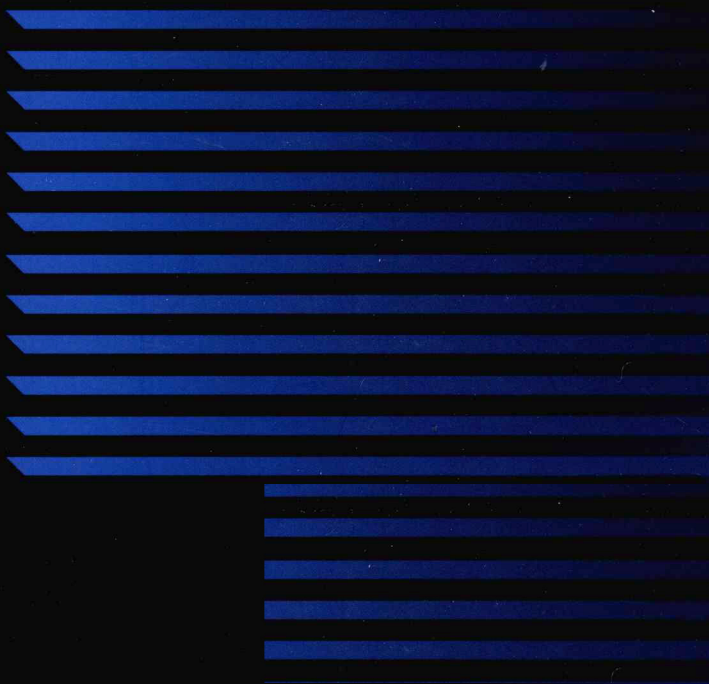


YR 语/言/与/认/知/译/丛

MINDS, BRAINS, COMPUTERS:
AN HISTORICAL INTRODUCTION TO THE FOUNDATIONS
OF COGNITIVE SCIENCE

心智、大脑与计算机： 认知科学创立史导论

◎ [美] R. M. 哈尼什 著
王 森 李鹏鑫 译



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

YR

语/言/与/认/知/译/丛

黄华新 盛晓明 主编

MINDS, BRAINS, COMPUTERS:
AN HISTORICAL INTRODUCTION TO THE FOUNDATIONS
OF COGNITIVE SCIENCE

心智、大脑与计算机： 认知科学创立史导论

© [美] R. M. 哈尼什 著
王 森 李鹏鑫 译



ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

图书在版编目 (CIP) 数据

心智、大脑与计算机:认知科学创立史导论 / (美)
R. M. 哈尼什著;王森,李鹏鑫译. —杭州:浙江大学
出版社,2010. 8

Minds, Brains, Computers: An Historical
Introduction to the Foundations of Cognitive
Science

ISBN 978-7-308-07273-1

I. ①心… II. ①哈… ②王… ③李… III. ①认知科学—研究
IV. ①B842. 1-09

中国版本图书馆 CIP 数据核字 (2010) 第 160502 号

浙江省版权局著作权合同登记图字:11-2008-52 号

This edition is published by arrangement with **Blackwell Publishing Ltd**, Oxford.
Translated by **Zhejiang University Press** from the original English language version.
Responsibility of the accuracy of the translation rests solely with **Zhejiang University
Press** and is not the responsibility of **Blackwell Publishing Ltd**.

心智、大脑与计算机:认知科学创立史导论 Minds, Brains, Computers: An Historical Introduction to the Foundations of Cognitive Science

[美] R. M. 哈尼什 著 王 森 李鹏鑫 译

策 划 曾建林
责任编辑 田 华
封面设计 刘依群
出版发行 浙江大学出版社
(杭州市天目山路 148 号 邮政编码 310007)
(网址: <http://www.zjupress.com>)
排 版 杭州中大图文设计有限公司
印 刷 德清县第二印刷厂
开 本 710mm×1000mm 1/16
印 张 27.75
字 数 500 千
版 印 次 2010 年 8 月第 1 版 2010 年 8 月第 1 次印刷
书 号 ISBN 978-7-308-07273-1
定 价 65.00 元

