

人工智能的进化

计算机思维离

人类心智

还有多远？

COMMON SENSE,

THE TURING TEST

AND THE QUEST

FOR REAL AI

让我们重回人工智能的起点，
检视人工智能的最初梦想。

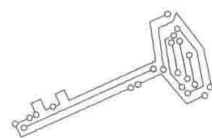
Hector Levesque

「加」赫克托·莱韦斯克
王佩 译

著

人工智能
有常识吗？
它能否
像我们一样
随机应变？

中信出版集团



Hector Levesque

人工智能的进化

计算机思维离

人类心智

还有多远？

「加」赫克托·莱韦斯克 著
王佩 译

图书在版编目 (CIP) 数据

人工智能的进化 / (加) 赫克托·莱韦斯克著; 王佩译. -- 北京: 中信出版社, 2018.4

书名原文: Common Sense, the Turing Test, and the Quest for Real AI

ISBN 978-7-5086-8567-0

I. ①人… II. ①赫… ②王… III. ①人工智能 - 普及读物 IV. ①TP18-49

中国版本图书馆 CIP 数据核字 (2018) 第 018479 号

Copyright © 2017 Massachusetts Institute of Technology

Simplified Chinese translation copyright © 2018 by CITIC Press Corporation

ALL RIGHTS RESERVED

本书仅限中国大陆地区发行销售

人工智能的进化

著 者: [加] 赫克托·莱韦斯克

译 者: 王 佩

出版发行: 中信出版集团股份有限公司

(北京市朝阳区惠新东街甲 4 号富盛大厦 2 座 邮编 100029)

承 印 者: 北京楠萍印刷有限公司

开 本: 880mm × 1230mm 1/32

印 张: 6.75 字 数: 122 千字

版 次: 2018 年 4 月第 1 版

印 次: 2018 年 4 月第 1 次印刷

京权图字: 01-2018-0780

广告经营许可证: 京朝工商广字第 8087 号

书 号: ISBN 978-7-5086-8567-0

定 价: 49.00 元

版权所有·侵权必究

如有印刷、装订问题, 本公司负责调换。

服务热线: 400-600-8099

投稿邮箱: author@citicpub.com

献 给

——

帕 特

前 言

本书旨在从人工智能（AI）的角度探讨人类的心智。研究人类思维的朋友可能会感到奇怪，人工智能怎么会懂得人类的心智呢？人工智能属于计算机科学，当然是和计算机打交道的，而心智是人类所特有的东西。当然，或许未来某天我们会像探讨“动物心智”一样，讨论一下“计算机心智”。但是，让人工智能解释人类心智，是不是有点儿像让天文学家解释蛀牙问题，有些风马牛不相及呢？

答案是否定的，这两者绝非风马牛不相及。因为计算机科学并非仅关注计算机本身，计算机科学主要关注的是计算，而计算是个过程，比如对一串数字进行分类排序、压缩音频文件或去除数字图像的红眼等。当然，这个过程通常要由电脑完成，但是，人工或者某种设备也可能完成。

人工智能（或至少一部分人工智能）的基础假设认为，日常思维（人们每天使用的思维）同样也是一种计算过程，在对此进行研究时，无须过多考虑是谁或是什么东西在进行这种思维过程。

这也正是本书的主要内容。

有一点引起了争议，同时也更有趣，那就是当今存在不止一种人工智能研究流派，而且各家都在著书立说。在不同的人工智能研究流派当中，确有一部分是在研究各种思维形式，其他则对其敬而远之。

究其原因，我们不妨先考虑一下金鸡独立这个动作。我们有多希望能将思维、规划以及解决问题的能力融为一体？人类个体能否做到这点呢？我们是否认为事先对这一主题进行研究会有所帮助？比如看看类似《傻瓜单腿站立指南》这种书？如果一位人工智能研究人员想造出能单腿站立的机器人（不管出于任何原因），他绝对认为涉及思维的人工智能没什么借鉴意义。

当今研究人工智能的众多流派中，绝大部分都已背离了 20 世纪 50 年代开始的对于思维（和规划、解决问题）的原始研究。包括研究方向和研究目标等涉及该研究的基本理念都已发生改变，而且这种变化在最近兴起的机器学习当中尤为突出。毫无疑问，从纯技术的角度来看，这种变化取得了巨大成功，而且极有可能超越人工智能的其他流派。虽然对机器学习的研究极度依赖数据统计，但它的本质与人工智能的原始研究仍然大不相同。

