



华中师范大学出版基金丛书
学术著作系列

认知科学导论

熊哲宏 著

C B J J

出版社

华中师范大学



认知科学导论

熊哲宏 著

C B J J

2002 · 武汉

出版社

华中师范大学



号 11 千 登 微 (带)

· 国家社会科学基金资助项目 ·

认知科学导论

熊哲宏 著

ISBN 7-307-04820-4

半 译

升 贵
升 贵

字 千 10
幅 年

本 平
次 函
数 中

华中师范大学出版社

书 图 中 华 师 范 大 学 出 版 社 印 刷 厂 印 刷

(鄂) 新登字 11 号

图书在版编目 (CIP) 数据

认知科学导论/熊哲宏 著.

—武汉: 华中师范大学出版社, 2002. 9

ISBN 7-5622-2549-4/B·64

I. 认…

II. 熊…

III. 认知科学—概论

IV. B.842.1

中国版本图书馆 CIP 数据核字 (2002) 第 0485014 号

认知科学导论

© 熊哲宏 著

华中师范大学出版社出版发行

(武昌桂子山 邮编: 430079)

新华书店湖北发行所经销

华中理工大学印刷厂印刷

责任编辑: 刘晓嘉

责任校对: 罗少琳

封面设计: 新视点

督 印: 姜勇华

开本: 850×1168 1/32

版次: 2002年9月第1版

印数: 1-1000

印张: 16 字数: 401千字

2002年9月第1次印刷

定价: 30.00元

本书如有印装质量问题, 可向承印厂调换。

引言

本书写作的主要目的至少有两个：一是尽可能全面系统地综合、概括、提炼认知科学自诞生以来，特别是 20 世纪 90 年代以来发展的新成就；二是在此基础上进一步探讨制约认知科学发展的深层哲学问题。

我们写作的第一个目的的重要性和意义是不言而喻的。认知科学一般被认为是 20 世纪世界科学标志性的新兴研究门类。它作为探究人脑或心智工作机制的前沿性尖端学科，已经引起了全世界科学家们的广泛关注。正如中国学者白春礼在《21 世纪科技发展趋势》一文中指出的那样：“研究最复杂客体的心理学和认知科学将成为后来崛起的高峰，并在更远一些年代里可能成为新的科学革命的中心。人脑的工作机理将是 21 世纪认知科学的前沿。正像工业革命用机器代替和扩展了人类的体力劳动能力一样，这方面的科学突破，将极大地扩展人类脑力劳动的能力。”（白春礼，1996）

但是，对于一般读者甚至某些特定领域的专家来说，他们对认知科学的总体概貌还不是十分清楚的。这除了认知科学本身还不够成熟的原因之外，还与迄今为止我们对认知科学尚缺乏统一的、整合性的探讨有关。虽然某些领域的专家也的确在从事认知科学意义上的研究，但他们往往在特定的专业领域各自为战，相互知道对方的工作进展很少，彼此之间缺乏沟通与对话，这种情况在中国还比较突出。本书试图为各个领域的专家，如心理学家，计算机科学家、神经科学家、语言学家、人类学家和哲学家等，勾画出一幅认知科学的总体概貌。由于本课题有相当的难

度，我们采取了如下研究策略：首先从包罗万象的认知科学文献中梳理出贯穿认知科学各个领域的基本问题或主要问题。发现认知科学问题是我们探讨的第一步。然后挖掘认知科学解决这些问题的基本前提或基本假定（例如人脑和计算机都是“物理符号系统”就是一个基本假定），并在此基础上理清认知科学为说明认知问题所构造的各种理论模型（包括贯穿于其中的研究“范式”）。总起来说，我们写作的第一个目的就是要使读者明确认知科学的基本问题，基本前提或假定，基本概念和理论框架。为此，本书忽略了认知科学的某些具体成果和具体观点。

