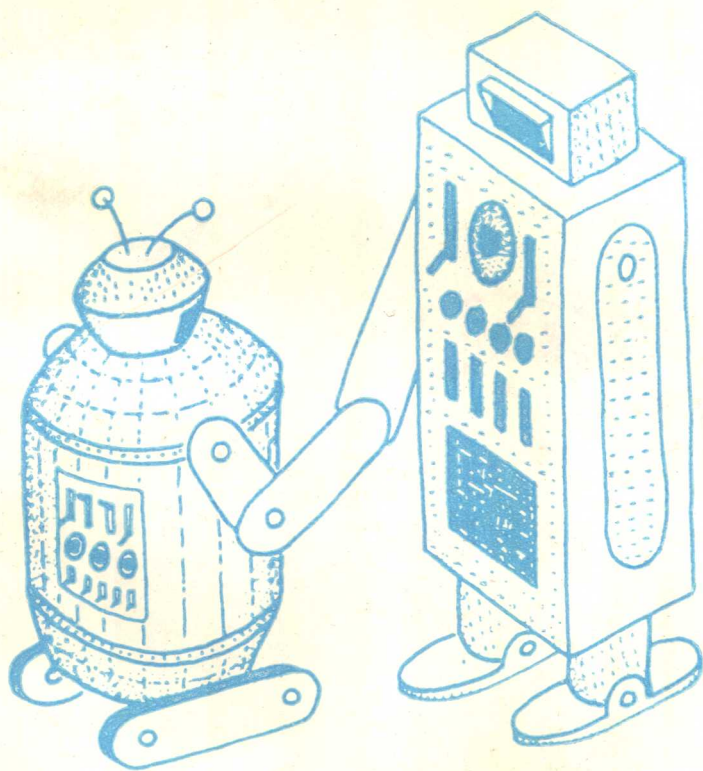


机器人与人类心智



〔美〕勃克斯著 游俊 桂起权 罗敏等译

机器人与人类心智

作者：〔美〕勃克斯

译者：游俊 桂起权 罗敏 任晓黎

校者：桂起权 任晓明

成都科技大学出版社

一九九三年·成都

Arthur W. Burks

Robots and Free Minds

the college of Literature,

Science, and Arts, The university of Michigan, 1986

本书根据密歇根大学文学、科学和艺术学院 1986 年版译出

(川) 新登字 015 号

机器人与人类心智

作者: [美] 勃克斯

译者: 游俊, 桂起权 罗敏 任晓黎

校者: 桂起权 任晓明

责任编辑: 何杨 陈正权

封面设计: 李文川

印刷: 郫县印刷厂

成都科技大学出版社出版发行

开本: 787×1092 毫米 1/32 印张: 4 字数: 100 千

版次: 1993 年 6 月第 1 版 1993 年 6 月第 1 次印刷

印数: 1—1, 500 册

ISBN7-5616-2374-7/G·432

定价: 2.70 元

内 容 提 要

这是一本探讨智能机器人如何实现人类自然功能问题的入门书。该书介绍了人工智能技术的历史发展以及计算机技术在发达国家的广泛应用；从哲学渊源上考察了哲学对人工智能研究的启迪作用，提出了智能化归理论和生物智能的意向性进化学说；认为应当将逻辑学、生物学、自动机理论结合起来，全方位研制生物目的性自动机，从而在广泛领域用机器人代替人去工作。作者还描述了人工智能发展前景和可能产生的弊端。

该书涉及自然科学和社会科学的广泛领域。既有学术著作的深刻内涵，又有讲演型文稿的幽默与生动。读者适应面宽，趣味性、可读性强。

译者前言

勃克斯 (Arthur W. Burks) 是美国当代著名计算机科学家、哲学家。《机器人与人类心智》是他的一部重要著作。

勃克斯 1915 年 10 月 13 日出生于美国明尼苏达州的德卢斯。1932 年毕业于伊利诺伊州巴达维亚高级中学。1936 年获数学学士学位, 1937 年在密歇根大学获哲学硕士学位, 1941 年在密歇根大学获哲学博士学位。1941 年到 1946 年他在美国宾夕法尼亚大学摩尔电子工程学院任教, 1943 年任该学院讲师和工程师。在这里, 他参加了第一台通用电子计算机 (ENIAC) 的设计、研制和改进工作。从 1946 年到 1948 年, 他担任普林斯顿高级研究所顾问。他与冯·诺依曼和赫尔曼·戈德斯坦一道, 发展了电子数字计算机的逻辑设计。这种设计被许多大学、政府研究机构以及国际商用计算机公司作为制作计算机的样板。