版权所有 翻印必究 印装差错 负责调换

浙江大学出版社发行部邮购电话 (0571)88925591

“语言与认知译丛”总序

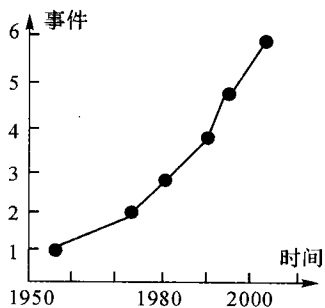
人类的心智(mind)和行为也许是宇宙间最顶端、最复杂也是最奇异的现象了,但人类只有通过自身的心智和行为才能认识和理解自己。无怪乎美国著名的认知神经科学家达玛西奥(A. Damasio)在研究意识时发出这样的感叹:“还有什么比知道如何知道更困难的事情呢?正因为我们有意识,才使我们能够,甚至不可避免地要对意识提出疑问,还有什么比认识到这一点更让人惊异和迷乱的呢?”“知道如何知道”——这正是认知科学的根本任务,而且也是促使其从哲学认识论中萌芽并最终在当代的哲学—科学研究中枝繁叶茂的根本动力。

认知研究已成为当前世界大国国家科技战略特别关注的领域之一。一个日益普遍的看法是:对心智的科学认识将在人类认识自身、科学技术、医学发展、经济增长、社会安全、人类幸福和生活品质的提高等人类和国家利益方面产生革命性的影响!世界众多一流大学或相应机构都在这个领域进行着你追我赶的研究,力图率先取得原创性的成果;加强和促进认知科学的发展同样符合我国的国家科技战略目标。《国家中长期(2006—2020年)科学和技术发展规划纲要》将“脑科学和认知科学”列为8个基础前沿研究领域之一,而且加快了对认知科学的资助和研究机构的规划部署。自“985工程”一期和二期实施以来,相继有一些高等院校和科研院所建立了以认知研究为重点的研究机构。浙江大学语言与认知研究中心(CSLC)就是“985工程”二期面向认知研究的人文社会科学与自然科学兼容的哲学社会科学创新基地之一。

认知科学有“一个长的过去,但只有一个相对短的历史”。也许正因为其历史短暂,其发展态势就显得尤为迅捷。自20世纪50年代“认知革命”发

生以来,认知科学一直处于高速发展的阶段。图中列出的一些重要的学术事件清楚地展示了这一点。面对这种情势,CSLC自项目启动伊始就怀有强烈的紧迫感。然而另一方面,当前认知科学的研究局面斑驳

1. “认知革命”(1956)
2. “认知科学期刊”(1977)
Sloan 报告(1978)
认知科学协会(1979)
3. “第二代认知科学”的兴起
(20世纪80年代)
4. 脑十年(1990—1999)
5. NBIC 和人类认知组计划
(2000)
6. “心智十年”倡议(2007)



加速发展的认知科学

陆离,这是历史上任何一个学科在其发展中都不曾有过的。至今认知科学还没有一个公认的统一的学科边界,还处在统一范式形成的前夜:研究的基本观念、维度、问题域和方法都复杂多样。为了在这个驳杂的局面中明确定位,形成特色,我们认为必须对当前认知研究的格局和趋势有一个较为全面的认识,从而根据自己的优势,在权衡慎思后提出自己的问题并开展深度研究,为推动认知科学在我国的发展尽自己的职责。基于这个考量,CSLC决定选译一些认知研究著作,作为系列丛书连续出版。对选译的著作,CSLC的设想非常简明:(1)根据CSLC文理兼容、偏向哲学社会科学的研究特色,选译著作应有很强的思想性;(2)这些著作的思想观念不求经典,但却是开拓新研究方向,融合新研究方法的创始之作。此动议萌生之时,CSLC就开始着手选题和组织翻译,历时两年余,“语言与认知译丛”首批作品开始陆续奉献于读者面前。译事辛苦,尽管各书译者都勤勤恳恳,几易其稿,但不足乃至错讹之处可能仍难避免,诚恳期望学界同仁和广大读者朋友批评指正。在此成书之际,CSLC尤其感谢浙江大学出版社的真情投入和热情支持。