本书的第二个目的导源于我们对认知科学怎样发展的深层理论思考。作为统一的、整合的认知科学，它与一般的经验实证科学不同。至少可以说，它的问题、方法、模型和理论都带有普遍性。例如，假如真有一种“通用智能理论”的话，那么它必须既能解释天然系统（人脑）又能解释人工系统（计算机）。这样，就使认知科学的研究带有一定程度的抽象理论层次。换句话说，真正的认知科学问题都内在地、固有地包含着它自身的哲学问题。于是在认知科学界，“认知科学的哲学问题”（Philosophical Issues in Cognitive Science）一词便不径而走。也有人喜欢用“认知科学的基础”（Foundations of Cognitive Science）这一术语。二者要研究和解决的问题是一致的。正是由于认知科学摆脱不了哲学问题，加之它倡导一种统一的、整合性的研究方式，以致普特南说“认知科学像哲学”（Peter Baumgartner *et al.*, 1995, p. 183）。当然，“像哲学”并不等于就是哲学。它不过是表明哲学思维方式对解决认知科学问题是有积极意义的。为此，本书不仅在每章的最后一节专门谈哲学问题，而且在叙述体例或章节安排——当然还够不上成熟的或理想的“理论框架”——上也体现了某种哲学的考虑。

至于本书的基本写作思路可概括如下：首先追溯认知科学的

起源和发展，弄清它的对象和范围、主要问题和研究方法；然后在指明认知科学的指导性隐喻——“计算机隐喻”——的前提下探讨认知科学的两大“范式”，并具体分析这两大范式如何解决某些认知科学的普遍问题（如智能的本质，常识知识问题，计算机和计算的本质，认知科学解释以及意识的地位）；最后展望认知科学的未来。

2	计算机隐喻	(75)
2.1	计算机模拟	(75)
2.1.1	计算机模拟的范例	(76)
2.1.2	两种“范式”的共同隐喻	(80)
2.1.3	人脑的好隐喻?	(86)
2.2	计算机的特征	(92)
(1) 2.2.1	计算机和计算的概念	(93)
(1) 2.2.2	人脑是计算机?	(99)
2.3	计算机与人性	(103)
(d) 2.3.1	新“智能主体”	(103)
2.3.2	新“计算客体”	(107)
(e) 2.3.3	“人机交互”的世界	(110)
3	符号系统范式	(113)
3.1	“物理符号系统假设”	(114)
(71) 3.1.1	“PSSH”及附带推论	(114)
(91) 3.1.2	人和计算机的信息处理与结构	(119)
(55) 3.1.3	人类思维的计算机模型——认知模型	(123)
3.2	计算心理学的性质	(131)
(85) 3.2.1	计算心理学的理论模型	(132)
(74) 3.2.2	产生式系统方法学	(135)
3.3	符号系统范式的地位和前景	(138)
(12) 3.3.1	PSSH的性质和应用	(139)
(52) 3.3.2	符号处理范式仍然具有生命力	(144)
(10) 3.3.3	联结主义比PSSH更有价值	(154)
3.4	符号系统范式的主要哲学问题	(164)
(80) 3.4.1	认知主义的迷雾	(164)
(57) 3.4.2	“表征”的种种难题	(168)

4	联结主义范式	(178)
4.1	联结主义模型的基本特征	(179)
4.1.1	人工神经网络发展简况	(180)
4.1.2	联结主义系统的构成和主要特点	(185)
4.2	联结主义模型的理论优势	(190)
4.2.1	“构架”(“硬件”)事关重大	(191)
4.2.2	神经网络的“自上而下方法”	(197)
4.2.3	联结主义“计算模型”的合理性	(201)
4.2.4	来自冷静的观察者的思考	(208)
4.3	人工神经网络的哲学涵义	(212)
4.3.1	联结主义是通往心智的皇家大道?	(212)
4.3.2	联结主义模型是大脑模型吗?	(218)
4.3.3	联结主义排除常识心理学吗?	(221)
5	“图灵测验”与智能的本质	(225)
5.1	“图灵测验”与机器思维	(225)
5.1.1	图灵“模拟游戏”的思想实验	(226)
5.1.2	图灵对“机器不能思维”论点的反驳	(227)
5.2	关于“图灵测验”有效性的论争	(233)
5.2.1	“图灵测验”是有用的测验观念	(234)
5.2.2	“图灵测验”是无用的测验观念	(237)
5.3	人工智能与人的智能	(242)
5.3.1	“通用智能理论”的特征	(243)
5.3.2	智能与意识	(247)
5.3.3	智能与理解力	(254)
5.3.4	智能测验的效度问题	(258)
6	常识知识问题	(265)
6.1	常识与“常识知识问题”	(265)

