



中国计算机学会学术著作丛书

心智计算

Mind Computation

史忠植 著

清华大学出版社





中国计算机学会学术著作丛书

心智计算

Mind Computation

史忠植 著



清华大学出版社
北京

内 容 简 介

本书旨在创建类脑计算的基本理论和构架。全书围绕心智模型 CAM,系统地论述心智计算的理论基础。全书内容由九章构成。第1章绪论,概要介绍与心智有关的哲学问题、生物基础、计算表征等问题;第2章讨论心智模型 CAM 的系统结构;第3章阐述记忆机理;第4章探讨意识的机理和功能;第5章讨论视觉感知;第6章阐述运动控制;第7章讨论心理语言和自然语言的处理;第8章探讨学习问题;第9章提出智能科学发展的路线图。

本书适合智能科学、计算机、自动化领域的研究人员和高校师生参考阅读。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

心智计算/史忠植著.--北京:清华大学出版社,2015

(中国计算机学会学术著作丛书)

ISBN 978-7-302-40733-1

I. ①心… II. ①史… III. ①人工智能—计算—研究 IV. ①TP183

中国版本图书馆 CIP 数据核字(2015)第 153006 号

责任编辑:薛 慧

封面设计:傅瑞学

责任校对:赵丽敏

责任印制:沈 露

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社总机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:清华大学印刷厂

经 销:全国新华书店

开 本:175mm×245mm 印 张:23.75 字 数:462 千字

版 次:2015 年 8 月第 1 版 印 次:2015 年 8 月第 1 次印刷

印 数:1~2000

定 价:69.00 元

产品编号:064121-01

评审委员会

中国计算机学会学术著作丛书

■ 名誉主任委员：张效祥

■ 主任委员：唐泽圣

■ 副主任委员：陆汝钤

■ 委员：（按姓氏笔画为序）

王 珊 吕 建 李晓明

林惠民 罗军舟 郑纬民

施伯乐 焦金生 谭铁牛

丛书序

第

一台电子计算机诞生于 20 世纪 40 年代。到目前为止,计算机的发展已远远超出了其创始者的想象。计算机的处理能力越来越强,应用面越来越广,应用领域也从单纯的科学计算渗透到社会生活的方方面面:从工业、国防、医疗、教育、娱乐直至人们的日常生活,计算机的影响可谓无处不在。

计算机之所以能取得上述地位并成为全球最具活力的产业,原因在于其高速的计算能力、庞大的存储能力以及友好、灵活的用户界面。而这些新技术及其应用有赖于研究人员多年不懈的努力。学术研究是应用研究的基础,也是技术发展的动力。

自 1992 年起,清华大学出版社与广西科学技术出版社为促进我国计算机科学技术与产业的发展,推动计算机科技著作的出版,设立了“计算机学术著作出版基金”,并将资助出版的著作列为中国计算机学会的学术著作丛书。时至今日,本套丛书已出版学术专著近 50 种,产生了很好的社会影响,有的专著具有很高的学术水平,有的则奠定了一类学术研究的基础。中国计算机学会一直将学术著作的出版作为学会的一项主要工作。本届理事会将秉承这一传统,继续大力支持本套丛书的出版,鼓励科技工作者写出更多的优秀学术著作,多出好书,多出精品,为提高我国的知识创新和技术创新能力,促进计算机科学技术的发展和进步作出更大的贡献。

中国计算机学会

2002 年 6 月 14 日

前言

脑

科学、认知计算是当前极为活跃的研究领域,欧盟、美国等都设立重大研究计划,投入巨资开展研究。作者倡导的智能科学,希望通过脑科学、认知科学、人工智能等学科交叉研究,提出类脑计算的途径,达到人类水平的人工智能目标。

心智计算(mind computation)是当前智能科学研究的热点。心智是指人的全部精神活动,包括思维、推理、记忆、学习、情感、决策、意志、意识等。心智计算是心理符号的计算,是大脑的信息加工过程。心智的可计算性是联结人类的精神活动与人工系统仿真的桥梁和纽带。心智的可计算性提供了用计算过程来解释人类行为的理论基础。通过长期研究,作者提出了心智模型 CAM(consciousness and memory),为类脑计算智能系统提供通用构架。