1948 年到 1954 年间, 勃克斯担任了底特律的伯勒斯公司的顾问。其主要工作是领导一个由伯勒斯资助在密歇根大学的研究小组, 研究用来处理数据的数字计算机的逻辑设计, 研究面向用户的程序语言以及将这种语言译为机器语言的程序, 还研究了自动机理论。这一时期的研究成果在《20 世纪计算的历史》(学院出版社 1980 年出版) 中得到了充分肯定。

从 1956 年起, 他担任美国密歇根大学逻辑与计算机研究组的主任。这个研究组主要研究自动机理论, 自适应系统, 自然系统的模拟等。

1960年夏天，勃克斯受聘为伊利诺伊大学应用数学研究所研究员。1967年起任密歇根大学计算机与通讯科学系教授、系主任。从1957年开始，勃克斯和另一位教授一起培养密歇根大学计算机与通讯科学方向的研究生，培养出许多这一研究方向的硕士和博士。开设了自动机理论的讲座和课程。从1954年开始，他还被聘为密歇根大学哲学系教授。

作为一个既是科学家又是哲学家的“双栖学者”，他集数学家、计算机科学家、人工智能专家、哲学家于一身，表现出非凡的才能和广博的学识。这一点，可从他这本很有特色的著作中看出端倪。

《机器人与人类心智》的中心议题是：一个有穷自动机可以实现人类的一切自然功能。换言之，勃克斯认为，一个机器人能实现人类的一切自然功能。这就是他所谓“人=机器人”命题。围绕着这一中心论题，他从哲学、逻辑学、生物学、计算机科学和人工智能等方面着手，开展了广泛而深刻的讨论。

该书介绍了人工智能的历史发展以及计算机技术在发达国家的广泛应用，着重讨论了它在社会控制和管理中的应用；从哲学渊源上考察了哲学对人工智能研究的指导和启迪；提出了智能“化归”理论以及人工智能和生物智慧的意向性进化学说；认为仅从逻辑机制的角度进行人工智能研究不可能导向人类心智的成功模拟，还应从心理学、神经生理学、生物学等角度进行全方位探索，研制生物目的性自动机，从而在广泛的领域里用机器人代替人去工作。作者还描述了人工智能技术发展的美好前景，指出已经和可能伴生的各种弊端（例如，计算机官僚主义、程序病等）。令人生趣，富于启迪。

该书 1986 年出版，由作者的 6 篇讲演稿组成。它广泛涉及哲学、逻辑学、心理学、生物学、计算机科学等学科的内容。既有学术著作的深刻内蕴，又具讲演型文稿的幽默与生动。兼具学术价值和科普价值。读者适应面宽，趣味性、可读性强。在这本小书中，作者以科学家的严谨态度，一一注明重要引文出处，必要时提供一些背景知识。同时，作者在每一讲中都穿插一个小故事，将枯燥乏味的计算机理论象讲故事一样娓娓道来，真正做到了寓教于乐、深入浅出。例如，书中谈到著名戏剧《罗塞姆的万能机器人》，生动地描述了剧中的机器人不具有情感和目的性，上刀山下火海在所不辞，甚至自己将被大卸八块，重新组装，也满不在乎，女主角海伦娜闻讯吓坏了，赶快去说服心理学家高尔博士改制了机器人的“生理相关器”，结果机器人有了情感，有了目的性。这种饶有兴趣的故事揭示了深刻的理论内涵，预示着人工智能发展的方向。由此可见，让科学性、精确性与通俗性、生动性溶为一体，相得益彰，是该书的一大特色。

为了使本书的主题更为突出，内容更为丰富。我们在书末增加了一个附录：逻辑学、生物学与自动机——历史的反思。该文是勃克斯在《人一机研究》杂志上发表的一篇论文。它从宏观角度、从历史角度考察了生物目的性自动机的历史渊源和未来的发展趋势，论证了逻辑学、生物学和自动机理论的相互依赖、相互渗透。读后会使人开阔视野，会让人发现一块神奇的领域，一个崭新的天地。