CSLC“语言与认知译丛”主编

黄华新 盛晓明

2008年9月

中文版序:认知科学与心智计算理论概述

在《心智、大脑与计算机》中文版出版之际,我愿意对书中详细探讨的有关心智计算理论的观点进行概述。

“心智计算理论”(computational theory of mind, CTM)在帕斯卡(Blaise Pascal, 1623—1662)、霍布斯(Thomas Hobbes, 1588—1679)以及莱布尼茨(Gottfried Wilhelm von Leibniz, 1646—1716)等人的著作中就有所论及,但它的真正繁荣是在巴贝奇(Charles Babbage, 1791—1871)十进制机器、楚泽(Konrad Zuse, 1910—1995)二进制机器(约 1940 年)以及图灵(Alan Mathison Turing, 1912—1954)抽象计算机(约 1936/1937 年)等开创性工作基础上而出现于 20 世纪后半叶。事实上,CTM 的繁荣也伴随着认知科学在 20 世纪六七十年代的诞生和发展。

学界通常对认知科学存在两种理解,而这两种理解都与 CTM 密切相关。将认知科学看作是研究认知的交叉学科,我称之为广义认知科学——正如这个名称所表达的——本书图 I.1 即为它的图解。但是很多学者对这种理解并不满意,认为认知科学并非仅是这些有关认知研究的学科总称——他们期望一种更为具体的能够涵括这种研究取向的定义。因此也就有了狭义认知科学,将认知看作是一种系列计算,也就是 CTM。

但倘若认知最终被证明并非是计算的,那么对认知科学的这种理解恐怕会不幸地失去其研究主题,而这似乎是难以接受的。最好的办法是允许并且接受认知科学是“研究认知的交叉学科”这样的广义理解,同时也接受对认知科学的狭义理解,即目前最好的认知理论也就是将认知看作是一种系列计算。那么就拥有了如下三个问题:

Q1: 认知是计算,这种观点如何产生?

Q2: 这种观点究竟确切地说明了什么?

Q3: 这种观点存在什么难题,它具有怎样的前景?

I. CTM 的历史:从亚里士多德到图灵

CTM 的出现受到很多智识传统和学科的重要影响。这里,我将简要回顾从亚里士多德到图灵对于 CTM 产生所作的一些贡献。这部分内容主要有两个主题:一个与可内省的常识心理学相关;另一个与非可内省的非常识生物学相关。我将前者称为“软件”主题,而将后者称为“硬件”主题。

亚里士多德 在我看来,这两个主题均来自亚里士多德(Aristotle,公元前 384—前 322):

软件主题:亚里士多德对记忆和回忆的讨论,都是基于记忆内容的联想链接;

硬件主题:亚里士多德将脑看作是血液的散热器——思维由心脏产生(因此人们说“用心学习”)。

A. 截止到图灵:软件主题

(i)联想主义 17—19 世纪,英国经验主义者霍布斯、洛克(John Locke, 1632—1704)、休谟(David Hume, 1711—1776)、密尔(James Mill, 1773—1836)等对亚里士多德有关记忆联想结构的讨论,作了最为突出的探讨。本书图 1.1 标识了经验联想主义关于心灵观念的一般观点。联想主义主要持有如下假设:

- (1)心灵通过联想原则由简单观念构成复杂观念(同样参考图 1.1);
- (2)观念基于联想原则随时间相互联结(见图 1.2)。

联想主义一直是心理学的主导理论,直到 20 世纪由于下列因素的出现导致了联想主义走向衰落:

- (1)人们对内省主义方法的强烈不满;
- (2)脑神经科学研究的快速发展(见下文);
- (3)“科学”心理学观念的出现(见下文)。

(ii)行为主义 在当时人们看来,更为“科学”的心理学观念应该是行为主义。纽约《时代周刊》(1942)曾宣称行为主义是“探索人类心智的新时代”。它起始于华生(John Broadus Watson, 1878—1958)《行为主义者心目中的心理学》(1913),该文被评为《心理学评论》出版史上最为重要的一篇文章。华生在文中所讨论的一些主要问题成为后来行为主义的核心观点,直到 1960 年以后行为主义的衰落。他的主要观点如下:

- (1)心理学是一门客观的自然科学;

