

独创性声明

学位论文题目： 人工智能发展中的哲学问题思考

本人提交的学位论文是在导师指导下进行的研究工作及取得的研究成果。论文中引用他人已经发表或出版过的研究成果，文中已加了特别标注。对本研究及学位论文撰写曾做出贡献的老师、朋友、同仁在文中作了明确说明并表示衷心感谢。

学位论文作者： 

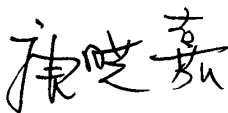
签字日期： 2013 年 4 月 7 日

学位论文版权使用授权书

本学位论文作者完全了解西南大学有关保留、使用学位论文的规定，有权保留并向国家有关部门或机构送交论文的复印件和磁盘，允许论文被查阅和借阅。本人授权西南大学研究生院（筹）可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

（保密的学位论文在解密后适用本授权书，本论文：☒不保密，☐保密期限至 年 月止）。

学位论文作者签名： 

导师签名： 

签字日期： 2013 年 4 月 7 日

签字日期： 2013 年 4 月 7 日



目 录

摘 要	I
Abstract	III
第 1 章 绪论	1
1.1 目前人工智能的研究情况	1
1.2 研究意义	4
1.3 基础研究思路与方法	4
第 2 章 人工智能的产生	6
2.1 智能的起源	6
2.2 智能的本质	6
2.3 人工智能的产生	7
第 3 章 人工智能的基本内容	10
3.1 人工智能的派别	11
3.1.1 符号主义	11
3.1.2 连接主义	12
3.1.3 行为主义	12
3.2 人工智能发展的分歧	13
3.2.1 弱人工智能	13
3.2.2 强人工智能	14
3.2.3 弱人工智能与强人工智能的关系	15
3.3 人工智能发展的技术	16
第 4 章 人工智能发展中的哲学思考	18
4.1 哲学对人工智能发展的影响	18
4.1.1 灵魂说	18
4.1.2 “人是机器”	19
4.1.3 信息哲学	20
4.2 人工智能与人类智能的关系	20
4.2.1 模拟人类智能	20
4.2.2 延伸人类智能	21
4.2.3 人类智能的组成部分	22

第 5 章 人工智能的发展方向	24
5.1 自然语言的理解	24
5.1.1 什么是自然语言	24
5.1.2 自然语言逻辑	24
5.2 不确定性推理	27
5.2.1 不完全归纳推理	27
5.2.2 模糊逻辑	29
第 6 章 结语	32
参考文献	33
致 谢	35

人工智能发展中的哲学问题思考

逻辑学专业高师硕士研究生 程石

指导教师 唐晓嘉 教授

摘 要

从 1956 年的达特矛斯会议上约翰·麦卡锡提出人工智能的概念至今,短短 50 多年,人工智能的发展速度令世人惊服。当进入 21 世纪,人工智能发展到现在,其成果和应用已经遍及人类日常生活、工业生产、科学研究等方方面面。小到手机中的智能语音搜索软件,大至工业生产线上智能机器人;简单如能够早已进入千家万户的全自动洗衣机,复杂如平均每秒运算超过 1000 万次的超级计算机,他们在各个方面极大的提升了人类认知和改变世界的能力,所表现出来的强大功能在某些方面已经将发明他们的人类远远抛与脑后,成为人类社会不可或缺的一部分,不断地改变和推动着人类社会进程。

虽然人工智能发展势头迅猛,但是在研究的道路上也遇到一些困难,如不加以解决,势必影响其今后的研究进展。为了更好的发展人工智能,使其更好的为人类服务,就必须从本质上认识人工智能,分析其发展的道路,发现其前进道路上的问题,找到解决这些问题的方法。本文综合分析前人的研究成果,从哲学的角度,运用辩证唯物主义和系统分析方法,从什么是智能入手,分析智能的起源和本质,认识人工智能是如何产生的;在整理人工智能发展的现状,分别从人工智能的派别,弱人工智能和强人工智能,人工智能发展的技术等方面进行剖析后,笔者认为目前人工智能发展所遇到的瓶颈:如何让机器拥有人类的认知、学习、推理和修正等能够体现人类智能的创造性思维——比如根据人类时时刻刻在根据自然条件主动修正、调整自己行为的能力,或在所获信息不够充分的情况下采取的某种尝试性的选择,亦或在所获信息向冲突时如何对信息进行取舍从而采取相应的行动等。

笔者认为,人工智能发展到如今要有本质上的突破,从哲学方面来看,今后的发展将会将重点放在以下两个方面:(1)如何让机器能够使用人类自然语言与人类进行沟通和交流,能够准确识别不同语境下的自然语言。而自然语言中天然存在着各种歧义。人工智能发展至今已经能够解决分析单个词语的意义,而在识别自然语言中的长句及其所蕴含的真正意义上仍有不足。要解决这一点,需要对人类自然语言进行提炼,并进行逻辑分析,这种分析不能只停留在自然语言的表面,而是要真正理解自然语言所要表达的内涵。20 世纪的哲学也发展出观念论、指称论、使用论、言语行为理论、真值条件论等各种意义理论,本文力图通过自然语言中的交际性研究自然语言逻辑中的核心——语用推理来分析人工智能在自然语言理解方面需要解决的问题;(2)如何让机器拥有人类智能最精华的部分——创造性,能够进行不

确定性推理。人类智能的本质特征是其创造性,这也是人类智能的最高表现形式。人类要进行创造性的活动,在进行必然性的演绎推理是必不可少的,而更重要的是具有某种不确定性的归纳、类比推理以及模糊推理等。不完全归纳推理作为人类认识世界、改造世界重要的方法论已经在人类的实践中取得了成功,如何让机器也能进行这类推理必将是今后研究的重点之一。另一方面,世界上存在着无数的模糊现象,它们在人脑思维中的反映成为各种的模糊概念和命题,人工智能要实现人类智能的功能,就必须能够处理各类模糊概念和命题,模糊逻辑为此奠定了基础,而且基于模糊逻辑的研究也已取得了重要的成果,但算法规则的提取和隶属函数的选择却十分费劲,这将是今后人工智能研究的热点。

关键词: 哲学、人类智能、人工智能

Philosophical Thinking in the Development of Artificial Intelligence

Postgraduate master of logic Shi CHENG

Supervisor Professor Xiaojia TANG

Abstract

From John McCarthy proposed the concept of artificial intelligence the on Dartmouth Conference in 1956, within short 50 years, the paces of development of the artificial intelligence shocked the world. In the 21st century, the achievements of artificial intelligence and the application has been throughout our daily life, all aspects of industrial production and scientific research. Small as voice search software in smart phone, large as intelligent robots on industrial production line ; simple as automatic washing machine in millions of households, complex as the supercomputer which could make over 10 million operations average per second, artificial intelligence greatly enhance human cognition and the ability to change the world in all aspects. Powerful artificial intelligence manifested far more than human in some ways, and become an integral part of the human society, continue to change and promote the process of development of human society.

Although development of artificial intelligence is fast, but research of artificial intelligence also encountered some difficulties witch is bound to affect its future progress if not be resolved. In order to improve the development of artificial intelligence, to make it better to serve human, we must recognize artificial intelligence essentially, analysis the trace of its development, found problems on the road ahead, to find the way to solve these problems. This paper a comprehensive analysis of previous research results, from a philosophical point of view, the use of the method of dialectical materialism and system analysis, starting from “what is the nature intelligence”, discuss the origins of intelligence; analyze the Development of Artificial Intelligence status quo from faction of artificial intelligence, weak artificial intelligence and strong artificial intelligence, artificial intelligence, the development of technical basis. I think the current artificial intelligence development bottlenecks encountered: how to make machines with human recognition knowledge, learning, reasoning and correction to reflect human intelligence, creative thinking - such as human at all times in the initiative amendment, under natural conditions, the ability to adjust their behavior, or take in the case of insufficient information obtained certain tentative choice, or that the information

obtained to the conflict between how the trade-offs in order to take the appropriate action.

From the philosophical point of view, the future development of the Artificial Intelligence (AI) will focus on the following aspects, if it calls for essential breakthroughs. Firstly, machines must be made to use natural human languages, to communicate with human beings and to identify the natural language in different contexts. A variety of ambiguity apparently exists in natural languages. Until now, the AI has been able to analyze the significance of the single word, while it shows some weakness in identifying the true meaning of long sentences of natural human languages. To solve this problem, it needs to refine and logically analyze human languages, which cannot stick to the surface of the language but to understand the connotation of it. In the 20th century, philosophy has given birth to such significant theories as Idealism and Speech Act Theory. This paper aims at analyzing the problems that the AI must solve from the perspective of Pragmatic Reasoning—the core of communicative logic of human languages. Secondly, it is another difficulty to enable machines to possess creativity—the essence of human intelligence—and to make uncertain reasoning. Creativity is the essential characteristics and the highest form of human intelligence. As long as human beings participate in creative activities, they must exert deductive reasoning. More importantly, they must have capacity to make inductive reasoning, analogical reasoning and fuzzy inference. As a vital methodology of understanding and reforming the world, incomplete inductive reasoning has achieved success in human practice. How to equip machines with this kind of ability is bound to be one of the priorities of the future research. On the other hand, there are countless vague phenomena in the world which become fuzzy concepts and propositions in human brains. To achieve the level of human beings, the AI must be able to deal with various kinds of fuzzy concepts and propositions. Fuzzy logic lays a foundation for it and relative studies on this field have achieved momentous results. In spite of the accomplishments, it is still strenuous to get algorithm extraction and the choice of Membership function. This is definitely the hot topic of AI research in the near future.

Keywords: philosophy、 human intelligence、 Artificial Intelligence

第1章 绪论

1956年的达特茅斯会议上约翰·麦卡锡第一次提出了人工智能的概念,仅仅经历了短短50多年的发展,人工智能就已渗透到人类社会生活的各个角落,甚至在许多方面表现出的能力大大地超过了人类的智能。人们一方面为其发展速度感到欣喜和震惊,另一方面,人工智能发展中产生的各种问题也引起了人们的注意、思考和讨论。越来越多的科学家及学者开始关注与研究人工智能发展中产生的问题,因为如果不能科学地解决这些问题,人工智能的走向将会越来越令人担忧,并可能对人类社会的发展造成无法预测的影响。人工智能涉及的学科甚多,包括哲学、心理学、生物学、系统学、社会学、信息学、计算机科学等,因而国内外的众多科学家和学者均从各自不同的领域对其展开了大量的研究。总体来说,对于人工智能,国外的研究起步比较早,基础扎实,在理论上和实际应用中都取得许多突破性的进展,并且对于人工智能发展中所产生的问题也进行了大量分析和研究,有大量的成果;我国对人工智能的研究虽然起步较晚,但是对人工智能所涉及到的理论的研究也有热烈的讨论,对人工智能产品的开发也进行的如火如荼,也取得了比较丰硕的成果。总的说来,人工智能作为一个处在人类科技发展前沿的新兴学科,它受到了前所未有的关注,发展势态良好。

1.1 目前人工智能的研究情况

国外关于人工智能发展问题的研究及著述有很多,如果简单的从就研究的结论来分类,总的来说大致有三类观点。第一,认为人工智能是人脑的模拟和扩展,其本身就是一个信息处理系统。斯坦福大学的费根鲍姆(Edward Feigenbaum)教授从知识工程的角度出发,他认为“知识就是力量,电子计算机则是这种力量的放大器,而能把人类知识予以放大的机器,也会把一切方面的力量予以放大。”麻省理工学院的P. H. 温斯顿教授认为“人工智能就是研究如何使计算机去做过去只有人才能做的富有智能的工作”。类似的观点在哈里·亨德森的《人工智能一大脑的镜子》¹和休伯特·德雷福斯的《计算机不能做什么》²中也都有所体现。第二,认为人工智能等同于甚至人类智能。人工智能虽然是有限的,但是其向人类智能的接近却是无限度的,随着其自身的不断突破与发展,最终将无限逼近于人类智能。人工智能之父西蒙就持这样的观点,他认为人工智能能够达到人类智能的水平。神经生物学家亨利·马克莱姆希望通过“蓝脑计划”(Blue Brain)建立人类大脑的模型,从而揭开人脑秘密。经过大量的研究分析,他认为计算机完成人脑的复制只是时间的问题。发明家雷·库兹韦尔(Ray Kurzweil)的预言听上去则有

