

分类号 N02

学号 09081039

UDC

密级 公 开

哲学硕士学位论文

人工智能中的不确定性问题研究

硕士生姓名 张昕

学 科 专 业 科学技术哲学

研 究 方 向 军事技术哲学

指 导 教 师 刘戟锋 教授

国防科学技术大学研究生院

二〇一一年十一月

人工智能中的不确定性问题研究

国防科学技术大学研究生院

The Research on Uncertainty of Artificial Intelligence

Candidate: Zhang Xin

Advisor: Liu Jifeng

A dissertation

Submitted in partial fulfillment of the requirements

for the degree of Master of Philosophy

in Philosophy of Science and Technology

Graduate School of National University of Defense Technology

Changsha, Hunan, P.R.China

(November, 2011)

独 创 性 声 明

本人声明所呈交的学位论文是我本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表和撰写过的研究成果，也不包含为获得国防科学技术大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文题目： 人工智能中的不确定性问题研究

学位论文作者签名： 张 明 日期： 2012 年 2 月 12 日

学位论文版权使用授权书

本人完全了解国防科学技术大学有关保留、使用学位论文的规定。本人授权国防科学技术大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档，允许论文被查阅和借阅；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。

（保密学位论文在解密后适用本授权书。）

学位论文题目： 人工智能中的不确定性问题研究

学位论文作者签名： 张 明 日期： 2012 年 2 月 12 日

作者指导教师签名： 刘 斌 祥 日期： 2012 年 2 月 12 日

目 录

摘 要.....	i
ABSTRACT	ii
第一章 绪论.....	1
1.1 选题依据.....	1
1.2 研究方法和意义.....	3
1.2.1 研究方法.....	3
1.2.2 研究意义.....	3
1.3 国内外研究.....	5
1.3.1 人工智能研究历程.....	5
1.3.2 国内外研究现状	6
1.4 主要内容和创新点	9
第二章 不确定性在人工智能中的重要地位	10
2.1 传统人工智能研究存在不确定性缺失的问题.....	10
2.1.1 符号主义、联结主义、行为主义面临的困境	10
2.1.2 智能衡量标准难以确定.....	14
2.2 不确定性是真实世界的本质属性	16
2.2.1 确定性是对真实世界的近似刻画	16
2.2.2 不确定性辩证地反映了真实世界	18
2.2.3 不确定性主要体现为时间和空间上的连续性	20
2.3 不确定性的引入：人工智能的新发展	21
第三章 人工智能逻辑基础的不确定性	24
3.1 逻辑的基础地位.....	24
3.2 刚性逻辑不能全部反映人类思维的本质.....	25
3.3 包容不确定性的柔性逻辑对人工智能的促进.....	26
3.3.1 模糊集和粗糙集在不确定性推理方面丰富了形式逻辑的内容	26
3.3.2 非单调逻辑在动态更新方面弥补了形式逻辑的不足.....	27
3.4 柔性逻辑需要进一步发展和完善	27
第四章 人工智能研究方法的不确定性.....	31
4.1 智能内在生成机制的研究方法探悉	31
4.1.1 信息获取的不确定性.....	32

4.1.2 主体认知的不确定性.....	34
4.1.3 形成知识的不确定性.....	37
4.1.4 主体决策的不确定性.....	38
4.1.5 执行结果的不确定性.....	39
4.2 面向物质基础的研究方法是人工智能的突破口	39
4.2.1 传统计算机的物质基础.....	41
4.2.2 人工智能的发展催生新型计算机的诞生	42
4.2.3 通过在硬件搭建中引入不确定态来实现硬不确定性人工智能	43
4.3 面向软件模拟的研究方法是人工智能的保证	44
第五章 人工智能技术开发的不确定性.....	46
5.1 技术开发中的风险和不确定性	46
5.2 技术开发过程何以开展.....	47
5.2.1 内在运行机理的不确定性导致技术开发困难	48
5.2.2 外生变量的不确定性导致技术开发困难.....	49
5.3 技术开发何以评判	51
第六章 人工智能应用前景的不确定性.....	52
6.1 人工智能发展的社会关系之辨	52
6.1.1 机器与机器的社会关系.....	52
6.1.2 机器与人的主体地位之争	54
6.2 人工智能发展引发的伦理问题	55
6.2.1 智能机器是否会导致人类劳动力过剩.....	55
6.2.2 智能机器人是否等同于人	56
6.2.3 智能机器人与人交战是否会引入新的战争模式	57
结 束 语	59
致 谢.....	61
参考文献	63
作者在学期间取得的学术成果.....	66

表 目 录

表 2.1 符号主义、联结主义、行为主义的区别	11
表 3.1 模糊集和粗糙集的比较	26
表 4.1 随机信息、不完备信息、中间态信息的区别.....	34
表 6.1 单体机器人和群机器人的区别.....	53

图 目 录

图 2.1 非计算机难解问题在思维中所占的比例 12

图 2.2 图灵测试 14

图 2.3 中文屋子 15

图 4.1 智能主体之间的简要关系..... 32

图 4.2 旋转圈 33

图 4.3 黑白点 33

图 4.4 缪勒·莱依尔错觉 35

图 4.5 菲克错觉 35

摘 要

不确定性问题是人工智能研究的重要方面，客观地反映了真实世界的本质特性。传统的人工智能研究存在不确定性缺失的问题，阻碍了人工智能的进一步发展。

论文选定人工智能中的不确定性问题作为研究对象，首先从哲学角度进行了阐述，认为不确定性体现为时间和空间上的连续性。接着分别从逻辑基础、研究方法、技术开发和应用前景四个方面展开了较为深入的研究。在逻辑基础方面，认为包容不确定性的柔性逻辑将对人工智能的发展起到促进作用；在研究方法方面，本文补充和完善了当前在理论上被普遍接受的“机制主义”研究方法，并创新地提出了从硬件上引入不确定态的 $\{0,1,\alpha\}$ 的三态模型以及在软件模拟中考虑到社会性因素是新型智能机发展的重要方向；在人工智能技术开发方面，认为内在运行机理的不确定性导致开发困难，而外界环境的不确定性导致过程控制困难；在人工智能应用前景方面，展望了随着人工智能的发展可能引起的社会关系改变和一系列可能遇到的伦理问题，为人工智能的发展提供了新的思路。

关键词：不确定性；人工智能；逻辑基础；研究方法；技术开发；应用前景

ABSTRACT

The uncertainty problem is an important aspect of artificial intelligence research, which objectively reflects the nature feature of the world. Traditional research on artificial intelligence has the issues of missing uncertainty problem, hindering further development of artificial intelligence.

This thesis chooses the uncertainty problem in artificial intelligence as the study object. First, we elaborated that uncertainty reflects the continuity of time and space from a philosophical point of view. Then, we carried out an in-depth study on the uncertainty problem in artificial intelligence from four aspects: logic basis, research method, technology development, and application prospect. On the logical basis of artificial intelligence, we argue that the development of uncertainty-inclusive flexible logic will play a catalytic role for artificial intelligence; on the research methods of artificial intelligence, we supplement and perfect the widely accepted “mechanism” theory and try to put forward that the introduction of uncertain state into the hardware, resulting in $\{0,1,\alpha\}$ tri-state model, and the consideration of social factors into software simulation will be important directions for the development of new intelligent machines; on the technology development of artificial intelligence, we argue that the uncertainty of inner execution mechanism leads to the difficulties of development and the uncertainty of the external environment leads to difficulties of process control; on the application prospects of artificial intelligence, we introduce that the development of artificial intelligence may change the social relationship and encounter ethical problem. This provides a new idea for the development of artificial intelligence.

Key Words: uncertainty; artificial intelligence; logic basis; research method; technology development; application prospect

第一章 绪论

一切思维不过是计算（即加加减减）。——霍布斯

用符号语言与运算可以表示任何事物。——布尔

我们可以把宇宙目前的状态看作是过去的结果和将来的原因。如果一个有理性的人知道某时刻生物界的一切力和所有生物的相互位置，而他的才智又足以分析一切资料，那么它就能用一个方程式表达宇宙中最庞大的物体和最轻微的原子运动。对他来说，一切都是确定的，将来与过去都呈现在他眼前。

——拉普拉斯

爱因斯坦坚信物理学能够给我们关于客观存在的世界的知识。作为原子的量子现象领域中经验的结果，我和别的许多物理学家一起，逐渐转变到另一种观点上来，认为情况并不是那样的。我们认为，在任何既定时刻，关于客观世界的知识只是一种粗糙的近似。

——波恩

不确定性是我们在其中生活的宇宙的一个基本特征。——霍金

1.1 选题依据

论题的选择源于刘戟锋教授提出的一个有创造性的观点：“人工智能的发展——从形而上学到辩证法”。把人工智能的发展与形而上学和辩证法相联系，粗略看来有些难以理解，但仔细揣摩，这个提法是耐人寻味的。

随着科学技术的发展，人类逐渐形成了世界是物质的，物质是运动的，运动是可认识的观念，其中能够被认识的部分体现为确定性，而暂时无法认识或尚未被认识的部分，则展现出不确定性的状态。对于真实世界而言，不确定性是事物的本质属性，确定性只是其中的特例。在学习辩证法后，人们往往排斥形而上学，事实上，形而上学的方法本身是很有用的，为了认识事物，人类必须通过形而上学的方法将事物做静止的、确定性的近似处理。但是，这并不意味着在世界观上承认真实世界是静止的，是确定性的，而这正是形而上学的局限，它将一种近似的处理方法推广到了对一切事物的认识上，而辩证法恰好在这方面有所弥补，既包容了形而上学的处理方法，又在世界观上承认事物的运动性。

笔者认为不确定性体现为时间和空间上的连续性，是真实世界的本质属性。在人工智能研究中，不确定性起到了比较重要的作用。首先从人工智能的逻辑基础上看，传统的计算机基于冯·诺依曼体系结构，存在着一维性、串行处理和形式化等特征，很难处理非数值的自然语言方面的信息或模拟人类的形象思维。这是因为传统计算机的逻辑基础是刚性逻辑，采用了非 0 即 1，非“开”即“关”，

非此即彼的确定性的处理方式。而未来的计算机，无论是量子计算机还是生物计算机，都把重心放在了对不确定性问题的研究和实现上，其逻辑基础是柔性逻辑或泛逻辑，不再排斥模糊态和中间态，而是努力通过硬件或软件的方法来完成对不确定态的计算，模仿并刻画人类的形象思维。

其次从人工智能的研究方法来看，以符号主义、联结主义、行为主义为代表的传统的人工智能研究方法都是基于对确定性问题的研究，虽然相继取得了一系列的成果，但由于不确定性缺失的问题，无一例外的在模拟人类形象思维等方面陷入了困境，并且几十年来，不同方法仍然难以融合统一。随着不确定性走进人工智能领域，首先在理论上有所突破，“机制主义”研究方法从智能内在机制出发，把握到了不同方法的联系，并将三大主流方法统一到了同一个理论框架中。由此可见，不确定性和确定性是辩证统一的，不能强行分割。人工智能领域要想统一并真正有所突破，必须寻求一种包容确定性和不确定性的研究方法。

此外，从人工智能的技术开发来看，随着制作工艺越来越精良，技术手段越来越高超，机器的性能已经可以达到千万亿次、万万亿次的水平，出现了一系列经典的优化算法，使得基于确定性基础的传统人工智能的研究已经登峰造极，几乎达到了基于冯·诺依曼体系结构的经典计算机在存储、计算、搜索等方面的极限，要想在技术方面进一步寻求突破不但困难，而且会耗费极大的人力和物力。未来的人工智能的发展必须站在新的视角，开拓新的领域来实现质的飞跃，因此对不确定性问题的研究无疑成为重点。由于人工智能中的不确定性问题研究与传统的人工智能研究有较大的不同，无论是硬件搭建还是软件模拟都需要进行较大的改动，呼吁新兴技术的诞生。相信不久以后，新技术和新方法将如雨后春笋般层出不穷，并和原有技术、成果相得益彰，进一步丰富和发展人工智能。

综上所述，从逻辑基础、研究方法、技术开发这三个重要的方面来看，人工智能的发展很可能经历从研究人工智能中的确定性问题发展到研究人工智能中的不确定性问题的过程。在人工智能诞生之初的二三十年时间里，为了能用科学的手段更好的研究人工智能，往往对研究对象进行确定性的处理，忽略了其时间和空间上的连续性。虽然在研究过程中运用形而上学的方法对研究对象做出确定性的近似可以简化问题，但并不能解决所有的问题，随着研究的深入将遇到无法逾越的障碍，要想破冰前进，需要在世界观上承认事物的运动性、变化性和中间态，也就是说在世界观上辩证地看待事物，而不是一成不变的看待事物。因而，站在哲学的高度，“人工智能的发展——从形而上学到辩证法”是一个新颖而不失道理的论断。

但是，如果论文选题拟定为：“人工智能的发展——从形而上学到辩证法”，将使得题目范围比较广，需要相当厚实的哲学底蕴和相关领域技术作为支撑，理

解难度大，驾驭起来非常困难。因此，在选题的时候，放弃了原本让人眼前一亮的题目，而选取了与本科所学的计算机专业结合较紧密，在理论上把握相对容易，国内外具有相关研究支撑的题目，将人工智能和不确定性综合起来把握，通过对人工智能中的不确定性问题研究来阐明刘戟锋教授的观点。

1.2 研究方法和意义

1.2.1 研究方法

为了更好地阐述观点，论文主要采取了分类研究和案例研究两种研究方法。分类研究方法的好处在于使得文章结构清晰，条理清楚，让人一目了然地看到整体和部分之间的联系，分而治之更好地处理问题。分类研究的方法在论文脉络组成上有比较明显的体现：论文从一开始就通过分类的思想将人工智能研究分为了基于确定性基础的传统人工智能研究和基于不确定性基础的新兴的人工智能的研究，由此引入了人工智能中的不确定性问题研究这一论题；接着从哲学上将人工智能中的不确定性问题归纳为时间和空间上的连续性；然后在具体研究过程中，将不确定性问题分解为逻辑基础的不确定性、研究方法的不确定性、技术开发的不确定性和应用前景的不确定性四类，按照从整体到部分的思想逐类进行探讨，从而较为系统地展开研究。

案例研究方法的好处在于论有据、说有理，能为论文提供有力的支撑，并且通过将抽象的观点、概念具体化，有助于更好地理解和掌握相关知识。论文在内容组织上多处使用了案例研究法，主要在三、四、五、六章中有所体现。以第四章为例，为了更好地阐述智能的内在生成机制，论文以被国内外广泛接受的“机制主义”研究方法为主例，以日常生活中人们熟知的事例作为辅例，细致分析了信息获取的不确定性、主体认知的不确定性、形成知识的不确定性、主体决策的不确定性和执行结果的不确定性。类似的，在其余各章，论文都有针对性地添加了生动的事例，或者通过具体的案例来解释相对抽象的理论，使得文章更加翔实，即使没有相关知识背景的读者也能够轻松阅读。

1.2.2 研究意义

传统的人工智能研究建立在确定性的基础上，虽然在研究初期取得了一系列成果，但随着技术的发展和研究的深入，当前的研究成果并没有使机器拥有人类所期待的智能。究其原因在于人类的思维除了可形式化的逻辑思维外还包括不可形式化的形象思维。当前的计算机虽然拥有强大的计算、搜索和存储能力，但对不可形式化的问题却难以下手。因此，要想解决人工智能遇到的困难，需要在研

究中考虑到真实世界中的不确定性因素，把确定性研究和不确定性研究有机统一起来，制造出真正“智能”的机器。

人工智能作为一门前沿的学科，吸引着各国多领域专家的重视，有着深远的研究意义。人的能力通常可以划分为体能和智能两类，基于这样的分类，人类制造或运用的工具也随之划分为两类，分别是辅助提高体能的工具和辅助提高智能的工具。伴随着人类的发展，劳动从纯粹的体力劳动到出现了体力劳动和脑力劳动的分工，再到开始更多地偏向脑力劳动。同时，劳动工具开始从辅助提高体能的工具逐渐向辅助提高智能的工具过渡的过程，而人工智能无疑大大地推进了这个过程。

之前，人类的绝大多数研究成果都偏向于对体能的提升，从而在较大程度上替代人类的体力劳动，虽然人类也希望能够制造出替代智力劳动的工具，并且早在几千年前就开始对大脑进行了探索，但迫于条件的限制，研究进展缓慢。直到人类研制出了计算机这一伟大的劳动工具，才为进一步研制人类智能的替代品提供了基础上的保证。基于冯·诺依曼体系结构的计算机在相应软件的支持下，通过程序员预先编写的算法，施展出了强大的搜索、计算、存储等方面的能力。因此，人类开始期待，能否通过对机器的不断改进，让机器像人一样拥有智能，从而不仅仅能够替代人类的体力劳动，还可以分担一部分智力劳动，对人工智能的研究也因此产生。“人工智能”这一正式的提法诞生于 1956 年的达特茅斯会议上，从此，人工智能作为一个新兴的前沿领域得到了来自不同国家的多个领域的专家的重视，这是因为人工智能不仅仅能带来丰硕的经济效益，还将给国家安全带来重大的影响。

军事技术体系的核心是武器技术，刘戟锋教授在《军事技术论》一书中，首次提出了军事技术的四要素包括了打击力、防护力、机动力和信息力^[1]，而人工智能的相关研究成果将可以覆盖这四个要素，全面提高一个国家的军事技术水平。例如：在打击敌方运动目标时，普通的导弹可能发射数枚才能击中，并且还存在着一击不中丧失战机的情况。但安装了智能制导的导弹将能跟踪追击目标，直到将目标击落，大大提高了打击力；装备了人工智能柔性外骨骼的士兵，不但移动速度更加快，疲惫感、恐惧感大大减轻，还可以承担普通人百倍以上的攻击，将大大减少战场伤亡率，使得防护力和机动力都得到了提高；智能苍蝇将以假乱真的飞入敌方重要场所录音或摄像，获取机要文件，这将在一定程度上提高信息力。

随着计算机的普及和核武器、生化武器、声波武器、动能武器等一系列毁伤效应大、争议也较大的新型武器的问世，人类战争将经历从物理战到心理战的转变^[2]，新型武器将作为威慑性武器对其他国家产生心理上的威慑，因此，即使不能真正在战争中使用，也可以通过拥有来起到威慑的作用，不战而屈人之兵。以计

算机为基础的人工智能相关技术有着先天具备的信息力优势，因此无论开展的是物理战还是心理战，只要能够运用到信息的地方，就有人工智能的用武之地。

人工智能的研究有着重要的意义，并取得了一系列的研究成果。但不容忽视的是，传统的人工智能研究基于确定性的基础，然而在真实世界中，确定性和不确定性是同时存在、辩证统一的，因此以符号主义、联结主义和行为主义为代表的研究都因为不确定性缺失的问题在经历短时间的辉煌后进入了瓶颈状态。因此，要想进一步深入研究人工智能，需要对人工智能中的不确定性问题进行探索，从一个新的角度对原有的研究进行补充，进一步丰富和完善人工智能的内容，突破单纯确定性研究无法逾越的障碍。

本文选取了对人工智能中的不确定性问题进行研究，首先在选题上具备一定的前沿性和新颖性，特别是从哲学入手对不确定性问题进行了系统的分析和讨论，对相关领域前期的工作进行了梳理，并提出了个人的见解和想法，丰富和完善了人工智能理论，有一定的启发性和开拓性。

1.3 国内外研究

人工智能作为一门新兴的学科，从真正建立到发展的时间不过 50 余年，并且在研究过程中，与哲学、脑科学、计算机技术、心理学、认知学等学科都有着较为紧密的联系。因此在研究过程中，国内外专家都习惯于将人工智能归纳到智能科学与技术学科中，而不是作为孤立的研究对象，以便更好地展开相关工作。

1.3.1 人工智能研究历程

人类对智能学的研究可以追溯到 4 万年前原始人在洞穴壁上留下的传递信息的符号以及结绳记数的估算方法，并在公元前 3500 年诞生了最早的智能工具——算盘。公元前的古文明虽然在今天看起来比较落后，但在思维方面却很活跃，中西方的哲人纷纷借助有限的知识，尝试对智能进行了思考和探究，碰撞出思想的火花。例如公元前 770——公元前 221 年间，先秦诸子开始对逻辑展开研究，并通过“辩术”进行语言的交锋，探讨概念间的异同，类的区分以及矛盾的关系，涌现出了以惠施、公孙龙、墨翟、荀况为代表的思想家，呈现出了百家争鸣的学术盛况。与之同时在西方，一方面以阿尔克梅翁、希波克拉底、赫罗菲拉斯、埃拉西斯特拉图斯为代表的医学家、生理学家通过解剖等方法从自然科学角度对大脑展开了研究；另一方面，以苏格拉底、柏拉图、亚里士多德为代表的哲学家发现了与智能息息相关的“数”，并通过理性的思考，为善恶、正误等寻求确定性的标准并建立了逻辑学，运用了矛盾律、排中律和归纳法对范畴和命题展开了研究。因此，站在智能学研究的起点来看，中国古代和西方相比，虽然研究方法不如西

方系统、严谨，并且缺乏实证性的手段，但是在思想上却各有千秋，几乎是同步发展。可是，这一良好的发展局面并没有持续多久，因为政治、军事等多方面的原因，在该领域的发展停滞了很长的时间，直到近代才复苏。但从近代开始，生产过最早的智能工具的中国却没有保持住发展的态势，逐渐淡出历史舞台，其他国家取而代之地继续在理论和实践上取得了一个个里程碑^[3]。

近代的智能世界同样丰富多彩，在思想方面，1637年笛卡尔提出了“我思故我在”，1748年拉·梅特利认为“人是机器”，1820年拉普拉斯提出了“一切已确定”；在理论方面，15-16世纪之间产生了基于可能性的概率论，1673年莱布尼兹设想了思维机器的模型，1847年布尔完成了逻辑数学化，1859年达尔文丰富和完善了进化论，1890年詹姆斯提出了神经网络学说；在实践方面，1642年诞生了计算机的前身：数学数字计算器，1861年卜洛柯初步发现了大脑皮层的特定分工，1872年卡加尔和高尔基通过试验提出了神经元学说^[3]。上述的研究成果催生了基于{0, 1}逻辑的计算机的诞生，并随着对以人脑为主的自然智能研究的深入，如何开发人脑并发展人脑引起了相关领域专家的注意，为人工智能的诞生奠定了基础。