在人脑的智能活动中,意识和记忆发挥关键作用。在庆祝Science 创刊 125 周年之际,该刊公布了 125 个最具挑战性的科学问题,发表在 2005 年 7 月 1 日出版的专辑上。在今后 1/4 个世纪的时间里,人们将致力于研究解决这些问题。其中问题 2 是“意识的生物学基础是什么?”问题 15 是“记忆如何存储和提取?”因此,心智模型 CAM 将会随着意识、记忆科学问题研究的不断深入,与时俱进,不断发展和完善。

全书围绕心智模型 CAM,系统地论述心智计算的理论基础。全书内容由九章构成。第 1 章绪论,概要介绍与心智有关的哲学问题、生物基础、计算表征等问题;第 2 章讨论心智模型 CAM 的系统结构;第 3 章阐述记忆,包括长时记忆、短时记忆、工作记忆的工作机理和表示方式;第 4 章

探讨意识的机理和功能,从工程的角度展示意识实现的可能方案;第5章讨论视觉感知,侧重讨论物体知觉和空间知觉;第6章阐述运动控制的神经结构、脑电信号分析和运动的神经编码,以及运动的决策模式;第7章讨论心理语言和自然语言的处理,语言认知将把心理语言模型和神经科学结合起来;第8章探讨学习问题,重点论述强化学习、深度学习、内省学习以及脑认知数据分析;第9章概述类脑计算的最新进展,提出智能科学发展的路线图。

本书总结了作者和中国科学院计算技术研究所智能科学实验室多年的科研成果,也吸取了国内外同行的研究成果和有关文献的精华,在此谨向这些成果和文献的作者表示感谢,他们的丰硕成果和贡献是本书学术思想的重要源泉。在本书撰写过程中,作者与斯坦福大学心智与脑计算中心麦克伦特(McClelland J L)教授、圣路易斯华盛顿大学范埃森(David van Essen)教授、德国海德堡大学迈耶(Meier K)教授、南加州大学罗森勃卢姆(Rosenbloom P S)教授、美国密歇根大学莱尔德(Laird J E)教授、加拿大滑铁卢大学伊莱亚史密斯(Eliasmith C)教授、德国德累斯顿工业大学巴德尔(Baader F)教授、美国卡内基梅隆大学米切尔(Mitchell T M)教授、美国西北大学福布斯(Forbus K D)教授、美国密歇根州立大学翁巨扬(Weng J)教授等进行讨论和交流,这对本书学术思想的确立和发展,发挥了重要作用,在此谨向上述学者表示衷心的感谢。

本书所述的研究工作得到国家重点基础研究发展计划“973”项目“脑机协同的认知计算模型”(项目编号:2013CB329502)、国家自然科学基金重点项目“基于云计算的海量数据挖掘”(批准号:61035003)、“基于感知学习和语言认知的智能计算模型研究”(批准号:60435010)等的支持。

由于作者水平有限,加上心智计算还处于初创阶段,一些概念和学术思想仍有待深入研究,书中不妥和错误之处在所难免,恳请各位专家和广大读者不吝指正。

作 者

2015年3月

目录

第 1 章 绪论	1
1.1 心智	1
1.2 心智的哲学问题	2
1.3 心智的生物学基础	5
1.4 心智的智能科学问题	7
1.5 心智的结构	15
1.6 心智的模块性	16
1.7 心智的社会	19
1.8 自动机理论	20
1.8.1 概述	20
1.8.2 有限状态自动机	23
1.8.3 概率自动机	24
1.8.4 细胞自动机	25
1.9 图灵机	26
1.10 心智的计算理论	27
第 2 章 心智模型 CAM	32
2.1 概述	32
2.2 心智建模标准	35
2.3 认知心智建模	39
2.3.1 物理符号系统	39
2.3.2 ACT-R	39
2.3.3 Soar	41
2.4 联结心智建模	43
2.4.1 联结机制	43
2.4.2 自适应谐振理论	45

2.5 智能体心智建模	48
2.6 CAM 系统结构	49
2.7 CAM 认知周期	53
第3章 记忆	55
3.1 概述	55
3.2 动态描述逻辑基础	56
3.2.1 基本概念	57
3.2.2 动态描述逻辑语义	59
3.2.3 动态描述逻辑推理	59
3.3 长时记忆	61
3.3.1 语义记忆	63
3.3.2 情景记忆	69
3.3.3 程序性记忆	72
3.4 短时记忆	74
3.4.1 短时记忆编码	74
3.4.2 信息提取	76
3.4.3 CAM 的短时记忆	78
3.5 工作记忆	80
3.5.1 工作记忆模型	81
3.5.2 工作记忆和推理	82
3.5.3 工作记忆的神经机制	83
3.6 遗忘理论	84
3.7 记忆的生理机制	88
3.8 记忆-预测理论	91
3.8.1 恒定表征	91
3.8.2 大脑皮质区的结构	91
3.8.3 大脑皮质区如何工作	92
第4章 意识	94
4.1 概述	94
4.1.1 意识的基本要素	97
4.1.2 意识的属性	100
4.2 意识理论	101
4.2.1 意识的剧场模型	101