些疯狂,他认为“21 世纪结束之前,人类将不再是地球上最有智慧或最有能力的生命实体。”^①具有人工智能的计算机将具有与人类智能并驾齐驱的各种能力,甚至是意志和感情。美国科学家艾什比认为,要制造一个综合能力的机器脑在原则上没有什么问题,所需要的只是时间和技术进步。他强调,这种脑一旦制造出来,决不只是简单的机械执行和模仿,它还能够自己学习发展自己的智慧。库兹韦尔的“奇点理论”宣称:在一个所谓的“奇点”来临之时,机器将可通过人工智能进行自我完善,甚至超过人类本身,从而开启一个新的时代。“人工大脑之父”雨果·德·加里斯在《智能简史》³一书中就指出:人工智能已经在神经和智能上产生了指数级的初级进化,并在将来某一天达到“奇点”状态,最终“失去控制”,以人类无法想象的速度自我进化,最终超越人类智能。在清华大学演讲中,加里斯教授更是预言:“可能 20 年、30 年后人工智能机器就可以和人做朋友了,但 50 年后,人工智能将成为人类最大的威胁。世界最终会因人工智能超过人类而爆发一场战争,这场智能战争也许会夺去数十亿人的生命。”但是这样的观点并不为大多数专家所接受,就连计算机之父冯·诺依曼和图灵都表示计算机不会超过人类的智能。1980 年,美国著名语言哲学家约翰·塞尔提出了“中文房间”的模型,指明了计算机能够模拟但无法超越人类智能。在这个模型中,人类可以通过建立一种动态的标准,通过不断修改充实这个标准来使机器拥有更多的智能。但由于标准的制定者为人类,它并不会比人类更聪明,只会随着人类对自我认识的提高而提高,因而最终可以想象机器模拟人类智能是一个向人类智能无限接近的过程。随着以互联网诞生为标志的第三次工业革命的到来,数据挖掘和知识发现开始崭露头角。以数据库为基础,通过运用多种学习手段,从大量数据中提炼出抽象的知识来揭示数据背后的客观世界蕴含的内在本质和联系,实现自动获取知识的知识发现系统是一个极富应用前景的研究方向。目前对数据挖掘技术研究的热点主要集中在提高算法效率、数据的时序性和互联网知识发现几个方面。此外,智能接口也是人工智能发展的热点之一。智能接口包括有语音识别、自然语言理解和图像识别等。目前这方面的研究成果已经大量的投入使用,并且正在迅速提升性能。

中国拥有一套古老的数学体系,其中《九章算术》以计算为中心,密切联系实际,主要解决人们生产、生活中的数学问题,确定了中国古代数学的框架,但是我国一直缺乏完善的数学理论体系,直到五四运动以后才真正开始了近代数学研究。20 世纪 70 年代后期,在计算机技术大发展的背景下,数学家吴文俊继承和发展了中国古代数学的传统的算法化思想,转而研究几何定理的机器证明,提出的数学机械化设想,虽然是数学领域的研究,但对人工智能的许多领域却有着深

^① (美)雷·库兹韦. 灵魂机器时代——当计算机超过人类智能时. 世纪出版集团上海译文出版社[M]. 2006 年 4 月. 第 2-3 页

远影响。随着国家人工智能相关产业的政策扶植,人工智能的研究取得了长足的进步,在理论研究方面已经达到了世界水平。1981年起,相继成立了中国人工智能学会(CAAI)、中国软件行业协会人工智能协会、中国智能机器人专业委员会以及中国智能自动化专业委员会等学术团体,1989年首次召开中国人工智能联合会议(CJCAI),中国科学家在人工智能领域取得一些在国际上有影响的创造性成果:我国科学院院士、清华大学李衍达教授首创——知识表达的情感适应模型,通过人机合作,由计算机提供候选模型,人进行情感选择,从而建立有效的信息模型。李德毅教授提出了定性和定量转换“云模型”,能够用语言值表示和处理随机不定性和模糊不定性,并成功将其应用于数据挖掘和智能控制等领域。王守觉教授则在人工神经网络的硬件化实现及高维仿生信息学方面的进行了大量研究,研制成功了半导体神经网络硬件系列,并提出了“仿生模式识别”理论,为解决机器形象思维问题提供了一条新途径。另一方面,大量的著述与论文开始着力于进行人工智能发展中出现的一些具体问题的研究,如蔡曙山在《哲学家如何理解人工智能——塞尔的“中文房间争论”及其意义》^①一文中从语言哲学家约翰·塞尔的“中文房间”模型入手,经分析得出“机器智能能够不断接近人类智能,但永远不可能超过人类智能”^②的结论;蔡自兴在《人工智能对人类的深远影响》^③一文中从经济、社会、文化等多个方面对人工智能进行了剖析,分析了人工智能对人类和人类社会方方面面的改造;童天湘的《从“人机大战”到人机共生》^④从计算机与人对弈出发,分析人工智能的发展,他认为通过生物技术和人工智能技术的结合,人类会创造出“人机共生”的智能机器人;戴汝为的《从基于逻辑的人工智能到社会智能的发展》^⑤从人工智能发展的历史进程角度切入研究,他认为采用信息网络、多媒体现代技术的“信息空间综合集成研讨体系”将会成为涌现社会智能的可操作的技术系统;王雨田在《归纳逻辑与人工智能相结合的研究问题》^⑥中以归纳的研究为例,指出当前人工智能的发展中亟须引入有关归纳的哲学、科学哲学、逻辑学、心理学、认知科学引入知识工程、机器学习、不确定性推理等等前沿科学技术的研究基础理论的研究;郑祥福在《人工智能的四大哲学问题》^⑦中认为当代西方哲学的认知转向是和人工智能的研究协调发展的,人工智能的哲学问题已不在是人工智能的本质问题,而是关于人的意向性问题、概念框架问题、语境问题和日常化认识问题,并就这些问题给出了基本的解决思路;盛晓明、项后军在《从人工智能看科学哲学的创新》^⑧从人工智能的历史发展出发,在对科学哲学中几个主要学派分析比较后认为科学哲学不仅应该经常回到“活的”科学史中去并不断地进行理论上的大胆创新与整合,才能够迎接新兴科学的挑战,才能有伟大的将

^① 蔡曙山.哲学家如何理解人工智能——塞尔的“中文房间争论”及其意义[J].自然辩证法研究.2001(11)第17卷.18页

来；刘普寅、李洪兴在《软计算及其哲学内涵》^①一文中通过论述基于模糊逻辑系统、神经网络、遗传算法等软计算技术对于研究非线性复杂系统以及处理智能信息的有效性，认为软计算在智能信息处理技术中将发挥十分重要的作用。

1.2 研究意义

关于人工智能发展问题的研究进行了许多年，成果颇丰，但是也存在着不足。国内外科学家、学者从各自不同的领域都对人工智能发展问题做了深入研究，然而他们的研究普遍缺乏唯物、辩证、历史的世界观和方法论，另一方面，对于人工智能发展的最核心的问题，比如何为人类智能？人工智能是什么？人类智能与人工智能的本质区别是什么？人工智能究竟走向何方？等，仅仅依赖于某一学科，或仅对人工智能某一方面进行研究，这样得出的结论是值得商榷的。只有运用辩证唯物主义科学的世界观与方法论才能突破瓶颈，找到解决人工智能发展中的各类问题的正确途径。

恩格斯说：“不管自然科学家采取什么样的态度，他们还是得受哲学的支配。问题只在于：他们是愿意受某种坏的时髦的哲学的支配，还是愿意受一种建立在通晓思维的历史和成就的基础上的理论思维的支配。”^①

关于人工智能发展中的各类问题国内外学者做了大量的研究，本文旨在综合前人的研究成果，从哲学的角度对人工智能发展中的问题进行深入思考，结合自我的分析研究，明确哲学在人工智能发展中所起的作用，进一步丰富人工智能发展的哲学思考，力图找出当前人工智能发展道路上的难点并尝试给出一种具有可操作性的解决方案，从而促进目前重点的研究并理清今后的发展思路。希望本文能对人工智能未来的理论研究及具体实践提供一定的理论依据，合理规范人工智能技术的开发，使人工智能更好地、更安全地为人类服务、推动社会发展。

1.3 基本研究思路与方法

仅用了不到 60 年的时间，人工智能已经取得了巨大的进步。笔者之前本科阶段学习计算机科学与技术，对人工智能方面的应用有非常浓厚的兴趣，平时也对此有初步的思考。但是之前的思考仅仅是停留在人工智能的实际应用技术方面，对它所蕴含的深层次的理论知识并未做太多涉猎。考上西南大学政治与公共管理学院哲学系下逻辑学专业高师研究生后，拜师于众多哲学大家，在哲学拓展了自己的视野，丰富了自己的知识储备后，对人工智能的思考得以继续深入。

要分析人工智能发展中的哲学问题，首先应理清“什么是智能”、“什么是人

^① 恩格斯. 自然辩证法[M]. 人民出版社. 1972 年. 187 页

工智能”这些基本的概念。本文涉及研究从智能的起源和本质切入，通过对人工智能从萌芽到迅猛发展的历史过程的反思，进而阐述其研究的派别、发展的分歧和发展技术方向等基本内涵。在此基础上，分别从“灵魂说”、“人是机器”和“信息哲学”三个方面阐述哲学在人工智能的发展中施加的影响；再通过梳理人工智能与人类智能之间的关系，分析目前人工智能发展中的哲学理论瓶颈，旨在提出目前人工智能发展到如今要有质的突破，从哲学方面来看，今后的发展将会将重点放在人工智能机器对自然语言的理解和使人工智能机器能够进行不确定性推理的两个方面的结论。

本文旨在从哲学的角度探讨人工智能发展的问题，写作中主要运用了以下两种方法：

（1）辩证唯物主义方法论。辩证唯物主义是科学的方法论，它不仅要求我们实事求是，还要有开放性、灵变性、多维性及创造性的思维，分析问题时要充分考虑影响事件的各个因素，关注事件的联系与发展，从多方面、多角度解决问题。唯物辩证方法是我们研究事物的根本方法，研究人工智能发展的问题，把握住研究方向与科学性，必须坚持辩证唯物主义方法论。

（2）系统分析方法。系统分析是把一个复杂的项目看成为系统工程，通过系统分析，确定问题的本质和起因，有效地提出解决方案，并通过一定标准对这些方案进行比较，帮助研究者在复杂的问题和环境中做出科学抉择。本文将人类智能和人工智能视为一个系统，从整体的角度，梳理人类智能与人工智能两者之间的关系，分析它们之间的相互作用，以达到深刻认识人工智能的目的。

第2章 人工智能的产生

2.1 智能的起源

智能是人在实践中对客观事物进行观察、记忆、分析、判断及有目的地行动和有效地解决世纪问题的综合能力，是人类特有的、区别于一般生物的主要特征。人特有的智能，是客观世界长期发展的产物，同时也是社会劳动的产物。

