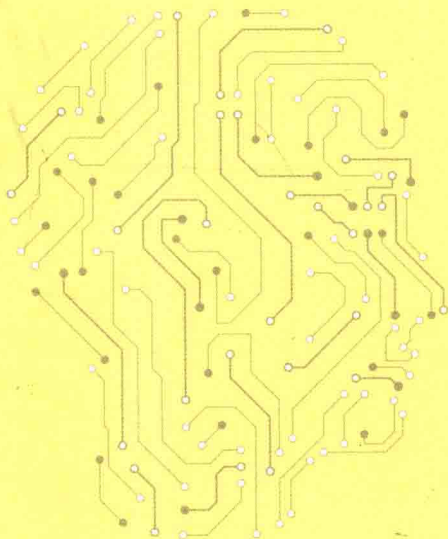


TURING 图灵原创

人工智能简史

A Brief History of Artificial Intelligence

尼克 著



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

TURING 图灵原创

人工智能简史

A Brief History of Artificial Intelligence



尼克 著

人民邮电出版社

北京

图书在版编目(CIP)数据

人工智能简史 / 尼克著. — 北京: 人民邮电出版社, 2017. 12

(图灵原创)

ISBN 978-7-115-47160-4

I. ①人… II. ①尼… III. ①人工智能—简史 IV.
①TP18

中国版本图书馆CIP数据核字(2017)第263722号

内 容 提 要

本书全面讲述人工智能的发展史, 几乎覆盖人工智能学科的所有领域, 包括人工智能的起源、自动定理证明、专家系统、神经网络、自然语言处理、遗传算法、深度学习、强化学习、超级智能、哲学问题和未来趋势等, 以宏阔的视野和生动的语言, 对人工智能进行了全面回顾和深度点评。

本书作者和书中诸多人物或为师友或相熟相知, 除了详实的考证还有有趣的轶事。本书既适合专业人士了解人工智能鲜为人知的历史, 也适合对人工智能感兴趣的大众读者作为入门的向导。

◆ 著 尼 克
责任编辑 张 霞
责任印制 彭志环

◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路11号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
大厂聚鑫印刷有限责任公司印刷

◆ 开本: 880×1230 1/32
印张: 10.125

字数: 243千字

2017年12月第1版

印数: 1-5 000册

2017年12月河北第1次印刷

定价: 49.00元

读者服务热线: (010)51095186转600 印装质量热线: (010)81055316

反盗版热线: (010)81055315

广告经营许可证: 京东工商广登字 20170147号

站在巨人的肩上
Standing on Shoulders of Giants



iTuring.cn

前 言

历史素有两种写法：以人为主和以事为主。所有的传记都是以人为主的；而各种专史，如战争史，则多以事为主。所谓历史是人民创造的还是英雄创造的，我个人的偏好还是以人为本。八卦的历史，读者自然喜欢，对作者也有好处，就像一战后英国首相劳合·乔治对他的耶路撒冷总督说的那样：有争执，咱们政治家才派得上用场，如果他们停下来不打了，你就失业了。

人工智能到底是什么？给一门学科界定范围很难，尤其是这门学科还在快速变化中。即使是数学这样的成熟学科，有时我们也理不清边界，而像人工智能这样朝令夕改的，更是不容易闹清楚了。人工智能的定义素无共识。在大学里，机械系、电子系、计算机系，甚至哲学系都有人干人工智能。让这些人对这门学科取得共识谈何容易。从实用主义（哲学的“实用主义”，不是日常用语“实用主义”）看，一个学科就是学科共同体共同关注的东西。有些毛边可以宽容，演变。这种外延式的定义要比从上帝视角给一个内涵式定义更为实用。

一般认为，人工智能起源于 1956 年在达特茅斯学院召开的夏季研讨会。国内关于达特茅斯会议和神经网络早期历史的各种段子很多源于我几年前的两篇博客，后来被《上海书评》转发。经过修订，我把它们重新编

为本书的两章：“达特茅斯会议：人工智能的缘起”和“神经网络简史”。“计算机下棋”一章的大部分也在《南方周末》发表过。“自动定理证明兴衰纪”的核心内容在《中国计算机学会通讯》连载过。

明尼苏达大学的查尔斯·巴贝奇研究所一直在做计算机科学的口述历史，采访了很多对计算机科学有影响的人，其中有相当一批是人工智能学者。大部分的采访都有录音。除了翻阅各种文献外，我听了近 100 小时的采访录音，许多人工智能老一代革命家临终时话都说不利索，听这种东西除了兴趣，还得有体力。