本书的目的就是回过头来，重新思考人工智能这一概念（即

现在所谓的“老式人工智能”），同时解释为什么这一概念出现 60 年以来，我们依然对于研究人类思维兴致不减，而且在此过程中，我们不仅推出了几项实用技术，还从中学到很多东西。

一本关于人工智能的著作可以承载多种内容。它可以是人工智能的教科书，可以是关于人工智能新技术的调查，可以讲述人工智能的历史沿革，甚至可以回顾电影当中对于人工智能的刻画。而这本书并非以上提及的种种。这本书讲述了人工智能背后的理念和假说、人工智能的理论基础，讨论了人工智能看待事物的独特方式及原因，并介绍了人类的心智及其如何产生智能行为。

本书全部文本只用了大约 150kb（千字节）的存储空间，看到这么一本小书，让人多少有些沮丧。相比而言，我电脑里一秒钟的维瓦尔第音乐会视频所占的空间都比它的两倍还大。因此，单纯从原始数据来看，我电脑里半秒钟的视频与本书分量相当。但是在半秒钟的时间里，乐队指挥都来不及举起指挥棒。

正如古语所说：一画（帧）胜千言。然而，我认为还有另外一点值得注意，与图画相比，语言是更为紧凑的信息载体。几百个字呈现在视频当中，或许只是一瞬间，却也能将炖牛肉的方法解释清楚。由于人类的进化，我们已经能够随心所欲地使用这些超紧凑型的信息载体。我们通过在他人的耳边低语，就能对其日后的行为产生影响。人类是如何做到这一点的，正是我们现在所探究的问题。我希望读者能够喜欢我的观点，也能感受到我在探究过程中的激动心情。

目 录

前 言 v

第 1 章 什么是人工智能? 001

自适应机器学习 003

老式人工智能 006

具有常识的程序 008

图灵测试 010

中文房间理论 012

第 2 章 巨型拼图之谜 017

疑问接踵而来 019

留给我们的困难 023

一条解决的途径 030

第 3 章 知识与行为 033

超越刺激与反应 036

知识与信念 041

意向立场 043

智能行为 046

能力与表现 048

第 4 章 真智能还是假智能？ 051

谈话机器人 055

投机取巧不可行 060

威诺格拉德模式 063

我们从中得到的经验 070

GOFAI 的回归 072

第 5 章 有经验的学习与没有经验的学习 075

我们如何学习词语？ 078

我们如何学习事实？ 081

我们如何学习行为？ 083

我们如何超越经验？ 085

第 6 章 书本智慧与市井智慧 089

语言的影响力 092

书本智慧 094

海伦·凯勒的智慧 095

书本中的市井智慧 099

第 7 章 长尾理论与培训的界限 103

长尾现象带来的难题 106

如何处理突发事件 110

无意识与有意识 113

威诺格拉德模式测试 115

是不是技巧还不够完美? 118

第 8 章 符号与符号处理 119

符号的代数运算 121

符号的逻辑运算 125

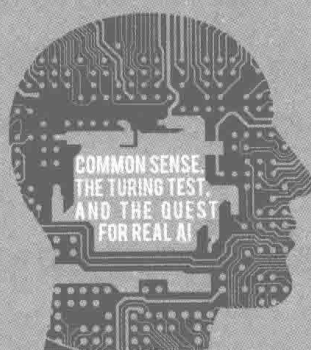
符号的意义 129

一切都源自图灵机 131

第9章	基于知识的系统 137
	符号能够解决的问题 139
	用符号表示无形 141
	知识表示假说 143
	假设是真的吗? 144
	知识表示与推理 146
	谁才是唯一的选择? 149
第10章	人工智能技术应用 151
	人工智能的未来 155
	自动化是好是坏? 159
	超级智能与奇点 161
	真正的风险 164
	超越进化 166
致 谢	169
注 释	173
参考文献	193

第 1 章

什么是人工智能？



本书旨在从人工智能的角度探讨人类的心智。对于人类的心智，我们会有很多话说，但首先，我们转向人工智能领域本身，并简要介绍本书将会深入探讨的一些想法和理念。

自适应机器学习

当下，人工智能在科技领域受到普遍关注。各大报纸、杂志的科技和商业版块，几乎每个星期都会谈到人工智能，它们告诉我们，诸如微软、IBM（国际商用机器公司）、谷歌和苹果等大型科技公司都在人工智能的研发领域大量投入，其他公司也都竞相效仿。2015 年 11 月，丰田公司宣布在人工智能领域投资 10 亿美元；2015 年 12 月，埃隆·马斯克（Elon Musk）宣布新成立一家名为 OpenAI 的非营利性企业，研究人工智能，并对其另外注资