20世纪无疑是智能学的孕育期和发展期，可以以1956年“人工智能”这一术语的正式提出作为分界。1956年前为孕育时期，在理论方面，海森堡和哥德尔分别从物理和数学角度提出了不确定性原理和不完备性定理，冯·诺依曼提出了“在计算机内部存储器中存储指令和程序”的工作原理以及基于此原理的传统计算机的体系结构；在实践方面，沃森和克里克发现了DNA双螺旋结构，进一步揭开了自然智能的面纱，随着宾西法里亚大学成功研制出了EDVAC（1945年）和ENIAC（1946年），凝聚着人类智能的计算机正式诞生并从此飞速发展^[3]。围绕着计算机的诞生，人们开始对人工智能产生了浓厚的兴趣，围绕着机器能否拥有思维，在理论界和学术界都展开了激烈的争论，取得了丰硕的成果。直到1956年达特茅斯会议的成功召开，人工智能才正式走入了人类的视野，并在不同领域不同学科的专家的共同努力下迅速发展，比较有代表性的是1956年提出了问题求解器，1957年结合现代语言学着手对机器翻译自然语言的研究，1961年针对问题解决对搜索方法进行了广泛研究，1961年提出了神经网络算法，1969年开始对专家系统进行研究并于1977年提出了知识工程的概念等等，并且随着模糊集、粗糙集等理论的完善，人们开始关注人工智能中的不确定性问题，并开辟出新的研究领域。

1.3.2 国内外研究现状

人工智能中的不确定性问题仍然贯穿于智能科学与技术的研究中。从

《2009-2010 智能科学与技术学科发展报告》可以看到，该学科研究仍然以两个方面为主，一是对自然智能的深入研究，即继续深入理解脑的工作机理；二是对人工智能的研究，深化自然智能机制的模拟研究，这两个方面相辅相成，互为促进。

在自然智能研究方面，国际学术界有人呼吁将 21 世纪看作“脑的世纪”，欧美等国推出了“脑的十年”研究计划，希望能够把握人类的感觉、认知过程；日本提出了“国际人类前沿科学计划”研究知觉和认知、记忆和学习、思维和语言以及运动和行为等方面的内容^[25]；我国也着重从医学上系统地对大脑进行研究，希望能够了解大脑不同部分的分工以及记忆学习过程在大脑中是如何进行的。但是由于大脑的结构过于复杂以及基于伦理和道德等方面的原因导致相关试验难以开展，所以虽然取得了神经元学说、离子通道理论、脑功能成像等进展，但对脑的宏观和微观层次的结构仍然难以把握，对脑的高级认知功能和脑神经机理所知甚微，对“人的意识如何从脑中产生和如何反作用于脑”仍然难以给出令人信服的解答。

在人工智能研究方面，国际人工智能联合大会在 2007 年提出了约束满足、知识表示与推理、学习、多智能体、自然语言处理、规划与调度、机器人学、搜索、不确定性问题、网络与数据挖掘这十个仍然是该领域研究重点的问题。虽然不确定性问题只占了十分之一的比重，但由于其他九个问题在研究中都需要考虑不确定性，由此“不确定性”仍然是当前人工智能研究中亟待解决的重要方面。

在人工智能不确定性问题研究方面，与国际水平相比，我国在方法论和理论方面有一定的优势，打破了“诸子百家”之后的沉寂。例如，中国工程院李德毅院士在 2005 年完成的专著《不确定性人工智能》中进行了总结归纳，认为不确定性人工智能主要包括模糊集、粗糙集、人工神经网络、混沌和分形、群体智能等多个研究方向；创立了可同时包容模糊性和随机性的云模型，其在不确定性推理中的成功运用，在国内外处于领先地位^[5]。以汪培庄为代表的科学家成功研制了一台模糊推理机，推动了模糊计算机技术的发展^[6]。中国学者基于粗糙集的粒计算研究已取得了长足的发展，并成为人工智能研究中一个飞速发展的新方向^[7]。在人工神经网络方面，马义德等人开展了 PCCN 理论及应用研究^[8]，王国胤等人提出了适合表达三值/多值逻辑的神经元模型^[9]，史忠植等人提出了神经网络模型信息处理与变换机理的新的表示方法^[10]；中国科学院生物物理研究所研究员郭爱克院士及其研究小组在国际上首次发现果蝇“行为抉择”的工作机理，对进一步了解人脑的决策机制有启发式意义^[6]；广东工业大学蔡文教授创立了被称为“可拓学”的矛盾问题智能求解的理论和方法，成为由我国科学家原创，横跨哲学、数学和工程的新学科^[11]；中国科学院半导体研究所王守觉院士创立了多维空间仿生信息学，提出了“基于高维空间的方式模式识别方法”，并自主研制了神经网络计算机来

完成实现^[12]；北京邮电大学钟义信教授提出了以智能生成机制的仿真为核心的“人工智能机制主义研究方法”，使国际人工智能三大主流研究方法在机制主义框架内实现统一，形成了“人工智能统一理论”^[13]。

但在实践方面，对相关技术的开发应用仍然是我国学者的薄弱环节，更多的成果停留在对国际上已有的应用进行改进或者引进。与此同时，以美国和日本为主的国家，非常重视人工智能的相关应用，并开发了以智能机器人为代表的一系列新奇的产品，并且欧美等国家还在着力加大对人工智能不确定性方面的研究，尝试解决传统人工智能研究遇到的瓶颈和困难，并努力把相关应用赋予军事用途，巩固和加强国防。

此外，不确定性的引入促进了人工智能的进一步发展，这使得研究者在探讨人工智能的哲学思考，也不遗余力。例如，美国圣菲研究所 2006 年出版了《剑桥五重奏——机器能思考吗》，书中五位世界级的科学文化巨人围绕着“机器能思考吗”这一论题展开了广泛深入的探讨，认为让除人以外的机器实现思考看似可能，但却在哲学和生物学上存在着阻碍^[14]。人工大脑研究先驱美国犹他州立大学计算机系教授 Hugo de Garis 于 2007 年出版了《智能简史——谁会代替人类成为主导物种》，认为人工智能的不断发展进化，将会给人类带来社会、宗教、军事等多方面的问题，但从物竞天择的角度出发，在全宇宙的范围来看，人工智能将可能会有着比人类更强的适应能力^[15]。虽然人们已经认识到不确定性对实现人工智能的重要性，并且在实践层面形成了初步的理论，但是其哲学含义还未得到很好的研究。已有的工作主要集中在对不确定性本身的哲学意义探索，如诺贝尔奖获得者伊·普里高津的著作《确定性的终结》^[16]与《从混沌到演化》^[17]，我国学者金观涛教授的著作篇章《发展的哲学——论“矛盾”和“不确定性”》^[18]等。中国社会科学院李坚在 2006 年发表的《不确定性问题初探》探寻不确定性的概念，并分析了在量子力学、复杂性科学、社会科学中的不确定性问题^[19]。谢千河等人在 2009 年发表的《不确定性的哲学发展历程及分析》对不确定性的来龙去脉进行了较为深入的分析^[20]。武汉科技学院讲师张智明在 2007 年发表的《人工智能技术及其哲学思考》中提出了人工智能首先是对人类认识论的研究深化；其次，人工智能扩大了认识论的研究领域，最后，人工智能的发展为树立和落实“科学发展观”提供了条件^[21]。华中师范大学赵泽林在 2009 年发表的《人工智能的基础哲学问题探究》中探寻了智能发生的奥秘，尝试建立关于智能的动态发展观，并在哲学分析中合理地描绘智能本身的面貌，尝试合理地阐释智能的结构，为人工智能的实现找到合适的智能模型^[22]。

本文以这些哲学研究成果为基础，将不确定性引入人工智能领域进行哲学上的有益探索，并在研究过程中发现，人工智能的不确定性还包括了多个方面的内

容，虽然不同方面有所关联，但仍然可以大致划分为逻辑基础、研究方法、技术开发、应用前景四个部分，以便分块展开研究。这个研究思路和以往的有所不同，但脉络更加清晰，可以较好地提炼出底层和顶层的方法，让人一目了然，更好的梳理前期工作，更有侧重地展开后续研究。

1.4 主要内容和创新点

论文除绪论外共分为五章，主要对人工智能中的不确定性问题进行了分类并展开了系统的研究。首先肯定了不确定性在人工智能研究中的重要地位，并通过主流研究方法陷入困境以及智能衡量标准难以确定为例，阐述了传统人工智能研究中不确定性缺失的问题，进而提出不确定性是真实世界的本质属性，包括了时间和空间上的连续性，明确提出了人工智能要想有突破性的发展，必须引入不确定性，前沿的研究更多的开始在不确定性方面开辟新的领域并取得了初步的成果。这部分内容主要体现在论文的第二章：不确定性在人工智能中的重要地位。

接着，论文将人工智能中的不确定性问题分为了逻辑基础的不确定性、研究方法的不确定性、技术开发的不确定性和应用前景的不确定性四个主要方面，分别通过第三章、第四章、第五章、第六章来进行较为深入的讨论，既包容了之前学者的重要观点，又尝试提出了个人的见解和想法，系统地对人工智能中的不确定性问题进行了探讨。

本文主要有两个创新点：首先论文对人工智能中的不确定性问题进行了哲学上的阐述。认为物质世界的微观存在状态是离散的、间断的、非连续的，但在宏观表现上却是连续的，人作为宏观存在物，在观察感知物质世界时，不确定性可以理解为时间和空间上的连续性。并且，为了更好地认识物质世界，人类需要采用形而上学的手段，靠着确定性、间断性的方法来把握，从而逐步逼近真理。

其次创新提出了通过从硬件上构建三态模型，从而在硬件上真正实现不确定性的方法。新构建的三态模型将经典布尔系统中的取值范围由 $\{0, 1\}$ 二态推广到 $\{0, 1, \alpha\}$ 三态，其中 α 为拟人因子，标示为不确定态， α 在中间运算的过程中，始终保持着不确定态，只有在最后输出的时候，才会转变为0或者1，最终得到确定性的结果，通过引入 α 对选择进行控制，使得多种选择可以并行出现，以便主体根据情形需要做出更优化的选择。

第二章 不确定性在人工智能中的重要地位

智能按不同的标准通常可以划分为先天智能和后天智能，感性智能和理性智能，个体智能和群体智能，自然智能和人工智能^[23]。其中人工智能在 1956 年美国达特茅斯（Dartmouth）会议上正式走进人们的视野，从诞生起就显示出了较强的学科交叉性和前沿性，其研究成果对多个领域都有深远的影响^[24]。

由于人类对自然智能的认识还不完善，不同科学和学科背景的学者对人工智能有着不同的理解，对人工智能的定义有所不同，例如：人工智能是让理解、推理和行为成为可能的计算；人工智能是用机器模仿和执行人脑的功能，由此产生相应的思维活动；人工智能通过使机器做事让人过得更好，从而超越人类智能^[25]。

虽然定义各有侧重，但目前普遍认为人工智能的研究包括了理解自然智能（特别是人类智能）的认知机理与决策机理；探索、模拟和实现自然智能（包括认知机理和决策机理）的方法和途径；根据经济和社会发展的需要，研制具有一定智能水平的机器系统；把智能系统应用于国家经济建设和社会服务等领域，促进科学技术和经济社会发展的智能化等主要内容。

随着信息技术的飞速发展，计算机已越来越多地融入到人类的经济、社会 and 生活中，如何用计算机更好地替代人类的工作，甚至创造出像人一样的智能机器，成为人工智能追求的目标之一。

2.1 传统人工智能研究存在不确定性缺失的问题

在人工智能 55 年的发展过程中，围绕人工智能的基础理论和研究成果，出现了符号主义、联结主义和行为主义三大主流研究方法^[26]。符号主义、联结主义和行为主义虽然在特定的方向上取得了一系列的成果，但是目前还没有一种研究方法能够彻底地替代另外两种研究方法，并且这三种研究方法都存在一定的局限性，究其原因是不确定性缺失的问题。

2.1.1 符号主义、联结主义、行为主义面临的困境

符号主义又被称为功能主义、逻辑主义和计算机学派，在人工智能发展的前 20 年占据了主导地位。符号主义认为认知是一种符号处理的过程，人类的思维过程也可用某种符号来描述。无论人脑和计算机在硬件和软件层次上有多大的区别，但在计算机理论层次上，都有产生、操作和处理抽象符号的能力，因此把人脑和计算机都看作操作、处理离散符号的形式化系统。

20 世纪 80 年代以后，人工智能学界发生了一场人工神经网络(Artificial Neural

Networks, ANN) 的“革命”，对符号主义方法产生了很大的冲击，逐渐形成了联结主义学派。不同于物理符号系统假设，联结主义认为人类的认知活动主要基于大脑神经元的活动^[27]。

联结主义又被称为结构主义和仿生学派，其代表人物是美国生理学家沃伦·麦卡洛克 (Warren McCulloch) 等人，认为模拟人的智能要依靠仿生学，特别是模拟人脑并建立脑模型。人类思维的基本单元是神经元而不是符号，智能是相互联结的神经元竞争和协作的结果。联结主义认为结构与功能、行为密切相关，不同的结构表现出不同的功能和行为。人脑不同于计算机，用计算机模拟人脑应着重于结构模拟，即模拟人脑的神经网络结构。

以控制论方法进行人工智能研究由来已久，行为主义认为智能与表达、推理、知识等无关，主要取决于感知和行为，并提出了智能行为的“感知——动作”模式^[28]。

行为主义又被称为进化主义和控制学派，早期的工作重点主要是模拟人在控制过程中的智能行为和作用，研究自寻优、自适应、自校正、自镇定、自组织和自学习等内容。20 世纪 80 年代，诞生了智能控制系统和智能机器人，使行为主义方法在人工智能研究中掀起高潮。

虽然符号主义、联结主义、行为主义都是对人类智能的模拟，但在人工智能的基本理论问题，诸如研究方法、侧重点、研究原理、研究内容、知识结构等方面都有所不同，如表 2.1 所示。

表 2.1 符号主义、联结主义、行为主义的区别

研究方法	侧重	研究原理	研究内容	知识结构	代表人物	主要成果
符号主义	功能	人的认识是对符号计算和推理的过程	基于逻辑的知识表示和推理	规范知识	纽厄尔 西蒙 麦卡锡	归纳推理 启发式算法 专家系统 知识工程
联结主义	结构	人的智能是由人脑生理结构和工作模式决定的	神经网络	经验知识	Hopfield 麦克莱伦 德	脑模型 多层网络中反向传播的 B-P 算法
行为主义	行为	智能取决于对环境的感知和行为，不需要推理、表示和知识	模拟人的各种控制行为	常识知识	布鲁克	智能控制系统 智能机器人

对于符号主义来说，首先在知识结构上需要规范的知识，要求所处理的信息准确、完善，能够符号化，即所给的信息是确定性的，但确定性的知识在我们生活中仅是极小的一部分，所以符号主义很难在不确定性的知识，如：感觉、情感、视觉等方面有所进展。其次，从符号主义所需的形式化方法上来看，形式化是一个语言的抽象过程，而人的形象思维、灵感、顿悟不需要语言，因而不能被形式化^[29]，而形式化是计算机发展的第一极限，即计算机只能处理可形式化的信息。并且，即使是思维中可形式化的部分，例如抽象思维，也未必是可计算的，而可计算性是计算机发展的第二极限，即计算机不能实现其中不可计算的部分。此外，可计算性问题包括两类，第一类是只需要低次多项式时间的计算机非难解问题，第二类是非多项式时间的计算机难解问题。计算机难解问题是计算机发展的第三极限，将导致运算所需时间和空间的急剧膨胀^[30]。例如：如果一个算法的复杂度是 $O(n^2 2^n)$ ，当 $n=1$ 时，复杂度是 2；当 $n=2$ 时，复杂度是 16；以此类推，当 $n=10, 100, 1000$ 时，复杂度将呈指数级增长，即使当前的计算机能够算出来，也需要几年甚至几十年的时间，导致失去了解决问题的意义。如图 2.1 所示，在人类思维中，其实只有极少的一部分（红色标注处），才是计算机实际可以模拟和解决的，即符号主义可能取得成果的部分，而其它大部分区域，却是当前符号主义研究方法无能为力的。

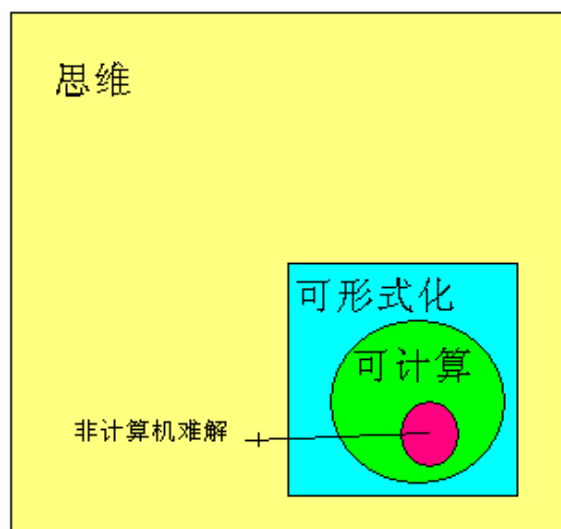


图 2.1 非计算机难解问题在思维中所占的比例

联结主义侧重于从结构上模拟人脑，而人脑是迄今为止最复杂的系统之一。虽然随着科学的进步，对于脑的研究逐渐深入，但要彻底了解脑的结构、功能以及工作机理还有相当长的路要走，联结主义也因此在此初期取得一系列成果后，和符号主义一样步入了瓶颈状态。

在传统计算机体系结构框架下，即使人类能完全解开大脑之谜，联结主义也未必会有很大的突破^[31]。人脑在进行处理的时候，是并行工作的，但传统的计算机结构仍然是基于冯·诺依曼结构的。冯·诺依曼结构是串行的，即使进行并行处理，其最基本的功能单元在执行时仍然是串行的。此外，一味从结构上模仿，并不意味着能具备相应的功能。例如：1965 年，我国科学家人工合成了胰岛素（一种蛋白质），但这种合成物却没有表现出任何生命的迹象，使得看似有应用前景的试验仅仅成为了证明从无机物可以合成有机物的例子。因此，要想在人工智能

方面有所突破，在模仿的过程中必须有所取有所不取，能够根据实际情况进行突破和创新。以仿生学为例，飞机的成功研制是由于创造性的运用了螺旋桨做动力，而不是照搬鸟飞翔时扑动翅膀的“扑翼机”路线^[32]；火箭升空虽然模仿了水母、墨鱼等海洋生物的反冲作用，但在动力方面却大不相同。此外，仅仅靠硬件还无法模拟人的大脑，还需要相应软件的协助，但一旦要转换成程序语言的形式，又不可避免的回到了在符号主义中面临的形式化的难题。

行为主义从一开始就认为智能取决于对环境的感知和行为，不需要知识、表示和推理，但这个出发点并不符合人类的发展规律。对于人类来说，思维是人区别于其它动物的一个重要的标志。思维的存在，使人类能表示、推理、学习相关知识，这些对感知、行为同样有着比较重要的指导作用。例如，昆虫虽然能够感知环境并做出反应，但更多的是条件反射，并不是思维层次上的智能。人工智能的目标是希望能够制造出有思维参与的真正智能的机器，抛开了知识、表示和推理，不可能得到真正意义上的智能，所以当前行为主义在发展中也遇到了障碍。

美国科学史家、科学哲学家托马斯·库恩（Thomas Kuhn, 1922-1996）认为在科学的发展观上，只看到科学知识的静态积累或科学假说的不断证伪和更替都是不符合科学发展的客观事实的。他主张把科学看作由“科学共同体”按照“范式”进行的专业活动，并描绘了一种常规时期和革命时期相互交替的科学发展模式。科学的发展就是范式的建立、转换和替代的革命过程，科学革命的结果最终是以新范式彻底地取代旧范式而告终^[33]。

在人工智能领域，库恩的理论同样适用，因为符号主义、联结主义和行为主义都具备特定的知识结构、系统的研究方法、各有侧重的研究内容，都曾在人工智能领域得到过相当专家的认可和研究者的追捧，可以在一定程度上看作是三种不同的“范式”。但在人工智能的变革中，却暂时没有看到新范式彻底地替代原范式的革命过程的出现。因为目前在人工智能领域，是多种方法齐头并进，各种理论相互补充的状态，新兴的方法往往是在原有方法失败的方面取得成果，而不是在其成功的方面完成突破。

三种方法之所以会在发展到一个让人欣喜的阶段后又进入瓶颈状态，究其原因是在研究过程中出现了不确定性缺失的问题，例如：符号主义希望把一切都符号化，而符号化首先要求知识是确定性的（包括准确、完整等要求）；联结主义着重于尽可能精确的模拟人脑（人脑中有很多复杂的行为，很难确定性的模拟出来）；行为主义直接排除了思维的因素，而思维的神奇之处恰恰在其不确定性的方面^[34]。因此，要想真正突破人工智能的难题，必须考虑人工智能中的不确定性问题，把确定性和不确定性有机统一。

2.1.2 智能衡量标准难以确定

英国数学家阿兰·图灵（Alan Turing, 1912-1954）于 1950 年在《思维》杂志上发表了著名的论文：《机器能思维吗？》，提出了不要问“机器能否思维”，而应当问：“机器能否通过智能测试。”为此，他提出了一个判断机器是否具有人类智能的测试标准——图灵测试（Turing Test）。

图灵测试简单地说就是将测试者和被测试者分别置身在互不能见的房间里，其中被测试者是人和机器。测试者将提出多个问题，并通过给出的回答来判断究竟是人还是机器在作答。如果机器回答的问题让测试者认为是人在作答，那么这台机器就顺利通过了图灵测试。为了一目了然，本文将引用其他学者绘制的图示来说明，如图 2.2 所示。