- (2)心理学的任务是预测和控制行为；
- (3)人类与动物心理是连续的,并无本质不同；
- (4)内省和意识不属于心理学的研究内容。

华生的观点在后来 45 年里得到著名心理学家赫尔(Clark Leonard Hull,1884—1952)、奥斯古德(Charles Egerton Osgood,1916—1991)、托尔曼(Edward Chase Tolman,1886—1959)等人的深入研究。最为极端的是斯金纳(Burrhus Frederic Skinner,1904—1990),他在《语言行为》(1957)中尝试将他早年动物操作条件(“斯金纳箱”)的工作扩展到人类语言领域。

(iii)认知主义 1959年,年轻的语言学家乔姆斯基(Chomsky,1928—)出版了一部与斯金纳观点针锋相对的著作,被历史学家称之为“华生行为主义宣言之后最有影响力的心理学著作”。最重要的是,乔姆斯基进一步提出在解释行为时需要考虑心灵具有其本身的主动操作原则的观点。

在米勒(George Armitage Miller,1920—)、加兰特尔(Eugene Galanter)与普利布拉姆(Karl H. Pribram)的《行为的规划和结构》(1960)中贯彻了这种观点,被普遍认为开启了认知主义对行为主义的回应。该著作提出了一种 TOTE(Test-Operation-Test-Exit,测试—操作—测试—输出)单元,用以替代行为主义的“反射弧”,如本书图 2.2 所示。

这些单元具有如下主要特征:

- (1)单元包含着有机体关于世界的信息表征；
- (2)涉及一种控制单元序列操作的系统；
- (3)它们能够以无限的方式彼此嵌套。

下面是分别与计算机关键概念的心理学的类比:

- (1)表征知识的数据结构；
- (2)转换控制程序；
- (3)有助于提高编程效率及灵活性的子程序。

B. 截止到图灵:硬件主题

亚里士多德的“硬件”猜想在历史上并没有获得支持,但有关“硬件”的理解在亚里士多德之后并没有立即走向正确,而是先变得更加离奇。

预览 神经科学的历史,在很大程度上可以看作是尝试寻找脑的心理功能与其结构间映射的历史,因此长期伴随着定位论和整体论两种取向之间的争论。最初,人们确立了脑在思维(广义含义)和行为中的作用。随后确定的问题是,思维产生于脑质还是脑质中的脑孔(脑室)。接着,从脑的功能和解剖特征上对脑进行了粗略划分,如划分了小脑和脑皮质层。进而又

发现了脑皮质层自身各个区域的不同子功能,以及神经系统的基本细胞要素。最后确立的是这些细胞要素的结构和功能。我们就以这样的顺序分三步介绍:

(i) **脑质与脑孔(脑室)** 在第一个千年早期(公元 100—400 年左右),理论家已经认识到脑是思维的器官这一重要原理,但却认为发生在脑内的那些孔道(脑室)之中。基于这种观点,“生命精气”储存在脑室之中,当需要时则通过中空的神经作用肌肉执行,以及获得感知(他们并没有说明这个过程如何进行)。这种观点以这种或那种形式持续了大约 1500 年(正确理论的出现着实不易!)。甚至笛卡尔也持有一种变更的脑室理论——松果体悬挂在中脑室,同时接受心灵和身体的作用。到了 18 世纪,思维才被完全确立为由脑质发生,但发生在脑的什么部位呢? 又是如何发生的呢?

(ii) **在什么部位发生?** 对于这一问题有两种回答,直到今天它们各自依然产生着影响。定位论起始于高尔(Franz Joseph Gall, 1758—1828),他提出这种观点的原因部分地出于其杰出的神经解剖学工作,部分地出于其哄骗性的骨相学说。高尔是最先用解剖学案例证明不同的认知功能大体上与脑皮层粗略的不同解剖部位相关,但是他认为可以通过这些不同部位的特征如测量脑颅突起部位的大小,推断个体具体的人格特征,如本书图 3.5 所示,这就走得太远了。

布洛卡(Paul Broca, 1842—1880)和维尔尼克(Carl Wernicke, 1848—1905)对于失语症(患者因脑损伤而丧失语言能力)历史性的研究,为脑功能定位论作出了重要贡献。随后,神经解剖学家如布罗德曼(Korminian Brodmann, 1868—1918)等提出了更为详细的局部脑皮层与其功能的对应关系,如本书图 3.5 所示。

