

钟义信 著

高等人工智能原理

— 观念 · 方法 · 模型 · 理论



科学出版社

014032849

TP18
258

高等人工智能原理

——观念·方法·模型·理论

钟义信 著



科学出版社

北京



北航

C1721021

TP18
258

内 容 简 介

本书旨在创建和阐明高等人工智能的基本理论。全书内容由 5 篇 10 章构成。第一篇总论:高等人工智能研究的科学观与方法论,介绍自然智能理论研究的启迪和人工智能研究方法的变革。第二篇高等人工智能的基础理论,介绍全信息理论和知识理论。第三篇高等人工智能的主体理论,介绍感知、注意与记忆(第一类信息转换原理)和意识、情感、理智与行为(第二类信息转换原理)。第四篇高等人工智能与现行人工智能的关系,介绍物理符号系统:规范知识支持的机制主义方法,人工神经网络:经验知识支持的机制主义方法,感知-动作系统:常识知识支持的机制主义方法。第五篇应用问题择要,介绍有关应用的几个共性课题。

本书适合信息领域的研究人员、大学教师、在读博士和硕士研究生及大学高年级本科学生,以及广大高新技术企业研究人员等参考阅读。

图书在版编目(CIP)数据

高等人工智能原理:观念·方法·模型·理论/钟义信著. —北京:
科学出版社, 2014. 3

ISBN 978-7-03-040011-6

I. ①高… II. ①钟… III. ①人工智能 IV. ①TP18

中国版本图书馆 CIP 数据核字(2014)第 041711 号

责任编辑:牛宇锋 / 责任校对:刘亚琦

责任印制:张 倩 / 封面设计:蓝正设计

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京通州皇家印刷厂 印刷

科学出版社发行 各地新华书店经销

*

2014 年 3 月第 一 版 开本:720×1000 1/16

2014 年 3 月第一次印刷 印张:28 1/4

字数:538 000

定价:128.00 元

(如有印装质量问题,我社负责调换)



中国科学院科学出版基金资助出版

本书的研究和写作先后得到国家自然科学基金
68872014,69171023,69982001,60496327,60575034,
60873001,以及 70711120412(国际合作)等项目的资助。

它们是孕育《高等人工智能原理》的学术温床。

作者谨致崇高的敬意和衷心的感谢!

上述基金项目系列研究成果以
“构建信息科学理论基础,创新人工智能核心理论”为题
于 2012 年获得中国人工智能学会“吴文俊人工智能科学技术奖”的
最高奖:成就奖

谨以此书献给国家、母校、师长、同学、朋友、家人
以及热爱信息与智能科学的人们!

前言

当您拿起这本《高等人工智能原理》的时候,您(特别是比较熟悉人工智能相关领域的读者)很可能会提出这样的问题:为什么作者要在现在提出“高等人工智能”的概念和阐述“高等人工智能”的原理呢?

这正是作者希望在前言交代的基本问题,这些问题包括:什么是智能?什么是人工智能?人们应当怎样研究智能和人工智能科学技术?什么是高等人工智能?它与现行人工智能在科学观念、研究方法、基本概念、系统模型、工作原理上有什么不同?为什么本书要定名为“高等人工智能原理”?本书的基本出发点和主要目标是什么?与迄今国内外出版的各种人工智能学术著作相比,本书与众不同的学术观点、学术思想、学术特色和可能的学术贡献是什么?

(一) 智能和人工智能含义的再理解

宇宙的起源、生命的本质、智能的奥秘,是在现代科学技术地形图上耸立着的三座高峰。攀上这些高峰,将可能更新人们对世界的认识。由此,我们可以体会智能科学技术研究的复杂与艰难,同时也可以领略智能科学技术研究的价值与意义。

智能是研究智能科学技术,特别是研究“高等人工智能理论”必须首先准确把握的基础概念。相信人人都会同意:如果在基础概念的理解上差之毫厘,那么,研究的结果就很有可能失之千里。这就是作者开宗明义——应当对智能和人工智能含义进行“再理解”的原因。

那么,什么是智能?