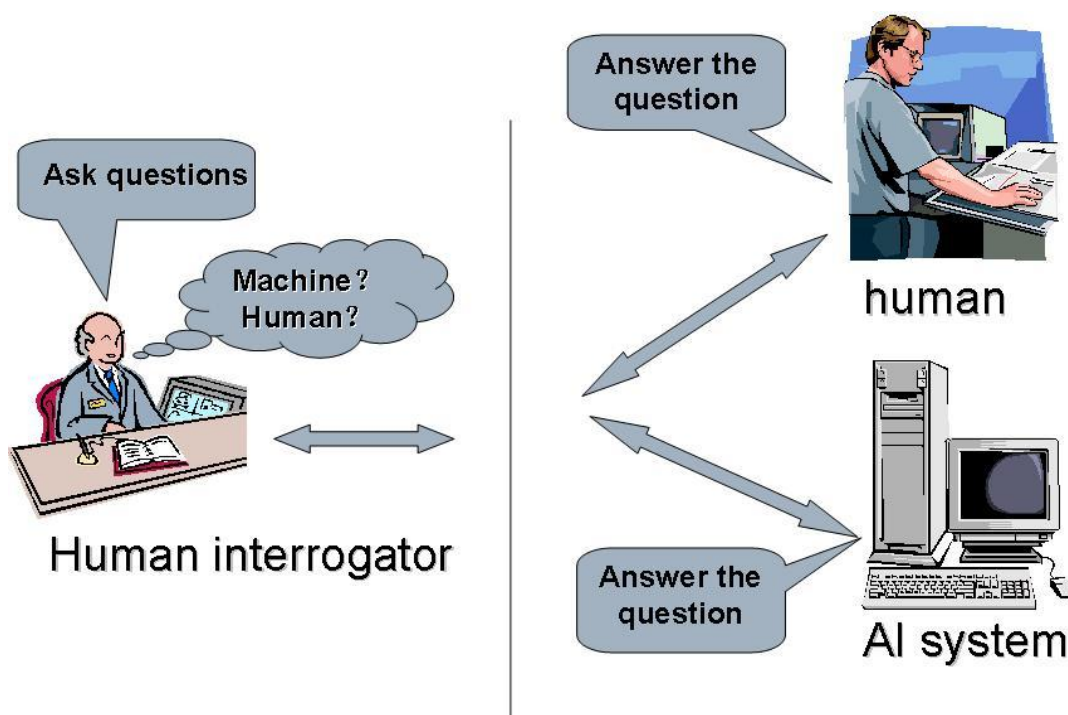


图 2.2 图灵测试

但仍有不少人认为即使通过了测试，也不能说机器有智能，比较著名的是塞尔·西尔勒假想的“中文屋子”的实验，认为即使有机器通过了图灵测试，也不一定说明机器就真的像人一样有思维和意识^[35]。

中文屋子的整个过程包括三步：一是将中文字符送进屋子，二是对照操作规程把中文字符转换为另外的中文字符，三是将新的中文字符送出房间^[36]。塞尔认为，只要按照操作规程完成操作，即使对中文一窍不通的人仍然能够看上去“理解”中文语句的意思。为了更形象地理解中文屋子的思想，本文引用了相关学者

绘制的图示来说明，如图 2.3 所示。



图 2.3 中文屋子

不过图灵对此做出了反驳，认为要知道机器有没有智能，需要站在机器的角度，即：让人变成机器，而不是凭着人类的主观感受来猜想机器是否真正具备智能^[37]。

图灵测试至今在智能判定方面还影响着相当一部分人的思想，但也引起了不少的质疑，例如以美国信息论专家香农等为代表的学者认为图灵测试是行为主义性质的，通过测试的机器只能说在行为上等价于人类，但无法判断是基于“思维”解决问题还是基于背诵题库来解决的^[38]。除此之外，还有人提出“图灵测试”本身是不确定性的，不符合科学测量的方法。因为人类感官存在着局限性，所以相关的科学测量需要借助仪器，但是在“图灵测试”中，判定机器是否有智能的标准是测试者——人，而人在获取信息的过程中必然会带着主观的判断，从而导致测试结果是不确定性的^[39]。

此外人类对智能的衡量标准也是不确定性的，即在不同的时期，人类对智能有着不同的标准，检测出来的智能具有偶然性和临时性，导致在一个时期被认定为智能的行为在另一个时期被认定为不属于智能。正如美国认知科学教授道格拉斯·理查·赫夫斯台特（Douglas R. Hofstadter）所指出的：“有些时候，当我们朝着人工智能方向前进了一步之后，却仿佛不是造出了某种大家都承认的确是智能的东西，而只是弄清了实际智能不是哪一种东西。”例如，1997 年当 IBM 深蓝

(Deep Blue) 计算机采用启发式搜索方法战胜了国际象棋冠军加里·卡斯帕罗夫时，人们认为计算机拥有了智能，但现在，更多人认为取胜原因仅仅因为计算机借助了强大的计算、存储和搜索能力，并不能证明计算机拥有智能。再如：目前对智能提出的要求是机器要有创造力，可能在若干年以后，当机器真正拥有了创造力，人类还会提出更进一步的智能标准，导致即使拥有创造力也不能代表机器拥有智能^[32]。

当前在人工智能领域，对于图灵测试能否作为智能评判的标准以及智能到底是什么仍然没有统一的标准，是一个不确定性的问题，虽然暂时还无法给出确定性的答案，但不同的专家和研究者仍然继续按照自己的思路进行着不懈的探索。其实，在人工智能领域，还有许多类似的不确定性问题值得人类去思考和探究，不确定性已经成为人工智能发展需要考虑的重要方面。

2.2 不确定性是真实世界的本质属性

随着研究的深入和技术的提升，曾经看起来是确定性存在的现象在当前看来呈现出了不确定性的特征；曾经认为是确定性的理论也被一个个颠覆。无数的现象和试验结果说明，所谓的真理不过是在一定时空范围内被证明确定性为真的理论，并不排除在另外的情况下将有不可预期的现象发生。在越来越多的不确定性现象面前，人们逐渐认识到：真实世界唯一确定的一件事情是“世界是不确定性的，确定性是不确定性的特例^[22]”。

2.2.1 确定性是对真实世界的近似刻画

早在人工智能诞生之前，人类在认知上就多次排斥着不确定性，反复地坚持着确定性的认知观，并且取得了丰硕的成果，特别是在统计力学、量子力学出现之前，牛顿的经典力学奠定了 18 世纪物理学的基础，确定性融入了大多数人的认知观，而不确定性却一直未能真正地走进科学的殿堂。直到随着人们研究范围的扩大和深入，在将相关定理、公式推广的过程中遇到了难以解决的问题，才不得不一次次修正原先单纯的确定性的认知，逐渐在研究中注意并接受了不确定性。

例如近代哲学之父、法国哲学家、科学家和数学家勒内·笛卡尔（Rene Descartes, 1596-1650）思想的出发点从“上帝”发展为“找到一个确定性的知识基础”。笛卡儿认为数学是唯一能够提供精确知识的学问，主张把数学和物理学结合起来，精确地表示出所有的科学解释。在笛卡尔看来，数学是确定性的知识基础，包括人类思维、意识在内的任何东西都可以抽象为数学符号来求解，即一切都是确定性的。英国著名唯物论哲学家霍布斯（Thomas Hobbes, 1688-1679）说：“一切思维不过是计算（加加减减）。”传统的计算是确定性的，在他看来，思

维可以完全形式化，也是确定性的。德国哲学家、物理学家、数学家莱布尼兹（Leibniz, 1673）沿着霍布斯的思路尝试通过数学描述思维来打破思维的神秘性。莱布尼兹坚信，基于符号化的方法可以建立“普遍逻辑”和“逻辑演算”，世界上的一切都可以解释清楚。此外，布尔（George Boole, 1815-1864）还提出了逻辑代数，并为逻辑推理建立了形式语言，认为：用符号语言与运算可以表示任何事物，包括思维。在他们看来，思维是由人脑产生的，而人脑是一个确定性存在的物质，因此思维本身也是确定性存在的，只要是确定性存在的东西就可以通过数学这门学问来解决。

笛卡尔、霍布斯、莱布尼兹、布尔认为数学能够确定性地表示包括思维在内的任何事物，但这种想法无疑过于乐观。基于刚性逻辑基础的传统数学虽然解决了一系列繁杂的问题，并且看似能够很好地描绘真实世界，但不容忽视的是，传统数学描绘的只是真实世界中确定性的部分，对于模糊的、中间态的、非恒常的事物以及人类的形象思维却无能为力，因而基于研究的需要，以模糊数学、非单调逻辑等为代表的一系列不确定性方法逐渐补充到传统的确定性的数学中。

同样的，以物理学为代表的自然科学也不可避免的经历了数学所遇到的障碍。例如法国数学家、天文学家拉普拉斯（P.Laplace, 1749-1827）曾说：“如果准确的获得了宇宙的完整信息，就能够决定它在未来和过去任意时刻的状态。”这种机械决定论的观点，重在强调世界上的一切都已确定，被一种不可抗拒的力量决定着，任何改变现状的努力都是徒劳的。

18世纪被称为经典物理学时期，英国物理学家、数学家、天文学家艾萨克·牛顿（Isaac Newton, 1643-1727）在1687年发表了《自然哲学数学原理》，对万有引力和三大定律进行了表述，奠定了物理世界经典力学的科学观点。牛顿被看作是一位形而上学的机械唯物主义者，认为时空是绝对的，并且一切可观测的物理量在原则上都可以无限精确的测量，测量的基础在于物理规律的确定性。当时普遍认为数学-物理方法可以解释包含人类智能在内的一切问题，即宇宙的一切已被确定。牛顿奠定的经典力学在很长一段时间内统治了人类的认识，解决了包括天体运动在内的宏观物体的低速运动问题，但是随着科学研究的逐步深入，经典力学逐渐暴露出一定的局限性。

例如：1827年，苏格兰植物学家罗伯特·布朗（Robert Brown, 1773-1858）发现了花粉在水中的不规则运动——布朗运动，表明物质的分子始终处于无规则的热运动中。最开始，人们尝试用经典力学来解释分子运动的现象，试图分析相关的碰撞模型，均以失败告终。最后，大量分子的热运动不得不通过统计力学来计算出分子整体的平均分布律和速度分布律，但要想精确到单个分子的速度和状态是不可能的，分子的热运动体现出了随机性。也就是说，基于确定性基础的经

典力学虽然能够成功的处理单个或多个物体的碰撞模型，但当运用推广到大量的微小的粒子时，却不再有效，以布朗运动为代表的大量微粒的运动模型为人们展现出了一种不确定性的状态，让统计力学走入人们的视野。

再如，在经典力学中研究对象可以确定性的分为“波”和“粒子”，这是一种非此即彼的状态。但 1905 阿尔伯特·爱因斯坦（Albert Einstein, 1879-1955）在研究光电效应时发现光在具有普遍上认可的“波”的性质的同时还具有“粒子”的特性，这是一种亦此亦彼的“模糊”状态。在随后的研究中，科学家发现包括电子在内的一切物质都存在波粒二象性，波粒二象性成为物质的一个普遍的规律，奠定了量子力学的基础。但随着研究的深入，科学家开始质疑波粒二象性是否真正存在，因为暂时没有任何一个实验能够同时观察到物质的波动性和粒子性，不同的观察手段只能观察到其中的一种性质，导致物体究竟是什么性质仍然得不到确定性的解释。

虽然爱因斯坦发现了相对论，对量子力学做出了重大贡献，但他还是认为“上帝不和我们掷骰子”，继续保留着“世界是确定性的”观点。1927 年，德国理论物理学、原子物理学家韦纳·海森堡（Werner Heisenberg）提出了量子力学的基本原理——“测不准原理”，认为：“一个微粒的某些物理量（位置和动量、方位角和动量矩、时间和能量等）不可能同时具有确定性的数值，即一个量越确定，另一个量的不确定性越大，再精密的仪器也不能测出完全精确的值，物理学不能做出超越统计学范围的预测。”并逐渐被人们接受。

此外：“光是直线传播的”，“时间是一维线形运动的”曾被普遍的接受，但广义相对论的出现，让人们看到了光线竟然会在一定时空范围内扭曲，时间也会出现倒流或穿越；曾经认为生活的宇宙是独一无二的全部真实世界的存在，但在量子学出现后，越来越多的科学家开始相信所能观测的宇宙及其上发生的一切现象不过是真实世界的冰山一角，作为物理存在的整体，真实世界是多重宇宙组成的，包含着大量还不为人所知的平行宇宙^[40]。

虽然数学家、物理学家、哲学家对世界是确定性的做出了种种假设和乐观的判断，还试图通过各种方法进行证明，但直到今天，人们还未能揭开思维的神秘面纱。曾经被认为可以精确刻画任何事物的数学和物理方法不但没有精确地刻画出人类的思维和意识，反而连自身也因为无法回避不确定性而几度陷入危机，不确定性逐渐引起了越来越多人的关注。

2.2.2 不确定性辩证地反映了真实世界

不确定性曾经一度被排斥在科学殿堂之外，很少有学者承认不确定性的存在，但是不确定性却像真实世界的影子一般一直存在于人类的认知中，并且可以追述

到几千年前。例如，早在中国古代，著名的思想家老子通过对日常生活的观察，明确提出了“有无相生，难易相成，长短相形，高下相倾”的观点，强调了即使是“有无”、“难易”、“长短”、“高矮”这些相对立的事物，也存在一个渐变的过程，事物之间的差别是模糊的，没有确定性的界限，在他的眼里，不确定性体现为对立事物之间渐变的状态和模糊的过渡。

战国时期，《孙子兵法》在《势》篇谈到：“声不过五，五声之变不可胜听也；色不过五，五色之变不可胜观也；味不过五，五味之变不可胜尝也；战势不过奇正，奇正之变不可胜穷也。奇正相生，如环之无端，孰能穷之。^[41]”其意思是虽然乐音只有宫、商、角、徵、羽五个音阶，但通过高低音阶交替、音符长短变化等方式可以演奏出无穷种声音，其丰富的表现力并不局限于确定性的初始条件，任何细微的变化都可能带来完全不同的效果。色彩调节、口味搭配、战术变换也是同样的道理。《虚实》篇中谈到：“五行无常胜，四时无常位。日有短长，月有死生^[41]。”意思是木、火、土、金、水五行相生相克，并无定数；春、夏、秋、冬四季循环更替，没有界线；白昼的长短，月亮的盈亏随时间变化，体现了循环延展，相互包含，更替推移的不确定性思想。

东汉哲学家王充在论述世间万物生成机理的时候，明确地反对了自然目的论和预成论的观点，认为世上所有的事物都是偶然自生的，不存在预先设定好的目标和方案，即使有人为的目的和行为，其最终结果如何，仍然是受一些偶然因素左右的，在他的眼里，不确定性体现为随机性、偶然性^[20]。

在诗词中，更是通过虚数、比喻等方法模糊地体现了想要表达的感觉，给人以意犹未尽的无限想象空间，同样是不确定性的体现。例如在描绘女子动人之处时有：“回眸一笑百媚生，六宫粉黛无颜色。”“幽兰露，如啼眼。”再如描写感伤之情时有“杨柳岸，晓风残月”，“一个是水中月，一个是镜中花。想眼中能有多少泪珠儿，怎禁得起秋流到冬尽，春流到夏。”虽然诗词中的语句没有具体的说清想要表达什么情感，但却能让读者产生共鸣，按照个人的审美和理解来构想画面。

中国古代对不确定性的理解比较抽象，而西方的阐述相对更加容易理解。在近代西方哲学中，客观唯心主义建立者格奥尔格·黑格尔（Georg Wilhelm Friedrich Hegel, 1770-1831）把认识作为整体看待，提出了正反合三段式运动。依黑格尔的意见，在最好的思维中，思想的“真”和“假”不再是界限分明的对立物；不存在完全真的或者完全假的事物，任何部分事物都不十分真，真理不再是一个静止的点，而是一个不断建立自身的运动主体，在黑格尔眼中，不确定性体现为非恒常性。

德国哲学家、天文学家伊曼努尔·康德(Immanuel Kant, 1724—1804) 在《纯

粹理性批判》中回答道：“我们只能知道自然科学让我们认识到的东西。”康德认为事物是具体的和物化的，而语言是抽象的，在认识事物的过程中，人们根本不可能认识到事物的真性，只能认识到事物的表象。康德的著名论断就是：“知性为自然立法。”即：事物的特性与观察者有关^[20]，在康德的眼中，不确定性体现为难以掌控的主观性。

其实，不确定性存在于人类的日常生活的方方面面，存在于真实世界的每个角落，并不只是在科学技术高度发展的情况下才走入人类视野，在不同时期它都曾影响过部分人的思想，但是大多数人只有在遇到困境时才会留意不确定性的真实存在。

2.2.3 不确定性主要体现为时间和空间上的连续性

自然科学等方面的研究成果逐渐为人们展现出了真实世界的本质特征是不确定性的，确定性是不确定性的特例。从哲学角度来看，真实世界由物质组成，而其中的物质一直做着永不停息的运动，这样的运动有的已经被人类探悉，而有的还暂时还未被掌握。为了能更好地开展研究，人们必须使用形而上学的方法对真实世界进行近似刻画，即做出确定性的处理，以便有针对性地了解其在特定时刻所处的状态，进而发现其中的规律。

形而上学的方法在研究中有着重大的意义，通过形而上学方法的近似，人类可以将研究对象看作在该时刻是静止不变的，并且能够从其存在的空间中分离出来。如果不采取形而上学方法的近似，而是精益求精地要求在不断运动和渐变状态中研究事物，将导致永远无法捕捉到当前的状态，难以探悉潜在的规律，并且研究结果也很难得到实际的应用。

形而上学的近似处理虽然有助于更好地研究真实世界，但这并不意味着人们可以将这种研究方法无限地扩展到对该事物甚至整个真实世界的看法，即不能在世界观上认为真实世界是确定性的，而应该辩证地看到不确定性才是真实世界的本质特性。

世界的本原是物质的，在对物质世界的探究中，人类逐渐发现了物质世界的微观状态是离散的、间断的、非连续的，呈现出来的是原子、分子的甚至更微小的粒子构成。但不容忽视的是，物质世界的宏观表现是连续的，人作为宏观存在物，在观察认识物质世界的时候，感受到的也是其在时间和空间上的连续性，例如：拔地而起的高楼、红白相间的衣服等等。

本文认为不确定性主要体现为时间和空间上的连续性。正是因为这种宏观上的连续性的存在，才使得物质世界在时间和空间上展现出来了运动性和模糊性的状态。不确定性作为真实世界的本质特征是始终存在、无法被消除的。但是，为

了研究的方便，人类往往通过形而上学的手段，假定物质世界时离散的、间断的、确定性的，从而使用各种方法来逐步逼近真理。换句话讲，不确定性是事物真实存在的属性，而确定性只是人类认识物质世界的一种手段，这与运动和静止的关系类似，即运动是绝对的，静止是相对的，运动是存在的，而静止本身是不存在的，静止不过是人类用于认识事物存在的一种假定和处理方式。

首先假定研究事物所处的空间一定，那么从时间层面上看，不确定性主要体现在运动性，即随着时间的变化，研究对象是不断运动的。在这里，运动性既可以包括同一平面上的线性运动和非线性运动，又可以包括不同平面或多维空间中的复杂运动。以最简易的同一平面上的线性运动为例，有学者会提出：“线性运动是否应该属于确定性的，而不应该划分到不确定性中。”但是，不容忽视的是，线性运动并不仅仅是单一方向轴上的延伸运动，还可能是折返运动，很难确定性的把握住下一时刻的位置，导致人们无法捕捉到其任意时刻的状态，因而很难预知事物的未来发展，很难推测事物的曾经状态，很难掌握事物内在运行的机理，导致在研究中出现了不可知的现象，需要借助概率的方法来推测最可能的结果以及随着时间的变化对事物的认识呈现出否定之否的发展趋势。

其次假定研究事物所处的时间轴不变，从空间层面上看，不确定性体现为连续性。当然，这里的连续性是指人类肉眼观察或通常思考所能达到的层面上的连续性，但如果深入到分子构造或其他更细微的结构中剖析，再具有连续性的事物也将显现出分隔、断续的特征。由于存在连续性，使得真实存在的事物并不都是理想状态下的层次分明的状态，并不都能从其存在的空间中剥离出来，而是更多地展现出了一种渐变的、亦此亦彼的模糊状态，使得人们在很多时候难以对研究对象进行确定性的划分和归类、难以对事物的真伪、好坏、明暗、浓淡进行定性的区分，很难分辨出事物究竟在什么时刻完成了量变到质变的转换，只能借助模糊数学等包容亦此亦彼的方法来进行相对准确的研究。

2.3 不确定性的引入：人工智能的新发展

谈到人工智能，人们会不约而同地想到“图灵测试”，并且大部分人认为图灵测试是判断机器是否拥有智能的确定性的标准。但是生活在同时包容确定性和不确定性的真实世界而不是理想世界的我们，不得不反问一下：这个智能衡量标准真的是确定性的吗？如果不是，那么人工智能中还将面临多少不确定性的问题呢？我们的答案是智能衡量标准是不确定性的，并且在人工智能领域，还有许多不确定性的问题有待研究，不确定性为人工智能研究开辟了一个崭新的领域，以李德毅院士为代表的专家将其称为“不确定性人工智能”。

当不确定性人工智能走进人们的视野后，引起了多方面的关注。在不确定性

人工智能出现之前，虽然也研究知识表示和推理、自然语言的处理、机器人学、搜索、数据挖掘等方面的内容，但在思考和建模的过程中，使用的数学、物理方法都是建立在确定性的基础上；而现在，人们开始更多的关注不确定性，并引入了模糊集、粗糙集、非单调逻辑、混沌与分形、核函数、云模型等一系列适应于解决不确定性问题的研究方法。在国家“973”、“863”、自然科学基金等项目的大力支持下，我国在不确定性人工智能研究方面取得了令人欣喜的成绩，特别是理论方面的研究，目前处于世界领先水平，引起了国际上的高度重视。

在理论研究方面，中国科学技术大学研究生院陈霖院士及其研究小组提出的以整体论为核心的“拓扑性质初期视觉理论”得到了国际认知学科主流学术界的认可，并于2009年当选国际认知科学学会主席；中国科学院生物物理研究所研究员郭爱克院士及其研究小组在国际上首次发现果蝇“行为抉择”的工作机理，对进一步了解人脑的决策机制有启发式意义；中国工程院李德毅院士创立的同时包容模糊性和随机性的定性定量转换模型——云模型以及在不确定性推理中的成功运用，在国内外处于领先地位；广东工业大学蔡文教授创立了被称为“可拓学”的矛盾问题智能求解的理论和方法，成为由我国科学家原创，横跨哲学、数学和工程的新学科；中国科学院半导体研究所王守觉院士创立了多维空间仿生信息学，提出了“基于高维空间的方式模式识别方法”，并自主研制实现了神经网络计算机；北京邮电大学钟义信教授提出了以智能内在生成机制的研究为核心的“人工智能机制主义研究方法”，使国际人工智能三大主流研究方法在机制主义框架内实现统一，形成了“人工智能统一理论”^[6]。