4.2.2	意识的还原论理论	104
4.2.3	神经元群组选择理论	107
4.2.4	意识的量子理论	109
4.2.5	意识的方块模型	111
4.2.6	综合信息理论	112
4.3	注意	114
4.3.1	注意的功能	114
4.3.2	选择性注意	116
4.3.3	注意分配	121
4.3.4	注意系统	122
4.4	元认知	122
4.4.1	元认知知识	123
4.4.2	元认知体验	123
4.4.3	元认知监控	123
4.4.4	元认知训练	124
4.5	动机	124
4.5.1	概述	124
4.5.2	动机理论	125
4.6	CAM 的意识子系统	130
4.6.1	觉知模块	131
4.6.2	注意模块	132
4.6.3	全局工作空间模块	133
4.6.4	动机模块	134
4.6.5	元认知模块	136
4.6.6	内省学习模块	137
第 5 章	视觉感知	138
5.1	皮质视觉区	138
5.2	视觉计算理论	140
5.2.1	马尔的视觉计算理论	140
5.2.2	格式塔视觉理论	143
5.2.3	双视觉通路	144
5.2.4	拓扑视觉理论	144
5.3	特征捆绑	145
5.3.1	时间同步理论	146

5.3.2	特征捆绑的形式模型	147
5.3.3	特征整合理论	147
5.3.4	神经网络模型	148
5.4	物体识别	149
5.4.1	视觉表象	149
5.4.2	物体底层特征提取	149
5.4.3	关系编码	151
5.4.4	学习识别网络	152
5.4.5	连接搜索	153
5.5	视觉空间认知	154
5.6	视觉有效编码	159
第6章	运动控制	164
6.1	运动控制的神经结构	164
6.2	大脑皮质运动区	166
6.3	基底神经节	167
6.4	运动控制通路	172
6.5	脑电信号分析	175
6.5.1	脑电信号分类	175
6.5.2	脑电信号分析方法	177
6.6	运动的神经编码	178
6.6.1	概述	178
6.6.2	熵编码理论	180
6.6.3	贝叶斯集群编码	184
6.6.4	贝叶斯集群解码	185
6.7	脑机接口	186
6.7.1	概述	186
6.7.2	脑机接口技术	188
6.7.3	P300 脑机接口系统	191
6.8	脑机融合	194
第7章	语言认知	197
7.1	心理词典	197
7.2	语言输入的知觉分析	199
7.2.1	口语输入	199

7.2.2	语音编码.....	200
7.2.3	韵律认知.....	201
7.2.4	书面输入.....	208
7.2.5	单词识别.....	210
7.2.6	言语产生.....	211
7.3	乔姆斯基的形式文法	213
7.3.1	短语结构文法.....	213
7.3.2	上下文有关文法.....	214
7.3.3	上下文无关文法.....	215
7.3.4	正则文法.....	216
7.4	扩充转移网络	218
7.5	概念依赖理论	221
7.6	语言理解	223
7.6.1	概述.....	223
7.6.2	发展阶段.....	225
7.6.3	基于规则的分析方法.....	228
7.6.4	基于语料的统计模型.....	232
7.6.5	机器学习方法.....	234
7.7	脑语言功能区	237
7.7.1	经典语言功能区.....	237
7.7.2	语义相关功能区.....	239
7.7.3	音韵相关功能区.....	239
7.7.4	拼字相关功能区.....	239
7.7.5	双语者脑语言功能区.....	240
7.8	语言理解的神经模型	240
7.8.1	失语症.....	240
7.8.2	经典定位主义模型.....	242
7.8.3	记忆-整合-控制模型	244

第8章 学习..... 246

8.1	概述	246
8.2	强化学习	248
8.2.1	强化学习模型.....	248
8.2.2	Q 学习	251
8.2.3	部分感知强化学习.....	253