图灵大概是第一个对智能做出深刻思考的智者。他 1936 年的文章“可计算的数”奠定了计算机科学的理论和实践基础，也把相关的哲学思考推进了一大步，以至于哲学家蒙克（Ray Monk）把他列为有史以来最伟大的十位哲学家之一。图灵 1950 年在哲学杂志《心》（*Mind*）上发表的文章“计算机与智能”是传世之作，但这篇文章没有靠谱的中文翻译，我将我的译文和一篇图灵小传附在书后作为附录。图灵小传的一个早期版本曾出现在我的《哲学评书》一书中，但新版本融入了一些近几年关于图灵研究的新成果。“人工智能”这个词组的出现和达特茅斯会议有关。但英国学术圈在 1956 之前和之后的很长一段时间一直在用“机器智能”的说法，这和图灵 1950 年的文章有关。一般认为，这篇文章是这个学科的源头。但后来发现图灵 1948 年在英国国家物理实验室（NPL）写过一个内部报告，题为“智能机器”，其中提到了“肉体智能”（embodied intelligence）和“无肉体智能”（disembodied intelligence）的区分。机器人学家布鲁克斯（Rodney Brooks）认为图灵 1948 年的报告比图灵 1950 年的文章更加重要，它从某种意义上预示了后来符号派和统计派之争。这段历史我也列在附录里，放在图灵小传之后，因为我觉得先读读图灵的生平也许会有助于理解他的思想。

本书每一章几乎都可单独阅读，大部分内容，对于受过高中教育的人，应该都不难懂。但第 10 章是个例外，这一章企图以严肃的态度探讨人工智能。我以一种浓缩的方式讲述了图灵机、丘奇-图灵论题、相似性原则和超计算。没有计算理论，很多人工智能的基础问题实在是拎不清。如果读者觉得吃力，可以跳过这一章。

我常用的一种历史研究工具是谷歌的 Ngram。谷歌扫描了三千多万本书，把书中出现的词组的词频统计结果公布。以时间为横轴、词频为纵轴画一条曲线，就可看出特定的词在不同历史时间段的兴衰，从而得出某些洞察。例如，通过比较“United States are”和“United States is”在历史上出现的频率，就可看出美国人是何时开始认同美国作为一个统一的国家的。很明显，南北战争之后，“United States is”开始变得更常用。我们通过统计若干人工智能中关键词的 Ngram，可以感知人工智能在不同阶段的宏观发展。我曾经写过一篇“计算历史学”（见《哲学评书》）介绍 Ngram。大数据为历史学提供了有力而令人信服的工具。

科普有一种写法：用一些貌似通俗的语言去解释复杂的原理。我一直不大相信这种方法，无论作者是内行还是专业科普作家。我压根就没见过一本可以把量子力学解释清楚的科普书。即使简单如图灵机，也鲜有适当的普及读物。倒是那些讲历史和八卦的书引人入胜，安德鲁·霍奇斯的《艾伦·图灵传：如谜的解谜者》是内行写作的典范，而数学家索罗门·费佛曼的太太安妮塔·费佛曼的两本逻辑学家传记是我心目中的标杆。戴森（Freeman Dyson）一直是我喜欢的作者，他也时不时为《纽约书评》写写八卦，趣味和我接近，我总是从阅读他的文章的过程中收获良多。即使我不懂他的数学和物理的领域，也能时有洞察。我的书单上还有蒙克的所有传记，它们既高级又有趣。就像蒙克所说，历史可以帮助内行

了解知识的进化并获得新的视角，同时也为外行人了解专业知识提供入门的台阶或向导。

读大科学家写的科普著作，最有意思的倒不是那些对成熟思想的通俗叙述，而是那些对不成熟看法的披露，还有不好意思写到正经学术论文里的自负和牢骚。恰因为这个原因，我也喜欢多依奇（David Deutsch）的几本书。

我们很少有机会在学科发展之初就能把学科脉络梳理清楚。过去有过几个这样的时间段，例如 1900 年到 1950 年的逻辑学，1945 年到 2000 年的分子生物学和 1950 年到当下的语言学。本书除了想梳理始于 20 世纪 40 年代的人工智能的历史外，还有一个作者隐含的心愿：作为人工智能的科普。哈代曾说科学（尤其是数学和理论物理，也许理论计算机科学）和艺术的原创需要一等的智力，解释和欣赏（例如乐评家和书评家）是二等的智力活儿。本书假想的对象是那些有能力但又是外行的人。丘成桐曾说（大意）：要想做大学问，必须先培养对学问的感情。除了科普，我还希望能帮助一小撮内行人或准内行人培养感情。我尽可能地列出了相关的参考文献供进一步学习。人工智能毕竟不是超弦理论，凭着一些智力还是可以自学的。