但是，由于我国在基础工业技术、制造技术、材料工程等方面的基础相对薄弱，导致先进的理论还不能马上付诸于应用，目前在技术研究和应用研究方面的创新贡献相对较少，比较有代表性的是机器人足球、机器博弈、智能汽车和机器鱼。在机器人足球方面，我国选派的参赛队伍多次在国际机器人足球竞赛 FIRA 和 Rob-Cup 上获奖，成为国际机器人足球界实力最强的国家之一。机器博弈是人工智能最早开展研究的领域，在2006年我国人工智能学会机器博弈专业委员会举办的计算机博弈锦标赛暨中国象棋人机大战中，参赛机器取得了较好的战绩。智能汽车通过模仿人类对环境的感知、信息融合、规划和决策来进行驾驶，取得了阶段性的成果^[6]。机器鱼的研究综合了仿生学、人工智能、控制论、电子学等多学科知识，通过工程手段模仿鱼类运动，例如：国防科大机电与工程学院研究人员在“尼罗河魔鬼”等鱼类的柔性长鳍上找到了灵感，着手研究波动推进的仿鱼智能水下机器人，为开发高度移动性、适应性、稳定性的无噪声、对环境扰动小的智能航行器和水下机器人奠定了基础，一旦突破了现有材料的限制，将对国家安全防御等方面起到重要的促进作用^[42]。

由于不确定性人工智能兴起的时间比较短，前沿的理论和研究方法几乎为全世界共享，所以国内外基本上处在相同的起点开展研究。在理论研究上，我国跟上了国际的步伐，并在部分领域提出了有建树的原创性观点，但在技术和应用研究方面仍然在较大程度上局限于对国际上已有工作的优化，创新程度较低，还需要进一步提高，使创新理论能尽快的和富有应用前景的实践相结合，打造中国原创性的应用产品。

考虑到不确定性在人工智能研究中的重要地位，本文将对人工智能中的不确定性问题分类，从逻辑基础的不确定性、研究方法的不确定性、技术开发的不确定性和应用前景的不确定性四个方面展开研究。

第三章 人工智能逻辑基础的不确定性

逻辑是人工智能发展的重要理论基础，在人工智能发展早期，人们狭隘地认为逻辑基础是数理逻辑；但当人工智能研究中期出现了理论危机之后，开始有人认为逻辑和智能没有必然的联系，人工智能的发展可以不必借助知识和推理；随着研究的进一步深入，人们发现逻辑仍然是人工智能的重要理论基础，当前人工智能只能使用数学符号化的逻辑。但人工智能除了要处理理想世界中的确定性问题外，还需要解决现实世界中的一系列不确定性的问题，而经典逻辑并不能解决非二值、非精确性、非完备性的问题，因此要想进一步发展人工智能，需要能包容不确定性新的逻辑基础^[43]。

3.1 逻辑的基础地位

逻辑学诞生至今，已有两千多年的历史，可以大致分为形式逻辑学和辩证逻辑学两个方面的内容。逻辑学经历了两次飞跃，首先是数理逻辑的建立，数理逻辑又称为标准逻辑，主要通过符号语言和逻辑演算进行表达，先后出现了命题演算、公理集合论、递归函数论、模型论、证明论等内容，使数理逻辑逐渐走向成熟。但由于标准逻辑命题是基于二值数学化的形式逻辑，只能处理确定性的问题，排除了一切形式的矛盾和不确定性，因此又被称为刚性逻辑学时期。但在人工智能研究中，确定性的问题只是局部问题，更多的不确定性问题需要新的逻辑基础来指导。人工智能的研究推动了逻辑学的第二次飞跃，其主要任务是通过真实世界中存在的种种不确定性特征来突破形式逻辑的局限性，建立数理辩证逻辑基础，促使逻辑学进入柔性逻辑时期。

虽然在形式化问题上经历了组合爆炸的危机，但逻辑始终是人工智能发展的重要基础。人工智能学会副理事长何华灿教授打了个形象的比方：“逻辑是思维和智能的 DNA，各种形式的智能都是逻辑的语义解释和实现。”在自然界中，从最低等的生物到最具智能的人类，虽然形态各异、种类千差万别，但无一例外的都是通过 DNA 来描述的；同样的，小到思维的细节，大到生态系统的自适应，都能找到有一定规律的语义来表达，一旦找到这个语义，就有可能进行符号化的描述并在人工智能中得以实现^[43]。

在人工智能研究中，对应于广义的智能定义，逻辑基础同样也是一个广义的概念，例如：逻辑基础存在于生物生长和免疫的过程中；存在于人类的抽象思维中；存在于群体协作中；存在于生态系统自适应过程中。逻辑基础本身一直存在，关键是如何找到并尽可能的模拟它，从而促进人工智能及相关学科快速发展。

3.2 刚性逻辑不能全部反映人类思维的本质

刚性逻辑又被称为形式逻辑、标准逻辑、经典数理逻辑，有着二值性、排中律、恒常性的特点，主要是对全信息和确定性世界的表述。经典数理逻辑用精确的方法研究形式逻辑，将基于概念、规则、推理的自然语言表述过程抽象化为形式语言的表述和符号演算，并在人工智能初期取得了数学定理证明等一系列的研究成果。但由于标准逻辑的立论点是对理想世界的全信息确定性的刻画，排除了一切不确定性因素，所以随着研究的深入，显露出了一定的局限性^[44]。

第一，刚性逻辑命题的真值域是二值的，即“非此即彼”的状态，排除了“亦此亦彼”的过渡态、模糊态，例如：刚性逻辑只能对“帽子是红色还是黄色”做出“是”或者“否”的回答，却不能对“红中带黄”的不确定态进行描述。第二，刚性逻辑具有矛盾律和排中律的特性，即一个命题和它的否命题不可能同时为真，并且原命题和它的否定中必然有一个为真，但在实际生活中，很可能一个命题和其否命题同时为真，例如，“浏阳河很窄又很宽”，这是一个辩证的说法，取决于参考物的选取，与大海相比，浏阳河显得窄，但和其支流相比，又显得很宽。这是在日常生活中容易遇到也容易理解的问题，但刚性逻辑对类似问题却无法下手。第三，刚性逻辑具有封闭性的特点，要求推理所需的信息是确定不变的全信息，很难对不完全的信息和非恒常的知识系统进行描述，体现出时不变性、演绎性和单调性的特征。

在人工智能研究中，要想让机器拥有智能，智能的判定标准成为首先要解决的问题，对此，图灵提出了图灵测试，但因为该标准存在较大的主观性，所以自提出以来，相继有哲学家、认知科学家、计算机专家尝试对标准进行修订或者提出新的测试标准，其中最为著名的是1980年美国语言哲学家塞尔（John R.Searle）提出的“中文屋子”模型^[45]。中文屋子在本质上模拟的是一段计算机程序，程序的改变在塞尔看来不过是中文屋子中操作规程的改变，只能体现编写操作规程的“人”的智能，并不能反映只完成输入输出的“机器”的智能。

中文屋子反驳了强人工智能的观点并认为：“当前基于形式逻辑的二进制计算机系统不可能有任何智能，计算机系统是否具有智能，与其使用的逻辑基础密切相关。”塞尔构造的中文屋子模型其逻辑基础就是形式逻辑，在模型中，计算机只需要完成字符的匹配工作，并不要求理解字符的含义，对任何问题的回答都只有“真”和“假”两个确定性的取值。但在人类的真实思维中，还有许多不确定的内容，这些却很难用真假来做确定性的区分。

从中文屋子的例子不难看出，二值逻辑本身是排斥一切矛盾和不确定性的，基于二值的形式逻辑并不能反映人类思维的本质，因此基于形式逻辑基础上的传

统计算机结构不可能让机器真正拥有智能。

当然，塞尔的中文屋子虽然否定了当前逻辑基础下计算机拥有智能的可能性，但并非对未来的计算机是否拥有智能做出了一个消极的判断。逻辑基础在人工智能研究中起着基础性的地位，只有突破刚性逻辑的禁锢，才有可能让人工智能研究展露新的生机。因此，在后续的研究中，研究者逐渐意识到刚性逻辑（形式逻辑）的不足，并且将目光逐渐转向对柔性逻辑（数理辩证逻辑）的研究中。

3.3 包容不确定性的柔性逻辑对人工智能的促进

柔性逻辑是对刚性逻辑（形式逻辑）的柔性化，包容并拓展了原有形式逻辑的内容。柔性逻辑突破了刚性逻辑的二值特征，引入了柔性的命题连接词和柔性量词，拓展了包括假设推理、类比推理、进化推理在内的一系列柔性推理模式，又被称为数理辩证逻辑。

与刚性逻辑相比，柔性逻辑有两方面的优势，首先，柔性逻辑将形式逻辑中的精确推理推广为包容不精确性的推理，这个过程包括了把二值逻辑向三值逻辑、多值逻辑、概率理论、模糊集、粗糙集、连续值逻辑等方面发展的过程^[46]，并将形式逻辑中要求完备的信息放宽为可以在信息不完全的情况下进行推理，通过先验或者后验的信息对缺省条件进行假设性的补充，更恰当的反映了客观事物，对于解决实际问题有较大的帮助。其次，柔性逻辑将形式逻辑从一维的静态逻辑推广到多维的动态逻辑，有助于解决人工智能中的非恒常性问题。本文将以前研究较为深入的模糊集、粗糙集和非单调逻辑为例进行探讨。

3.3.1 模糊集和粗糙集在不确定性推理方面丰富了形式逻辑的内容

在由精确推理向不精确推理的过程中，模糊集和粗糙集是两种重要的逻辑基础，分别解决了同类对象隶属程度的表述和不同类对象在知识分类方面的问题。当前不少人感到区分模糊集和粗糙集有一定的困难，对此，本文通过表 3.1 区分。

表 3.1 模糊集和粗糙集的比较

类别	研究对象	研究重点	研究工具	计算过程	所需数据
模糊集	同类对象	隶属程度	隶属度函数	主观	先验数据
粗糙集	不同类对象	知识分类	集合的上下逼近	客观	不需先验数据

顾名思义，模糊集是用来表达模糊性概念的集合，由 1965 年系统科学家扎德（L.A.Zadeh）首次在《模糊集合论》（《Fuzzy Sets》）中提出并创立。模糊集出现之前的集合都是具有某种属性的对象的全体，属性的划分是确定性的，并且在对应过程中，定义域中的一个值只能在值域中找到一个唯一与之对应的值。但在

模糊集中，定义域中的一个值可能会对应值域中的两个值，但到底更偏向于哪一个值，将通过隶属度来考虑。隶属度体现了元素对集合隶属的程度，其取值范围是 $[0, 1]$ 。通过引入隶属度，使得元素和集合的关系从“属于”和“不属于”变成了“属于”、“不属于”和“隶属度上属于”三种情况。

1982 年波兰学者 Z. Pawlak 提出了粗糙集理论。粗糙集从知识分类入手，着手通过不确定性本身提供的信息来研究不确定性。其主要思想是将不确定或不精确的知识用知识库已有的知识来刻画，具体做法是在保持分类能力不变的情况下，通过知识约简（“交”与“并”的过程），推导出概念的分类规则，得到不可区分的关系和等价类划分所确定的近似域，从而找到数据中隐含的规律，把不确定性的知识放入对应的集合中。

3.3.2 非单调逻辑在动态更新方面弥补了形式逻辑的不足

在研究知识动态更新和解决非恒常问题方面，非单调逻辑是一个重要的基础。普通逻辑具有单调性。函数的单调性指：“从有序集 X 到有序集 Y 的函数，满足条件：若 $X_1 \leq X_2$ ，则 $f(X_1) \leq f(X_2)$ ，(或 $f(X_1) \geq f(X_2)$)，我们称 $f(X)$ 为单调递增(或递减)函数，简称为单调函数。”经典逻辑都具备单调性，简要地说，经典逻辑一旦往一个方向发展后，就不会再回退，是一个不可逆的过程。非单调逻辑属于非标准逻辑的一种，是在人工智能研究的大背景下产生的，非单调逻辑随着前提的增加，显示出一种不确定性的状态，其发展是可逆的，具有定义域的开放性和值域的不确定性的特点，类似于人类常识推理非单调的特点^[47]。

非单调逻辑是处理不确定性问题的一种重要的方法。例如：新知识有可能对原有的知识进行补充，这个时候体现的是一种相容的状态，原有的推理成立，继续单调的发展；但新知识也有可能会与原有的知识冲突，即出现了不兼容的问题，这个时候原有的结论将不再成立，推理将不再具有单调性，需要回过头对知识进行更新或再划分。

3.4 柔性逻辑需要进一步发展和完善

虽然柔性逻辑是研究人工智能中的不确定性问题的重要基础，目前也取得了一系列的成果，但由于相关理论发展比较晚，有的暂时还只提出了理论模型，没有足够的应用支撑，因此，即使是应用相对较多的模糊集、粗糙集和非单调逻辑也或多或少存在着一些需要进一步完善的地方。

模糊集虽然能比较好的处理亦此亦彼的模糊关系，有利于对不确定性人工智能中人的形象思维进行模拟，但存在一个比较大的局限性，即隶属度的问题。本文认为隶属度可以分为两类，一种是确定性的隶属度，一种是不确定性的隶属度。

可以通过全面的统计或科学的测量得到的隶属度，是确定性的隶属度；而受主观影响较大、调查不全面或所选的被调查者不具有群体代表性所得的隶属度是不确定性的隶属度。确定性的隶属度可以较真实的反映同类事物的隶属程度，而不确定性的隶属度推广意义不大，不适用于正式的、科学的研究。

其次，隶属度还需要前期经验知识的积累，是先验的。离开了经验，很难获得隶属度，而没有隶属度，模糊集的后期工作将难以开展。因此，能否得到一种相对客观、灵活性更强的隶属度函数，成为能否促进模糊集发展的关键。当前，隶属度函数的确立还没有一个统一的方法，比较常用的有模糊统计法、例证法、专家经验法和二元对比排序法。

模糊统计法认为不同的试验者，对同一个集合的边界认定不同，使得对一个确定元素是否属于该集合有着不同的认识。例如：有人认为 100 是大数，所以将 150 放入“大数”这个集合；而有的人认为 10000 才是大数，所以不会将 150 放在“大数”这个集合中。但在每次统计中，“150”这个数是确定性的，而“大数”这个模糊集的边界值是不确定性的。这时“150”对于“大数”集合的隶属频率将等于“150”在“大数”集合中出现的次数除以试验的总次数。随着试验次数的增加，隶属频率将趋于稳定，这个稳定值就是其隶属度。例证法是为了确定某一元素在集合中的隶属度，从而通过几种“语言真值”让认知主体进行判断，最后通过统计得到离散的隶属函数。例如：在对“上海 2010 年的房价高吗”进行判断时，所给出的语言真值分别是：“高得离谱”、“比较高”、“价格适中”、“房价偏低”，然后让用户在这四个语言真值中进行选择，从而确定“上海房价高”的隶属度。专家经验法是首先根据经验从整体上把握，确定粗略的隶属度函数，然后再根据实际效果进行“学习”，完成修正。二元对比排序法是按照某一特征对多个事物进行两两比较来确定其排列顺序，由此得到隶属函数。

这四种主要的方法都是从不同的角度出发，对目标进行模糊性的刻画。但是本文认为不同的方法有不同的适用性，如果能针对不同的事物在不同的阶段综合选取适应的方法，将得到较为客观和相对准确的隶属度。

模糊统计法和例证法的缺点在于过于主观，并且模糊统计法可能导致计算量非常大，不适合于及时地处理和更新；而例证法首先在“语言真值”的确定上就过于依赖智能主体，很难找到最准确的语言真值，因此是四种方法中局限性最大的一种。专家经验法通过回馈修正，灵活性、针对性比较强，但由于初始阶段需要大量知识的积累，仍不能避免先验的问题。二元对比排序法在算法上容易实现，并且可以弥补普通情况下模糊集需要先验知识的局限性。在二元对比排序中，只要有元素，就可以通过两两排序，得到一个粗略的隶属度，随着排序元素增多，隶属度也逐步趋于稳定。因此，在实际运用中，为了得到相对客观的结果，可以

在隶属度确定的初始阶段（即使没有先验知识也没有关系）使用二元对比排序法，然后结合专家经验法和模糊统计法进行回馈和修正。

粗糙集与模糊集相比有一个重要的优势在于不需要先验数据，得到的数据比较客观，但其本身仍然存在一定的局限性。例如：粗糙集在应用于智能控制中，通常会形成决策表并从中总结出“**If** 条件 **P** 满足 **Then** 采取策略 **Q**”的控制规则，在条件不是很复杂的情况下，可以较快的约简得到相应的策略。但是如果将应用推广到地震预报、冲突分析、近似推理、投资分析、模型建立等需要大规模的数据，并且数据之间的关系错综复杂的问题上，不可避免的会出现数据之间的冲突或同一数据具有过多的属性值的情况，如何高效的进行约简并做出正确的决策成为比较困难的问题，如果解决不好这个问题，将导致组合爆炸，粗糙集的优势也难以体现出来。

由于经典的粗糙集理论是基于完备的系统来进行推导和制定策略的，因此在处理不完备系统的时候，容易因为样本数据丢失、损坏和不全面影响到推导和决策。目前通常采用的方法是对缺失的数据进行处理，将不完备的系统尽可能补充成为完备的系统，建立扩展粗糙集模型，因此如何对缺失数据进行处理将在一定程度上影响到粗糙集决策的优劣。

此外，粗糙集要求所处理的信息是离散的，因此当前在病例诊断、股票分析、信贷决策等应用中逐渐展现出优势。但是，人们除了需要处理离散的信息还不可避免的要处理连续的信息，对于粗糙集来说，连续的信息将导致很难进行类属的划分，即使划分出来，也将因为分类过细导致复杂性大大增加，失去了不确定性问题处理的低代价的优势，因此在处理连续数据的时候，可以考虑将粗糙集的思想 and 模糊集的理论结合起来，相辅相成解决问题。

非单调逻辑在不确定性人工智能的研究中虽然较好地符合了人类常识推理的特性，但由于更偏向于思维层次的否定之否逼近真理的过程，而不是一种具体的操作方法，因此不能单独使用。要想解决不确定性的问题，必须根据待处理问题的特点，有针对性的将非单调逻辑和概率理论、模糊集、粗糙集等具体方法有机结合起来，从而不断地完善、逐步求精，最终得到满意的结果。

在处理不确定性问题方面，这些新兴的逻辑都在不同方面具有优势，但由于应用还比较少，有的局限性还暂时不为人所知，因此本文对应用比较广的模糊集、粗糙集、非单调逻辑提出了个人的建议，而不是对这些方法的否定。在研究人工智能中的不确定性问题的过程中，很难出现一种统揽全局的十全十美的逻辑，相反，正如其不确定性的特征一样，这是一个包容性很强的领域，允许不同方法取长补短，相互完善，共同为人工智能的发展搭建坚实的逻辑基础。

作为客观事物发展规律和人类思维规律的抽象性表述的逻辑在人工智能中具

有基础性的作用。随着人工智能的进一步发展，需要包容确定性和不确定性在内、相容并且能较准确描述人类认知和思维的统一的逻辑基础，何华灿教授于 2005 年首次在《泛逻辑学原理》一书中提出了“泛逻辑概念”，从底层自下而上为人工智能研究搭建更具包容性的基础，在国际上处于领先水平。泛逻辑实质上仍是以柔性逻辑（数理辩证逻辑）为核心，把刚性逻辑（形式逻辑）作为柔性逻辑的特殊形式来处理。泛逻辑学为了将人工智能搭建在互不相容的逻辑基础上，对柔性逻辑连接词等进行了统一，并努力搭建起了刚性逻辑和柔性逻辑之间的联系，尝试建立能包容一切逻辑形式和推理模式的逻辑基础，进一步促进人工智能的发展。

第四章 人工智能研究方法的不确定性

符号主义、联结主义、行为主义曾经是传统人工智能研究中的主要研究方法，但随着对人工智能中的不确定性问题研究的深入，人们开始关注和探索包容不确定性的新的研究方法。由于在该领域起步晚、研究难度大并且具有一定的颠覆性，因此尚未与实际应用大量结合，取得的进展更多的局限在理论方面。目前，我国学者钟义群教授提出了针对智能内在生成机制的机制主义研究方法较好的将符号主义、联结主义和行为主义统一在了同一理论框架下。

4.1 智能内在生成机制的研究方法探悉

《2009-2010 智能科学与技术学科报告》把机制主义作为我国在人工智能领域的突破性理论，引起了国内外的普遍重视^[6]。

机制主义认为虽然功能、结构、行为是一个系统的重要属性，但对于智能系统来说，揭示其本质的不是功能、结构或行为，而是“智能生成机制”，而智能生成机制是一个“信息——知识——智能”的转换过程。在机制主义框架下，三种研究方法可以实现比较完美的互补和统一，例如：联结主义获得的经验性知识经过验证可以成为符号主义所需的规范性知识；符号主义获得的规范性知识经过普及可以成为行为主义所需的常识性知识^[12]。

在机制主义框架下，整个智能生成过程可分为四步：首先是从本体论信息到认识论信息的信息获取过程；接着是从认识论信息到知识的认知过程；然后是从知识到智能策略的决策过程；最后是从智能策略到智能行为的执行过程^[12]。

本文认为机制主义之所以能和谐地将三种研究方法结合起来，不容忽视的一点是机制主义充分考虑到了不确定性的存在，并把不确定性融入到了传统的人工智能研究中。因此，对照智能生成过程的四个步骤，分为信息获取的不确定性、主体认知的不确定性、形成知识的不确定性、主体决策的不确定性以及执行结果的不确定性五个方面进行讨论。

在信息获取过程中，不确定性包括了信息的随机性、中间态和不完备性；在认知过程中，不确定性包括了感觉的不确定性、知识结构的不确定性、记忆的不确定性、思维的不确定性；在知识形成过程中，不确定性不仅包括了随机性、模糊性、不完备性，还包括了知识的不协调性和非恒常性；在决策过程中，由于决策本身就是在两个或两个以上备用方案中选择一个的过程，所以也是不确定性的。虽然从智能策略到智能行为的执行过程是确定性的，所谓的确定性是在策略制定后会产生相应确定性的行为，但执行的结果是不确定性的。例如：农民根据以

往春天种植棉花丰收的经验，做出了在春天种棉花的决策并付诸实践，即使主观上辛勤耕种，但可能当年遇到干旱或棉铃虫害，导致颗粒无收。

整个智能生成过程，对于单个智能主体来说，是完整的一环，但由于智能主体的智能行为并不会孤立于环境而单独存在，所以对于多个智能主体来说，以上过程会更加复杂，充满了不确定性，例如智能主体 1 的智能执行结果或许只是智能主体 2 的本体信息获取中的一部分；在多个智能主体形成知识、策略以及执行的过程中，不同主体会相互影响、干扰、促进、牵制等等，相互间的简要关系如图 4.1 所示。