有很多理由认为,当代定位论者继承的这一传统可以追溯到两千年前的脑室定位论。“整体论”是与定位论相对立的另一种观点,代表性的人物有与布洛卡同时期的弗楼伦(Marie-Jean-Pierre Flourens, 1794—1867)以及较近的拉什利(Carl Spencer Lashley, 1890—1958),但在生物学理论中从未占有主导地位。

(iii) **如何发生?** 关于心灵机制的最大争论,主要体现在以发明水银染色技术而闻名的高尔基(Camillio Golgi, 1843—1926)的“神经网络理论”与卡哈尔(Santiago Ramony Cajal, 1852—1934)的“神经元理论”之间的争论。

争论的焦点是:神经系统与血液和氧化的循环系统一样,是一种完全连续网络(高尔基:神经网络),还是类似于人的骨骼带有间隙和连接点(卡哈尔:神经元)? 显然,当代人们的使用习惯(神经元)已充分说明神经元理

论最终是正确的,尽管直到临近 20 世纪时因谢灵顿(Charles Sherrington, 1857—1952)对于突触(卡哈尔的“节点”)的研究才取得决定性的证据。到了 1950 年前后,有关神经元的总体特征大体上都已经确定。

II. CTM 的现状:从图灵到 2000 年

图灵机 在 1936/7 年,图灵发表了《论可计算数及其在判定难题中的应用》一文,尽管标题有些模糊,但这也可能是计算机科学中最为重要的一篇文献。在文中,图灵精确地论证了“(自动)计算机器”(为了纪念图灵,我们现在称之为“图灵机”)原理,并且进一步揭示了“通用图灵机”的概念——它可以做任何(其他)图灵机所能做的计算。随后,众多数学家和计算机科学家设计了在现实中能够执行更多任务的形式机器,但都被证明与图灵机是等价的。如果问(通用)图灵机究竟具有怎样的计算能力,可以粗略地认为:它能够计算任何能够计算的所有计算。

图灵测试 1950 年,图灵发表了另一篇论文,题为《计算机器与智能》,对认知科学和 CTM 都产生了重要影响。在该文中,图灵探讨了机器具有智能的可能性,并提出了著名的“图灵测试”(他称之为“模拟游戏”)——粗略地讲,就是当人与具有某种智能层次的机器,如计算机,经过一段时间交谈后,将不能区分是与机器,还是与另外一个人在交谈。

A. 两个主题的融合(I):一篇作为转折点的文献

1943 年,麦卡洛克(Warren McCulloch, 1898—1969)和皮茨(Walter Pitts, 1923—1969)在 MIT 发表了著名的但有些艰涩(甚至以专业标准看也是如此)的题为《神经活动内在概念的逻辑演算》一文。该文的重要性体现在两个方面:一是使多个学科产生了相互联系;二是促进了随后 CTM 的发展。

(i)使哪些学科联系在了一起? 这些学科主要包括神经科学、计算机科学和心理学(幸运地还包含些许哲学)。

神经科学 他们首先总结了当时神经科学所取得的一些主要成果,并“理想化”了神经元的一些重要特征(如本书图 4.1 所示):

- (1)神经元激活是一种“全一或一无”的过程;
- (2)在潜伏附加期内,有固定数目的突触被激活,以便能够在任何时刻对神经元予以激活;
- (3)在神经系统内,最重要的延迟是突触延迟;

(4)在任何抑制性突触活动时刻,神经元绝不会产生兴奋;

(5)神经网络的结构不随时间发生变化。

这种“理想化”的神经元模型对应着二值(真、假)命题逻辑的逻辑环路或逻辑闸,如本书图 4.4 和图 4.5 所示。

计算机科学 随后证明,当这种环路系统辅以可无限扩展的存储机制,那么系统就可以具有(通用)图灵机的计算能力。

心理学 提出所有的心理学问题归根结底都可以还原于神经系统的二值、开-关逻辑。

描述层次 该文还尝试对“信息加工系统”用三个不同描述层次对信息进行了统一和系统化。

(ii)产生的影响 麦卡洛克和皮茨的这篇文献,最终对认知科学的诞生以及进一步提出 CTM 产生了重要影响,是在冯·诺依曼的 EDVAC 报告中唯一引用的文献——EDVAC 报告实质上是所有随后数字计算机的蓝图,也是罗森布拉特(Frank Rosenblatt)感知器研究的基础——最终发展为联结主义研究,稍后我们再回到这个问题。

