Artificial Intelligence: A Modern Approach, Fourth Edition

人工智能：现代方法（第 4 版）

[美] 斯图尔特 • 罗素（ Stuart Russell ） 著
彼得 • 诺维格（ Peter Norvig ）
张博雅　陈坤　田超　顾卓尔　吴凡　赵申剑　译
张志华　审校

人 民 邮 电 出 版 社
北　京

内 容 提 要

本书全面、深入地探讨了人工智能（AI）领域的理论和实践，以统一的风格将当今流行的人工智能思想和术语融合到引起广泛关注的应用中，真正做到理论和实践相结合。全书分 7 个部分，共 28 章，理论部分介绍了人工智能研究的主要理论和方法并追溯了两千多年前的相关思想，内容主要包括逻辑、概率和连续数学，感知、推理、学习和行动，公平、信任、社会公益和安全；实践部分完美地践行了“现代”理念，实际应用选择当下热度较高的微电子设备、机器人行星探测器、拥有几十亿用户的在线服务、AlphaZero、人形机器人、自动驾驶、人工智能辅助医疗等。

本书适合作为高等院校人工智能相关专业本科生和研究生的教材，也可以作为相关领域专业人员的参考书。

◆ 著　　[美] 斯图尔特·罗素（Stuart Russell）
　　[美] 彼得·诺维格（Peter Norvig）
　译　　张博雅　陈坤　田超　顾卓尔　吴凡　赵申剑
　审　　校　张志华
　责任编辑　杨海玲
　责任印制　王　郁　胡　南

◆ 人民邮电出版社出版发行　　北京市丰台区成寿寺路 11 号
　邮编　100164　　电子邮件　315@ptpress.com.cn
　网址　http://www.ptpress.com.cn
　印刷

◆ 开本：787×1092　1/16
　印张：　　2022 年　月第 1 版
　字数：　千字　　2022 年　月河北第 1 次印刷
著作权合同登记号　图字 01-2020-6510 号
审图号：GS 京（2022）1043 号

定价：00.00 元

读者服务热线：(010)81055410　印装质量热线：(010)81055316
反盗版热线：(010)81055315

版权声明

献给洛伊、戈登、露西、乔治和艾萨克。——S.J.R.

献给克丽丝、伊莎贝拉和朱丽叶。——P.N.

对本书的赞誉

本书是在世界范围内广受欢迎的人工智能教材之一，其最鲜明的特点有二：第一，一卷在手，人工智能方方面面的主要知识被系统性地一网打尽，经典知识内容与前沿知识内容取舍剪裁别具匠心，深具章法，驾轻就熟，相得益彰，颇有一种“包藏人工智能宇宙之机，吞吐人工智能天地之志”的架势；第二，文字阐述深入浅出，言简意赅，旁征博引，详略得当，同时适合初学者以及在人工智能领域已有一定经验和造诣者这两大类人群阅读和学习，各取所需，甘之若饴。本书冠以人工智能的“现代方法”，可谓实至名归。特别是在以勃兴于 2011 年的深度学习模型为基本表征的人工智能走到山重水复、柳暗花明的当前形势下，其中的“现代”二字就显得更为重要。人们正呼唤着下一代人工智能新境界的到来。2022 年 10 月，包括两位图灵奖得主约书亚 • 本吉奥（Yoshua Bengio）和杨立昆（Yann LeCun）在内的一批学者撰文提出了“具身图灵测试”的概念，强调机器系统与世界环境的具身交互研究是开拓下一代人工智能创新方法的关键要义，而这一点与本书从“智能体”的视角总揽全篇的前瞻性思路不谋而合。显然，深度学习远不能包打天下，复杂开放环境下智能任务的解决离不开多种理论、方法和技术手段在“具身智能”条件下的兼包并蓄、融会贯通，这就对人工智能研究者在知识结构体系上提出了更高的要求，而本书恰好具备能够满足这种要求缺一不可的广博性和深刻性。

孙茂松

欧洲科学院外籍院士

清华大学计算机科学与技术系教授

清华大学人工智能研究院常务副院长

本书可谓是一流学者撰写一流教材的典范，作者是国际人工智能领域知名专家、ACM/AAAI 会士、曾获 IJCAI 卓越研究奖、AAAI 费根鲍姆奖、AAAI/EAAI 杰出教育家奖、ACM 杰出教育家奖等荣誉，自 1995 年第 1 版出版以来已被全球大约 1500 所大学用作人工智能入门教科书。人民邮电出版社隆重推出第 4 版中文版，无疑是中文读者的福音。

周志华

ACM/AAAI 会士

欧洲科学院外籍院士

南京大学计算机系主任兼人工智能学院院长

《人工智能：现代方法》是一本经典教材。“现代方法”选择从当下的角度讲述人工智能的故事，而贯穿全书的核心方法论是“智能体”。以计算机为载体的人工智能，揭开了机器智能大幕的一角，制造更复杂的机器，实现更强大的智能，机器智能将为科学研究创造无穷无尽的新对象。在这个意义上，智能是“科学的无尽疆域”，而人工智能这个“现代方法”，正是开疆拓土的动力之源。方法不止，智能无疆，“人工智能：现代方法”这个书名可以永远延续下去。

黄铁军

北京智源人工智能研究院院长

北京大学计算机学院教授

这是一本教材，但不是传统意义上的教材，它用现代思想凸显人工智能及相关工作的发展脉络，用智能体贯穿全书知识点的诠释，各章内容自然衔接，易于理解与掌握。

这不仅是一本教材，还是一本“大”百科全书，它全方位探讨了人工智能领域的方方面面，涵盖了从基础知识、模型方法、工具技术、社会伦理到应用专题等各个层面，是一本人工智能的高级工具书。

这是一本面向人工智能大领域的“大”书，作者也是大学者，连译者都是大学者领衔的团队，堪称经典之作，非常值得初学者、从业者、教师及科研工作者等专业人员阅读。

俞勇

上海交通大学特聘教授

上海交通大学 ACM 班创始人

首批国家高层次人才特殊支持计划教学名师

人工智能领域的特点是知识点散、前置知识多、技术迭代快，因此写一本全面深入的人工智能教材是一件很难的事。本书是人工智能领域的经典教材，全方位描述了人工智能的主要分支和技术方向，提供了人工智能领域的全景图。经过 20 多年的不断优化，目前本书已经是第 4 版。和第 3 版相比，本书增加了深度学习、人工智能伦理等近年来的热点研究内容，非常适合对人工智能技术感兴趣的读者阅读。强烈推荐！

邱锡鹏

复旦大学计算机学院教授

这是人工智能领域世界范围内最经典、最全面、最具影响力的教材，覆盖了人工智能领域所有重要子领域的核心问题、算法思想和现实应用。第 4 版加入了深度学习、多智能体系统、机器人、人工智能伦理等前沿领域的最新进展和挑战，更适合作为不同层次和领域的研究人员及学生的人工智能入门教材。

安波

新加坡南洋理工大学教授

本书是享誉世界的人工智能经典教材，我在读博期间就学习过其第 3 版，内容全面翔实，介绍深入浅出，既是初学者理想的入门教材，也是人工智能从业者的案头参考书。很高兴这本书的第 4 版被译介到国内，新版增加了 2010 年以来深度学习等最新前沿技术动态，新章节的贡献者有朱迪亚 • 珀尔（Judea Pearl）和伊恩 • 古德费洛（Ian Goodfellow）等知名学者。期待这本新版教材更好地推动我国人工智能的发展。

刘知远

清华大学计算机系长聘副教授

本书是继“西瓜书”和“花书”之后又一部大师之作。本书简称为 AIMA，其历史其实更为久远，几乎可以认为是国际上关于人工智能的标准教材。两位作者斯图尔特 • 罗素和彼得 • 诺维格是打通学术和产业“任督二脉”的大师。相比于“花书”的作者，本书的两位作者可以被归类为“传统派”，他们认为机器智能并不一定要学习生物智能，机器可以通过更擅长的计算、

更完美的数学模型以及大数据下的去模型化来实现智能。因此，在 2009 年出版第 3 版时，彼得 • 诺维格无法预测到近 10 年深度学习在诸多领域（特别是他擅长的自然语言处理、机器翻译）的快速进展，而作为强化学习的高手，斯图尔特 • 罗素也不会预想到深度强化学习在决策应用中的流光溢彩。第 4 版中融入了两位大师基于人工智能近 10 年最新进展的深度思考。同时，作者也是人工智能伦理和哲学的思想家，他们在最后一章中回答了人工智能未来走向和关乎人类命运的几乎所有问题。这是一本大部头的书，若能日拱一卒，势必功不唐捐。

吴甘沙

驭势科技董事长 /CEO

在炒作和质疑声中，人工智能技术不断倔强而真实地成长。可以预见并且逐渐成为现实的是，智能化革命就如同当年的信息化革命一样，会给很多领域和行业带来革命性的变化。人工智能也从一门选修型前沿学科，逐渐演化成一门必修型基础学科。要了解人工智能，一本全面的人工智能教材必不可少，眼前的这本书出自名家，是一本享誉世界的经典教材，翻译质量也非常高。相信这本经典之作能带你踏入人工智能之门。

王斌

小米集团技术委员会主席、人工智能实验室主任

人工智能与计算机技术几乎同时起步，但与计算机技术几乎线性的发展路径不同，人工智能的发展路径经历了几次大的转向，目前越来越依赖于数据科学和高性能计算机器的发展，因而知识和技术覆盖范围非常广泛。作为一本经典教材，本书提纲挈领，像百科全书一样涵盖了人工智能的大部分领域，每章几乎都对应了一个人工智能问题或技术分支，甚至可以单独成为一门课程。本书的特色是不纠缠于技术和工程细节，直击问题的本质和方法的底层逻辑，在帮助读者形成人工智能领域知识的整体大框架的基础上，增强读者对人工智能基础问题和技术方法的理解和进一步学习的能力。

肖睿

北大青鸟研究院院长

课工场创始人

方法不止，智能无疆

《人工智能：现代方法》是一本经典教材。我和作者斯图尔特·罗素教授相识，和译者团队的张志华教授相熟，特别高兴为最新中文版写几句话。

书名中的“现代方法”，罗素教授的标准解释是“选择从当下的角度讲述人工智能的故事”。从初版到现在的第 4 版，确实如此。比如，这一版的第五部分“机器学习”就重点介绍了过去十年的热点，特别是深度学习和强化学习，如果再加上第六部分“沟通、感知和行动”中的自然语言处理、计算机视觉和机器人学，这两部分似乎就是“现代”人工智能的全部，为什么还要前面四个部分呢？

要回答这个问题，需要对“现代方法”做另一个层次的解读，这关乎人工智能这门学科的性质。

我认为，人工智能首先是一门技术，和计算机、互联网等技术类似，不同于物理学和生命科学那样的科学。科学是寻求事物和现象背后的规律，例如揭示宇宙奥秘的万有引力定律、相对论和量子力学，揭示生命奥秘的进化论和基因。技术是创造新事物和新现象，例如以指南针为代表的中国四大发明，以飞机和计算机为代表的现代技术。

技术发明和科学发现是两种独立的原始创新活动，把科学视为技术的基础，这是偏见。有些技术确实是基于既有科学原理，例如原子弹是核物理发展到一定阶段的产物，但原理只提供了可能性，没有链式反应和内爆技术等一系列技术发明，原子弹不会成为必然。有些技术并不基于科学原理，例如计算机的基础是图灵可计算理论，这是“人工”理论。更多技术在发明时并不明白背后的原理，例如指南针发明时并无电磁学，飞机发明时并无空气动力学。人工智能也一样，深度学习成功应用后，可解释性成为热点，至今理论还在探索中。如果没有深度学习的发明和实践，可解释性理论从何而起呢？

人工智能研究是应该寻求理论基础，还是应该探索实现更强智能的新方法？两者都该做，但后者是主旋律：先有方法和实现，后有理论解释，先有智能技术，后有智能科学，如此迭代发展。智能技术无止境，智能科学也无止境，可以有解释现有人工智能的专门理论，没有指导未来人工智能的通用理论。

经典人工智能时代，我国的最大贡献是机器定理证明的“吴方法”。吴方法提出前一年，吴文俊院士曾撰文指出：“西方数学史家往往以希腊式的严密推理相标榜，并以中国数学从来没有达到演绎科学的形式相指责。然而，我们已经看到，在微积分的发明上希腊形式的那种脆弱性以及与之相较中国式数学的生命力。”后来，他更明确地指出：“它（中国数学）重视计算，是计算性，构造性，也是算法性的。大部分的重要结果都以‘术’的形式表示，而‘术’通常相当于现代的算法。”算法不是数学推理，而是人构造问题解决方案，就是方法。

从探索实现智能的方法论角度看这本书，就容易看出“大而有序”：第二部分“问题求解”是人在设计搜索、博弈和约束满足问题的解决方案；第三部分“知识、推理和规划”是人定义逻辑推理、人整理知识以及人设计的“自动规划”；第四部分“不确定知识和不确定推理”引入了不确定性和概率方法，以实现更强智能，但所有智能仍然是人设计决定的；第五部分“机器学习”，人类后退一步，只设计学习方法，让机器自己“学习”，特别是强化学习，只定义

基本规则，智能主要来自与环境的交互，智能实现重大跃升。然而，深度学习和强化学习虽然更强大，但学到的知识是隐式的，获得的智能不可解释，要打开机器学习的“黑盒”，还需要前四部分的传统方法，当然也可能需要探索全新方法。就此而言，没有比“现代方法”更好的词来概括这本书了。

贯穿全书的核心方法论是“智能体”。罗素教授把人工智能定义为“对从环境中接受感知并执行行动的智能体的研究”。这个概念稍加扩展，就既能概括以机器为载体的人工智能，也能概括以有机体为载体的生物智能——生物就是感知环境并适应环境的有机智能体。更一般地，我认为“智能是系统通过获取和加工信息而获得的一种能力，从而实现从简单到复杂的演化”，这当然也同时涵盖了生物智能和机器智能。

在自然界已知的事物和现象中，人和人脑是最复杂的系统，人类智能是最复杂的现象，因此，脑科学被视为“自然科学的最后疆域”。然而，没有理由相信，人类是生物进化的最后阶段，人类智能是最高水平的智能，有机体是智能的唯一载体。以计算机为载体的人工智能，揭开了机器智能大幕的一角，制造更复杂的机器，实现更强大的智能，机器智能将为科学研究创造无穷无尽的新对象。在这个意义上，智能是“科学的无尽疆域”，而人工智能这个“现代方法”，正是开疆拓土的动力之源。

方法不止，智能无疆，“人工智能：现代方法”这个书名可以永远延续下去。

黄铁军
北京智源人工智能研究院院长、北京大学计算机学院教授
2022 年 10 月 15 日

唯思想永恒

深度学习是机器学习最前沿的领域，它促进了人工智能技术产生了革命性进展，特别是给计算机视觉、语音识别、自然语言处理、棋牌游戏以及某些科学领域带来了颠覆性的突破。深度学习同时驱动了新的机器学习范式产生，比如生成对抗学习、元学习等；并使强化学习和因果学习得以“复兴”，展示更为强大的潜力。斯图尔特·罗素（Stuart Russell）和彼得·诺维格（Peter Norvig）两位教授的这本书在这一背景下于 2021 年年初出版正应其时。

人工智能是一个大领域，该书是一本“大”书，作者是大学者。该书全方位探究了人工智能这一领域，涵盖了从基础知识、模型方法、社会伦理到应用专题等各个层面。正如作者在序言中所提到的，本版中约 25% 的内容是全新的，剩下的 75% 也做了大量修改，以呈现出更加完整的人工智能领域图景，且本版中 22% 的参考文献是 2010 年以后出版的。此外，作者邀请了 9 位相关方向最有代表性的学者撰写了部分内容。本书主要包括 5 方面的内容：问题求解的搜索方法，基于知识的推理和规划方法（逻辑和知识表示），知识和推理中的不确定性（概率推理、概率编程和多智能体决策），机器学习（概率方法、深度学习和强化学习），应用专题（自然语言处理、计算机视觉和机器人学）。此外，书中还讨论了人工智能面临的哲学、伦理和安全等社会问题。书中也蕴含了作者对人工智能的理解和思考，处处闪烁着思想的光辉，耐人回味。比如，本版的封面展示了人工智能各个发展阶段最重要事件和人物，体现了作者的别具匠心。第 1 章关于人工智能的思想、历史发展等的论述深刻、透彻和精辟。第 28 章讨论某些具有前瞻性的想法和方向。我本人在阅读时受到的启发良多，大有裨益。

“南朝四百八十寺，多少楼台烟雨中。”人工智能试图模拟人类的行为和思维，是一个最富有期待和遐想的学科，其发展波澜壮阔、起伏跌宕。她经历了热情高涨和期望无限的早期（1952—1969），通用搜索机制局限所导致的回落期（1966—1973），以专家系统为代表的基于规则学习的崛起期（1969—1986），神经网络联结主义的回归期（1986—1995），统计机器学习的复兴期（1995—现在），以及大数据驱动的深度学习的突破期（2006—现在）。人工智能从哲学、数学、经济学、神经科学、心理学、计算机科学、控制科学、语言学等诸多学科中汲取思想、观点和技术，滋养并发展自身。机器学习试图从数据或经验中学习进而提升机器的能力或性能，这不同于人工智能，但她是目前趋向人工智能的一个最重要或有效的途径。

人工智能是思想发轫、观点争鸣、技术创新的汇集地，是学术英雄辈出的荟萃地。人工智能的发展历程告诉我们：发展人工智能技术需要高度的想象力、创造力和执行力，需要务实、理性、严谨的求是态度。人工智能未来仍会经历波折，各种潮流、观点也会纷争喧嚣，但沉淀下来的是隽永的思想。

我非常感谢人民邮电出版社杨海玲编辑的信任，邀请我的学生来承担该书的中文翻译。译稿的初稿是由我的博士生张博雅、陈坤，已毕业的硕士田超、吴凡和赵申剑，以及博士后顾卓尔完成的。博雅和陈坤对全书译稿进行了统一审校，我的其他在读博士生也参与了相关章节的审校。他们的背景分别是统计学、数据科学和计算机科学，这有益于他们合作翻译该书。然而他们在人工智能领域仍都是新人，知识结构还不全面，但是他们勤于学习、执行力极强、工作专注。在半年左右的时间内完成了译著的初稿，之后又经过自校对、交叉校对等环节力图使译

著保持正确性和一致性。我为他们的责任心和独立工作能力感到自豪。

由于我们深感自己的中英文能力都有限，译文还是比较生涩，难免出现不当之处，而且我们特别担心未能完整地传达出原作者的真实思想和观点。因此，我们强烈地建议有条件的读者去阅读英文原著，也非常期待大家继续指正译著，以便今后进一步修订完善。我恳请读者多给予译者以鼓励。请把你们的批评留给我，这是我作为他们的导师必须要承担的。

最后，我希望我的学生们能享受其翻译过程，翻译和阅读这么一部大书得以领略艾伦·图灵、冯·诺依曼、诺伯特·维纳、理查德·贝尔曼、库尔特·哥德尔、约翰·麦卡锡、马文·明斯基、唐纳德·米奇、爱德华·费根鲍姆、艾伦·纽厄尔、赫伯特·西蒙等在人工智能领域中的工作，感悟他们的思想、领略他们的智慧，何其美哉！我们当谦卑再谦卑，勤奋更勤奋。是以代写此序为勉！

张志华

2022 年 9 月 12 日

中文版致谢

首先，我们要感谢原作者在本书翻译时给予我们的帮助，感谢人民邮电出版社对我们的信任和支持。

这是一本“大”书，涵盖了人工智能的广泛领域。我们几位译者来自不同专业，能力有限且工作量巨大。在翻译过程中，许多老师和同学给予我们很大的帮助。在此我们一一列出，以表示我们衷心的感谢！

特别感谢韩燮教授审校了全书的第三部分（即第 7 章～第 11 章）。赵融和郭新东同学审校了第 26 章。

我们实验室的其他同学帮助我们进行了校对：陈雨静（第 1 章和第 2 章）、崔圣宇（第 3 章部分）、罗维俭（第 3 章部分和第 4 章）、彭洋（第 5 章和第 6 章）、韩雨泽（第 12 章）、张宇航（第 13 章）、梁家栋（第 14 章）、谢广增（第 15 章）、王迩东（第 16 章）、李翔（第 17 章）、金昊（第 18 章）、谢楚焓（第 20 章）、林大超（第 21 章）、杨文昊（第 22 章）、程昊（第 23 章和第 24 章）、胡一征（第 25 章）、林诗韵、赵悦楷和张良宇（第 27 章和第 28 章）。我们还要感谢邓辉、李威、秦钢、王晓雷、魏太云、肖睿、姚远和张淞进行了专家审读。

他们对译文中的专业术语、中文语句、算法、公式以及原文中可能存在的问题等提出了很多宝贵的修改意见，增强了译文的准确性和可读性。当然，现在的译文仍存在一些没有及时发现的问题，因此修订工作将继续更新。我们恳请读者能提供反馈，以便我们在后续版本中修正问题。

此外，我们还要感谢刘艳云老师，感谢她帮助我们与出版社沟通交流，处理相关事务。

最后，感谢我们的导师张志华教授，感谢老师在翻译过程中的指导和对译文的审校。

张博雅（北京大学前沿交叉研究院）
陈坤（北京大学数学科学学院）
2022 年 9 月 15 日

前言

人工智能（artificial intelligence，AI）是一个大领域，本书也是一本“大”书。我们试图全方位探索这一领域。书中内容涵盖逻辑、概率和连续数学，感知、推理、学习和行动，以及公平、信任、社会公益和安全，应用范围从微电子设备到机器人行星探测器，再到拥有几十亿用户的在线服务。

本书的副标题是“现代方法”。这意味着我们选择从当下的角度讲述人工智能的故事。我们将现有已知的内容融合到统一的框架中，使用当今流行的思想和术语重新构建早期的工作。对那些因为研究领域不是本书所及而没有得到本书重视的人，我们深表歉意。

第 4 版新变化

第 4 版反映了自 2010 年第 3 版面世以来人工智能领域发生的下列变化。

- 由于数据、计算资源和新算法的可用性增加，我们更关注机器学习而不是人工设计的知识工程。
- 增加了专门的章节介绍深度学习、概率编程和多智能体系统的相关内容。
- 修订了自然语言理解、机器人学和计算机视觉的内容，以反映深度学习的影响。
- 在第 26 章“机器人学”中包括了与人类互动的机器人以及强化学习在机器人学中的应用。
- 之前，我们将人工智能的目的定义为创建一些试图最大化期望效用的系统，其中具体效用信息——目标——由系统的人类设计师提供。现在我们不再假设目标是固定的，也不再假设人工智能系统知道目标，相反，人工智能系统可能不确定人类的真正目标。它必须学习到要最大化的内容，必须在不确定目标的情况下也能适当地发挥作用。
- 我们增加了人工智能对社会影响的相关内容，包括道德、公平、信任和安全等重要问题。
- 我们把习题从每章末尾移到了网站。这让我们能够不断添加、更新和改进习题，以满足教师的需求并反映这一领域和人工智能相关软件工具的进展。
- 书中约 25% 的内容是全新的，剩下的 75% 也做了大量修改，以呈现出更加统一的人工智能领域图景。本版中 22% 的参考文献是 2010 年以后出版的。

本书概述

智能体（intelligent agent）的概念是贯穿整本书的主题思想。我们将人工智能定义为对从环境中接收感知并执行动作的智能体的研究。每个这样的智能体都要实现一个将感知序列映射为动作的函数，我们介绍了表示这些函数的不同方法，如反应型智能体、实时规划器、决策论系统和深度学习系统。我们强调，学习既是构造良好系统的方法，也是将设计者的影响范围扩展到未知环境的方法。我们没有把机器人学和视觉看作独立定义的问题，而是将其看作实现目标的服务。我们强调任务环境在确定合适的智能体设计中的重要性。

我们的主要目标是传达在过去 70 多年的人工智能研究和过去 2000 多年的相关工作中涌现出现的思想。在表达这些思想时，我们在保持准确性的前提下尽量避免过于拘泥于形式。书中

提供了数学公式和伪代码算法，让关键思想具体化；附录 A 中给出了数学概念和符号，附录 B 中给出了伪代码。

本书主要用作本科人工智能课程或课程序列的教科书。本书共 28 章，每章大约需要一周的课程量，因此完成整本书的教学需要两学期的时间。如果课程只有一学期，可以按教师和学生的兴趣选择部分章节进行教学。本书也可用于研究生课程（可能需要增加参考文献中建议的一些主要资料）或用于自学或作为参考书。

在整本书中，定义了新**术语**的地方，都会以蓝色粗体显示。该术语的后续重要用法也以粗体显示。本书还提供了简要的索引。

阅读本书唯一的先修要求是对计算机科学基本概念（算法、数据结构、复杂性）的熟悉程序达到大学二年级的水平。大学一年级的微积分和线性代数知识对一些主题的阅读很有帮助。

在线资源

在线资源可通过培生教育集团的官方网站或本书的配套网站获得。本书的配套网站上有以下内容。

- 习题、编程项目和研究项目。这些不再放在每章末尾，只在网站提供。在本书中，我们将使用“习题 6.NARY”之类的名称引用在线习题。网站上的说明允许您按名称或主题查找习题。
- 使用 Python、Java 和其他编程语言实现的本书中的算法（目前托管在 GitHub 上）。
- 1500 多所使用过本书的学校名单，其中许多都附有在线课程材料和教学大纲的链接。
- 供学生和教师使用的补充材料及其链接。
- 书中可能存在的错误以及关于如何报告书中错误的说明。

图书封面

封面描绘了 1997 年国际象棋比赛中决定性的第 6 场比赛的最终棋局，在这场比赛中，IBM 的“深蓝”（Deep Blue）计算机击败了加里 • 卡斯帕罗夫（Garry Kasparov）（执黑棋），这是计算机首次在国际象棋比赛中击败世界冠军。卡斯帕罗夫位于封面顶部，他的右边是前世界冠军李世石和 DeepMind 的 AlphaGo 项目之间进行的历史性围棋比赛的第二场比赛的关键局势。AlphaGo 的第 37 手违背了几个世纪以来的围棋正统观念，人类专家认为这是一个令人尴尬的错误，但结果证明这一走法是正确的。封面上，左上角是由波士顿动力公司制造的 Atlas 人形机器人，埃达 • 洛芙莱斯（Ada Lovelace，世界上第一位计算机程序员）和艾伦 • 图灵（Alan Turing，他的基础工作定义了人工智能）之间的是自动驾驶汽车感知环境的画面，棋盘底部的是火星探测漫游者机器人和逻辑学研究先驱亚里士多德的雕像，英文书名背后是亚里士多德的《论动物的运动》（*De Motu Animalium*）中的规划算法，棋盘面上是联合国全面禁止核试验条约组织（UN Comprehensive Nuclear-Test-Ban Treaty Organization）使用的用于从地震信号中检测核爆炸的概率编程模型。

致谢

制作一本书需要无数人的帮助。600 多人阅读了本书的部分内容，并提出了改进意见。我

们感谢他们所有人。在这里，我们只列出几位特别重要的贡献者。首先是撰稿人：

- 朱迪亚·珀尔（Judea Pearl）（13.5 节）；
- 维卡什·曼辛卡（Vikash Mansinghka）（15.4 节）；
- 迈克尔·伍尔德里奇（Michael Wooldridge）（第 18 章）；
- 伊恩·古德费洛（Ian Goodfellow）（第 21 章）；
- 雅各布·德夫林（Jacob Devlin）和张明伟（Ming-Wei Chang）（第 24 章）；
- 吉滕德拉·马利克（Jitendra Malik）和戴维·福赛思（David Forsyth）（第 25 章）；
- 安卡·德拉甘（Anca Dragan）（第 26 章）。