在地球诞生之初，并没有生命的存在，也就没有智能可言，而一切物质都具有的反应特性是人类智能产生的物质基础，生物的反应形式（刺激感应、感觉和心理）是人类智能产生的前提。人类智能的产生经历了三个决定性的阶段：第一阶段，由无生命物质的反应特性到低等生物的刺激感应性。反应特性是指物质在受到其他物质作用时会发生相应的变化。比如岩石受到空气、阳光、水的作用会发生风化。到约31亿年前，在无生命的物质中产生了低等生物，这种生物的反应形式是刺激感应性。所谓刺激感应性，是指生物对外界环境的变化和作用的刺激感应能力。如植物的枝叶会伸向有阳光的地方，根会扎到有水肥的地方。生物的刺激感应性是比反应特性高级的反映形式，它具有自主性、选择性和趋利避害性。因此这种刺激感应性已不是简单的机械的、无选择性的物理或化学反应，表现出一定自主反应的萌芽；第二阶段，由低等生物的刺激感应的反映形式到高级动物的感觉和心理。随着生物的不断进化，逐步进化出感觉器官和神经系统。由于有了神经系统这种物质基础，因而动物有了感觉这样的反映形式。在刺激感应性的反映中，信息、物、刺激必须同时出现，而在感觉中，信息与物可以是间接的关系，表现出某种与刺激物的原型相分离的趋向。高级动物有了由神经系统进化而来的大脑，因而有了心理这种反映形式。高等动物能把感觉器官与感觉能力联系起来，形成对客观环境的反映，这就是动物心理。第三阶段，从动物的感觉和心理到人类智能的产生。智能不仅是自然界的产物，同时也是社会的产物。由猿进化到人，这个过程是通过不断改变客观世界的社会劳动完成的，在社会劳动过程中，需要深入事物内部，了解其规律和本质，而语言的产生，使大脑能够概括出各种感觉材料，提高了抽象思维能力。在劳动和语言的推动下，猿脑变成人脑，人脑比动物脑更大，构造更加复杂，出现了动物所没有的“语言中枢”和“额叶”等，从此出现了人所独具的高级反映形式——人类智能。

2.2 智能的本质

辩证唯物主义认为智能的生理本质是以人脑和神经系统的生理活动为基础的

心理活动，而智能是人在社会生活中对客观世界的反映，这是智能的社会本质。

首先，智能的生理本质。高度复杂的物质系统——人脑为人类智能奠定了生理基础。智能依赖于高度组织、高度复杂的以大脑为中心的中枢神经系统及周围神经系统，依赖于整个生理组织的生理活动。人体感官接收客观外界事物的客观信息，通过神经系统传送到人的大脑，大脑对信息进行储存、分析及创造。这是从信息的形式到人的意识和思想的转换，离开了人的这种生理活动，就不可能产生人类智能。

其次，智能的社会本质。人类智能必须建立在社会实践的基础上。第一，人脑不会自动地产生智能，它只是人所具有的能对客观世界产生能动反映的生理组织，而人脑产生智能是以社会实践中各种信息的输入为前提的；第二，人类智能是在社会实践中穿透客观事物表象，认识事物的内在本质和规律的过程；第三，人类智能有严格的逻辑结构和学习机制；第四，通过对事物本质和规律的认识人类智能可以预测客观事物的发展趋势，进而指导社会实践。

辩证唯物主义对于人类智能本质的研究，为人工智能的发展提供了科学的哲学视野。

2.3 人工智能的产生

人工智能是通过了解智能的实质，用机械和电子装置来模拟人的思维活动并产生类似于人类智能反映的行为。人工智能从胚胎到萌芽至最后诞生，经历了较为漫长的历史发展过程。

17 世纪的英国资产阶级革命时期，近代自然科学处于形成和发展的初期，机械力学占中心地位，人们主要运用机械力学的原理来论证世界的物质统一性。哲学家托马斯·霍布斯（Thomas Hobbes）创立了机械唯物主义的完整体系，认为宇宙是所有机械地运动着的广延物体的总和。霍布斯“把世界看作由因果链组成的大机器。世界只有物体存在，物体由因果关系连接为整体。物体分为两类：自然物体和人工物体”^①。他把逻辑的思维看作是观念的加或减的机械运算，力图以机械运动原理来解释人的情感和欲望，以此来确定社会动乱和安宁的根源。18 世纪，朱利安·拉美特利、德尼·狄德罗、克劳德·爱尔维修、保尔·霍尔巴哈等哲学家的出现，把机械唯物主义的发展带到了顶峰。拉美特利从物质具有运动力和创造力的基本观点出发，批判地继承了笛卡尔的“动物是机器”的思想，进一步得出“人是机器”的结论。他认为人是有感觉、有精神的、活的机器，人的身体状况决定了人的心灵状况，人的机体组织是由物质的机械规律支配的自动机。霍尔巴赫则

^① 赵敦华. 西方哲学简史[M]. 北京大学出版社. 2006 年. 206 页

认为宇宙中的一切事物，都是由物质构成的，不但无生命的东西如此，就连人也是一个由不同物质组织而成的有机整体，是一部非常复杂、由大量物质配合而成、随着各种性质、比例、活动方式而变化的机器。

机械唯物主义认为，世界的存在就是机械运动，任何的存在物，包括人、动物和其它，其规律就是机械规律。机械唯物主义者虽然没有办法预料到计算机或是人工智能的出现，但却对计算机和人工智能的产生提供了一些启示。

18 世纪，英国发起了技术发展史上的一次巨大革命，开创了机器代替手工工具的时代，从此揭开了机器与人之间的斗争。第一次世界大战以后，各国工业迅猛发展，人们在烦躁的体力劳动中产生一种愿望，就是创造出一种机器，能够代替人进行各种工作，这种想法成为机器人产生的客观要求。

1936 年，被称为计算机科学之父的英国数学家阿兰·麦席森·图灵（Alan Mathison Turing）设计了图灵机，这是一种抽象的计算模型，图灵从理论上证明了人们用纸笔进行的任何数学运算过程都可以用这个机器进行模拟。图灵机的设计蕴含了构造某种具有一定的智能行为的人工系统以实现脑力劳动部分自动化的思想，这正是人工智能的研究目标。1941 年采用电子方式处理数据的电子计算机的出现，为人工智能的实现奠定了技术基础。1947 年，图灵在一次关于计算机的会议上作了题为“智能机器”（Intelligent Machinery）的报告，详细地阐述了他关于思维机器的思想，第一次从科学的角度指出：“与人脑的活动方式极为相似的机器是可以制造出来的。”图灵这一新奇却不乏深刻的思想远远超出了当时人们的想象力。同年，美国橡树岭国家实验室在研究核燃料的工作中，设计产生了世界上第一台主从遥控的机器人，从此拉开了机器人研究的序幕。1947 年 10 月，数学家诺伯特·维纳（Norbert Wiener）写出划时代的著作《控制论》，该书的副标题是“关于在动物或机器中控制或通讯的科学”，这就是他对控制论所下的定义。维纳揭示了机器中的通信和控制机能与人的神经、感觉机能的共同规律，他认为人类智能源于反馈机制的作用，而反馈机制可以用机器进行模拟。控制论的概念跨接了许多领域，将通讯理论、控制论与人类神经系统的工作原理联系了起来，为人工智能的发展指明了方向。1950 年 10 月，图灵在《Mind》杂志上发表了一篇划时代的论文——《计算机与智能》，在这篇文章中他首次提出了“机器思维”的概念，而在另一篇作为补充说明的文章中，图灵进一步写道：“‘你无法制造一台替你思考的机器’，这是人们一般会毫无疑义接受下来的老生长谈。……我的论点是：与人脑的活动方式极为相似的机器是可以制造出来的。这些机器有时会出现错误，但有时它们也会提出非常新颖的语句，而且总的来说，它们输出的东西将与人脑输出的东西同样值得注意。”图灵提出一种检验机器是否被认为是具有思维的方法：即在互相不接触并且不知情的情况下，让计算机和人通过某种特殊的方式进

行一系列的问答，如果在足够长的时间内，人无法根据这些问答对对方是人还是计算机做出有效判断，那就可以认为这台计算机具有了与人相当的智力，从另一个角度说，这台计算机能够进行思维。这个假想即是著名的“图灵测试”(Turing Testing)。图灵不仅提出了“机器思维”的概念，第一次从行为主义的角度对人工智能给出了定义，还大胆预言随着计算机的发展，机器的智能将不断提高，最终将出现智能机器。

1955 年被称为人工智能之父的约翰·麦卡锡 (John McCarthy) 联合“信息论”的发明者克劳德·艾尔伍德·香农、框架理论的创始者马文·明斯基及 IBM 的杰出计算机设计师纳撒尼尔·罗切斯特发起了达特茅斯项目。1956 年，四人在美国新罕布什尔州汉诺威市的达特茅斯学院组织举办了达特茅斯夏季人工智能研讨会 (Dartmouth Conference)，这项涉及数十名来自数学、心理学、神经学、计算机科学与电气工程等各种领域的学者、耗时 2 个月的雄心勃勃的研究计划，其目标是“基于‘我们能够精确、全面地描述人类智能中的学习以及其他特征、并制造出机器对其模拟’这一构想，继续阔步前进”。这次会议上，马文·明斯基的神经网络模拟器 Snare (主要学习如何通过迷宫)、麦卡锡的下棋程序尤其是 $\alpha - \beta$ 搜索法，以及赫伯特·西蒙和艾伦·纽厄尔带去的已能证明数学名著《数学原理》一书第二章 52 个定理中的 38 个定理的启发式程序“逻辑理论家”(Logic Theorist) 成为会议的 3 个亮点。会议首次提出了“人工智能”这一术语，确立了研究目标，标志人工智能学科的诞生。此后，人工智能就作为一门新兴学科驶入了发展的高速公路，人们从不同的角度、基于不同的学科开展着研究，在各个方面取得了巨大的成就。

第3章 人工智能的基本内容

自“人工智能”在1956年达特茅斯会议上被提出至今，仅仅才过去50多年，但从其产生到发展成熟，却历经了较为漫长和波折的过程。

人工智能的研究从狭义的认识来说是让计算机理解并实现人类智能的研究，包括逻辑思维、形象思维等，这些思维能力通常都可以在博弈的过程中体现。所以虽然图灵为人工智能的发展指明了方向，但最初的研究大多局限在研究计算机是否可以胜过人的博弈问题，尤其是计算机博弈这个问题上，大多数研究者认为若是掌握了博弈的本质，就触摸到了人类智能行为的核心。1965年，英特尔微处理器制造公司创始人之一的戈登·摩尔（Gordon Moore）发现了计算机第一定律——摩尔定律，即每个新的芯片都产生于前一个芯片产生后的18至24个月内，并且其容量大体上是前一个芯片的两倍。摩尔定律是表述硬件计算潜能的定律，作为一个伟大的科技趋势，被证明一直有效，以至于倍数增加得到的数字非常巨大，若是此定律持续有效，那么其计算能力相对于时间周期将呈指数式的上升，而其惊人的存储能力及运算能力就足以产生高智能机器。芯片存储的增加和运算速度的提高为人工智能的发展打下了技术基础。

1969年召开了第一届国际人工智能联合会议，随着之后一系列学术活动的举办和社会团体的成立，人工智能的研究得到了快速的进步与发展。而20世纪70年代后期至80年代初期，两个根本的问题的悬而未决使人工智能的发展曾一度落入低谷，一是所谓的交叉问题，即传统方法只能模拟人类深思熟虑的行为，但不包括人与环境的交互行为；二是所谓的扩展问题，即传统人工智能方法只适合与建造领域狭窄的专家系统，不能推广到规模更大、领域更宽的复杂系统中去。1977年，美国人工智能专家E. A. 费根鲍姆提出了知识工程（Knowledge Engineering）的概念。人工智能与计算机技术的结合产生了所谓“知识处理”的新课题，即要用计算机来模拟人脑的部分功能，首先必须获取知识，并能把知识表示在计算机中，能运用它们来解题。因此，知识的获取、知识的表示和知识的运用也就成了知识工程主要研究内容。整个20世纪80年代，专家系统和知识工程在全世界得到迅速发展，使得计算机系统的应用从数值处理、数据处理阶段进入知识处理阶段。1982年，D. E. 鲁梅哈特和J. L. 麦克莱兰领导了一个研究小组探讨平行分布加工模型，并在1986年出版《平行分布加工：认知的微观结构的探索》一书中，对具有非线性连续转移函数的多层前馈网络的误差反向传播算法（又称BP算法）进行了详尽的分析，实现了明斯基关于多层网络的设想，重新激发了人工神经网络的研究热情。数据挖掘技术的出现为人工智能的发展注入了新的活力，数据挖