8.2.4	基于动机的强化学习	254
8.2.5	Soar 系统的强化学习	261
8.3	深度学习	264
8.3.1	概述	264
8.3.2	人脑视觉机理	265
8.3.3	自编码器	265
8.3.4	受限玻耳兹曼机	267
8.3.5	深度信念网络	269
8.3.6	卷积神经网络	271
8.4	内省学习	276
8.4.1	概述	276
8.4.2	内省学习一般模型	278
8.4.3	内省学习的元推理	279
8.4.4	失败分类	280
8.4.5	内省过程中的基于案例推理	280
8.5	脑认知数据分析	281
8.5.1	脑功能成像	281
8.5.2	脑神经语义	282
8.5.3	脑功能连接性分析	283
第 9 章	类脑计算	286
9.1	概述	286
9.2	蓝脑计划	288
9.2.1	脑神经网络	288
9.2.2	脑皮质模型	289
9.2.3	超级计算模拟	292
9.3	欧盟人脑计划	295
9.3.1	概述	295
9.3.2	峰电位时序相关可塑性	298
9.3.3	统一脑模型	301
9.4	美国脑计划	303
9.4.1	人类连接组项目	303
9.4.2	MoNETA	304
9.4.3	惠普忆阻器	306
9.4.4	Neurocore 芯片	307

9.5	大脑模拟系统 Spaun	308
9.6	神经形态芯片	311
9.6.1	神经形态芯片的发展简史	311
9.6.2	IBM 的 TrueNorth 神经形态系统	313
9.6.3	英国 SpiNNaker	315
9.7	智能科学发展路线图	316
9.7.1	初级类脑计算	317
9.7.2	高级类脑计算	320
9.7.3	超脑计算	322
参考文献		326
索引		355

第1章

绪论

21

世纪的智能科学,与脑科学、认知科学一起,开始成为科学研究的主流学科。人类的科学事业所面临的挑战之一就是认识心智与大脑的关系,这也是智能科学探索的目标。心智问题是智能科学研究的最基本、最核心的问题之一^[533]。本章重点讨论心智有关的哲学问题、生物基础、计算表征等问题。

1.1 心智

英文“mind”是个多义词,相对于物质或身躯时可以指心,或者精神;在谈论理智时可以指健全的心智,正常的神志;也可以用作智力,理解力,想法,意见,意向,见解,主意,记忆,回忆等;在哲学中“mind”一般译作心灵。在智能科学中“mind”被译作心智。心智是指一系列认知能力组成的总体,这些能力可以让个体具有意识、感知外界、进行思考、做出判断以及记忆事物。心智是人类的特征,但是其他生物可能也具有心智。

心智的现象和心理方面长期以来就相互交织在一起。现象学的心智概念,是一个作为意识经验的心智概念,作为一个有意识地经验到的心理状态的概念。这是心智的最扑朔迷离的方面。另一个是心智的心理学概念,这是一个作为因果的,或者作为行为的解释基础的心智概念。在这种意义上,一种状态是心理的,如果它在行为的形成方面,扮演着恰当的因果角色,或者在行为的解释方面,至少扮演着一个恰当的角色。

按照现象学的概念,心智被它感觉的方式所刻画;而按照心理学的概念,心智被它的所作所为所刻画。这两个心智概念之间不存在相互竞争的问题。其中的任何一个都不是心智的正确分析。它们涉及不同的现象领域,二者都是相当真实的。

一特定的心理概念通常可以被分析为一个现象的概念、一个是心理的概念,或者作为二者的组合。例如,感觉,在它的核心意义上,最好被看作是一个现象的概念:有一种感觉就是有某种感觉的状态。另一方面,学习和记忆的概念最好被看作是心理的。大致地说,某些东西是学习,是因为它适当地调整其行为能力,以对某种环境刺激做出反应。一般而言,心智的现象特征,被具有那种特征的主体的样子所刻画;而一个心理的特征,则被在行为的因果关系中或者在行为的解释中所关联到的角色所刻画。

软计算创始人扎德(Zadeh L A)指出^[510]:模糊逻辑的原型本质上即为人类的心智。它可效仿人类进行推理、形成概念从而提供处理模糊信息粒以及词语计算的方法。因而允许利用不精确性、不确定性、部分真值以及近似表达的现实诸领域,其作用显得尤为重要。以致得出结论:“模糊信息粒化是人类认识的组成部分,对人工智能该结论有着意味深长的含义;若不把模糊信息粒化方法作为工具,人工智能不能达到自己的目的。”扎德强调的是人类心智的模糊信息处理能力。