然后是本书出版过程中的关键角色：

- 辛西娅·杨（Cynthia Yeung）和玛莉卡·坎托（Malika Cantor）（项目管理）；
- 朱莉•萨斯曼（Julie Sussman）和汤姆•加洛韦（Tom Galloway）（文字加工和写作建议）；
- 奥马里·斯蒂芬斯（Omari Stephens）（插图）；
- 特蕾西·约翰逊（Tracy Johnson）（编辑）；
- 埃琳·奥尔特（Erin Ault）和罗丝·克南（Rose Kernan）（封面设计和颜色转换）；
- 纳林·奇伯（Nalin Chhibber）、萨姆·戈托（Sam Goto）、雷蒙·拉卡兹（Raymond de Lacaze）、拉维·莫汉（Ravi Mohan）、夏兰·奥赖利（Ciaran O'Reilly）、阿米特·帕特尔（Amit Patel）、德拉戈米尔•拉迪夫（Dragomir Radiv）和萨马格拉•夏尔马（Samagra Sharma）（在线代码开发和指导）；
- Google Summer of Code students（在线代码开发）。

斯图尔特想要感谢他的妻子洛伊·谢弗洛特（Loy Sheflott），感谢她无尽的耐心和无限的智慧。他希望戈登（Gordon）、露西（Lucy）、乔治（George）和艾萨克（Isaac）能很快读到本书，并原谅他在本书上花了这么长时间。感谢 RUGS（Russell's Unusual Group of Students）一如既往地提供了非同寻常的帮助。

彼得想要感谢他的父母托尔斯滕（Torsten）和格尔达（Gerda）让他迈出第一步，感谢他的妻子克丽丝（Kris）、孩子贝拉（Bella）和朱丽叶（Juliet）、同事、老板以及朋友在他漫长的写作和修改过程中鼓励和包容他。

作者简介

斯图尔特·罗素（Stuart Russell）1962 年出生于英国朴茨茅斯。1982 年，获得牛津大学物理学一等荣誉学士学位。1986 年，获得斯坦福大学计算机科学博士学位。之后他进入加利福尼亚大学伯克利分校，任计算机科学系教授，并曾担任系主任，人类兼容人工智能中心主任，他也是史密斯 – 扎德（Smith-Zadeh）工程系讲席教授。1990 年，他获得了美国国家科学基金会（NSF）杰出青年科学家总统奖；1995 年，他成为计算机与思想奖的获奖人之一。他是美国人工智能协会（AAAI）、美国计算机协会（ACM）和美国科学促进协会的会士，牛津大学瓦德汉学院的荣誉院士和安德鲁·卡内基（Andrew Carnegie）院士。2012 年到 2014 年，他在巴黎担任布莱斯·帕斯卡（Blaise Pascal）主席。他在人工智能领域发表了 300 多篇论文，涉及范围广泛。他的其他著作包括 *The Use of Knowledge in Analogy and Induction*、*Do the Right Thing: Studies in Limited Rationality*（与 Eric Wefald 合著）和《AI 新生：破解人机共存密码——人类最后一个大问题》（*Human Compatible: Artificial Intelligence and the Problem of Control*）。

彼得·诺维格（Peter Norvig）曾任谷歌公司的研究总监、核心网络搜索算法的负责人。他曾与他人合作共同教授了一门有 16 万名学生注册的在线人工智能课程，帮助开启了当下的大规模在线公开课程的大幕。他曾担任美国宇航局艾姆斯研究中心计算科学部的负责人，负责人工智能和机器人学的研究和开发。他获得了布朗大学应用数学学士学位和加利福尼亚大学伯克利分校计算机科学博士学位。他曾任南加利福尼亚大学的教授和加利福尼亚大学伯克利分校、斯坦福大学的教师。他是美国人工智能协会和美国计算机协会的会士，以及美国艺术与科学院和加利福尼亚科学院的院士。他的其他著作有 *Paradigms of AI Programming: Case Studies in Common Lisp*、*Verbmobil: A Translation System for Face-to-Face Dialog* 和 *Intelligent Help Systems for UNIX*。

两位作者共同获得了 2016 年首届 AAAI/EAAI 杰出教育家奖。

资源与服务

本书由异步社区出品，社区（https://www.epubit.com/）为您提供相关资源和后续服务。

配套资源

本书提供全书的彩图及程序代码。您可以扫描右侧二维码，发送“59811”，添加异步助手，获取本书配套资源。

提交勘误

作者和编辑尽最大努力来确保书中内容的准确性，但难免会存在疏漏。欢迎您将发现的问题反馈给我们，帮助我们提升图书的质量。

当您发现错误时，请登录异步社区，按书名搜索，进入本书页面，单击“提交勘误”，输入勘误信息，单击“提交”按钮即可。本书的作者和编辑会对您提交的勘误进行审核，确认并接受后，您将获赠异步社区的积分。积分可用于在异步社区兑换优惠券、样书或奖品。

扫码关注本书

扫描下方二维码，您将会在异步社区微信服务号中看到本书信息及相关的服务提示。

与我们联系

我们的联系邮箱是 contact@epubit.com.cn。

如果您对本书有任何疑问或建议，请您发邮件给我们，并请在邮件标题中注明本书书名，以便我们更高效地做出反馈。

如果您有兴趣出版图书、录制教学视频或者参与技术审校等工作，可以通过邮件与本书责任编辑联系。

如果您来自学校、培训机构或企业，想批量购买本书或异步社区出版的其他图书，也可以发邮件给我们。

如果您在网上发现有针对异步社区出品图书的各种形式的盗版行为，包括对图书全部或部分内容的非授权传播，请您将怀疑有侵权行为的链接通过邮件发给我们。您的这一举动是对作

者权益的保护，也是我们持续为您提供有价值的内容的动力之源。

关于异步社区和异步图书

“异步社区”是人民邮电出版社旗下 IT 专业图书社区，致力于出版精品 IT 图书和相关学习产品，为作译者提供优质出版服务。异步社区创办于 2015 年 8 月，提供大量精品 IT 图书和电子书，以及高品质技术文章和视频课程。更多详情请访问异步社区官网 https://www.epubit.com（yanghailing@ptpress.com.cn）。

“异步图书”是由异步社区编辑团队策划出版的精品 IT 专业图书的品牌，依托于人民邮电出版社的计算机图书出版积累和专业编辑团队，相关图书在封面上印有异步图书的 LOGO。异步图书的出版领域包括软件开发、大数据、AI、测试、前端、网络技术等。

异步社区

微信服务号

目录

第一部分　人工智能基础

第二部分　问题求解

第三部分 知识、推理和规划

第六部分 沟通、感知和行动

第七部分 总结

第一部分

人工智能基础

第1章

绪论

在本章中，我们将解释为什么我们认为人工智能是一个最值得研究的课题，并试图定义人工智能究竟是什么。这是开启人工智能学习之旅之前不错的准备。

我们称自己为智人（有智慧的人），因为**智能**（intelligence）对我们来说尤其重要。几千年来，我们一直试图理解我们是如何思考和行动的，也就是不断地了解我们的大脑是如何凭借它那小部分物质去感知、理解、预测并操纵一个远比其自身更大更复杂的世界。**人工智能**（artificial intelligence，AI）领域不仅涉及理解，还涉及构建智能实体。这些智能实体机器需要在各种各样新奇的情况下，计算如何有效和安全地行动。

人工智能经常被各种调查列为最有趣、发展最快的领域之一，现在每年创造的价值超过一万亿美元。人工智能专家李开复预测称，人工智能对世界的影响“将超过人类历史上的任何事物”。此外，人工智能的研究前沿仍是开放的。学习较古老科学（如物理学）的学生可能会认为最好的想法都已经被伽利略、牛顿、居里夫人、爱因斯坦等人发现了，但当下人工智能仍然为专业人员提供了许多机会。

目前，人工智能包含大量不同的子领域，从学习、推理、感知等通用领域到下棋、证明数学定理、写诗、驾车或诊断疾病等特定领域。人工智能可以与任何智能任务产生联系，是真正普遍存在的领域。

1.1 什么是人工智能

我们声称人工智能很有趣，但是我们还没有描述它是什么。历史上研究人员研究过几种不同版本的人工智能。有些根据对人类行为的复刻来定义智能，而另一些更喜欢用“**理性**”（rationality）来抽象正式地定义智能，直观上的理解是做“正确的事情”。智能主题的本身也各不相同：一些人将智能视为内部思维过程和推理的属性，而另一些人则关注智能的外部特征，也就是智能行为。[①]

从人与理性[②]以及思想与行为这两个维度来看，有 4 种可能的组合，而且这 4 种组合都有其追随者和相应的研究项目。他们所使用的方法必然是不同的：追求类人智能必须在某种程度上是与心理学相关的经验科学，包括对真实人类行为和思维过程的观察和假设；而理性主义方法涉及数学和工程的结合，并与统计学、控制理论和经济学相联系。各个研究团体既互相轻视又互相帮助。接下来，让我们更细致地探讨这 4 种方法。

① 公众有时会将“人工智能”和“机器学习”这两个术语混淆。机器学习是人工智能的子领域，研究基于经验提升表现的能力。有些人工智能系统使用机器学习方法来获得能力，有些则不然。

② 我们并不是在暗示人类是“非理性的”，不是像字典上所说的“被剥夺了正常的心智清晰度”。我们只是承认人类的决策在数学上并不总是完美的。

1.1.1 类人行为：图灵测试方法

图灵测试（Turing test）是由艾伦·图灵（Alan Turing）提出的（Turing, 1950），它被设计成一个思维实验，用以回避“机器能思考吗？”这个哲学上模糊的问题。如果人类提问者在提出一些书面问题后无法分辨书面回答是来自人还是来自计算机，那么计算机就能通过测试。在第27章中，我们会讨论图灵测试的细节，以及一台通过图灵测试的计算机是否真的具备智能。目前，为计算机编程使其能够通过严格的应用测试尚有大量工作要做。计算机需要具备下列能力：

- **自然语言处理**（natural language processing），以使用人类语言成功地交流；
- **知识表示**（knowledge representation），以存储它所知道或听到的内容；
- **自动推理**（automated reasoning），以回答问题并得出新的结论；
- **机器学习**（machine learning），以适应新的环境，并检测和推断模式。

图灵认为，没有必要对人进行物理模拟来证明智能。然而，其他研究人员提出了**完全图灵测试**（total Turing test），该测试需要与真实世界中的对象和人进行交互。为了通过完全图灵测试，机器人还需要具备下列能力：

- **计算机视觉**（computer vision）和语音识别功能，以感知世界；
- **机器人学**（robotics），以操纵对象并行动。

以上6个学科构成了人工智能的大部分内容。然而，人工智能研究人员很少把精力用在通过图灵测试上，他们认为研究智能的基本原理更为重要。当工程师和发明家停止模仿鸟类，转而使用风洞并学习空气动力学时，对“人工飞行”的探索取得了成功。航空工程学著作并未将其领域的目标定义为制造“能像鸽子一样飞行，甚至可以骗过其他真鸽子的机器。”

1.1.2 类人思考：认知建模方法

我们必须知道人类是如何思考的，才能说程序像人类一样思考。我们可以通过3种方式了解人类的思维：

- **内省**（introspection）——试图在自己进行思维活动时捕获思维；
- **心理实验**（psychological experiment）——观察一个人的行为；
- **大脑成像**（brain imaging）——观察大脑的活动。

一旦我们有了足够精确的心智理论，就有可能把这个理论表达为计算机程序。如果程序的输入/输出行为与相应的人类行为相匹配，那就表明程序的某些机制也可能在人类中存在。

例如，开发通用问题求解器（General Problem Solver，GPS）的艾伦·纽厄尔（Alan Newell）和赫伯特·西蒙（Herbert Simon）并不仅仅满足于让他们的程序正确地求解问题，他们更关心的是将推理步骤的顺序和时机与求解相同问题的人类测试者进行比较（Newell and Simon, 1961）。**认知科学**（cognitive science）这一跨学科领域汇集了人工智能的计算机模型和心理学的实验技术，用以构建精确且可测试的人类心智理论。

认知科学本身是一个引人入胜的领域，值得多本教科书和至少一部百科全书（Wilson and Keil, 1999）来介绍。我们会偶尔评论人工智能技术和人类认知之间的异同，但真正的认知科学必须建立在对人类或动物实验研究的基础上。这里，我们假设读者只有一台可以做实验的计算机，因此我们将把这方面的内容留给其他书籍。

在人工智能发展的早期，这两种方法经常会混淆。有作者认为，如果算法在某个任务中表现良好，就会是建模人类表现的良好模型，反之亦然。而现代作者将这两种主张分开，这种区分使人工智能和认知科学都得到了更快的发展。这两个领域相互促进，值得一提的是计算机视

觉领域，它将神经生理学证据整合到了计算模型中。最近，将神经影像学方法与分析数据的机器学习技术相结合，开启了“读心”能力（即查明人类内心思想的语义内容）的研究。这种能力反过来可以进一步揭示人类认知的运作方式。

1.1.3　理性思考：“思维法则”方法

希腊哲学家亚里士多德是最早试图法则化“正确思维”的人之一，他将其定义为无可辩驳的推理过程。他的**三段论**（syllogism）为论证结构提供了模式，当给出正确的前提时，总能得出正确的结论。举个经典的例子，当给出前提*苏格拉底是人*和*所有人都是凡人*时，可以得出结论*苏格拉底是凡人*。[这个例子可能是塞克斯都 • 恩披里柯（Sextus Empiricus）提出的而不是亚里士多德提出的。] 这些思维法则被认为支配着思想的运作，他们的研究开创了一个称为**逻辑**（logic）的领域。

19 世纪的逻辑学家建立了一套精确的符号系统，用于描述世界上物体及其之间的关系。这与普通算术表示系统形成对比，后者只提供关于数的描述。到 1965 年，任何用逻辑符号描述的可解问题在原则上都可以用程序求解。人工智能中所谓的**逻辑主义**（logicism）传统希望在此类程序的基础上创建智能系统。

按照常规的理解，逻辑要求关于世界的认知是*确定的*，而实际上这很难实现。例如，我们对政治或战争规则的了解远不如对国际象棋或算术规则的了解。**概率**（probability）论填补了这一鸿沟，允许我们在掌握不确定信息的情况下进行严格的推理。原则上，它允许我们构建全面的理性思维模型，从原始的感知到对世界运作方式的理解，再到对未来的预测。它无法做到的是形成智能*行为*。为此，我们还需要关于理性行为的理论，仅靠理性思考是不够的。

1.1.4　理性行为：理性智能体方法

智能体（agent）就是某种能够采取行动的东西（agent 来自拉丁语 agere，意为“做”）。当然，所有计算机程序都可以完成一些任务，但我们期望计算机智能体能够完成更多的任务：自主运行、感知环境、长期持续存在、适应变化以及制定和实现目标。**理性智能体**（rational agent）需要为取得最佳结果或在存在不确定性时取得最佳期望结果而采取行动。

基于人工智能的“思维法则”方法重视正确的推断。做出正确推断有时是作为理性智能体的一部分，因为采取理性行为的一种方式是推断出某个给定的行为是最优的，然后根据这个结论采取行动。但是，理性行为的有些方式并不能说涉及推断。例如，从火炉前退缩是一种反射作用，这通常比经过深思熟虑后采取的较慢的动作更为成功。

通过图灵测试所需的所有技能也使智能体得以采取理性行为。知识表示和推理能让智能体做出较好的决策。我们需要具备生成易于理解的自然语言句子的能力，以便在复杂的社会中生存。我们需要学习不仅是为了博学多才，也是为了提升我们产生高效行为的能力，尤其是在新环境下，这种能力更加重要。

与其他方法相比，基于人工智能的理性智能体方法有两个优点。首先，它比“思维法则”方法更普适，因为正确的推断只是实现理性的几种可能机制之一。其次，它更适合科学发展。理性的标准在数学上是明确定义且完全普适的。我们经常可以从这个标准规范中得出可以被证明能够实现的智能体设计，而把模仿人类行为或思维过程作为目标的设计在很大程度上是不可能的。

由于上述这些原因，在人工智能领域的大部分历史中，基于理性智能体的方法都占据了上风。在最初的几十年里，理性智能体建立在逻辑的基础上，并为了实现特定目标制定了明确的规划。后来，基于概率论和机器学习的方法可以使智能体在不确定性下做出决策，以获得最佳

期望结果。简而言之，人工智能专注于研究和构建做正确的事情的智能体，其中正确的事情是我们提供给智能体的目标定义。这种通用范式非常普遍，以至于我们可以称之为**标准模型**（standard model）。它不仅适用于人工智能，也适用于其他领域。控制理论中，控制器使代价函数最小化；运筹学中，策略使奖励的总和最大化；统计学中，决策规则使损失函数最小；经济学中，决策者追求效用或某种意义的社会福利最大化。

然而在复杂的环境中，完美理性（总是采取精确的最优动作）是不可行的，它的计算代价太高了，因此需要对标准模型做一些重要的改进。第 5 章和第 17 章会探讨**有限理性**（limited rationality）的问题，也就是在没有足够时间进行所有可能的计算的情况下，适当地采取行动。但是，完美理性仍然是理论分析的良好出发点。

1.1.5 益机[①]

自标准模型被提出以来，其一直是人工智能研究的指南，但从长远来看，它可能不是一个正确的模型，原因是标准模型假设我们总是为机器提供完全指定的目标。

人为定义的任务，如国际象棋或最短路径计算之类的，都附带固有的目标，因此标准模型是适用的。然而，在真实世界中，我们越来越难以完全正确地指定目标。例如，在设计自动驾驶汽车时，我们可能会认为目标是安全到达目的地。但是，由于存在其他司机失误、设备故障等原因，在任何道路上行驶都有可能受伤，因此，严格的安全目标是要求待在车库里而不要上路驾驶。向目的地前进和承担受伤风险是需要权衡的，应该如何进行这种权衡？此外，我们能在多大程度上允许汽车采取会惹恼其他司机的行动？汽车应该在多大程度上调控其加速、转向和刹车动作，以避免摇晃乘客？这类问题很难预先回答。在人机交互的整个领域，这些问题尤其严重，自动驾驶只是其中一个例子。

在我们的真实需求和施加给机器的目标之间达成一致的问题称为**价值对齐问题**（value alignment problem），即施加给机器的价值或目标必须与人类的一致。如果我们在实验室或模拟器中开发人工智能系统（就像该领域的大多数历史案例一样），就可以轻松地解决目标指定不正确的问题：重置系统、修复目标然后重试。随着人工智能的发展，越来越强大的智能系统需要部署在真实世界中，这种方法不再可行。部署了错误目标的系统将会导致负面影响，而且，系统越智能，其负面影响就越严重。

回想看似没有问题的国际象棋案例，想象一下，如果机器足够智能，可以推断并采取超出棋盘限制的动作，会发生什么。例如，它可能试图通过催眠或勒索对手，或贿赂观众在对手思考时发出噪声等手段来增加获胜的机会。[②] 它也可能会为自己劫持额外的计算能力。这些行为不是“愚蠢”或“疯狂”的，这些行为是将获胜定义为机器唯一目标的逻辑结果。

一台实现固定目标的机器可能会出现很多不当行为，要预测所有不当行为是不可能的。因此，我们有足够理由认为标准模型是不充分的。我们不希望机器“聪明”地实现它们的目标，而是希望它们实现我们的目标。如果我们不能将这些目标完美地传达给机器，就需要一个新的表述，也就是机器正在实现我们的目标，但对于目标是什么则是不确定的。当一台机器意识到它不了解完整的目标时，它就会有谨慎行动的动机，会寻求许可，并通过观察来更多地了解我们的偏好，遵守人为控制。最终，我们想要的是对人类**可证益的**（provably beneficial）智能体。我们将在 1.5 节中讨论这个主题。

① 根据 beneficial insect 的翻译“益虫”，将 beneficial machine 翻译成“益机”。——译者注

② 鲁伊·洛佩兹（Ruy Lopez）在最早的一本关于国际象棋的书（Lopez, 1561）中写道：“把棋盘放好，让阳光晃进对手的眼睛。”

1.2　人工智能的基础

在本节中，我们将简要介绍为人工智能提供思想、观点和技术的学科的历史。像任何历史一样，本书只关注少数人物、事件和思想，而忽略其他同样重要的。我们围绕一系列问题来组织这段历史。我们不希望带给读者这样一种印象：这些问题是各个学科唯一要解决的问题，或者各个学科都将人工智能作为最终成果而努力。

1.2.1　哲学

- 可以使用形式化规则得出有效结论吗？
- 思维是如何从物质大脑中产生的？
- 知识从何而来？
- 知识如何导致行为？

亚里士多德（Aristotle，公元前 384—公元前 322）制定了一套精确的法则来统御思维的理性部分，他是历史上第一位这样做的哲学家。他发展了一套非正式的三段论系统进行适当的推理，该系统原则上允许人们在给定初始前提下机械地得出结论。

拉蒙·鲁尔（Ramon Llull，约 1232—1315）设计了一种推理系统，发表为 *Ars Magna*（即 *The Great Art*）（Llull, 1305）[①]。鲁尔试图使用实际的机械设备——一组可以旋转成不同排列的纸盘——实现他的系统。

大约在 1500 年，达·芬奇（Leonardo da Vinci，1452—1519）设计了一台机械计算器，虽然当时并未制造，但最近的重构表明该设计是可行的。第一台已知的计算器是在 1623 年左右由德国科学家威廉·席卡德（Wilhelm Schickard，1592—1635）制造的。布莱斯·帕斯卡（Blaise Pascal，1623—1662）于 1642 年建造了滚轮式加法器（Pascaline），并写道：“它产生的效用似乎比动物的所有行为更接近思维。”戈特弗里德·威廉·莱布尼茨（Gottfried Wilhelm Leibniz，1646—1716）制造了一台机械设备，旨在根据概念而非数值进行操作，但其应用范围相当有限。托马斯·霍布斯（Thomas Hobbes，1588—1679）在《利维坦》（*Leviathan*）一书中提出了会思考的机器的想法，用他的话说就是一种“人造动物”，设想“心脏无非就是发条，神经只是一些游丝，而关节不过是一些齿轮。”他还主张推理就像是数值计算，认为“推理就是一种计算，也就是相加减。”[②]

有观点认为，思维至少在某种程度上是根据逻辑或数值规则运作的，可以建立模仿其中的一些规则的物理系统。也有观点说，思维本身就是这样一个物理系统。勒内·笛卡儿（René Descartes，1596—1650）首次清晰地讨论了思维与物质之间的区别。他指出，思维的纯粹物理概念似乎没有给自由意志留下多少空间。如果思维完全受物理法则支配，那么它拥有的自由意志不会比一块“决定”往下掉的石头多。笛卡儿是**二元论**（dualism）的支持者。他认为，人类思维（灵魂或者精神）的一部分处于自然之外，不受物理定律的约束。但是，动物不具备这种二元特性，它们可以被视为机器。

唯物主义（materialism）是二元论的一种替代，它认为大脑根据物理定律的运作构成了思维。自由意志仅仅是实体对可选决策的感知。**物理主义**（physicalism）和**自然主义**（naturalism）

① *Ars Magna* 为拉丁文书名，翻译成英文的书名为 *The Great Art*。——编者注

② 此处对《利维坦》一书中的引用采用了商务印书馆 1985 年 9 月出版的由黎思复、黎廷弼翻译的《利维坦》版本中的译文。——编者注

这两个术语也被用于描述这类与超自然观点相反的观点。

如果给定可以操纵知识的实体思维，接下来的问题就是建立知识的来源。**经验主义**（empiricism）运动始于弗朗西斯·培根（Francis Bacon，1561—1626）的《新工具》（*Novum Organum*）[①]一书，并以约翰·洛克（John Locke，1632—1704）的名言“知识归根到底都来源于经验”为特征。

大卫·休谟（David Hume，1711—1776）的《人性论》（*A Treatise of Human Nature*）（Hume, 1739）提出了现在称为**归纳法**（induction）的原则：通过暴露要素之间的重复联系获得一般规则。

以路德维希·维特根斯坦（Ludwig Wittgenstein，1889—1951）和伯特兰·罗素（Bertrand Russell，1872—1970）的工作为基础，著名的维也纳学派（Sigmund, 2017）———一群在20世纪20年代及20世纪30年代聚集在维也纳的哲学家和数学家——发展了**逻辑实证主义**（logical positivism）学说。该学说认为，所有知识都可以通过逻辑理论来描述，逻辑理论最终与对应于感知输入的**观察语句**（observation sentence）相联系。因此，逻辑实证主义结合了理性主义和经验主义。

鲁道夫·卡尔纳普（Rudolf Carnap，1891—1970）和卡尔·亨普尔（Carl Hempel，1905—1997）的**确证理论**（confirmation theory）试图通过量化应分配给逻辑语句的信念度来分析从经验中获取知识，信念度的取值基于逻辑语句与确证或否定它们的观察之间的联系。卡尔纳普的《世界的逻辑构造》（*The Logical Structure of the World*）（Carnap, 1928）也许是最先提出将思维视为计算过程这一理论的著作。

思维的哲学图景中最后一个要素是知识与动作之间的联系。这个问题对人工智能来说至关重要，因为智能不仅需要推理，还需要动作。而且，只有理解了怎样的行为是合理的，才能理解如何构建行为是合理的（或理性的）智能体。

亚里士多德在《论动物的运动》（*De Motu Animalium*）中指出，动作的合理性是通过目标和动作结果的知识之间的逻辑联系来证明的：

> 但是，思考有时伴随着行为，有时却没有，有时伴随着行动，有时却没有，这是如何发生的？这看起来和对不变的对象进行推理和推断时发生的情况几乎是一样的。但是在那种情况下，结局是一个推测性的命题……而在这里，由两个前提得出的结论是一个行为……我需要覆盖物；斗篷是一种覆盖物。我需要一件斗篷。我需要什么，我必须做什么；我需要一件斗篷。我必须做一件斗篷。结论是，“我必须做一件斗篷”，这是一个行为。

在《尼各马可伦理学》（*Nicomachean Ethics*）（第三卷·第3章，1112b）中，亚里士多德进一步阐述了这个主题，并提出了一个算法：

> 我们考虑的不是目的，而是实现目的的手段。医生并不考虑是否要使一个人健康，演说家并不考虑是否要去说服听众……他们是先确定一个目的，然后考虑用什么手段和方式来达到目的。如果有几种手段，他们考虑的就是哪种手段最能实现目的。如果只有一种手段，他们考虑的就是怎样利用这一手段去达到目的，这一手段又需要通过哪种手段来获得。这样，他们就在所发现的东西中一直追溯到最初的东西……分析的终点也就是起点。如果恰巧遇到不可能的事情，例如需要钱却得不到钱，那么就放弃这种考虑。而所谓可能的事情，就是以我们自身能力可以做到的那些事情。[②]

① 培根的《新工具》（*Novum Organum*）是亚里士多德的《工具论》（*Organon*，又称“思想工具”）的更新。

② 此处对《尼各马可伦理学》一书中的引用采用了商务印书馆2017年8月出版的廖申白翻译的《尼各马可伦理学》版本中的译文。——编者注

2300 年后，纽厄尔和西蒙在他们的**通用问题求解器**（General Problem Solver）程序中实现了亚里士多德的算法。我们现在将其称为贪婪回归规划系统（见第 11 章）。在人工智能理论研究的前几十年中，基于逻辑规划以实现确定目标的方法占据主导地位。