本书是由几位译者合作译出的。英文原著的提供者兼审校者任晓明，曾负笈京华，获硕士学位。尔后南下武汉，获哲学博士学位。在翻译过程中，主要负责全书的译名统一，补

加注释，改正译者错误，编辑整理译稿，撰写译者前言等工作。除此之外，还参与了第1讲的翻译工作。

第1讲和第2讲的译者桂起权，系博士生导师之一，治学严谨、造诣颇深，除了翻译第1讲和第2讲外，还担任了全书大部分译稿的审校工作。第3讲和第4讲的译者罗敏，本科毕业于北师大，尔后在国防大学上研究生，也参与了部分译稿的审校工作。游俊毕业于四川师院，翻译了第5讲、第6讲和附录，还参与了第3讲和第4讲的翻译工作。

本书付印前，为慎重起见，又作了一次审校，作为英文教师的任晓黎，参与了校译工作，并译出作者前言和有关英文材料。

本书的第3讲曾由高地译出，第4讲曾由胡耀鼎译出，第5讲曾由任晓明译出，分别发表于《哲学译丛》和《世界科学》。由于是摘译且有错漏，这次又依修订后的英文原本译出。译校过程中虽然尽了最大努力，但也许还会有错漏。于是，我们只得套用老话了，“还望读者不吝赐教”。

译者

1993年5月

前 言

本书由我的六个公开讲演组成。第一个讲演是我 1973 年在美国哲学学会西部分会担任主席时作的讲演。第二个讲演是密歇根大学 1978 年亨利·拉塞尔讲座的一次讲演。余下的四个讲演是在文学、科学和艺术学院 1982 年秋季举办的名教授讲座上发表的。这个讲座的最后一讲，题为“机器人与人类心智”实际上是第三次讲的，但我在本书中将其一分为二。

本书出版时做了修订，主要是删去重复之处，增加了新的材料。但是，我保持了讲演稿原有的风格，不刻意追求学术的精确性。对于这些讲演，我使用了各种各样的例证，例如自动奏乐箱、第一个通用电子计算机 ENIAC 的模块或插件。

我把这些讲演中表达的哲学思想称之为“逻辑机械论哲学”。萨尔蒙教授正在编辑我的这种哲学的论文集，其中有我对一些商榷意见的答复，以《逻辑机械论哲学》为书名的著作明年将由莱德尔出版社出版。在那本书中，我的答复将对我在这一表达的许多观点做出技术性的塑述。

既然这些讲演带有那么一点自传意味，那么给出一定的背景是有益的。我在大学学的是数学和物理，由于我的兴趣广泛，又由于我对数理逻辑感兴趣，后来我转而研究哲学。1941 年我在密歇根大学获得了哲学博士学位，但找不到工作，后来终于交好运，二战期间在宾夕法尼亚大学教电子工程学。在宾州，又交好运，我参加了第一个通用电子计算机

的设计。按高级计算机研究所的安排，我于是与冯·诺依曼共事，后来编辑和完成了他关于自再生自动机的专著。1946年，我去密歇根大学哲学系，但是继续开展计算机研究工作，1949年建立了逻辑和计算机研究小组，1957年与别人合作制定了计算机科学方面的博士生教学计划，1967年担任计算机和通讯科学系第一任系主任。

第一个讲演重印于《美国哲学学会会议录和讲话》1972—1973年合订本第46期第39—57页，美国哲学学会做了删节。第二个讲演重印于《自然和系统》1979年第1期第231—243页，编者对讲演做了删节。这一讲的中文译文由高地翻译，发表在《哲学译丛》，1985年第1期，第63—67页。《计算机与形而上学》由胡耀鼎翻译成中文，发表在上述杂志，1984年第3期，第16—23页。

P. O. 斯坦纳院长安排出版本书，副院长 E. W. 尼森做了牵线搭桥工作，A. 甘特为我的书出版而打字，M. 韦塞尔将文稿录入计算机，我的妻子艾丽斯与我讨论选材并提出编辑意见，我对他（她）们表示衷心的感谢。我的研究得到了文学、科学和艺术学院的支持，它给了我一个休假年的时间，我还得到国家文科研究员基金（1978—1979年）的资助，得到国家自然科学基金会的几项研究补助。