本书写作得到白硕、陈利人、宫力、洪涛、刘江、马少平、毛德操、施水才和赵伟等诸位师友的帮助和指点，特此致谢。乌镇智库的同仁为本书提供了必要的数据库，我的助理冰冰为我提供了多方面的支持，一并谢过。

目 录

第 1 章	达特茅斯会议：人工智能的缘起	1
	1. 背景	1
	2. 达特茅斯会议	6
	3. AI 历史的方法论	9
	4. 会议之后	14
	5. 预测未来：会有奇点吗？	19
第 2 章	自动定理证明兴衰纪	24
	1. 自动定理证明的起源	24
	2. 罗宾逊和归结原理	32
	3. 项重写	34
	4. 阿贡小组和马库恩	35
	5. 符号派的内部矛盾：问答系统和归结原理的失落	37
	6. 几何定理证明与计算机代数	39
	7. 定理证明系统和竞赛	44
	8. 哲学问题	46
	9. 现状	49
	10. 结语	51

2 // 目 录

第 3 章	从专家系统到知识图谱	60
	1. 费根鲍姆和 DENDRAL	60
	2. MYCIN	64
	3. 专家系统的成熟	65
	4. 知识表示	66
	5. 雷纳特和大知识系统	70
	6. 语义网	73
	7. 谷歌和知识图谱	75
第 4 章	第五代计算机的教训	79
	1. 背景	79
	2. 理论基础：逻辑程序和 Prolog	82
	3. 五代机计划和五代机研究所	85
	4. 并发 Prolog	88
	5. 美国和欧洲对日本五代机计划的反应	90
	6. 结局和教训	94
	7. 日本还有机会吗：日本下一代人工智能促进战略	95
第 5 章	神经网络简史	97
	1. 神经网络的初创文章	97
	2. 罗森布拉特和感知机	103
	3. 神经网络的复兴	107
	4. 深度学习	111
第 6 章	计算机下棋简史：机定胜人，人定胜天	116
	1. 机器下棋史前史	116
	2. 跳棋插曲	118

3. 计算机下棋之初	119
4. “深蓝”	124
5. 围棋和 AlphaGo	125
第 7 章 自然语言处理	128
1. 乔治敦实验	128
2. 乔姆斯基和句法分析	129
3. ELIZA 和 PARRY	136
4. 维诺格拉德和积木世界	143
5. 统计派又来了	149
6. 神经翻译是终极手段吗?	151
7. 问答系统和 IBM 沃森	152
8. 回顾和展望	154
第 8 章 向自然学习：从遗传算法到强化学习	159
1. 霍兰德和遗传算法	159
2. 遗传编程	164
3. 强化学习	166
4. 计算向自然学习还是自然向计算学习	172
5. 计算理论与生物学	173
第 9 章 哲学家和人工智能	177
1. 德雷弗斯和《计算机不能干什么》	177
2. 塞尔和中文屋	184
3. 普特南和缸中脑	187
4. 给哲学家一点忠告	190

4 // 目 录

第 10 章 人是机器吗? ——人工智能的计算理论基础	195
1. 丘奇-图灵论题: 为什么图灵机是最重要的发明?	197
2. 相似性原则: 另一个重要但不太被提及的计算理论思想	201
3. 超计算	205
4. BSS 实数模型	206
5. 量子计算	208
6. 计算理论的哲学寓意	211
7. 丘奇-图灵论题、超计算和人工智能	212
第 11 章 智能的进化	216
1. 大脑的进化	216
2. 能源的摄取和消耗	218
3. 全社会的算力作为文明的测度	220
4. 人工智能从哪里来?	222
5. 人工智能向哪里去: 会有超级智能吗?	223
第 12 章 当我们谈论生死时, 我们在谈论什么?	230
附录 1 图灵小传	237
附录 2 人工智能前史: 图灵与人工智能	249
附录 3 冯诺伊曼与人工智能	255
附录 4 计算机与智能	261
参考文献	293
人名对照	308

第 1 章

达特茅斯会议：人工智能的缘起

What is past is prologue.