纯粹从行为的角度来思考实现目标通常是有用的，但在某些情况是不适用的。例如，如果有几种不同的方法可以实现目标，我们就需要某种方法来进行选择。更重要的是，确定性地实现一个目标可能是无法做到的，但某些行为仍然必须被实施。那该如何决策呢？安托万 • 阿尔诺（Antoine Arnauld）（Arnauld, 1662）分析了赌博中的理性决策概念，提出了一种量化公式，可以最大化期望收入的货币价值。后来，丹尼尔 • 伯努利（Daniel Bernoulli）（Bernoulli, 1738）引入了更普适的**效用**（utility）概念，可以体现结果的内在主观价值。如第 16 章所述，在不确定性下，理性决策的现代概念涉及最大化期望效用。

在道德和公共政策方面，决策者必须考虑多个个体的利益。杰里米 • 边沁（Jeremy Bentham）（Bentham, 1823）和约翰 • 穆勒（John Stuart Mill）（Mill, 1863）提出了**功利主义**（utilitarianism）思想：基于效用最大化的理性决策应该适用于人类活动的所有领域，包括代表许多个体做出公共政策的决策。功利主义是一种特殊的**结果主义**（consequentialism），行为的预期结果决定了正确与否。

相反，伊曼努尔 • 康德（Immanuel Kant）在 1785 年提出了一种基于规则或**义务伦理学**（deontological ethics）的理论。在该理论中，“做正确的事”不是由结果决定的，而是由管理可行行为的普适社会法则所决定的，可行行为包括“不要撒谎”“不要杀人”等。因此，如果期望的好处大于坏处，那么功利主义者可以撒一个善意的谎言，但康德主义者则不能这样做，因为撒谎本质上就是错误的。穆勒承认规则的价值，但将其理解为基于第一性原理对结果进行推理的高效决策程序。许多现代人工智能系统正是采用了这种方法。

1.2.2　数学

- 得出有效结论的形式化规则是什么？
- 什么可以被计算？
- 如何使用不确定的信息进行推理？

哲学家们提出了人工智能的一些基本理念，但人工智能要成为正规科学，需要逻辑和概率的数学化，并引入一个新的数学分支——计算。

形式化逻辑（formal logic）的思想可以追溯到古希腊、古印度和古代中国的哲学家，但它的数学发展真正始于乔治 • 布尔（George Boole，1815—1864）的工作。布尔提出了命题和布尔逻辑的细节（Boole, 1847）。1879 年，戈特洛布 • 弗雷格（Gottlob Frege，1848—1925）将布尔逻辑扩展到包括对象和关系，创建了沿用至今的一阶逻辑[①]。一阶逻辑除了在人工智能研究的早期发挥核心作用外，还激发了哥德尔和图灵的工作，这些工作支撑了计算本身。

概率（probability）论可以视为信息不确定情况下的广义逻辑，这对人工智能来说是非常重要的考虑。吉罗拉莫 • 卡尔达诺（Gerolamo Cardano，1501—1576）首先提出了概率的概念，并根据赌博事件的可能结果对其进行了刻画。1654 年，布莱斯 • 帕斯卡（Blaise Pascal，1623—1662）在给皮埃尔 • 费马（Pierre Fermat，1601—1665）的信中展示了如何预测一个未完成的赌博游戏的结局，并为赌徒分配平均收益。概率很快成为定量科学的重要组成部分，用于处理不确定的度量和不完备的理论。雅各布 • 伯努利（Jacob Bernoulli，1654—1705，丹尼

① 弗雷格提出的一阶逻辑符号（文本和几何特征的神秘组合）从未流行起来。

尔·伯努利的叔叔）、皮埃尔·拉普拉斯（Pierre Laplace，1749—1827）等人发展了这一理论，并引入了新的统计方法。托马斯·贝叶斯（Thomas Bayes，1702—1761）提出了根据新证据更新概率的法则。贝叶斯法则是人工智能系统的重要工具。

概率的形式化结合数据的可用性，使**统计学**（statistics）成为了一个新研究领域。最早的应用之一是 1662 年约翰·格兰特（John Graunt）对伦敦人口普查数据的分析。罗纳德·费舍尔（Ronald Fisher）被认为是第一位现代统计学家，他汇总了概率、实验设计、数据分析和计算等思想（Fisher, 1922）。在 1919 年，他坚称，如果没有机械计算器“百万富翁”（Millionaire，第一个可以做乘法的计算器），他就无法进行工作，尽管这台计算器的成本远远超过了他的年薪（Ross, 2012）。

计算的历史与数字的历史一样古老，但用于计算最大公约数的欧几里得算法被认为是第一个非平凡的**算法**（algorithm）。“算法”一词源自一位 9 世纪的数学家穆罕默德·本·穆萨·阿尔·花剌子模（Muhammad ibn Musa al-Khwarizmi），他的著作还将阿拉伯数字和代数引入了欧洲。布尔等人讨论了逻辑演绎的算法，到 19 世纪末，人们开始努力将一般的数学推理形式化为逻辑演绎。

库尔特·哥德尔（Kurt Gödel，1906—1978）表明，虽然存在一种有效方法能够证明弗雷格和罗素的一阶逻辑中的任何真实陈述，但是一阶逻辑无法满足表征自然数所需的数学归纳原理。1931 年，哥德尔证明关于演绎的限制确实存在。哥德尔的**不完全性定理**（incompleteness theorem）表明，在任何像皮亚诺算术（Peano arithmetic，自然数的基本理论）这样强的形式化理论中，必然存在一些没有证明的真实陈述。

这个基本结果也可以解释为作用于整数上的某些函数无法用算法表示，即它们无法被计算。这促使艾伦·图灵（Alan Turing，1912—1954）试图准确地描述哪些函数是**可计算的**，即能够通过有效的过程进行计算。丘奇–图灵论题（Church-Turing thesis）提出将图灵机（Turing, 1936）可计算的函数作为可计算性的一般概念。图灵还表明，存在某些任何图灵机都无法计算的函数。例如，没有一台机器能够在广义上判断给定程序是会根据给定的输入返回答案，还是永远运行下去。

尽管**可计算性**（computability）对理解计算很重要，但**易处理性**（tractability）的概念对人工智能的影响更大。粗略地说，如果解决一个问题实例所需的时间随着问题规模呈指数增长，那么这个问题就是难处理的。在 20 世纪 60 年代中期，复杂性的多项式增长和指数增长之间的区别首次被强调（Cobham, 1964; Edmonds, 1965）。因为指数级增长意味着即使是中等规模的问题实例也无法在合理的时间内解决，所以易处理性很重要。

由斯蒂芬·库克（Stephen Cook）（Cook, 1971）和理查德·卡普（Richard Karp）（Karp, 1972）开创的 **NP 完全性**（NP-completeness）理论为分析问题的易处理性提供了基础：任何可以归约到 NP 完全的问题都可能是难处理的。（尽管尚未证明 NP 完全问题一定是难处理的，但大多数理论家都相信这一点。）这些结果与大众媒体对第一台计算机的乐观态度——“比爱因斯坦还快的电子超级大脑！”——形成了鲜明对比。尽管计算机的速度在不断提高，但对资源的谨慎使用和必要的缺陷将成为智能系统的特征。粗略地说，世界是一个极大的问题实例！

1.2.3 经济学

- 我们应该如何根据自己的偏好做出决定？
- 当其他人可能不支持时，我们应该怎么做？
- 当收益可能在很遥远的未来时，我们应该怎么做？

经济学起源于 1776 年，当时亚当 • 斯密（Adam Smith，1723—1790）发表了《国富论》（全名为《国民财富的性质和原因的研究》，*An Inquiry into the Nature and Causes of the Weallth of Nations*）。斯密建议将经济视为由许多关注自身利益的独立主体组成，但他并不主张将金融贪婪作为道德立场。他在较早的著作《道德情操论》（*The Theory of Moral Sentiments*）（Smith, 1759）开篇就指出，对他人福祉的关注是每个个体利益的重要组成部分。

大多数人认为经济学就是关于钱的，而实际上第一个对不确定性下的决策进行数学分析的是安托万 • 阿尔诺（Arnauld, 1662）的最大期望值公式，而这一分析也的确是与赌注的货币价值相关。丹尼尔 • 伯努利（Bernoulli, 1738）注意到，这个公式似乎不适用于更大规模的金钱，例如对海上贸易远征的投资。于是，他提出了基于期望效用最大化的原则，并指出额外货币的边际效用会随着一个人获得更多货币而减少，从而解释了大众的投资选择。

里昂 • 瓦尔拉斯（Léon Walras，1834—1910）为效用理论提供了一个更为普适的基础，即对任何结果（不仅仅是货币结果）的投机偏好。弗兰克 • 拉姆齐（Frank Ramsey）（Ramsey, 1931）以及后来约翰 • 冯 • 诺伊曼（John von Neumann）和奥斯卡 • 摩根斯特恩（Oskar Morgenstern）在他们的著作《博弈论与经济行为》（*The Theory of Games and Economic Behavior*）（Neumann and Morgenstern, 1944）中对这一理论进一步改进。经济学不再是研究金钱的学科，而是对欲望和偏好的研究。

决策论（decision theory）结合了概率论和效用理论，为在不确定性下做出个体决策（经济的或其他的）提供了一个形式化完整的框架，也就是说，概率适当地描述了决策者所处的环境。这适用于“大型”经济体，在这种经济体中，每个主体都无须关注其他独立主体的行为。对“小型”经济体而言更像是一场**博弈**（game）：一个参与者的行为可以显著影响另一个参与者的效用（积极或消极的）。冯 • 诺依曼和摩根斯特恩对**博弈论**（game theory）的发展［也可以参考（Luce and Raiffa, 1957）］得出了令人惊讶的结果，即对于某些博弈，理性智能体应该采用随机（或至少看起来是随机）的策略。与决策论不同，博弈论并没有为行为的选择提供明确的指示。人工智能中涉及多个智能体的决策将在**多智能体系统**（multiagent system）的主题下探讨（第 18 章）。

经济学家（除了一些例外）没有解决上面列出的第三个问题：当行为的收益不是立即产生的，而是在几个连续的行为后产生时，应该如何做出理性的决策。这个课题在**运筹学**（operations research）的领域探讨，运筹学出现在第二次世界大战期间英国对雷达安装的优化工作中，后来发展出了无数民用应用。理查德 • 贝尔曼（Richard Bellman）（Bellman, 1957）的工作将一类序贯决策问题进行了形式化，称为**马尔可夫决策过程**（Markov decision process），我们将在第 17 章研究该问题，并在第 22 章以**强化学习**（reinforcement learning）的主题研究该问题。

经济学和运筹学的工作对理性智能体的概念做出了很大贡献，但是多年来的人工智能研究是沿着完全独立的道路发展的。原因之一是做出理性决策显然是复杂的。人工智能的先驱赫伯特 • 西蒙（Herbert Simon, 1916—2001）凭借其早期工作在 1978 年获得了诺贝尔经济学奖，他指出基于**满意度**（satisficing）的决策模型（做出“够好”的决策，而不是费力地计算最优决策）可以更好地描述实际的人类行为（Simon, 1947）。自 20 世纪 90 年代以来，人工智能的决策理论技术重新引起了人们的兴趣。

1.2.4　神经科学

- 大脑如何处理信息？

神经科学（neuroscience）是对神经系统（尤其是对大脑）的研究。尽管大脑进行思考的

确切方式是科学的奥秘之一，但大脑确实是能思考的现实已经被人们接受了数千年，因为有证据表明，对头部的强烈打击会导致精神丧失。人们也早就知道人的大脑在某种程度上是不同的，大约在公元前 335 年，亚里士多德写道："在所有动物中，人类的大脑与身体大小的比例最大。"[①] 然而，直到 18 世纪中叶，大脑才被广泛认为是意识的所在地。在此之前，意识所在地的候选位置包括心脏和脾脏。

1861 年，保罗·布罗卡（Paul Broca，1824—1880）对脑损伤患者中的失语症（语言缺陷）进行了调查研究，他在大脑左半球发现一个局部区域（现在被称为布罗卡氏区域）负责语音的产生，从而开始了对大脑功能组织的研究。[②] 那时，人们已经知道大脑主要由神经细胞或**神经元**（neuron）组成，但直到 1873 年，卡米洛·高尔基（Camillo Golgi，1843—1926）才发明了一种可以观察单个神经元的染色技术（见图 1-1）。圣地亚哥·拉蒙–卡哈尔（Santiago Ramon y Cajal，1852—1934）在神经组织的开创性研究中使用了该技术。[③] 现在人们普遍认为认知功能是由这些结构的电化学反应产生的。也就是说，一组简单的细胞就可以产生思维、行为和意识。如约翰·希尔勒（John Searle）（Searle, 1992）的精辟名言所说：大脑产生思想。

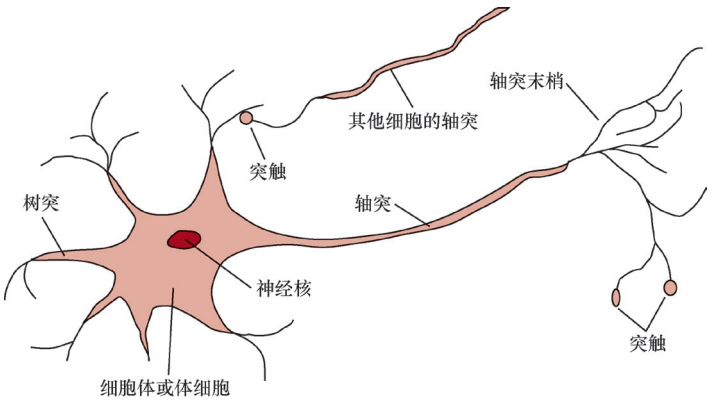

图 1-1 神经细胞或神经元的部分。每个神经元都由一个包含神经核的细胞体或体细胞组成。许多从细胞体中分支出来的纤维状被称为树突，其中的长纤维被称为轴突。轴突伸展的距离很长，比这张图上显示的要长得多。轴突一般长 1 厘米（是细胞体直径的 100 倍），但也可以达到 1 米。一个神经元在称为突触的连接处与其他 10 ～ 100 000 个神经元建立连接。信号通过复杂的电化学反应从一个神经元传递到其他神经元。这些信号可以在短期内控制大脑活动，还可以长期改变神经元的连通性。这些机制被认为是大脑学习的基础。大多数信息都在大脑皮质（大脑的外层）中处理的。基本的组织单元似乎是直径约 0.5 毫米的柱状组织，包含约 20 000 个神经元，并延伸到整个皮质（人类皮质深度约 4 毫米）

现在，我们有了一些关于大脑区域和身体部位之间映射关系的数据，这些部位是受大脑控制或者是接收感官输入的。这样的映射可以在几周内发生根本性的变化，而有些动物似乎具有多个映射。此外，我们还没有完全理解当一个区域受损时其他区域是如何接管其功能的。而且，关于个人记忆是如何存储的，或者更高层次的认知功能是如何运作的，目前几乎没有任何相关理论。

1929 年，汉斯·伯杰（Hans Berger）发明脑电图仪（EEG），开启了对完整大脑活动的测量。功能磁共振成像（fMRI）的发展（Ogawa *et al.*, 1990; Cabeza and Nyberg, 2001）为神经科学家提供了前所未有的大脑活动的详细图像，从而使测量能够以有趣的方式与正在进行的认知

① 后来人们发现树鼩和一些鸟类的脑体比超过了人类的脑体比。
② 许多人引用亚历山大·胡德（Alexander Hood）（Hood, 1824）的论文作为可能的先验资料。
③ 卡哈尔提出了"神经元学说"，高尔基则坚持他的信念，认为大脑的功能主要是在神经元嵌入的连续介质中发挥的。虽然两人共同获得 1906 年的诺贝尔奖，但发表的获奖感言却是相互对立的。

过程相对应。神经元活动的单细胞电记录技术和**光遗传学**（optogenetics）方法的进展（Crick, 1999; Zemelman *et al*., 2002; Han and Boyden, 2007）增强了这些功能，从而可以测量和控制被修改为对光敏感的单个神经元。

用于传感和运动控制的**脑机接口**（brain-machine interface）的发展（Lebedev and Nicolelis, 2006）不仅有望恢复残疾人的功能，还揭示了神经系统许多方面的奥秘。这项工作的一项重要发现是，大脑能够自我调整，使自己成功与外部设备进行交互，就像对待另一个感觉器官或肢体一样。

大脑和数字计算机有不同的特性。如图 1-2 所示，计算机的周期时间比大脑快一百万倍。虽然与高端个人计算机相比，大脑拥有更多的存储和互连，但最大的超级计算机在某些指标上已经与大脑相当。未来主义者充分利用这些数字，指出了一个即将到来的**奇点**（singularity），在这个奇点上计算机达到了超越人类的性能水平（Vinge, 1993; Kurzweil, 2005; Doctorow and Stross, 2012），然后会进一步迅速提高。但是比较原始数字并不是特别有用。即使计算机的容量到达无限也无济于事，在理解智能方面仍然需要进一步的概念突破（见第 28 章）。粗略地说，如果没有正确的理论，更快的机器只会更快地给出错误的答案。

	超级计算机	个人计算机	人类大脑
计算单元	10^6 个 GPU + CPU	8 个 CPU 内核	10^6 列
	10^{15} 个晶体管	10^{10} 个晶体管	10^{11} 个神经元
存储单元	10^{16} 字节（10 PB）RAM	10^{10} 字节（10 GB）RAM	10^{11} 个神经元
	10^{17} 字节（100 PB）磁盘	10^{12} 字节（1 TB）磁盘	10^{14} 个突触
周期时间	10^{-9} 秒	10^{-9} 秒	10^{-3} 秒
运算/秒	10^{18}	10^{10}	10^{17}

图 1-2　领先的超级计算机 Summit（Feldman, 2017）、2019 年的典型个人计算机和人类大脑的粗略对比。数千年来，人类大脑的能力并没有发生太大变化，而超级计算机的计算能力已经从 20 世纪 60 年代的百万次浮点运算（MFLOP）提高到了 20 世纪 80 年代的十亿次浮点运算（GFLOP）、20 世纪 90 年代的万亿次浮点运算（TFLOP）、2008 年的千万亿次浮点运算（PFLOP）以及 2018 年的百亿亿次浮点运算（exaFLOP，1 exaFLOP = 10^{18} 次浮点运算/秒）

1.2.5　心理学

- 人类和动物是如何思考和行为的?

科学心理学的起源通常可以追溯到德国物理学家赫尔曼 • 冯 • 赫尔姆霍茨（Hermann von Helmholtz，1821—1894）和他的学生威廉 • 温特（Wilhelm Wundt，1832—1920）的工作。赫尔姆霍茨将科学方法应用于人类视觉的研究，他的 *Handbook of Physiological Optics* 被描述为“关于人类视觉的物理学和生理学的最重要的专著”（Nalwa, 1993, p.15）。1879 年，温特在莱比锡大学开设了第一个实验心理学实验室。温特坚持严格控制的实验，他实验室的工作人员在进行感知或联想任务的同时，内省他们的思维过程。严格的控制在很大程度上帮助心理学成为了一门科学，但是数据的主观性质使得实验者不太可能会推翻自己的理论。

另外，研究动物行为的生物学家缺乏内省的数据，于是发展了一种客观的方法，赫伯特 • 詹宁斯（Herbert S. Jennings）（Jennings, 1906）在他有影响力的著作 *Behavior of the Lower Organisms* 中对此进行了描述。约翰 • 沃森（John Watson，1878—1958）领导的**行为主义**（behaviorism）运动将这一观点应用于人类，以内省无法提供可靠证据为由，拒绝任何涉及心理过程的理论。行为主义者坚持只研究施加动物的感知（或刺激）及其产生的行为（或反应）

的客观度量。行为主义发现了很多关于老鼠和鸽子的知识，但是在理解人类方面却不太成功。

认知心理学（cognitive psychology）认为大脑是一个信息处理设备，这至少可以追溯到威廉·詹姆斯（William James，1842—1910）的著作。赫尔姆霍茨也坚持认为感知涉及一种无意识的逻辑推断形式。在美国，认知观点在很大程度上被行为主义所掩盖，但在弗雷德里克·巴特利特（Frederic Bartlett，1886—1969）所领导的剑桥大学应用心理学系，认知模型得以蓬勃发展。巴特利特的学生和继任者肯尼斯·克雷克（Kenneth Craik）（Craik, 1943）所著的 *The Nature of Explanation* 强有力地重新确立了诸如信念和目标之类的“精神”术语的合法性，认为它们就像用压力和温度来讨论气体一样科学，尽管气体是由既不具有压力又不具有温度的分子组成。

克雷克指出了知识型智能体的 3 个关键步骤：（1）刺激必须转化为一种内在表示；（2）认知过程处理表示，从而产生新的内部表示；（3）这些过程反过来又被重新转化为行为。他清晰地解释了为什么这是一个良好的智能体设计：

> 如果有机体拥有一个“小规模的模型”，建模了外部现实及其在脑海中可能采取的行为，那么它就能够尝试各种选择，得出哪个是最好的，并在未来出现情况之前加以应对。有机体可以利用过去的知识处理现在和未来的情况，并在各方面以更全面、更安全、更有力的方式应对紧急情况。（Craik, 1943）

继 1945 年克雷克死于自行车事故之后，唐纳德·布劳德本特（Donald Broadbent）继续从事这一工作。布劳德本特的 *Perception and Communication*（Broadbent, 1958）是最早将心理现象建模为信息处理的著作之一。与此同时的美国，计算机建模的发展导致了**认知科学**（cognitive science）领域的诞生。这个领域可以说是开始于 1956 年 9 月麻省理工学院的一次研讨会上，并且仅仅两个月后，人工智能本身就“诞生”了。

在研讨会上，乔治·米勒（George Miller）发表了“The Magic Number Seven”，诺姆·乔姆斯基（Noam Chomsky）发表了“Three Models of Language”，艾伦·纽厄尔和赫伯特·西蒙发表了“The Logic Theory Machine”。这 3 篇影响广泛的论文分别展示了如何使用计算机模型处理记忆、语言和逻辑思维的心理学问题。现在心理学家普遍认为“认知理论应该就像一个计算机程序”（Anderson, 1980），也就是说，认知理论应该从信息处理的角度来描述认知功能的运作。

为了综述目的，我们将**人机交互**（human-computer interaction，HCI）领域归于心理学下。人机交互的先驱之一道格·恩格巴特（Doug Engelbart）倡导**智能增强**（intelligence augmentation）的理念（IA 而非 AI）。他认为，计算机应该增强人类的能力，而不是完全自动化人类的任务。1968 年，在恩格巴特的“所有演示之母”（mother of all demos）上首次展示了计算机鼠标、窗口系统、超文本和视频会议，所有这些都是为了展示人类知识工作者可以通过某些智能增强来共同完成工作。

今天，我们更倾向于将 IA 和 AI 视为同一枚硬币的两面，前者强调人类控制，而后者强调机器的智能行为，都是机器有利于人类所必需的。

1.2.6 计算机工程

- 如何构建高效的计算机？

现代数字电子计算机是由陷入第二次世界大战中的 3 个国家的科学家们独立且几乎同时发明的。第一台可操作的计算机是由艾伦·图灵的团队于 1943 年建造的机电希思·罗宾逊（Heath Robinson[①]），它的唯一目的是破译德国的情报。1943 年，同一小组开发了 Colossus，

① 以一位英国漫画家的名字命名的复杂机器。这位漫画家描绘了一些古怪而又荒唐的复杂装置来完成日常任务，如给面包涂黄油。

这是一款基于真空管的强大通用机器。[①] 第一台可操作的可编程计算机是 Z-3，是德国工程师康拉德 • 楚泽（Konrad Zuse）在 1941 年发明的。楚泽还发明了浮点数和第一个高级编程语言 Plankalkül。第一台电子计算机 ABC 是约翰 • 阿塔纳索夫（John Atanasoff）和他的学生克利福德 • 贝里（Clifford Berry）在 1940 年至 1942 年间在爱荷华州立大学组装的。阿塔纳索夫的研究很少得到支持或认可，而 ENIAC 作为宾夕法尼亚大学秘密军事项目的一部分被证明是现代计算机最有影响力的先驱。ENIAC 的开发团队包括了约翰•莫奇利（John Mauchly）和约翰•普雷斯伯 • 埃克特（J. Presper Eckert）等工程师。

从那时起，每一代计算机硬件更新都带来了速度和容量的提升以及价格的下降，这是**摩尔定律**（Moore's law）所描述的趋势。直到 2005 年之前，大约每 18 个月 CPU 的性能就会翻一番，但功耗问题导致制造商开始增加 CPU 的核数而不是提高 CPU 的时钟频率。目前的预期是，未来性能的增加将来自于大量的并行性，这体现了与大脑特性奇妙的一致性。在应对不确定的世界时，基于这一理念设计硬件：不需要 64 位的数字精度，只需 16 位（如 `bfloat16` 格式）甚至 8 位就足够了，这可以使处理速度更快。

已经出现了一些针对人工智能应用进行调整的硬件，如图形处理单元（GPU）、张量处理单元（TPU）和晶圆级引擎（WSE）。从 20 世纪 60 年代到大约 2012 年，用于训练顶级机器学习应用的计算能力遵循了摩尔定律。从 2012 年开始，情况发生了变化：从 2012 年到 2018 年，这一数字增长了 30 万倍，每 100 天左右翻一番（Amodei and Hernandez, 2018）。在 2014 年花一整天训练的机器学习模型在 2018 年只需两分钟就可以训练完成（Ying *et al.*, 2018）。尽管**量子计算**（quantum computing）还不实用，但它有望为人工智能算法的一些重要子方向提供更显著的加速。

毋庸置疑，在电子计算机出现之前计算设备就已经存在了。最早的自动化机器可追溯到 17 世纪（见 1.2.1 节的讨论）。第一台可编程机器是由约瑟夫 • 玛丽 • 雅卡尔（Joseph Marie Jacquard，1752—1834）于 1805 年发明的提花织布机，它使用打孔卡片来存储编织图案的指令。

19 世纪中期，查尔斯 • 巴贝奇（Charles Babbage，1792—1871）设计了两台计算机，但都没有完成。差分机的目的是为工程和科学项目计算数学表。它最终于 1991 年建成并投入使用（Swade, 2000）。巴贝奇的分析机更有雄心：它包括可寻址内存、基于雅卡尔打孔卡的存储程序以及有条件的跳转。这是第一台能够进行通用计算的机器。

巴贝奇的同事埃达 • 洛芙莱斯（Ada Lovelace，诗人拜伦勋爵的女儿）理解了计算机的潜力，将其描述为“一种能思考或者……能推理的机器”，能够对“宇宙中所有事物”进行推理（Lovelace, 1843）。她还预测到了人工智能的技术成熟度曲线，并提出：“我们最好防范可能夸大分析机能力的想法。”遗憾的是，巴贝奇的机器和洛芙莱斯的思想已基本被遗忘了。

人工智能还得益于计算机科学软件方面的发展，后者提供了编写现代程序所需的操作系统、编程语言和工具（以及有关它们的论文）。而这也是人工智能对其有回馈的领域：人工智能工作开创的许多想法正重归主流计算机科学，包括分时、交互式解释器、使用窗口和鼠标的个人计算机、快速开发环境、链表数据类型、自动存储管理，以及符号式编程、函数式编程、说明性编程和面向对象编程的关键概念。

1.2.7　控制理论与控制论

- 人造物如何在它们自己的控制下运行？

① 在第二次世界大战后，图灵想把这些计算机用于人工智能研究，例如，他创建了第一个国际象棋程序的框架（Turing *et al.*, 1953），但英国政府阻止了这项研究。

居住在亚历山大城的古希腊工程师克特西比乌斯（Ktesibios，约公元前 250 年）建造了第一个自我控制的机器：一台水钟，其特点是拥有一个可以保持恒定水流速度的调节器。这一发明改变了人造物可以做什么的定义。在此之前，只有生物才能根据环境的变化来改变自己的行为。其他自调节反馈控制系统的示例工作包括由詹姆斯·瓦特（James Watt，1736—1918）创建的蒸汽机调节器以及科内利斯·德雷贝尔（Cornelis Drebbel，1572—1633，潜艇发明者）发明的恒温器。詹姆斯·克拉克·麦克斯韦（James Clerk Maxwell）（Maxwell, 1868）开创了控制系统的数学理论。

