

认知相关性与智能模型构造的系统观点

危 辉 著



科学出版社

认知相关性与智能模型 构造的系统观点

危 辉 著

科 学 出 版 社

北 京

内 容 简 介

人或者高级哺乳动物的智能呈现为一个整体,各种认知行为或认知技能是相互协同的。由于对生物智能系统的认知尚不全面和技术手段的限制,人工智能研究处于一种较为割裂的状态,也就是在设计人工智能系统时对认知原型的相关性和系统性考虑不足,以满足单项目标为目的。本书从认知系统的整体性出发,探讨了从认知信息加工角度人工智能研究应该注意的相关性、系统性和发展性问题。我们认为统一的认知结构应该是实现系统性人工智能建模的关键,其中表示问题是联系发展问题和智能应用问题的之处。神经科学对神经信息编码、皮层信息加工等机制的研究为人工智能的表示问题打开了深入的大门,我们在此基础上探讨了构建言语计算模型、认知加工动力学过程、视皮层对概念语义的支撑等问题。

本书适合从事计算机科学、人工智能与认知科学的研究人员阅读。

图书在版编目(CIP)数据

认知相关性与智能模型构造的系统观点/危辉著. —北京:科学出版社, 2012

ISBN 978-7-03-034986-6

I. ①认… II. ①危… III. ①人工智能—研究 IV. ①TP18

中国版本图书馆 CIP 数据核字(2012) 第 134149 号

责任编辑:王丽平 房 阳/责任校对:钟 洋

责任印制:钱玉芬/封面设计:耕者设计工作室

科学出版社出版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

铭浩彩色印装有限公司印刷

科学出版社发行 各地新华书店经销

*

2012 年 6 月第 一 版 开本: B5 (720 × 1000)

2012 年 6 月第一次印刷 印张: 11 1/4

字数: 213 000

定价: 45.00 元

(如有印装质量问题, 我社负责调换)

前 言

传统的符号化人工智能采用的是“再现智能”的方法，设计的模型往往是对一种以某种形式表现、应用于某一局部领域的智能现象的模拟，对模型的认知意义并不刻意追求。这种指导思想对智能就事论事，缺乏对其本质的一个深刻洞察，没有将此智能现象作为人所有智能的一部分放在人的心智体系中进行全面系统的考察、放在认知发展全过程的前因后果序列中考察、放在无法回避的认知约束下进行考察。我们已经不满足淤陷其中的形式推理，而转向实现更朴素和本源的基于语义的推理。要想根本性地提高人工智能系统的性能，实现向“解释智能”的转变就必须将人工智能放在认知科学这个交叉学科的背景下进行修正，后者通过适当地使用还原和整合的策略说明高层次的智能现象如何突现。认知是包括知识获取、存储、回忆和使用的全部精神活动。如果对认知取信息加工的观点，那么认知的相关性就是认知结构在所加工信息的形态、加工流程、结构的组成、学习对结构的影响等方面的依赖关系。对这些相关性的研究对于知识获取、知识表示和知识应用的分阶段模型和联合模型的结构以及构造过程研究具有重大意义。认知的相关性约束要求我们考虑智能的发生、发展和完善的过程，以不割裂的观点看待智能行为的模拟，这就是智能模型构造的系统观点。通过对认知发展过程的研究使我们能更清楚、全面地看到智能的本质，而不仅仅是它的外表。认知相关性要求和系统的智能观是建立在统一的认知结构的基础之上，这个结构的基础和统一的体现是对信息的直接表达，这是一种对知觉信息的表达，是对语义的直接表达，它是心智的计算理论中心智表达和心智计算的核心，通过基于神经网络的直接表达的实现能够找到整体主义和还原主义、天赋论和建构论的结合点。语言是最具代表性的智能行为，对个人言语的计算结构研究能最大限度的体现认知心理学对于言语的信息加工算法、生理心理学对言语加工模型的体系结构、计算神经科学对于这个加工结构的实现、发展心理学对计算模型的进化作用。对言语计算结构的研究是认知相关性和智能模型构造的系统观点的实践。