- 6.1.1 常识的概念 (265)
- 6.1.2 “常识知识问题”是个什么问题? (268)
- 6.2 常识知识问题的尝试性探索 (270)
- 6.2.1 构造“常识逻辑” (270)
- 6.2.2 用“常识”装备 Cyc 系统 (273)
- 6.3 常识知识问题的性质 (277)
- 6.3.1 威伦斯基:“没有人能指出常识问题是什么” (277)
- 6.3.2 德雷福斯:“常识知识问题将引起人工智能
的崩溃” (280)
- 6.3.3 豪杰兰德:“勒内特的计划毫无希望” (282)
- 6.4 “常识知识问题”的主要哲学问题 (284)
- 6.4.1 人工智能应不应该使用“逻辑”? (285)
- 6.4.2 “基于规则的专家系统”概念是合理的吗? (291)
- 6.4.3 认知科学能否形成关于“相关性”的普遍
理论? (299)
- 7 “中文屋论证”与第一人称观点 (305)
- 7.1 塞尔的“中文屋论证” (305)
- 7.1.1 早期论证:句法不足以满足语义 (305)
- 7.1.2 近期论证:从第一人称的观点看 (316)
- 7.2 对塞尔“中文屋论证”的各种反驳 (319)
- 7.2.1 丹尼特:“中文屋不是一种论证” (319)
- 7.2.2 霍夫施塔特:“塞尔的魔鬼”懂中文 (322)
- 7.2.3 福多:“句法特性能保持语义特性” (325)
- 7.2.4 彭罗斯:“像教义一样地令人可疑” (328)
- 7.3 “中文屋论证”隐含的认知科学哲学问题 (330)
- 7.3.1 形式符号操作是不是“信息处理”? (330)
- 7.3.2 放弃还是坚持功能主义? (334)

8	认知科学解释	(343)
8.1	解释的层次	(343)
8.1.1	马尔的三层次解释理论	(344)
8.1.2	经典认知科学的解释层次	(348)
8.1.3	联结主义的解释层次	(355)
8.2	解释的模型	(359)
8.2.1	表征模型	(360)
8.2.2	计算模型	(365)
8.2.3	解释模型中本体论和现象学之分野	(369)
8.3	解释的机制	(371)
8.3.1	因果机制	(372)
8.3.2	还原机制	(383)
9	意识在心智研究中的地位	(395)
9.1	认知科学背景下的意识问题	(395)
9.1.1	意识研究的方法学问题	(397)
9.1.2	“他人心理”问题	(405)
9.1.3	“可感受特性”或主观体验问题	(407)
9.1.4	“人工意识”问题	(410)
9.2	计算模型中的意识	(412)
9.2.1	符号处理式的意识模型	(413)
9.2.2	计算模型能解释意识吗?	(416)
9.3	网络模型中的意识	(424)
9.3.1	意识的脑并行模型	(424)
9.3.2	基于功能模块和中央系统的模型	(427)
9.3.3	神经认知研究模型	(429)
10	认知科学的未来	(432)
10.1	面临的挑战	(432)