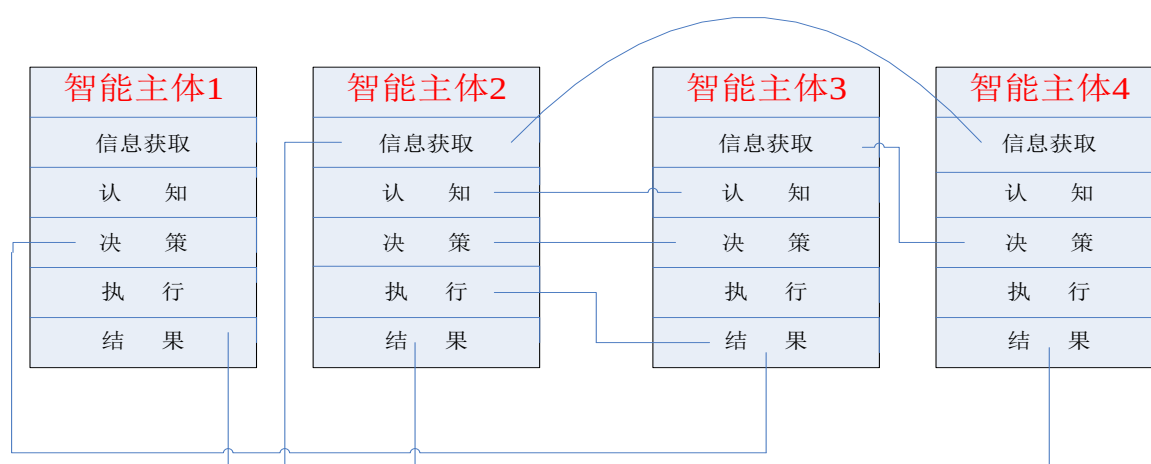


图 4.1 智能主体之间的简要关系

4.1.1 信息获取的不确定性

在从本体论信息到认识论信息的“信息获取”过程中，事物的本体论信息指的是事物对其运动状态及变化方式的自述；而事物的认识论信息指的是认识主体关于该事物运动状态及变化方式的表述。认识主体要想完全掌握事物，必须把握住事物的全信息。但是在实际过程中，认知主体（人）的各种感觉器官，例如视觉、听觉、味觉、嗅觉、触觉等都不可能把握住全部信息，必然会对信息有所筛选。以视觉为例，如图 4.2 所示，调整人眼距离图的远近，不同的人会对圆圈的旋转状态做出不同的判断；再如图 4.3 所示，不同的人会数出不同的黑白点的数目，其原因在于人的感官在对信息筛选的过程中存在着不确定性。虽然 Shannon 在《信息论》中认为：“信息是消除不确定性的东西”，但在智能主体筛选的过程中实际上已经出现了不确定性。

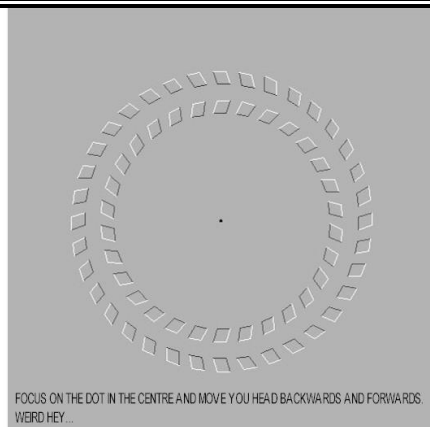


图 4.2 旋转圈

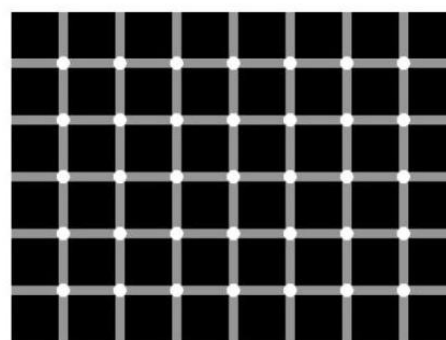


图 4.3 黑白点

在本体论层面上，智能主体所获取的信息是不确定性的，通常可以分为三类：随机信息、中间态信息和不完备信息，分别体现了随机性、中间态和不完备性。

随机性是由于客观条件不充分或偶然因素的干扰，已经明确的几种结果在观测中按照一定的概率出现，可以归纳到时间上的运动性的范畴。随机事件介于必然发生事件和不可能发生事件之间^[5]。例如在遗传的过程中，基因在复制的时候总会出现随机性的改变，使得儿女在保留父母的重要特征的基础上，在细节上总存在或多或少的差别；在做决策的过程中，既定的方案很可能因为突然遇到的人或事物（如：领导、海啸、传染病）等而改变，而这些人或事物的出现是随机性的。虽然有时候随机性会带来不好的方面，但更多时候，随机性为生活中带来了意想不到的惊喜。随机性使得基因突变、物种多样化，人从猿类中进化出来；随机性使得每个人都将面临各种各样的机遇和挑战，做出不同的选择，推动历史发展成现在的模样。

中间态是由于事物类属划分的不分明或对象在概念正反之间处于亦此亦彼的形态而引起的判断上的不确定性^[5]，可以归纳到空间上模糊性的范畴，例如：美和丑、秋和冬、难和易，但美丑、秋冬、难易并没有一个准确的界限，不同的人因为主观感受的不同，对这些概念的分界也存在差别。如果强行进行划分，把模糊的事物精确化，将导致类似“秃头悖论”的现象发生，将会大大增加方法的复杂性，耗费更多的时间和空间，并且所得到的结果也未必具有实际意义。模糊性的存在使人和动物能够在最短的时间内把握住环境的主要特征并采取相应的行为；使得音乐、绘画等艺术更加震撼心灵，唤起人们内心不同的情感；使得文学作品“只可意会、不可言传”，给人以无尽的想象空间。

不完备性指由于条件限制、认知能力等原因导致不能获得事物的全部信息^[5]，也可以归纳到空间上的模糊性的范畴，包括了“灰色信息”和“未确知信息”等内容，其中“灰色信息”是指获取信息中的未知部分，而“未确知信息”是指因

条件不足导致对获取信息在认知上的不确定性。不完备性有两种不同的情况，一种情况本身是确定性的，只是限于当时条件的限定表现出不确定性，随着条件的成熟可以从“黑盒”或“灰盒”状态变为“白盒”状态；另一种情况是不完备信息本身就是不确定性的，现在无法确定将来也不可能确定。三种类型的不确定信息的区别如表 4.1 所示。

表 4.1 随机信息、不完备信息、中间态信息的区别

信息类型	本身是否确定	导致不确定性的原因	是否为全信息
随机信息	确定	客观条件，主观因素	是
不完备信息	确定 / 不确定	客观条件，主观因素	否
中间态信息	不确定	类属难以划分	是

值得注意的是为了简化，所以在研究时预先假定了智能主体获取的是一维信息，而这个一维信息就有以上三种可能的形式，对于实际情况而言：智能主体往往在同一时刻就会获取多维的信息，形成由信息构成的庞大的信息系统。在信息系统中包括了视觉、听觉、嗅觉、味觉、触觉等多维信息。仅就视觉信息而言，就可能包括了一个巨大的场景，其中有植物色彩的信息、动物种类的信息、人物表情的信息、环境布置的信息等等，还可能融入其它系统的干扰信息和冲突信息。这样一来，其复杂度可想而知，在筛选、提取、辨认、识别的过程中也不可避免地会导致更多的不确定性因素。

4.1.2 主体认知的不确定性

不确定性是真实世界客观存在的本质属性，并不是由智能主体引起的，但由于认知是客观事物在智能主体中的映射，因而主体认知也体现出了不确定性，主要包括了两个层面的内容。首先从单个智能主体角度出发，由于感官、知识结构、记忆、思维存在着不确定性，导致主体在认知过程中体现出不确定性；其次从社会角度出发，按照不同的评价和参考标准，即使智能主体自己认为对该事物的认知是客观的、确定性的，但是在其他人看来，仍然可能认为该智能主体的认知是不确定性的。这是因为社会由不同的人组成，而每个人都是不同的智能主体，在认知方面都不可避免的存在着不确定性。因此，本文主要站在单个智能主体层面上进行探讨。

(1) 感觉的不确定性

感觉是指智能主体接受外界以及身体内部组织和器官的刺激的特性。根据刺激来源不同，可以分为体外刺激和体内刺激两部分。其中，来自体外的刺激由眼睛、耳朵、鼻子、舌头、身体各部分的感官感觉到，体现为视觉、听觉、味觉、

嗅觉和触觉；来自体内各组织和器官的刺激包括了舒服、疲惫、快乐、伤心等内在的感受。智能主体的生存离不开环境，在与环境交互的过程中，感觉将本体论信息主体的各种特征和属性尽可能全面综合地传送给智能主体，辅助主体形成对事物的认知^[48]。

“耳听为虚，眼见为实”，但感官的不确定性将导致“眼见未必为实”，这是真实存在的“视觉后象”给认知主体带来的错觉。后象是指光刺激视觉器官后，视觉细胞在刺激消失后仍在短暂时间内保持着兴奋的现象，使人产生错觉。但错觉并非就是坏事，例如人们观看的电视和电影并不是连续的，只是当胶片以每秒 24 张的速度播放时，后象带来了视觉的残留，让人产生画面是连续的错觉，没有视觉后象的存在，将无法享受观影的乐趣。其实，人的眼睛还有许多特点，例如：缪勒·莱依尔错觉利用了人眼在看向外延升的图形时会产生距离较大的眼动，在看向内收敛的图形时会引起较小的眼动的特点，如图 4.4 所示；菲克错觉利用了人眼做上下运动时比做水平运动困难，导致看垂直线比看水平线吃力的特点，如图 4.5 所示。

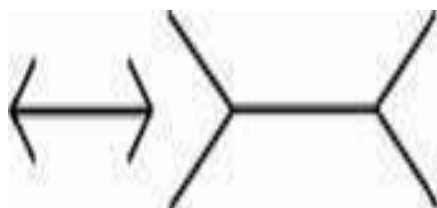


图 4.4 缪勒·莱依尔错觉



图 4.5 菲克错觉

再如：当 50 个人排成一队玩一对一的传话游戏时，由于看不到前面人的口型，只能靠听觉辨别，往往第 50 个人最后说出的话与第一个人刚开始说出的话大相径庭。这是因为在传话过程中，语句的长短、传话人的心态以及音色、语调、音量，听话人的听力、身体状态、注意力、记忆力等导致了听觉的不确定性。

同样的，嗅觉、味觉、触觉也具有不确定性，“香”、“辣”、“冷”这些概念都是笼统的，很难定量的分析出什么是香、辣、冷，这是因为感觉器官无法精确的定义出有多少个令人愉悦的气体分子，辣度达到了多少个斯高尔威单位（SHU），水温是零下摄氏几度。

虽然感官无法测出分子数目、斯高尔威单位、摄氏温度，但至少还可以通过仪器精确的测量出，尽可能逼近真值。但是即使感官能够像仪器一样精确的测量出真值，但由于不同主体对同一个概念界定的标准和范围不同，也很难确定性的划分出所有主体都认可的“香”、“辣”、“冷”。对于舒服、疲惫、快乐、伤心这些感受来说，连能让人信服的测量仪器都没有，纯粹依靠主观的感觉，要界定清楚不同的状态更加困难。此外，人作为情感复杂的生物，有时候在认知的过

程中还伴随着思维的参加，可能出现隐瞒真实感觉的现象，将导致认知过程出现更多的不确定性。例如：即使很不喜欢生日礼物的颜色和款式，但为了考虑到送礼的朋友的感受，仍然表现得非常开心。这时，站在观察者的角度，将很难判定主体的真实感觉。

（2）知识结构的不确定性

人们往往无法捕捉到环境的全信息，必然会对信息有所筛选，而该筛去什么，保留什么，常常建立在主体先前活动和经验的基础上。主体的地域特征、经验知识、背景，甚至个人兴趣、性格、价值观念等都会影响着认识主体对本体论信息的把握，导致了认知主体在把握信息时的不确定性。以对人和计算机的认识为例，随着人类知识的扩充，在认知过程中经历了“人是机器——人不是机器——人等于机器——人不等于机器”的路线^[32]，其中的每一个步骤从当时的知识水平来看都是能得到普遍认可的，但随着科学的进步，知识结构的更新，导致原来认为是正确的知识被新知识替代，呈现出否定之否的辩证发展过程。知识结构的不确定性还会导致关注点的不同，使得认知主体会从同一个全信息中提取不同的部分信息。例如：当看到一只苏格兰折耳猫时，小孩更多关注的是猫的可爱，想购买猫的人更多关注的是猫的价格，而兽医更多关注的是猫的身体状况。

（3）记忆的不确定性

记忆是对输入信息的编码、储存和提取的过程，但记忆并不牢固，存在一个遗忘过程，德国心理学家艾宾浩斯（Hermann Ebbinghaus, 1850-1909）通过大量实验，绘制了描述遗忘过程的艾宾浩斯遗忘曲线。在记忆过程中，由于不同的认知主体生理特点、生活经历、记忆习惯、记忆方式、记忆特点、记忆时间、兴趣方向等方面的不同，大脑通常会抓住认知主体最感兴趣的部分，其它细枝末节将随着时间流逝而遗忘。例如参加朋友的生日宴会，在刚开始能够回忆起宴会上水晶灯的造型、嘉宾穿着的服装、餐桌上佳肴的美味、敬酒时说的打趣话等等，但随着时间推移，记忆内容的色彩、动感、声音等元素开始衰退或出现偏差，很可能到最后只记得自己通过参加宴会认识了心爱的女孩，其它的都忘记了。

此外，人类还存在着一种潜意识的记忆，可以通过反复或强制的手段进入人脑，从而冲淡、更改原有的记忆，也将导致记忆的不确定性。例如，人脑可能出于某种目的假想了一个并未真正发生的场景，但通过反复回味、强迫来相信该情景的真实存在，让大脑逐渐构造出一个有声有色、身临其境的场景，最后连构造者本身也逐渐相信该场景的真实存在，并牢牢的留存在记忆中。

（4）思维的不确定性

思维是人脑对客观现实的概括和间接的反映，包括逻辑思维和形象思维两种形态。逻辑思维往往是确定性的，更多强调的是理性，可以抽象为符号语言并通

过计算机来模拟实现；而形象思维包括了联想、创造、顿悟等形式，包含了更多感性上的元素，没有确定性的规律可寻，但形象思维在人脑中却占据了相当大的比例，是人脑智能的重要体现，导致了人类思维的不确定性。

逻辑思维需要步步推理，是一个线形运动的过程；形象思维是把所接触到的形象性材料进行瞬间的整合，是一个并行的、跳跃的过程，有利于智能主体快速地从整体上大致把握信息，而并不要求所把握的信息如逻辑思维所需要的那么精确。

4.1.3 形成知识的不确定性

智能主体通过认知过程获得了知识，但所形成的知识具有不确定性，体现为随机性、中间态、不完备性、不协调性和非恒常性^[5]。知识中的随机性、中间态、不完备性和获取信息过程中的类似，对于智能生成机制第二步所形成知识的来说，更重要的是不协调性和非恒常性。

(1) 不协调性

知识的不协调指知识内在的矛盾，按不协调的程度可以依次分为冗余、干扰和冲突，是不确定性在空间层面上的体现。冗余是指人们在解决问题的过程中遇到的重复的知识。冗余知识可以作为增强系统鲁棒性的备份，但也会导致资源的浪费，加大系统的时空开销，例如在 Internet 搜索“冗余”一词，将出现一千多万条相关结果，由于内容大致相似，其冗余程度可想而知。一般来说，在冗余知识互不矛盾、互为补充的情况下，冗余可以增加知识的鲁棒性，但会占据更多的存储空间并且导致搜索和选择的困难。通常情况下，冗余对知识和系统的影响程度不如干扰和冲突大。

干扰是指对当前待解决的问题不但没有帮助，反而会对其他知识起到阻碍、抑制的作用，甚至导致错误结果的现象。例如：小道消息会干扰人们对正式发布消息真实性的判断；磁暴、电磁波会干扰有用信号，导致接收到的信息不完全、模糊不清甚至无法接收。在一个复杂系统中，干扰是不可避免的，其影响的程度不但和干扰源的性质和强度有关，还与系统本身的稳定程度、抗干扰能力有关。

冲突指知识之间相互抵触或完全对立。冲突是矛盾，而矛盾存在于一切事物中，知识也不例外。任何事物都有两面性，新旧知识的冲突并非只是导致系统崩溃的坏事，还可能促使知识体现出一种非单调的逻辑，鼓励人们通过研究探索得到更完全的信息，来修改和补充原有的知识，甚至推翻原有的知识重新构建更新的知识，使得人类在对立统一中逐渐逼近真理，但逼近真理的过程往往伴随着不确定性。

(2) 非恒常性

非恒常性是不确定性在时间层面上运动性的体现，主要指知识随时间的变化而变化的特性。在变化过程中，可能伴随着从不知道到知道，从不深刻到深刻，不断更新的否定之否的过程；也可能因为记忆的衰退、遗忘导致已有的知识逐渐淡出人脑；还可能因为智能主体观念的彻底改变，例如当智能主体从绝对的唯心主义者变成了辩证唯物主义者，很可能会推翻原有的知识构建。

4.1.4 主体决策的不确定性

决策是智能主体在多个备选方案中选择一个的过程，从定义上看，多选一是概率问题，所以决策必然伴随着不确定性，此时的不确定性主要由人引起。在从知识到智能策略的决策过程中，智能策略的生成离不开对问题和环境的充分掌握，不同智能主体的出发点和目标不同，即使所有的初始条件都相同，也可能做出迥异的决策；同一主体在不同时刻、不同环境、不同心境、不同知识结构的情况下，做出的决策也会大不一样。例如在秦二世时期，大臣赵高“指鹿为马”，大臣们根据知识的储备和感官的判断都知道是鹿而不是马。但大多数人为了保住爵位和性命，纷纷做出了“指鹿为马”的决策。

为了做出更优化的决策，不仅需要经验的积累，还需要尽可能多的掌握信息。在平时生活中，可能随意做出的决策无关痛痒，但在战争中，决策的关系极大，可能改变个人甚至一个国家的命运。

以战国时期“马陵之战”为例，对阵双方主帅是魏将庞涓和齐将孙臆。为避魏军锐气，孙臆想到只能智取，不可硬夺，便设计了“增兵减灶”的计策。庞涓照搬了兵法中“百里而争利，则擒三军将，劲者先，罢者后，其法十一而至；五十里而争利，则蹶上军将，其法半至；三十里而争利，则三分之二至。是故军无辎重则亡，无粮食则亡，无委积则亡^[41]”的道理，认为齐军百里兴师入魏境，战线太长，粮草难以供给，必不能久。所以当他看到齐军第一天造十万个灶，第二天减为五万，第三天减为三万的表面现象时，就做出了率骑兵日夜兼程追赶齐军的错误决策，导致命丧马陵道，所率骑兵的性命也因主帅的错误决策而葬送。“怒可复喜，愠可复悦，亡国不可以复存，死者不可以复生。”决策的重要性可见一斑。

《孙子兵法》在《谋攻篇》中谈到：“知彼知己，百战不殆；不知彼而知己，一胜一负；不知彼不知己，每战必殆。”进而在《地形篇》中补充到：“知己知彼，胜乃不殆；知天知地，胜乃可全^[41]。”意思是在作战中，要尽可能掌握更多准确的信息，具体来看，是在完全知晓己方“道”、“天”、“地”、“将”、“法”的情况下，通过“乡间”、“内间”、“反间”、“死间”、“生间”等情报人员的灵活运用，以便更充分掌握敌情，做出正确决策，减少部队伤亡，力

争达到“不战而屈人之兵”的全胜状态。

4.1.5 执行结果的不确定性

做出适宜的决策并不是一件容易的事情，从智能策略到智能行为的执行过程本身是确定性的，而要想执行结果能朝着预期的方向发展更是受限于多方面的条件，导致最终结果充满了不确定性，同时包括了时间上的运动性和空间上的连续性两个方面的内容。对于单个人来说，执行的最终结果会受到自然环境、社会因素等各方面个人不可控的客观因素的影响，差之毫厘，失之千里。对于博弈双方来说，单方面的决策即使是智能主体觉得最优化的，其执行结果不但与环境等客观条件息息相关，还取决于对手的决策。而当把情况拓展到多个博弈方时，情况将会变得更加复杂。

由于不确定性在人工智能的研究中起着重要的作用，所以对于一种研究方法来说，如果只做确定性的考虑，虽然能取得阶段性的成绩，但在更深层次的研究中，将面临巨大的挑战。以符号主义为例，其发展从自动推理到启发式搜索，再到基于知识领域的专家系统和知识工程，体现了知识（特别是专业领域知识）在高效解决各类复杂问题中的重要性。但是，基于领域知识的推理常常会使用到不完全、非二值、非精确、不确定、有矛盾和可变化的知识，常识推理在其中扮演着重要角色，这些暴露了二值逻辑的局限性，从而诱发了 20 世纪 80 年代中期人工智能的理论危机，导致人们不得不寻求新的智能模拟方法来化解，而新兴的智能模拟方法将或多或少考虑到人工智能的不确定性特征。

4.2 面向物质基础的研究方法是人工智能的突破口

1996 年美国加州大学的扎德（L. Zadeh）教授提出了“硬计算”和“软计算”的概念，顾名思义，硬计算是对精度的探求，要求通过在运算中准确的取值从而得到精确的结果，有利于数学建模和精确的控制行为；而软计算则无需精确的数学或逻辑模型便可对输入数据进行处理得到结果，允许取值不确定、不精确和不全真，具有较高的容错能力，能够用代价较低的解决方案较好的模拟自然智能系统。按照这样的划分，逻辑计算、经典集合论等将归于硬计算，而人工神经网络、遗传算法、概率论、云模型、混沌模型等将归类于软计算，即传统的人工智能更多的使用了硬计算的方法，而在当前处理人工智能的不确定性问题方面更多使用的是软计算的方法。

类似于扎德教授提出的软计算和硬计算，本文将不确定性人工智能分为软不确定性人工智能和硬不确定性人工智能两类。软不确定性人工智能是指通过软件的方法来体现不确定性的特征，而硬不确定性人工智能是指从硬件结构上来体现