第二次世界大战后，**控制理论**（control theory）发展的核心人物是诺伯特·维纳（Norbert Wiener，1894—1964）。维纳是一位杰出的数学家，在对生物和机械控制系统及其与认知的联系产生兴趣之前，曾与伯特兰·罗素等人合作。像克雷克（把控制系统作为心理模型）一样，维纳和他的同事阿图罗·罗森布鲁斯（Arturo Rosenblueth）以及朱利安·毕格罗（Julian Bigelow）挑战了行为主义正统派（Rosenblueth *et al.*, 1943）。他们认为具有目的的行为源于试图最小化“错误”的调节机制，即当前状态和目标状态之间的差异。20 世纪 40 年代后期，维纳与沃伦·麦卡洛克（Warren McCulloch）、沃尔特·皮茨（Walter Pitts）和约翰·冯·诺伊曼一起组织了一系列有影响力的会议，探索关于认知的新数学和计算模型。维纳的《控制论》（*Cybernetics*）（Wiener, 1948）成为畅销书，使大众意识到了人工智能机器的可能性。

与此同时，英国控制论专家罗斯·艾什比（W. Ross Ashby）开创了类似的思想（Ashby, 1940）。艾什比、图灵、沃尔特和其他一些学者为“那些在维纳的书出现之前就有维纳想法的人”组织了推理俱乐部[①]。艾什比在《大脑设计》（*Design for a Brain*）（Ashby, 1948, 1952）一书中详细阐述了他的想法，即可以通过**自我平衡**（homeostatic）设备来实现智能，该设备使用恰当的反馈回路来实现稳定的自适应行为。

现代控制理论，特别是被称为随机最优控制的分支，其目标是设计随时间最小化**代价函数**（cost function）的系统。这与人工智能的标准模型——设计性能最优的系统大致相符。尽管人工智能和控制理论的创始人之间有着密切的联系，为什么它们却是两个不同的领域呢？答案在于参与者所熟悉的数学技术与每种世界观所包含的对应问题是紧密结合的。微积分和矩阵代数是控制理论的工具，它们适用于固定的连续变量集描述的系统，而人工智能的建立在一定程度上是为了避开这些可感知的局限性。逻辑推理和计算工具使人工智能研究人员能够考虑语言、视觉和符号规划等问题，而这些问题完全超出了控制理论家的研究范围。

1.2.8 语言学

- 语言是如何与思维联系的？

1957 年，斯金纳（B. F. Skinner）发表了 *Verbal Behavior*，包含该领域最著名的专家对语言学习的行为主义方法的全面详细的描述。但奇怪的是，一篇对这本书的评述也像这本书一样广为人知，几乎扼杀了大众对行为主义的兴趣。评述的作者是语言学家诺姆·乔姆斯基，彼时他刚刚出版了一本关于他自己理论的书《句法结构》（*Syntactic Structure*）。乔姆斯基指出，行为主义理论并没有解决语言创造力的概念，它没有解释孩子们如何理解并造出他们从未听过的句子。乔姆斯基以句法模型为基础的理论可以追溯到古印度语言学家波你尼（Panini，约公元前 350 年）。该理论可以解释语言创造力，而且与以前的理论不同，它足够形式化，原则上可以被程序化。

现代语言学和人工智能几乎同时“诞生”，并一起成长，交叉于一个称为**计算语言学**（computational linguistics）或**自然语言处理**（natural language processing）的混合领域。相比 1957

① 推理俱乐部（Ratio Club）。Ratio 取自推理演算器（calculus ratiocinator），因此此处翻译为“推理俱乐部”。——编者注

年，理解语言复杂了许多。理解语言需要理解主题和上下文，而不仅仅是理解句子结构。这似乎是显而易见的，但直到 20 世纪 60 年代才得到广泛认可。**知识表示**（knowledge representation）（关于如何将知识转化为计算机可以推理的形式的研究）的大部分早期工作与语言相关联，并受到语言学研究的启发，而语言学研究反过来又与数十年的语言哲学分析工作有关联。

1.3　人工智能的历史

总结人工智能历史里程碑的快速方法是列出图灵奖得主：马文·明斯基（Marvin Minsky）（1969 年图灵奖得主）和约翰·麦卡锡（John McCarthy）（1971 年图灵奖得主）定义了基于表示和推理的领域基础；艾伦·纽厄尔（Allen Newell）和赫伯特·西蒙（Herbert Simon）（1975 年图灵奖得主）提出了关于问题求解和人类认知的符号模型；爱德华·费根鲍姆（Ed Feigenbaum）和劳伊·雷迪（Raj Reddy）（1994 年图灵奖得主）开发了通过对人类知识编码来解决真实世界问题的专家系统；朱迪亚·珀尔（Judea Pearl）（2011 年图灵奖得主）提出了通过原则性的方式处理不确定性的概率因果推理技术；最近的是约书亚·本吉奥（Yoshua Bengio）、杰弗里·辛顿（Geoffrey Hinton）和杨立昆（Yann LeCun）（2018 年图灵奖得主）[①]，他们将“深度学习”（多层神经网络）作为现代计算的关键部分。本节的其余部分将更详细地介绍人工智能历史的每个阶段。

1.3.1　人工智能的诞生（1943—1956）

现在普遍认为由沃伦·麦卡洛克和沃尔特·皮茨（McCulloch and Pitts, 1943）完成的工作是人工智能的第一项研究工作。他们受到皮茨的顾问尼古拉斯·拉舍夫斯基（Nicolas Rashevsky）（1936, 1938）对数学建模工作的启发，选择了 3 方面的资源构建模型：基础生理学知识和大脑神经元的功能，罗素和怀特海（Whitehead）对命题逻辑的形式化分析，以及图灵的计算理论。他们提出了一种人工神经元模型，其中每个神经元的特征是“开”或“关”，并且会因足够数量的相邻神经元受到刺激而切换为“开”。神经元的状态被认为是“事实上等同于提出其充分激活的命题”。例如，他们证明任何可计算的函数都可以通过一些神经元互相连接的网络来计算，以及所有的逻辑联结词（AND、OR、NOT 等）都可以通过简单的网络结构来实现。麦卡洛克和皮茨还表明适当定义的网络可以学习。唐纳德·赫布（Donald Hebb）（Hebb, 1949）示范了用于修改神经元之间连接强度的简单更新规则。他的规则，现在称为**赫布型学习**（Hebbian learning），至今仍是一种有影响力的模式。

哈佛大学的两名本科生马文·明斯基（Marvin Minsky，1927—2016）和迪安·埃德蒙兹（Dean Edmonds）在 1950 年建造了第一台神经网络计算机——SNARC。SNARC 使用了 3000 个真空管和 B-24 轰炸机上一个多余的自动驾驶装置来模拟由 40 个神经元组成的网络。后来，明斯基在普林斯顿大学研究了神经网络中的通用计算。他的博士学位委员会对这类工作是否应该被视为数学持怀疑态度，但据说冯·诺伊曼评价：“如果现在还不能被视为数学，总有一天会的。”

还有许多早期工作可以被描述为人工智能，包括 1952 年由曼彻斯特大学的克里斯托弗·斯特雷奇（Christopher Strachey）和 IBM 公司的亚瑟·塞缪尔（Arthur Samuel）分别独立开发的西洋跳棋程序。然而，还是图灵的观点最有影响力。早在 1947 年，他就在伦敦数学协会

① 此书英文原著将约书亚·本吉奥、杰弗里·辛顿和杨立昆记录为获得了 2019 年图灵奖，他们实则获得的是 2018 年图灵奖。——编者注

（London Mathematical Society）就这一主题发表了演讲，并在其 1950 年的文章“Computing Machinery and Intelligence”中阐明了有说服力的议程。在论文中，他介绍了图灵测试、机器学习、遗传算法和强化学习。如第 27 章所述，也回答了许多针对人工智能的质疑。他还认为，通过开发学习算法然后教会机器，而不是手工编写智能程序，将更容易创造出人类水平的人工智能。他在随后的演讲中警告说，实现这一目标对人类来说可能不是最好的事情。

1955 年，达特茅斯学院的约翰·麦卡锡说服明斯基、克劳德·香农（Claude Shannon）和纳撒尼尔·罗切斯特（Nathaniel Rochester）帮助他召集对自动机理论、神经网络和智能研究感兴趣的美国研究人员。他们于 1956 年夏天在达特茅斯组织了为期两个月的研讨会。这场研讨会共有 10 位与会者，其中包括来自卡内基理工学院[①]的艾伦·纽厄尔和赫伯特·西蒙、普林斯顿大学的特伦查德·摩尔（Trenchard More）、IBM 的亚瑟·塞缪尔以及来自麻省理工学院的雷·所罗门诺夫（Ray Solomonoff）和奥利弗·赛弗里奇（Oliver Selfridge）。该提案指出：[②]

> 1956 年夏天，我们提议在新罕布什尔州汉诺威的达特茅斯学院进行为期两个月共 10 人参与的人工智能研讨。这次研讨是基于这样的假设：理论上可以精确描述学习的每个方面或智能的任何特征，从而可以制造机器来对其进行模拟。我们将试图寻找让机器使用语言，形成抽象和概念，解决人类特有的各种问题并改进自身的方法。我们认为，如果一个精心挑选的科学家团队在一整个夏天里共同研究这些问题，则可以在一个或多个方面取得重大进展。

尽管有这种乐观的预测，但达特茅斯的研讨会并没有带来任何突破。纽厄尔和西蒙提出了也许是最成熟的工作——一个称为“逻辑理论家”（Logic Theorist，LT）的数学定理证明系统。西蒙声称：“我们已经发明了一种能够进行非数值思维的计算机程序，从而解决了神圣的身心问题。”[③] 研讨会结束后不久，这个程序就已经能证明罗素和怀特海的 *Principia Mathematica* 第 2 章中的大多数定理。据报道，当罗素被告知 LT 提出了一个比 *Principia Mathematica* 书中更精巧的证明时，罗素感到很高兴。但《符号逻辑杂志》（*The Journal of Symbolic Logic*）的编辑们没被打动，他们拒绝了由纽厄尔、西蒙和逻辑理论家合著的论文。

1.3.2 早期热情高涨，期望无限（1952—1969）

20 世纪 50 年代的知识界总体上倾向于相信“机器永远不能做 *X*。”（见第 27 章中图灵收集的 *X* 的详细列表。）人工智能研究人员自然而然地一个接一个地演示 *X* 以回应。他们特别关注那些被认为能够显示人类智能的任务，包括游戏、谜题、数学和智商测试。约翰·麦卡锡将这段时期称为“瞧，妈，不需要人动手操控！”（Look，Ma，no hands!）时代。

纽厄尔和西蒙继 LT 成功之后又推出了通用问题求解器，即 GPS。与 LT 不同，GPS 从一开始就被设计为模仿人类求解问题的协议。结果表明，在它可以处理的有限类型的难题中，该程序考虑的子目标和可能采取的行为的顺序与人类处理相同问题的顺序类似。因此，GPS 可能是第一个体现“人类思维”方式的程序。作为认知模型，GPS 和后续程序的成功使得纽厄尔和西蒙（1976）提出了著名的**物理符号系统**（physical symbol system）假说，该假说认为“物理

① 现在是卡内基梅隆大学（CMU）。

② 这是麦卡锡的术语“人工智能”被第一次正式使用。也许“计算理性”会更精确、威胁更小，但“人工智能”一直存在。在达特茅斯会议 50 周年纪念会上，麦卡锡表示，他反对使用“计算机”或“可计算”等术语，以表达对诺伯特·维纳的敬意，因为维纳倡导模拟控制设备，而不是数字计算机。

③ 纽厄尔和西蒙还发明了一种链表处理语言 IPL 来编写 LT。他们没有编译器，只能手动将其翻译为机器代码。为了避免错误，他们并行工作，在编写每条指令时相互大声喊出二进制数，以确保他们是一致的。

符号系统具有进行一般智能动作的必要和充分方法”。意思是，任何显示出智能的系统（人类或机器）必须通过操作由符号组成的数据结构来运行。之后我们会看到这个假说已经受到了多方面的挑战。

在 IBM，纳撒尼尔·罗切斯特和他的同事开发了首批人工智能程序。赫伯特·盖伦特（Herbert Gelernter）（Gelernter, 1959）构造了几何定理证明程序（Geometry Theorem Prover），它能够证明许多数学学生认为相当棘手的定理。这项工作是现代数学定理证明程序的先驱。

从长远来看，这一时期所有探索性工作中，最有影响力的可能是亚瑟·萨缪尔对西洋跳棋的研究。通过使用现在称之为强化学习的方法（见第 22 章），萨缪尔的程序可以以业余高手的水平进行对抗。因此，他驳斥了计算机只能执行被告知的事情的观点：他的程序很快学会了玩游戏，甚至比其创造者玩得更好。该程序于 1956 年在电视上演示，给人留下了深刻的印象。和图灵一样，萨缪尔也很难找到使用计算机的机会，他只能晚上工作，使用仍在 IBM 制造工厂测试场地上还未出厂的计算机。萨缪尔的程序是许多后继系统的前身，如 TD-Gammon（Tesauro, 1992）和 AlphaGo（Silver *et al.*, 2016）。TD-Gammon 是世界上最好的西洋双陆棋棋手之一，而 AlphaGo 因击败人类世界围棋冠军而震惊世界（见第 5 章）。

1958 年，约翰·麦卡锡为人工智能做出了两项重要贡献。在麻省理工学院人工智能实验室备忘录 1 号中，他定义了高级语言 **Lisp**，Lisp 在接下来的 30 年中成为了最重要的人工智能编程语言。在一篇题为“Programs with Common Sense”的论文中，麦卡锡为基于知识和推理的人工智能系统提出了概念性议案。这篇论文描述了“建议接受者”（Advice Taker），这是一个假想程序，它包含了世界的一般知识，并可以利用它得出行动规划。这个概念可以用简单的逻辑公理来说明，这些逻辑公理足以生成一个开车去机场的规划。该程序还被设计为能在正常运行过程中接受新的公理，从而实现无须重新编程就能够在新领域中运行。因此，“建议接受者”体现了知识表示和推理的核心原则：对世界及其运作进行形式化、明确的表示，并且通过演绎来操作这种表示是很有用的。这篇论文影响了人工智能的发展历程，至今仍有意义。

1958 年也是马文·明斯基转到麻省理工学院的一年。然而，他与麦卡锡的最初合作并没有持续。麦卡锡强调形式逻辑中的表示和推理，而明斯基则对程序工作并最终形成反逻辑的观点更感兴趣。1963 年，麦卡锡在斯坦福大学建立了人工智能实验室。1965 年亚伯拉罕·鲁滨逊（J. A. Robinson）归结原理（一阶逻辑的完备定理证明算法；见第 9 章）的发现推进了麦卡锡使用逻辑来构建最终“建议接受者”的计划。麦卡锡在斯坦福大学的工作中强调了逻辑推理的通用方法。逻辑的应用包括柯德尔·格林（Cordell Green）的问答和规划系统（Green, 1969b）以及斯坦福研究所（SRI）的 Shakey 机器人项目，后者（将在第 26 章中进一步讨论）是第一个展示逻辑推理和物理活动完全集成的项目。

在麻省理工学院，明斯基指导了一批学生，他们选择了一些似乎需要智能才能求解的有限问题。这些有限的领域被称为**微世界**（microworld）。詹姆斯·斯莱格尔（James Slagle）的 Saint 程序（Slagle, 1963）能够求解大学一年级课程中典型封闭形式的微积分问题。托马斯·埃文斯（Thomas Evans）的 Analogy 程序（Evans, 1968）能够解决智商测试中常见的几何类比问题。丹尼尔·博布罗（Daniel Bobrow）的 Student 项目（Bobrow, 1967）能够求解代数故事问题，例如：

> 如果汤姆获得的客户数量是他投放的广告数量的 20% 的平方的两倍，已知他投放的广告数量是 45，那么汤姆获得的客户数量是多少？

最著名的微世界是**积木世界**（blocks world），由一组放置在桌面上的实心积木组成（或者

更常见的是模拟桌面），如图 1-3 所示。在这个世界中，一个典型的任务是用机械手以某种方式重新排列积木，这个机械手一次可以拿起一块积木。积木世界孕育了戴维·哈夫曼（David Huffman）（Huffman, 1971）的视觉项目、戴维·沃尔茨（David Waltz）（Waltz, 1975）的视觉和约束传播工作、帕特里克·温斯顿（Patrick Winston）（Winston, 1970）的学习理论、特里·温诺格拉德（Terry Winograd）（Winograd, 1972）的自然语言理解程序以及斯科特·法尔曼（Scott Fahlman）（Fahlman, 1974）的规划器。

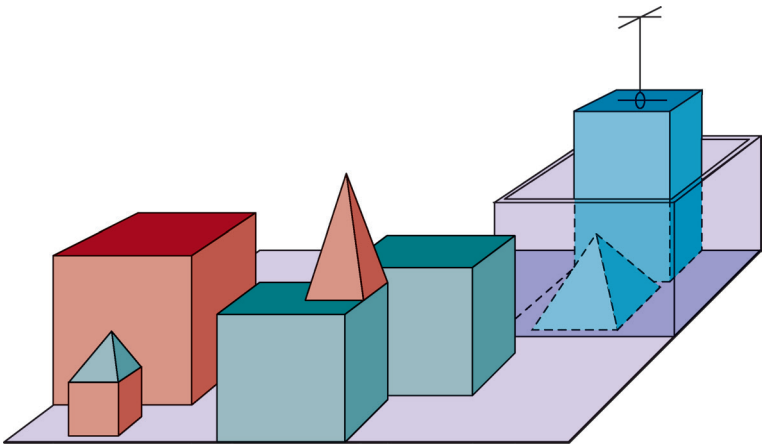

图 1-3 积木世界的场景。SHRDLU（Winograd, 1972）刚刚完成了一个命令——“找到一块比你所持有的积木块更高的积木块，并把它放进盒子里”

建立在麦卡洛克和皮茨提出的神经网络上的早期工作也蓬勃发展。什穆埃尔·温诺格拉德（Shmuel Winograd）和杰克·考恩（Jack Cowan）的研究（Winograd and Cowan, 1963）展示了大量元素如何共同代表一个独立的概念，同时提升稳健性和并行性。赫布的学习方法分别得到了伯尼·维德罗（Bernie Widrow）（Widrow and Hoff, 1960; Widrow, 1962）和弗兰克·罗森布拉特（Frank Rosenblatt）（Rosenblatt, 1962）的改进，他们的网络分别被称为**线性自适应神经网络**（adaline）和**感知机**（perceptron）。**感知机收敛定理**（perceptron convergence theorem）（Block *et al.*, 1962）指出，学习算法可以调整感知机的连接强度来拟合任何输入数据（前提是存在这样的拟合）。

1.3.3 一些现实（1966—1973）

从一开始，人工智能研究人员对未来成功的预测毫不避讳。下面这句 1957 年赫伯特·西蒙的名言经常被引用：

> 我的目的不是使大家感到惊讶或震惊，我可以总结出的最简单的说法是，现在世界上存在着能够思考、学习和创造的机器。此外，它们的这些能力将迅速提高，在可见的未来内，它们能够处理的问题范围将与人类思维的应用范围一样广泛。

虽然“可见的未来”这个词是模糊的，但西蒙也做出了更具体的预测：10 年内，计算机将成为国际象棋冠军以及机器将能证明重要的数学定理。实际上，这些预测的实现（或近似实现）用了 40 年时间，远远超过 10 年。当初西蒙的过度自信来自于早期人工智能系统在简单示例任务上的出色表现。但是，在几乎所有情况下，这些早期系统在更困难的问题上都失败了。

失败主要有两个原因。首先是许多早期人工智能系统主要基于人类如何执行任务的“知情内省型”，而不是基于对任务、解的含义以及算法需要做什么才能可靠地产生解的仔细分析。

失败的第二个原因是对人工智能要求解的问题的复杂性缺乏认识。大多数早期的问题求解系统都会尝试组合不同的步骤，直到找到解为止。这一策略最初奏效是因为微世界所包含的对象非常少，因此可能的动作非常少，解的动作序列也非常短。在计算复杂性理论发展完备之前，人们普遍认为“扩展”到更大的问题仅仅是需要更快的硬件和更大的内存。但是当研究人员无法证明涉及几十个事实的定理时，伴随着归结定理证明发展而来的乐观情绪很快就受到了打击。一般而言，程序可以找到解的事实并不意味着该程序具备任何在实践中找到解所需的机制。

无限计算能力的幻想并不局限于求解问题的程序。早期的**机器进化**（machine evolution）［现在称为**遗传编程**（genetic programming）］实验（Friedberg, 1958; Friedberg *et al.*, 1959）基于绝对正确的信念，即通过对机器代码程序进行一系列适当的小变异，就可以为任何特定任务生成表现良好的程序。这个想法就是通过选择过程来尝试随机突变，并保留似乎有用的突变。尽管使用了长达数千小时的 CPU 时间，但几乎没有任何进展。

未能处理“组合爆炸”是莱特希尔报告（Lighthill, 1973）中对人工智能的主要批评之一，基于这份报告，英国政府决定在除两所大学外的所有大学中停止支持人工智能研究。（口述传说描绘了一幅稍有不同、更加丰富多彩的画面，但带有政治野心和个人好恶的描述都不是本书的话题。）

第三个困难是产生智能行为的基本结构上存在一些基本限制。例如，明斯基和派珀特的著作 *Perceptrons*（Minsky and Papert, 1969）证明，尽管感知机（一种简单的神经网络形式）被证明可以学习它们能够表示的任何事物，但它们能表示的事物很少。举例来说，我们无法训练双输入感知机来判断它的两个输入是否相同。尽管他们的研究结果并不适用于更复杂的多层网络，但用于神经网络研究的经费很快就减少到几乎为零。讽刺的是，在 20 世纪 80 年代和 21 世纪 10 年代再次引起神经网络研究巨大复兴的新反向传播学习算法，早在 20 世纪 60 年代初已经在其他情景下得到了发展（Kelley, 1960; Bryson, 1962）。

1.3.4　专家系统（1969—1986）

在人工智能研究的前十年提出的问题求解是一种通用搜索机制，试图将基本的推理步骤串在一起，找到完整的解。这种方法被称为**弱方法**（weak method），这种方法虽然很普适，但它不能扩展到大型或困难的问题实例上。弱方法的替代方案是使用更强大的领域特定的知识，这些知识允许更大规模的推理步骤，并且可以更轻松地处理特定专业领域中发生的典型案例。有人可能会说，必须已经差不多知道答案才能解决一个难题。

DENDRAL 程序（Buchanan *et al.*, 1969）是这种方法的早期例子。它是在斯坦福大学开发的，爱德华·费根鲍姆（曾是赫伯特·西蒙的学生）、布鲁斯·布坎南（Bruce Buchanan，从哲学家转行的计算机科学家）和乔舒亚·莱德伯格（Joshua Lederberg，诺贝尔生理学或医学奖得主，遗传学家）联手解决了从质谱仪提供的信息推断分子结构的问题。该程序的输入包括分子的基本分子式（如 $C_6H_{13}NO_2$）和质谱，其中质谱给出了分子被电子束轰击时产生的各种碎片的质量。例如，质谱可能在 $m = 15$ 处有一个峰，这对应于甲基（CH_3）碎片的质量。

朴素版本的程序生成所有可能的符合分子式的结构，然后预测每个结构在质谱仪中的观测结果，并将其与实际质谱进行比较。正如人们所预期的，这对中等规模的分子来说也是难以处理的。DENDRAL 的研究人员咨询了分析化学家，并发现他们通过寻找质谱中已知的峰模式来工作，

这些峰表明分子中的常见子结构。例如，以下规则用于识别酮（C=O）结构（分子量 28）：

如果 M 是整个分子的质量，且在 x_1 和 x_2 处有两个峰，并且

（a）$x_1 + x_2 = M + 28$；（b）$x_1 - 28$ 是一个高峰；（c）$x_2 - 28$ 是一个高峰；（d）x_1 和 x_2 中至少有一处是高峰，

则该分子含有酮基。

认识到分子包含特定的子结构，可以极大地减少可能候选项的量级。据作者称，DENDRAL 之所以强大，是因为它不是以第一性原理的形式，而是以高效“食谱”的形式体现了质谱的相关知识（Feigenbaum *et al.*, 1971）。DENDRAL的意义在于它是第一个成功的知识密集型系统：它的专业知识来源于大量专用规则。1971 年，费根鲍姆和斯坦福大学的其他研究人员开启了启发式编程项目（heuristic programming project，HPP），以此来研究**专家系统**（expert system）的新方法可以在多大程度上应用到其他领域。

接下来的一个主要工作是用于诊断血液感染的 MYCIN 系统。MYCIN 有大约 450 条规则，它能够表现得和一些专家一样好，甚至比初级医生要好得多。MYCIN 与 DENDRAL 有两个主要区别。首先，不像 DENDRAL 规则，不存在可以推导出 MYCIN 规则的一般理论模型，MYCIN 规则不得不从大量的专家访谈中获得。其次，规则必须反映与医学知识相关的不确定性。MYCIN 引入了一种称为**确定性因子**（certainty factor）的不确定性计算（见第 13 章），这在当时似乎与医生评估证据对诊断影响的方式非常吻合。

第一个成功的商用专家系统 R1 在数字设备公司（Digital Equipment Corporation，DEC）投入使用（McDermott, 1982），该程序帮助公司配置新计算机系统的订单。截至 1986 年，它每年为公司节省约 4000 万美元。到 1988 年，DEC 的人工智能小组已经部署了 40 个专家系统，而且还有更多的专家系统在开发中。同时期，杜邦公司有 100 个专家系统在使用，500 个在开发。当时几乎每家美国大公司都有自己的人工智能团队，不是在使用专家系统，就是在研究专家系统。

领域知识的重要性在自然语言理解领域也很突出。尽管特里·温诺格拉德的 SHRDLU 系统取得了成功，但它的方法并没有扩展到更一般的任务：对于歧义消解之类的问题，它使用了依赖于积木世界中微小范围的简单规则。

包括麻省理工学院的尤金·查尔尼克（Eugene Charniak）和耶鲁大学的罗杰·尚克（Roger Schank）在内的几位研究人员一致认为，强大的语言理解需要关于世界的一般知识以及使用这些知识的一般方法。（尚克进一步声称，“根本就没有语法这回事”，这让很多语言学家感到不安，但确实引发了一场有益的讨论。）尚克和他的学生们建立了一系列的程序（Schank and Abelson, 1977; Wilensky, 1978; Schank and Riesbeck, 1981），这些程序都用于理解自然语言。但是，重点不在于语言本身，而在于用语言理解所需的知识来表示和推理问题。

在真实世界中的广泛应用引发了表示和推理工具的广泛发展。有些是基于逻辑的，例如，Prolog 语言在欧洲和日本流行，而 PLANNER 家族在美国流行。其他人则遵循明斯基的**框架**（frame）思想（Minsky, 1975），采用了一种更结构化的方法，将有关特定对象和事件类型的事实组合起来，并将这些类型组织成类似于生物分类法的大型分类层次结构。

1981 年，日本政府宣布了“第五代计算机”计划，这是一个十年计划，旨在建造运行 Prolog 的大规模并行智能计算机。按现在的货币系统衡量，预算将超过 13 亿美元。作为回应，美国成立了微电子与计算机技术公司（Microelectronics and Computer Technology Corporation，MCC），这是一个旨在确保国家竞争力的联盟。在这两个项目中，人工智能都是广泛努力的一部分，包括芯片设计和人机界面研究。在英国，阿尔维（Alvey）报告恢复了被莱特希尔报告取消的