作 者

2012 年 5 月

目 录

前言

第 1 章	引言：人工智能中的一些基本问题	1
1.1	人工智能受到的批评	1
1.2	人工智能中的几个基本问题	2
1.3	人工智能需要一个更广泛的背景	4
1.4	智能的心理表现与生理基础	6
1.5	本书的逻辑演进关系	7
第 2 章	认知相关性及其在人工智能背景下的意义	8
2.1	从形式推理到语义推理	8
2.1.1	形式推理及其困难	8
2.1.2	语义推理	12
2.2	认知相关性	13
2.2.1	模式相似性	14
2.2.2	语义相关	15
2.2.3	连贯性	16
2.2.4	继承性	17
2.2.5	协作性	18
2.3	认知相关性对人工智能的意义	19
2.3.1	系统和整体的智能观	19
2.3.2	知识的依赖性为什么重要	20
2.4	人工智能模型在认知科学背景下应满足的约束	21
2.5	本章小结	23
第 3 章	人工智能的关键：表示	24
3.1	从知识表示到表示：人工智能认识论上的进步	24
3.1.1	知识表示与表示是不同的	24
3.1.2	表示反映了对客观真实的认识	27
3.1.3	人工智能认识论上的进步	29
3.2	表象式直接表示	30
3.2.1	非符号化直接表示	30
3.2.2	视知觉研究	30

3.2.3	表象式直接表示	31
3.2.4	相关的问题	35
3.3	实践: 基于结构学习和迭代自映射的自联想记忆模型	37
3.3.1	结构学习与权值学习不同	37
3.3.2	网络结构设计	37
3.3.3	网络结构和迭代过程的数学描述	42
3.3.4	实验结果	46
3.3.5	相关的问题	47
3.4	本章小结	48
第 4 章	智能模型构造的发展观	49
4.1	常识: 知识获取的瓶颈	49
4.1.1	常识问题的由来	49
4.1.2	造成常识问题的原因	51
4.2	对策: 智能系统构造的发展观	51
4.3	发展心理学与智能系统构造	53
4.3.1	依赖性源于发展性	53
4.3.2	发展心理学的启示	54
4.4	本章小结	56
第 5 章	实例: 言语的计算结构研究	57
5.1	本研究与传统计算语言学的不同	57
5.2	言语的心理语言学框架	58
5.3	基于知觉加工模式的发展式分词算法	61
5.3.1	从认知心理学的角度看待自然语言理解问题	62
5.3.2	基于语言发展规律的分词模型	62
5.3.3	基于知觉加工双向过程的分词算法	63
5.3.4	实例分析	65
5.3.5	相关讨论	67
5.4	基于参照的对词结构操作语义的归纳学习	68
5.4.1	语言获得过程中的发现学习	68
5.4.2	基于参照的词结构的操作语义	69
5.4.3	操作语义的表达	70
5.4.4	基于参照的归纳发现学习算法	71
5.4.5	相关讨论	74
5.5	本章小结	75

第 6 章 基于连通结构与矩阵特征向量的联想记忆双层模型	77
6.1 引言	77
6.2 直接表达	78
6.3 动力学过程与连通结构	80
6.4 随机连接的双层网络结	81
6.5 单个神经元的动力学过程算法	83
6.5.1 单个神经元的动力学特征	83
6.5.2 基于矩阵特征值的单个神经元结构学习算法	85
6.6 自主的、并行分布式连通求取算法	88
6.7 实验结果	90
6.7.1 回忆与纠错实验	90
6.7.2 连通结构建立的速度与规模	91
6.7.3 学习次数与连通程度的对比	91
6.8 容量与纠错率分析	92
6.9 本章小结	94
第 7 章 神经系统动力学模型及动力学过程对认知操作的广泛表达意义	96
7.1 心理活动的外观与内部基础	96
7.2 基于激励均势扩散模式的网络结构组织模型	97
7.3 神经元动力学行为的数学描述	99
7.4 激励浸润与网络集群动力学过程的数学描述	102
7.5 基于结构特征的记忆编码算法	106
7.6 网络的性能分析	108
7.6.1 稳定性分析	108
7.6.2 回忆成功率分析	110
7.6.3 记忆容量分析	112
7.7 动力学过程对认知操作的广泛表达意义	112
第 8 章 人工智能的神经系统动力学融合表示模型研究	116
8.1 引言: 不能融合的传统人工智能模型	116
8.2 基于神经系统动力学的直接表达	116
8.3 研究目标、内容与方法	118
第 9 章 初级皮层提供的表征基础	121
9.1 引言	121
9.1.1 研究意义	122
9.1.2 国内外研究情况综述	123
9.1.3 本章工作	124