《现代汉语词典》对智能的解释非常简明:智慧和能力。

《韦氏大词典》的解释是“Capacity for understanding and for other forms of adaptive behavior”(理解能力以及各种适应行为的能力)。

《牛津词典》的解释则是“Power of seeing, learning, understanding, and knowing”(观察、学习、理解和认识的能力)。

显然,这些工具书对智能的释义都有道理,而且通俗易懂。然而,辞书是为广大非专业读者服务的,因此,不可能满足科学研究的需要。但是,智能概念是整个智能理论的基础,如果这么基础的概念都没有深刻定义,那么整个智能理论的研究就可能存在比较大的局限。为了科学研究的需要,这里给出两个不同层次而又互相密切关联的智能概念(也称为定义)。

人类智能

总地来说,人类智能就是人类认识世界和优化世界的能力。

具体地说,人类智能就是人类为了追求不断改善生存发展条件这一永恒目的而凭借已有的先验知识去发现和定义所处环境中需要解决而且可能解决的问题,并预设求解的目标;在获得问题和目标这些信息的基础上,提取必要的专门知识,进而在目标引导下运用信息和专门知识来制订求解策略,并把策略转化成行为,从而解决问题实现目标的能力;在解决老问题之后又发现新问题和解决新问题的能力。

需要注意,目的是指永恒的战略追求,目标是指求解具体问题的具体追求,后者是前者的局部体现。先验知识是先前已经具备的综合知识。求解问题所必要的专门知识是指直接与问题和目标相关的具体知识,以及在先验知识不足以支持问题求解的情况下通过学习所获取的新知识。它们是先验知识的子集。

人工智能

人工智能是人造机器所拥有的智能。

具体来说,人工智能是指针对人类设计者给定的问题、领域知识、目标等初始信息,机器提取求解问题所需要的专门知识,进而在目标引导下运用信息和专门知识制订求解策略,并把策略转化成行为,从而解决问题实现目标的能力。

同样需要注意的是,一般求解问题所需要的专门知识可以从领域知识中提取,但在领域知识不足以支持问题求解的情况下,就需要通过学习来提取新的知识。

不难看出,这里给出的两个智能概念与各词典所作的解释之间不存在矛盾,因为上述人类智能和人工智能的概念显然都需要观察、学习、理解、认识和适应能力的支持。然而这两个概念却更为明确、更为深入,也更具有科学研究的可操作性:不仅阐明了两类智能概念的内涵和相互关系,而且揭示了生成智能的科学途径,使它们不仅可望而且可及。这是科学研究特别期待的。

概念表明,一方面,人类智能和人工智能两者是相通的,表现在两者都要通过获取信息、提取知识和在目标引导下制订策略解决问题。另一方面,它们之间又存在原则的差别:人类具有自身的目的和一定的先验知识,因而在面对具体环境的时候能够自主发现和定义需要解决的问题并预设求解问题的目标,而且,人类可以通过获取信息和提取知识创造性地制订策略解决问题;而机器则没有自身固有的目的和知识,因此机器本身不能自主地发现和定义问题及预设求解目标,它的问题、知识和目标都是人类设计者事先给定的,机器只能在人类给定的框架内解决问题;而且,由于人类所给定的“问题-领域知识-目标”不一定充分合理,机器解决问题的“创造性”也会受到局限。

应当指出,“面对具体环境,根据永恒目的和先验知识发现和定义问题,并预设求解目标”的能力是人类创造力的首要前提,没有这个前提就谈不上人类智能。这

是人类智能特有而机器所没有的能力,而且通常在人类的思维过程中完成,因此,可以称为隐性智能;相应地,可以把“获取信息、提取知识、制订策略和解决问题”的能力称为显性智能。可见,人类智能同时包含隐性智能和显性智能以及它们之间的交互;人工智能只有显性智能。