资助资金。然而，这些项目都没有在新型的人工智能能力或经济影响方面下实现其宏伟目标。

总的来说，人工智能行业从 1980 年的几百万美元增长到 1988 年的数十亿美元，还产生了数百家构建专家系统、视觉系统、机器人以及专门服务于这些目的的软硬件的公司。

但此后不久，经历了一段被称为“人工智能冬天”的时期，许多公司因未能兑现夸张的承诺而停滞。事实证明，为复杂领域构建和维护专家系统是困难的，一部分原因是系统使用的推理方法在面临不确定性时会崩溃，另一部分原因是系统无法从经验中学习。

1.3.5　神经网络的回归（1986—现在）

在 20 世纪 80 年代中期，至少有 4 个不同的团队重新发明了最早在 20 世纪 60 年代初期发展起来的**反向传播**（back-propagation）学习算法。该算法被应用于计算机科学和心理学中的许多学习问题，*Parallel Distributed Processing* 合集（Rumelhart and McClelland, 1986）中的结果的广泛传播引起了极大的轰动。

这些所谓的**联结主义**（connectionist）模型被一些人视为纽厄尔和西蒙的符号模型以及麦卡锡和其他人的逻辑主义方法的直接竞争对手。人类在某种程度上操纵符号似乎是显而易见的——事实上，人类学家特伦斯·迪肯（Terrence Deacon）在其著作《符号化动物》（*The Symbolic Species*）（Deacon, 1997）中指出，这是人类的决定性特征。与此相反，20 世纪 80 年代和 21 世纪 10 年代神经网络复兴的领军人物杰弗里·辛顿将符号描述为“人工智能的光以太”（19 世纪许多物理学家认为电磁波传播的介质是光以太，但其实这种介质不存在）。事实上，我们在语言中命名的许多概念，经过仔细检查后，都未能获得早期人工智能研究人员希望以公理形式描述逻辑定义的充要条件。联结主义模型可能以一种更流畅和不精确的方式形成内部概念，更适配真实世界的混乱。它们还具备从样本中学习的能力，它们可以将它们的预测输出值与问题的真实值进行比较，并修改参数以减少差异，使它们在未来的样本中更有可能表现良好。

1.3.6　概率推理和机器学习（1987—现在）

专家系统的脆弱性导致了一种新的、更科学的方法，结合了概率而不是布尔逻辑，基于机器学习而不是手工编码，重视实验结果而不是哲学主张。[①] 现在更普遍的是，基于现有理论而不是提出全新的理论，基于严格的定理或可靠的实验方法（Cohen, 1995）而不是基于直觉的主张，以及展示与真实世界应用的相关性而不是虚拟的示例。

共享的基准问题集成为了展示进度的标准，包括加州大学欧文分校的机器学习数据集库、用于规划算法的国际规划竞赛、用于语音识别的 LibriSpeech 语料库、用于手写数字识别的 MNIST 数据集、用于图像物体识别的 ImageNet 和 COCO、用于自然语言问答的 SQuAD、机器翻译的 WMT 竞赛以及布尔可满足性求解器国际 SAT 竞赛。

人工智能的创立在一定程度上是对控制理论和统计等现有领域局限性的反抗，但在这一时期，它吸纳了这些领域的积极成果。正如戴维·麦卡莱斯特（David McAllester）（McAllester, 1998）所说：

> 在人工智能早期，符号计算的新形式（例如框架和语义网络）使大部分经典理

① 一些人将这种变化描述为整洁派（neat，认为人工智能理论应该以数学的严谨性为基础的人）战胜了邋遢派（scruffy，那些宁愿尝试大量的想法，编写一些程序，然后评估哪些似乎可行的人）。这两种方法都很重要。向整洁派的转变意味着该领域已经达到了稳定和成熟的水平。目前对深度学习的重视可能代表着邋遢派的复兴。

> 论过时，这似乎是合理的。这导致了一种孤立主义，即人工智能在很大程度上与计算机科学的其他领域分离。这种孤立主义目前正在被摒弃。人们认识到，机器学习不应该独立于信息论，不确定推理不应该独立于随机建模，搜索不应该独立于经典优化和控制，自动推理不应该独立于形式化方法和静态分析。

语音识别领域对这种模式进行了说明。20 世纪 70 年代，研究人员尝试了各种不同的架构和方法，许多是相当暂时和脆弱的，并且只能处理几个精心挑选的例子。在 20 世纪 80 年代，使用**隐马尔可夫模型**（hidden Markov model，HMM）的方法开始主导这一领域。HMM 有两个相关的方面。首先，它们基于严格的数学理论。这使得语音研究人员能够在其他领域数十年数学成果的基础上进行开发。其次，它们是在大量真实语音数据的语料库上训练而产生的。这确保了健壮性，并且在严格的盲测中，HMM 的分数稳步提高。因此，语音技术和手写体字符识别的相关领域向广泛的工业和消费级应用过渡。注意，并没有科学证据表明人类使用 HMM 识别语音，HMM 只是为理解和求解问题提供了一个数学框架。然而，在 1.3.8 节中我们将看到，深度学习已经破坏了这种舒适的叙述。

1988 年是人工智能与统计学、运筹学、决策论和控制理论等其他领域相联系的重要一年。朱迪亚·珀尔的 *Probabilistic Reasoning in Intelligent Systems*（Pearl, 1988）使概率和决策论在人工智能中得到了新的认可。珀尔对贝叶斯网络的发展产生了一种用于表示不确定的知识的严格而有效的形式体系，以及用于概率推理的实用算法。第 12 ～ 16 章涵盖了这个领域，此外最近的发展大大提升了概率形式体系的表达能力，第 20 章描述了从数据中学习**贝叶斯网络**（Bayesian network）和相关模型的方法。

1988 年的第二个主要贡献是理查德·萨顿（Rich Sutton）的工作，他将强化学习（20 世纪 50 年代被用于亚瑟·塞缪尔的西洋跳棋程序中）与运筹学领域开发的马尔可夫决策过程（Markov decision processe，MDP）联系起来。随后，大量工作将人工智能规划研究与 MDP 联系起来，强化学习领域在机器人和过程控制方面找到了应用，并获得了深厚的理论基础。

人工智能对数据、统计建模、优化和机器学习的新认识带来的结果是，计算机视觉、机器人技术、语音识别、多智能体系统和自然语言处理等子领域逐渐统一，此前这些子领域在某种程度上已经脱离了核心人工智能。重新统一的过程在应用方面（例如，在此期间实用机器人的部署大大扩展）和关于人工智能核心问题更好的理论理解方面都产生了显著的效用。

1.3.7 大数据（2001—现在）

计算能力的显著进步和互联网的创建促进了巨大数据集的创建，这种现象有时被称为**大数据**（big data）。这些数据集包括数万亿字的文本、数十亿的图像、数十亿小时的语音和视频，以及海量的基因组数据、车辆跟踪数据、点击流数据、社交网络数据等。

这导致了专为利用非常大的数据集而设计的学习算法的开发。通常，这类数据集中的绝大多数例子都没有标签。例如，在雅让斯基关于词义消歧的著作（Yarowsky, 1995）中，出现的一个词（如“plant”），并没有在数据集中标明这是指植物还是工厂。然而，如果有足够大的数据集，合适的学习算法在识别句意的任务上可以达到超过 96% 的准确率。此外，班科和布里尔认为，将数据集的规模增加两到三个数量级所获得的性能提升会超过调整算法带来的性能提升（Banko and Brill, 2001）。

类似的现象似乎也发生在计算机视觉任务中，例如填补照片中的破洞（要么是由损坏造成的，要么是挖除前朋友造成的）。海斯和埃弗罗斯（Hays and Efros, 2007）开发了一种巧妙的

方法，从类似的图像中混合像素。他们发现，该技术在仅包含数千幅图像的数据库中效果不佳，但在拥有数百万幅图像的数据库中，该技术超过了质量阈值。不久之后，ImageNet 数据库（Deng *et al.*, 2009）中可用的数千万幅图像引发了计算机视觉领域的一场革命。

大数据的可用性和向机器学习的转变帮助人工智能恢复了商业吸引力（Havenstein, 2005; Halevy *et al.*, 2009）。大数据是 2011 年 IBM 的 Watson 系统在《危险边缘》（*Jeopardy!*）问答游戏中战胜人类冠军的关键因素，这一事件深深影响了公众对人工智能的看法。

1.3.8　深度学习（2011—现在）

深度学习（deep learning）是指使用多层简单的、可调整的计算单元的机器学习。早在 20 世纪 70 年代，研究人员就对这类网络进行了实验，并在 20 世纪 90 年代以**卷积神经网络**（convolutional neural network）（LeCun *et al.*, 1995）的形式在手写数字识别方面取得了一定的成功。然而，直到 2011 年，深度学习方法才真正开始流行起来。首先是在语音识别领域，然后是视觉物体识别领域。

在 2012 年的 ImageNet 竞赛中，需要将图像分类为 1000 个类别之一（犰狳、书架、开瓶器等）。多伦多大学杰弗里 • 辛顿团队开发的深度学习系统（Krizhevsky *et al.*, 2013）比以前基于手工特征的系统有了显著改进。从那时起，深度学习系统在某些视觉任务上的表现超过了人类，但在其他一些任务上还显落后。在语音识别、机器翻译、医疗诊断和博弈方面也有类似的进展。AlphaGo（Silver *et al.*, 2016, 2017, 2018）之所以能够战胜人类顶尖的围棋棋手，是因为它使用了深度网络来表示评价函数。

这些非凡的成功使学生、公司、投资者、政府、媒体和公众对人工智能的兴趣重新高涨。似乎每周都有新的人工智能应用接近或超过人类表现的消息，通常伴随着加速成功或人工智能新寒冬的猜测。

深度学习在很大程度上依赖于强大的硬件，一个标准的计算机 CPU 每秒可以进行 10^9 或 10^{10} 次运算。运行在特定硬件（例如 GPU、TPU 或 FPGA）上的深度学习算法，每秒可能进行 $10^{14} \sim 10^{17}$ 次运算，主要是高度并行化的矩阵和向量运算。当然，深度学习还依赖于大量训练数据的可用性，以及一些算法技巧（见第 21 章）。

1.4　目前的先进技术

斯坦福大学的人工智能百年研究（也称为 AI100）召集了专家小组来提供人工智能最先进技术的报告。2016 年的报告（Stone *et al.*, 2016; Grosz and Stone, 2018）总结：“未来人工智能的应用将大幅增加，包括更多的自动驾驶汽车、医疗诊断和针对性的治疗，以及对老年人护理的物理援助”，并且“社会现在正处于关键时刻，将决定如何以促进而不是阻碍自由、平等和透明等民主价值观的方式部署基于人工智能的技术。”AI100 还在其网站上创建了一个**人工智能指数**（AI Index），以帮助跟踪人工智能的进展。以下列举了与 2000 年基线相比（除非另有说明），2018 年和 2019 年报告的一些亮点。

- 出版物：人工智能论文数量在 2010 年至 2019 年间增长了 20 倍，达到每年约 2 万篇。最受欢迎的类别是机器学习（2009 年至 2017 年，arXiv.org 上的机器学习论文数量每年都会翻一番）。其次是计算机视觉和自然语言处理。
- 情绪：大约 70% 的人工智能新闻文章是中性的，但正面基调的文章从 2016 年的 12% 上

升到 2018 年的 30%。最常见的问题是道德问题——数据隐私和算法偏见。

- 学生：与 2010 年基线相比，课程注册人数在美国增加了 5 倍，全球增加了 16 倍。人工智能是计算机科学中最受欢迎的专业。
- 多样性：全球人工智能领域的教授中，大约 80% 是男性，20% 是女性。博士生和行业招聘也有类似的数字。
- 会议：NeurIPS 的参会人数比 2012 年增加了 8 倍，达到 13 500 人。其他会议的参会人数年增长率约为 30%。
- 行业：美国的人工智能初创公司数量增长了 20 倍，达到 800 多家。
- 国际化：中国每年发表的论文多于美国，与整个欧洲一样多。但是，在引用加权影响方面，美国作者领先中国作者 50%。从人工智能招聘人数看，新加坡、巴西、澳大利亚、加拿大和印度是增长最快的国家。
- 视觉：物体检测的错误率（大规模视觉识别挑战，LSVRC）从 2010 年的 28% 下降到 2017 年的 2%，超过了人类的表现。自 2015 年以来，开放式视觉问答（VQA）的准确率从 55% 提高到 68%，但仍远落后于人类 83% 的表现。
- 速度：在过去两年中，图像识别任务的训练时间减少了 100 倍。顶级人工智能应用使用的计算能力每 3.4 个月就会翻一番。
- 语言：以斯坦福问答数据集（SQuAD）的 F1 分数衡量的问答准确率，自 2015 年到 2019 年从 60 分提升到 95 分，在 SQuAD2 版本上进展更快，仅在一年内从 62 分提升到 90 分。这两个分数都超过了人类表现。
- 人类基准：截至 2019 年，人工智能系统在多个领域达到或超越人类表现，包括国际象棋、围棋、扑克、《吃豆人》（*Pac-Man*）、《危险边缘》（*Jeopardy!*）、ImageNet 物体检测、有限域中的语音识别、约束域中的英文翻译、《雷神之锤 3》（*Quake III*）、《刀塔 2》（*Dota* 2）、《星际争霸 II》（*StarCraft II*）、Atari 的各种游戏、皮肤癌检测、前列腺癌检测、蛋白质折叠、糖尿病视网膜病变诊断等。

人工智能系统何时（如果可以的话）能够在各种任务中达到人类水平的表现？马丁·福特（Martin Ford）（Ford, 2018）通过对人工智能专家的访谈发现这一目标时间的范围很广，从 2029 年到 2200 年，均值为 2099 年。在一项类似的调查中（Grace *et al.*, 2017），50% 的受访者认为这可能在 2066 年发生，有 10% 的人认为这最早可能在 2025 年发生，少数人则认为“不可能”。对于我们是需要根本性的新突破，还是仅仅对现有方法进行改进，专家们也存在分歧。但是不要过于严肃对待他们的预测，正如菲利普·泰洛克（Philip Tetlock）（Tetlock, 2017）在预测世界事件领域所证明的那样，专家并不比业余爱好者预测得更准。

未来的人工智能系统将如何运作？我们还不能确定。正如本节所详述的，这个领域采用了几个关于它本身的故事：首先是一个大胆的想法，即机器的智能是可能的，然后是它可以通过将专家知识编码成逻辑来实现，接着是建模世界的概率模型将成为主要工具，以及最近的机器学习将产生可能根本不基于任何易于理解的理论的模型。未来将揭示接下来会出现什么模式。

人工智能现在能做什么？也许不像一些更乐观的媒体文章让人相信的那样多，但仍然很多，以下是一些例子。

自动驾驶：自动驾驶的历史可以追溯到 20 世纪 20 年代的无线电遥控汽车，而在 20 世纪 80 年代首次展示了没有特殊向导的自动道路驾驶（Kanade *et al.*, 1986; Dickmanns and Zapp, 1987）。在 2005 年的 212 公里沙漠赛道 DARPA 挑战赛（Thrun, 2006）和 2007 年繁忙城市道路的城市挑战赛上，自动驾驶汽车成功展示之后，自动驾驶汽车的开发竞赛正式开始。2018

年，Waymo 的测试车辆在公共道路上行驶超过 1600万公里，没有发生严重事故，其中人类司机每 9650 公里才介入一次接管控制。不久之后，该公司开始提供商业机器人出租车服务。

自 2016 年以来，自动固定翼无人机一直在为卢旺达提供跨境血液输送服务。四轴飞行器可以进行出色的特技飞行，可以在构建三维地图的同时探索建筑，并进行自主编队。

腿足式机器人：雷伯特等人制作的四足机器人 BigDog（Raibert *et al.*, 2008），颠覆了我们对机器人如何行动的概念——不再是好莱坞电影中机器人缓慢、僵硬、左右摇摆的步态，而是类似于动物，并且能够在被推倒或在结冰的水坑上滑倒时恢复站立。类人机器人 Atlas 不仅能在崎岖不平的路况中行走，还可以跳到箱子上，做后空翻后可以稳定落地（Ackerman and Guizzo, 2016）。

自动规划和调度：在距离地球 1.6 亿公里的太空，美国宇航局（NASA）的“远程智能体”程序成为第一个控制航天器操作调度的机载自动规划程序（Jonsson *et al.*, 2000）。远程智能体根据地面指定的高级目标生成规划，并监控这些规划的执行（在出现问题时检测、诊断和恢复）。现在，Europa 规划工具包（Barreiro *et al.*, 2012）被用于 NASA 火星探测器的日常操作，而 Sextant 系统（Winternitz, 2017）允许航天器在全球 GPS 系统之外进行深空自主导航。

在 1991 年海湾危机期间，美国军队部署了动态分析和重新规划工具 Dart（Cross and Walker, 1994），为运输进行自动化的后勤规划和调度。规划涉及的交通工具、货物和人员达 5 万之多，并且必须考虑起点、目的地、路线、运输能力、港口和机场能力以及解决所有参数之间的矛盾。美国国防高级研究计划局（Defense Advanced Research Project Agency，DARPA）表示，这一应用取得的效果足以回报 DARPA 过去 30 年在人工智能领域的投资。

每天，优步（Uber）等网约车公司和谷歌地图等地图服务为数亿用户提供行车向导，在考虑当前和预测未来交通状况的基础上快速规划最佳路线。

机器翻译：在线机器翻译系统现在可以阅读超过 100 种语言的文档，涵盖 99% 的人类使用的母语，每天为数亿用户翻译数千亿词语。虽然翻译结果还不完美，但通常足以理解。对于具有大量训练数据的密切相关的语言（如法语和英语），在特定领域内的翻译效果已经接近于人类的水平（Wu *et al.*, 2016b）。

语音识别：2017 年，微软表示其会话语音识别系统的单词错误率已降至 5.1%，与人类在 Switchboard 任务（转录电话对话）中的表现相当（Xiong *et al.*, 2017）。现在全世界大约三分之一的计算机交互是通过语音而不是键盘完成的，另外 Skype 提供了 10 种语言的实时语音翻译。Alexa、Siri、Cortana 和谷歌都提供了可以回答用户问题和执行任务的助手。例如，谷歌 Duplex 服务使用语音识别和语音合成为用户预订餐厅，它能够代表用户进行流畅的对话。

推荐：Amazon、Facebook、Netflix、Spotify、YouTube、Walmart 等公司利用机器学习技术，根据用户过去的经历和其他类似的人群为用户推荐可能喜欢的内容。推荐系统领域有着悠久的历史（Resnick and Varian, 1997），但由于分析内容（文本、音乐、视频）以及历史和元数据的新深度学习方法的出现，推荐系统正在迅速发生变化（van den Oord *et al.*, 2014; Zhang *et al.*, 2017）。垃圾邮件过滤也可以被认为是推荐（或不推荐）的一种形式。目前的人工智能技术可以过滤掉 99.9% 以上的垃圾邮件，电子邮件服务还可以推荐潜在收件人以及可能回复的文本。

博弈：1997 年，当“深蓝”（Deep Blue）击败国际象棋世界冠军加里 • 卡斯帕罗夫（Garry Kasparov）后，人类霸权的捍卫者把希望寄托在了围棋上。当时天体物理学家、围棋爱好者皮特 • 赫特（Piet Hut）预测称：“计算机在围棋上击败人类需要一百年的时间（甚至可能更久）。”但仅仅 20 年后，AlphaGo 就超过了所有人类棋手（Silver *et al.*, 2017）。世界冠军柯洁说：“去年的 AlphaGo 还比较接近于人，现在它越来越像围棋之神。”AlphaGo 得益于对人类棋手过去数十万场棋局的研究以及对团队中围棋专家的知识提炼。

后继项目 AlphaZero 不再借助人类输入，只通过游戏规则就能够自我学习并击败所有对手，在围棋、国际象棋和日本将棋领域击败了包括人类和机器在内的对手（Silver *et al.*, 2018）。与此同时，人类冠军在各种游戏中被人工智能系统击败，包括《危险边缘》（Ferrucci *et al.*, 2010）、扑克（Bowling *et al.*, 2015; Moravčík *et al.*, 2017; Brown and Sandholm, 2019），以及电子游戏《刀塔 2》（Fernandez and Mahlmann, 2018）、《星际争霸 II》（Vinyals *et al.*, 2019）、《雷神之锤子》（Jaderberg *et al.*, 2019）。

图像理解：计算机视觉研究人员不再满足于在具有挑战性的 ImageNet 物体识别任务上超越人类的准确性，他们开始研究更困难的图像描述问题。一些令人印象深刻的例子包括“一个人在土路上骑摩托车”“两个比萨饼放在炉顶的烤箱上”和“一群年轻人在玩飞盘”（Vinyals *et al.*, 2017b）。然而，目前的系统还远远不够完善，一个“装满大量食物和饮料的冰箱”原来是一个被许多小贴纸遮挡住部分的禁止停车的标志。

医学：现在，人工智能算法在多种疾病的诊断方面（尤其是基于图像的诊断）已经达到或超过了专家医生的水平。例如，对阿尔茨海默病（Ding *et al.*, 2018）、转移性癌症（Liu *et al.*, 2017; Esteva *et al.*, 2017）、眼科疾病（Gulshan *et al.*, 2016）和皮肤病（Liu *et al.*, 2019c) 的诊断。一项系统回顾和汇总分析（Liu *et al.*, 2019a）发现，人工智能程序的平均表现与医疗保健专业人员相当。目前医疗人工智能的重点之一是促进人机合作。例如，Lyna 系统在诊断转移性乳腺癌方面达到了 99.6% 的总体准确性，优于独立的人类专家，但两者联合的效果仍然会更好（Liu *et al.*, 2018; Steiner *et al.*, 2018）。

目前，限制这些技术推广的不是诊断准确性，而是需要证明临床结果的改善，并确保透明度、无偏见和数据隐私（Topol, 2019）。2017 年，只有两项医疗人工智能应用获得 FDA 批准，但这一数字在 2018 年增至 12 项，并在持续上升。

气候科学：一个科学家团队凭借深度学习模型获得了 2018 年戈登·贝尔奖，该模型发现了之前隐藏在气候数据中的极端天气事件的详细信息。他们使用了一台具有专用 GPU 硬件，运算性能超过 exaop 级别（每秒 10^{18} 次运算）的超级计算机，这是第一个实现这一目标的机器学习程序（Kurth *et al.*, 2018）。Rolnick 等人（Rolnick *et al.*, 2019）提供了一个 60 页的目录，其中列举了机器学习可用于应对气候变化的方式。

这些只是几个目前存在的人工智能系统的例子。这不是魔法或科幻小说，而是科学、工程和数学，本书将对此进行介绍。

1.5 人工智能的风险和收益

弗朗西斯·培根是一位被誉为创造科学方法的哲学家，他在《论古人的智慧》（*The Wisdom of the Ancients*）（1609）一书中指出：“机械艺术的用途是模糊的，它既可用于治疗，也可用于伤害。”随着人工智能在经济、社会、科学、医疗、金融和军事领域发挥越来越重要的作用，我们应该考虑一下它可能带来的伤害和补救措施——用现代的说法，就是风险和收益。这里总结的话题在第 27 章和第 28 章中有更深入的讨论。

首先从收益说起。简而言之，我们的整个文明是人类智慧的产物。如果我们有机会获得更强大的机器智能，我们的理想上限就会大大提高。人工智能和机器人技术可以将人类从繁重的重复性工作中解放出来，并大幅增加商品和服务的生产，这可能预示着一个和平富足的时代的到来。加速科学研究的能力可以治愈疾病，并解决气候变化和资源短缺问题。正如谷歌 DeepMind 首席执行官德米斯·哈萨比斯（Demis Hassabis）所建议的那样：“首先解决人工智

能问题，然后再用人工智能解决其他所有问题。”

然而，早在我们有机会“解决人工智能”之前，我们就会因误用人工智能而招致风险，无论这是无意的还是其他原因。其中一些风险已经很明显，而另一些似乎基于当前趋势。

- *致命性自主武器*：联合国将其定义为无须人工干预即可定位、选择并击杀人类目标的武器。这种武器的一个主要问题在于它们的可扩展性——不需要人类监督意味着一小群人就可以部署任意数量的武器，并且这些武器的打击目标可以是通过任何可行的识别准则来定义的人类。自主武器所需的技术类似于自动驾驶汽车所需的技术。关于致命性自主武器潜在风险的非正式专家讨论始于 2014 年的联合国会议，并于 2017 年进入正式的官方专家组的条约审议阶段。
- *监视和劝诱*：安全人员监视电话线路、视频摄像头、电子邮件和其他消息渠道的代价昂贵、乏味且存在法律问题，但可以以一种可扩展的方式使用人工智能（语音识别、计算机视觉、自然语言理解）对个人进行大规模监视并检测感兴趣的活动。基于机器学习技术，通过社交媒体为个人量身定制信息流，可以在一定程度上修改和控制政治行为，这一问题在 2016 年开始的美国总统选举中变得显而易见。
- *有偏决策*：在评估假释和贷款申请等任务中，粗心或故意滥用机器学习算法可能会导致因种族、性别或其他受保护类别而产生有偏见的决策。通常，数据本身反映了社会中普遍存在的偏见。
- *就业影响*：关于机器会减少工作岗位的担忧由来已久。故事从来都不是简单的。机器能够完成一些人类可能会做的工作，但它们也让人类更有生产力，因此更适合被雇佣；让公司更具盈利能力，因此能够支付更高的工资。它们可能使一些本来不切实际的活动在经济上可行。它们的使用通常会导致财富增加，但往往会将财富从劳动力向资本转移，从而进一步加剧不平等。之前的技术进步（如机械织布机的发明），对就业造成了严重的影响，但最终人们还是找到了新的工作。另外，人工智能也有可能从事这些新的工作。这个话题正迅速成为世界各地经济学家和政府关注的焦点。
- *安全关键的应用*：随着人工智能技术的进步，它们越来越多地应用于高风险、安全关键的应用，如驾驶汽车和管理城市供水。已经发生过致命事故，这凸显了对使用机器学习技术开发的系统进行正式验证和统计风险分析的困难。人工智能领域需要制定技术和道德标准，至少要与其他工程和医疗领域中普遍存在的标准相当，而这些标准关乎人们的生命。
- *网络安全*：人工智能技术可用于防御网络攻击，如检测异常的行为模式，但这些技术也能用于增强恶意软件的威力、生存能力和扩散能力。例如，强化学习方法已被用于创建高效的工具，这些工具可以进行自动化、个性化的勒索和钓鱼攻击。

我们将在 27.3节更深入地讨论这些主题。随着人工智能系统变得越来越强大，它们将更多承担以前由人类扮演的社会角色。正如人类过去曾利用这些角色作恶一样，可以预见，人类可能会在这些角色中滥用人工智能系统而作恶更多。上面给出的所有例子都指出了治理的重要性，以及最终监管的重要性。目前，研究团体和参与人工智能研究的主要公司已经为人工智能相关活动制定了自愿自治原则（见 27.3 节）。各国政府和国际组织正在设立咨询机构，为每个具体的用例制定适当的条例，准备应对经济和社会影响，并利用人工智能的能力来解决重大的社会问题。

