第九章 脑、心智与认知

9.5 脑与认知

9.5.1 认知概念

认知是指人们获得知识或应用知识的过程，或者为了一定的目的，在一定的心理结构中进行的信息加工过程。与感知信息涉及感觉器官和神经处理过程不同，认知是人脑对接受外界输入的信息进行加工处理并转换为内在的心理活动，进而支配人的行为的过程。

认知无论是对哲学、心理学，还是认知科学及人工智能，都是一个基础的、重要的而又复杂、困难的科学研究对象，上述认知的定义来自普通心理学。此外，包括哲学、脑科学和神经科学、人工智能等交叉学科的研究形成了认知科学。

哲学家对于认知的理解非常宽泛，认为人类所有活动都可以看作是认知过程，这个过程涉及记忆、推理、判断、决策等活动。从古希腊到现代的哲学家几乎都对认知现象进行过思考，并提出了一些重要的哲学思想，如亚里士多德的认知“目的论”、笛卡儿的认知“预示论”、洛克的认知“观念论”、哈特莱的认知“实在论”、休谟的认知“实证论”、康德的认知“先验论”、现代的认知“计算论”等。

美国心理学家浩斯顿等人将认知归纳为5种主要类型：①认知是信息的处理过程；②认知是心理上的符号运算；③认知是问题求解；④认知是思维；⑤认知是一组知觉、记忆、思维、判断、推理及语言使用等相关活动。心理学历史上的行为主义学派曾长期坚持认为，只运用客观的实验和行为方法研究客观的可观察行为，后来发展的认知心理学则认为，应该运用信息加工观点来研究认知活动，就是将人脑与计算机进行类比，将人脑看作计算机的信息加工系统，这种理论实际上受到了人工智能的影响。

脑科学和神经科学比较关心的问题是脑和中枢神经系统作为一种特殊的物质系统，是怎样生长出认知能力的。人工智能认为，认知是系统生成智能的基础，因此希望理解清楚认知的机制，为人造系统生成智能打下基础。人工智能要达到类人智能程度，必须在认知方面有所突破，只是到目前为止，人工智能还停留在感知智能层面。

9.5.2 认知控制

执行功能是指脑的控制系统，它使人们能够组织思想和行为、安排任务的优先次序和规划，以及进行决策等。其中一些能力的发展贯穿了整个童年和青少年时期，并可用于准确预测个体今后发展的多种结果。

“执行功能”是心理学和神经科学领域的术语，有时又被称作认知控制，具体是指监督与协调其他高级心理机能的多组分体系。这一涵盖性术语用于描述注意、心理灵活性、计划、解决问题、文字推理、工作记忆，以及在不同任务间切换的能力。

执行功能源自现代心智的进化，而现代心智与前额皮层有关——相较于最近的灵长类祖先，人脑的前额皮层要发达得多。执行功能所涉及的过程，对指导目的性行为及处理新情况的能力都十分重要。

19世纪中叶，铁路工人菲尼斯·盖奇在一次意外中额叶受损，从而导致决策能力受到了影响。自此人们便认识到了前额皮层对于执行功能的重要性。随后研究人员对第一次世界大战中额叶受损的士兵展开了调查研究。研究表明，他们很难掌握新的任务。最终人们就相关的研究得出结论：执行功能对于抽象的高层次思维十分重要。

20世纪60年代，心理学家鲁利亚提出，额叶负责编程、监督并调节人们的行为。近来这一颇具影响力的观点被重新系统地表述为“监督注意系统模型”。根据该模型，执行功能包括含多个协调目标和行为交互子系统。其中一个子系统会在处理常规场景时被激活，并监督竞争脑内资源的自动反应，选择最合适的反应并抑制其他反应。而当人处于非常规的情境时，监督系统便会启动，根据需要转移注意力，以生成适当的新反应，并根据需要抑制和激活自动反应。

2001年，另一个重要的模型出现了。该模型基于脑内的信息处理是一种竞争过程的观点。该模型认为，前额皮层负责监督注意、记忆、情绪与运动等多个脑内系统的活动模式，并保持那些实现当前目标和行为所必需的模式。为此前额皮层在感觉输入、内部状态与动作输出之间建立彼此的映射，并在适当的神经通路内放大活动，以便执行当前任务。这在映射较弱或不断变化即遇到新情况时尤为重要。

第十一章 从脑科学到人工智能

11.0 学习导言

在人工智能发展初期，人工智能与神经科学和心理学密不可分，许多早期的先驱者跨越了这两个领域。历史证明，这些学科之间的协同发展是非常有效的。人工智能领域和神经科学领域是长期交织在一起的。近年来，神经科学和人工智能等相关领域都在快速发展。