掘是指从数据库中提取隐含的、未知的、非平凡的及有潜在应用价值的信息或模式。它能对庞杂的数据进行科学的梳理和细分,从中发现潜在的价值规律,人工智能技术所包含的知识表示、推理和检索技术都在数据挖掘中得到体现。目前对数据挖掘技术研究的热点主要集中在提高算法效率、数据的时序性和互联网知识发现几个方面。此外,智能接口也是人工智能发展的热点之一。智能接口包括有语音识别、自然语言理解和图像识别等。现在,人工智能正向高度智能化方面发展,不同学科的相互交叉和融合为人工智能的发展带来了更多的突破和挑战。

3.1 人工智能的派别

人工智能的研究发展的 50 多年历史中,不同学科背景的科学家及学者从各自领域着手,对人工智能做出了各自的理解,提出不同的观点,逐渐产生了不同的学术流派,其中对人工智能研究影响较大的是符号主义、连接主义和行为主义三大学派。

3.1.1 符号主义

符号主义 (Symbolicism),又被称作逻辑主义 (Logicism)、心理学派 (Psychologism) 或计算机学派 (Computerism),是一种基于逻辑推理的智能模拟方法,其原理主要为物理符号系统假设和有限合理性原理,早期的人工智能研究者大多属于此类,学派的代表人物有赫伯特·西蒙、艾伦·纽厄尔和尼尔逊等。

该学派认为:人工智能源于数理逻辑,人类认知和思维的基本单元是符号,认知过程就是在符号表示上的一种运算,而数理逻辑试图找出构成人类思维或计算的最基础的机制,例如推理中的“代换”、“匹配”、“分离”,计算中的“运算”、“迭代”、“递归”,用计算的方法来代替人们思维中的逻辑推理过程。十七世纪莱布尼茨就曾经设想过能不能创造一种“通用的科学语言”,可以把推理过程象数学一样利用公式来进行计算。计算机出现后,数理逻辑和计算机科学具有完全相同的宗旨及形式化方法,于是符号主义者开始寻求在计算机上实现逻辑演绎系统的方法,通过研究人类认知系统的功能机理,用某种符号来描述人类的认知过程,并把这种符号输入到能处理符号的计算机中,就可以模拟人类的认知过程,从而实现人工智能。西蒙、艾纽厄尔和约翰·肖共同开发的启发式程序——逻辑理论家,证明了《数学原理》一书中 38 条数学定理,这意味着数理逻辑用于描述智能行为和实现逻辑演绎系统,用计算机研究人的思维,模拟人类智能活动是可行的。之后,符号主义走过了一条启发式算法——专家系统——知识工程的发展道路,尤其是专家系统的成功开发与应用,使人工智能研究取得了突破性的进展。我们可以把符号主义的思想简单的归结为“认知即计算”。

符号主义作为人工智能的主流派别，为人工智能的发展作出了突出贡献。特别是知识工程的产生与应用，成为指导具体研制各类智能系统的一般方法和基本工具，成为一门具有方法论意义的科学。

3.1.2 连接主义

连接主义 (Connectionism)，又称作仿生学派 (Bionicsism)、生理学派 (Physiologism) 或结构模拟学派，该学派以人工神经网络模型为核心，认为人工智能是一种基于神经网络、网络间连接机制与学习算法的智能模拟方法。学派的代表人物是心理学家麦卡洛克和数理逻辑学家皮兹。

连接主义认为人工智能源于仿生学，把人的智能归结为人脑中大量的简单单元通过复杂的相互连接并进行活动的结果。连接主义者希望通过研究搞清楚大脑的结构以及它进行信息处理的过程和机理，揭示人类智能的奥秘，从而真正实现人类智能在机器上的模拟。连接主义具有代表性的成果是由生理学家麦卡洛克 (McCulloch) 和数理逻辑学家皮茨 (Pitts) 于 1943 年创立的脑模型，也称为“McCulloch-Pitts 神经元模型”，即 M-P 模型。他们总结了神经元的一些基本生理特性，提出神经元形式化的数学描述和网络的结构方法，从此开创了神经计算的时代，为人工智能创造了一条用电子装置模仿人脑结构和功能的新途径。以感知机 (Perceptron) 为代表的脑模型研究在 20 世纪 60 年代出现过热潮，由于受到理论及技术条件的制约，在 20 世纪 70 年代后期至 80 年代初期又陷入低潮。1982 年，美国物理学家霍普菲尔特提出了离散的神经网络模型，1984 年他又提出了连续的神经网络模型，使神经网络可以用电子线路来仿真，开拓了神经网络用于计算机的新途径。1986 年，鲁梅尔哈特等人提出了多层网络中的反向传播算法，即 BP 算法，使多层感知机的理论模型有所突破。

连接主义以人工神经网络为其典型代表性技术，因此，我们可以把连接主义的思想简单地称为“神经计算”。由于和其他学科的交叉融合，使这一技术在图像处理、模式识别等领域取得了重要的突破，为实现连接主义的智能模拟创造了条件。

3.1.3 行为主义

行为主义 (Actionism)，又称作进化主义 (Evolutionism) 或控制论主义 (Cyberneticsism)，其理论核心为控制论以及感知-动作模式控制系统。该学派认为智能行为的基础是“感知——行动”的反应机制，学派代表人物是罗德尼·布鲁克斯。

行为主义是 20 世纪末以人工智能新学派的面孔出现的，行为主义认为人工智能源于控制论。1948 年诺伯特·维纳发表了著名的《控制论——关于在动物和机

器中控制和通讯的科学》，他在书中指出：“控制论是在自控理论、统计信息论和生物学的基础上发展起来的，机器的自适应、自组织、自修复和学习功能是由系统的输入输出反馈行为决定的”^①。控制论把神经系统的工作原理与信息理论、控制理论、逻辑以及计算机联系起来。控制论推进了机器人研究，机器人是“感知——行为”模式，通过系统与环境的交互，从运行环境中获取信息，从而做出相应的反映行为。这一学派的代表首推罗德尼·布鲁克斯 (Rodney A. Brooks)，他研制了一个由 150 个传感器和 23 个执行器构成的像蝗虫一样能做六足行走的机器人试验系统。这个机器人被看作是新一代的“控制论动物”，虽然不具备人的推理能力，但其应付复杂环境的能力却大大超过了原有的机器人，总的来说就是一个模拟昆虫行为的、基于感知——行为模式的控制系统。

以上三个人工智能学派虽有各自有其片面性，但都对人工智能的发展做出了巨大贡献。随着各个学科的交叉融合，三个学派也将长期共存与合作，取长补短，共同创造人工智能发展的新天地。

3.2 人工智能发展的分歧

人工智能发展研究的众多争议问题中，弱人工智能和强人工智能的分歧最为突出也最为研究者关注。依据对程序化的计算机在人工智能研究中的作用所持有的不同态度，人工智能分为强人工智能和弱人工智能，弱人工智能持工具主义的态度，认为程序是用来解释理论的工具，而后者持有实在论的态度，认为程序本身就是解释对象——心灵。细分人工智能的强弱有助于理解它们之间的关系，以期从中找到更加适合人工智能发展的方法及途径。

3.2.1 弱人工智能

国内外的研究对弱人工智能所持观点基本是统一的：能真正地进行推理和解决问题的智能机器是不可能被制造出来的。因为机器只能执行人的指令，其被赋予的某种智能是被设定的，机器仅仅是通过运行设定好的计算机程序而对特定的外界刺激做出相对的反应，而数字计算机只是一种科研工具，它的价值仅体现在作为实验工具的精确性上。因为智能机器不会具有自主意识，并不会主动地去学习知识，仅仅是在人类给定的范围内进行确定的数学运算，所以并不真正拥有智能。

弱人工智能所表现出的行为方式是人类预先设定好的，其目的性是生命以外的，是程序性的目的，其知识是有限的，运算过程也是处在一个相对封闭的环境

^① 约翰·华生 (John B. Watson). 行为主义心理学[M]. 李维译. 浙江教育出版社. 1998 年. 56 页

之内。即使它的能力在今后的发展中得到非常大的提升,这样的提升也仅仅是知识量的增加、运算速度提高和运算范围的扩大。人类依靠弱人工智能产品极大的减轻了人脑的工作量,使自己从简单重复劳动中解放出来,并将有限的精力投入到更重要的研究当中。这样的人工智能完全是受人类掌控的,所以它作为人类手中的工具对人类来说相对安全。

3.2.2 强人工智能

与弱人工智能的观点相反,强人工智能认为真正能推理和解决问题的智能机器是可以被制造出来的,并且机器将具有知觉及自我意识,“带有正确程序的计算机确实可被认为具有理解和其他认知状态,在这个意义上,恰当编程的计算机其实就是一个心灵”^①。强人工智能有类人的人工智能和非类人的人工智能之分,前者认为机器的思考和推理与人的思维相似,而后者则认为机器具有和人完全不一样的知觉和意识,其推理方式也和人完全不一样。

目前强人工智能在发展遇到了哲学上的瓶颈。强人工智能在实现的时候遇到了意向性、知识框架等问题。首先谈谈意向性问题,布伦塔若(Franz Brentano)认为“每一精神活动所具有的‘意向性’特性,即超越性,指向实在事物的特性”^②。机器是否具有意向性,用什么标准来判定机器拥有意向性是一个很重要的问题。图灵在著名的图灵测试中通过设定一种对话模式来判定测试对象是人还是机器,能够通过测试的机器则会被判定成人类。但是1980年,美国著名语言哲学家约翰·塞尔提出了“中文房间”的模型,它指明计算机能够模拟但无法超越人类智能,在这个模型中,人类可以通过建立一种动态的标准,通过不断修改充实这个标准来使机器拥有更多的智能。但由于标准的制定者为人类,它并不会比人类更聪明,只会随着人类对自我认识的提高而提高,因而最终可以想象机器模拟人类智能是一个向人类智能无限接近的过程。塞尔把意向性作为大脑最根本的特性,他认为“思维并不是那种非常机械的东西,因为虽然大脑是一台机器,但它的操作是由能量的传输来规定的,……,(机器计算)它代表的是一个抽象的、我们用来执行机器运算的数学过程。机器计算与思维不同,它不是由能量传输来规定的。”^③“脑功能产生心灵的方式不能是一种单纯操作计算机程序的方式”^④,从哲学上否定了强人工智能的实现。

其次是知识框架问题,也就是人工智能知识背景问题。这个问题的最早是由麦卡锡和海斯从技术层面提出的,主要是为了让机器能够应付不断变化的环境。

^① (英)玛格丽特·A·博登编著.人工智能哲学[M].刘西瑞等译.上海译文出版社.2006: 73 页

^② 刘放桐等编著.新编现代西方哲学[M].人民出版社.2000.220 页

^③ 蔡曙山.关于哲学、心理学和认知科学的12个问题——与约翰·塞尔教授的对话[J].学术界.2007(3).13 页

^④ (美)塞尔.心、脑和科学[M].杨音莱译.上海译文出版社.1991年.30—31 页

但是随着研究的深入,这个技术层面问题逐渐发展到认识论层面。认识论里面一个最基本的问题就是“知识的来源在哪里?知识体系的标准和基础是什么?知识的确定性依据是什么?”^①古往今来,知识的追求一直都是认识论的中心。人工智能要建立对客观世界的知识框架是一个巨量的工程,它需要对客观世界有准确完善的描述,同时还要建立起一套知识体系并正确的运用。目前这是人工智能发展的难点之一。