9.2 背景知识与生理依据	125
9.2.1 早期视觉系统简介及视皮层基本结构	125
9.2.2 视网膜水平层次结构简介	126
9.2.3 外膝体基本结构	131
9.2.4 初级视皮层结构简介	132
9.2.5 早期视觉系统通路结构简介	135
9.3 计算模型设计与实现	139
9.3.1 早期视觉系统模型	139
9.3.2 视网膜、外膝体层的模拟	140
9.3.3 方位柱的模拟实现	142
9.3.4 颜色通道的模拟实现	144
9.4 实验系统设计与分析	146
9.4.1 计算模型的设计验证	146
9.4.2 过程与结果的验证	149
9.4.3 高阶功能探索实验	155
9.5 讨论	157
第 10 章 进一步的研究	159
10.1 基于生物视觉结构的视知觉研究	159
10.2 言语的计算结构研究	161
参考文献	163

第 1 章 引言：人工智能中的一些基本问题

一旦某些心智功能被程序化了，人们很快就不把它看做“真正的思维”的一种本质成分。智能所固有的核心是存在于那些尚未程序化的东西之中：人工智能是尚未做到的事情。

—— Larry Tesler

1.1 人工智能受到的批评

人工智能是一个不断进步的领域，如无特别说明，本书所指的人工智能是指建立在“物理符号系统假设 (physical symbol systems hypothesis)”基础上的传统人工智能。人工智能发展 40 年，由于没能实现对它的种种乐观的预言，受到了很中肯的批评，其中的几例言简意赅、一针见血。一则来源于著名的计算机理论家 Hubert Dreyfus：“声称非单调推理是人工智能的进步，就如同声称爬上一棵树是向着登上月球的目标迈进了一步。” (Claiming progress in nonmonotonic reasoning is progress towards AI is like claiming that climbing a tree is progress towards reaching the moon.)

另一则寓言是用来形容计算机视觉的研究现状的：

“警察问一个在路灯下找东西的人：‘你在找什么？’

‘我在找钥匙。’

‘你的钥匙是丢在这儿了吗？’

‘不是。’

‘那你为什么要在这儿找呢？’

‘因为这儿比别处亮。’”

这则寓言可以扩展到对整个传统人工智能现状的描述，那就是“用已有的工具去解决好解决的问题” (赵南元，1994)。

在早期的一些狭窄的领域，符号主义方法取得了许多有限的成功 (如博弈、逻辑定理证明、专家系统等)，它们中的大多数被设计成针对特别的任务，而不是基于一个 general 的理论根基。这些早期的成功使人们信心十足，期望使用同样的人工智能 (artificial intelligence, AI) 技术最终能在像机器视觉、自动翻译、语音识别、常识表示等领域获得同样的成功，然而事实证明这将思维简单化了，且也过于乐观了。AI 方法中主要能触及的成功，如专家系统的开发，然而它们仅对一类相对狭

窄的问题是有效的。开发一个受限域的智能系统和通用的智能系统根本就是两回事，因为两者的主要区别不在知识量的多寡上，也就是说知识工程的方法无效。从实现智能系统的灵活性、适应性、健壮性与通用性角度来说，后者需要更多关于智能理论上的突破。

在人工智能的两种主要研究方法——结构模拟和功能模拟中，所采取的指导思想或哲学观是不同的：前者是对智能进行解释——怎样的器件结构和流程规则会导致某一特定智能行为的出现；而后者是对智能进行重演——根据已知的结构化知识来再现智能行为的外观。在文献 (Aparicio and Levine, 1994) 中，将这两者的区别总结得非常精辟，分别是以 “How we know” 和 “What we know” 为构造智能模型的出发点。可以说功能模拟仅是 “知其然，而不知其所以然”，它是在当前对生物智能的生理和心理结构研究成果远不足以支持工程模型开发的前提下的一种妥协性选择，但我们应当清楚地认识到我们的理想是 “解释智能”。在开发智能系统时应当尽可能地借鉴神经生理学和认知心理学的研究成果。也就是说，人工智能的研究不仅是计算机科学、自动控制等专业的事情，应该更多地考虑生理学家和心理学家的意见。