10.1.1	生态学的挑战	(433)
10.1.2	社会学的挑战	(437)
10.1.3	现象学的挑战	(439)
10.1.4	解释学的挑战	(441)
10.2	新近的趋势	(446)
10.2.1	日益关注认知的社会方面	(446)
10.2.2	进化和发展的新视野	(451)
10.2.3	统一认知构架的探索	(458)
10.3	未来的走向	(462)
10.3.1	“PSSH”还能走多远	(463)
10.3.2	联结主义曙光初露	(470)
10.3.3	两大“范式”相互融合的可能性	(479)
	参考文献	(489)
	后记	(497)

1 认知科学的性质

认知科学作为 20 世纪世界科学的新兴研究门类，作为探究人脑或心智工作原理的前沿性尖端学科，它的产生决不是偶然的，而是有着深远的社会经济、科学技术（特别是计算技术）和思想文化的背景。认知科学的历史还不算太长，甚至可以说是一门相当年轻的科学，但它已经历了艰难曲折的发展道路和阶段，为揭示人脑的工作机制这一最大的宇宙之谜作出了不可磨灭的贡献。认知科学虽然还远未臻成熟，但它已开拓了自己独特的研究对象和领域，形成了自己独有的研究方法。时至今日，任何一个严肃的科学家——不管他工作在哪个专门领域——都不能无视认知科学的存在。事实上，越来越多的科学家都自觉或不自觉地加入了认知科学的研究阵营，并为拥有或获得“认知科学家”的头衔而自豪。

本章从追溯认知科学的起源和发展入手，继而较为详尽地探讨认知科学的对象和范围，并在明确认知科学的学科性质基础上概述认知科学的研究方法，使读者对认知科学这门 20 世纪标志性的新学科有一个总体的了解。

1.1 认知科学的起源

认知科学的起源是复杂的。一般地说，在 20 世纪 40 年代到 60 年代，认知科学的基本观点最初散见于各自分离的特殊学科之中，然后逐步综合统一起来。当然，在认知科学的起源中，各门科学的地位和作用是不一样的。按照黎黑的观点，“从长远的观点来看，最重要的认知研究源自数学和电子工程，而且很少甚

至没有涉及心理学及其问题。随着第二次世界大战期间现代数字计算机的发明,“人工智能”(artificial intelligence)领域应运而生”。(黎黑,李维译,1998,第703页)对认知科学的起源作出主要贡献的学科和其他因素有:计算机的发明和被称为“人工智能”研究方案的出现;心智哲学中以普特南和福多为代表的“功能主义”理论的确立;心理学和语言学反驳以斯金纳为代表的激进行为主义的“认知革命”,等等。美国认知科学家伽德纳(H.Gardner)在其《心智的新科学:认知革命史》(1985)中,对认知科学的起源和发展作了很好的论述。(伽德纳,张锦等译,1991)

1.1.1 计算机的发明和“人工智能”的产生

美国哲学家塞尔(John R.Searle)指出:“在认知科学史中,计算机是关键性的。事实上,如果没有数字计算机,就不会有认知科学。”(Peter Baumgartner *et al.*, 1995, p. 204) 认知科学家已基本达成共识的一个看法是,自第二次世界大战以来数字计算机的广泛使用,在我们形成思考心智的新方式——现今被称为是“认知科学的中心前提”——方面起了关键作用:从“信息处理”的观点研究人脑是可能的。计算机在认知科学所包含的所有学科中已成为心智的强有力隐喻。

众所周知,英国数学家、逻辑学家图灵(A.M.Turing)在第二次世界大战期间为计算机制造特别是计算机理论作出过不可磨灭的贡献。1939年二战爆发后,图灵开始为英国政府工作,协助破译德军的无线电密码。1940年,图灵和他的密码破译小组建造了一台原始的计算机。“他们不动声色地坐在战争游戏的赌桌前,用计算机审视着希特勒手中的每一张牌。迷惑不解的德军将军们从未真正搞清:他们的超级战车究竟在哪里出了问题。……图灵的工作具有关乎英国生死存亡的重要意义,连丘吉

