



《千脑智能》, [美]杰夫·霍金斯著, 廖璐、熊宇轩、马雷译, 浙江教育出版社 2022年9月出版, 定价: 109.90元

2016年2月, 一个阳光明媚的中午。

美国工程院院士、计算机科学家与神经科学家杰夫·霍金斯坐在 Numenta (一家神经科学研究公司) 的办公室里, 一边思考大脑如何运作以及人工智能未来发展相关问题, 一边等着和妻子珍妮特吃午饭。

霍金斯下意识地伸手去拿桌子上的咖啡杯, 就在指尖划过咖啡杯边缘那一刻, 他忽然跳了起来, 跑出门去, 并一头撞上来找他的妻子。

“大脑新皮质知道所有东西的位置。”霍金斯嚷嚷着, 跑向 Numenta 负责研究的副总裁萨布泰·艾哈迈德, 并开始解释自己的想法, 但艾哈迈德不明白他在说什么。

那一刻, 这位“深耕”大脑 40 余年的“科技狂人”有了新发现——他在“神秘丛林”中发现了“新皮质中的地图状参考系”。

此后 3 年多时间里, 霍金斯通过实验验证发现, 并发表数篇论文, 一步步完善了自己的“千脑智能”理论, 并用 18 个月时间把这些思考写成《千脑智能》。

### 和先贤站在一起

霍金斯被视为科技界的“一代传奇”。

在北京智源人工智能研究院院长、北京大学教授黄铁军眼里, 霍金斯更是一位“智能时代的孤胆英雄”, 因为很难用科学家、企业家、发明家或其他什么称呼概括他。

1979年9月, 霍金斯从美国康奈尔大学电子工程系毕业, 进入英特尔公司工作。那一年, 诺贝尔奖得主、基因双螺旋结构发现者弗朗西斯·克里克在《科学美国人》上发表了《思考大脑》, 呼吁建立

## 他, 智能时代的孤胆英雄

● 本报记者 张双虎

一种包罗万象的大脑理论, 从而解释这种“极其神秘”的人体器官。

霍金斯被克里克的文章深深吸引, 决定寻找大脑运行背后的理论框架, 但英特尔虽然是“电脑”企业的领头羊, 却没有研究“大脑”的部门。

1986年, 霍金斯到美国加州大学伯克利分校攻读神经科学博士学位, 希望能弄清楚智能的运作方式。由于这项研究风险较大, 他想专注于相关理论研究的提案再次被拒绝。

“当时霍金斯不知道如何开展工作, 也找不到这个方向的导师。”黄铁军说, “作为博士生导师, 我想说这样的学生特别难得, 绝大多数博士生习惯于导师指定研究方向甚至论文题目, 而不是追求自己内心热爱的研究方向。因此, 很多学校的博士生获得学位的比例很高, 但原始创新率很低。”

霍金斯泡了两年图书馆, 读了过去 50 年神经科学领域最重要的数百篇论文, 以及心理学家、语言学家、数学家和哲学家对大脑和智能的看法。

“这才是博士研究生的必修课。”黄铁军说, “不是导师告诉你做什么, 而是求教于追寻这一问题的所有先贤, 经过这一关, 霍金斯就已经和他们站到一起了。”

在加州伯克利的两年中, 霍金斯虽然没找到答案, 但人类对这个问题的认知水平他已经了然于胸。

霍金斯返回工业界, 开启了掌上电脑的开发, 成立 Palm 公司并取得巨大成功。2002年, 霍金斯用创业积累的资金成立了红木神经科学研究所, 聘任 10 位全职科学家, 吸引了 100 多名访问学者, 专注新皮质理论研究。但顶级神经科学家都知道, 大脑是个巨大的“神秘丛林”, 他们更愿意聚焦于力所能及的新发现, 而不是霍金斯追求的大脑理论。