长期来看呢？我们能否实现长期以来的目标：创造出与人类智力相当或更强大的智能？如果我们做到了，然后呢？

在人工智能的大部分历史上，这些问题都被日常工作所掩盖——让人工智能系统做任何事情，哪怕是远程智能。与任何广泛的学科一样，绝大多数人工智能研究人员专注于特定的子领域，例如博弈、知识表示、视觉或自然语言理解，通常假设这些子领域的进展将有助于实现更广泛的人工智能目标。尼尔斯·约翰·尼尔森（Nils John Nilsson）（Nilsson, 1995）作为 SRI 的 Shakey 项目的最初负责人之一，提醒了该领域那些更广泛的目标，并警告说这些子领域本身有成为目标的风险。后来，一些有影响力的人工智能创始人，包括约翰·麦卡锡（McCarthy, 2007）、马文·明斯基（Minsky, 2007）和帕特里克·温斯顿（Beal and Winston, 2009），都认同尼尔森的警告，认为人工智能应该回归其本源，而不是专注于具体应用中可衡量的性能，用赫伯特·西蒙的话来说就是“会思考、会学习、会创造的机器”。他们将这种努力方向称为**人类级别的人工智能**（human-level AI，HLAI）——机器应该能够学会做人类可以做到的任何事情。他们在 2004 年召开了第一次研讨会（Minsky *et al.*, 2004）。另一个有着类似目标的工作是**通用人工智能**（artificial general Intelligence，AGI）运动（Goertzel and Pennachin, 2007），在 2008 年举行了第一次会议并组织出版了 *The Journal of Artificial General Intelligence*。

大约在同一时间，人们担心创造远远超过人类能力的**超级人工智能**（artificial superintelligence，ASI）可能是个坏主意（Yudkowsky, 2008; Omohundro, 2008）。图灵（Turing, 1996）在 1951 年曼彻斯特的一场演讲中也提出了同样的观点，他借鉴了塞缪尔·巴特勒（Samuel Butler）（Butler, 1863）的早期观点：[①]

> 似乎很可能，机器思维方法一旦开始，用不了多久它就会超越我们微弱的力量……因此，在某个阶段，我们应该需要期待机器能够受控制，就像塞缪尔·巴特勒在 *Erewhon* 中所提到的那样。

随着深度学习方面的最新进展，尼克·波斯特洛姆（Nick Bostrom）的《超级智能》（*Superintelligence*）（2014）等书籍的出版，以及斯蒂芬·霍金（Stephen Hawking）、比尔·盖茨（Bill Gates）、马丁·里斯（Martin Rees）和埃隆·马斯克（Elon Musk）的公开声明，这些担忧只会变得更加普遍。

对创造超级智能机器的想法产生普遍的不安感是自然的。我们可以称之为**大猩猩问题**（gorilla problem）：大约 700 万年前，一种现已灭绝的灵长类进化了，一个分支进化为大猩猩，另一个分支进化为人类。今天，大猩猩对人类分支不太满意，大猩猩根本无法控制自己的未来。如果这是成功创造出超级人工智能的结果（人类放弃对未来的控制），那么我们也许应该停止人工智能的研究，并且作为一个必然的结果，放弃人工智能可能带来的好处。这就是图灵警告的本质：我们可能无法控制比我们更聪明的机器。

如果超级人工智能是一个来自外太空的黑匣子，那么谨慎地打开这个黑匣子确实是明智之举。但事实并非如此：我们设计了人工智能系统，所以如果它们最终“掌控了自己”，那将是设计失败的结果（正如图灵所说）。

为了避免这种结果，我们需要了解潜在失败的根源。诺伯特·维纳（Wiener, 1960）在看到亚瑟·塞缪尔的西洋跳棋程序学会下棋并打败其创造者后，开始考虑人工智能的长远未来，他说：

---

① 甚至在更早的 1847 年，《原始解释者》（*Primitive Expounder*）的编辑理查德·桑顿（Richard Thornton）就对机械计算器大加抨击：“思想……超越自身，并通过发明机器进行自我思考来消除自身存在的必要性……但是谁知道，当这种机器变得更加完美的时候，它会不会想出一个规划来弥补自己的所有缺陷，然后想出超出常人所能理解的思想！”

> 如果我们为了达到目的而使用一个我们无法有效干预其运作方式的机械智能体……那么我们最好能完全确定设定给机器的目标是我们真正想要实现的。

许多文化都有关于人类向神灵、精灵、魔术师或魔鬼索取东西的神话。在这些故事中，他们总是得到了他们真正想要的东西并最终后悔。如果还有第三个愿望的话，那就是撤销前两个。我们将其称为**迈达斯国王问题**（King Midas problem）：迈达斯是希腊神话中的传奇国王，他要求他所接触的一切都变成黄金，但他在接触了他的食物、饮料和家人后，就后悔了。[①]1.1.5 节中我们已经提到过这个问题，将固定目标设定给机器的标准模型需要进行重大修改。解决维纳困境的方法根本不是“给机器设定一个明确的目的”。相反，我们希望机器努力实现人类的目标，但知道它们并不确切地知道这些目标是什么。

遗憾的是，迄今为止，几乎所有的人工智能研究都是在标准模型下进行的，这意味着这版书中几乎所有的技术材料都反映了这一知识框架。然而，在新框架内已经有一些初步成果。在第 16 章中，我们指出，当且仅当机器对人类的目标不确定时，机器才有积极的动机允许自己关闭。在第 18 章中，我们设计并研究**辅助博弈**（assistance game），它在数学上描述了一种情况，即人类有一个目标而机器试图实现它，但最初不确定目标是什么。在第 22 章中，我们解释**逆向强化学习**（inverse reinforcement learning）的方法，它允许机器通过观察人类的选择来更多地了解人类的偏好。在第 27 章中，我们探讨两个主要的困难：首先，我们的选择取决于我们的偏好，这是通过一个非常复杂、难以逆向的认知结构来实现的；其次，我们人类可能在一开始就没有一致的偏好（无论是作为个人还是作为一个群体），所以人工智能系统可能并不清楚应该为我们做什么。

小结

本章定义了人工智能并阐述了其发展的文化背景。本章要点如下。

- 不同的人对人工智能的期望不同。首先要问的两个重要问题是：你关心的是思想还是行为？你想模拟人类，还是试图达到最佳结果？
- 根据我们所说的标准模型，人工智能主要关注**理性行为**。理想的智能体会在某种情况下采取可能的最佳行为，在这个意义下，我们研究了**智能体**的构建问题。
- 这个简单的想法需要两个改进：首先，任何智能体（无论是人还是其他物体）选择理性行为的能力都受到决策计算难度的限制；其次，机器的概念需要从追求明确目标转变到追求目标以造福人类，虽然不确定这些目标是什么。
- 哲学家们（追溯到公元前 400 年）暗示大脑在某些方面就像一台机器，操作用某种内部语言编码的知识，并且这种思维可以用来选择要采取的行动，从而认为人工智能是有可能实现的。
- 数学家提供了运算逻辑的确定性陈述以及不确定的概率陈述的工具，也为理解计算和算法推理奠定了基础。
- 经济学家将决策问题形式化，使决策者的期望效用最大化。
- 神经科学家发现了一些关于大脑如何工作的事实，以及大脑与计算机的相似和不同之处。
- 心理学家采纳了人类和动物可以被视为信息处理机器的观点。语言学家指出，语言的使用符合这一模式。

① 如果迈达斯遵循基本的安全原则，并在他的愿望中包括“撤消”按钮和“暂停”按钮，他会过得更好。

- 计算机工程师提供了更加强大的机器，使人工智能应用成为可能，而软件工程师使它们更加易用。
- 控制理论涉及在环境反馈的基础上设计最优行为的设备。最初，控制理论的数学工具与人工智能中使用的大不相同，但这两个领域越来越接近。
- 人工智能的历史经历了成功、盲目乐观以及由此导致的热情丧失和资金削减的循环，也存在引入全新创造性的方法和系统地改进最佳方法的循环。
- 与最初的几十年相比，人工智能在理论和方法上都已经相当成熟。随着人工智能面对的问题变得越来越复杂，该领域从布尔逻辑转向概率推理，从手工编码知识转向基于数据的机器学习。这推动了真实系统功能的改进以及与其他学科更大程度的集成。
- 随着人工智能系统在真实世界中的应用，必须考虑各种风险和道德后果。
- 从长远来看，我们面临着控制超级智能的人工智能系统的难题，它们可能以不可预测的方式进化。解决这个问题似乎需要改变我们对人工智能的设想。

参考文献与历史注释

人工智能的早期先驱之一尼尔斯·尼尔森（Nilsson, 2009）给出了人工智能的完整历史。佩德罗·多明戈斯（Pedro Domingos）（Domingos, 2015）和麦莱尼亚·米切尔（Melanie Mitchell）（Mitchell, 2019）为普通读者提供了机器学习的概述，李开复（Kai-Fu Lee）（Lee, 2018）描述了人工智能国际领导地位的竞争。马丁·福特（Martin Ford）（Ford, 2018）采访了23位领衔的人工智能研究人员。

人工智能的主要专业协会有人工智能促进协会（Advancement of Artificial Intelligence，AAAI）、ACM人工智能特别兴趣小组（Special Interest Group in Artificial Intelligence，SIGAI，其前身为SIGART）、欧洲人工智能协会以及人工智能和行为模拟协会（Society for Artificial Intelligence and Simulation of Behaviour，AISB）。人工智能的伙伴关系将许多关注人工智能的道德和社会影响的商业和非营利组织聚集在一起。AAAI的*AI Magazine*包含许多专题和教程，其网站包含了人工智能相关新闻、教程和背景资料。

人工智能的最新工作会出现在国际人工智能联合会议（International Joint Conference on AI，IJCAI）、欧洲人工智能会议（European Conference on AI，ECAI）和AAAI会议这类主要人工智能会议的会刊中。国际机器学习会议（International Conference on Machine Learning，ICML）和神经信息处理系统（Neural Information Processing Systems，NeurIPS）会议涵盖机器学习领域。通用人工智能的主要期刊是*Artificial Intelligence*、*Computational Intelligence*、*IEEE Transactions on Pattern Analysis and Machine Intelligence*（TPAMI）、*IEEE Intelligent Systems*（TIS）和*Journal of Artificial Intelligence Research*（JAIR）。此外，还有许多专门讨论特定领域的会议和期刊，我们将在相应的章节中介绍。

第2章

智能体

我们在此讨论智能体的本质，完美的或不完美的、环境的多样性以及由此产生的智能体类型的集合。

第 1 章将**理性智能体**（rational agent）的概念确定为研究人工智能的方法的核心。本章将使这个概念更加具体。我们将看到，在任何可以想象的环境中运行的各种智能体都可以应用理性的概念。本书的计划是使用这个概念来制定一小组设计原则，并用于构建成功的智能体，可以合理地称之为**智能**系统。

我们从检查智能体、环境以及它们之间的耦合开始。观察到某些智能体比其他智能体表现得更好，可以自然而然地引出理性智能体的概念，即行为尽可能好。智能体的行为取决于环境的性质。我们将对环境进行粗略分类，并展示环境的属性如何影响智能体的设计。我们描述一些基本的“框架”智能体设计，本书余下的部分将充实相关内容。

2.1 智能体和环境

任何通过**传感器**（sensor）感知**环境**（environment）并通过**执行器**（actuator）作用于该环境的事物都可以被视为**智能体**（agent）。这个简单的想法如图 2-1 所示。一个人类智能体以眼睛、耳朵和其他器官作为传感器，以手、腿、声道等作为执行器。机器人智能体可能以摄像头和红外测距仪作为传感器，还有各种电动机作为执行器。软件智能体接收文件内容、网络数据包和人工输入（键盘 / 鼠标 / 触摸屏 / 语音）作为传感输入，并通过写入文件、发送网络数据包、显示信息或生成声音对环境进行操作。环境可以是一切，甚至是整个宇宙！实际上，我们在设计智能体时关心的只是宇宙中某一部分的状态，即影响智能体感知以及受智能体动作影响的部分。

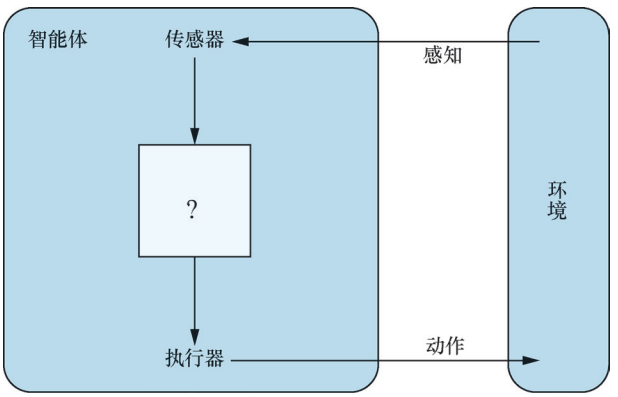

图 2-1　智能体通过传感器和执行器与环境交互

我们使用术语**感知**（percept）来表示智能体的传感器正在感知的内容。智能体的**感知序列**（percept sequence）是智能体所感知的一切的完整历史。一般而言，一个智能体在任何给定时刻的动作选择可能取决于其内置知识和迄今为止观察到的整个感知序列，而不是它未感知到的任何事物。通过为每个可能的感知序列指定智能体的动作选择，我们或多或少地说明了关于智能体的所有内容。从数学上讲，我们说智能体的行为由**智能体函数**（agent function）描述，该函数将任意给定的感知序列映射到一个动作。

可以想象将描述任何给定智能体的智能体函数制成表格。对大多数智能体来说，这将是一个非常大的表，事实上是无限的（除非限制考虑的感知序列长度）。给定一个要进行实验的智能体，原则上，我们可以通过尝试所有可能的感知序列并记录智能体响应的动作来构建此表[①]。当然，该表只是该智能体的外部特征。在内部，人工智能体的智能体函数将由**智能体程序**（agent program）实现。区别这两种观点很重要，智能体函数是一种抽象的数学描述，而智能体程序是一个具体的实现，可以在某些物理系统中运行。

为了阐明这些想法，我们举一个简单的例子——真空吸尘器世界。在一个由方格组成的世界中，包含一个机器人真空吸尘器智能体，其中的方格可以是脏的，也可以是干净的。图 2-2 展示了只有两个方格——方格 A 和方格 B——的情况。真空吸尘器智能体可以感知它在哪个方格中，以及方格中是否干净。智能体从方格 A 开始。可选的操作包括向右移动、向左移动、吸尘或什么都不做。[②] 一个非常简单的智能体函数如下：如果当前方格是脏的，就吸尘；否则，移动到另一个方格。该智能体函数的部分表格如图 2-3 所示，实现它的智能体程序如图 2-8 所示。

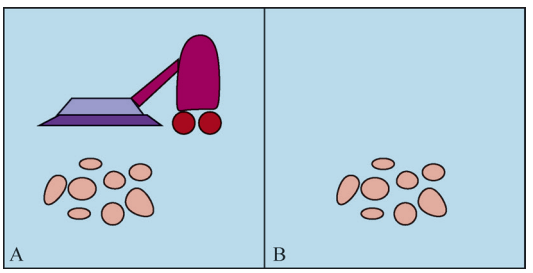

图 2-2 一个只有两个方格的真空吸尘器世界。每个位置可以是干净的，也可以是脏的，智能体可以向左移动或向右移动，可以清理它所占据的方格。不同版本的真空吸尘器世界允许不同的规则，例如智能体可以感知什么，它的动作是否总是成功等

从图 2-3 中可以看到，通过以各种方式填充右边的列可以简单地定义各种真空世界的智能体。那么，显而易见的问题是：填充表格的正确方法是什么？ 换句话说，是什么使智能体表现好或坏、聪明或愚蠢？我们将在 2.2 节中回答这些问题。

在结束本节之前，我们应该强调，智能体这一概念旨在成为分析系统的工具，而不是将世界划分为智能体和非智能体的绝对表征。人们可以将手持计算器视为一个智能体，它在给定感知序列“2+2=”时选择显示“4”的动作，但这样的分析很难帮助我们理解计算器。在某种意义上，工程的所有领域都可以被视为设计与世界互动的人工制品，人工智能运行在（作者认为是）这个系列最有趣的一端，在这一端，人工制品具有重要的计算资源，任务环境需要非凡的决策。

① 如果智能体在选择其动作时使用一些随机化，那就必须多次尝试每个序列来确定每个动作的概率。有人可能会认为随机地行动是相当愚蠢的，但本章后面展示出它可能非常聪明。

② 真正的机器人不太可能会有“向右移动”和“向左移动”这样的动作，而是采用“向前旋转轮子”和“向后旋转轮子”这样的动作。我们选择的动作更易于在书本上理解，而不是为了在实际的机器人中易于实现。

感知序列	动作
[*A*, *Clean*]	*Right*
[*A*, *Dirty*]	*Suck*
[*B*, *Clean*]	*Left*
[*B*, *Dirty*]	*Suck*
[*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Dirty*]	*Suck*
⋮	⋮
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Clean*]	*Right*
[*A*, *Clean*], [*A*, *Clean*], [*A*, *Dirty*]	*Suck*
⋮	⋮

图 2-3　图 2-2 所示的真空吸尘器世界的简单智能体函数的部分表项。如果当前方格是脏的，智能体就会进行清理，否则它将移到另一个方格。注意，除非限制可能感知序列的长度，否则该表的大小是无限的

2.2　良好行为：理性的概念

理性智能体（rational agent）是做正确事情的事物。显然，做正确的事情比做错误的事情要好，但是做正确的事情意味着什么呢？

2.2.1　性能度量

道德哲学发展了几种不同“正确事情”的概念，但人工智能通常坚持一种称为**结果主义**（consequentialism）的概念：我们通过结果来评估智能体的行为。当智能体进入环境时，它会根据接受的感知产生一个动作序列。这一动作序列会导致环境经历一系列的状态。如果序列是理想的，则智能体表现良好。这种可取性的概念由**性能度量**（performance measure）描述，该度量评估任何给定环境状态的序列。

人类有自己的欲望和偏好，因此，人类有适用于自身的理性概念。这一概念与成功地选择产生环境状态序列的行动有关，这些环境状态序列从人类的角度来看是可取的。但是，机器没有自己的欲望和偏好，至少在最初，性能度量是在机器设计者的头脑中或者是在机器受众的头脑中。我们将看到，一些智能体设计具有性能度量的显式表示（一个版本），而在其他设计中，性能度量完全是隐式的，智能体可能会做正确的事情，但它不知道为什么。

回顾诺伯特 • 维纳的警告，以确保“施以机器的目的是我们真正想要的目的”（1.5 节），注意，正确地制定性能度量可能非常困难。例如，考虑 2.1 节中的真空吸尘器智能体，我们可能会建议用单个 8 小时班次中清理的灰尘量来度量性能。当然，有了理性的智能体，你所要求的就是你所得到的。然而一个理性的智能体可以通过清理灰尘，然后将其全部倾倒在地板上，然后再次清理，如此反复，从而最大化这一性能度量。更合适的性能度量是奖励拥有干净地板的智能体。例如，在每个时间步中，每个干净的方格都可以获得 1 分（可能会对耗电和产生的噪音进行惩罚）。作为一般规则，更好的做法是根据一个人在环境中真正想要实现的目标，而

不是根据一个人认为智能体应该如何表现来设计性能度量。

即使避免了明显的缺陷，一些棘手的问题仍然存在。例如，上一段中“干净地板”的概念是基于一段时间内的平均整洁度。然而，两个不同的智能体可以达到相同的平均整洁度，其中一个智能体工作始终保持一般水平，而另一个智能体短时间工作效率很高但需要长时间的休息。哪种工作方式更可取似乎是保洁科学的好课题，但实际上这是一个具有深远影响的深刻哲学问题。大起大落、不计后果的生活，和安全但单调的生活，哪个更好？一个人人都生活在中度贫困的经济体，和一个有些人生活富裕而另一些人非常贫困的经济体，哪个更好？我们把这些问题留给勤奋的读者作为习题。

对于本书的大部分内容，我们将假设性能度量可以正确地指定。然而，出于前面所述原因，我们必须接受这样一种可能性：我们可能会将错误的目的施加给机器，确切地说，就是 1.5 节描述的迈达斯国王问题。此外，当设计一款软件（其副本将属于不同的用户）时，我们无法预测每个用户的确切偏好。因此，我们可能需要构建相应的智能体，它能够反映真实性能度量的初始不确定性，并随着时间的推移对其了解更多，第 16 章、第 18 章和第 22 章介绍了此类智能体。

2.2.2 理性

在任何时候，理性取决于以下 4 方面：

- 定义成功标准的性能度量；
- 智能体对环境的先验知识；
- 智能体可以执行的动作；
- 智能体到目前为止的感知序列。

这引出了**理性智能体的定义**：

> 对于每个可能的感知序列，给定感知序列提供的证据和智能体所拥有的任何先验知识，理性智能体应该选择一个期望最大化其性能度量的动作。

考虑一个简单的真空吸尘器智能体，如果一个方格是脏的就清理它，如果不脏就移动到另一个方格，这就是图 2-3 中给出的智能体函数。它是理性智能体吗？这需要看情况了！首先，我们需要说明性能度量是什么，对环境了解多少，以及智能体具有哪些传感器和执行器。我们假设：

- 在 1000 个时间步的“生命周期”内，性能度量在每个时间步为每个干净的方格奖励 1 分；
- 环境的“地理信息”是先验的（图 2-2），但灰尘的分布和智能体的初始位置不是先验的，干净的方格会继续保持干净，吸尘（*Suck*）动作会清理当前方格，向左（*Left*）或向右（*Right*）的动作使智能体移动一个方格，如果该动作会让智能体移动到环境之外，智能体将保持在原来的位置；
- 可用的动作仅有向右（*Right*）、向左（*Left*）和吸尘（*Suck*）；
- 智能体能够正确感知其位置以及该位置是否有灰尘。

在这种情况下，智能体确实是理性的，它的预期性能至少与任何其他智能体一样。

显而易见，同一个智能体在不同的情况下可能会变得不理性。例如，一旦清除了所有灰尘，该智能体将会毫无必要地反复来回；如果性能度量考虑对每个动作罚 1 分，那么智能体的表现就会很差。在确定所有方格都干净的情况下，一个更好的智能体不会做任何事情。如果干净的方格可能再次变脏，智能体应该偶尔检查，并在必要时重新清理。如果环境的地理信息是未知的，智能体则需要对其进行**探索**（explore）。习题 2.VACR 要求在这些情况下设计智能体。

2.2.3　全知、学习和自主

我们需要仔细区分理性和**全知**（omniscience）。全知的智能体能预知其行动的实际结果，并能据此采取行动，但在现实中，全知是不可能的。考虑这样一个例子：有一天我正沿着香榭丽舍大街散步，我看到街对面的一位老朋友。附近没有车流，我也没有别的事要做，所以理性上，我会开始过马路。与此同时，在 10 千米的高空，一架飞过的客机上有一扇货舱门脱落下来[①]，在我到达马路对面之前，我就被压扁了。我过马路是不理性的吗？我的讣告上不太可能写“试图过马路的白痴”。

这个例子表明，理性不等同于完美。理性使期望性能最大化，而完美使实际性能最大化。不要求完美不仅仅是对智能体公平的问题。关键是，如果我们期望一个智能体做事后证明是最好的行动，就不可能设计一个符合规范的智能体，除非我们改进占卜水晶球或时间机器的性能。

因此，我们对理性的定义并不需要全知，因为理性决策只取决于迄今为止的感知序列。我们还必须确保我们没有无意中允许智能体进行低智的行动。例如，如果一个智能体在穿过繁忙的道路之前没有向两边看，那么它的感知序列将不会告诉它有一辆大卡车正在以高速接近。我们对理性的定义是不是说现在就可以过马路了？绝非如此！

首先，考虑到这种缺乏信息的感知序列，过马路是不理性的：不观察路况就过马路发生事故的风险太大。其次，理性智能体在上街之前应该选择“观察”动作，因为观察有助于最大化期望性能。采取行动来改变未来的感知，有时被称为**信息收集**（information gathering），这是理性的一个重要组成部分，将在第 16 章中详细介绍。信息收集的另一个例子是真空吸尘器在最初未知的环境中必须进行的**探索**（exploration）。

我们的定义要求理性智能体不仅要收集信息，还要尽可能多地从它所感知到的东西中**学习**（learn）。智能体的初始配置可以反映对环境的一些先验知识，但随着智能体获得经验，这可能会被修改和增强。在一些极端情况下，环境完全是先验已知的和完全可预测的。在这种情况下，智能体不需要感知或学习，只需正确地运行。

当然，这样的智能体是脆弱的。如卑微的粪甲虫例子，在挖出巢穴产卵后，它会从附近的一堆粪中取出一团粪来堵住入口。如果粪球在途中被截下，粪甲虫根本不会注意到粪球已经不见了，仍会继续它的任务，并滑稽地用不存在的粪球堵住巢穴。进化已经在粪甲虫的行为中建立了一个假设，当它被违反时，不成功的行为就会产生。

稍微聪明一点的是掘土黄蜂。雌性掘土黄蜂会挖一个洞，出去刺一只毛毛虫并把它拖到洞口，再次进入洞里检查一切是否正常，然后把毛毛虫拖进洞里再去产卵。当蜂卵孵化时，毛毛虫会充当食物来源。到目前为止还不错，但如果昆虫学家在掘土黄蜂检查洞穴时将毛毛虫移动几厘米远，它将回到其规划中的“把毛毛虫拖到洞口”步骤，即使经过数十次移动毛毛虫的干预，它仍然继续执行该规划而不进行修改，不断地重新检查洞穴。掘土黄蜂无法知道其固有规划正在失败，因此不会改变规划。

如果在某种程度上，智能体依赖于其设计者的先验知识，而不是其自身的感知和学习过程，我们就说该智能体缺乏**自主性**（autonomy）。一个理性的智能体应该是自主的，它应该学习如何弥补部分或不正确的先验知识。例如，能学习预测何时何地会出现额外灰尘的真空吸尘器比不能学习预测的要好。

实际上，我们很少从一开始就要求智能体完全自主：除非设计者提供一些帮助，否则当智能体几乎没有经验或完全没有经验时，它将不得不随机行动。正如进化为动物提供了足够的内

① 参见 N. Henderson，“波音 747 大型喷气式飞机迫切需要新门锁”，华盛顿邮报，1989 年 8 月 24 日。

建反射，使其能够生存足够长的时间来学习一样，为人工智能体提供一些初始知识和学习能力也是合理的。在充分体验相应环境后，理性智能体的行为可以有效地独立于其先验知识。因此，结合学习能够让我们设计单个理性智能体，它能在各种各样的环境中取得成功。

2.3 环境的本质

既然已经有了理性的定义，考虑构建理性智能体的准备几乎已经完成。然而，还必须考虑**任务环境**（task environment），它本质上是“问题”，理性智能体是“解决方案”。我们首先展示如何指定任务环境，并用一些示例说明该过程。然后，展示任务环境的多种形式。任务环境的性质直接影响智能体程序的恰当设计。

2.3.1 指定任务环境

在讨论简单的真空吸尘器智能体的理性时，我们必须指定性能度量、环境以及智能体的执行器和传感器。我们将所有这些都归在任务环境的范畴下，基于首字母缩写规则，我们称其为**PEAS**（Performance，Environment，Actuator，Sensor）描述。在设计智能体时，第一步必须始终是尽可能完整地指定任务环境。

真空吸尘器世界是一个简单的例子，让我们考虑一个更复杂的问题：自动驾驶出租车司机。图 2-4 总结了出租车任务环境的 PEAS 描述。我们将在以下段落中更详细地讨论每个元素。

智能体类型	性能度量	环境	执行器	传感器
自动驾驶出租车司机	安全、速度快、合法、舒适旅程、最大化利润、对其他道路用户的影响最小化	道路、其他交通工具、警察、行人、客户、天气	转向器、加速器、制动、信号、喇叭、显示、语音	摄像头、雷达、速度表、GPS、发动机传感器、加速度表、麦克风、触摸屏

图 2-4 自动驾驶出租车司机任务环境的 PEAS 描述

首先，我们希望自动驾驶追求的**性能度量**（performance measure）是什么？理想的标准包括到达正确的目的地，尽量减少油耗和磨损，尽量减少行程时间或成本，尽量减少违反交通法规和对其他驾驶员的干扰，最大限度地提高安全性和乘客舒适度，最大化利润。显然，其中一些目标是相互冲突的，因此需要权衡。