由于隐性智能比显性智能更加复杂,目前还没有获得系统性的研究成果,而且,由于本书基本上定位于人工智能的研究,这里将主要关注显性智能的问题,而把隐性智能作为未来进一步研究的课题。

和其他学科类似,智能科学技术的研究领域也由相辅相成的两个方面构成。一方面是智能的原型研究,目的在于揭示和阐明智能的生成机理。智能的原型就是自然智能,即生物智能,包括人类的智能、动物的智能和植物的智能,但最受关注的是人类智能。因为人是万物之灵,拥有至高无上的智能。原型研究主要是指脑神经科学和认知科学研究,也与医学、生物学、人类学、社会学、环境学、哲学、语言学等学科密切相关。它们构成了智能科学技术的基础理论研究。另一方面是将原型智能转化为机器智能的研究,目的是在理解原型智能生成机理的基础上,在机器系统中尽可能地复现自然智能,制造具有一定智能水平的机器。这种机器系统的智能就称为“人工智能”或者“机器智能”。这是智能科学技术的应用基础研究和应用研究。

人工智能(artificial intelligence, AI)这个术语是 1956 年 McCarthy 在美国马萨诸塞州 Dartmouth 夏季研讨会期间为了表述“利用计算机模拟人类逻辑思维能力”这个当时的新生学科而创造的,它只关注和代表“计算机模拟人的逻辑思维能力”这种特殊的机器智能,而不代表全部机器智能。换言之, AI 只代表关于逻辑思维智能的模拟,而不代表关于形象思维和创造性思维这类智能的模拟。例如,它不代表“人工神经网络系统”,也不代表“感知-动作系统”。可见, AI 的实际含义只是其字面含义的一个特殊部分。

这样,就产生了 AI 的字面含义和实际含义不一致的矛盾,并在历史上产生过许多本不必要的矛盾。其中一个突出的矛盾事例就是 AI 拒绝接纳人工神经网络的研究,致使后者不得不与模糊逻辑研究和进化算法研究联合另立门户,称为“计算智能”(computational intelligence, CI),造成国际智能科学技术研究学术界的“欠和谐”。

为了解决历史上遗留下来的这些矛盾,也为了有利于今后智能科学技术研究与交流的和谐发展,我们有必要重新明确 AI 的含义,使它的实际含义和字面含义保持一致,即字面的 AI 和实际的 AI 都表达“全部人工智能”。因此,在本书以下的行文中,“人工智能”(AI)就完全等同于“机器智能”(machine intelligence, MA),把原来的“利用计算机模拟人类逻辑思维能力”的学科则称为“狭义的 AI”。

(二) 智能科学的重大意义

根据人类智能和人工智能的定义,可以作出如下判断。

智能是一切生物物种能力的最高体现:一种生物拥有的智能水平越高,它在生物物种谱系中所处的地位就越高。于是,智能科学技术是生命科学技术的制高点。同时,智能是人们对信息资源进行深度加工所能获得的最高级产物,因此,智能科学技术是信息科学技术的核心、前沿和制高点。

由此也可以说,智能科学技术是 21 世纪信息科学技术和生命科学技术这两大带头科学技术所共生的而且是最精彩的交叉科学技术。

思想指挥行动。对于一个人,智能水平越高,他取得成就的可能程度就越高;对于一个国家,智能科学技术水平越高,它的国民经济、社会文明、国计民生和国家安全的发展能力也就可能越强,在国际竞争中制胜的机会就越大。这是智能科学技术在整个科学技术体系和经济社会发展进程中的重大意义所在,也是写作本书的根本动力。

由智能的概念可以知道,人类的智能活动(显性智能)至少需要以下功能支持,而且缺一不可:信息获取(由感觉器官系统承担)、信息传递(由传导神经系统承担)、信息处理(由初级皮层承担)、知识生成(由高级皮层承担)、策略制订(由联合皮层和前额叶组织承担)及策略执行(由效应器官承担)。不难理解,“知识生成”和“策略制订”是整个智能活动过程的核心,是“智能”的主要承担者和体现者,因而可以称为“核心智能”。