B. 两个主题的融合(Ⅱ):神经-逻辑

“神经-逻辑”主要研究类神经网络的计算特征。

感知器及其缺陷 1958 年,罗伯森拉特发表论文讨论感知器问题,提出感知器是这样一种装置,“一旦被描述就产生感知”。感知器基本上是由麦卡洛克和皮茨神经元构成的一种网络,但除了具有神经元的上述五个特征外,网络还能够进行学习。一个简单的感知器包含三个部分:单元感觉输入层、单元联结层和单元反应输出层,如本书图 4.6 所示。

经过学习,感知器大致上能够与人类类似地识别出不同的刺激群,如男性面孔与女性面孔。但是在 1969 年,明斯基(Marvin Minsky)和帕佩尔特(Seymour Papert)证明,这种感知器在原则上被训练学习某些基本的判别(例如互斥问题,或者说要么 P,要么 Q,但不可两者同时),从而导致了感知器研究的冷落,研究者们的兴趣点开始转向由明斯基和帕佩尔特提出的更具普遍意义的数字范式。

C. 两个主题的融合(Ⅲ):数字

图灵机 上文已有所提及,图灵为证明可计算和非可计算论题,在 1936/7 年提出了图灵机模型,是一种(相对)简单的抽象形式“机器”。可看作主要由两个部分构成:带有程序的磁头和一条无限长的磁带,磁带是机器

的“记忆”，能够依据程序指令左右移动，读写及删除字符，如本书图 6.1 和图 6.2 所示。

图灵机具有一种特殊的操作循环：(从磁带中)读取、(在磁带上)印写、移动磁头(每次一个方格)、(按照磁头中的程序)进入下一种状态。正是基于这样一种机器，图灵论证了他的著名论题。

冯·诺依曼机 1945年6月，约翰·冯·诺依曼(John von Neumann, 1903—1957, 约翰并不是他最初的名字)完成了最初 EDVAC 设计报告草稿(这份报告草稿提交时尚不完整,有很多标识参考文献的空格,冯·诺依曼期望之后补充)。随后,EDVAC 计算机在宾夕法尼亚大学摩尔电气工程学院迅速流传开来,事实上成为以后所有计算机生产线的蓝图。如他本人所描述的,冯·诺依曼机具有如下特征:

- (1)它是全自动的;
- (2)具有存储数据和指令的记忆功能;
- (3)能够储存程序;
- (4)具有执行指令的控制机制;
- (5)具有逻辑和算法的“器官”;
- (6)具有输入和输出设置。

描述层次 值得注意的是,图灵机和冯·诺依曼机都可以看作是一种实际的“信息加工系统”,能够在两个层次上进行描述:物理环路(硬件)和运行的程序(软件)。

D. 两个主题的融合(IV):联结主义

只有一层可训练的麦卡洛克和皮茨“神经元”的简单感知器,被证明不能学习识别某些对象类别,因而人们停止了对感知器的研究。包含多层“神经元”的感知器或许能够学习识别那些类别,但在一段时期内没有人知道如何训练它们。大约在 1985 年,众多研究者开始关注所谓的“联结主义”或者“并行分布式加工”的研究,此时业已证明能够训练多层神经网络识别那些疑难的类别。

三层前馈网络 随后,联结主义者提出了很多不同的网络组织或者“结构”,但我们注意到有一种是应用最广的,那就是具有三个联结层次的前馈网络,它应用于著名的 NETtalk 演示项目中的示例,如本书图 10.2 所示。

在 NETtalk 机低端的“输入层”给机器呈现书面文本,接着这些刺激产生的激活便传递到中间层或“隐单元层”,最后在机器顶端的“输出层”开始“朗读”这个文本。也就是通过调节连接权值从而使文本层与发音层相匹