当然，还有一条重要而简单的原因就是人们对人工智能的期待在不断提高，就像本章开头所引用的话，致使这个领域的进展总是赶不上人们的希望，这种希望是科学进步的原动力。

1.2 人工智能中的几个基本问题

在文献 (Kirsh, 1991) 中作者首先问了几个问题：什么是人工智能的一个理论？它是一个结构还是一堆知识？它是否应该在一个抽象的模拟环境中通过研究它的性能来检测？我们是否能够先集中在一个小的问题集合或一个特定的分析层次上来研究人工智能？然后 Kirsh 提出了关于人工智能的五个核心问题，他具有强烈的认知科学背景，更多地是从心理学和哲学的视点，而不仅是从计算机科学的视点来考察智能模型的真实性和可行性，希望在方法学的高度框定人工智能的理论轨迹，他批评纯技术和实用主义的观点，所以具有深刻和长远的意义。他提出的五个基本问题如下。

1) 知识和概念化是否是人工智能的中心

此问题的正面论点是：Nilsson 认为 “人工智能问题的最重要部分就是发现一个适用的概念化知识结构”。智能需要可陈述性的知识和一些形如推理的计算，人工智能的核心便是研究世界的概念化，这个世界是智能系统进行认知活动时假设和使用的世界。人工智能的一个理论就是对一种认知技能所需知识的说明。人工智能的一个主要目标就是去发现认知行为的基本知识构成，这些基本知识就是那些

被认为是普遍真实的概念，就是那些我们知道同时我们也推断其余人也知道的概念，它们可以是常识也可以是领域知识。

2) 认知技能是可分离的吗

此问题的正面论点是：智能主体所面对的问题和任务能被形式化说明，因此也能在知识或概念水平上进行抽象地研究，这是一种有效的研究智能的方法。也即为了分析的需要，对智能系统的输入与输出都可以是形式化符号表示。至于这些概念化知识是如何建立在感知 — 运动基础上的过程可以与形式化过程分开来研究。

3) 认知的动力学过程是可语言化的吗

此问题的正面论点是：我们能用逻辑或某种现成的语言来追踪刻画一个处理认知任务的系统产生的信息状态的轨迹，也就是说无论是谓词演算还是如英语都可用作追踪认知过程中所做计算的语义。以逻辑学派的观点，一个智能主体在计算其下一步的行为时，他产生了一个用系统设计者的概念结构指派的，关于对象、功能和关系的信息状态轨迹。这个问题和问题 1) 是一脉相承的，承认它就等于承认知识的概念化的核心地位，且拓展到认为全部的知识都是可概念化和形式化的。

我们力求以一种 bottom-up 的方式来理解认知过程，而这种方法在处理涉及非常复杂和形式化的问题求解策略或推理机制时并没有引起人们多少兴趣，但我们必须首先确认一个事实：认知是基于神经系统的。也就是说，除了已有的对命题符号的操作之外，我们还能找到一个更接近于事物本质的实现方法——神经网络。对于认知过程的理解应当基于相继发生的神经活跃模式，而不能基于相继的可陈述性句子。这样做的目的是为了达到对诸如与真实环境的物理交流、活跃模式的动态特征等基本过程的更深入的理解。高级认知活动是众多低级神经活动的体现，那些问题求解策略、推理机制、语言等均是神经表示下的物理的真实。这些高级认知活动的符号形式化仅是一种抽象的描述，是站在概念和语言这一范畴上进行的抽象，那么显然这一符号形式化手段不能表示比概念更基本的认知层次，而这些认知层次对于概念及其语义的确立是至关重要的。认知不能被简约成对语言陈述的操纵和对毫无意义的纯语法结构的操纵。

4) 认知可在学习被分离的情况下进行研究吗

此问题的正面论点是：以纯自顶向下的方法，我们假设说明一个系统知道什么知识而不用说明系统是怎样知道这些知识的做法是可能的，即知识和知识获取能被分离开来。学习和概念化有足够的不同之处而可分开来研究。