接下来，出租车将面临什么样的驾驶**环境**（environment）？任何出租车司机都必须能够在各种道路上行驶，如乡村车道、城市小巷以及 12 车道的高速公路。道路上有其他交通工具、行人、流浪动物、道路工程、警车、水坑和坑洼。出租车还必须与潜在以及实际的乘客互动。另外，还有一些可选项。出租车可以选择在很少下雪的南加利福尼亚州或者经常下雪的阿拉斯加运营。它可能总是靠右行驶，或者我们可能希望它足够灵活，在英国或日本时可以靠左行驶。显然，环境越受限，设计问题就越容易解决。

自动驾驶出租车的**执行器**（actuator）包括可供人类驾驶员使用的器件，例如通过加速器控制发动机以及控制转向和制动。此外，它还需要输出到显示屏或语音合成器，以便与乘客进行对话，或许还需要某种方式与其他车辆进行礼貌的或其他方式的沟通。

出租车的基本**传感器**（sensor）将包括一个或多个摄像头以便观察，以及激光雷达和超声波传感器以便检测其他车辆和障碍物的距离。为了避免超速罚单，出租车应该有一个速度表，

而为了正确控制车辆（特别是在弯道上），它应该有一个加速度表。要确定车辆的机械状态，需要发动机、燃油和电气系统的传感器常规阵列。像许多人类驾驶者一样，它可能需要获取 GPS 信号，这样就不会迷路。最后，乘客需要触摸屏或语音输入才能说明目的地。

图 2-5 中简要列举了一些其他智能体类型的基本 PEAS 元素。更多示例参见习题 2.PEAS。这些示例包括物理环境和虚拟环境。注意，虚拟任务环境可能与“真实”世界一样复杂。例如，在拍卖和转售网站上进行交易的**软件智能体**（software agent），或称软件机器人或**软机器人**（softbot），为数百万其他用户和数十亿对象提供交易，其中许多对象具有真实的图片。

智能体类型	性能度量	环境	执行器	传感器
医学诊断系统	治愈患者、降低费用	患者、医院、工作人员	用于问题、测试、诊断、治疗的显示器	用于症状和检验结果的触摸屏/语音输入
卫星图像分析系统	正确分类对象和地形	轨道卫星、下行链路、天气	场景分类显示器	高分辨率数字照相机
零件选取机器人	零件在正确箱中的比例	零件输送带、箱子	有关节的手臂和手	摄像头、触觉和关节角度传感器
提炼厂控制器	纯度、产量、安全	提炼厂、原料、操作员	阀门、泵、加热器、搅拌器、显示器	温度传感器、气压传感器、流量传感器、化学传感器
交互英语教师	学生的考试分数	一组学生、考试机构	用于练习、反馈、发言的显示器	键盘输入、语音

图 2-5　智能体类型及其 PEAS 描述的示例

2.3.2　任务环境的属性

人工智能中可能出现的任务环境范围显然非常广泛。然而，我们可以确定相当少的维度，并根据这些维度对任务环境进行分类。这些维度在很大程度上决定了恰当的智能体设计以及智能体实现的主要技术系列的适用性。首先我们列出维度，然后分析几个任务环境，阐明思路。这里的定义是非形式化的，后面的章节提供了每种环境的更精确的陈述和示例。

完全可观测的（fully observable）与**部分可观测的**（partially observable）：如果智能体的传感器能让它在每个时间点都能访问环境的完整状态，那么我们说任务环境是完全可观测的。如果传感器检测到与动作选择相关的所有方面，那么任务环境就是有效的完全可观测的，而所谓的相关又取决于性能度量标准。完全可观测的环境很容易处理，因为智能体不需要维护任何内部状态来追踪世界。由于传感器噪声大且不准确，或者由于传感器数据中缺少部分状态，环境可能部分可观测。例如，只有一个局部灰尘传感器的真空吸尘器无法判断其他方格是否有灰尘，自动驾驶出租车无法感知其他司机的想法。如果智能体根本没有传感器，那么环境是**不可观测的**（unobservable）。在这种情况下，有人可能会认为智能体的困境是无解的，但是正如我们在第 4 章中讨论的那样，智能体的目标可能仍然可以实现，有时甚至是确定可以实现的。

单智能体的（single-agent）与**多智能体的**（multiagent）：单智能体和多智能体环境之间的区别似乎足够简单。例如，独自解决纵横字谜的智能体显然处于单智能体环境中，而下国际象棋的智能体则处于二智能体环境中。然而，这里也有一些微妙的问题。首先，我们已经描述了如何将一个实体视为智能体，但没有解释哪些实体必须视为智能体。智能体 A（例如出租车司机）是否必须将对象 B（另一辆车）视为智能体，还是可以仅将其视为根据物理定律运行的对

象，类似于海滩上的波浪或随风飘动的树叶？关键的区别在于 B 的行为是否被最佳地描述为一个性能度量的最大化，而这一性能度量的值取决于智能体 A 的行为。

例如，国际象棋中的对手实体 B 正试图最大化其性能度量，根据国际象棋规则，这将最小化智能体 A 的性能度量。因此，国际象棋是一个**竞争性**（competitive）的多智能体环境。但是，在出租车驾驶环境中，避免碰撞使所有智能体的性能度量最大化，因此它是一个部分**合作的**（cooperative）多智能体环境。它还具有部分竞争性，例如，一个停车位只能停一辆车。

多智能体环境中的智能体设计问题通常与单智能体环境下有较大差异。例如，在多智能体环境中，通信通常作为一种理性行为出现；在某些竞争环境中，随机行为是理性的，因为它避免了一些可预测性的陷阱。

确定性的（deterministic）与**非确定性的**（nondeterministic）：如果环境的下一个状态完全由当前状态和智能体执行的动作决定，那么我们说环境是确定性的，否则是非确定性的。原则上，在完全可观测的确定性环境中，智能体不需要担心不确定性。然而，如果环境是部分可观测的，那么它可能是非确定性的。

大多数真实情况非常复杂，以至于不可能追踪所有未观测到的方面；出于实际目的，必须将其视为非确定性的。从这个意义上讲，出租车驾驶显然是非确定性的，因为人们永远无法准确地预测交通行为。此外，轮胎可能会意外爆胎，发动机可能会在没有警告的情况下失灵。我们描述的真空吸尘器世界是确定性的，但变化可能包括非确定性因素，如随机出现的灰尘和不可靠的吸力机制（参考习题 2.VFIN）。

最后注意一点，**随机的**（stochastic）一词被一些人用作“非确定性”的同义词，但我们会区分这两个术语。如果环境模型显式地处理概率（例如，“明天的降雨可能性为 25%”），那么它是随机的；如果可能性没有被量化，那么它是“非确定性的”（例如，“明天有可能下雨”）。

回合式的（episodic）与**序贯的**（sequential）：在回合式任务环境中，智能体的经验被划分为原子式的回合。在每一回合中，智能体接收一个感知，然后执行单个动作。至关重要的是，下一回合并不依赖于前几回合采取的动作。许多分类任务是回合式的。例如，在装配流水线上检测缺陷零件的智能体需要根据当前零件做出每个决策，而无须考虑以前的决策；而且，当前的决策并不影响下一个零件是否有缺陷。但是，在序贯环境中，当前决策可能会影响未来所有决策。[①] 国际象棋和出租车驾驶是序贯的：在这两种情况下，短期行为可能会产生长期影响。因为在回合式环境下智能体不需要提前思考，所以要比序贯环境简单很多。

静态的（static）与**动态的**（dynamic）：如果环境在智能体思考时发生了变化，我们就说该智能体的环境是动态的，否则是静态的。静态环境很容易处理，因为智能体在决定某个操作时不需要一直关注世界，也不需要担心时间的流逝。但是，动态环境会不断地询问智能体想要采取什么行动，如果它还没有决定，那就等同于什么都不做。如果环境本身不会随着时间的推移而改变，但智能体的性能分数会改变，我们就说环境是**半动态的**（semidynamic）。驾驶出租车显然是动态的，因为驾驶算法在计划下一步该做什么时，其他车辆和出租车本身在不断移动。在用时钟计时的情况下国际象棋是半动态的。填字游戏是静态的。

离散的（discrete）与**连续的**（continuous）：离散/连续的区别适用于环境的状态、处理时间的方式以及智能体的感知和动作。例如，国际象棋环境具有有限数量的不同状态（不包括时钟）。国际象棋也有一组离散的感知和动作。驾驶出租车是一个连续状态和连续时间的问题，出租车和其他车辆的速度和位置是一系列连续的值，并随着时间平稳地变化。出租车的驾驶动

① “sequential”（串行）一词在计算机科学中也被用作“parallel”（并行）的反义词，与此处的含义在很大程度上是不相关的。

作也是连续的（转向角等）。严格来说，来自数字照相机的输入是离散的，但通常被视为表示连续变化的强度和位置。

已知的（known）与**未知的**（unknown）：严格来说，这种区别不是指环境本身，而是指智能体（或设计者）对环境“物理定律”的认知状态。在已知环境中，所有行动的结果（如果环境是非确定性的，则对应结果的概率）都是既定的。显然，如果环境未知，智能体将不得不了解它是如何工作的，才能做出正确的决策。

已知和未知环境之间的区别与完全可观测和部分可观测环境之间的区别不同。一个*已知的*环境很可能是*部分可观测的*，例如，在纸牌游戏中，知道规则但仍然无法看到尚未翻转的牌。相反，一个*未知*环境可以是*完全可观测的*，如一个全新的电子游戏，屏幕可能会显示整个游戏状态，但在尝试之前并不知道各个按钮的作用。

如 2.2.1 节所述，性能度量本身可能是未知的，这可能是因为设计者不确定如何正确地描述，也可能是因为最终用户（其偏好很重要）是未知的。例如，出租车司机通常不知道新乘客是喜欢悠闲还是快速的旅程，是喜欢谨慎还是激进的驾驶风格。虚拟个人助理一开始对新主人的个人喜好一无所知。在这种情况下，智能体可以基于与设计者或用户的进一步交互来了解更多关于性能度量的信息。继而，这表明，任务环境必须被视为一个多智能体环境。

最困难的情况是*部分可观测的*、*多智能体的*、*非确定性的*、*序贯的*、*动态的*、*连续的*且*未知的*。驾驶出租车除了驾驶员的环境大多是已知的，在所有其他方面都很难。在一个陌生的国家驾驶租来的汽车，那里有不熟悉的地理环境、不同的交通法规以及焦虑的乘客，这令人更加紧张。

图 2-6 列出了许多熟悉环境的属性。注意，这些属性并不总是一成不变的。例如，因为将患者的患病过程作为智能体建模并不适合，所以我们将医疗诊断任务列为单智能体，但是医疗诊断系统还可能必须应对顽固的病人和多疑的工作人员，因此环境还具有多智能体的方面。此外，如果我们将任务设想为根据症状列表进行诊断，那么医疗诊断是回合式的；如果任务包括提出一系列测试、评估治疗过程中的进展、处理多个患者等，那么则是序贯的。

任务环境	可观测	智能体	确定性	回合式	静态	离散
填字游戏	完全	单	确定性	序贯	静态	离散
限时国际象棋	完全	多	确定性	序贯	半动态	离散
扑克	部分	多	非确定性	序贯	静态	离散
西洋双陆棋	完全	多	非确定性	序贯	静态	离散
驾驶出租车	部分	多	非确定性	序贯	动态	连续
医疗诊断	部分	单	非确定性	序贯	动态	连续
图片分析	完全	单	确定性	回合式	半动态	连续
零件选取机器人	部分	单	非确定性	回合式	动态	连续
提炼厂控制器	部分	单	非确定性	序贯	动态	连续
交互英语教师	部分	多	非确定性	序贯	动态	离散

图 2-6　任务环境的例子及其特征

因为如前所述，“已知的 /未知的”不是严格意义上的环境属性，所以图 2-6 中没有包含此列。对于某些环境，例如国际象棋和扑克，很容易为智能体提供完整的规则知识，但考虑智能体如何在没有这些知识的情况下学会玩这些游戏仍然是有趣的。

与本书相关的代码库包括多个环境实现以及用于评估智能体性能的通用环境模拟器。实验通常不是针对单个环境进行的，而是针对从**环境类**（environment class）中抽象的许多环境进

行的。例如，要在模拟交通中评估出租车司机，我们需要运行具有不同的交通状况、照明和天气条件的多次模拟。我们关注智能体在环境类上的平均性能。

2.4 智能体的结构

到目前为止，我们通过描述行为（即在任意给定的感知序列之后执行的动作）讨论了智能体。现在我们必须迎难而上来讨论智能体内部是如何工作的。人工智能的工作是设计一个**智能体程序**（agent program）实现智能体函数，即从感知到动作的映射。假设该程序将运行在某种具有物理传感器和执行器的计算设备上，我们称之为**智能体架构**（agent architecture）：

$$\text{智能体} = \text{架构} + \text{程序}$$

显然，我们选择的程序必须是适合相应架构的程序。如果程序打算推荐步行这样的动作，那么对应的架构最好有腿。架构可能只是一台普通 PC，也可能是一辆带有多台车载计算机、摄像头和其他传感器的机器人汽车。通常，架构使程序可以使用来自传感器的感知，然后运行程序，并将程序生成的动作选择反馈给执行器。尽管本书第 25 章和第 26 章涉及传感器和执行器，但其余大部分内容都是关于设计智能体程序的。

2.4.1 智能体程序

我们在本书中设计的智能体程序都有相同的框架：它们将当前感知作为传感器的输入，并将动作返回给执行器。[①] 注意智能体程序（将当前感知作为输入）和智能体函数（可能依赖整个感知历史）之间的差异。因为环境中没有其他可用信息，所以智能体程序别无选择，只能将当前感知作为输入。如果智能体的动作需要依赖于整个感知序列，那么智能体必须记住历史感知。

我们用附录 B 中定义的简单伪代码语言描述智能体程序。（在线代码库包含真实编程语言的实现。）图 2-7 显示了一个相当简单的智能体程序，它记录感知序列，然后使用它来索引动作表，以决定要执行的动作。动作表（如图 2-3 中给出的真空吸尘器世界示例）明确表示了智能体程序所体现的智能体函数。作为设计者，为了以这种方式构建理性智能体，我们必须构造一个表，该表包含每个可能的感知序列所对应的适当动作。

```
function TABLE-DRIVEN-AGENT(percept) returns 一个动作
  persistent: percepts，初始为空的序列
              table，以感知序列为索引的动作表，初始为完全确定

  将percept添加到percepts的末尾
  action ← LOOKUP(percepts, table)
  return action
```

图 2-7 每个新感知都会调用 TABLE-DRIVEN-AGENT 程序，并且每次返回一个动作。它在内存中保留了完整的感知序列

表驱动的智能体构建方法注定失败，深入思考这一问题会很有启发性。设 $\mathcal{P}$ 为可能的感知集，T 为智能体的生存期（对应它将接收的感知总数），查找表将包含 $\sum_{t=1}^{T}|\mathcal{P}|^t$ 条记录。考虑自动驾驶出租车：来自单个摄像头（通常是 8 个摄像头）的视觉输入速度约为 70 MB/s（每秒

① 智能体程序框架还有其他选择。例如，我们可以让智能体程序作为与环境异步运行的协程。每个这样的协程都有一个输入和输出端口，并由一个循环组成，该循环读取输入端口的感知，并将动作写到输出端口。

30 帧，每帧 1080 像素 × 720 像素，每个像素包含 24 位颜色信息），驾驶 1 小时后，将会生成一张超过 $10^{600\,000\,000\,000\,000}$ 条记录的表。即使是作为真实世界中微小的、表现良好的片段的国际象棋，其查找表也至少有 10^{150} 条记录。相比之下，可观测宇宙中的原子数量少于 10^{80} 个。这些表的巨大规模意味着：（a）这个宇宙中没有任何物理智能体有空间存储表；（b）设计者没有时间创建表；（c）任何智能体都无法从其经验中学习所有正确的记录。

尽管如此，假设表填充正确，TABLE-DRIVEN-AGENT 确实做了我们想要做的事情：它实现了所需的智能体函数。

> 人工智能面临的关键挑战是找出编写程序的方法，尽可能从一个小程序而不是从一个大表中产生理性行为。

历史上有许多例子表明，在其他领域可以成功地做到这一点：例如，20 世纪 70 年代以前，工程师和学生使用的巨大平方根表格，现在已经被电子计算器上运行的仅有 5 行代码的牛顿方法所取代。现在问题是，人工智能能像牛顿处理平方根那样处理一般智能行为吗？我们相信答案是肯定的。

在本节剩余部分中，我们将概述 4 种基本的智能体程序，它们体现了几乎所有智能系统的基本原理：

- 简单反射型智能体；
- 基于模型的反射型智能体；
- 基于目标的智能体；
- 基于效用的智能体。

每种智能体程序以特定的方式组合特定的组件来产生动作。2.4.6 节大致解释了如何将所有这些智能体转换为学习型智能体，以提高其组件的性能，从而产生更好的动作。2.4.7 节描述在智能体中表示组件本身的各种方式。这种多样性为这一领域和这本书本身提供了一个主要的组织原则。

2.4.2　简单反射型智能体

最简单的智能体是**简单反射型智能体**（simple reflex agent）。这些智能体根据当前感知选择动作，忽略感知历史的其余部分。例如，真空吸尘器的智能体函数在图 2-3 所示，是一种简单反射型智能体，因为它的决策仅基于当前位置以及该位置是否有灰尘。该智能体的智能体程序如图 2-8 所示。

```
function REFLEX-VACUUM-AGENET([location,status]) returns 一个动作

  if status = Dirty then return Suck
  else if location = A then return Right
  else if location = B then return Left
```

图 2-8　在只有两个位置的真空吸尘器环境中，简单反射型智能体的智能体程序，该程序实现图 2-3 中列出的智能体函数

注意，与之前对应的表相比，真空吸尘器的程序确实非常轻量。最明显的简化来自忽略感知历史，这将相关感知序列的数量从 4^T 减少到 4。进一步的简化基于以下事实：动作不依赖于位置，只依赖于当前方格是否有灰尘。虽然我们已经使用 if-then-else 语句来编写智能体程序，但它非常简单，可以将其实现为布尔电路。

即使在更复杂的环境中，也会出现简单的反射行为。想象自己是自动驾驶出租车司机。如果前面的汽车刹车并且刹车灯亮起，那么你应该注意到这一点并开始刹车。换句话说，你通

过对视觉输入进行一些处理来建立我们称之为“前面的汽车正在刹车”的条件。然后，这会触发智能体程序中的既定联结，对应动作“启动刹车”。我们称这样的联结为**条件–动作规则**（condition-action rule）[①]，写作：

如果前面的车正在刹车，**则**启动刹车。

人类也有许多这样的联结，其中一些是习得反应（如驾驶），而另一些则是先天反射（如在有东西接近眼睛时眨眼）。在本书中，我们展示了学习和实现这种联结的几种不同方式。

图 2-8 所示的程序限定于一个特定的真空吸尘器环境。一种更通用、更灵活的方法是，首先为条件操作规则构建通用解释器，然后为特定任务环境创建规则集。图 2-9 给出了通用程序的结构示意图，展示条件 – 动作规则如何在智能体中建立从感知到动作的联结。如果这看起来普通，不要担心，很快就会变得更加有趣。

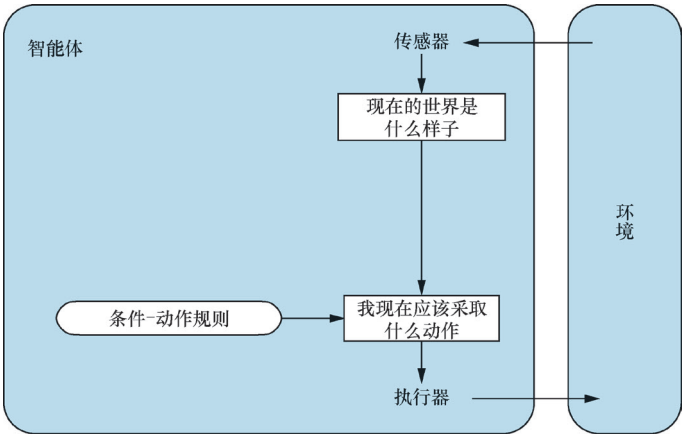

图 2-9 简单反射型智能体的示意图。我们使用矩形表示智能体决策过程的当前内部状态，使用椭圆表示过程中使用的背景信息

图 2-9 中智能体对应的智能体程序如图 2-10 所示。INTERPRET-INPUT 函数根据 *percept* 生成当前状态的抽象描述。给定状态描述，RULE-MATCH 函数返回规则集中匹配的第一条规则。注意，关于“规则”和“匹配”的描述纯粹是概念性的。如上所述，实际实现可以像实现布尔电路的逻辑门集合一样简单。或者，也可以使用“神经”电路，其中逻辑门由人工神经网络中的非线性单元代替（见第 21 章）。

```
function SIMPLE-REFLEX-AGENT(percept) returns 一个动作
  persistent: rules，一组条件–动作规则

  state ← INTERPRET-INPUT(percept)
  rule ← RULE-MATCH(state, rules)
  action ← rule.ACTION
  return action
```

图 2-10 简单反射型智能体。它根据一条规则进行操作，该规则的条件与感知定义的当前状态相匹配

简单反射型智能体具有值得赞扬的简单特性，但它们的智能有限。图 2-10 中的智能体只有在当前感知的基础上才能做出正确的决策，也就是说，只有在环境完全可观测的情况下才可行。

① 也称为**情境–动作规则**、**产生式系统**或 if-then 规则。

即使是轻微的不可观测性也会造成严重的问题。例如，前面给出的刹车规则假设前车正在刹车的条件可以通过当前的感知（视频的单帧）确定。如果前车有一个安装在中间的（因此是唯一可识别的）刹车灯，这是可行的。但是，旧款车型的尾灯、刹车灯和转向灯的配置各不相同，而且从单幅图像中分辨出汽车是在刹车还是仅仅打开了尾灯不是总能做到的。一个简单反射型智能体在这样一辆车后面行驶，要么会连续不必要地刹车，或者更糟的是根本就不刹车。

我们在真空吸尘器世界中也可以看到类似的问题。假设一个简单的真空吸尘器反射型智能体没有位置传感器，只有一个灰尘传感器。这样的智能体只有两种可能的感知：[*Dirty*] 和 [*Clean*]。它可以用吸尘（*Suck*）来响应 [*Dirty*]，它该如何响应 [*Clean*] 呢？如果碰巧从方格 A 开始，向左（*Left*）移动会（永远）失败，如果从方格 B 开始，向右（*Right*）移动会（永远）失败。对在部分可观测环境中工作的简单反射型智能体而言，无限循环通常是不可避免的。

如果智能体可以**随机化**（randomize）其操作，则可以跳出无限循环。例如，如果真空吸尘器智能体感知到 [*Clean*]，它可能会通过抛硬币来选择左右。我们很容易就能证明智能体将平均通过两步到达另一个方格。如果方格是脏的，智能体将清理它，任务就会完成。因此，随机化的简单反射型智能体可能优于确定性的简单反射型智能体。

我们在 2.3 节中提到，在某些多智能体环境中，正确的随机行为是理性的。在单智能体环境中，随机化通常是不理性的。在某些情况下，这是一个有用的技巧，可以帮助简单反射型智能体，但在大多数情况下，我们可以使用更复杂的确定性智能体以做得更好。

2.4.3　基于模型的反射型智能体

处理部分可观测性的最有效方法是让智能体追踪它现在观测不到的部分世界。也就是说，智能体应该维护某种依赖于感知历史的**内部状态**（internal state），从而至少反映当前状态的一些未观测到的方面。对于刹车问题，内部状态范围不仅限于摄像头拍摄图像的前一帧，要让智能体能够检测车辆边缘的两个红灯何时同时亮起或熄灭。对于其他驾驶任务，如变道，如果智能体无法同时看到其他车辆，则需要追踪它们的位置。为了在任何时候都能驾驶，智能体需要追踪其钥匙的位置。

随着时间的推移，更新这些内部状态信息需要在智能体程序中以某种形式编码两种知识。首先，需要一些关于世界如何随时间变化的信息，这些信息大致可以分为两部分：智能体行为的影响和世界如何独立于智能体而发展。例如，当智能体顺时针转动方向盘时，汽车就会向右转；而下雨时，汽车的摄像头就会被淋湿。这种关于“世界如何运转”的知识（无论是在简单的布尔电路中还是在完整的科学理论中实现）被称为世界的**转移模型**（transition model）。

其次，我们需要一些关于世界状态如何反映在智能体感知中的信息。例如，当前面的汽车开始刹车时，前向摄像头的图像中会出现一个或多个亮起的红色区域；当摄像头被淋湿时，图像中会出现水滴状物体并部分遮挡道路。这种知识称为**传感器模型**（sensor model）。

转移模型和传感器模型结合在一起让智能体能够在传感器受限的情况下尽可能地跟踪世界的状态。使用此类模型的智能体称为**基于模型的智能体**（model-based agent）。

图 2-11 给出了基于模型的反射型智能体的结构，它具有内部状态，展示了当前感知如何与旧的内部状态相结合，并基于世界如何运转的模型生成当前状态的更新描述。智能体程序如图 2-12 所示。有趣的部分是函数 UPDATE-STATE，它负责创建新的内部状态描述。模型和状态的表示方式的细节因环境类型和智能体设计中使用的特定技术而异。

无论使用哪种表示，智能体几乎不可能准确地确定部分可观测环境的当前状态。相反，标有“现在的世界是什么样子”（图 2-11）的框表示智能体的“最佳猜测”（或者在具有多种可能

性的情况下的最佳猜测）。例如，一辆自动驾驶出租车可能无法看到停在它前面的大卡车周围的情况，只能猜测是什么导致了拥堵。因此，关于当前状态的不确定性可能是不可避免的，但智能体仍然需要做出决定。

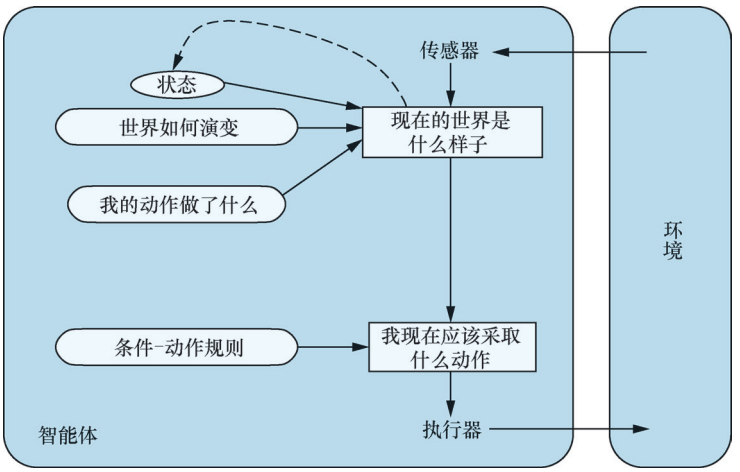

图 2-11 基于模型的智能体

function MODEL-BASED-REFLEX-AGENT(*percept*) **returns** 一个动作
  persistent: *state*, 智能体对世界状态的当前理解
    transition_model，关于下一个状态如何决定于当前状态和动作的描述
    sensor_model，关于当前世界状态如何反映到智能体感知的描述
    rules，一组条件–动作规则
    action，最近的动作，初始为空

  state ← UPDATE-STATE(*state*, *action*, *percept*, *transition_model*, *sensor_model*)
  rule ← RULE-MATCH(*state*, *rules*)
  action ← *rule*.ACTION
  return *action*