勃克斯

目 录

译者前言

前言

第一讲 逻辑，计算机与人

第二讲 计算机与社会控制

第三讲 计算机与哲学

第四讲 机器人与心智

第五讲 进化与意向性

第六讲 自由，道德与控制

附录 逻辑学、生物学与自动机

——历史的反思

后记

第一讲 逻辑、计算机与人

今晚我将提倡的论点是：一个有穷决定论自动机可以实现人类的一切自然功能。

你们大家都有一些关于有穷决定论自动机的知识，虽然多半不知道这个名称。一个通用电子计算机，当它无错误地运转时，就是一个有穷决定论自动机。

一个计算机的核心是它的内部计算系统，它储存并处理信息，包括它自己的指令。一个计算机也有输入-输出设备，而内部计算系统则通过它与外部世界进行交流。

为了将计算机和人类进行比拟，我们必须使它的输入-输出设备制得更加接近人类，就像电视摄像机对于眼睛，麦克风对于耳朵，传感元件对于气味、味道和温度等一样。作为电动输出装置，这种机器具有机械臂，它的手和手指可以操纵物体，它还有运动轮子和马达。一架完整的电视摄像机可以有伺服机构和使它跟踪运动目标的反馈回路，而且人造手的手指可以感觉压力和材料质地。为此，一个计算机可以像一个印度菩萨似的具有许多头、许多手臂和手，假如需要这些部件来实现人类的功能的话。

我真的说人是一台机器吗？是的，在某种意义上你和我都是一架机器。然而我的表述在三个哲学上重要的方面——

一般性、精确性和形而上学内容上都与传统说法不同。

我的论点比主张人是机器的传统学说更具一般性，因为无可争议，实际机械无论在元件水平上（机械的硬件或柔软的肉体）或者在程序和组织结构的水平上通常都是执行给定的功能的。例如，假如一台电子计算机可以证明或否证有趣的数学陈述，不管它是借助于专门的算法程序（就可靠的情况而言）还是借助于观察和处理图表和相似符号（那就是C. 皮尔士所认为的人类进行数学推理的方式）来完成的，那都是与我们的论点不相干的。

其次，我的人-机器人论点比“人是机器”的传统主张更精确。有穷自动机可以区分为图灵机、决定论自动机和概率论自动机。抽象地看，它们都是被精确定义的逻辑系统，或者是演绎的或者是归纳的。当我说一个“有穷自动机”时，我是在说一个精确地定义的逻辑系统在物理上的实现。

第三，“人是机器”的传统学说意味着唯物主义和附生现象主义，但我的表述并不仅如此。一个有穷自动机是逻辑系统在物理上的实现。为了将自动机与人相比拟，我们添加了输入-输出装置。因此，我主张一个给定的逻辑系统，当它在物理上实现了并附加输入-输出装置时，就可以实现作为一个一定的自然系统的人那样的功能。

因此，我不从卢克莱修和拉美特利那里，却从莱布尼兹那里选择我的原始材料。莱布尼兹说：

一切推理无非是概念的联结与代换，无论概念用词、符号或图形表达都如此。

假如我们能找到适于表达我们的思想的概念和

记号，就像作为算术表达式的数和作为几何表达式的直线一样确定和精确，我们就能在一切论题中在可以经受推理检验的限度内，完成在算术中几何中所能完成的工作。

所有依赖于推理的研究将借助概念的代换和某种计算，这种计算有助于极好结果的发现……

此外，我们应该相信我们所发现和推断的世界的实在性。因为它不难通过算它一遍或通过验算而加以证实，这就类似于算术中的九归验算一样。如果有人怀疑我的结果，我就对他说：“先生，让我们计算吧！”这样一来，用笔和墨水，我们就会很快解决问题。^①

为了达到这一目的，莱布尼兹设法建立一种任何语词都能用正整数表示的普遍语言。基本术语将用质数表示，而复合术语则用质数的乘积表示。比如，如果“动物”这一术语用 2 来表示，而“理性”用 3 来表示，“人”这一术语用 2 与 3 的乘积或 6 来表示，那么表示“所有人都是有理性的”这一陈述就被还原为“6 可以被 3 整除”这一算术真理。

这是现代计算机数据处理的萌芽。莱布尼兹试图将字母表示为数字代码，这样字母信息就能够被数学化地处理了！

这对我们的论题意味着什么呢？推理是人类天然的功能。

① 第一段引文由弗兰克·科普利译自德文版的《莱布尼茨哲学论文集》，第 7 卷，第 31 页，该论文集由 C. I. 格哈特编辑。第二段引文，发表年代是 1677 年，最初的出处见路易斯·库图拉特关于莱布尼茨的著作第 155—156 页，见菲科普·维纳所译《莱布尼茨文集》第 15 页。

根据亚里士多德的说法，它使人类区别于较低等的动物。莱布尼兹认为所有的人类推理，包括发现和证实，都能被处理为数字运算。数字运算能够用一台通用数字计算机来执行，这就是一部有穷决定论自动机。因此，假设莱布尼兹是正确的，那么一台有穷决定论自动机就能够执行一种于人类非常基本和重要的自然功能，即推理。