5) 所有认知技能是否能建立在一个统一的认知结构上

此问题的正面论点：Newell 等认为认知基本上是在一个单一的结构上进行的多个程序的结果。太多的人工智能和认知科学家均致力于开发独立的表示和控制机制来解决特定的认知任务，每一个都有其各自倾向的计算模型，然而它们很少满足一个约束，就是它们能被综合进一个统一的、一致的认知理论中来。问题求解、

决策、记忆、学习、技能，甚至感知和运动行为都能在一个结构上运行，它们工作时只遵循较少的原则。心理学已经接近这样一个可能性，即统一的认知理论，使得在一个统一机制的系统上能产生人的所有认知行为。

上述问题是彼此密切相关的，不可能单独讨论其中某一个问题。本书所进行的研究分别就上述问题持有自己的立场，除第五个问题之外均持否定观点。第 2 章和第 3 章涉及问题 1)、2) 与 3)，第 4 章和第 5 章涉及问题 4) 与 5)。

1.3 人工智能需要一个更广泛的背景

回顾人工智能的诞生，能看到学科交叉的历史。在 20 世纪 40 年代和 50 年代早期，一场“控制论”革命抓住了一些天才科学家的想象 (Aparicio and Levine, 1994)。John Neumann、Warren McCulloch、Norbert Wiener 和其他一些学者领导了一场旨在考察人脑和当时尚属新发明的计算机之间类比的运动。这些智慧先驱当时激发的思想是存在于一个宽正面的多学科领域中，并不存在后来被普遍化的关于不同类型的分类 (符号主义与连接主义)。例如，图灵关于计算的定义就来自他关于人类记忆和控制过程的概念，当时这些思想盛行于心理学界和语言科学界。作为一个渴望将神经生物学成果应用于计算机设计的学者，Von Neumann 如果不是因为短期生病而未能参加 1956 年夏季 Dartmouth 会议的话，那么 AI 今天可能会另有一番风味了。在这个会议中处于主导地位的是像 John McCarthy 这样的逻辑学者，很明显没有神经生物学家参加。实际上，作为一个新博士的 Marvin Minsky 是很少的几个在场的持有“神经观点”的人，也许是部分受到这次会议所强调的思想的影响，他以后的一代工作都极大地偏离了生物学制约因素。但 Newell 和 Simon 这样的心理学家参加了这次会议，所以可以说 AI 在一开始就至少受到了心理学的影响。在以后的 20 至 30 年的时间里，AI 除了很浅薄的一点源自认知心理学和语言学的影响之外，基本上是顺着 Dartmouth 会议的精神发展的，几乎再没有进行真正的学科间的交流。随后在那些感兴趣于用生物组织来实现或对认知功能建模的人和那些只考虑设计机器实现这些功能而不管其生物基础实质的人之间，发生了一场分裂。在当时由于对问题复杂程度和对于可供使用的技术、可供借鉴的神经生物学知识的缺乏，这一分裂被证明是正确的，这也可由整整 40 年经典人工智能理论与应用的实际进展所佐证。这一分裂的影响波及至今。在当前出现的连接方法和符号方法相互妥协的势头下，双方阵营都能客观评价自己的优势与缺点，不再以己之长较人之短，并放弃以此结果进行断言的草率态度。应该回到分裂发生前的那种相互借鉴的立场上。

人工智能不是对计算机进行研究，而是对思维和行为所体现的智能进行研究，计算机是实现智能理论程序化的工具。所有对人工智能所下的定义都指出了两点：

第一，了解人类智能的特点；第二，使机器具有更多的智能特性。

人工智能是以认知科学为其学科基础的，它应该建立在心理学、神经科学、语言学、人类学和哲学这个广泛的基础之上，虽然人工智能在计算机科学和控制科学的推动下取得了举世瞩目的成就，但后两者的学科基础作为人工智能的学科基础是远远不够的，我们尤其应该看到当人工智能汲取系统论、信息论和控制论的思想精髓之后，它更大的发展将取决于对人类智能本质的认识上的进步。当我们对一个有某种智能的系统进行其认知意义的考察时，人工智能就超越了工程科学的范围。而认知科学作为一个高度交叉的学科，所涉及的问题不是一般的复杂^①，但唯独有将这些问题落实到人工智能模拟算法和体系结构的设计与实现上才会使对思维问题的探索变得细致、严谨、统一和深刻。

