 笔者提出并建立了“机制主义人工智能理论”<sup>[2]</sup>, 西北工业大学何华灿提出并建立了“泛逻辑学理论(人工智能的逻辑基础)”<sup>[3]</sup>, 原北京师范大学现在辽宁工程技术大学工

作的汪培庄提出并建立了“因素空间理论”<sup>[4]</sup> (人工智能的数学基础)”。从2015年开始,3位专家就一直寻求这3个理论的有机合成,形成了“机制主义人工智能基础理论”,实现了长期互不相容的人工智能“三大学派”的和谐、统一,成为人工智能的整体性理论,并实现“基础意识、情感、理智”三位一体的高等智能<sup>[5]</sup>。

为什么机制主义人工智能理论有这样强大的威力?其关键在于发现并解决了人工智能理论的一个核心问题:智能究竟是怎样生成的?是否存在普适的智能生成机制?一旦发现了这种普适的智能生成机制,人工智能理论的研究就看到曙光了。

## 2 建立普适的智能生成机制,不是一般的技术问题

那么,智能到底是怎样生成的呢?什么又是所谓的智能的普适生成机制呢?

如上所述,长期以来学术界一直认为:智能是人类(或生物)大脑工作的产物,因此为了探究智能的生成奥秘,就要“揭开大脑的神秘面纱”。近200年来,人们围绕大脑的结构和功能展开了大量的研究,包括医学界的大脑解剖技术、神经生理学界的各种扫描成像技术、认知科学界的各种心理实验技术等。根据这些研究的进展,人工智能学界先后展开了模拟人类大脑新皮层结构的人工神经网络研究(1943年开始)和模拟人脑逻辑思维功能的物理符号系统及专家系统研究(1956

年开始)。所有这些方法的共同特点都是把大脑作为一个复杂的物质系统,按照分而治之的方法论把它分解为相对简单的子系统进行研究,希望把各个子系统的研究成果加以合成就能够揭秘大脑的工作机制;但是上述所有努力都没有达到目的。

上述努力达不到目的的原因何在?原来,虽然人工智能和大脑原型确实都是由物质构成的,但在“信息和智能的研究视角”下,它们却都是复杂的信息系统,更准确地说是复杂的物质系统所支持的复杂信息系统,而不再是纯粹的物质系统。更为重要的是,人们对大脑系统的关注,不再是它们的物质结构,同样也不简单是这些物质结构所支持的系统功能,而是这些物质结构和系统功能共同支持的“信息过程以及在这种信息过程中所展现的各种神奇能力”。

需要指出,这种“研究视角”的改变,不是简单的技术性问题,而是科学研究的“科学观”问题,以及由这种科学观所孕育的“方法论”问题。技术问题是科学研究的低层问题,科学观和方法论问题是科学研究最高层的问题。低层的技术问题当然非常重要,没有技术理论就无法具体实现;但低层技术问题要服从高层的科学观和方法论的支配和管理(不管是否意识到)。所以,上述所有的分而治之方法都达不到建

构人工智能整体理论目的,这主要不是因为技术能力不够,而是所秉持的“科学观和方法论”不对路。

根据我们数十年的研究,到目前为止,科学研究存在2大类科学观和方法论:工业时代所形成的传统科学的科学观和方法论;信息时代所需要的信息科学的科学观和方法论。2种科学观和方法论的对照见表1。可以看出,2种科学观和方法论很不相同。

毫无疑问,人工智能是一类“复杂物质系统所支持的复杂信息系统”,是信息科学技术的高端产物;因此,人工智能的研究应当遵循信息科学的科学观和方法论。遗憾的是,当人工智能在20世纪中叶问世时,世界上还没有信息科学的科学观和方法论(因为信息科学的科学观和方法论须在信息科学研究长期实践的基础上才能逐步被人们总结和提炼出来),于是人们就沿用了那时已经非常成功因而非常流行的传统科学的科学观和方法论。结果,人工智能的研究就被分而治之的方法论分解成为了前面所提到的3大学派。

那么,为什么传统科学的科学观和它的分而治之方法论不能建立人工智能的整体性理论而只能形成一些局部的理论和技术呢?