不确定性特征。两种方法各有优劣，软不确定性人工智能的方法在实现起来比较容易，但有一定的局限性；硬不确定性人工智能的方法虽然很可能起到立竿见影的效果，但是彻底的改变原有的冯诺依曼体系结构将导致计算机历史上的重大变革，具有较大的风险性，实现起来也非常困难，需要认知科学、脑科学、信息科学、心理学等相关领域的专家协同完成，并且适应于原有硬件结构的软件很可能不适应于新的硬件体系结构，需要重新开发相应的软件，实现周期比较长。在现有体系结构框架下，虽然提出了各种各样的方法策略，但基本上都是在软不确定性人工智能研究方面进行补充和完善，并没有从根本上寻求突破。要想彻底跳出传统人工智能研究的框架，需要在硬不确定性人工智能上寻求突破，通过物质基础的彻底改变，帮助机器产生“意识”。

在物质和意识第一性的问题上，马克思辩证唯物主义观点从“物”的角度出发，肯定了物质是意识产生的前提，起着基础性、本原性和决定性的作用，坚持物质决定意识；同时也肯定了意识不仅仅是从属、依赖于物质，还会对物质产生能动的反作用。

为了更好地解释清楚物质和意识的关系，有的学者将“实践”作为物质和意识辩证关系的桥梁，认为实践不仅仅是物质性的存在，还带有意识、思维等精神因素，实践是自我意识、社会、自然界进行沟通的多维平台，有利于缩短“物质”与“意识”之间的距离，更好地联系物质世界和精神世界^[49]；有的学者从实践论出发，用“实践”与“物质”的关系取代了原有“物质”和“意识”的关系，认为实践是对物质的能动作用和改造，是物质与意识的统一，这样的观点不再将“人”强行一分为二（即把人的肉体归为物质，而把思维归类于意识），而将人看作一个整体，“存在”由“人”和“物质”两类组成^[50]；还有的学者认为意识之所以能在人脑系统中产生出来，体现出原先“组分事物”（组成系统的物质）不具备的整体特征，是因为在系统产生的瞬间，“统性事物”（组分事物在构成系统瞬间产生的在物质、时空、结构、功能等方面你中有我、我中有你的部分）也随之产生^[51]。

虽然目前人类还不能完全了解智能的生成机理，但可以肯定的是智能的产生离不开人脑这个物质基础，而且很可能就是人脑系统中产生的“统性事物”，从而展现出任何一个脑细胞或神经元（组分事物）都不具备的能力。

“物质决定意识”的观点已经被越来越多的人接受，对于同样的物质基础来说，产生的意识即使有所不同，但不同之处更多体现为程度上的差异；但对于不同的物质基础来说，很可能产生出根本上不一样的意识，这是在人工智能研究中不容忽略的一点。人工智能希望用计算机来实现人类的智能，对于计算机来说，其物质基础是硬件和支撑硬件运行的软件。传统的计算机从诞生到现在，虽然在

硬件上材料、工艺等方面有了巨大的改进，但其最根本的硬件基础却始终是基于冯·诺依曼体系结构的，因此即使其上运行的软件随着技术的进步，界面上更加人性化，速度更优化，算法更精良，终究仍会受到硬件结构的限制，“MOORE 定律”难以维继已经很好地说明这一问题。

4.2.1 传统计算机的物质基础

传统计算机又称为经典计算机，基于冯·诺依曼体系结构，以经典物理学为信息处理的理论基础，经历了电子管、晶体管、集成电路、超大规模集成电路的发展历程。传统计算机以存储程序原理为基础，存储程序原理的基本思想是指令驱动以及指令和数据混合存储。程序执行时，中央处理器 CPU 按照程序指定的逻辑顺序把指令和数据逐条从存储器中取出并加以执行，自动完成程序描述的处理工作。存储器是按地址访问、线性编址的一维结构，每个单元的位数是固定的，指令在存储器中按其执行顺序存放并由指令计数器指明要执行的指令所在的单元地址。

冯·诺依曼在谈到人脑优越性时指出，大脑的语言不是数学语言，其巨大功能是通过并行加工信息的复杂系统来实现的，然而当前体系结构的计算机的计算在根本上都是按序串行实现的^[38]，因此即使使用了最高效的 CPU，也不可避免的要重复完成“取指令”和“执行指令”的循环操作。即它从主内存中取出一条指令和数据，处理这条指令，然后再做下一次的取值和处理操作。从本质上讲，冯·诺依曼体系结构存在两方面的一维性，即一维的计算模型和一维的存储模型，限制了计算过程并行性的开发和利用，迫使程序员受制于“逐字思维方式”的传统程序设计风格，语言臃肿无力，缺少必要的数学性质，使得程序复杂性无法控制，软件质量无法保证，生产率无法提高，在成就传统计算机辉煌的同时也限制了它的进一步的发展。

此外，传统的计算机通过增加时钟频率和数据位宽度来提升性能，这种性能的增加主要体现在 CPU 上，而不是内存数量的增加上，并且通过增加时钟频率的方法并不能等倍数的提升计算机的性能。例如，把计算机的时钟频率增加一倍或内存增加一倍并不会使计算机性能增加一倍。此外，冯·诺依曼计算机的结构不利于处理人工智能中的非数值问题，并且机器语言与自然语言在语义上存在的较大差别。机器语言对数据的处理需有严格的形式和访问方式，这与自然语言处理的随意性有很大的差别。如何消除形式与语义的间隔，是经典计算机面临的重大难题。通常使用计算机需要经历“分析问题——建立数学模型——设计算法——程序编写——程序执行”的过程，而且所编写的高级语言与机器语言同样存在严重的语义阻隔，在高级语言中各种形式的数据结构最终在存储器上都将转换成一

维的线性存储模型进行存储。随着人工智能对计算机性能的要求越来越高，从硬件方面突破人工智能的极限成为研究人员越来越关注的问题。

4.2.2 人工智能的发展催生新型计算机的诞生

为了克服这一困难，当前兴起了对量子计算机、生物计算机、光子计算机、纳米计算机等硬不确定性人工智能的研究，希望从硬件基础上颠覆原有的冯·诺依曼结构计算机模型。虽然相关研究还处于尝试阶段，但已经在特定领域展现出了传统计算机无法比拟的优势，目前进展较大的是量子计算机和生物计算机。

量子计算机是以量子物理作为信息处理基础的新型计算机。量子计算是用量子比特来代表信息，用量子力学理论来描述有关信息表示和信息处理的问题，其中信息的演变遵从薛定谔方程，信息计算和信息处理通过量子比特完成一系列么正变换，信息的提取是对量子系统的量子状态进行测量坍塌的结果。

经典计算机以比特作为信息单位进行编码的，一个比特可以表示为 0 或 1 两个状态，用于区分电容器带电与否。相应地，量子计算中量子信息的单位称为量子比特。它除了处于 0 态或 1 态外，还可处于叠加态。叠加态是 0 态和 1 态的任意线性叠加，既可以是 0 态又可以是 1 态，0 态和 1 态按照一定的概率同时存在，并在测量或与其它物体发生相互作用的过程中呈现出 0 态或 1 态。任何两态量子系统都可以用来制备量子比特，例如，用 0 代表基态，用 1 代表激发态，则当有两个量子比特时，它们可处的量子态就是 00, 01, 10, 11 四个状态的叠加态。

量子计算的原理是使量子态所处叠加态的各个基态通过量子门通用集合的作用相互干涉，从而使各个基态间的相对相位和概率幅度发生改变，造成所观测的各个基态的发生概率产生变化，最终使叠加态发生变化。量子算法中融入了量子力学的许多基本特性，如并行性、相干叠加性、纠缠性、测量坍缩等，极大地提高了计算的效率，已经逐步成为一种崭新的计算模式。例如：在求解大数质因子分解时，常规计算机所需的时间则为 $O(2^{n/2})$ ，当 n 变得很大时，常规计算机将显得无能为力，所以大数因子分解是计算机难解问题，已广泛用于密码通信中，有较好的加密功能。但对于量子计算机来说，只需 n 的多项式时间，就能解决该类传统计算机中的难解问题，有效的完成密码破译工作。如果量子计算机广泛应用在数据加密中，常规的加密算法将形同虚设，这从一个侧面展示出量子计算机的优势。

另一种具有广泛前景的计算机是生物计算机。生物计算机在结构上完全不同于经典计算机，它通过寻找与大脑近似的物质基础，借此产生能够与大脑相似的智能，促使计算机与人脑融合。生物计算机可以直接接受人脑的指挥，成为人脑的外延或扩充部分，改善和增强人的记忆力，还可用于解决大量非同寻常的问题，

例如传统计算机无论使用运算速度多快的机器，采用多么优化的算法，在密码匹配方面都远远比不上生物计算机。

2008 年 8 月，英国里丁大学制造出了世界上第一台生物控制的机器人——米特·格登（Meet Gordon）。格登是一个鼠脑机器人，培养皿中的营养液为鼠脑神经提供营养，通过突触传递电信号来进行链接、放电和组合。在神经元的协同工作下，鼠脑机器人可以通过无线蓝牙来遥控机器躯体完成简单的动作，实现自主学习、自适应的过程。它的创造者凯文·沃维克（Kevin Warwick）教授认为神经元都是靠神经递质来传递信息，老鼠和人类的智能差异在于神经元的数量而不是神经元的质量，鼠脑只有 100 万个神经元，而人脑却有 1000 亿个。鼠脑机器人的物质基础不再是传统的电脑，而是培养皿中的脑神经元，更类似于人脑的结构；神经元的相互关系不再是确定性的 0, 1 组合，而是在没有预先编入程序的情况下，让神经元通过感受外部环境的刺激，从而在相互之间建立起连接和沟通，自主的完成控制动作，体现出了虽然初级但却可以算作真正意义上的智能，大大鼓舞了人类对开展生物计算机研究的信心。

4.2.3 通过在硬件搭建中引入不确定态来实现硬不确定性人工智能

当前生物计算机虽然较真实的模拟了人脑结构，但总体上还处于试验阶段；量子计算机研究相对深入，但仍然存在着较多的不完善，例如量子计算机虽然具有多个量子状态，但是在实际引用或运算过程中仍然会转变为确定性的状态，即用 $\{0, 1\}$ 系统来进行构建，在硬件上仍不能完全体现出不确定性。

本文尝试提出，为了在硬件上真正实现不确定性，能否研制构建一个三态的模型，将经典布尔系统中的取值范围由 $\{0, 1\}$ 二态推广到 $\{0, 1, \alpha\}$ 三态，其中 α 为拟人因子，标示不确定性态。

假想的 $\{0, 1, \alpha\}$ 三态模型源于传统的 $\{0, 1\}$ 模型，所以可以按照原来对 $\{0, 1\}$ 模型的认识来理解这个新模型。三态模型与传统二态模型的不同之处在于引入了不确定态 α ， α 在中间运算的过程中，始终保持着不确定态，只有在最后输出的时候，才会转变为 0 或者 1，最终得到确定性的结果。

要想应用好三态模型，在解决实际问题的时候把哪一个因素选作不确定态 α 非常重要，将直接影响到能否得到期待得到的最优化的结果。提出这个假想的模型后，有学者质疑道：“如果 α 在中间过程中一直处于不确定态，这样的运算还有意义吗？”但不容忽视的是，人脑恰巧是类似的思维模式：人脑在没有特定条件限制的情况下，往往不知道下一步到底做何种选择，而是等待进一步的约束条件，从而根据情形的需要再做出确定性的选择，但即使当前状态被论证为最优的选择也未必是整个过程中最优化的选择。因此三态模型的好处在于，给每一种

方法都赋予了参与完成最后计算的机会，而不是在第一阶段就确定性的指定为某种方法，而将第一阶段相对不优化的方法直接排除在外。事实上，有的方法在第一阶段的赋值中，并没有任何优势，甚至还可能是看上去最差的选择，但很可能在第二或者第三阶段的时候这种选择将最逼近目标，完成起来代价最小，回报最高。因此，通过引入 α 对选择进行控制，使得多种选择可以并行出现，以便主体根据情形需要做出最优化的选择。

4.3 面向软件模拟的研究方法是人工智能的保证

对于计算机来说，虽然软件上的改变不可能产生革命性的变革，要想机器真正拥有智能，相适应的软件基础同等重要。其实当前人工智能从确定性人工智能向不确定性人工智能发展的过程中，由于硬件基础始终是冯·诺依曼体系结构，所以软件的进步起到了主要作用：从经典数学和逻辑到概率理论、混沌、云模型等一系列软不确定性人工智能的方法，是一个软件改进的过程，也产生了初期的成果，解决了原有算法无法解决的问题。。

值得注意的是，在模拟人类智能的时候，即使完全制造出了具备与人脑相同结构和功能的硬件，也未必能真正拥有智能。因为人类的智能在形成过程中一方面依赖于生物基础，而另一方面是社会实践的产物和结果。智能和劳动一样，是从猿到人的进化过程中，通过与自然和社会交互而逐渐发展起来的。亚里士多德在《政治学》中说：“人类天生是趋向于城邦生活的动物（政治动物）。”马克思补充到：“人即使不像亚里士多德所说的那样，天生是政治动物，无论如何也天生是社会动物。”对此，进化论的创立者达尔文通过科学研究明确的指出：“人是一种社会动物。”

人作为社会存在的产物，在社会中生活和发展，从根本上说是一个实践的过程，而实践的基础是通过劳动来实现的。恩格斯在《自然辩证法》中单列了一节阐述了“劳动在从猿到人的转变中的作用”，劳动促进“手”从劳动的器官成为劳动的产物；劳动改进了猿类的喉头，从而产生了语言；劳动和语言一起促进了脑和为其服务的器官的进一步发展，并在交互进化的过程中，逐渐形成了人类社会，使人彻底地从猿类中分离出来^[52]。

智能的发展对于个人来说是一个有限的过程，现代人的大脑从体积、容量等方面来看，和原始人类相差无几，也就是说生理基础是相当的。但是现代人却创造了超越原始人想象的繁荣，这个巨大的差异显然不是来自于生理上的差异，归根到底是来自于社会的^[53]，社会的存在让人不仅仅拥有了感觉、知觉、记忆、想象和思维等意识，还帮助人把这些意识凝练成了知识、观点、理论、方法，并在此基础上形成了更高级的价值、信仰等。社会可以看作是一个系统，而每个人都

是其中的组分事物，当组合在一起时，使整个系统产生了统性事物，从而拥有了单个人所不具备的功能，展现出个人智能难以达到的群体智能。

从原始人的集体狩猎，采摘野果为食到现在高度分化的体力工作和脑力工作，在社会中生存的每个人都不可避免的进行着劳动，并通过劳动促进社会和个人发展。从劳动形式来看，最开始的原始劳动是体力劳动和智力劳动的统一；接着智力劳动逐渐从体力劳动中分离出来；然后人类发展了机械工业用来替代体力劳动；到现在人类开始发展智能工业，希望能够替代脑力劳动^[54]，人工智能也因此成为需要。

人工智能研究的分支比较多，除了之前基于逻辑思维的确定性研究外，当前更多地集中在语言处理、智能控制、模式识别、知识提炼等方面，而这些方面都不约而同的需要置身于社会中考虑。社会的存在，使人类学会了用语言来进行沟通交流，用文字来记录、传承思想，弥补记忆的不足；使人的感官和肢体在社会的合作和竞争中变得更加发达；使人能够快速融入群体，汲取他人积淀下来的知识，分享共同劳动创造的物质财富和精神财富。

要想彻底的实现人工智能，硬件的改变虽然是突破口，但就当前发展来看，要想在短时间内完成硬不确定性人工智能的实现还比较困难，所以软不确定性人工智能的仍是近期努力的方向。当前研究较多的运用了软不确定性人工智能的方法，如模糊集、粗糙集、非单调逻辑等，但都存在着或多或少的不足，要想真正付诸于实用，还需要考虑到“社会”这个因素，把社会属性作为不同机器人交流的接口，使研制出来的智能机器人像人一样具备社会性。这样一来，即使每个机器人的智能水平有限，但他们可以通过“与生俱来”的社会属性根据预设目标组成一个“社会”，每个机器人都是社会中的组分事物，通过协作，使搭建的“社会”系统迸发出人类期待的统性事物，通过团队协作创造性的开展工作，更出色的完成人类赋予的任务。

第五章 人工智能技术开发的不确定性

技术开发(Technical Development)通常指把研究所得的发现或一般科学知识运用到产品或工艺上的活动,改进现有产品或者创造新的产品、生产过程或服务方式的技术活动,技术开发和技术创新没有本质的区别,重大的技术创新将给社会带来根本性的影响。1912年熊彼特(J.A.Schumpeter)在《经济发展理论》中认为技术创新包括了5个方面的内容,分别是新产品的引进、新技术的引进、新市场的开辟、新原料来源的掌握和新的工业产生方式。70年代后半期,美国国家科学基金会(National Science Foundation of U.S.A.)对技术创新定义是:“技术创新是将新的或改进的产品、过程或服务引入市场。”从这个定义可以清楚地看到,技术创新的概念拓宽了,从某种程度上看,技术转移等改进手段也融入到了技术创新的概念中。

由于人工智能是一门新兴的学科,所以在人工智能的研究中伴随着不断的技术开发(技术创新),技术开发过程不仅包括了新技术的开发,还包括了军转民、民转军的技术转移过程。在尝试阶段,很多创新成果未必能真正投入使用,即使能投入使用,也可能不能达到预期的经济或军事效应,使整个技术开发过程充满了不确定性,而在技术开发中,不确定性更多指的是风险。

5.1 技术开发中的风险和不确定性

国际著名的创新研究学者罗森堡(N.Rosenberg)作为研究技术创新中的不确定性的先驱之一指出:“不确定性是讨论技术创新与创新文化的一个起支配作用的关键词。”不确定性是技术开发或创新的重要特征,但在讨论技术创新中的不确定性到底是什么的问题时,主要有三类代表性的观点,第一种观点认为不确定性等同于风险,如果硬要加以区分,按照奈特(F.H.Knight)的观点,“风险”是可保的风险,即能根据事实的客观分类有能力计算出概率的情形,而“不确定性”是不可保的风险,无法完成客观的分类。也就是说同样是不确定性的东西,“不确定性”本身不仅仅包括了概率事件,还包括了灰色信息、不完备信息等内容。第二种观点认为不确定性是一种约束条件,是系统的外生变量,按照这种理解,在技术开发中,不确定性主要取决于环境等外在内容,而与技术开发的内在机制无关。第三种观点认为不确定性代表了某种不可能性,认为技术创新是一种偶然的、因为信息不全面的不可预知的过程^[55]。

在技术创新中,如果仅仅把不确定性看作外生变量是不全面的,因为不确定性不仅仅包括了外部环境的影响,还包括了内在运行机理本身的“黑盒”状态和

“矛盾”状态，“黑盒”状态让人无法确切的了解和掌握运行机理，从而导致了掌握信息的不完全性；“矛盾”状态是内在的推动力，推动系统向着否定之否的方向发展。其次，不确定性贯穿了技术创新的整个过程，而不只是因为掌握信息的不全面，从而导致了结果的偶然性和不可预知性。最后，没有必要将风险和不确定性强行分割，因为两者之间呈现的是“你中有我，我中有你”的模糊的状态，强行分割为可保风险和不可保风险并不能加强理解，并且不确定性归根到底最后还是作为风险来影响技术开发或创新的，因此在讨论技术开发的时候，本文将不确定性和风险等同起来考虑。

辩证的来看，技术创新中的不确定性（风险）并不仅仅是一种消极的因素，制约技术创新成功的可能性，导致技术创新可能不能取得预期的成果和达到可观的经济收入。不确定性（风险）从另一方面来看是成功的必经之路，促使新技术替代旧技术，新方法补充旧方法，是效应和利润生成的基础。没有经历风险的考验只可能原地踏步，保守的重复相同的劳动不可能得到丰厚的回报，规避创新注定会遭到淘汰。

人工智能技术创新中的风险主要体现在内在运行机理的不确定性导致开发困难和外界环境的不确定性导致开发困难两个方面的内容。其中外界导致的不确定性可以大致分为三个阶段进行考虑，分别是研究决策阶段、产品研发阶段和应用推广阶段，由于不同阶段的特点不同，所以导致各阶段风险的种类和程度也有所不同。

李建明博士汲取多家之长，首次提出了“军事技术创新风险”的概念，并着重分析了技术风险，认为在军事技术开发中，技术风险主要可以分为性能风险、费用风险和进度风险三类^[56]。其实，不仅仅是军事技术开发的风险包括以上三类，民用技术开发的风险同样也包括以上三类，只是不同风险的影响程度及其带来的收益或损失有所区别而已。

对于人工智能来说，其研究成果既适用于军用，也适用于民用，因此，在讨论其技术开发的风险，本来可以类似地从性能风险、费用风险和进度风险三个方面来衡量。但是本文在研究中经过认真考虑，认为风险更侧重于外生变量的考虑，然而不确定性不仅仅包括了外生变量还包括了内在运行机理。因此在考虑人工智能技术开发的不确定性时，重点从内在运行机理的不确定性和外生变量的不确定性两个方面进行讨论，这样的分类更适合于人工智能研究的特点。

5.2 技术开发过程何以开展

人工智能技术开发和其他产品和新技术开发类似，将遇到各类问题，例如：内在运行机理的不确定性和外生变量的不确定性将导致技术开发困难。

5.2.1 内在运行机理的不确定性导致技术开发困难

在程序员开发程序的过程中，当上千行代码编写完成后，点击“运行”发现程序报错，是一件非常头疼的事情，因为要想程序正常运行，必须查找出准确的错误点，因此很多情况下，代码的调试所花的时间比代码的编写还要多。不过当前 C++，JAVA 等都自带了“跟踪调试”的工具，可以帮助程序员迅速锁定错误发生的位置，并可以看到其内在的运行情况，使调试的不确定状态转化为确定的状态，大大加快了程序调试的进程。试想，如果无法看到内在的运行情况，全靠直觉感知或数万次的尝试在“黑盒”中找寻，将耗费极大的人力物力。

以“鼠脑”机器人为例，鼠脑机器人在没有人为干预和程序预先设置的情况下，可以通过培养皿中未失活的鼠脑神经元来控制机器完成一系列的动作，体现了虽然低级但可以成为“智能”的行为。通过研究发现，鼠脑神经元和人脑神经元在结构和功能上相差无几，并且都是通过神经递质传递信息来完成沟通、互联、自组织，所不同的是神经元在数量上却有数量级的差别，因此，可以期待，是否当培养皿中的神经元数量达到人脑数量级时，被鼠脑控制的机器人将体现出进一步的智能来。这个假设需要不断的尝试，因为虽然鼠脑机器人是人类开发的人工智能产品，人们也看到了鼠脑机器人表现出来的“智能”，但没有人能彻底了解清楚鼠脑机器人的内在运行机理——鼠脑神经元是如何在培养皿中识别和感知；在自主完成一系列基础动作的时候，神经元之间是如何进行交流的；神经元之间是如何分工的；传递了什么样的信息；为什么神经元间不能进行进一步的交流；人类如何在鼠脑机器人进行行动时提前获知它下一步要做什么。也就是说，鼠脑机器人表现出来的行为是能见的，确定可知的，但其内在的运行机理却呈现出了一个“黑盒”或“灰盒”的状态，对于技术开发者来说是一个不可知或者不全面的信息，无法知道该项技术开发周期多长，开发难度多大，从哪方面寻求突破口容易，选用什么方法节省成本，这些不确定性将大大增加开发的难度和风险。