传统人工智能讨论的问题是博弈、专家系统等层次的问题；神经科学讨论的是神经元、回路等层次的问题。在它们之间存在一个层次上的鸿沟，致使它们在方法、模型、策略等问题上难觅共同点，它们的研究目标定位是宏观和微观，互不相同。能填补这一鸿沟的是：神经心理学、神经病理学、认知神经科学、生理心理学等将宏观和微观联系起来的学科。我们会发现在抽象知识表示层与具体的实现层间寻找的桥梁是问题“如何实现知识的语义表示和直接表示”。心理学 (Matlin,1992) 对认知的定义是：认知是知识获取、存储、回忆与使用过程中包含的所有精神活动 (Cognition are all mental activities including acquisition、storage、retrieval and use of knowledge)。在认知背景下的人工智能，就应该以此为主线，以解释智能而非再现智能为目标。通过借鉴心理学与神经科学的研究成果，将思维过程、认知过程中的心理功能分析步骤与神经实现机制对应起来，并借用并行分布式模型来实现机器智能的目标。这当然是一个极富前瞻性和异常艰难的任务！图 1.1 展示这些借鉴关系。

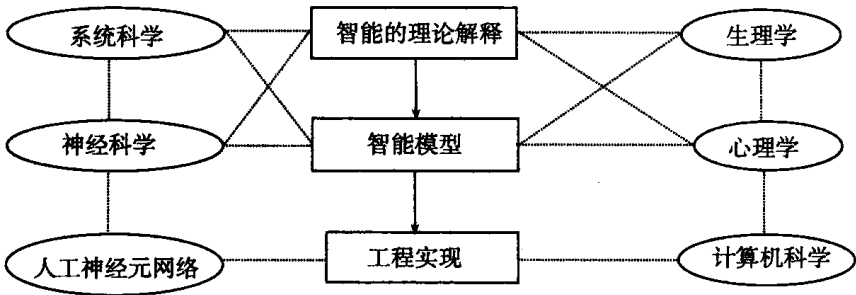


图 1.1 人工智能的借鉴关系

要做好学科间的交流必须：① 设定共同的讨论水平层次；② 各学科知识的综

① 人类智能的奥秘与物质的本质、宇宙的起源和生命的本质一道成为当今人类最感兴趣的四大问题。

合；③ 讨论一个问题时要从问题本身出发，不要带有不同学科各自的理论与方法背景。比如认识论关心：认识的对象与来源、认识的本质、认识的过程、认识的规律与认识的形式；神经科学关心：神经组织的功能；人工智能关心：认知的过程、思维、知识、算法、PDP 等。它们三者“神经活动”这个层次上能找到探讨问题的共同基础。

1.4 智能的心理表现与生理基础

认知心理学采用信息加工的观点，把人脑看成信息加工系统，认为认知就是信息加工。信息加工观点提出的基本问题是：① 信息加工通过哪些阶段？② 人类心理中信息是以什么形式表示的？心理过程能最终还原成生理过程和物理过程，但必须要处理好还原与整合的矛盾、突现与渐变的关系。图 1.2 表示认知的行为表现与内在机理间有一种相互需求和适应。

现代心理学理论中有一个决定论原理 (邵郊, 1996)，其基本内容是：人类的一切精神活动都是由一定的生理机能所决定的，精神与神经是互相统一的。也就是说，对应到具体问题上，如形象思维的存在，能从人脑中找到物质基础。