心智状态是指信念、能力、意向、期望、动机、承诺等心理状态,它是决定智能社会行为与个体行为的重要因素。人类心智与观念和意识有关,它是人类意识发展到一定阶段的产物。人类学家认为,一切生物都有某种心智。心智的发展经历四个阶段:①简单反射阶段,如眼睛受到强光的刺激,瞳孔就收缩,这不受意志的支配;②条件反射阶段,巴甫洛夫的著名试验表明,信号刺激能够使狗流出唾液;③工具阶段,黑猩猩能够使用棍子打下树上的果实;④符号阶段,就是能够使用语言符号与外界进行交流,这只有人类才能做到。由此可见,动物心智处于相对低级的阶段,而人类的心智是心智发展到最高阶段的产物,人类智能的产生和发展是与人类所特有的符号语言分不开的。

1.2 心智的哲学问题

很长的时间里人们都试图从哲学、宗教、心理学和认知科学的角度来理解心智究竟是什么,并尝试探究心智的独特性质。研究心智的著名哲学家包括柏拉图(Plato)、亚里士多德(Aristotle)、笛卡儿(Descartes R)、莱布尼茨(Leibniz G W)、康德(Kant I)、海德格尔(Heidegger M)、塞尔(Searle J R)、丹尼特(Dennett D)等人。包括弗洛伊德(Freud S)和詹姆士(James W)在内的心理学家也从心理学的角度建立了有关人类心智本质的一系列有影响力的理论,以便对心智进行描述和

定义。在 20 世纪晚期和 21 世纪初期的认知科学领域内,科学家已经建立并发展了多种途径和方法来对心智和与心智有关的现象进行描述。在人工智能领域也开始尝试结合控制论和信息论来对非人类心智存在的可能性进行探索,并研究将人类精神现象在机器上实现的方法。

近年来,心智哲学(心灵哲学)迅猛发展,一跃而成为哲学领域具有基础意义和举足轻重地位的学科。如果说近代哲学向现代哲学的运动经过了所谓的“哥白尼式革命”,语言哲学由此取代认识论在近代哲学中的重心基础地位而成为现代哲学的重心和标志,那么在当今哲学的百花园中,心智哲学则成了重中之重、基础之基础。如果说本体论、认识论问题的解决离不开语言哲学的话,那么同样,语言哲学问题的解决则有赖于有关心智哲学的探讨和发展,例如,要说明语言的意义、指称、性质、特征等,当然要诉诸于心理状态的意向性,尽管不是唯一地由之所决定的。心智哲学方面的论著数量之多、问题之深广、见解之新颖别致、争论之激烈、进展之快确实令人震惊。

心智哲学研究的主要是心理现象的形式、范围、本质、特征、心与身的关系、心理内容及其根源等,并对常识心理解释模式和心理学进行哲学反思^[519]。随着认识的深入,心智哲学已改变了或正在改变传统的心理观。由于心理现象是宇宙结构图景中的重要组成部分,因此心智哲学的最新探索也触及到了一些根本的宇宙观问题,如随附性、依赖性、决定、协变、还原、规律等。有迹象表明,心智哲学向纵深的发展将是引起未来哲学变革的重要根源和动力之一。由此看来,心智哲学不是一个狭隘的心智学问,而是一个有着相对稳定的硬核、模糊的边界、开放的性格和远大的前途的广博的研究领域。当前,心智哲学焦点问题如下。

1. 心-身问题

心-身问题,即心理现象的本质、心与身的关系问题,当前的争论主要集中在还原主义、功能主义、“实现”和物理主义的二难困境等问题上。

围绕心智的本质所产生的问题主要是它与大脑和神经系统的关系,构建了心-身问题。心身二分法讨论了心智是否在某一程度上独立于人体的肉体(二元论),而肉体来自于并且可以被认作是包括神经活动在内的物理现象(物理主义),也讨论了心智是不是与我们的脑以及脑活动保持一致。另一个问题则讨论是否只有人类才拥有心智,或者是部分或全部动物以及所有生物也拥有心智,或者甚至人造机器也可能拥有心智。

无论心智与肉体的关系究竟如何,人们普遍都认为心智使个体具有主观感觉,并且对其周围环境存在意向性,可以通过一定的媒介感知并回应刺激,同时拥有意识,可以进行思考和感觉。