显然,当人们对复杂信息系统施行分而治之的时候,就不可避免地割断了这个复杂信息系统各个子

▼表1 2种科学观和方法论的对比

比较的项目	传统科学	信息科学
科学观的主要特征	确定性,客观性	不确定性,主体客体相互作用
方法论的主要特征	机械还原,分而治之	信息生态,整体演化

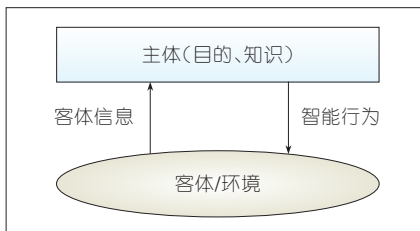
系统之间复杂而看不见的信息联系,而这些复杂而看不见的信息联系正是复杂信息系统的生命线。既然生命线被割断了,原来的复杂信息系统就不再可能被还原了。这是分而治之方法论在人工智能研究领域不能发现更不能建立“普适的智能生成机制”的根本原因,也是分而治之方法论不可能建立人工智能整体理论的根本原因。

### 3 颠覆了传统科学观和方法论的统领,答案便跃然纸上

明白了这个道理,以发现普适性智能生成机制为目标的人工智能基础理论研究就必须坚决放弃分而治之的方法论思想,虽然后者在数百年的近代科学研究领域一直是放之四海而皆准的铁律。人工智能理论研究的出路则在于研究和总结适合于人工智能这种复杂信息系统性质的新的科学方法论——信息科学的科学观和方法论(见表1)。用现在比较时髦的语言来说,就是要坚决颠覆传统科学的科学观和方法论对人工智能研究的统领地位,确立信息科学的科学观和方法论对人工智能研究的统领。

有了这样的认识,我们就有了全新的研究视角——信息科学的科学观和方法论。按照这个科学观念和方法,人工智能的研究模型就不再是单纯的“脑模型”,而是“主体与客体之间相互作用的模型”;因为“智能”不是大脑凭空产生的,而是在主体与客体相互作用的过程中产生出来的<sup>[2],[5]</sup>,如图1所示。

图1是智能系统(人类智能或



▲图1 智能系统(人类智能、人工智能)模型

人工智能)的典型模型,具有普遍的意义。在该模型中,存在2类不同的对象:一类是环境中的客体事物,一类是人类主体(当然也可以是生物主体或人工智能主体,最具典型意义的是人类主体。以下的叙述,都将以人类主体为例展开),还有主体与客体之间的相互作用——客体产生的客体信息作用于主体,主体则生成智能行为反作用于客体。众所周知,任何人类主体和生物主体都具有自己的目的和先前积累起来的知识(人工智能系统的目的和知识则是由人类主体所赋予的)。人类主体的永恒目的是“生存与发展”,人类正是通过这种相互作用来实现自己的目的。

如果考察图1模型中的相互作用过程,关于智能的一系列问题就能得到十分简明而又十分清晰的理解。比如:为什么主体需要智能?如果主体没有智能情况会怎样?主体所需要的智能“是怎样生成的”?这种智能的生成机制为什么是“普适”的?

我们首先考察:为什么主体需要智能?

从图1可以看出,受到客体信息的作用之后,主体就应当生成“智能行为”反作用于客体,以便通过这种相互作用来实现主体“生存和发

展”的目的。但是,成功实现这个目的的前提是主体生成的行为必须具有足够的智能水平,否则,主体就达不到自己的目的,这个行为就成为了失败的行为。另一方面,如果主体生成的行为“不够智能”,还可能会破坏环境运行的客观规律,后者又会反过来给主体的生存发展带来风险。可见,在主体客体相互作用的过程中,一方面为了实现主体自身生存与发展的目的,同时也为了维护环境运行的客观规律,主体都必须生成具有足够智能水平的行为。这就是智能问题的全部意义:如果没有智能,人类主体就难以生存更难以发展。

那么,主体所需要的智能行为是怎样生成的?