过去只是序幕。

——William Shakespeare (莎士比亚)

1. 背景

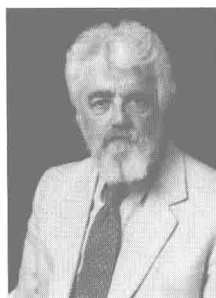
现在一说起人工智能的起源，公认是 1956 年的达特茅斯会议。殊不知还有个前戏，1955 年，美国西部计算机联合大会（Western Joint Computer Conference）在洛杉矶召开，会中还套了个小会：学习机讨论会（Session on Learning Machine）。讨论会的参加者中有两个人参加了第二年的达特茅斯会议，他们是塞弗里奇（Oliver Selfridge）和纽厄尔（Allen Newell）。塞弗里奇发表了一篇模式识别的文章，而纽厄尔则探讨了计算机下棋，他们分别代表两派观点。讨论会的主持人是神经网络的鼻祖之一皮茨（Walter Pitts），他最后总结时说：“（一派）企图模拟神经系统，而纽厄尔则企图模拟心智（mind）……但殊途同归。”这预示了人工智能随后几十年关于“结构与功能”两个阶级、两条路线的斗争。

开聊达特茅斯会议之前，先说说 6 个最关键的人。首先，会议的召集者麦卡锡（John McCarthy）当时是达特茅斯学院的数学系助理教授。

1954 年，达特茅斯学院数学系同时有 4 位教授退休，这对达特茅斯这样的小学校而言真是不可承受之轻。刚上任的年轻系主任克门尼(John Kemeny)之前两年才在普林斯顿大学逻辑学家丘奇(Alonzo Church) 门下取得了逻辑学博士，于是跑到母校求援。这么说起来，克门尼算是图灵的师弟，他战时和物理学家费曼一起工作，还一度当过爱因斯坦的数学助理，后来一头扎在计算机研究里，和麦卡锡一起琢磨出了分时系统。他 1955 年在《科学美国人》杂志上写过一篇文章“把人看作机器”(Man Viewed as a Machine)，介绍了图灵机和冯诺伊曼^①的细胞自动机(最早叫“自生机”)，文章的简介提到“肌肉机器”(muscle machine)和“大脑机器”(brain machine)。所谓“大脑机器”就是人工智能的另一种说法而已。克门尼最为人知的工作应该是发明了老少咸宜的编程语言 BASIC。现在估计已经没人知道 BASIC 语言发明人曾是 LISP 语言发明人的老板。克门尼是天生的官僚，后来位居达特茅斯学院数学系主任和校长，美国三里岛核电站出事时，总统委托他当调查委员会主席，这是后话。

克门尼从母校数学系带回了刚毕业的 4 位博士前往达特茅斯学院任教，麦卡锡是其中之一。麦卡锡后来发明的 LISP 语言中最重要的功能 Eval 实际就是丘奇的 λ 演算，而且他后半生致力于用数理逻辑把常识形式化。大家由此猜测他可能也是丘奇的学生，但其实不是，他学的压根就不是逻辑。他的老师是失去双手的代数拓扑学家莱夫谢茨(Lefschetz)。但麦卡锡对逻辑和计算理论一直有强烈兴趣，他 1948 年本科毕业于加州理工学院，在学校主办的 Hixon 会议上听到冯诺伊曼关于细胞自动机的讲座，后来他刚到普林斯顿大学读研究生时就结识了冯诺伊曼，在老冯影响下开始对在计算机上模拟智能产生兴趣。

① 我故意没有在“冯”和“诺伊曼”之间加那个讨厌的点儿，因为在更多时候，查找参考文献时，他的姓是列在 V 下，而不是 N 下。



麦卡锡 (1927—2011)

达特茅斯会议的另一位积极的参加者是明斯基。他也是普林斯顿大学的数学博士，和麦卡锡在读书时就相熟。他的主业也不是逻辑，尽管他后来写过一本很有影响力的计算理论的书，还培养过好几个计算理论的博士，其中就有图灵奖获得者布鲁姆 (Manuel Blum)。布鲁姆目前和他老婆 (Lenor Blum，就是实数计算模型 BSS 的 B)、儿子一家三口都在卡内基梅隆大学任教。明斯基的理论情结和丘奇关系也不大，他的老师塔克 (Albert Tucker) 是莱夫谢茨的学生，主要做非线性规划和博弈论，多年来担任普林斯顿大学数学系主任，出身数学世家，儿子、孙子也都是数学家。按辈分论，麦卡锡还是明斯基的师叔。塔克的另一名出色的学生后来得了诺贝尔经济学奖，他就是心灵美丽的纳什。纳什比明斯基小一岁，但比他早 4 年拿到博士学位，也算是明斯基的师兄了。明斯基的博士论文是关于神经网络的，他在麻省理工学院 150 周年纪念会议上回忆说是冯诺伊曼和麦卡洛克 (Warren McCulloch) 启发他做了神经网络。有人还找过他麻烦，质疑说神经网络的研究算数学吗，倒是老冯力挺说：现在不算，但很快就得算。倒是明斯基自己后来和神经网络结下梁子，那段故事见本书第 5 章“神经网络简史”。明斯基的熟人都认为他是无所不通的天才，他的忘年交沃尔弗拉姆 (Stephen Wolfram) 称，他晚年计划写本神学的书，但去世时书还没影子。