在信息科学技术的体系中,信息获取功能的承担者称为“传感系统”,信息传递功能的承担者称为“通信系统”,信息处理功能的承担者称为“计算系统”,知识生成和策略制订功能的承担者称为“人工智能系统”,策略执行功能的承担者则称为“控制系统”。其中,传感系统和控制系统是智能系统与外部世界之间的两端接口(输入环节和输出环节),通信系统是智能系统与这两端接口之间的联络中介,计算系统是智能系统的预处理,人工智能系统则是智能系统的核心。

人类智力能力的进化规律和信息科学技术的发展规律都是由简单走向复杂、由表层走向核心。具体地说,人类智力能力的进化是沿着由两端(感觉器官和执行器官)走向中介(传导神经系统)再至大脑前端(初级皮层),最后到达大脑新皮层而臻于成熟。同样,信息科学技术领域的发展也由两端(传感系统和控制系统)走向中介(通信系统)再至前端(计算系统),最后走向核心(人工智能系统)才能趋于完善。

第二次世界大战结束后的半个多世纪,传感、控制、通信、计算等领域突飞猛进,取得了长足的进步,信息科学技术的发展呈现出“万事齐备,只待智能”的态势。

传感、控制、通信、计算等科学技术的进步为处于核心前沿和制高点的人工智能科学技术的发展准备了必要的基础,同时也为它们自身的智能化及整个科学技术和经济社会的智能化孕育了强大的社会需求。事实上,“智能化”已经成为当今世界社会各行各业普遍而强烈的共同呼唤。这就是如今“智能 ABC”,……,“智能 XYZ”的呼声响彻环球大地的原因!

智能科学技术已经成为当代科学技术关注的焦点,智能科学技术登上科学技术核心舞台的时机已然到来。

(三) 智能科学技术的研究方法

智能科学技术本身所固有的基础性和深刻性,智能科学技术研究工作所特有的复杂性和前沿性,智能科学技术应用所具有的普遍性和广泛性,以及智能科学技术应用可能给社会发展带来的革命性和转型性,这一切都决定了智能科学技术研究(包括自然智能的研究和人工智能的研究)的重要性,绝不可以等闲视之。

从科学研究的纵深角度看,由于智能科学技术研究的对象实质是信息(而不是人们所熟悉的物质和能量),是以信息为主导因素的开放复杂系统(而不是人们所熟悉的封闭系统和简单系统),是通过新颖的生成机制所演化的奇妙智能(而不是人们所熟悉的普通信息处理机制和普通的信息能力),深深地触及了学术研究的科学观和方法论问题,而且都不是传统科学观和方法论所能解释清楚的问题。因此,不应当期望智能科学技术的研究在传统的科学观和方法论框架内就能求得满意的结果,而必须从根本的科学观和方法论的考察做起。

换言之,智能科学技术的研究需要新的科学观和新的方法论。

从科学研究的横断角度看,智能科学技术的研究内容既涉及信息科学技术(包括信息理论、知识理论、智能理论、决策理论,以及传感技术、通信技术、存储技术、计算技术和控制技术等),又涉及生命科学技术(神经生理科学,特别是脑神经科学),还涉及人类学、社会学、心理学、思维科学、认知科学、哲学等众多学科,自然天成地形成了一个以智能科学为核心的学科群,因此,智能科学技术的研究不应当期望局限在某些个别学科领域内就可以解决问题,而应当在整个学科群的综合视野内探寻智能生成的根本规律。

或者说,智能科学技术的研究应当是学科群的研究。

以下就从“科学观与方法论的深度”和“自然天成学科群的广度”两个角度考察现有的人工智能研究的现状,探求高等人工智能研究的生长规律。

(四) 人工智能研究面临的严峻挑战

智能系统是一类以信息为主导特征的开放性复杂信息系统,它与以物质和能量为主导特征的物质系统有着迥然不同的性质和运动规律。然而,人们所熟悉的传统科学观和方法论基本上是在近代物质科学研究过程中逐渐形成和发展起来的,它们已经不能完全适应智能科学技术研究的需要。人工智能的发展亟须新的科学观和方法论。