因此, 霍金斯又把红木神经科学研究所捐给加州大学伯克利分校, 并创立了研究公司 Numenta, 自己带领团队专注于大脑理论研究。

### “千脑”并作

“这本书以及相关理论讲的是大脑,

大脑如何学习、大脑如何运作、大脑如何理解这个世界。”霍金斯解释说, “这个理论也和人工智能未来的发展方向相关。”霍金斯介绍说, 要学习大脑, 就是要知道如何实现身体的运动、世界是如何运作的, 以及大脑如何学习这个世界。大脑使用类似地图的结构建立关于世界的模型, 而且大脑的新皮质能同时建立数千个模型, 所以就像有上千个“大脑”同时在工作一样, 这也是《千脑智能》书名的由来。

“大脑是通过运动来学习的。”霍金斯在书中写道, “所以不管是‘学习’手里的咖啡杯, 还是一辆自行车, 或是一个抽象的概念——比如什么是千脑智能、什么是深度学习, 都是通过运动进行学习。”

首先, 我们要了解运动才能了解大脑, 但是在这方面还没有详尽的理论告诉我们大脑和运动之间的关系。霍金斯解释说, 如果手指在杯子侧面, 然后向顶部移动, 大脑就会预测我将感觉到杯嘴的圆弧。大脑在手指接触杯嘴前就作出了预测, 大脑需要什么信息才能作出这种预测? 答案显而易见, 大脑需要知道这两件事: 它所接触的是什么物体, 以及手指将到达杯子的什么位置。

“大脑需要知道手指相对于杯子的位置, 这一发现意味着, 新皮质中一定有神经元负责表征手指在一个参考系中的位置, 这个参考系同样与杯子相关。”霍金斯进一步解释, 当身体的不同部位, 如指尖、手掌、嘴唇可能会同时接触咖啡杯, 每一个接触杯子的部位都会基于它在杯子上的独特位置对其感觉进行单独预测。

大脑不是在作一个预测, 而是同时作几十甚至几百个预测。大脑通过视觉、听觉、味觉等各种感觉进行预测, 并利用参考系定义一个物体, 从而学习和认识事物。

无论是人、猫、狗还是其他动物, 体验的是同样的世界, 对于哺乳动物来说, 需要建立起数十亿个同样的世界的模型, 进而理解看不到的世界, 比如基因、分子、宇宙。人们通过建立参考系进行学习, 同时, 学习的结果又丰富了我们建立的参考系, 它是一个互相促进、互相塑造

的过程。

### 先理解大脑再创造 AI

“我相信, AI 的未来将基于这些大脑原理。”霍金斯在书中写道, 弄清楚人类大脑的工作方式, 然后对大脑进行逆向工程, 并开发模拟器功能的软件, 才能实现真正的人工智能。

在书中, 霍金斯用 16 个章节介绍了千脑智能理论、人工智能的未来和人类智能的未来。他认为该理论是理解大脑方面向前迈出的一大步, 虽然还有很多事情要做, 但这是完整大脑理论中缺失最大的部分。

早期的人工智能不研究大脑, 纯粹依靠逻辑和算法, 整个智能是人设计出来的, 不研究智能发生背后的脑的机制是什么。现在热门的人工智能是大家在讨论深度学习、强化学习算法的时候, 关注到背后执行这些算法的是一个神经网络。

“当说到神经网络的时候, 一定是和脑建立了联系。但今天讲的人工神经网络, 是人设计的一个网络, 只是参考了生物神经网络的一些最基本的知识, 并未参照生物神经网络的信息加工原理。”黄铁军解释说, “比如皮质柱的结构在人工智能里是没有的。而人工智能的神经元节点结构是人设计的, 算法也是人设计的, 尽管它参考了脑, 但和真正大脑的距离还很远。而千脑理论讲脑的信息处理以及脑底层的基本结构, 这是人工智能下一步发展的方向。”

比尔·盖茨也推荐了这本书, 他说: “更深入地了解大脑中被称为新皮质的部分是我们开发真正通用人工智能的关键, 这正是《千脑智能》这本书的核心内容。”

“如果霍金斯是对的, 通用人工智能也许就不再遥不可及。”清华大学基础科学讲席教授、脑与智能实验室首席研究员刘嘉说。

“如果成功的话, 千脑智能理论对脑科学的意义, 就像进化论对生物科学的意义一样重大。”黄铁军说, “我期望这个新学说能够成功, 至少作为大脑原理的一部分长存下来, 让这位追逐梦想 40 余年的孤胆英雄稍感慰藉。”