塞弗里奇被后人提及不多，但他真是人工智能学科的先驱，他在麻省理工学院时一直和神经网络的开创人之一麦卡洛克一起在维纳（Norbert Wiener）手下工作，他是维纳最喜欢的学生，但没读完博士学位。维纳《控制论》一书的第一个读者就是塞弗里奇。塞弗里奇是模式识别的奠基人，他写了第一个可工作的 AI 程序。他后来在麻省理工学院参与领导 MAC 项目，这个项目后来一分为二：计算机科学实验室和人工智能实验室。但分久必合，现在这两个项目又合并了，变成了 MIT CSAIL。顺便给女读者添点料：塞弗里奇的爷爷就是英国第二大百货店塞尔福里奇（Selfridges）的创始人。所谓“顾客永远是对的”（The customer is always right.）就出自塞尔福里奇，他本是美国人，后到英国创业，发财后老婆就死了，于是勾搭上一对匈牙利双胞胎歌舞演员，出入赌场，赔光了家业。他的故事 2013 年还被有意思的英国人拍成了电视剧。塞尔福里奇百货几经周转，现在的主人是美国百货公司希尔斯（Sears）。塞尔福里奇百货和隔壁的哈罗德百货支撑着牛津街的零售业，现在大概一半顾客来自中国二线城市。

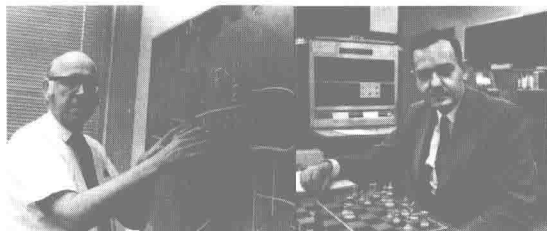
信息论的创始人香农（Claude Shannon）被麦卡锡拉大旗做虎皮也请到会上打酱油。其实麦卡锡和香农的观点并不一致，平日相处也不睦。香农的硕士、博士论文都是讲怎么实现布尔代数的，当时麻省理工院校长布什（Bush）亲自指导。博士毕业后他去了普林斯顿高等研究院，曾和数学家外尔（Hermann Weyl）、爱因斯坦、哥德尔等共事。战争中，他一直在贝尔实验室做密码学的工作，图灵在 1943 年曾秘访美国，和同行交流破解德国密码的经验，其间和香农曾有会晤，一起聊过通用图灵机。战后香农去英国还回访过图灵，一起讨论过计算机下棋。香农内向，从没说过这段往事，直到 1982 年接受一次采访时才提起。1950 年香农在《哲学杂志》发表过一篇讲计算机下棋的文章，为计算机下棋奠定了理论基础。香

农比其他几位年长十岁左右，当时已是贝尔实验室的大佬。



香农 (1916—2001)

另外两位重量级参与者是纽厄尔和司马贺 (Herbert Simon)。纽厄尔是麦卡锡和明斯基的同龄人，他硕士也是在普林斯顿大学数学系读的，按说普林斯顿大学数学系很小，他们应有机会碰面，但那时纽厄尔和他俩还真不认识。他们第一次见面，纽厄尔回忆是在 IBM，而麦卡锡回忆是在兰德公司。纽厄尔的硕士导师就是冯诺伊曼的合作者、博弈论先驱摩根斯顿，纽厄尔硕士毕业后就迁往西部加入著名智库兰德公司。他在兰德开会时认识了塞弗里奇，并受到对方做的神经网络和模式识别的工作的启发，但方法论走的却完全是另一条路。



纽厄尔 (1927—1992) 与
司马贺 (1916—2001)

司马贺比他们仨都大 11 岁 (怀特海比罗素也大 11 岁)，那时是卡内基理工学院 (卡内基梅隆大学的前身) 工业管理系的年轻系主任，他在兰德公司学术休假时认识了纽厄尔。司马贺后来把纽厄尔力邀到卡内基梅隆