尔首相都深切关注着他的工作，并亲自探望慰问了他。1942年整整1年的时间里，图灵小组用他们的原始设备破获了大量来自德国的情报，并从中找到了他们迫切需要更新设备的新理由：一些截获的情报谈到了辎重队、宇宙射线簇射和氰化物的工业产量。”（罗林斯，刘玲等译，1998，第19~20页）

早在1935年，图灵便开始着手构思一台抽象的机器——理论上的万能机器，并定义了“计算”的逻辑基础。他使用了“计算机”一词，但在图灵所处的年代，这个术语意指受雇于一些公司进行计算工作的男人或女人（实际上常常是女人）。

图灵设想了这样一台计算机怎样执行任务。计算机有着一支铅笔和一张无限长的以竖条分隔成一框一框的纸。图灵称这种纸带为“磁带”，每框为一个“方格”。他设想：任何时刻计算机的行为由它所观察到的符号和当时它的“心智状态”来决定。基于这种观察和心智状态，计算机做以下三件事之一：移动到相邻的方格；改变当前的方格；或者进入新的心智状态。图灵假定计算机使用的“字母表”是有限的，因为它无法区分无限数目的字母。他还假定计算机可能的“心智状态”也是有限的。

在图灵看来，计算机程序由一个精确的指令集组成，这些指令交由与观察任务有关的人执行。比如，“如果所读的方格中是0，则用1代替，然后右移一个方格”。他假设“可计算”的概念——他还证明了存在“不可计算”的问题——和他所设想的计算机处理这一长卷纸带是相称的。到目前为止，就任何制造出来的计算机而言，图灵的设想或观点都是正确的。这样，图灵便确立了现今称为“图灵机”（Turing machine）的概念：一种想象的、能计算任何严格可定义的计算过程或算法的计算机。西蒙对此评论说：“1936年英国逻辑学家图灵把现在所知的处理机称为图灵机时，他通过演示机器如何操作符号而完成了向形式化迈进的第一步。”（西蒙，曹南燕译，1998，第250页）

图灵还最早洞悉了“模拟”的思想：只要我们对某一过程的了解足够透彻，足以用计算机语言作出完整的描述，那么计算机就能模仿这一过程。所以，只要我们能将任何一台机器的工作完全描述在一张行为表上（某种情况下这样做，某种情况下那样做），那么，只要我们让计算机按照那张行为表亦步亦趋地去做，就能使计算机模拟那台机器。

根据“图灵机”的概念，与其他的机器不同，计算机的物理部件似乎并不特别重要，一旦我们赋予它创造自己所能理解的随机信息的能力，或换种说法，一旦它能以某种方式进行自我交流，就成了可编程控制的机器。通过自我更新，这就变成了一台万能信息处理机。图灵深信：一旦我们表达出任何一个我们可描述的问题，并设计出解决这一问题的详细步骤，我们就能利用计算机来解决这一问题。正因为计算机能实现我们精确描述出来的任何事情（能运作任何“软件”），它就不是一般的“机器”，而是处理任何信息的“万能机”。

随着数字计算机的发明，人们在关于人类和机器的概念上发生了一场革命。西蒙（H.A.Simon）在回顾自己作为人工智能的创始人之一时写道：“这一突而且永久的变化是因为阿伦·纽厄尔、克利夫·肖和我隐约看到了当时刚出现在公众面前的电子计算机的革命性应用。我们抓住了机会，把计算机作为一种通用符号处理机（因此也用于思维）而不只是快速算术机来使用。1955年底我们发明了程序计算机使用的LIST处理语言，并用这种语言创造了逻辑理论家——第一个通过选择性搜索解决非数字问题的计算机程序。正是这两项成果使我们被定为人工智能的创始人。”（西蒙，曹南燕译，1998，第243~244页）

把新出现的数字计算机看作通用符号处理机，是因为这些先驱们发现了人脑和计算机在工作原理上的根本相似性：即二者能够通过使用二进制编码来运行。计算机具有或者让电流通过或者