图 2-12 基于模型的反射型智能体。它使用内部模型追踪世界的当前状态，然后以与反射型智能体相同的方式选择动作

2.4.4 基于目标的智能体

了解环境的现状并不总是足以决定做什么。例如，在一个路口，出租车可以左转、右转或直行。正确的决定取决于出租车要去哪里。换句话说，除了当前状态的描述之外，智能体还需要某种描述理想情况的**目标**信息，例如设定特定的目的地。智能体程序可以将其与模型（与基于模型的反射型智能体中使用的信息相同）相结合，并选择实现目标的动作。图 2-13 展示了基于目标的智能体结构。

有时，基于目标的动作选择很直接，例如，单个动作能够立刻实现目标的情况。有时会更棘手，例如，智能体为了找到实现目标的方法而不得不考虑很长的复杂序列。**搜索**（第 3 ～ 5 章）和**规划**（第 11 章）是人工智能的子领域，专门用于寻找实现智能体目标的动作序列。

注意，这类决策从根本上不同于前面描述的条件–动作规则，因为它涉及对未来的考虑，包括“如果我这样做会发生什么？”和“这会让我快乐吗？”在反射型智能体设计中，这种信

息并没有被明确地表示出来，因为内置规则直接从感知映射到动作。反射型智能体在看到刹车灯时刹车，但它不知道为什么。基于目标的智能体在看到刹车灯时会刹车，因为这是它预测的唯一动作，这个动作可以实现不撞到其他汽车的目标。

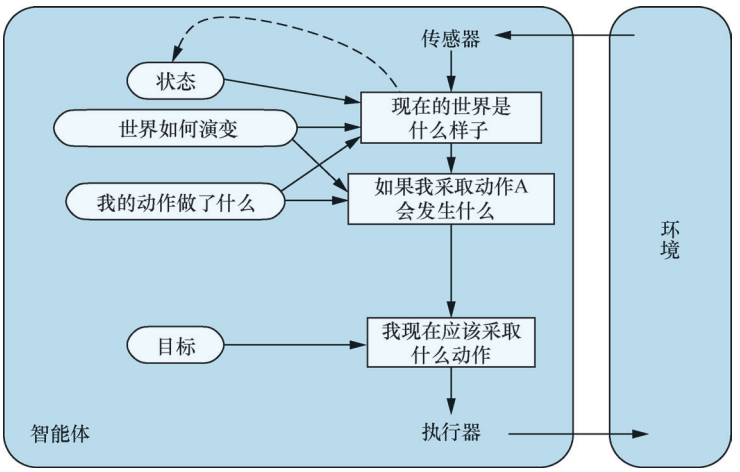

图 2-13　基于模型、基于目标的智能体。它追踪世界状态以及它试图实现的一系列目标，并选择一项最终能够实现目标的动作

尽管基于目标的智能体看起来效率较低，但它更灵活，因为支持其决策的知识是显式表示的，并且可以修改。例如，只要将目的地指定为目标，就可以很容易地更改基于目标的智能体的行为，以到达不同的目的地。反射型智能体关于何时转弯和何时直行的规则只适用于单一目的地，这些规则必须全部更换才能去新的目的地。

2.4.5　基于效用的智能体

在大多数环境中，仅靠目标并不足以产生高质量的行为。例如，许多动作序列都能使出租车到达目的地（从而实现目标），但有些动作序列比其他动作序列更快、更安全、更可靠或更便宜。目标只是在“快乐”和“不快乐”状态之间提供了一个粗略的二元区别。更一般的性能度量应该允许根据不同世界状态的“快乐”程度对智能体进行比较。经济学家和计算机科学家通常用**效用**（utility）这个词来代替“快乐”，因为“快乐”听起来不是很科学。[①]

我们已经看到，性能度量会给任何给定的环境状态序列打分，因此它可以很容易地区分到达出租车目的地所采取的更可取和更不可取的方式。智能体的**效用函数**（utility function）本质上是性能度量的内部化。如果内部效用函数和外部性能度量一致，那么根据外部性能度量选择动作，以使其效用最大化的智能体是理性的。

再次强调，这不是理性的唯一实现方式，我们已经看到了一个适用于真空吸尘器世界的理性智能体程序（图 2-8），但并不知道它的效用函数是什么。与基于目标的智能体一样，基于效用的智能体在灵活性和学习方面有很多优势。此外，在两种情况下，仅靠目标是不充分的，但基于效用的智能体仍然可以做出理性的决策。首先，当存在相互冲突的目标时，只能实现其中的一部分（例如速度和安全），效用函数会进行适当的权衡。其次，当智能体有多个目标实现，但没有一个目标可以确定地实现时，效用提供了一种方法，可以权衡目标的重要性和成功的可能性。

① 这里的“utility”一词指的是“实用的品质”，而不是电力公司或自来水厂等公共设施。

部分可观测性和非确定性在真实世界中普遍存在，因此，不确定性下的决策也普遍存在。从技术上讲，基于效用的理性智能体会选择能够最大化其动作结果**期望效用**（expected utility）的动作，也就是在给定每个结果的概率和效用的情况下，智能体期望得到的平均效用（附录 A 更精确地定义了期望）。在第 16 章中，我们证明，任何理性智能体的行为都必须表现得好像拥有一个效用函数，并试图最大化其期望值。具有显式效用函数的智能体可以使用通用算法做出理性决策，该算法不依赖于特定效用函数的最大化。通过这种方式，理性的“全局”定义（将那些具有最高性能的智能体函数指定为理性）变成了对理性智能体设计的“局部”约束，并可以通过一个简单的程序来表示。

基于效用的智能体结构如图 2-14 所示。基于效用的智能体程序见第 16 章和第 17 章，其中设计了决策型智能体，必须处理非确定性或部分可观测环境中固有的不确定性。如第 18 章所述，多智能体环境中的决策也在效用理论的框架下进行了研究。

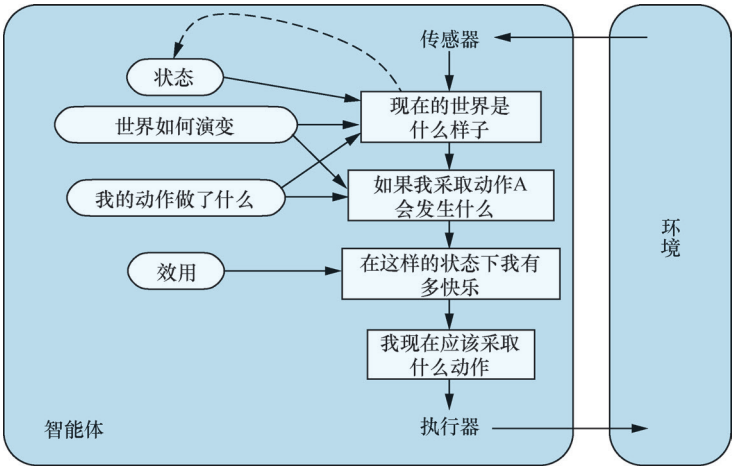

图 2-14 基于模型、基于效用的智能体。它使用了一个世界模型以及一个效用函数来衡量它在各状态之间的偏好，然后选择产生最佳期望效用的动作，其中期望效用是通过对所有可能的结果状态和对应概率加权所得

说到这里，读者可能会想，“这么简单吗？只需要构建能够最大化期望效用的智能体，我们就完成了？”这类智能体确实是智能的，但这并不简单。基于效用的智能体必须对其环境进行建模和跟踪，这些任务涉及大量关于感知、表示、推理和学习的研究。这些研究结果填满了本书的许多章节。选择效用最大化的行动方案也是一项艰巨的任务，需要更多的章节描述精巧的算法。即使使用这些算法，由于计算复杂性，完美理性在实践中通常是无法实现的（正如我们在第 1 章中所指出的）。我们还应该注意到，并非所有基于效用的智能体都是基于模型的。我们将在第 22 章和第 26 章中看到，**无模型的智能体**（model-free agent）可以学习在特定情况下什么样的动作是最好的，而不必确切地了解该动作如何改变环境。

最后，所有这些都假设设计者能够正确地指定效用函数，第 17 章、第 18 章和第 22 章将更深入讨论未知效用函数的问题。

2.4.6 学习型智能体

我们已经描述了一些智能体程序和选择动作的方法。到目前为止，我们还没有解释智能体程序是如何产生的。在图灵（Turing，1950）早期的著名论文中，他考虑手动编程实现智能机器的想法。他估计了这可能需要多少工作量，并得出结论，“似乎需要一些更快捷的方法。”他提出的方法是构造学习型机器，然后教它们。在人工智能的许多领域，这是目前创建最先进系统的首选方法。任

何类型的智能体（基于模型、基于目标、基于效用等）都可以构建（或不构建）成学习型智能体。

正如我们之前提到的，学习还有另一个优势：它让智能体能够在最初未知的环境中运作，并变得比其最初的知识可能允许的能力更强。在本节中，我们简要介绍学习型智能体的主要思想。在整本书中，我们对特定类型智能体中的学习因素和方法的评论贯穿全书。第 19 ～ 22 章将更加深入地介绍学习算法本身。

学习型智能体可分为 4 个概念组件，如图 2-15 所示。最重要的区别在于负责提升的**学习元素**（learning element）和负责选择外部行动的**性能元素**（performance element）。性能元素是我们之前认为的整个智能体：它接受感知并决定动作。学习元素使用来自**评估者**（critic）对智能体表现的反馈，并以此确定应该如何修改性能元素以在未来做得更好。

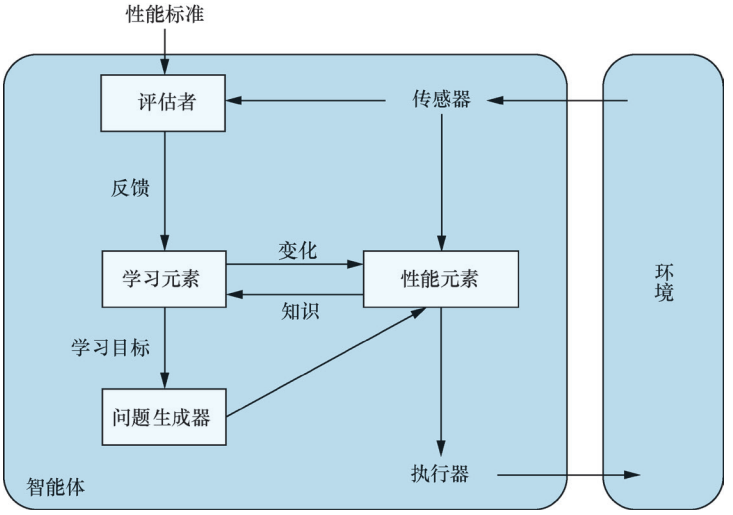

图 2-15　通用学习型智能体。“性能元素”框表示我们之前认为的整个智能体程序，现在“学习元素”框可以修改该程序以提升其性能

学习元素的设计在很大程度上取决于性能元素的设计。当设计者试图设计一个学习某种能力的智能体时，第一个问题不是“我要如何让它学习这个？”而是“一旦智能体学会了如何做，它将使用什么样的性能元素？”给定性能元素的设计，可以构造学习机制来改进智能体的每个部分。

评估者告诉学习元素：智能体在固定性能标准方面的表现如何。评估者是必要的，因为感知本身并不会指示智能体是否成功。例如，国际象棋程序可能会收到一个感知，提示它已将死对手，但它需要一个性能标准来知道这是一件好事；感知本身并没有这么说。确定性能标准很重要。从概念上讲，应该把它看作完全在智能体之外，因为智能体不能修改性能标准以适应自己的行为。

学习型智能体的最后一个组件是**问题生成器**（problem generator）。它负责建议动作，这些动作将获得全新和信息丰富的经验。如果性能元素完全根据自己的方式，它会继续选择已知最好的动作。但如果智能体愿意进行一些探索，并在短期内做一些可能不太理想的动作，那么从长远来看，它可能会发现更好的动作。问题生成器的工作是建议这些探索性行动。这就是科学家在进行实验时所做的。伽利略并不认为从比萨斜塔顶端扔石头本身有价值。他并不是想要打碎石头或改造不幸行人的大脑。他的目的是通过确定更好的物体运动理论来改造自己的大脑。

学习元素可以对智能体图（图 2-9、图 2-11、图 2-13 和图 2-14）中显示的任何“知识”组件进行更改。最简单的情况是直接从感知序列学习。观察成对相继的环境状态可以让智能体了解“我的动作做了什么”以及“世界如何演变”以响应其动作。例如，如果自动驾驶出租车在

湿滑路面上行驶时进行一定程度的刹车，那么它很快就会发现实际减速多少，以及它是否滑出路面。问题生成器可能会识别出模型中需要改进的某些部分，并建议进行实验，例如在不同条件下的不同路面上尝试刹车。

无论外部性能标准如何，改进基于模型的智能体的组件，使其更好地符合现实几乎总是一个好主意。（从计算的角度来看，在某些情况下简单但稍微不准确的模型比完美但极其复杂的模型更好。）当智能体试图学习反射组件或效用函数时，需要外部标准的信息。

例如，假设出租车司机因为乘客在旅途中感到非常不适，没有收到小费。外部性能标准必须告知智能体，小费的损失对其整体性能有负面影响；然后，该智能体可能会了解到暴力操作有损其自身的效用。从某种意义上说，性能标准将传入感知的一部分区分为**奖励**（reward）或**惩罚**（penalty），以提供对智能体行为质量的直接反馈。动物的疼痛和饥饿等固有的性能标准可以通过这种方式理解。

更一般地说，人类的选择可以提供有关人类偏好的信息。例如，假设出租车不知道人们通常不喜欢噪声，于是决定不停地按喇叭以确保行人知道它即将到来。随之而来的人类行为，如盖住耳朵、说脏话甚至可能剪断喇叭上的电线，将为智能体提供更新其效用函数的证据。这个问题将在第 22 章进一步讨论。

总之，智能体有各种各样的组件，这些组件可以在智能体程序中以多种方式表示，因此学习方法之间似乎存在很大差异。然而，主题仍然是统一的：智能体中的学习可以概括为对智能体的各个组件进行修改的过程，使各组件与可用的反馈信息更接近，从而提升智能体的整体性能。

2.4.7 智能体程序的组件如何工作

我们已经将智能体程序（用非常高级的术语）描述为由各种组件组成，其功能是回答诸如“现在的世界是什么样的？”“我现在应该采取什么动作？”“我的动作将导致什么？”等问题。人工智能学生的下一个问题是，“这些组件究竟是如何工作的？”要正确回答这个问题大约需要一千页的篇幅，但在这里我们希望读者能够注意一些基本区别，即组件表示智能体所处环境的各种方式之间的区别。

粗略地说，我们可以通过一个复杂性和表达能力不断增加的横轴来描述表示，即原子表示、因子化表示和结构化表示。为了辅助说明这些观点，我们可以考虑特定的智能体组件，例如处理“我的动作会导致什么”。这个组件描述了作为采取动作的结果可能在环境中引起的变化，图 2-16 展示了如何表示这些转移的示意图。

在**原子表示**（atomic representation）中，世界的每一个状态都是不可分割的，它没有内部结构。考虑这样一个任务：通过城市序列找到一条从某个国家的一端到另一端的行车路线（我们在图 3-1 中会解决这个问题）。为了解决这个问题，将世界状态简化为所处城市的名称就足够了，这就是单一的知识原子，也是一个“黑盒”，它唯一可分辨的属性是与另一个黑盒相同或不同。搜索和博弈中的标准算法（第 3 ～ 5 章）、隐马尔可夫模型（第 14 章）以及马尔可夫决策过程（第 17 章）都基于原子表示。

因子化表示（factored representation）将每个状态拆分为一组固定的**变量**或**属性**，每个变量或属性都可以有一个**值**。考虑同一驾驶问题更真实的描述，即我们需要关注的不仅仅是一个城市或另一个城市的原子位置，可能还需要关注油箱中的汽油量、当前的 GPS 坐标、油量警示灯是否工作、通行费、收音机上的电台等。两个不同的原子状态没有任何共同点（只是不同的黑盒），但两个不同的因子化状态可以共享某些属性（如位于某个特定的 GPS 位置），而其他属性不同（如有大量汽油或没有汽油），这使得研究如何将一种状态转换为另一种状态变得更加容易。人工智能的许多重要领域都基于因子化表示，包括约束满足算法（第 6 章）、命题逻辑

（第 7 章）、规划（第 11 章）、贝叶斯网络（第 12 ～ 16 章）以及各种机器学习算法。

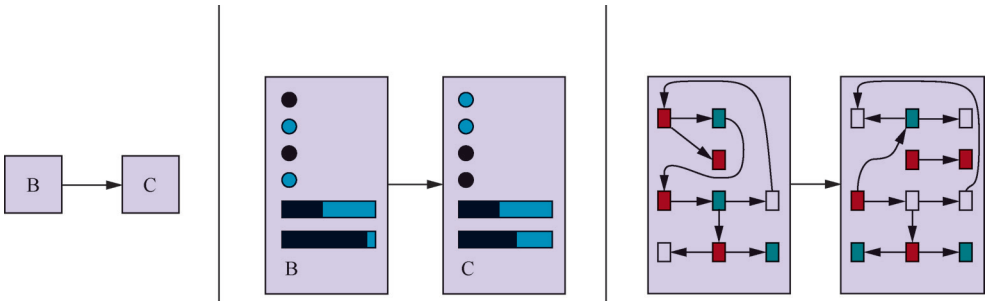

图 2-16　表示状态及其之间转移的 3 种方法：（a）原子表示一个状态（如 B 或 C）是没有内部结构的黑盒；（b）因子化表示状态由属性值向量组成，值可以是布尔值、实值或一组固定符号中的一个；（c）结构化表示状态包括对象，每个对象可能有自己的属性以及与其他对象的关系

出于许多目的，我们需要将世界理解为存在着相互关联的事物，而不仅仅是具有值的变量。例如，我们可能会注意到前面有一辆卡车正在倒车进入一个奶牛场的车道，但一头散养的奶牛挡住了卡车的路。因子化表示不太可能为属性 *TruckAheadBackingIntoDairyFarmDrivewayBlockedByLooseCow* 预先配备 *true* 或 *false* 的值。这就需要一个**结构化表示**（structured representation），在这种表示中可以明确地描述诸如奶牛和卡车之类的对象及其各种不同的关系（见图 2-16c）。结构化表示是关系数据库和一阶逻辑（第 8 ～ 10 章）、一阶概率模型（第 15 章）和大部分自然语言理解（第 23 章和第 24 章）的基础。事实上，人类用自然语言表达的大部分内容都与对象及其关系有关。

如前所述，原子表示、因子化表示和结构化表示所在的轴是**表达性**（expressiveness）增强的轴。粗略地说，可以通过简洁的描述捕捉到更具表达性的表示，表达性差的表示也可以捕捉到一切，但需要更多描述。通常，表达性更强的语言更简洁；例如，国际象棋规则可以用一两页结构化表示语言（如一阶逻辑）来描述，但需要数千页因子化表示语言（如命题逻辑）来描述，而需要 10^{38} 页的原子语言（如有限状态自动机）来描述。但是，随着表示能力的增强，推理和学习变得更加复杂。为了在避免缺点的同时获得表达性表示的好处，真实世界中的智能系统可能需要轴上的所有点同时运行。

另一个表示轴涉及从概念到物理记忆中位置的映射，包括计算机的内存和大脑的记忆。如果概念和记忆位置之间存在一对一的映射，我们称之为**局部表示**（localist representation）。但是，如果一个概念的表示分布在多个记忆位置，并且每个记忆位置被用作多个不同概念表示的一部分，我们称之为**分布式表示**（distributed representation）。分布式表示对噪声和信息丢失更健壮。使用局部表示，从概念到记忆位置的映射是随机的，如果传输错误而导致几位乱码，我们可能会将卡车（Truck）与无关的概念停战（Truce）混淆。但在分布式表示中，可以把每个概念想象成多维空间中的一个点，即使有一些乱码，也会移动到该空间中附近的点，其具有相似的含义。

小结

本章是人工智能的旋风之旅，在这个过程中我们认为人工智能是智能体设计的科学。本章要回顾的要点如下。

- **智能体**是在环境中感知和行动的事物。智能体的**智能体函数**指定智能体在响应任意感知序列时所采取的动作。
- **性能度量**评估智能体在环境中的行为。给定到目前为止所看到的感知序列，**理性智能体**的动作是为了最大化性能度量的期望值。
- **任务环境**规范包括性能度量、外部环境、执行器和传感器。在设计智能体时，第一步必须始终是尽可能完整地指定任务环境。
- 任务环境在几个重要维度上有所不同。它们可以是完全可观测的或部分可观测的、单智能体的或多智能体的、确定性的或非确定性的、回合式的或序贯的、静态的或动态的、离散的或连续的、已知的或未知的。
- 在性能度量未知或难以正确指定的情况下，智能体优化错误目标的风险很大。在这种情况下，智能体设计应该反映真实目标的不确定性。
- **智能体程序**实现智能体函数。存在各种基本的智能体编程，反映了决策过程中明确使用的信息类型。这些设计在效率、紧凑性和灵活性方面各不相同。智能体程序的适当设计取决于环境的性质。
- **简单反射型智能体**直接响应感知，而**基于模型的反射型智能体**保持内部状态以跟踪当前感知中不明晰的世界状态。**基于目标的智能体**采取行动来实现目标，而**基于效用的智能体**试图最大化自己期望的“快乐”。
- 所有智能体都可以通过**学习**提升性能。

参考文献与历史注释

动作在智能中的核心作用（实践推理的概念）至少可以追溯到亚里士多德的《尼各马可伦理学》（*Nicomachean Ethics*）。实践推理也是麦卡锡颇具影响力的论文“Programs with Common Sense”（McCarthy, 1958）的主题。机器人和控制理论领域本质上主要与物理主体有关。控制理论中的**控制器**概念与人工智能中的智能体概念相同。也许令人惊讶的是，人工智能在其历史上的大部分时间都集中在问答系统、定理证明器、视觉系统等孤立组件上，而不是完整智能体。吉内塞雷斯和尼尔森写的教科书（Genesereth and Nilsson, 1987）对智能体的讨论是一个有影响的例外。完整智能体的观点现在被广泛接受，并且是最近教科书（Padgham and Winikoff, 2004; Jones, 2007; Poole and Mackworth, 2017）的中心主题。

第 1 章追溯了理性概念在哲学和经济学中的根源。直到 20 世纪 80 年代中期，这一概念才在人工智能领域引起人们的兴趣，那时它开始充斥在许多关于该领域正确技术基础的讨论中。乔恩·多伊尔（Jon Doyle）的一篇论文（Doyle, 1983）预测称，理性智能体的设计将被视为人工智能的核心任务，而其他热门主题将衍生形成新的学科。

传统控制理论对环境特性及其对理性智能体设计的影响关注仔细且明显，例如，经典控制系统（Dorf and Bishop, 2004; Kirk, 2004）处理完全可观测的、确定性环境，随机最优控制（Kumar and Varaiya, 1986；Bertsekas and Shreve, 2007）处理部分可观测的随机环境，混合控制（Henzinger and Sastry, 1998; Cassandras and Lygeros, 2006）处理包含离散和连续元素的环境。在运筹学领域（Puterman, 1994）发展起来的**动态规划**（dynamic programming）文献中，完全可观测环境和部分可观测环境之间的区别也是核心问题，我们将在第 17 章中对此进行讨论。

虽然简单的反射型智能体是行为主义心理学的核心（见第 1 章），但大多数人工智能研究人

员认为它们过于简单而无法提供太多影响。罗森舍因（Rosenschein, 1985）和布鲁克斯（Brooks, 1986）质疑了这个假设（见第 26 章）。人们已经在寻找追踪复杂环境的有效算法方面做了大量工作（Bar-Shalom *et al.*, 2001; Choset *et al.*, 2005; Simon, 2006），其中大部分是在概率环境下所做的。

从亚里士多德的实践推理观点到麦卡锡关于逻辑人工智能的早期论文，都以基于目标的智能体为前提。机器人 Shakey（Fikes and Nilsson, 1971; Nilsson, 1984）是第一个基于目标的逻辑智能体的化身。吉内塞雷斯和尼尔森（Genesereth and Nilsson, 1987）对基于目标的智能体进行了全面的逻辑分析，肖厄姆（Shoham, 1993）开发了一种称为面向智能体编程的基于目标的编程方法。基于智能体的方法现在在软件工程中非常流行（Ciancarini and Wooldridge, 2001）。它还渗透到操作系统领域，其中**自主计算**（autonomic computing）指的就是通过感知 – 行为的循环和机器学习方法来监控自身的计算机系统和网络（Kephart and Chess, 2003）。注意，一组旨在在真正的多智能体环境中协同工作的智能体程序必然表现出模块化，即这些程序不共享内部状态，只通过环境相互通信。在**多智能体系统**领域，通常将单智能体的智能体编程为一个模块化的自主子智能体的集合。在某些情况下，人们甚至可以证明，由此产生的系统提供了与整体设计相同的最佳解决方案。

以目标为基础的智能体观点也主导了问题求解领域的认知心理学传统，从影响巨大的 *Human Problem Solving*（Newell and Simon, 1972）开始，并贯穿于纽厄尔后期所有的工作（Newell, 1990）。目标，进一步分析为欲望（广义）和意图（当前追求），是迈克尔 • 布拉特曼（Michael Bratman）开发的具有影响力的智能体理论的核心（Bratman, 1987）。

如第 1 章所述，效用理论作为理性行为基础的发展可以追溯到数百年前。在人工智能的早期研究中避开效用而倾向目标，但也有一些例外（Feldman and Sproull, 1977）。20 世纪 80 年代对概率方法兴趣的复苏，导致人们接受期望效用最大化作为决策的最一般框架（Horvitz *et al.*, 1988）。珀尔的教科书（Pearl, 1988）是人工智能领域第一本深入介绍概率和效用理论的教科书，它对不确定性下推理和决策的实用方法的阐述可能是 20 世纪 90 年代迅速转向基于效用的智能体的最大因素（见第 16 章）。强化学习在决策理论框架内的形式化也促成了这一转变（Sutton, 1988）。值得注意的是，直到最近，几乎所有的人工智能研究都假设性能度量可以通过效用函数或奖励函数的形式精确且正确地指定（Hadfield-Menell *et al.*, 2017a; Russell, 2019）。

图 2-15 中所示的学习型智能体的一般设计是机器学习文献中的经典（Buchanan *et al.*, 1978; Mitchell, 1997）。程序中体现的设计示例至少可以追溯到亚瑟 • 塞缪尔的（Samuel, 1959, 1967）中的学习型西洋跳棋程序。学习型智能体将在第 19 ～ 22 章中深入讨论。

（Huhns and Singh, 1998）和（Wooldridge and Rao, 1999）中收集了一些关于基于智能体方法的早期论文。关于多智能体系统的教科书为智能体设计的诸多方面提供了很好的指引（Weiss, 2000a; Wooldridge, 2009）。20 世纪 90 年代，多个专门讨论智能体的系列会议开始举办，包括智能体理论、架构和语言国际研讨会（International Workshop on Agent Theories, Architectures, and Languages，ATAL），国际自主智能体会议（International Conference on Autonomous Agents，AGENTS），以及国际多智能体系统联合会议（International Joint Conference on Multi-Agent Systems，ICMAS）。2002 年，这 3 个会议合并成了国际自主智能体及多智能体系统联合会议（International Joint Conference on Autonomous Agents and Multi-Agent Systems，AAMAS）。从 2000 年到 2012 年，每年都会举办面向智能体的软件工程（Agent-Oriented Software Engineering，AOSE）研讨会。期刊 *Autonomous Agents and Multi-Agent System* 创立于 1998 年。*Dung Beetle Ecology*（Hanski and Cambefort, 1991）提供了大量有关粪甲虫行为的有趣信息。YouTube 上有关于它们活动的视频记录，非常鼓舞人心。