此外，由于内在运行机理难以被研究者探索清楚，因此在最终产品没有亮相之前，研究者无法对其阶段性的性能进行准确地评估，导致在研究过程中，由此引起的性能风险将很难避免，成为了技术开发中最难控制、最主要的风险。性能风险指的是对技术开发期待实现的性能目标造成不利影响的不确定性，通常可以从两个角度来考虑。如果从单项技术研究的角度来看，性能风险主要指所开发的技术是否达到了预期的超前性；如果站在整个系统的角度来看，性能风险主要指系统中单项或某些技术的超前性给整个系统带来的不确定性。一项技术在着手开发之前，即使其决策者的洞察力不够强、眼光不够长远，都应该是在该领域尚未出现的新技术，多少在性能等方面有一定的优势。但是，经常会发现一些项目在研制成功后，发现其通常在性能上不能达到预期较好的效果，或者即使达到了，

但该技术却不再领先。很多时候，人们把这类风险归于过程控制的风险，但如果追根溯源，这还是由内在运行机理的不确定性导致的性能风险。因为，研究开始后，人们能够尽可能控制的更多的是外生变量，而对内在运行机理的控制是比较困难的。其次，在一个系统中，每项技术都是该系统的内生变量，需要把握住项目之间的内在联系。因为，即使单项或某些新型技术的开发达到了预期的目标，也可能因为其超前性给整个系统带来性能风险。在技术开发的过程中，矛盾是无处不在的，特别是对于一个技术体系来说，部分技术所追求的性能可能是相互制约的，因此要把握什么技术是系统中最核心的技术。以潜艇这个系统为例，如果把增强打击力作为潜艇技术开发的重点，即让潜艇携带更多的火炮，这将导致潜艇变得极为笨重，下潜和上浮变得十分缓慢，机动力受到严重的影响，在交战中很可能因为机动力的丧失导致还未来得及发挥打击力的优势就被击沉到海底。

内在运行机理的不确定性是导致技术开发风险的根本原因，这是因为如果能将内在运行机理的“黑盒”变为“白盒”，使其确定性的为技术开发者所知，那么“摸索”的过程将简化为“对症下药”的过程，使得人们能够有针对性地做出正确的决策，采用适当的方法，提供充足的资金支持，少走弯路，将在很大程度上降低外生变量的不确定性。此外，在一个由多个项目组成的系统中，每个项目都是该系统的内生变量，因此要求决策者和研究者充分把握项目之间的内在联系和相互制约和促进的条件，从而在技术开发中有针对性地侧重或规避，努力达到整个系统的最优状态。

5.2.2 外生变量的不确定性导致技术开发困难

外生变量对技术开发的影响主要体现在研究决策阶段、产品研发阶段和应用推广阶段。人工智能作为新兴的研究方向，其技术开发或创新引起了各国的高度重视，不仅仅是因为在民用中潜在的经济效益和巨大的市场，更重要的是技术开发的产品在军事上的重大应用。民用技术的开发，更多是企业的行为，侧重于经济效益；而军用技术的开发，更多是集一国之力，通过充足的资金，集聚最优秀的人才为国家的战略需求做好保障。虽然技术开发从外在来看，存在不同的阶段，各阶段风险程度不同，但综合来看，风险的种类主要包括了技术水平、外界环境、组织管理、资金支持几个方面的内容。对于军用技术开发来说，由于是集一国之力（例如，两弹一星的开发），所以各方面和各阶段的风险都比民用技术开发相对少一些，所以本文主要对民用技术开发的各阶段风险进行分析。

在研究决策阶段，风险主要与政府扶持力度、宏观经济、消费者需求、行业景气度、竞争对手策略、行业技术水平、对市场信息的掌握程度、领导重视程度、领导者素质、创新资金的支持、企业实力相联系；在产品研发阶段，风险主要与

科研人员的实力、技术累计程度、技术先进程度、技术的难度与复杂度、技术的协作能力、技术成熟程度、项目组织管理能力、合作伙伴实力、研发项目进展控制能力、资金供应及时性、资金运营能力、开发者待遇、开发环境营造等方面相联系；在推广应用阶段，风险主要与新产品质量性能、新产品成本、新产品与市场兼容情况、企业生产规模、原材料和零部件供应难易程度、新产品定价、市场规模、消费者需求变动、竞争对手数量、竞争对手实力、企业信誉和知名度、企业营销能力、竞争者的不当竞争、新产品寿命周期、地方政策相联系^[57]。可以看到，风险在各阶段无处不在，因此导致了在民用技术创新的过程中，每年约有 4 成的项目因为各种因素夭折，而如期完成的项目中，能够盈利的项目还不到两成。

上述分类是非常细致的，其实在真正的技术开发中，不同的因素影响程度有一定的差异，并且有的因素例如：开发者待遇、新产品成本、科研人员的实力等是在一定程度上可控的，而另外一些因素，例如：竞争对手的数量、竞争对手的实力等等是控制相对比较困难的。但综合来考虑，技术开发时，外生变量影响最大的还是费用风险和进度风险。例如：如果资金充足，资金运行顺畅，可以给开发者较好的待遇，吸引高水准的研发团队，技术的先进程度将有所提高，技术开发的难度和复杂度将下降，促使项目能够如期完成，产品开发成功率将较大程度的增加；如果进程控制得当，将能够保证该项技术开发不拖沓时间，将在一定程度上节省资金，并让该项技术拥有更长的超前期。

此外，技术创新在广义上还包括了技术转移，即科技成果的转化，是指技术从一个地方以某种形式转移到另一个地方，包括了国家之间的技术转移，部门之间的转移，以及军用技术转民用技术和民用技术转军用技术等内容。

技术转移与商品转移不同，首先技术转移存在级差，是高势位向低势位的转移，即技术转移往往是拥有先进技术的国家、企业等向技术相对落后的国家、企业的转化过程；其次在技术转移的过程中存在着不完全渡让性，即使用权的转移，技术拥有者可以根据政治或者经济的利益将拥有的技术重复的转让给不同的国家或者企业、个人，但并不影响拥有者本身继续掌握和使用该项技术。在技术的转移中，通常有嫁接型和移植型两种基本的模式。嫁接型是指关键技术或工艺的部分转移，成功率较低，支付的成本也比较小；移植型是指对全部技术的转移过程，成功率高但相应的成本比较大。在技术转移过程中，风险仍然是存在的，例如：技术渡让方可能对关键技术仍然有所隐瞒，导致接受方仍然不能完全掌握该技术；接受方没有能熟练掌握该技术的人才，导致即使拥有了设备也无法发挥应有的效用。

5.3 技术开发何以评判

民用技术开发的最终目的是通过技术商品化推广，能够使企业获利，所以如果没有取得预期的收益，即使成功开发出了相应的技术，也可以认为是失败的。对于民用技术开发来说，失败影响到了一个企业的利益甚至生存，但对于军用技术开发来说，更侧重于国防建设，而不以盈利为主要目的，其结果对国家起着比较重大的影响，所以本节将以军用技术为例进行讨论。

在军用技术开发过程中，判断技术开发是否成功，不仅取决于是否按期研制出预期的成果，还与其他国家的技术开发情况和国际局势有着紧密的关系，这些因素都将影响着本国该项军用技术开发的结果评定。例如：在武器的研究过程中，一直存在“矛”和“盾”的关系，希望用我方研制的“矛”去进攻敌方研制的“盾”，用我方研制的“盾”去防御敌方研究的“矛”。如果按照决策我方按期研制出了“矛”，但发现敌方有了更先进的“盾”，我方所研制的“矛”只能应对我方“盾”的弱点，但对敌方的“盾”却无计可施，导致最终不能达到击败敌方的期望，这将导致看似成功研制的项目失败，失败原因在于对手在该项研制中的实力强于我方，导致我方新技术失去优势。再如，即使我方研制的“矛”能让敌方的“盾”防不胜防，但却四、五十年都没有战争，使得“矛”根本无法用于战场，等到四、五十年后需要交战时，曾经先进的武器已经变得颇为落后，无法进攻敌方最新研制的“盾”，导致武器装备的浪费，同样导致原本成功的研制项目失败，原因在于国际局势不适应于技术发挥效能。

由此可以看到，技术的成功与否，不仅仅取决于单方面的预期，还受着更多因素的影响。对于人工智能技术开发来说，最后的结果还将受限于更多的因素，例如：创新成果是否被大众认定为人工智能方面的成果？以图灵测试为例，图灵创新的提出了一种检验机器是否具有智能的标准，但却受到很多学者的质疑，认为该标准不能检验机器是否具有智能。因此，人工智能领域的技术创新，最终结果还取决于当时对“智能”产品定义的标准是什么，才能界定出创新成果是否属于人工智能方向的创新。其次，人工智能领域的技术创新最终能否真正付诸于应用开发还取决于是否符合道德和伦理的标准，否则很难被大众接受，导致成果无法应用推广，即使是很先进的成果，如果不能应用推广，也只能接受失败的命运。

第六章 人工智能应用前景的不确定性

伴随人工智能中的不确定性问题的逐步得到解决，将促进机器的智能得到快速的发展，更多的融入到人类的日常生活、军事国防等领域。随着人工智能研究的深入和应用的推广，也引发了一系列值得注意的问题，例如：计算机能否像人一样思维？未来的社会关系将发生怎样的变化？人工智能是否会如克隆技术一般引发一系列的伦理问题？

首先来看一下人们比较关注的问题，即：计算机能否像人一样思维？这是从人工智能诞生起就常问常新的话题。一方面人类期待计算机能够拥有非凡的思维能力，甚至能在替代人类体力劳动的同时分担人类的脑力劳动；另一方面，人类又担心有朝一日计算机拥有了类人或超过人的思维能力后会不甘屈居人下，导致人类丧失主体地位。但就当前基于传统体系结构的计算机来说，它依旧受制于可形式化、可计算以及非计算机难解问题这三个极限，要想在此基础上拥有包括创造力在内的思维很可能只是一个美好的想法。但是，这并非意味着人们只能对人工智能的发展抱以消极的态度。事实上，伴随着量子计算机、生物计算机取得的进展，随着逻辑基础、研究方法、技术开发的变更，让未来的计算机拥有真正的思维也未必不是一件可行的事情，但是其思维能力究竟能够达到什么水平，却很难预料。如果最终在这个问题上得到了突破，那么下面的问题将继续引发我们的思考。

6.1 人工智能发展的社会关系之辨

随着人工智能的产品付诸于实用，在与人相处的过程中，将引入一些新的社会关系，例如：智能机器与智能机器组成的社会关系以及智能机器与人组成的社会关系，但归根结底可以归纳为人与技术的关系。

6.1.1 机器与机器的社会关系

随着人工智能的发展，未来机器人的数量将大规模的增加，智能也有较大程度的提高。值得注意的是，人类着力于研究高智能的机器人（单体机器人），但正如不同的人有不同的知识结构，从事不同的劳动，有不同的创新能力，所研制的机器人智能再高，也只能在某些特定领域展露头脚，而不可能是一款在任何领域都能独当一面的全能型的机器人。并且，即使人类在将来有能力设计一款结构复杂的全能机器人，也会因为高昂的研究、维护经费导致无法普及，难以全面的发挥其效能。

为了解决类似问题，越来越多的研究者把目光投向群机器人的设计中。群机器人由大规模的自治的简单机器人组成，具有分布式系统特征，机器人之间通过蚁群算法、微粒群算法等群体优化算法通信，进行交互和协调，使群机器人系统迸发出群体智能，高效率完成复杂的任务^[58]。

群机器人的灵感来源于生物学，在自然界中，蚂蚁、蜜蜂等动物虽然低等，但是通过群居，组建了庞大的社会，通过分工协作，可以筑造坚固美观的蚁巢、蜂窝；还可以把比自身大数倍、数百倍的动物作为猎物，体现出了比个体智能强大得多的群体智能。类似的，群机器人系统根据任务的不同，由一定数量的同构自治机器人组成。单个的自治机器人结构简单，造价较低，并且能力有限，但都具有局部感知和通信的能力，可以按照既定的目标与环境和群内的机器人联系，通过算法完成自组织和自组装。并且，在危险的环境中，群机器人系统为了完成任务，可以利用数量的优势和功能上的冗余在不影响群机器人稳定性的情况下，通过牺牲个别自治机器人来达到目的。与单体机器人相比，群机器人的特点如表 6.1 所示。

表 6.1 单体机器人和群机器人的区别

类型	组成	个体能力	个体结构	造价	容错力	效率
单体机器人	单个个体	高	复杂	高	弱	低
群机器人	多个个体	低	简单	低	强	高

由此可以看到，群机器人与单体机器人相比具有高容错力、造价低廉、生产周期短、高效率等优势，在未来人工智能的发展中，群机器人很可能是发展的重点，将建立起机器与机器之间的社会关系，即在智能机器人的研究中，除了着力于开发单体机器的智能外，还可以将更多的精力用于研究如何提高机器人的群体智能上。在由机器组成的社会中，首先将由人类预先设置的目标进行启动，但是在启动之后，根据任务需要，智能的机器将完成自组织和自组装的过程，从而完成向社会所需角色的转变。

在机器组成的社会中，机器与机器的关系虽然不可能像人类社会关系那么复杂（事实上没有必要全盘复制人类社会关系），但也需要分工协作。在机器社会中，有负责决策的中心机器；有扮演类似于人的眼睛、鼻子、嘴巴、耳朵等感官的“信息嗅探”机器；有大量按指令进行操作的“劳动者”机器；还有一部分看似冗余的机器，平时可以处于闲置状态，也可以根据任务安排和情况需要扮演特定角色的“救场”机器。在危险的环境或紧急的情况下，在机器社会中可以做出“舍卒保帅”的决策，还可以对因系统故障、电路老化、接口短路导致不能胜任工作、失控的机器进行淘汰或作为完成任务过程中的“死士”、“探路者”，从

而尽可能保证从事关键任务的机器正常运行以及大部分机器的“全身而退”。

在平常工作中，一个群里的自治机器人还需要学习人类社会中相互帮助的品质，让机器同样懂得“关心他人”。对此，可以在群机器人系统中添加专门探测机器状态的“大夫”机器人，通过接口通信、扫描来判断所检测的机器是否存在线路老化、接口松动、电压不稳定等症状，并发出警戒或对其进行力所能及的检修。

机器与机器的社会关系是一个有意思的设想。如何实现机器搭建的社会，让机器在拥有人类社会优良品质的同时，防止机器产生嫉妒、埋怨等不良的情绪，还需要人类做进一步的探索，从而让人类心中期待的完美社会的愿望首先在机器社会中之得以实现。

6.1.2 机器与人的主体地位之争

人类制造智能机器首先考虑的是需求：人类之所以要制造智能机器，不是希望失去主导的地位，让机器来替代全部体力和智力活动；而是希望在掌握主动权的同时让机器协助、分担一部分体力和智力活动，把机器的优势和人类本身的优势相结合，形成以人为主，以机器为辅的良好系统。当前计算机利用了速度、存储、记忆、搜索、耐力等方面的优势，更多分担的是人类的体力劳动，而对于智力劳动的帮助较少，智能机器的研究有利于机器更好的与人理解、沟通，更有利于分担人类的智力劳动。

伴随着智能机器人的大规模出现，社会关系将呈现出多元化的趋势，除了传统的人与人的社会关系外，还存在机器与机器的机器社会以及机器与人组成的“人一机”社会。

1950年，著名科普作家、科幻小说家阿西莫夫(Isaac Asimov, 1920-1992)在《我，机器人》中提出了机器人三定律(Three Laws of Robotics)，并于1985年在《机器人与帝国》一书中，增加了第零法则补充完善，成为了人工智能、机械伦理学等领域专家普遍认可的重要理论：“第零法则：机器人不得伤害人类整体，或袖手旁观坐视人类整体受到伤害；第一法则：在不违背零法则下，机器人不得伤害人类，或袖手旁观坐视人类受到伤害；第二法则：在不违背第一法则下，机器人必须服从人类的命令；第三法则：在不违背第二法则下，机器人必须保护自己。”可以看出，三定律希望智能机器能在保证人类整体利益最大化的情况下保证单个人的利益，并随时准备牺牲自己来保全人类。

不过，三定律很可能是人类一厢情愿的想法。在三定律的限定下，机器人即使拥有了比人类更高的智能，也无法与人平起平坐，只能担任“奴仆”或“保镖”的角色，人和机器人的关系显示出了严重的不平等。

如果站在理想世界的角度来看，万物生而平等，任何一种生物都没有资格为其他生物制定法则、区分尊卑。但在真实的世界中，却一直存在着弱肉强食、智能驾驭力量的现象。作为已知智能最高的生物，人类不断地通过牺牲其他生物的生命或占有其栖息地来发展自我，制定了“人是万物之灵”的规则并将自己推上了最尊贵的主体地位。很难想象，如果有一天真的存在一批机器人机缘巧合的拥有了至高无上的智能，他们是否还会甘于屈从于人而不与人类争夺驾驭世界的权利。或许假想机器人的智能会超过人类是一件杞人忧天的事情，但可以通过转换观念进行设想：如果有朝一日，一群外星人降临地球，构成他们身体的不再是蛋白质、核酸、碳等物质，而是硅、铁、铜等元素，并且长得就像人类发明的计算机，如果他们拥有比人类更高的智能，那么他们会与人类怎样相处呢？

虽然目前还没有制造出智能程度高到需要用机器人三定律来规范行为的机器人，但随着科技的进步，人类将怎样定位与智能机器的关系以及如何保持人和智能机器的关系已成为广泛讨论的问题，融入相关元素的电影《黑客帝国》、《人工智能》、《我，机器人》、《夺命手机》、《源代码》等都从不同角度对此进行了精彩的诠释，反映了不同时期不同的人对此的关注和给出的回答。

虽然人类希望建立以机器为辅，人机结合的社会，但由于在决策确定的情况下，执行结果仍然存在着不确定性，因此，在真正的建立过程中，机器将在生产活动中占据多大的比例，替代人类多少任务的工作却难以界定。并且，更多人担心的是，当机器发展到一定水平后，是否还将甘于人类工具或者仆人的地位，智能的机器是否会通过自组织和自组装进行革命，来实现统治地位，这将引发一系列的伦理问题。

6.2 人工智能发展引发的伦理问题

6.2.1 智能机器是否会导致人类劳动力过剩

就业难是很多人面临的问题，高技术岗位的人才供不应求，但其他岗位，特别是技术难度小，偏体力的岗位出现了明显的供过于求的现象，导致大规模的人失业，造成了劳动力大量剩余，加重了社会的负担，对社会治安和稳定也造成了影响。

随着人工智能的发展，未来的智能机器不仅可以替代人类的大量体力劳动，还有能力替代一部分智力工作。与机器的速度、记忆、耐力、准确度等方面的优势相比，人类的优势更多的体现在创造力。如果一项工作对创造力和形象思维等方面要求较低，完全可以用机器来替代。机器可以日以继夜的高效率工作，并且有着人类无法超越的精确性，一台机器完全可以替代数十人甚至上百人的工作。

当前智能机器的发展程度还不够高，造价比较高昂，同时存在着一定的安全隐患，所以普及程度不是很高，但随着技术的进步这些问题都将得到解决。如果智能机器普遍的运用到了公司、工厂，凭着速度、精确度、耐力、记忆力等方面的优势，很有可能会替代工人的劳动，将导致用人单位大规模裁员，许多人无法找到工作，生存压力加大。

人类开发人工智能的目的是为了更好的服务于人类，劳动是人类重要的社会属性，但随着智能机器参与到人类的劳动中，很难控制好机器替代人类劳动的比例，有可能导致较大比例的人都不再劳动，造成劳动力的大量剩余，很可能对社会的发展带来极大的影响。

人类离开劳动是一件很可怕的事情，可能产生无法预测的后果。在人类从猿类进化的过程中，包括了手的解放、语言的产生、智能的发展以及社会的发展等多个方面的内容，但无一例外的源于劳动或为了有助于更好的劳动。长时间脱离必要的劳动，人类将怎么度过闲散的时间？是否会导致相互之间交流不畅、言语不连贯，甚至手脚乏力、大脑反映迟钝以及身体各部分机能退化？因此，在未来人工智能的应用中，如何在保证智能机器减轻人类劳动负担的同时规避劳动力过剩的问题，使需要细致规划和深思熟虑的。

6.2.2 智能机器人是否等同于人

英国量子物理学和量子计算领域专家戴维·多伊奇（David Deutsch）对人工智能的看法是：“人工智能是具有人类心智属性的计算机程序，它具有智能、意识、自由意志、情感等，但它运行在硬件上而不是人脑中。”如果机器真的拥有了智能，产生了意识、思维甚至情感，并且在机器之间也成功地搭建起了类似人类社会的社会关系，那么在对智能机器人的定位上将不能仅仅因为其不是自然的生命而作为普通机器来看待，例如：人类在打算报废智能机器的时候，它很可能会表达出害怕和不舍，并且通过思考，还可能向着扭转局面的方向做出努力，那么报废这样一个类人的有思维的机器人，是否会有“杀人”的罪恶感；同样的，对机器人进行买卖，或者把它们派遣到危险的环境中替代人完成随时可能“丧命”的工作时，是否也会有同样的感觉呢？

要解决这个问题，首先要判定出一个拥有社会属性的智能机器人是否等同于人，即：“思维和社会关系的存在是否等同于人的存在。”在这里，本文特别强调了“拥有社会属性”这几个字，因为之前已谈到智能机器人是否等同于人的问题时，总会有学者反驳道“人具有社会性，智能机器人再聪明，不具备社会属性当然不能与人等同。”事实上，通过接口互联、传感器通信等方法可以实现机器之间的相互交流，并且随着相关技术的提高，组建机器与机器的社会以及形成人

和机器的多元化关系也是可能实现的，即让机器拥有社会属性是可能的。在这一前提下，关键问题简化为“有思维的机器是否等同于人”的问题，这又涉及到了“人”和“生命”的标准将如何界定。