具体来说,近代形成和逐渐发展起来并曾在近几百年科学技术发展中屡试不爽而且发挥了巨大积极作用的“分而治之,各个击破”和“系统结构决定论(认为系统的结构是系统能力的决定性因素)”、“系统功能主导论(认为系统的功能是系统能力的主导因素)”及“系统行为表现论(认为系统的行为是系统能力的主要表现方式)”等传统的科学观和方法论已经不能完全适应智能科学技术研究的需要,因而在相当程度上制约了智能科学技术的发展,使智能科学技术经历了一段复杂曲折的探索过程。

简要回顾半个多世纪以来人工智能发展的历程,其中各种成功的经验和不成功的教训都发人深省,并且都可以追溯到科学观方法论的根源。

面对自然智能和人工智能这类复杂的研究对象,古代流行的笼而统之的物质观和囫圇吞枣的整体论研究方法肯定不能解决问题,因而人们普遍认为,必须运用研究物质系统时所采用的“分而治之,各个击破”的近代科学研究方法论来处理。问题是:应当怎样“分”?

首先,按照“分而治之”和“系统结构决定论”的传统思想,人们相信,既然人类的智力功能定位在大脑皮层,后者是一个大规模的生物神经网络,那么,只要能够把大脑皮层的结构模拟出来,就等于模拟了人类大脑智能系统的能力。这样,就催生了“人工神经网络”的研究方向。1943年以来,人们在这个方向上取得了不少令人鼓舞的进展,如模式识别、故障诊断、组合问题优化等。但是,由于工业技术条件的限制,人工神经网络只能是大脑神经网络结构大大简化的模拟,人工神经网络的结构复杂性与大脑神经网络的结构复杂性相差悬殊;更重要的是,人工神经网络的工作机制(学习算法)与大脑神经网络的工作机制相去甚远,因此,人工神经网络研究的前景颇受挑战和质疑。

于是,按照“分而治之”和“功能主导论”的思维观念,人们转而利用已经成长起来的电子计算机系统模拟当时人们最为关注的人类大脑逻辑思维功能。1956年以来,人们在这个方向上也取得了许多令人振奋的成果,如击败国际象棋世界冠军卡斯帕罗夫的深蓝(Deeper Blue)专家系统等。但是,模拟人类的逻辑思维功能需要大量而系统的知识做基础,而知识的获取、表达和推理都存在巨大的困难,于是,

逻辑思维功能模拟的方法陷入了“知识瓶颈”的困境。而且,功能模拟方法所研究的智能系统不考虑基础意识功能和情感表达功能,也是这个研究方法的重要缺陷。

在这种情势下,“分而治之”和“行为表现论”的观念和方法应运而生,人们撇开智能系统的结构和功能,把研究注意力转向模拟智能系统的“刺激-响应”行为表现,认为只要把智能系统在某种环境下的“刺激-响应”行为关系模拟出来了,就等于模拟了这个系统的智能。麻省理工学院(MIT)人工智能实验室推出的爬行机器人就是这种行为模拟方法的代表性成果。但是,行为模拟方法很难模拟深层的智能,这是它的先天局限。

虽然目标都是同为人工智能的研究,但是由于研究的方法思路各不相同,上述人工智能研究的结构模拟方法、功能模拟方法、行为模拟方法相继发展起来之后,很少相互沟通与合作,一直未能形成合力。相反,它们之间究竟“孰优孰劣”的争论却时有发生,终于形成互不认可、鼎足而立、各自为战的研究格局,在一定程度上延缓了人工智能研究的发展。

既然社会对智能科学技术已经涌现出强烈的需求,而目前人工智能的研究现状又处于三足鼎立的分离状态,那么,智能科学技术工作者就自然地肩负着一项神圣的使命:在已有发展成果的基础上寻求人工智能研究的新方法,把人工智能研究的“分力”转化为“合力”,促进人工智能的新发展。