10 亿美元。

资本市场为何会出现这样的喧嚣？如果仔细研究一下投资者对于这数十亿美元的预期，你会发现，这里的人工智能技术似乎与科幻小说和电影中想象的人工智能大相径庭。基本没人谈论诸如电影《银翼杀手》(*Blade Runner*) 中所看到的人形智能机器人，或者电影《2001：太空漫游》中的高能无形智能机器（哈尔 9000 计算机）。

当下，众多科技公司摩拳擦掌准备开发的这种人工智能，更应该被称为“自适应机器学习”（adaptive machine learning），即 AML。广义来说，AML 的设计理念是让计算机系统通过对大量数据进行分析，掌握智能行为方式。确切地说，目前人们对于人工智能的兴奋点主要在于挖掘并应用我们通常所说的“大数据”。

在此，我们不去深究任何 AML 技术如何开展工作这样的细节问题，只讨论其背后的理念。

出于论证的需要，假设我们想让计算机系统对猫进行识别。那么，该系统的工作就是将给定的众多图像分成两类：一类图像上有猫的样子，另一类图像上没有。接下来的问题就是如何建立这样一个系统。过去，人工智能程序员往往会编写出一款程序，在图像中搜寻猫的一些具体特征。例如，搜索猫的脸：泛绿色或泛黄色的眼睛、杏仁状的竖直瞳孔、倒置的粉色三角形一样的鼻子、胡须等；也可能会搜索猫的轮廓：小小的脑袋、三角形的耳朵、四条腿和向上翘起的尾巴，还可能会搜索猫身上毛皮的独特

颜色。如果该程序在某个图像当中发现了足够多的上述特征，就会将该图像标记为一只猫；如果没有，则会将其忽略。

但是 AML 提出了一种完全不同的理念。首先，你向系统提供大量数字图像，其中一些是猫的图像，另一些不是。然后告诉系统通过某种方式压缩所有图像数据，即在众多图像的众多补丁当中搜索一组“特征”。这种特征可能是某一特定颜色和亮度，也可能是某个区域，区域的边缘在亮度和颜色上明显与其他部分不同。这一理念的目的是通过找到这样一组特征，进行重新组合并构建与原始图像相似的内容。然后，你让系统根据这些特征继续进行抽象分析，在这些特征当中寻找普遍性，随后如此继续进行几个层次的分析。

这一切都是在不告知系统具体要搜索什么的情况下完成的。如果原始图像里有很多猫的图像，那么系统就很可能在更高层面上分析出一些与猫相关的特征。而且重要的是，这些特征全部由原始图像决定，而非人工智能程序员决定如何在图像中识别猫。正如斯坦福大学的吴恩达（Andrew Ng）所说：“你将海量数据输入程序，让数据说话，让软件自动从数据中学习。”

人们发现，AML 在这种所谓“无监督”的环境下运转极其良好，完全超出几十年前人工智能研究人员做出的预测。AML 的成功可以归结于以下三点：拥有可供分析处理的海量数据（通过在线、专业存储设备或传感装置获取），掌握功能强大的能够处理这些数据的计算技术，以及高速计算机。而这些条件，在 30 年前并

不具备。

当然，不会有人愿意斥资数十亿美元，仅仅是为了识别猫。但是你想一下，如果这些不是猫而是乳腺的影像，其中一些还藏有医生难以发现的肿瘤，结果会怎样？或者假设这些根本就不是可视数据，而是一段段录音，其中包含有关人士所说的几句话？或者假设这是银行的交易数据，其中还涉及欺诈或洗钱？或者是人们在线购物过程中浏览和购买商品的历史记录，抑或是与汽车前风挡玻璃处视觉数据相对应的脚踏板和方向盘运动轨迹？计算机系统能够自动分析这些海量数据，从中获得规律，并应用于极具经济和社会效益的领域。

老式人工智能

虽然 AML 技术的发展趋势让人心潮澎湃、激动异常，但是这并不是本书关注的重点。事实上，随着本书逐渐展开，你会发现，除了最后的第 10 章以外，我们基本不会讨论任何人工智能技术。

本书关注的是一个与众不同的人工智能发展愿景，该愿景于 20 世纪 90 年代由约翰·麦卡锡（John McCarthy）、马文·明斯基（Marvin Minsky）、艾伦·纽厄尔（Allen Newell）与赫伯·西蒙（Herb Simon）等人首先提出，我们且称之为老式人工智能（Good Old-fashioned AI），或简称 GOF AI。这些科研人员提出的另外一种智能并不致力于分析海量数据，而是专注于常识。