3.2.3 弱人工智能与强人工智能的关系

无论是弱人工智能还是强人工智能都是赋予机器人类的智能及其行为方式,从而代替人类进行劳动。马克思说过,“动物只是按照它所属的那个物种的尺度和需要来进行塑造,而人则懂得按照任何物种的尺度来进行生产,并且随时随地都能用内在固有的尺度来衡量对象;所以,人也按照美的规律来塑造物体”^②弱人工智能“根本没有意向状态,它只是受电路和程序支配简单地来回运动而已”^③,弱人工智能不会自我学习及自我升级,其所表现出的智能及行为完全处于人类的控制之下,相对安全,但较之强人工智能相对简单,缺乏主动性。强人工智能需要的可能仅仅是一个启动的命令便会自动执行,这无疑减轻了人类的工作强度,但其在产生自我意识并主动分析判断之后可以不完全听命于人,这无疑在安全性产生很大不确定性。

人类智能不仅是表现为具有感知、思维能力以及自我意志,当面对外界环境的刺激时,人类的行为还具有一定目的性的判断。而人工智能的行为也具有目的性,最初是为人类所设计,但当其出现自我判定的目的性时,这样的智能机器还能够为人类所控制吗?强人工智能机器如果开始拥有以维持自己的生命为目的的主动行为,并放弃其被设计之初的目的,这时拥有强人工智能的机器可以说已经非常接近于“人”了。当它需要被销毁时,或许会像电影《终结者》中描绘的那样做出反人类的行为。到那个时候,人和机器之前谁设计谁,谁控制谁也许就不再像现在这样是一个单方面的行为了。

如今,人工智能的研发和制作在技术上已有了飞速的发展,反而是对于人工智能应该怎样认知、其发展方向该如何确定,这些才是人类面临的最困难的事情。在人工智能的发展道路上,比起弱人工智能,强人工智能无疑会带来更大的效率与成果,但弱人工智能凭借其安全性与可靠性依然有其重要意义。二者的争议没有减缓在人工智能在各自领域的发展,反而积极的推动着人工智能的进步。

^① 谢礼圣等.西方哲学认识论与直观问题的凸现——从古希腊到近代科学[J].河南师范大学学报(哲学社会科学版).2005(4).第6页

^② 马克思.1844年经济学——哲学手稿[M].人民出版社.1972年.50-51页

^③ (英)玛格丽特·A·博登编著.人工智能哲学[M].刘西瑞,王汉琦译.上海译文出版社.2005年.83页

3.3 人工智能发展的技术

本着为人类服务这个相同的目的,人工智能的研发有着多样的技术手段。计算机技术井喷式的发展给人工智能研发打下了坚实的技术基础,依靠着先进的科学技术,目前人工智能的研发主要集中在以下三个方面。

一是系统仿真,“就是根据系统分析的目的,在分析系统各要素性质及其相互关系的基础上,建立能描述系统结构或行为过程,且具有一定逻辑关系或数学方程的仿真模型,据此进行实验或定量分析,以获得正确决策所需的各种信息”^①。系统仿真的基本方法是建立系统的结构模型和量化分析模型,并转换为计算机可执行的仿真模型。针对连续系统和离散系统,分别主要采用连续系统仿真方法和离散系统仿真方法。

系统仿真是一种人为的试验手段,对于物理和数学模型难以解决的问题,通过提供仿真模型来预测、分析和评价,还可以将复杂系统拆解成若干相对简单的子系统加以分析,对于系统中隐藏的问题也能及时发现。

二是人脑模,拟亦称“心理模拟”,是以人脑活动作为模仿对象的自动机器设计。由此设计而成的自动机器,在信息传入传出系统,控制系统,反馈器传导,网状结构等方面,都与人脑及其活动有一定程度的类似。最早的电子计算机,即由模拟神经系统活动设计而成。

人脑模拟比较有代表性的有2005年亨利·马克莱姆领军的“蓝脑计划”,该计划的目的是从实验数据逆向打造哺乳动物的大脑。项目组已经完成了人脑新皮层部分的神经元计算工作,并已绘制出一份3D神经元活动模拟图,科研人员希望研究最终能知晓并模拟出整个人脑的神经元的活动情况,借此了解现实世界为我们带来的感官讯息是如何破译并储存到人脑中的,从而揭开人类意识产生之谜。

三是人工神经网络(Artificial Neural Networks),也称作神经网络(NNs),或是连接模型(Connectionist Model),是一个由大量处理单元相互联接而形成的非程序化、具有自适应性的信息处理系统。其原理是通过模拟人脑神经网络储存及处理信息的过程,调整其内部各个处理单元之间相互连接的关系,从而达到处理信息的目的。人工神经网络是并行分布的,具有速度快、自适应强等特点。BP网络(前向多层反传神经网络)是目前应用最为广泛的人工神经网络模型之一。

人工神经网络是在现代神经学的研究基础上发展而来的,人类神经系统的基本功能单位是神经元,大量神经元的相互连接构成神经网络,从外界接受的信息就存储于神经元之间的连接中,而随着连接强度的变化产生了思维。人工神经网络的本质就是通过网络变换及强度变化,完成信息的分布式存储及并行协同处理。得利于其独特的技术机理,人工神经网络克服了人工智能在处理直觉、非结构化

^① 汪应洛主编.系统工程.机械工业出版社[M].2003年第3版.98页

信息方面的缺陷，并具有以下优点：第一，具有实时学习的特点。人工神经网络可以通过学习准则进行学习。比如在进行字母识别时，将不同的字母和对应的应识别的结果输入人工神经网络系统，当做出错误判断时，人工神经网络会通过自学习功能，减少下次犯同样错误的可能性，最终达到迅速、准确地判断和识别此字母。第二，具有逻辑推理及联想记忆的能力。第三，人工神经网络具有自适应及自组织能力。2000年美国国家科学基金会的一份研究报告中提到“聚合技术以认知科学为先导。因为一旦我们能够以如何、为何、何处、何时这4个层次上理解思维，我们就可以用纳米科技来制造它，用生物技术和生物医学来实现它，最后用信息技术来操纵和控制它，使它工作。”^①不难想象，下一代人工神经网络将具有多么强大的处理能力。人工神经网络具有自学习能力，若是只对系统做出了某些规则限制，而具体内容随着系统所处环境（输入信号情况）可以发生改变时，系统可以自动发现环境特征及规律。

人工神经网络具有多学科交叉技术的特点，在模式识别、智能控制、组合优化等领域得到了广泛的应用，并推动着人工智能技术不断发展。

^① Mihail C. Roco and William Sims Bainbridge, eds. *Converging Technologies for Improving Human Performance*, Virginia: National Science Foundation. Arlington. 2002

第4章 人工智能发展中的哲学思考

自上世纪50年代人工智能诞生以来,其发展的速度和水平早已大大的超出了人们的预期。在高速发展的同时,各种各样的问题也不断的产生。“人工智能是人类的助手”;“人工智能可以达到专家级水准”;“人工智能将达到人类智能的水平”;“人工智能会超过人类,机器人最终会统治世界”等等问题直接影响了人工智能的发展,这些问题随着人工智能发展越发地困扰着世人。而人工智能发展中最核心的问题就是何为人工智能的问题。

弗里德曼在《制脑者:创造堪与人脑匹敌的智能》一书中指出:“在研究各种独立的人类智能方面,人工智能项目可以说是其中的最杰出代表,有些成果是非常令人惊奇的,但这些研究的对象是过于具体化的功能,所以没法从它们中总结出规律性,另一个问题是它们不会直接从周围环境中汲取所需,而只能坐在那儿,直到人们给它们信号,然后也仅仅是复制这些信号而全然不知它的意义。它们中没有有一个程序能从周围环境中学习或适应环境,而这些哪怕最简单的生物也会具有的功能,却被我们人工智能学者忽略了。”^①

辩证唯物主义在解释人类智能时认为物质决定意识,客观世界的存在决定人类智能产生;人类智能是人在社会实践过程中人脑对客观世界的能动反映。人工智能虽然源于现代控制论、信息论等,但在辩证唯物主义看来这个过程是用机械和电子来模拟人类的记忆、思维和判断方式等的一个客观过程,这缺少了最重要的“主观能动性”。

4.1 哲学对人工智能发展的影响

4.1.1 灵魂说

自古以来灵魂在很多宗教思想都存在,是指人类超自然及非物质的组成部份。许多宗教都认为,灵魂居于人或其他物质躯体之内对其起主宰作用,亦可脱离这些躯体而独立存在。

毕达哥拉斯认为数是世界的本原,是构成实物世界的基础,“数的和谐”是自然界的一切现象和规律都必须遵从的准则。灵魂作为实物世界中的一种存在,“灵魂是一种和谐”^②,也应服从数的关系。哲学家柏拉图认为灵魂是独立的,不能加以分解的,有生命和自发性,是精神世界的、理性的、纯粹的,因为有追求世界的欲望,而堕落到地上,被圈入于肉体中,注定要经过一个净化的阶段,灵魂会

^① (美) 弗里德曼. 制脑者: 创造堪与人脑匹敌的智能[M]. 张陌, 王芳博译. 三联书店. 2001年. 10页

^② 苗力田主编. 古希腊哲学[M]. 中国人民大学出版社. 1989年. 64页

轮回转世,且灵魂是人的本性。在他看来,“人不是灵魂与身体的复合,而是利用身体达到一定目的之灵魂。”^①其后的新柏拉图主义都有对灵魂作描述。柏拉图的学生亚里士多德的著作中曾将灵魂与心理学一同讨论,他认为灵魂是形式,肉体只是质料,灵魂才是实体。灵魂分为3个部分,有理性、感觉和营养的机能,灵魂是“潜在的具有生命的自然形体的形式。”^②如今依然有许多人相信灵魂的存在,邓肯·麦克道高(Duncan MacDougall)将人死前死后的体重记录下来,对大量的数据进行分析对比,发现人在死后体重变轻了21克,突然消失的这点重量会不会就是传说中的灵魂呢?

人工智能发展中核心问题就是是否存在类人的意志活动,重新审视前人对灵魂的探索,不由地引发我们的大量的联想:“灵魂由什么构成?如果机器有了灵魂,那么与心理药物相对应的是什么?与疼痛相对应的是什么?从一间整洁的办公室所得到的生理和心理上的满足所对应的又是什么?”^③假设我们可以将某个人脑中的信息、活动复制并输入到某台机器之中,使之按照人的思维来活动,那就可以说复制了这个人的灵魂,这台机器从某种意义上说就成为了拥有机械身体的“人”。

灵魂说虽然本质上属于唯心主义观点,但它却为人工智能的发展提供了一种思路。

4.1.2 “人是机器”

始于16世纪的文艺复兴对近代早期欧洲的学术产生了深刻的影响,并在文化、知识、社会和政治各个方面都引发了革命。这一时期牛顿经典力学的出现席卷了整个科学界并将影响扩大到了哲学界。机械唯物主义思想在这样的大背景下应运而生。机械唯物主义肯定世界是物质的和运动的,物质按规律运动,意识是物质的反映。同时用机械力学原理来解释一切现象和过程,把自然界和社会的变化过程归结为数量增减、位置变更,把运动看作是外力的推动。

拉·梅特里根据大量医学、解剖学和生理学材料完成了《人是机器》一书,证明人的机体状况决定了人的心灵状况,特别着重证明了思维是大脑的机能和道德源于机体的自我保存的要求。他指出:“人是动物,因而也是机器,不过是更复杂的机器罢了。”^④并公开表明了唯物主义和无神论的立场,驳斥了心灵为独立的精神实体的唯心主义观点,论证精神对物质的依赖关系。马克思指出:“人直接地是自然存在物,人作为自然存在物,而且作为有生命的自然存在物,一方面具有自然力、生命力,是能动的自然存在物;这些力量作为天赋和才能、作为欲望存在