有鉴于此,国内外不少有识之士注意到了这个矛盾和解决这个矛盾的社会需求,并主动担负起了“化分力为合力”的任务。其中,最具代表性的工作是人工智能权威学者 Nilsson 在 1998 年出版的 *Artificial Intelligence: A New Synthesis*, 以及作为人工智能后起之秀的 Russell 和 Norvig 在 1995 年出版后又在 2003 年和 2006 年再版的长篇巨著 *Artificial Intelligence: A Modern Approach*。这两部学术著作分别宣称是人工智能的“新集成”和“新途径”,而且都不约而同地试图以 Agent 能力扩展为线索,把现有结构模拟、功能模拟、行为模拟三种人工智能研究成果拼装成 Agent 系统上。

不过,这种“拼装”只是表面的黏合和堆积,并没有揭示结构模拟、功能模拟和行为模拟三种主流方法内在的本质联系。因此,寻求“化分力为合力”的任务远远没有完成。

(五) 我们的发现和进展

我们的研究发现,面对人工智能这类开放的复杂信息系统,发端于物质系统的“分而治之”方法论存在的主要问题是:把复杂信息系统分解为若干子系统的时候,丢失了各个子系统之间相互联系相互作用的信息,而这些信息恰恰是开放复杂信息系统的活的灵魂和生命线;因此,按照“分而治之”方法论分别完成各个子系统的

研究之后,却怎么也不可能通过对这些子系统的“拼装”恢复原有复杂信息系统的面貌和性质。这便明确宣告传统的“分而治之”和机械的“还原论”方法论在开放的复杂信息系统研究领域的失效!

传统方法论失效的事实既令人沮丧,又令人振奋。这是因为,只有丢掉对于传统科学观和方法论的幻想,才会促使我们不得不义无反顾并全力以赴地探寻新的科学观和探索新的方法论。结果发现,“信息观、系统观、机制观的三位一体”才是人工智能这类开放复杂信息系统所需要的科学观,“基于信息转换的机制主义方法论”才是开放的复杂信息系统所需要的方法论。

首先,关于“信息观”。“智能”系统(包括自然智能系统和人工智能系统)是以信息为主导特征的复杂系统,没有信息便不可能有智能,信息不同就可能导致系统的智能水平的不同。因此,信息的因素和信息转换的规律就成为智能系统的核心灵魂和整体命脉,而智能系统所涉及的物质因素和能量因素则只能看做信息运动这个核心过程的支持性条件。于是,在研究和探索智能系统奥秘的时候,首先应当抓住信息和信息转换的本质特征,而不应把物质结构和能量形式作为本质的特征。这是研究智能这类开放复杂信息系统与研究物质系统和能量系统的根本区别。

其次,关于“系统观”。对于智能系统这类开放复杂的信息系统,必须保持信息内涵的完整性和信息转换过程的系统性,这种性质是正确认识这类复杂信息系统的基本前提和根本保证。在这里,信息内涵的完整性就是指信息的形式(语法信息)、内容(语义信息)和价值(语用信息)的三位一体(全信息);而信息转换的系统性就是指信息转换在时空领域的整体规律。如果这种信息内涵的完整性和信息转换过程的整体性被“分而治之”方法所不恰当地肢解和割断,甚至被丢失,就势必使信息系统的本性遭到破坏,并可能导致智能能力的丧失。

最后,关于“机制观”。对于任何智能系统,“智能生成的机制”才是贯穿整个智能系统全局的灵魂和生命线;至于智能系统的“结构”和“功能”则都是为实现和保障智能生成的机制这个核心灵魂和生命线而服务的;系统的“行为”则是智能生成机制所实现的结果表现。因此,“智能生成的机制”是比“系统的结构、功能和行为”更具本质意义的特征。抓住了智能生成的机制,才算真正抓住了智能系统的核心本质。