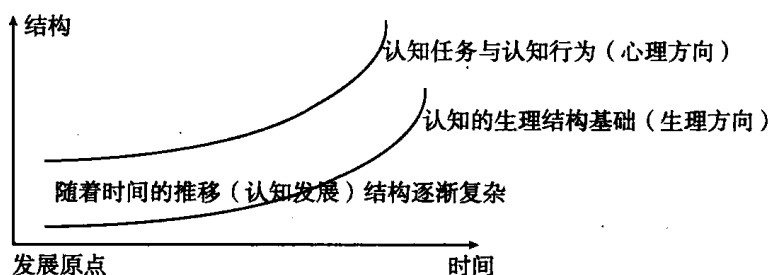


图 1.2 两条主线不是平行的对应关系：表现 ↔ 机理，只有综合起来才能找到认知心理的生理解释

如果我们扩大一点讨论的范围，示意图 1.3 的原理是普遍存在的 (危辉, 1998)。计算机科学研究人员，对纯生理的成果不一定关心，纯心理的描述又缺乏对算法细

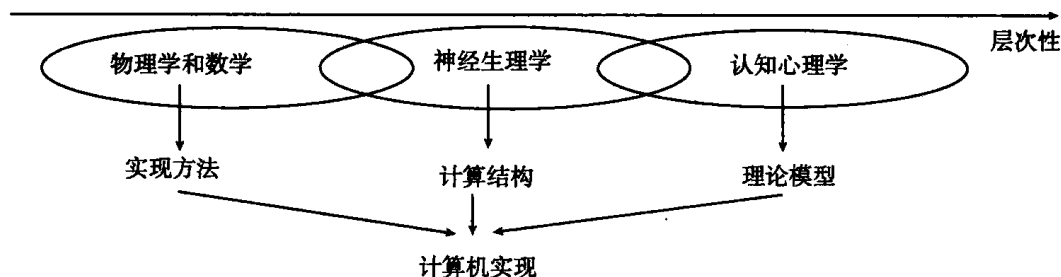


图 1.3 智能模型的基础

节的支持,二者结合起来就可能找到恰当的计算结构,参见示意图 1.3,表示学科间的层次关系和交叉性。

1.5 本书的逻辑演进关系

本书按下述逻辑关系组织。第 1 章说明由当前人工智能的成果总结得出的几个基本问题,以及围绕它们进行的批评,作者认为这些批评产生的原因是形式化方法作了太多的省略和抽象,以致远离了思维的客观真实。要想使人工智能有更广泛和更深层次的应用就必须了解心智的本质,而要模拟人的认知行为就必须深入进行学科间的借鉴,把认知科学作为人工智能的学科基础。第 2 章说明形式推理的不足,要想实现由形式推理向语义推理的转变必须认识到我们努力的方向在哪,而认知行为的相关性正是智能行为模拟时所不能回避的问题,这种相关性的表现在哪些方面以及如果在设计人工智能模型时考虑到这些相关性会对模型产生怎样的影响。第 3 章说明由认知相关性要求统一认知结构是重要的,它的核心是表示机制,是建立在知觉表示基础上的直接信息表示,并且作者进行了基于视觉表象和结构学习的人工神经网络表示模型的实践。第 4 章说明由认知相关性的要求要用一种系统的、整体性的观点来看待智能的性质,要用发展的观点来进行智能模型的构造,通过发展心理学可以得到许多关于智能模型构造过程和方法的启示。第 5 章说明语言是智能行为中最具系统性、整体性和发展性的代表,关于言语的计算结构研究能最全面地体现关于人工智能模型的设计如何建立在认知科学的基础上,以及说明作者对言语的心理模型的思考,并且给出了两个关于基于知觉模式的分词算法和基于归纳和参照的词结构操作语义的学习算法。此处要增加第 6~8 章的内容,最后一章说明针对上述问题的进一步的研究。

第2章 认知相关性及其在人工智能背景下的意义

关于人类的知识可以提出两个问题：第一，我们知道什么？第二，我们是怎样知道这些知识的？回答第一个问题的是科学……但是对于第二个问题——即我们怎样得到我们的知识——心理学在各门科学中却最为重要……我们的知识的全部素材都是由个别人生活中的心理事件构成的。因此，在这个领域内，心理学占有最高的地位。

——Bertrand Russell,《人类的知识》，P67

2.1 从形式推理到语义推理

在所有的人工智能领域都会遇到问题求解，解决问题的方法包括归纳、类比、演绎等多种。这里所说的推理并不严格地指逻辑学中的演绎，它是泛指建立在偏序关系上的思维的前继和后续过程，其中的偏序关系可以是因果关系、时序关系、顺序关系、激励关系、队列关系等，只要是能用有限状态自动机的状态转换机制刻画的思维或说问题求解方式，推理就映射为状态之间的转换。这样问题求解的方法就可以归为这种广义的推理机制。