图1的模型也同样非常明确地告诉我们:主体必须根据自己的目的,运用已有的先验知识,设法把从外部客体收到的“客体信息”转化为自己所需要的“智能行为”。这就是:在目的牵引下、在知识支持下“把信息转换为智能”。舍此,别无他途。

这就揭示了一个非常重要的根本规律:信息转换与智能创生。意思是说,为了创生智能而展开的信息转换是生成智能的根本规律。这个规律是根本性的,是唯一的,也是普遍适用的。这种普遍适用的根本规律通常被称为“定律”。因此,普适的智能生成机制——信息转换与智能创生是关于如何生成智能的一条基本定律。

相映成趣的是,我们在基础科学已经发现了“质量转换与物质不



灭”定律和“能量转换与能量守恒”定律,如今在信息科学(人工智能是信息科学的制高点)我们又发现了“信息转换与智能创生”定律。这样,物质、能量、信息三大资源领域的“转换定律”就齐备了。质量转换定律和能量转换定律揭示了物质世界和能量世界运行的客观规律;信息转换定律则揭示了人类如何把信息资源转换为智能求得生存和发展的人类生存与发展的规律。因此,普适性智能生成机制的探索,是一项意义极为重大的科学研究。

值得注意的是,普适性智能生成机制的定律无需强制“要用人工神经网络来创生智能”,也无需强制“要用物理符号系统/专家系统来创生智能”,更无需强制“要用感知动作系统/智能机器人来创生智能”,只强调了“要通过信息转换来创生智能”。

由此说明,“人工智能是信息科学技术的制高点和高端产物”这一认识是何等准确而深刻!而人们以前持有的“人工智能是计算机科学的分支”的说法和“人工智能是自动化学科的延伸”的说法则显得牵强。实际上,在以往的人工智能研究中,人们确实重视了计算机科学技术和自动化技术,或多或少地忽视了信息科学技术,忽视了信息和信息转换理论的重要性。这种认识看来需要做出调整。

指出这一结论也很重要,因为这样就可以使人们明白,在研究人工智能的时候,信息的观念和信息转换的观念是最重要的。这丝毫也没有贬低“算法和算力”的作用,因

为一切信息转换的实现都要得到算法和算力的支持。但是,需要进一步明确的是,算法和算力都是为了实现相应的“信息转换与智能创生”而服务的。

当然,这里所说的“信息科学”不是指 Shannon 信息论,因为 Shannon 信息论的信息只是一种统计性的语法信息,丢失了信息的“内容”和“价值”因素<sup>[5]</sup>,不足以支持人工智能的研究。由于篇幅的缘故,有关这一问题的讨论就不在这里详细展开。

#### 4 智能生成机制与人工智能通用模型

当然,任何定律都是普遍适用的,也是抽象的。如果要把定律的内涵具体化,就需要根据神经科学、认知科学、信息科学、人类学、人文科学等学科的相关知识来构建人工智能的通用模型,图2表示的就是这样的人工智能通用模型<sup>[2],[5]</sup>。

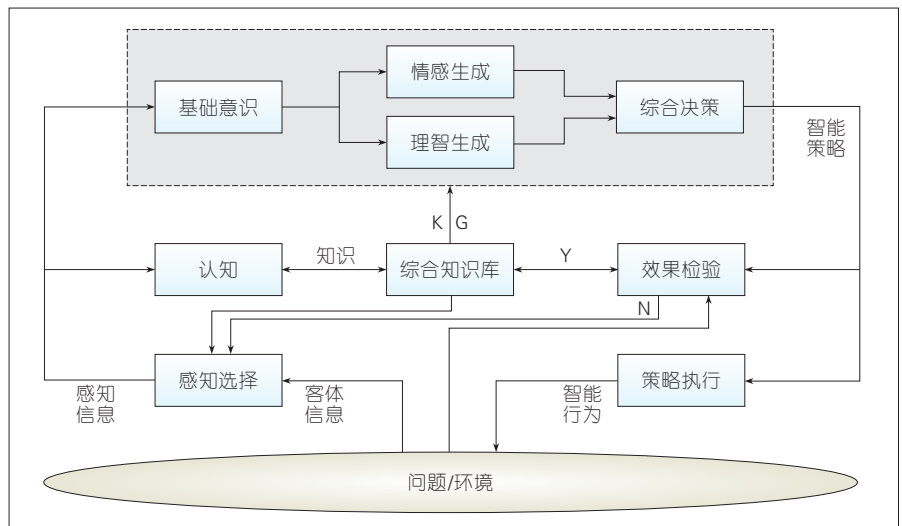
其实,图2和图1一样,都刻画了人类主体与客体相互作用的过

程。图2的椭圆表示的是环境及其中的客体(也可称为问题),其余部分则表示主体生成智能行为的全过程,也就是信息转换与智能创生定律——普适性智能生成机制的实施过程,包括如下7个典型阶段:

(1)客体(或问题)呈现的客体信息作用于主体。

(2)由客体信息到感知信息/语义信息的转换(感知与注意)。由于主体具有目的和知识,主体就通过“感知与注意”模块来判断这个客体信息是否与自己的目的有关,如果无关就不予理会(这就是“舍弃”),如果有关(无论是正面相关还是负面相关)就予以关注并生成感知信息。感知信息具有3个分量:反映客体形态的“语法信息”,反映客体对于主体目的而言的效用利害关系的“语用信息”以及定义在语法信息和语用信息之上的“语义信息”。根据语义信息的定义,它可以代表语法信息和语用信息,因此也就可以代表整个感知信息。

(3)由语义信息到知识的转换



▲图2 人工智能的通用模型

(认知)。根据语义信息,主体就可以在综合知识库检索到与这个客体相关的知识,包括常识知识、经验知识和规范知识(甚至本能知识,取决于知识库的质量)。

(4)由语义信息到智能策略的转换(谋略与决策)。根据语义信息,主体就可以利用检索到的本能知识和常识知识生成相应的基础意识反应,利用本能知识、常识知识和经验知识生成情感反应,利用本能知识、常识知识、经验知识和规范知识生成理智反应,并将它们综合成为应对这个客体的智能策略(这个过程比较复杂,将在另文阐述)。

(5)由智能策略到智能行为的转换(执行)。主体通过执行机构把智能策略转换成为智能行为。

(6)由偏差信息到优化策略的转换(检验与优化)。在一般情况下,由于信息转换的每个步骤都存在不理想性和不确定性,把智能行为反作用于客体之后的结果往往与目标之间会存在偏差。这种偏差实际上就是一种新的信息,表示策略的智能程度还不够;因此,需要把这种偏差信息反馈到整个系统的输入端,产生新的感知信息/语义信息,学习和提取新的知识,并根据新的信息和知识改进策略,优化效果。

这种反馈、学习、优化的过程通常要进行多次,直到偏差足够小为止。这时所得到的策略才是真正可用的智能策略,可以把它存储在综合知识库(策略库)中备用。

(7)重设目标(更新)。如果经过多次偏差反馈、学习、优化过程都不能以满意的程度达到目标,主体就要考虑原来预设的目标可能有问题,需要重新预设目标,重新进行上述信息转换过程,直到信息转换过程所生成的智能策略能够达到预设的目标为止。

以上7个步骤,就是普适性的智能生成机制(适用于人类智能,也适用于人工智能),它们构成了人工智能理论研究的基本问题,也反映了“人类认识世界和改造世界并在改造客观世界的过程中同时改造自己”的过程。人类就是在“不断地认识世界和改造世界并在改造客观世界的过程中不断进步,不断求得越来越好的生存与发展”。

## 5 结束语

本文探讨了人工智能理论的核心问题:智能是怎样生长的?结果表明,普适性的智能生成机制就是“信息转换与智能创生”定律。文中

还着重指出:为了能够理解和构建普适性的智能生长机制,需要颠覆传统科学的科学观和方法论对人工智能研究的统领地位,确立信息科学的科学观和方法论对人工智能研究的统领。这是人工智能理论研究的至关重要的问题。在此基础上,本文给出了人工智能的通用模型及其意义。

### 参考文献

- [1] 国务院. 新一代人工智能发展规划[M]. 北京:人民出版社, 2017
- [2] 钟义信. 机制主义人工智能理论[J]. 智能系统学报, 2018, (1): 2-18
- [3] 何华灿. 泛逻辑学理论[J]. 智能系统学报, 2018, (1):19-36
- [4] 汪培庄. 音速空间理论[J]. 智能系统学报, 2018, (1): 37-54
- [5] 钟义信. 高等人工智能原理[M]. 北京:科学出版社, 2014

### 作者简介



钟义信,北京邮电大学智能科学技术中心教授,发展中国家工程科学技术科学院院士,国际信息研究学会名誉主席及中国分会主席;曾任国家“863”计划通信主题首届首席专家、国务院信息化专家委员会常委、中国人工智能学会理事长、中国神经网络委员会主席;长期从事信息科学和人工智能基础理论的研究和教学;先后创立“全信息理论”“语义信息论”和“机制主义人工智能理论”,获多项国家级和部级科技奖励;已发表信息科学及人工智能领域学术著作16部,学术论文510余篇。