按照开放复杂信息系统的“信息观、系统观、机制观”的理念,智能系统(包括自然智能系统和人工智能系统)智能(其实也包括基础意识、情感等智能要素)生成的核心机制便是以全信息为基础的信息转换。于是,水到渠成,“信息转换”便成为智能系统这类开放复杂信息系统的基本方法论。

事实上,按照“以信息观、系统观、机制观的三位一体为科学观,以基于(全)信息转换的机制主义为方法论”,我们已经发现,智能生成的共性核心机制就是信息-知识-智能转换,简称为“信息转换”;也发现并阐明了与此相关的“全信息理论”和

“知识生态学结构”，因而建立了与结构模拟、功能模拟和行为模拟方法在学术理念上全然不同的“智能生成的机制模拟方法”；进而还发现了“人工智能的结构模拟、功能模拟、行为模拟方法分别是机制模拟方法在不同知识条件下的和谐特例”，从而形成了“人工智能统一的方法和理论”。不仅如此，我们还欣喜地发现，“基础意识、情感、智能的生成机制其实都是在各自条件和各自目标下的信息转换”。这是意义重大的发现！

综合以上这些发现和进展，一个以信息观-系统观-机制观的三位一体为科学观念，以基于(全)信息转换的机制主义为方法理念的“高等人工智能理论”便浮出了水面，与现有的人工智能理论研究相比，高等人工智能理论研究的理念、方法、深度和广度全然面貌一新。

由此发现，按照传统“分而治之”方法论形成的结构模拟、功能模拟、行为模拟三种方法之间之所以不能互相认可，不能形成合力，根本原因就在于它们各自依据的“科学观和方法论”存在严重的局限性和片面性，掩盖和割断了事物内在的信息联系，因此难以真正触及智能问题的深层本质，也难以互相沟通，而一旦运用新的更为先进的科学观和方法论阐明了智能的深层本质，它们之间内在的本质联系便会豁然显现！而且，更加深刻和更加科学的理论体系也便由此矗立在我们的面前。

(六)《高等人工智能原理》的特色

当我们利用上述“复杂信息系统的科学观和方法论”重新审视人工智能理论的时候，就可以清晰地看到：目前人们研究的“人工智能”不仅在研究方法上处于“鼎足三分，各自为战”的状态，形成了少有沟通甚至互不认可的三个人工智能理论学派，而且在研究内涵上也都局限于“既没有意识又没有情感”的不完整的智能系统模型。一句话：方法上不沟通，模型上不完整，究其根源则是观念上不合理！

另一方面，“复杂信息系统的科学观和方法论”也已明确昭示：只要抓住“智能生成机制”这个统领全局的系统灵魂和智能系统“信息转换”这个贯通全局的基本原理，不仅结构主义、功能主义和行为主义方法可以统一于机制主义方法，信息理论、知识理论和智能理论可以互相和谐沟通，而且智能系统的感知、注意、记忆、意识、情感、理智、决策理论就都可以被统一的方法贯通描述，呈现出“一通百通”的景象！

于是，我们看到，通过一部学术专著来系统总结这一系列崭新成果以便引起更多研究人员关注，深化扩展这些成果的时机显然已经成熟。这就是本书写作的具体动机。

本书曾经定名为“人工智能统一理论”，而最终定名为“高等人工智能原理”。这是因为，本书的成果不仅在于用机制主义方法和信息转换原理统一了现有的结

构主义、功能主义、行为主义三足鼎立的人工智能理论,更重要的是,与迄今出版的所有人工智能学术著作相比,本书所研究的高等人工智能理论无论在智能系统基本模型、智能系统研究的科学观念、智能系统研究的方法理念,还是智能理论研究成果的深度和广度诸多重要方面都有了实质性的突破与提升,使得原有“人工智能”的概念框架已经难以包容这些崭新的内容。

高等人工智能和现有人工智能之间的基本联系和主要差别如下表所示。

比较项目	现有人工智能	高等人工智能
科学观念	信息观	信息观、系统观、机制观的三位一体
学术理念	形式化的 Shannon 信息概念	形式、内容、价值三位一体的“全信息”概念
科学方法	结构、功能、行为;孤立模拟	机制模拟;结构、功能、行为的统一模拟
学术途径	信息处理技术	信息转换原理
系统模型	逻辑思维(无意识无情感)的智能	意识、情感、理智三位一体的智能