2.1.1 形式推理及其困难

为了明确在逻辑中将思维的模式进行了何种程度的抽象，让我们回顾一个说明什么是基于形式的推理的例子（汉密尔顿，1986）。

例 2.1 如果苏格拉底是人，那么苏格拉底有死。

苏格拉底是人。

所以 苏格拉底有死。

这是逻辑上适当的论证。但是考察论证

苏格拉底是人。

所以 苏格拉底有死。

结论也可以认为是由前提推得的，然而其所以如此是由于词“人”和“有死”的意义，而不是由于单纯的逻辑演绎。让我们用符号来表示这些论证：

第一式： $A \rightarrow B, A$ 。所以 B 。

第二式： A 。所以 B 。

使第一式有效的在于“形式”。任何具有同一形式的论证应该都有效，这是关于“如果……那么……”这样命题的逻辑直觉。第二式有效依赖于具体的语义。

逻辑学研究命题的形式要甚于研究具体的命题，然而对人工智能而言这样做是不够的，这样做只是推迟了困难的出现：落实到知识工程上知识获取的困难成为瓶颈。在人工智能近半个世纪的历史中，逻辑学派占有统治地位，它的理论基础是现代形式逻辑，它的特点是(郑文辉, 1994)：采用形式化的演算方法，只涉及符号和形式，而不考虑符号的意义；逻辑规律用符号语言表述；形式系统需要丰富而充实的定理体系；形式逻辑已彻底摆脱了认识论、心理学，最后成为纯粹的形式化问题。由于使一个形式化的论证过程得以继续的演绎规则，如分离规则、归结反演等，都是相对现成的，所以知识与概念化问题成为人工智能的核心。这个思想在知识工程上得到了最好的体现：认知是推理，智能由两部分构成，一个可陈述性的知识库和一个推理机，推理机通常是相对简单和独立于领域的，它只负责将当前的状态在知识库的某些片段的指导下按规则得到结果，灵活性的要求使得知识的表达必须采取明呈、可陈述的方式。

然而将智能形式化的方法遇到了很大的反驳：① Brooks 置疑，为什么假设智能就一定需要概念呢？(否定这个问题也就意味着否定了 1.2 节中的问题 (3))；② 智能机器设计者认为的概念化、智能机器保有的概念系统、概念系统被形式化的方法，这三者是不同的，其不同之处在于一个深刻的哲学问题：怎样才叫掌握了概念 (grasp a concept)？我们不能认为在机器的记忆部件中编码了和机器的设计者相同的概念化结构就足以掌握概念了 (Aparicio and Levine, 1994)，也可以说是一个智能系统拥有了知识库，不等于它就拥有了这些知识所表达的信念，前者是一个指派，后者是需要系统具备概念的语义基础的。然而往往就默认系统包含了知识库也就有了知识库所蕴涵的信念。逻辑学派认为坐实概念 (grounding concepts) 和创造概念化 (inventing conceptualizations) 是可以分开来的，后者才是人工智能至高至上的任务：study agents' concepts structure without concern for their grounding。然而仅有符号是不足以掌握概念的，这需要一个关于此概念的行为的、感知的、推理的技能的大量背景知识，使得需要知道这个概念如何被参考，它如何在推理、感知和行为中被使用的人工智能任务都能完成，而不仅仅是完成区分概念或表述它们间的公理化关系的任务。

例如，在一个知识库中用常元 a 、 b 、 c 、 d 、 e 分别表示概念钥匙、斧子、手枪、炸药和密码；用谓词常元 $\text{Can_open_door}(x)$ 表示用 x 可打开一扇门，那么命题 $\text{Can_open_door}(a)$ 、 $\text{Can_open_door}(b)$ 、 $\text{Can_open_door}(c)$ 、 $\text{Can_open_door}(d)$ 和 $\text{Can_open_door}(e)$ 都是成立的。常元 a 、 b 、 c 、 d 、 e 用符号的不同表示了钥匙、斧子、手枪、炸药和密码间概念的不同，这当然丢失了大部分的语义信息；自然从命