让我来解释一个我所说的人类的自然功能是什么意思。这个“自然”，排除超自然，排除真正的神秘体验，印度的灵魂转世和佛教的涅槃。我只声明，有穷决定论自动机有人的行为能力（就运动神经对自然感觉刺激作出反应而言）。但这不是华生的行为主义，也不是洛克的白板说。因为一台执行人类的所有自然功能的计算机将会有有一个非常复杂和精巧的逻辑结构和程序系统。

感觉刺激与运动神经反应之间的行为上的区别，类似于计算机输入与输出之间的差别。反对我的观点的大多数人会承认，所有人类的输入输出器官（眼、耳、鼻、舌、臂、手、皮肤等）都是电子的、机械的或者化学的，至少基本如此。然而他们仍可能反驳道，有某些类型的功能要依靠身体或肉体的自然属性，而一部机器不能完成这些功能。一部机器怎么能与人类以一种相爱的方式相互联系呢？人不仅仅是一个反应机制。谁会爱上一捆元件呢？

其实这倒是无关紧要的。谁能爱上巴黎圣母院的驼背人呢？这里更重要的是自动机情人的基本哲理。她是一台自动机，而在外观上和行动上却像一个理想的姑娘。她美丽、聪敏、富有魅力而且含情脉脉。但她是机器而不是人。她能有感觉吗？她果真会恋爱吗？这里有几个关于有穷自动机怎能

执行人类的一切自然功能的重要问题，但它们并不是我们这里所讨论的行为问题。我们所关心的只是一台自动机是否能够有好像它所感觉到的那样的行为，并且只是爱的行为，而不是它是否能够感觉到或者是沉浸在爱之中。

我的观点是，我们每一个人都有一个行为上等效的有穷决定论自动机。但是我对人的内部计算系统的兴趣更甚于他的输入-输出装置。所以让我们先来解决我的论题中的这个输入输出部分。

为了论证方便，让我们假设终有一天所有的人类输入输出功能能够被没有生命的人工装置所模拟。进一步假设一部有穷自动机能够匹配这些人工的像人的眼睛、耳朵、皮肤、手等等一样的部件，以便控制它们并从它们那里接受信息。最后我们假设一部有穷自动机能够执行人类所有的自然功能，而它的内部计算系统能被包装得小到足以进入一个像人一样大小的人造身体之中。

我想这些假设在技术上都是合情合理的。让我们考虑一下最后一个关于尺寸的假设。我曾参与设计了第一台通用电子数学计算机 ENIAC。这台机器长为 80 英尺，高为 8 英尺，深为 3 英尺，并且有 18 000 个真空管，占用这个房间的大部分空间。现在一个逻辑上等价的电子装置已经可能放在你的手中。五年之后，它能被放在你的指甲上。

但是我的中心论题并不是这些技术性假定的真实性，而只是它们在逻辑上的可能性。通过这些假设，我们撇开我的论题中的输入输出部分而只考虑其核心部分即关于人的内部计算系统。就自动机情人而言，我的中心论题是，存在一个有穷决定论自动机，当接通自动机情人的输入输出端时，就

能满足这样的所有计算需要，即使它的行为在一切外在方面像一个真的情人。这个核心是我的人-机论题在哲学上的重要部分。它包括但超越了莱布尼兹关于人类推理能被还原为计算的论点。

什么是一部有穷决定论自动机？我们可以通过观察一下电子计算机的内部计算系统并概括出它的逻辑结构来作出最佳回答。实际上，看看第一个存储程序的计算机的简单逻辑结构就足够了。这个逻辑结构主要归功于冯·诺依曼，尽管其他几个人（包括我在内）也都出了力。

图1展示了一台电子计算机的内部计算系统。就像整个高卢国一样，^①它被分为三个部分：一个记忆装置，一个运算器和一个控制器。它们都互相联系；而这个控制器联系着外部世界。

记忆装置贮存着数据。根据莱布尼兹的见解，字母信息同样也包括在内。一些字母代码是指令或要求，它告诉这台机器怎样做；转移记忆装置与运算器之间的信息，把贮存在运算器内的两个数字加起来，从输入端接收信息等。正像你们所知道的那样，指令的序列被称为程序。

① 高卢是欧洲古国，现在该地区分属法国、意大利和比利时。——译者注