^① 赵敦华.西方哲学简史[M].北京大学出版社[M].2006年7月.61页

^② 苗力田主编,古希腊哲学[M].中国人民大学出版社[M].1989年.480页

^③ (美)雷·库兹韦.灵魂机器时代——当计算机超过人类智能时.世纪出版集团上海译文出版社[M].2006年.172页

^④ (法)拉·梅特里[M].人是机器.商务印书馆.1981年.65页

于人身上；另一方面，人作为自然的、肉体的、感性的、对象性的存在物，和动植物一样，是受动的、受制约的和受限制的存在物。”^①

拉·梅特里的研究从医学和解剖学的角度为人工智能中仿真机器人以及智能机器的发展开辟了新的道路。

4.1.3 信息哲学

以互联网诞生为标志的第三次计算机革命，催生了各种计算机与信息的理论。“仅就数量而言，似乎在一夜之间，计算机告诉我们的非欧几何新定理就比过去许多年的积累还要多。”^②以信息和计算机科学以及信息和通信技术等相关的技术为代表的计算科学、信息论、系统论、控制论、信息与通信技术以及人工智能在信息科学领域井喷式的爆发正导致一场大变革，这便是所谓的“计算机革命”或“信息转向”，这场变革不仅发生在科学领域，信息哲学的突起引起了哲学领域的又一个浪潮，最终成为一个独立于传统哲学之外的全新的研究领域。

牛津大学哲学家弗洛里迪（Luciano Floridi）认为信息哲学主要是进行两方面的研究，一个是信息的基本原理及其本质的研究。信息哲学的任务是要发展一个以分析、评价和解释信息的各种原理和概念为目的的整合的理论体系；另一方面是信息理论和计算方法对哲学问题的详细阐述和应用。黑格尔说过，“哲学是对于对象的思维性考察。”^③通过对信息技术发展中产生的各种问题进行必要的哲学反思，信息哲学得到了蓬勃发展。

人工智能是“是对复杂信息处理问题的研究”^④。信息哲学大力推进了人工智能的理论发展及实践应用，同时人工智能也在验证信息哲学研究中出现的问题，推动着信息哲学的创新。

4.2 人工智能与人类智能的关系

4.2.1 模拟人类智能

人工智能的发展中最基本的问题就是什么是人工智能。由此而衍生出的问题各种各样，如怎样发展人工智能，人工智能究竟能发展到什么程度，人工智能是否终会超越人类智能等等。就目前的发展情况来看，已经有大量的智能产品进入人类的日常工作、学习和生活中，它们各自扮演着不同的角色，帮助人类认识和改造世界。一些人会因为人工智能的某些运行方式像人类智能就它等同与人类智能；而经历不到60年的发展历程后人工智能产品所展现出的某些能力就大大的超

^①马克思恩格斯全集[M]. 第42卷. 人民出版社. 1979年. 167页

^②吴文俊. 王者之路——机器证明及其应用[M]. 湖南科技出版社. 1999年. 82页

^③西方哲学原著选读（下卷）[M]. 商务印书馆. 1981年. 373页

^④（英）玛格丽特·A·博登编著. 人工智能哲学[M]. 刘西瑞，王汉琦译. 上海译文出版社. 2005年. 143页

过了人类本身又催生了人工智能会全面复制并超越人类智能的论调。

这个时候就需要运用辩证唯物主义的思维来分析上述问题。辩证唯物主义认为,人具有认识世界和改造世界的能力是人与其他一切事物区别开来的最本质的特征。正是因为人类智能的主观能动性和创造性,世界才会在人类的改造下变得越来越美好。随着人类社会的不断发展,人类的实践活动对人类的体力和脑力都提出了更高的要求,生产力推动人们在解放了手脚以后进一步解放自己的大脑。1947年,控制论的提出揭示了机器中的通信和控制机能与人的神经、感觉机能的共同规律,认为人类智能源于反馈机制的作用,而反馈机制可以用机器进行模拟。这为人工智能模拟人类智能打下了理论基础。

人们在研究人类智能处理信息过程的时候,发现人脑中的神经细胞通过发送“兴奋”和“抑制”两种状态来传递信息,后来人们发现了半导体,正好可以通过利用电脉冲,以半导体“导电”和“不导电”两种状态来分别代表人脑中神经细胞的“兴奋”和“抑制”,从而模拟人脑中信息传递的过程。但是这种模拟只是简单的模拟处理数据的过程,只是一个简单的按照程序设定的方式处理信息的过程,纯粹是一种毫无意识的机械活动。这样看来,人工智能只是因为生产力发展到一定水平后对人类提出了更高的要求的情况下产生的,人工智能是对人类智能中某种机械运算能力的一种模拟,而不是真正意义上的人类智能。

4.2.2 延伸人类智能

首先从人类发展的历史来看,正像达尔文发现有机界的发展规律一样,马克思发现了人类社会的发展规律“我们首先应当确定一切人类生存的第一个前提,也就是一切历史的第一个前提,这个前提是:人们为了能够‘创造历史’,必须能够生活。但是为了生活,首先就需要吃喝住穿以及其他一些东西。因此第一个历史活动就是生产满足这些需要的资料,即生产物质生活本身,而且这是这样的历史活动,一切历史的一种基本条件,人们单是为了能够生活就必须每日每时去完成它,现在和几千年前都是这样”^①。恩格斯也说过“人们首先必须吃、喝、住、穿,然后才能从事政治、科学、艺术、宗教等等”^②人类从最初的要活下来到后来要活的好的诉求是随着生产力的不断提高而产生的。人类在不断提升自己生产力的同时,生产力也作用于人类,催促着生产工具的不断改进,直至智能机器的产生。人类在不断的实践中发明了各种各样的工具来拓展自己的能力,斧头可以砍柴,弓箭可以捕猎,而智能机器则可以帮助人类的大脑处理部分繁琐的重复劳动。智能机器是在人类主观能动的创造,体现了人的本质。

^①马克思恩格斯选集[M].第1卷.人民出版社.1995年.第78—79页

^②马克思恩格斯选集[M].第3卷.人民出版社.1995年.第776页

生产力决定着人类社会的存在和发展。人类的生产劳动经历了无意识的使用工具到使用智能的机器,虽然这些机器在原理和功能等方面有着天差地别的不同,但或多或少都扩展了人类自身的能力,归根结底都是对人类肢体的延伸,是对人类智能的扩展。

4.2.3 人类智能的组成部分

人工智能在扩展人类智能的工作中展现出惊人的能力。于是,“机器将会统治世界”的讨论不绝于耳。人工智能如何发展,向何处发展,发展到何程度影响着人工智能的发展。丰富的想象是科学发展的动力,但是在关于人工智能如何发展的问题上绝不能仅仅依靠想象。运用辩证唯物主义世界观和方法论来分析人工智能,我们应该认识到:人工智能只是人类智能的组成部分,其产生和发展是有条件限制的。

首先,人工智能在不断发展。辩证唯物主义认为人类智能和人类社会是不断发展的。恩格斯说,“人的智力是按照人如何学会改变自然界而发展的”。^①随着人类改造世界的深入人类智能也不断地发展,并在各种工具的助力之下获得了极大的延伸。人工智能就像钳子延伸手,轮子延伸脚一样延伸了人类的大脑,扩展着人脑的功能,弥补着人脑的缺陷。但是,人工智能发展毕竟是人类智能发展的组成部分,是人类智能发展的某种实践和体现。

其次,人工智能发展需要物质基础。辩证唯物主义认为人类智能是客观物质世界长期发展的产物,是人脑对客观世界的能动反映。人工智能是对人脑的模拟,其本身也是客观物质世界发展的产物。人工智能发展依赖于人类社会出现的手段、工具和社会条件。人工智能的客观物质性决定人工智能发展只有在一定的物质条件下才能存在和发展。

第三,人工智能没有目的性。辩证唯物主义认为人类智能具有一种主动的目的性。人类活动总是有一定的动机,总是为了满足自身某种主动的需要。人工智能执行的程序也有“目的”,但这个目的不是来自机器本身,而是来自于人。没有满足自身主动的目的性的人工智能只能是人类智能的补充。

第四,人工智能没有能动性。辩证唯物主义认为人类智能不仅能反映事物的现象,更重要的是能够把握事物的本质。人类通过改变世界、改变环境来满足人类发展的需要。人工智能也在不断改变世界,但是人工智能的能动性本质上是人类智能的一种手段、工具,它并不能透过现象发现事物的本质,而只是人类智能能动性的一种表现。

第五,人工智能缺乏个性。辩证唯物主义认为矛盾是事物发展的根本动力,

^①恩格斯.自然辩证法[M].人民出版社.1972年.209页

具有普遍性和特殊性。矛盾的普遍性存在于一切事物的发展过程中，而矛盾的特殊性是指每个事物各自具有其特点。人类实践中对客观世界的反映是人类智能的普遍本质，而各个活动主体各自的特点就凸显出个别性。人工智能以机械或电子运动来模拟人类智能的共同性，但是却不能模拟出每一个个体的人不同的个性。人工智能这种个性的缺失注定它只能随人类智能的发展而发展。

第六，人工智能缺乏创造性。辩证唯物主义认为人类智能的重要特征是其具有创造性，它不仅能解决问题，更能提出问题。尽管人工智能在解决一些问题方面有时有着超越人类智能的表现，但是它始终是按照设定的程序在处理数据。就算处理数据的能力再强，缺乏人类智能创造性的人工智能也只是人类智能的延伸。

综上所述，辩证唯物主义在肯定了人工智能发展的同时，认为人工智能发展要受客观条件限制。人工智能是人类智能的不完全的模拟，是人类智能的延伸和补充，是人类智能的组成部分。人工智能超越人类智能的论调并没有理论和实践作为基础。

第5章 人工智能的发展方向

人工智能发展到现在,已经实现了模拟人脑进行的必然性推理,并且这种能力已经将人类远远抛在身后。即便是小学生手中简单的计算器的运算能力都能大大超过人类,而拥有复杂的高强度数学运算能力的“深蓝”则将棋王变成自己的手下败将。但是对于人类的认知、学习、推理和修正等能够体现人类智能的创造性思维(比如根据人类时时刻刻在根据自然条件主动修正、调整自己行为的能力,亦或在所获信息不够充分的情况下采取的某种尝试性的选择)这方面仍然进展缓慢。人工智能要提高这方面的能力,笔者认为主要需要从以下两个方面研究。

5.1 自然语言的理解

5.1.1 什么是自然语言

人类区别于动物的本质特征之一就是语言,人类智能和语言有着密不可分的联系。斯大林说过:“语言是工具、武器,人们利用它来互相交际,交流思想,达到互相了解。语言是直接和思维联系的,它把人的思维活动的结果,认识活动的成果,用词及由词组成的句子记载下来、巩固起来,这样就使人类社会中思想交流成为可能了。”自然语言是一种自然随着人类历史文化不断发展而产生的语言,人们日常思考学习交流等运用到的都是自然语言(比如汉语、英语、西班牙语等)。相对与人类创造出来的机器语言(如汇编语言、C语言、java语言等),自然语言是人类日常进行交流和思维的主要工具,具有天然的通俗易懂性。同时,交际中的自然语言始终是在某种语言环境下面进行的。语境从狭义来说可以是某一个句子出现时的上下文;从广义上来说包括了如时间、语气、地点、讲话者和听话者所处的不同知识背景、立场,这些都有可能对同样的一句话产生不一样的理解。

通过对自然语言本身进行结构性分析可以发现自然语言有三个不同的层次:语形——由口语中的语音、书面语言中的字、词等这些自然语言中的符号通过串联、排列组合构成;语义——主要是指词语、句子及其它们的组合所要表达的意义和各个语言符号与其所指代的对象之间的关系;语用——把自然语言与其被使用的语境联系起来,即在某个具体的语言环境下分析自然语言所表达的意义。如上面的“我想起来了”这句话,在不同的情况下便有“我想起(某件事情)来了”和“我想(从沙发上)起来了”两种不同的意思。

5.1.2 自然语言逻辑

要实现人机对话,要计算机能够理解人类的自然语言,要理解人类的推理,

需要对人类自然语言进行提炼,并进行逻辑分析,这种分析不能只停留在自然语言的表面,而是要真正理解自然语言所要表达的内涵,其重点应放在对自然语言语用层面的分析上,这也催生了一个新的逻辑学分支——自然语言逻辑。