毋庸置疑，在给出“人”的定义时，首先想到的是“人”是自然存在物，是有生命的，而不是被其他生物制造出来的机器。所以，即使一台机器逼真到与人长得一模一样，也不可能与人等同。由此可见，对“人”的定义并不仅仅是外表上的判断。那么，如果一台机器具有与人相当的智能，能够进行思考，那么能与人类等同吗？可以想象大多数人的回答仍然是“不能”，但仔细思考会发现医学上仍然还使用着“脑死亡”的生命判定标准^[59]，即思维活动的停止标志着生命的终结。那么对于一台机器来说，如果它能像人一样思维，看上去不也是有生命的吗？不能因为智能机器人不是自然存在物或外表与人不同而加以歧视。实际上，包含“自然存在物”、“生命”在内的任何定义都是人类在认识的不同阶段给出的，是主观上的判断，而宇宙从来没有对生命和非生命作过确定性的定义。随着人类观念的更新和研究的深入，人对事物的定义也不断进行着补充和更替，是一个辩证的过程。所以，很可能在“人”或“生命”的定义范围更加广泛时，能够思维并具备社会属性的机器人可以等同于人，这是因为智能的产生并不局限于一定要产生于人脑这个物质基础，而是可以扩展到任意的“硬件”结构。

虽然当前智能机器只能划分到人工生命的范畴中，但是人工生命的价值是否值得人类去珍重；一个能思维的，具有精神的机器人到底应该看作是人还是机器仍是值得进一步商榷的问题。

6.2.3 智能机器人与人交战是否会引入新的战争模式

1997年，美国国防部启动了 Demo III 计划研制战术移动的半自主机器人来完成侦查、布雷扫雷、路况维护等任务^[60]；2000年美国国防部高级研究项目局(DARPA)出资 5000 万美元研究开发了增强人体机能的可穿戴式智能外骨骼系统来提高战士的行军和作战能力^[61]。此外，由于无人机具有价格低廉、不惧伤亡、隐蔽性好、生命力强、起降简单、操作灵活等特点，美国、俄罗斯、以色列、英国等近 30 个国家还紧锣密鼓的展开了无人机的研究，通过研制靶机、侦察无人机、诱饵无人机、电子对抗无人机、攻击无人机、战斗无人机等机型来完成战斗任务^[62]。

人工智能之所以引起了各国专家的关注，不仅仅因为经济利益的丰厚和能更好的服务于人类，其在军事领域的应用也是极重要的一方面。人工智能的产品将大规模的应用到未来的战场上，除了在武器中运用部分智能引导、追踪等方面的装置外，智能机器人很可能会直接加入战斗，保护己方战士和装备，消灭对方的有生力量，摧毁对方的关键设施。由于现阶段对机器人的定义更多偏重于机器而

不是人，因此，在未来的战争中，如果双方都有机器人参战，将形成战斗力对称的场面；但如果参战双方只有一方具备智能机器人参战的能力，那另一方很可能因为遭受严重的损失，无数生命可能会葬送在智能机器人的手里，交战双方形成人和机器战斗——生命和非生命抗争的不对称局面。

一支训练有素的智能机器人部队即使完全脱离人的指挥，也有实力进行物理战或心理战，完成以一敌百、以一敌千的作战任务。首先在硬件结构方面，高科技材料打造的铜头铁臂自然比人类的血肉之躯坚固上万倍，并且机器人没有疼痛感、疲惫感、饥饿感和畏惧感，对生化武器的防御力较高，可以胜任恶劣、危险条件下的全天候作战。因此，在打击力、防护力、机动力等军事技术基本要素方面，人类是无法与智能机器人相比的。其次，智能机器人有着与生俱来的搜索、计算、记忆、存储能力，特别在人类赋予其智能后，更是如虎添翼，在信息力要素方面优势明显，可以通过灵活的制定策略、精确打击，让对手的作战网络瘫痪；发布虚假消息、篡改作战指令，让对手军心涣散。不难想象，大多数人类士兵在与一群掌握高尖端技术机器交战时，比与一支精良的人类部队交战更不寒而栗。

人与机器交战的不对称局面有可能会成为未来战争的新模式，这将颠覆数千年来战争的模式。从有战争起，无论是在冷兵器时代、火器时代还是核武器时代；无论是物理战、信息战还是心理战，都有一个确定性的规律可循，即：对抗的双方都是人。从原始社会的肉搏或采用简单加工的石头作战到当前高精尖的核武器、生化武器的运用，都是人直接参与作战或者人借助工具（千变万化的武器）参与作战。如果真的有一天，某国派遣了一支纯粹由智能机器人组成的机器人部队参与作战，而其对手是由人类组成的部队，这将导致对抗双方由“人-人”变为了“人-机器”，由于这些智能的机器是人造的产品，那么人与机器交战相当于人在与劳动制造的工具交战，这将是一个非常尴尬的场景。

由此可见，人工智能在多学科多领域的专家的努力下，即使解决了逻辑基础、研究方法、技术开发等方面的难题，由于人工智能研究的特殊性，其应用开发还将面临伦理和道德上的考验，仍然不能保证能够一帆风顺的付诸于实用。

结 束 语

人们通过形而上学的研究方法对真实世界进行了近似的确定性的刻画，但随着研究的逐步深入，逐渐认识到了真实世界的本质是不确定性的，确定性不过是不确定性的特例。不确定性在宏观上主要体现为时间和空间上的连续性，在人工智能研究中具有重要的地位，将有效弥补传统人工智能研究中的不足。

论文首次从逻辑基础、研究方法、技术开发、应用前景四个方面对人工智能中的不确定性问题进行了较为深入的研究。在人工智能逻辑基础方面，作为思维和智能的 DNA，逻辑基础起到了不容忽视的作用，贯穿于研究的每个细节中，要想让计算机真正拥有智能，选取适宜的逻辑基础是一个重要的方面。但是具有二值性、排中律、恒常性特征的刚性逻辑并不能全部反映人类思维中非形式化和非数值的部分，导致基于形式逻辑的二进制计算机系统不可能真正具备智能。因此，包容不确定性的柔性逻辑、泛逻辑由于包容了不精确推理、不完全推理、动态更新等特征，将成为未来人工智能发展的重要逻辑基础。

在人工智能研究方法方面，不同学者站在不同的视角，采用了不同的研究方法，其中影响力最大的符号主义、联接主义和行为主义的方法都曾经在特定的领域取得过辉煌的成就，但都因为不确定性缺失的问题而陷入瓶颈状态。虽然目前暂无一种统揽全局的研究方法，但在理论方面，从智能内在生成机制出发的“机制主义”方法较好的将三大主流方法统一在了同一个理论框架下。笔者认为机制主义的方法包容了不确定性，并把不确定性和机制主义结合起来探讨，丰富和扩展了机制主义的内容。为了更好的对人工智能中的不确定性问题进行研究，本文尝试提出了“硬不确定性人工智能”和“软不确定性人工智能”的分类，认为面向物质的硬不确定性人工智能的研究方法是人工智能发展的突破口，可以通过在硬件搭建中引入不确定态来实现；此外，面向软件的模拟的软不确定性人工智能的研究方法是人工智能发展的重要保证。当前在软不确定性人工智能研究中，虽然出现了一系列新兴的方法，但尚未考虑到对人的社会属性进行模拟，因而在后续研究中，如何通过软件模拟引入社会属性，让机器更逼近人类智能是努力的方向。

在人工智能技术开发中，内在运行机理的不确定性使得技术开发类似于一个“黑盒”状态，研究者很难查找错误和控制产品的性能；外生变量的不确定性主要指技术开发的风险，其中影响最大的是费用风险和进度风险。对于人工智能来说，其技术即使能按预期开发出来，还面临着对该项技术的评判，如果不符合人类道德和伦理需求将不能付诸于实用，导致技术即使成功开发也将面临失败。

在人工智能应用前景方面，随着智能机器的大量出现，机器能否真正拥有思

维成为人类关注的问题。一旦机器拥有了真正意义上的智能，社会关系可能不再简单的局限于人与人组成的人类社会关系，还可能会包括人与机器组成的“人-机”社会以及机器与机器组成的机器社会，但归根到底，是人与技术的关系。在机器智能水平较低的当前，毋庸置疑的是以人为主，机器为辅；人为主人，机器为奴仆的关系，但随着机器智能水平的逐步提高，智能机器人是否仍会甘心屈居人下，这将带来一系列的伦理问题。例如：智能机器人是否会替代大量人的劳动，导致人类劳动力过剩，造成社会资源的浪费，导致社会的不稳定。具备社会性的高智能机器人到底应该判定为“人”还是“机器”，如果智能机器人可以判定为“人”，那么买卖、伤害机器人是否符合道德标准；如果智能机器人仍然只能判定为“机器”，有朝一日纯粹的机器人部队与人类交战，将呈现出对弈双方分别是人和机器的新模式，这将彻底改变数千年来传统的对弈双方是人、人直接或者借助武器（冷兵器、火器、核武器）交战的模式，可能引起难以预测的后果。

综上所述，人工智能中的不确定性问题研究是一个包容性较强的前沿领域，它与计算机技术、哲学、军事技术等学科紧密联系，其产品也将迅速的投入社会应用，并在日常生活、经济、国防等方面起到重要的作用。虽然该领域的研究起步较晚，相关理论还缺乏大量试验和应用的支撑，有的还让停留在假想的基础上。但相信在多领域专家的合作努力下，只要彻底解决了人工智能中的不确定性问题，模仿人类意识并制造出真正具备智能的机器并非只是一个梦想。

致 谢

论文的完成离不开多方面对我的关心和帮助。首先衷心的感谢我的导师刘戟锋教授。

入学之初，我先闻师名，才见师面。早在师兄师姐的耳口相传中，我便知道了导师是“博学之儒，多才之士”，“胸襟开阔，卓有见识”。初次见面，导师送给我他撰写的《从物理战到心理战》，并建议我尽快补充科技哲学方面的基础知识，尤其要认真研读恩格斯的著作《自然辩证法》。导师对马恩的思想研究很深入，记得1997年出版了他对恩格斯军事技术思想研究的专著《哲人和将军》，然而在看到书名时，我首先想到的却是导师——身在和平年代，他是军事技术哲学领域公认的专家，是哲人；如果身在经受战火洗礼的岁月，慧眼独具、胸中自有十万兵的他必然是运筹帷幄、决胜千里的将军。导师的工作繁忙，经常受邀讲学和访问，但他却能坚持为全体研究生讲授军事技术和自然辩证法相关的课程。每次看到其他同学使用的教材都是我导师的著作时，我总觉得很自豪。

导师在传道、授业、解惑的过程中，总能做到倾囊相授，一直期盼着学生能够“冰出于水而寒于水，青出于蓝而胜于蓝”，总是告诫我们要抓住点滴积累灵感，尝试提出有建树的观点。师兄师姐虽然多次被其他老师夸赞为“名师出高徒”，但仍然感叹望尘莫及；而我的基础不如他们好，阅读的文献不如他们多，每当想到古人“负笈千里，从师之殷；立雪程门，敬师之至”的精神更是感到惭愧。

导师是慈师，也是严师。当我为论文选题担忧时，他细心的考虑到了我本科修习的是计算机专业，带着我选定了比较擅长的人工智能方向的题目；当我试图依赖他帮忙拟定论文的全部框架时，他坚持“授之以鱼不如授之以渔”，每次寥寥数语，迫使我独立思考，大量阅读资料；在我撰写论文遇到困难的时候，他有针对性的设想了一个“木卫二”的科幻故事，启发我对宇宙、时间、生命和智能的思考，让我从中受益。

感谢导师组曾华锋教授、朱亚宗教授、赵阳辉副教授、高嘉社副教授、何兆勇副教授、李佩琼副教授、刘艳琼讲师和盖立阁讲师对我的关心指导和严格要求。曾华锋教授平时工作繁忙，但总能为我们带来精心准备的授课并耐心的解答学生提出的问题，还带领其他老师尽力为我们营造更好的学习环境，通过读书笔记、读书报告会等形式增大我们的阅读量，培养浓厚的学术研讨氛围，还邀请了国外知名教授来校多次授课，拓宽了我们的视野。朱亚宗教授是我们专业德高望重的前辈，总用一颗无私的心一视同仁地对待每一名学生，为我们修改文章，与我们探讨新的观点，和我们分享前沿的知识。赵阳辉副教授美而雅致，特别是穿上旗袍后，更显出了一袭诗书香的大家闺秀的风范；她总给我一种静若井水的感觉，

但她对学生却很关心，每次见到我，她总会关心我最近的情况，论文撰写得怎么样，工作忙不忙等等，让我感到很温暖。高嘉社副教授为人随和，善于倾听我们的想法，能很好的调动课堂的气氛，在他的鼓励下，本来在课上不怎么喜欢交流的我也能大胆的说出自己不太成熟的见解。何兆勇副教授为我们直接授课虽然不多，但是经常参加我们的交流和研讨，为我们提出中肯的建议，关心和帮助我们，总给我们一种很亲切很豪爽的感觉。李佩琼副教授上课擅长采用不同的形式，课堂经常充满了笑声，他多次夸奖我的课件做得好看，讲课讲得不错，还给了我“A”，让我的信心和兴趣都有了显著的增加。刘艳琼讲师是从斯坦福大学深造回国的高材生，文理皆通、基础厚实、眼光独到，经常在课堂上为我们带来耳目一新的知识，并在开题、论文进展报告会上为我们提出中肯的建议。盖立阁讲师是和我们接触最多的老师之一，他总是一丝不苟地督促我们记好读书笔记、认真准备读书报告会、按期做好开题、预答辩等工作；他喜欢读书，喜欢和我们分享他的心得和经验，还经常把好书慷慨的分送给我们阅读。不熟悉他的每个学生都会抱怨他不通情理、过于严厉；但熟悉他的每个学生都会发自肺腑的评价：盖老师不可说不严，但对学生的关心不可说不多，对我们的帮助不可说不大。

感谢黄立波、王璐、屈婷婷、魏万强、黄嘉、谢大勇、王姝、刘慧、陈波、林涵、陈海洋、刘佚丹、黄松平、李建民、刘昱东、张煌、徐振兴、王雯霞、李思、田野等同学在平日生活中对我的照顾，在读书报告会、开题报告会中为我提出的宝贵意见，帮助我进一步开展研究。

感谢参考文献中所有作者的工作，你们闪光的观点为我的论文写作注入了灵感，提供了详实的研究资料。

参考文献

- [1] 刘戟锋.军事技术论.兵器工业出版社[M],1991.
- [2] 刘戟锋、曾华锋等.从物理战到心理战.吉林科学技术出版社[M],2007.
- [3] 冯天瑾.智能学简史[M].科学出版社,2007.
- [4] Von Neumann. The Computer and the Brain [J]. Yale University Press, New Haven, 1958(5).
- [5] 李德毅等.不确定性人工智能[M].国防工业出版社,2005.
- [6] 中国科学技术协会.2009-2010 智能科学与技术学科发展报告[M].中国科学技术出版社,2010.
- [7] 刘清,黄兆华. G-逻辑及其归结推理[J].计算机学报, 2004, 27(7): 865-872.
- [8] 马义德,李廉,王亚馥等.脉冲耦合神经网络原理及其应用[M].北京:科学出版社,2006.
- [9] Wang G Y, Shi H B. TMLNN:Triple-valued or Multiple-valued Logical Neural Network. IEEE Transactions on Neural Networks, 1998 (6): 1099-1117
- [10] 史忠植.神经网络[M].北京:高等教育出版社,2009.
- [11] 杨春燕,蔡文.可拓工程[M].北京:科学出版社, 2007.
- [12] 王守觉.多维空间仿生信息学入门[M].北京:国防工业出版社, 2008.
- [13] 钟义信.机制主义: 人工智能的统一理论[J].电子学报.2006(2):317-321
- [14] 约翰·卡斯蒂著,胡运发等译.剑桥五重奏——机器能思考吗[M].上海:上海科学技术出版社, 2006.
- [15] 雨果·德·加里斯.智能简史——谁会代替人类成为主导物种[M].北京:清华大学出版社,2007.
- [16] 伊利亚·普里戈金著.湛敏译.确定性的终结:时间、混沌与新自然法则[M].上海:上海科技教育出版社, 2009.
- [17] 伊利亚·普里戈金著.从混沌到演化[M].北京:北京大学出版社,2007.
- [18] 金观涛.系统的哲学[M].北京:新星出版社, 2005.
- [19] 李坚.不确定性问题初探.博士学位论文.社会科学院科学技术哲学, 2006.
- [20] 谢千河,王鹏.不确定性的哲学发展历程及分析[J].内江师范学院学报.2009(24):59-60
- [21] 张智明.人工智能技术及其哲学思考[J].机电产品开发与创新,20(5): 2007.
- [22] 赵泽林.人工智能基础哲学问题探究.学位论文.华中师范大学马克思主义哲学.
- [23] 龚启荣,林邦瑾.宇宙智能、人类智能、人工智能.贵州大学学报(自然科学

- 版) [J].2002(2):59-64
- [24] 吴鹤龄,崔林编.ACM 图灵奖:1966-2006(第三版)计算机发展史的缩影[M].高等教育出版社,2008.
- [25] 蔡自兴,徐光佑.人工智能及其应用(第三版)[M].清华大学出版社,2003.
- [26] 熊才权.人工智能研究方法途径[J].舰船电子工程.2005(3):14-17
- [27] Veelenturf L.Analysis and Application of Artificial Neural Networks [M]. Prentice Hall,1995.
- [28] D. O. Hebb.The organization of behavior[M].Wiley, New York,1949.
- [29] John McCarthy.Formalizing Commonsense:Papers by John McCarthy [M]. Ablex Publishing Corporation,1990.
- [30] Searle John.Minds,Brains and Programs[J].Behavioral and Brain Sciences,1980.
- [31] Pylyshyn Zenon.Computation and Cognition:Toward a Foundation for Cognitive Science[M]. Cambridge, Mass: Bradford Books/MIT Press,1984.
- [32] 杜建华.“人=机器人”论题辨析.学位论文.南开大学科学技术哲学,2005
- [33] 托马斯·库恩.科学革命的结构[M].金吾伦、胡新和译.北京大学出版社,2003.
- [34] Dreyfus Hubert.What Computers Still Can't Do[M].Cambridge,Mass:MIT Press,1992.
- [35] 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,2007:7-17
- [36] 蔡曙山.哲学家如何理解人工智能——塞尔的“中文屋子”及其意义.自然辩证法研究.2001(11):18-23
- [37] Drew Mcdermott. Mind and Mechanism[M]. The MIT Press, 2001.
- [38] 刘融.脑科学研究的哲学问题初探.学位论文.华中科技大学科学技术哲学,2009.
- [39] 黄志鑫.“图灵测试”的历史回顾与思考.学位论文.浙江大学科学技术史,2009.
- [40] 戴维·多伊奇.真实世界的脉络[M].广西师范大学出版社,2002.
- [41] 吴九龙等编.孙子校释[M].军事科学出版社,1991.
- [42] 王光明.仿鱼柔性长鳍波动推进理论与实验研究.学位论文.国防科大控制科学与工程,2011.
- [43] 何华灿,何智涛.对智能科学逻辑基础研究的战略思考.2005年第11届中国人工智能何华灿等.辩证逻辑的数学化趋势.河池学院学报.2007(2):6-10
- [44] 学术年会:42-52
- [45] John R.Searle, Is the Brain's Mind a Computer Program[J], Scientific American,

1990: 26-31

- [46] Attoh-Okine, N.O., Orji, C. Emerging artificial intelligence methodologies in uncertainty analysis and modeling [J]. Southcon/95. Conference Record, 1995: 409-415
- [47] 武金, 刘冬冬. 非单调逻辑: 困境与出路[M]. 毕节学院学报. 2007(5):44-47
- [48] Ian Ravenscroft. Philosophy of Mind[M]. Oxford University Press, 2005.
- [49] 张保权. 辩证看待物质和意识的相互关系[J]. 兰州学刊. 2010(11): 41-44
- [50] 朱保信. 从实践与物质的关系重新审视马克思主义哲学本体论[J]. 吉首大学学报(社会科学版). 2003(3):27-30
- [51] 温勇增. 系统意识观——马克思主义哲学意识观的继承和新发展[J]. 系统科学学报. 2009(17):100-106
- [52] 邬焜. 从信息尺度看人类社会的本质与进化[J]. 社会科学研究. 2005(2):51-54
- [53] 恩格斯. 自然辩证法[M]. 北京:人民出版社, 1984.
- [54] 吴波. 马克思的劳动观与人的自由全面发展. 学位论文. 河南大学马克思主义哲学, 2009.
- [55] 吴永忠. 论技术创新的不确定性. 自然辩证法研究. 2002(6):37-40
- [56] 李建明. 军事技术创新风险论. 学位论文. 国防科学技术大学科学技术哲学, 2011.
- [57] 游达明, 范钧. 技术创新多阶段风险分析. 企业技术开发. 2008(12):34-36
- [58] 薛颂东, 曾建潮. 群机器人研究综述. 模式识别与人工智能[J]. 2008(4):178-186
- [59] 龚园. 人工智能的哲学思考. 湖北经济学院学报(人文社会科学版)[J]. 2010(3):16-17
- [60] 辛贾. 美国陆军的未来战术移动机器人计划. 机器人技术与应用[J]. 2001(4):33-35
- [61] 张蕊等. 可穿戴式柔性外骨骼人机智能系统可靠性及应用伦理问题研究[J]. 机电产品开发与创新. 2008(9):19-21
- [62] 戎皓, 谈国军. 军用无人机的发展现状及趋势. 现代商贸工艺[J]. 2008(4): 111-112

作者在学期间取得的学术成果

- [1] 张昕.不确定性与人工智能.社科纵横（理论版）.已录用
- [2] 张昕.不确定性人工智能研究方法及其局限性.学理论.2011(12): 59-60
- [3] Xin Zhang, et al. Strengthen Study Motivations of College Students via Learning Environment Improvement. Advances in Intelligent and Soft Computing. 2011, Volume 140. 353-358（EI&ISTP 检索）
- [4] Xin Zhang, et al. Analysis on the Food Safety Management of Student Canteen and Relevant Solutions. International Conference on Agricultural and Natural Resources Engineering. 2011:136-139（ISTP 检索）
- [5] 张昕.全国散文作家论坛征文大赛二等奖.中国散文学会.2011.（2011 中国散文作家精品集收录）
- [6] 张昕.从不确定性视角看人工智能的研究方法.徐州教育学院学报.审稿中.
- [7] 张昕.不确定性人工智能研究的思考.科学技术哲学研究.审稿中.