在人工智能领域，近些年取得的研究成果是惊人的。然而历史上的人工智能研究，由于对脑功能机制探究的技术手段欠缺，以及人工智能和脑及神经科学领域相互交流较少，人工智能研究实际上长期处于盲人摸象的状态。在如今的神经科学领域，新的脑成像等技术的出现为揭示智能形成机制提供了可能性，人类有望对哺乳动物大脑实现革命性理解。因此，从脑科学和神经科学出发研究人工智能也具备了更充分的条件。机器智能和人类智能之间的差距是巨大的。为弥补这一巨大差距，从脑和神经科学中汲取的灵感会变得越来越重要。神经科学为新类型的算法和架构提供了丰富的灵感来源，这些算法和架构独立于并补充了基于数学和逻辑的方法和思想。神经科学也能够对现有的人工智能技术进行验证。如果一个已知的算法随后被发现在大脑中执行，那么它作为机器智能系统一个组成部分的合理性就得到了强有力的支持。本章主要从传统人工神经元到现代类脑计算等方面来学习和理解脑与神经科学对人工智能发展的基础性作用和意义。

11.9 脑与神经科学对人工智能的作用

人工智能与脑科学两个领域从截然不同的角度对智能进行处理：神经学研究大脑是如何工作的并且研究潜在的生物学原理，而人工智能更实用，它不受人类发展的束缚。可以将人工智能看作是应用计算神经科学解决实际问题。将智能变为机器算法，并将其与人脑进行比较，类脑计算方面的研究可能会使人们对大脑中的一些秘密产生深刻见解。毫无疑问，今后人工智能研究必须向人脑学习，发展类脑计算等技术。

脑与神经科学的进展，特别是借助新技术与新设备的研究支持研究者通过不同的实验方法（如生物解剖、电生理信号采集与分析、光遗传技术、分子病毒学、功能影像分析等）得到对脑的多尺度、多类型的生物证据，正在尝试从不同侧面揭示生物智能的结构和功能基础。

其实，几十年来的人工智能研究，囿于观察、测量手段等方面的限制，一直无法深入理解人脑的运行机理，都是围绕基于人脑表现出的智能特征开展各个方向的研究工作。随着CT等大型医疗设备的应用，人类可以观察到脑的内部运行机制，深入到脑神经元层次观察诸如观看一段视频或者听到某个熟悉的声音时，大脑是如何做出反应的，那些区域发挥什么作用，脑神经元处理外界信息的过程如何。但它才刚开始揭开神经元如何有助于高智能发展。神经学家通常只对他们研究背后的机制有模糊的概念。由于人工智能研究依赖于严谨的数学运算，该领域可以提供一种方法对这些模糊的概念进行阐明，使他们变为可验证的假设。

关注大脑的记忆和想象力，理解大脑中的哪些区域参与其中，存在何种机制，随后使用这些知识帮助我们思考，如何在人工智能系统中实现同样的功能。另一方面是，智能究竟是什么，这也包括自然智能和人类的智力。利用可以完成有趣任务的人工智能算法，可以了解如何看待大脑本身。也可以将这些人工智能系统作为模型，了解大脑中正在发生什么。

在微观层面，生物神经元和突触的类型、数目等在不同脑区中具有较大差异，且能够根据任务的复杂性实现结构和功能的动态适应。在介观层面，特异性脑区内部的连接模式和随机性网络背景噪声的有效融合，使生物神经元在保持了特定网络功能的同时，兼顾了动态网络可塑性。在宏观层面，不同脑区之间的协同使高度智能的类人认知功能得以实现。如哺乳动物脑的强化学习认知功能，长时、短时记忆功能等都是通过不同脑区功能的协同实现更复杂的认知功能。

未来要实现更强大的人工智能技术，需要计算模型融合来自微观、介观、宏观多尺度脑结构和信息处理机制的启发。

深思公司创始人、AlphaGo的开发者之一哈比乌斯认为，只有了解大脑，才能开发出更强的人工智能。他强调神经科学是实现强人工智能的基础。神经科学通过两种方式帮助人类实现这个目标。其中之一是将神经科学作为算法和架构理念的灵感来源。

关于开发通用人工智能系统的可行性，人脑是现存的模板和参考对象。ChatGPT已经具备通用人工智能系统的雏形，但它并非以人脑为原型构建。未来以人脑为原型的更强大的通用人工智能系统能否产生还是未知数。

11.10 本章小结

对于人工智能而言，从最初模拟生物神经元的人工神经元模型，到人脑强化学习启发的机器强化学习，脑的智能作为人工智能的模拟对象，从脑的智能特征模拟产生的算法模型到结构上模拟大脑的类脑计算，促进了很多人工智能技术的产生和发展。但是，人们对脑的智能机制了解得很少，迄今无法整体认识清楚脑的思维、认知等智能机制及自我意识等是如何产生的。实际上，脑不会脱离身体和环境单独产生智能。未来要发展像人一样的通用人工智能，应综合考虑脑、身体与环境的相互作用机制，进而设计出在灵活适应环境基础上不断进化的机器智能系统。