传统逻辑不考虑具体的思维内容,抛开语言环境,只是对现有的语言材料抽象提取共同的一般性推理,它主要分析的是自然语言的语形和语义层次,其最终目标是将自然语言形式化。例如:所有四川人是中国人;所有成都人是四川人;则所有成都人是中国人,通过对其提炼则可以得到: $A \subset B$; $C \subset A$; 则 $C \subset B$ 。

自然语言逻辑作为逻辑学下面的分支从本质上来说也是一种研究推理的科学,但是它的研究重点与传统逻辑有所不同。它不仅仅是对自然语言的语形和语义进行推理,更重要的是它还要对自然语言的语用层面进行推理分析,它“要以自然语言的语形与语义分析为基础,进一步通过揭示自然语言的使用者与语言符号及其意义之间的三维关系来对自然语言进行逻辑分析,它要揭示在各种不同的语言环境下对语言及其意义理解的诸逻辑问题,它要联系自然语言的具体语言环境来揭示自然语言中的逻辑”^①。因此,自然语言逻辑的核心就是语用推理。

所谓语用推理,有很多学者给出了自己的解释,比如“所谓语用推理,就是在交际过程中,听话人为了准确理解说话人的意图,将话语与语境结合而进行的推理”^②;再如“语用推理是处理自然语言语用关系的逻辑方法。它是在具体语境因素的制约下,由已知话语前提的语法意义推出未知的语用意义结论的过程”^③等,综合来说就是在分析一句话所要表达的意思的时候,要结合这句话所处的语境,而不是简单的从字面上来理解,要抓住这句话所蕴含的真正意义。

语用推理有几个特点:

第一,语用推理有对语境的依赖性。如果一个推理仅从字面和结构本身进行分析而不用考虑它所要表达的意义即可达成,这种推理可以归类为语形推理;如果一个推理需要考虑前提和结论所要表达的意义才可达成,这种推理可以归类为语义推理;如果一个推理需要放在某种具体的语言环境下,需要结合语境因素才能进行分析,则这样的推理就可以归类为语用推理。所以,语用推理是不能离开语境而存在的。

第二,语用推理往往缺少某些前提条件。笔者认为语用推理就是在推理的时候需要根据具体的语言环境补充完善某些在自然语言中被省略掉的前提条件从而得到准确结论的推理。比如“甲说:今天中午我们吃什么好呢?乙回答:昨天晚上去的那家店不错,我们再去一次。”中,甲要通过乙的回答推理出准确的结论,

^①胡泽洪. 语言逻辑与语用推理[J]. 学术研究. 2003 (12)

^②何向东, 袁正校. 隐涵、意图与语旨行为[J]. 探索与争鸣. 1999 (增刊)

^③许毅力. 自然语言的语用推理方法[J]. 哲学研究. 1988 (11)

要准确理解乙所要表达的意思,就必须联系谈话所处的具体语言环境,补充一些前提,比如“昨天晚上我们去吃了某某火锅”,从而得到这样一个推理:“因为昨天晚上我们去吃了某某火锅,今天中午再去一次,所以我们今天中午去吃那一家火锅”。这便是一个语用推理。从上面的例子来看,语用推理是在省略或缺少某种必要的前提的条件下进行的推理,要正确的进行这样的推理,就必须结合当前推理所处的语境,补充进合适的前提,这样才能得到结果。

第三,由于语用推理需要结合当前具体的语境进行,所以它反映的是由具体的和特殊的前提推出的具体的和特殊的结论;而语形推理和语义推理反映的则是由一般的和普遍的前提推出的结论的一般性和普遍性。传统逻辑主要研究是语形推理和语义推理,它只考虑推理的形式和结构,并不分析推理所处的具体语境,因此这样的推理具有一般性、普遍性,是对具体内容的提炼和升华,具有抽象性。例如“所有A都是B,所有C都是A,所以,所有C都是B”这一推理形式,不管处于什么样的语言环境,不管其中的A、B、C所指什么具体内容,这一推理形式都是普遍适用的。语形推理和语义推理中由前提推出结论的过程是明确的,作为演绎推理,只要在推理的过程中遵守了推理规则,该推理就是普遍有效的。语用推理则有所不同,因为在推理中必须加入具体的语境因素,要联系当前推理前提所处的具体环境补充相应的条件来推出结论,因此,语用推理不具有普遍性:每一个语用推理都是某一具体语境中的推理,离开了该具体语境,就无法或不能由前提得出结论;或者在同样的一个前提下,因为语境不同,补所充或增加的前提不同,可以得出不同的甚至相反的结论。比如“我想起来了”,要理解说这句话的人究竟要表达什么意思,就必须在具体的语言环境中进行分析。给这个推理补充一个前提,如果是“在沙发上坐太久了”,则表达的是“我想(站)起来了”;如果补充的前提是“你知道我上次借给你的钢笔放在哪里的吗?”,则表达的是“我想起来(钢笔是放在某某地方)了”。语用推理由于增加了推理过程中具体的语境因素,同一句话在不同的语境下可以形成不同的推理,可以得出不同的结论。因此,语用推理的前提和结论都不具有唯一性和确定性,这样的推理过程也不具有一般性和普遍性。

第四,语用推理具有相对的完整性。从说话者的角度来分析语用推理,会发现它的前提和结论之间的逻辑联系是完整的,是可以得出准确结论的;但是从听者的角度来看,则需要运用具体的语言环境和语用推理规则去分析推理的前提和结论之间的逻辑联系,从而理解说话者所要表达的意思。因此语用推理中所缺少的前提和条件也只是针对听话者而言的,对于说话者可以认为语用推理并不存在。

第五,语用推理进行过程中的交际双方需要有共同的语境背景基础该推理才具有有效性,推理所得出的结论需要交际双方共同认可才有意义。比如“甲说:

今天晚上在哪里见面？乙回答：和昨天一样。”如果甲昨天晚上并没有和乙见面，则不能完成推理，也就无法得到乙所要表达的准确意义。

人工智能发展到现在，在处理自然语言中的语形推理和语义推理方面已经取得了长足的进度，并有很多成果已经投入到使用，极大的方便了人类的活动。比如机器翻译，是利用计算机将一种自然语言翻译成另一种自然语言，从最初只能实现单词之间的翻译，到目前已经能够处理一些较长的句子，但是在准确率上仍然需要提高。目前已经被大量使用的各类智能拼音输入法中很有发展潜力的一个功能就是长句识别功能，即用户只需要输入拼音字母，并不需要音调、字与字之间分隔等便能输入较长的完整的句子，由计算机将拼音字母翻译成汉语。虽然有时仍会出现歧义或别字，但是这表明计算机已经开始尝试理解人类自然语言，是人工智能在理解人类自然语言研究方面取得的重大突破。

如何能够准确理解自然语言所表达的含义是未来人工智能发展中需要解决的难点，这也是目前世界上最火热的研究之一。

5.2 不确定性推理

人类智能的最高表现形式就是创造性，这是建立在必然性推理和基于不完全归纳以及模糊推理等的不确定性推理基础上的，而种种不确定性推理则是人类创造力的源泉。人工智能要实现模拟人类智能，就必须要对不确定性推理进行深入的研究。

5.2.1 不完全归纳推理

首先是不完全归纳推理，这样的推理的前提是有限的，而结论超出了前提的范围，因此这样的推理在前提真的情况下只在一定程度上支持结论，但并不保证结论一定为真，整个推理并不具有必然性，因而它是一种扩展性推理。

具体来讲，这样的归纳大致分为两类：

一、简单枚举法，即在一类事物中，根据已观察到的部分对象都具有某种属性，并且没有遇到任何反例，从而推出该类事物都具有该种属性的结论。其逻辑形式为：

S₁ 是 P

S₂ 是 P

S₃ 是 P

.....

S_n 是 P

S₁, S₂, S₃,, S_n 是 S 类的部分对象，并且其中没有 S 不是 P

所以，所有 S 是 P

简单枚举法是归纳方法中最为简单的一种，但是其在指导人类认识改变世界的活动中的意义是无法忽视的，它有助于在有一定数量的样本材料，但未能找到充分根据的时候对事物做出一个初步的概括，推出一个或然性的结论，而这样的结论可以作为下一步研究的起点。在进行研究时产生的各种假说往往都用到了简单枚举法，比如著名的“哥德巴赫猜想”——“所有大于 5 的奇数都可以分解为三个素数之和”显然，要列出所有大于 5 的奇数是不可能的，这只能根据有限的样本数量提出；我国民间谚语中的“瑞雪兆丰年”是劳动人民根据在年复一年的劳动中所获得的经验的总结，是简单枚举法在人们实际生产劳动中的具体运用之一。简单枚举法作为一种扩展性推理，为保证推理过程的可靠性，必须注意要尽量扩大样本数量和考察有无反例的存在。人们不能因为甲脸上有麻子、乙脸上也有麻子就断定所有人脸上都有麻子，这显然是不合理的。因此扩大考察样本的范围和数量对简单枚举法有重要的意义，但是这样大量的重复劳动费时费力，因此简单枚举法在实际应用中有着它不可避免的缺陷。

二、科学归纳法，它是以科学分析为主要依据，由某类中部分对象与其属性之间所具有的客观因果联系，推出该类的全部对象都具有某种属性的归纳推理。其形式为：

S1 是 P

S2 是 P

S3 是 P

.....

Sn 是 P

S1, S2, S3,, Sn 是 S 类的部分对象，它们与 P 之间有因果联系

所以，所有 S 是 P

科学归纳法和简单枚举法都属于不完全归纳推理，在前提中都只是针对一类事物中的有限部分对象进行考察，结论则都是对一类事物全体的断定，结论的知识范围超出了前提。同时，科学归纳法与简单枚举法也有不同的地方：首先是两者使用不同的推理根据。简单枚举法的推理仅仅根据已考察到的有限部分对象都具有的某种属性，并且考察中没有遇到任何反例就做出的推理；科学归纳推理则不仅仅停留在对认识事物的经验的简单重复上，而是深入所考察事物内部进行的科学分析，试图找出考察对象与某种属性之间内在客观因果联系，在此基础上进行推理，从而得出结论。其次是两者的前提数量的意义不同。对于简单枚举法的归纳推理来说，其结论的可靠性是由前提中考察对象的范围和数量决定的，理论上前提中考察的对象范围越广、数量越多，结论就越可靠。但是简单枚举法的考

察对象不可能是无穷数量,所以永远无法得到事物完整的认识。对于科学归纳推理来说,其结论的可靠性并不基于前提中考察对象的范围和数量,即便前提考察对象的范围不大、数量不多,甚至在只考察一、两个典型的情况下,只要充分认识考察对象与某种属性之间的客观因果联系,也能推理出可靠的结论。第三是科学归纳法的结论较简单枚举法的结论更具可靠性。虽然它们都属于或然性推理,结论范围都超出了前提,但是科学归纳法重点考察的是考察对象与属性之间内在的客观因果联系,所以科学归纳法的推理强度比简单枚举法的要大,这也就意味着科学归纳法比简单枚举法的结论的可靠性大。正是由于科学归纳法的重点是放在考察对象与属性之间内在客观因果联系,所以它在人类探究事物本质,发现事物运行内在规律方面有重要作用。

虽然由不完全归纳法所推理出的结论没有必然性,但是其在人类认识世界,改造世界时的作用仍然无法忽视。首先,人类在认识世界的时候没有能力在了解了整个世界的运行规律以后再进行活动,必然是在了解局部之后从这个局部得到的某种经验的重复中来建立某种规律性的认知,举一反三说的正是这个道理。其次,运用不完全归纳推理指导人类实践具有它的必要性和必然性。人类有理由从经验的不断重复中依据确认世界上存在某种类似与客观因果规律来建立起对事物的一种确定性和规律性的认识。第三,人类有能力在有限的范围内建立起合理的归纳方法和归纳逻辑,人工智能的发展已经在慢慢实现归纳逻辑的这种可能性,而恩格斯很早就指出了“社会一旦有技术上的需要,则这种需要比十所大学更能把科学推向前进。”^①第四,人类实践的成功也在一定程度上证明了不完全归纳推理的作用,证明了其真理性。