配,训练 NETtalk 学习基本英文的发音。如果这一过程是靠手动完成的,那么我们自然地会称之为“程序运行”的步骤。三层前馈机器的这些特征表明,它们同样具有图灵机的计算能力。

描述层次 我们注意到,联结网络与感知器一样,具有类似的生物硬件,当给予与人类相似的心理任务时,它们通过调整各层之间的联结权值而能够很好地学习完成任务。

E. 心智计算理论(结果)

上述这些内容反映了什么“心智理论”呢?对这个问题的回答有助于进行历史的回顾,并了解我们现在所处的历史位置。

(i) 心智表征理论(RTM) 至少从英国经验主义那里就具有一种“外在(out there)”心灵的概念,RTM(representational theory of mind)主要包含两个中心假设:

(RTM)

- (1) 认知的心理状态是表征的——我们通过表征世界而思维世界;
- (2) 思维过程最终都是这些心理表征的转换或“操作”。

对于经验主义而言,这些表征通常称之为“观念”,主要被认为是与它们相似的关于某事物的图像。我们在前面了解到,联想主义者认为思维主要是通过联想原则控制的“观念”连续过程。后来人们认为,表征的“相似性”理论与单纯的思维“联想”理论对于人类所有认知和行为的解释是不充分的。例如,我们已经了解到米勒、加兰特尔与普利布拉姆如何利用复杂的 TOTE 构建他们的心智理论。

(ii) 心智计算理论(CTM) 如果 TOTE 单元可以被看作是具体的程序示例,那么像脑一样操作 TOTE 单元的硬件便被看作是运行程序的硬件。因此,出现了著名的类比:心灵之于脑正如软件之于硬件。也有人不仅仅把这看作是一种类比,还把它当作一种科学宣言:心灵就是脑中运行的程序,并有这样的推论:想要得到人工心灵,唯一需要做的事就是找到合适的运行程序的硬件。既然现代数字计算机等价于通用图灵机,在原则上可运行执行任何计算的程序,这就意味着如果我们获得正确的心灵算法,那么就可以在数字计算机上编程使之具有心灵。我们如何知道它具有心灵呢?回答是如果能够通过适当的图灵测试则就证明计算机具有智能。这样我们就有了如下几种观念,塞尔(John R. Searle)曾称之为“强人工智能”:

- (1) 编程了的数字计算机,如果能够通过图灵测试,那么机器就具

有了智能、认知、心灵等；

(2)人脑基本上就是麦卡洛克—皮茨逻辑环路网络(十无限容量的存储=通用图灵机)；

(3)因此,可以通过找到我们脑内运行的程序,探究人类的智能、认知、心灵等。

请注意,这组命题并不是连贯的,前两个也许是正确的,但第三个是错的。世界上,智能(认知、心灵等)的来源可能有两种,一种出自程序,另一种不是,而人脑恰恰可能属于非程序的那种智能。因此,有人说图灵是CTM之父,这是错误的。图灵想要得到的是机器的智能,而CTM追求的是智能的机器。但CTM吸引人的地方是,很难想象还有另一种智能、认知、心灵等的来源。其他的任何一种理论看起来都具有神秘色彩,如福多(Jerry Fordor)所言:“这是城镇里的唯一游戏。”

那么,用计算进行重新装扮RTM,就得到了CTM:

(CTM)

(1)认知状态是具有内容的计算心理表征的计算关系。

(2)认知过程(认知状态的改变)是具有内容的计算心理表征的计算操作。

我们由此就获得了CTM的工作状态。

(iii)数字CTM与联结CTM 注意到,在介绍CTM时我们只提到了“计算”状态、表征、程序等。尽管CTM从经验上产生自传统数字计算机,但是并不保证一定要求如此——联结主义机器同样与比如冯·诺依曼机一样,是一种计算机。一种理论,如果要求计算结构和表征都是数字的,是一种有关数字的CTM理论,称之为心智数字计算理论(DCTM);同样的,如果要求计算结构和表征都是联结主义的,是一种有关联结主义的CTM理论,则称之为心智联结计算理论(CCTM)。尽管当下人们对哪一种理论是正确的或者是更正确的存在很大争论,但它们都属于CTM的一种类别,而CTM又属于RTM的一种类别。