可以看出,“高等人工智能原理”的“高等”主要体现在:新的科学观念,新的学术理念,新的科学方法,新的学术途径,新的智能系统模型,新的智能科学原理。这就使得本书很难用传统的“人工智能”概括,而“高等人工智能”则比较贴切地体现了上述这些新颖特色。

另一方面,本书将明确定位于原理的探索和阐述,而不在于具体的技术细节。因为,对于任何学科,揭示和阐明学科的基本原理都是首要的和根本的任务,而技术实现则可以在基本稳定的基本原理基础上不断深入并且与时俱进。这也是本书的一个重要考虑。况且,作为“高等人工智能理论”领域的初创性著作,它的主要目标和任务也只能是阐明总体的框架和基本的学术原理,而它的许多技术实现方面的重要问题需要今后进一步深入研究。

作者明白,由于“高等人工智能理论”本身的深刻性、复杂性、创新性和困难性,想要通过一部学术专著一劳永逸是绝对不可能的事情。本书的意义主要在于也仅仅在于正式明确地提出了“高等人工智能原理”的科学观念、方法理念、基本原理和研究纲要,而更加深刻、更加辉煌和更有意义的研究必然有待未来。因此,本书只能算是高等人工智能理论研究的一个开篇和引论。

本书的写作得到了一系列国家自然科学基金项目的资助,得到了众多国内外学术师长的指引和教诲,使作者有机会能与大量国内外学术同行特别是中国人工智能学会和北京邮电大学智能科学技术研究中心各位同仁进行友好探讨和深入切磋,同时得到了北京邮电大学校内外各届不计其数先后参与 Unified Theory of Artificial Intelligence 课程的博士生和硕士生的热烈反馈,得到了科学出版社各位领导与编辑的热心支持与帮助,得到了家人无微不至的关爱与亲切的鼓励。假如

没有这一切,本书的写作和问世肯定都不可能。在此,作者谨对以上所有朋友表示由衷的感谢和诚挚的敬意!

在本书的写作过程中,虽然作者处处尽心尽力,希望能够做得尽善尽美,但是,由于高等人工智能理论本身的博大精深且作者学术水平有限,书中难免存在不足之处。在此,作者竭诚欢迎并恳切希望广大读者不吝批评和指正,共同为高等人工智能理论的发展作出积极的努力,为我国和整个人类社会的智能化作出有益的贡献。

作 者

2013 年秋

于北京

目 录

前言

第一篇 总论:高等人工智能研究的科学观与方法论

第 1 章 自然智能理论研究的启迪	3
1.1 脑神经科学研究简介	3
1.1.1 人类大脑与智能系统	3
1.1.2 脑的组织学	5
1.1.3 脑组织的细胞学	8
1.2 认知科学研究简介	14
1.2.1 感知	15
1.2.2 注意	16
1.2.3 记忆	18
1.2.4 思维	21
1.2.5 语言	23
1.2.6 情绪	24
1.3 脑科学与认知科学的融通:“全信息”科学观	25
1.3.1 脑神经科学与认知科学:存在“理论的断裂”	25
1.3.2 认知科学研究:需要“全信息”,也能生成“全信息”	29
1.4 小结与评注	36
参考文献	37
第 2 章 人工智能研究方法的变革	39
2.1 人工智能研究鸟瞰	39
2.1.1 人工智能的基本概念	39
2.1.2 “人工智能”含义的辨析	45
2.1.3 人工智能研究的历史与现状	47
2.2 科学研究方法的进化	55
2.2.1 科学方法论的进化	55
2.2.2 科学方法论演进概要	56
2.3 概念与方法的重审:开放复杂信息系统的科学方法论	61
2.3.1 人工智能研究遭遇的科学方法论问题	62