“人工智能在发展中为了在机器的智能模拟中克服对归纳模拟的困难而有所突破,应该将归纳逻辑等有关的基础理论与机器学习、不确定推理和神经网络学习模型与归纳学习中已有的成果结合起来只有这样,才能在已有的归纳学习成果上,在机器归纳和机器发现上取得新的进展和突破”^②。

5.2.2 模糊逻辑

第二是模糊逻辑,即是研究模糊概念、模糊命题和模糊推理等的逻辑理论,是一种建立在多值逻辑的基础之上,运用模糊集合的方法来研究模糊思维、语言形式及其规律的学科。人类所面临的世界到处都存在着模糊的现象,这些现象经过人类感官传达到大脑后便在人类的思维里形成了各种各样模糊的概念和命题,比如“这里离学校已经很近了”、“这个人很高”、“这杯水很冷”、“在电影院附近

^①马克思恩格斯选集[M].第4卷.人民出版社.1995年.505页

^②王雨田.归纳逻辑和人工智能[M].中国纺织大学出版社.1995年.1-2页

碰头”等。传统的经典二值逻辑在处理这些命题的时候显得力不从心，而这些问题又是人工智能发展中不得不面对的，模糊逻辑正是在这样的背景下应运而生。

模糊逻辑首先是由美国加州大学伯克利分校的数学家扎德 (L. A. Zadeh) 提出的。1965 年他在《信息与控制》上发表了论文《模糊集合》，提出了模糊集合的概念，给其的定义为：

设 X 是由点或对象构成的一个空间, X 的一个类属性元素用 x 表示, 则有 $X = \{x\}$ 。在 X 中, 一个模糊集 A 由一个具有构成集合元素资格的函数 $f_A(x)$ 进行特征描述。该函数与 X 中的每一个点, 即区间 $[0, 1]$ 中的一个实数相联系。 $f_A(x)$ 在 x 上的值表征集合 A 中 x 所具有的构成 A 的元素资格的程度。当 A 在其所表示的一般意义上构成一个集合时, 其元素资格的函数就只可能有两个值, 即 0 和 1: 当 x 属于 A 时, $f_A(x)=1$; 当 x 不属于 A 时, $f_A(x)=0$ 。因此, 在这种情况下, $f_A(x)$ 就简化为人们所熟悉的集合 A 的特征函数。但是, 如果建立一种新的表示限制性的关系, 即 $0 < \alpha < 1$, $0 < \beta < 1$, $\alpha > \beta$, 那么一种三值逻辑的可能性就是成立的: (1) “ x 属于 A ” 如果 $f_A(x) \geq \alpha$; (2) “ x 不属于 A ” 如果 $f_A(x) \leq \beta$; (3) “ x 具有一种相对于 A 的中间状态” 如果 $\beta < f_A(x) < \alpha$, 这样就导致了一种具有三种真值的逻辑的出现, 即: $T(f_A(x) \geq \alpha)$, $F(f_A(x) \leq \beta)$, 以及 $U(\beta < f_A(x) < \alpha)$ ^①。

扎德认为“把真假作为语言变量处理得到模糊语言逻辑或简称模糊逻辑”^②, 他创造性的提出用隶属度表示真实度来代替经典二值逻辑中用布尔值表示真假的方法。这一理论为模糊逻辑的发展奠定了基础。

相对于传统的经典二值逻辑中一个命题非真必假、非假必真, 没有“半真半假”这样的中间值的可能, 模糊逻辑做出了根本性的修正, 肯定了中间值的存在及其意义。在模糊逻辑的视角下, 为了更准确的表征认知主体实际推理或者应该推理的方式, 一个命题不仅能够被推出“为真”或“为假”的结论, 还可以推出这个命题“非常接近于为真”、“几乎为假”, 或者还可以推出这个命题相对于和它相关的某个命题“在更大程度上”“为真”或者“为假”。模糊逻辑推广了经典的二值逻辑, 可以具有无穷多个中间状态, 成为了处理具有非精确性和非确定性问题的有效工具。人类思维中本身就存在其模糊性, 模糊技术以此为出发点, 运用模糊逻辑的方法, 对实践中的所遇到的模糊信息进行量化, 其中最重要的也是最困难的一步是利用专家知识和实际经验来定义相应模糊集的隶属函数, 即通过设定某种隶属函数来描述自然语言中的非确定成分, 并且通过作用于这种隶属函数的数学算子来分析量化语言中语气的轻重。因此“模糊集合论将经典集合论中用点刻划客观对象的方法推广为用几何曲线来描述客观对象集中的基本元素, 这无

^①L. A. Zadeh. Fuzzy Sets[J]. Information and Control. 1965 (8)

^②L. A. Zadeh. Fuzzy Sets[J]. Information and Control. 1965 (10)

疑将使模糊集的描述能力同经典集合相比有本质的提高”^①。在模糊理论研究中最基本的研究对象是隶属函数，它是依靠专家知识和实际经验来确定的，这其中不可避免的包含有主观的因素，但在这里主观因素并不会导致依此建立的理论没有可靠性。与此相反，正是由于在理论的建立中利用到主观性，模糊思维才能准确的反映人脑的思维特征，从而使得模糊逻辑理论在人工智能的发展中取得了非常成功的应用。目前模糊逻辑在自动控制、医疗诊断、气象预测等方面有比较广泛应用和重要的成果。

^①刘普寅，李洪兴. 软计算及其哲学内涵[J]. 自然辩证法研究. 2000年第5期第16卷. 27页

第6章 结语

人工智能自从上世纪 50 年代被提出以来,其经历的发展可用日新月异来形容。人工智能发展至今,从其取得的举世瞩目的成就可以看出,人工智能是人类智能发展的结晶,是人类智能进步道路上迸发出的美丽的火花。它通过模拟人类智能认识世界、改变世界的方式,极大的延伸和扩展了人类的能力,成为人类生活生产过程中不可或缺的一份子。

本文是以人工智能发展的问题作为研究对象而完成的。通过对大量相关文献的分析并结合实例,在综合前人研究成果的基础上,运用辩证唯物主义世界观和方法论,把人类智能和人工智能互动的发展以整体形式加以研究,从分析人类智能产生和发展入手,探讨人工智能发展的目的和价值,梳理人类智能和人工智能的相互关系,从而认识到人工智能是人类智能的模拟与延伸,是人类智能的组成部分。同时,通过对人工智能最新研究前沿的分析,发现人工智能目前发展中在哲学领域里所缺失的理论支撑,通过人类自然语言 and 不确定性推理两方面的分析,笔者认为今后人工智能的发展必将走向对自动进行推理运算和自我学习的研究;另外在处理自然语言方面目前已经取得了一定的进展,今后这方面应用有着广大的前景,必然也是人工智能发展的重点。

参考文献

- [1] (美) 哈里·亨德森. 人工智能——大脑的镜子[M]. 候然译. 上海科学技术文献出版社. 2008 年
- [2] (美) 休伯特·德雷福斯. 计算机不能做什么[M]. 宁春岩译. 三联书店. 1986 年
- [3] (美) 雨果·德·加里斯. 智能简史——谁会取代人类成为主导物种[M]. 清华大学出版社. 2007 年
- [4] 蔡曙山. 哲学家如何理解人工智能——塞尔的“中文房间争论”及其意义[J]. 自然辩证法研究. 2001 (11)
- [5] 蔡自兴. 人工智能对人类的深远影响[J]. 高技术通讯. 1995 (6)
- [6] 童天湘. 从“人机大战”到人机共生[J]. 自然辩证法研究. 1999 (9) 第 13 卷
- [7] 戴汝为. 从基于逻辑的人工智能到社会智能的发展[J]. 复杂系统与复杂性科学. 2006 (6) 第 3 卷第 2 期
- [8] 王雨田. 归纳逻辑与人工智能相结合的研究问题[J]. 哲学研究. 1992 年第 4 期
- [9] 郑祥福. 人工智能的四大哲学问题[J]. 科学技术与辩证法. 2005(10) 第 22 卷第 5 期
- [10] 盛晓明、项后军. 从人工智能看科学哲学的创新[J]. 自然辩证法研究. 2002(2) 第 18 卷第 2 期
- [11] 刘普寅、李洪兴. 软计算及其哲学内涵[J]. 自然辩证法研究. 2000(5) 第 16 卷第 5 期
- [12] (美) 雷·库兹韦. 灵魂机器时代——当计算机超过人类智能时, 世纪出版集团上海译文出版社[M]. 2006 年 4 月
- [13] 恩格斯. 自然辩证法[M]. 人民出版社. 1972 年
- [14] 赵敦华. 西方哲学简史[M]. 2006 年
- [15] 约翰·华生 (John B. Watson). 行为主义心理学[M]. 李维译. 浙江教育出版社. 1998 年
- [16] 马克思. 1844 年经济学——哲学手稿[M]. 人民出版社. 1972 年 [17] (英) 玛格丽特·A·博登编著. 人工智能哲学[M]. 刘西瑞, 王汉琦译. 上海译文出版社. 2005 年
- [17] 瞿心昱. 基于仿人脑认知计算模型的机器人视觉学习方法[D]. 浙江工业大学博士学位论文. 2012 年
- [18] (美) 弗里德曼 (David Freedman). 制脑者: 创造堪与人脑匹敌的智能[M]. 张陌, 王芳博译. 三联书店. 2001 年
- [19] 苗力田主编. 古希腊哲学[M]. 中国人民大学出版社. 1989 年
- [20] 赵敦华. 西方哲学简史[M]. 北京大学出版社[M]. 2006 年 7 月
- [21] 苗力田主编. 古希腊哲学[M]. 中国人民大学出版社[M]. 1989 年
- [22] (美) 雷·库兹韦. 灵魂机器时代——当计算机超过人类智能时. 世纪出版集团上海译文出

- 版社[M]. 2006 年
- [23] (法) 拉·梅特里[M]. 人是机器. 商务印书馆. 1981 年
- [24] 马克思恩格斯全集[M]. 第 42 卷. 人民出版社. 1979 年
- [25] 吴文俊. 王者之路——机器证明及其应用[M]. 湖南科技出版社. 1999 年
- [26] 西方哲学原著选读(下卷)[M]. 商务印书馆. 1981 年
- [27] (英) 玛格丽特·A·博登编著. 人工智能哲学[M]. 刘西瑞, 王汉琦译. 上海译文出版社. 2005 年
- [28] 胡泽洪. 语言逻辑与语用推理[J]. 学术研究. 2003 (12)
- [29] 何向东, 袁正校. 隐涵、意图与语旨行为[J]. 探索与争鸣. 1999 (增刊)
- [30] 许毅力. 自然语言的语用推理方法[J]. 哲学研究. 1988 (11)
- [31] 马克思恩格斯选集[M]. 第 4 卷. 人民出版社. 1995 年
- [32] 王雨田. 归纳逻辑和人工智能[M]. 中国纺织大学出版社. 1995 年
- [33] L. A. Zadeh. Fuzzy Sets[J]. Information and Control. 1965(8)
- [34] 刘普寅, 李洪兴. 软计算及其哲学内涵[J]. 自然辩证法研究. 2000 年第 5 期第 16 卷

致 谢

论文几经易稿，能够写到这一部分实属不易。本文选题及研究过程中得到唐晓嘉教授的悉心指导，唐老师多次询问研究进程，并为我指点迷津，帮助我开拓研究思路，她的精心点拨、热忱鼓励使我能够顺利完成论文。在此期间，唐老师更让我感受到了一个学者一丝不苟的作风，严谨求实的态度。虽历时三载，却给以终生受益无穷之道。对唐老师的感激之情是无法用言语表达的。

同时，还要感谢何向东教授、邹顺康教授、王静教授还有已故的彭自强教授等老师对我的教育培养。他们细心指导我的学习与研究，在此，我要向诸位老师深深地鞠上一躬。

感谢我的同学彭语良、袁晖三年来对我学习、生活的关心和帮助。

最后，向我的家人致谢，感谢他们对我的理解与支持。

程 石

于川北医学院