III. CTM 的未来

共性难题

两种CTM的理论似乎都面对一些严重困难——有一些可以说是所有

心智理论都共有的难题。

首先是意识难题。如何能够使用计算的术语解释我们全部的意识经验——无论是哪一种机器，我们如何对其“编程”？

其次是表征或内容难题。如何用计算的术语解释某一符号关涉其事实所关涉的内容？典型的情形是，程序员们说如果 CTM 是正确的就行得通，但并不是所有人都对此满意。

数字 CTM 难题

软件难题 如何解释数字模型在众多层面上与人类心理相似性上的失败？我们能够列出很长的清单，表明普通人的心理功能对于数字计算机来说是完全陌生的。例如，(i)当有人受到微小损伤时，并不会“崩溃”，他会说“只是稍微有些不便”，以及(ii)我们似乎并不是通过“地址”检索记忆中的信息，而是通过内容进行检索——是关于什么的信息。

硬件难题 如何解释脑与计算机硬件(如芯片)在结构和功能上存在的巨大差异？如何想象一些标准程序以每秒百万次的速度在脑内运行？这只比迄今速度最快的数字计算机慢一点点。

联结主义 CTM 难题

软件难题 如何解释联结主义模型在众多层面上与人类心理相似性上的失败？例如，我们能够处理那些复杂的并且即时的加工任务，如语言加工、计划和演绎推理。

硬件难题 尽管联结主义模型是受“生物学”的启发而提出的，但它们仅仅粗略地模拟神经系统的结构和功能——如何使之更加精确呢？例如，训练三层网络使用的路径，并找不到与之对应的生物学知识。

展 望

数字和联结主义 CTM 的优缺点现在都已经显现得很清楚了。如果能够找到一种方式，将两者作为整体充分地结合在一起，那么 CTM 的前景无疑是非常乐观的：CCTM 应用于感知、记忆以及其他“直觉性”任务；DCTM 应用于推理、计划和其他“高级心理”任务。不过，目前还没有人知道究竟如何做才能实现。

R. M. 哈尼什

2010 年 4 月于美国亚利桑那大学

前 言

认知科学是研究认知的科学,由心理学、计算机科学、神经科学、语言学 xv
以及哲学等不同领域的学科所构成。促使这些学科结合在一起的是这样一种理念,即脑(神经科学)是一种计算装置(计算机科学),而认知(心理学)研究其软件——运行在脑中的程序。本书是关于认知科学基础的历史导论。这里的每个概念都很重要。说它是历史的,是因为我们将追溯一些关键概念的历史——基本上从 19 世纪(偶尔会回到亚里士多德)到现在。说它是基础的,是因为本书主要关注当代认知科学中研究心—身关系的一般(计算的)框架,并未涉及一些特殊的竞争理论及其包含的实验结果细节。同时,由于认知科学是一门交叉学科,我们将较为详细地介绍哲学、心理学、神经科学和计算机科学对这门学科的一些贡献。

导论:什么是认知科学,描述了关于这门学科的广义和狭义两种理解,并提供一种整合的观点。

第一部分:历史背景,追溯哲学、心理学以及神经科学对认知科学的一些贡献,直到大约 1950 年数字计算机的出现。

第二部分:心智的数字计算理论,主要关注一种特殊的人工智能程序——SHRDLU,并讨论一些数字结构和标准的知识表征方式。根据这些计算知识考察它们所蕴含的心智理论,以及关于 SHRDLU 的一些最重要的问题。

第三部分:心智的联结计算理论,进一步关注联结主义机器的心智模型。我们将考察两种联结主义计算解释程序——Jets & Sharks 网络模型以及 NETtalk 模型。我们还论及联结主义结构、学习程序和表征程式等相关内容。在讨论这些问题之后,将转向探讨它们可能蕴含的心智理论以及 xvi
一些重要的相关问题。

结语: 认知科学的计算,或者究竟什么是计算机? 在分析计算机和计算两个重要概念后,我们试图对它们进行整合。结果表明,整合后的特征同样适用于数字模型和联结主义模型。