

“十三五”国家重点图书出版规划



中国人工智能学会推荐

人工智能丛书

# 人工智能原理

Principles of Artificial Intelligence

王文敏

高等教育出版社

“十三五”国家重点图书出版规划

中国人工智能学会推荐

人工智能丛书

# 人工智能原理

Principles of Artificial Intelligence

王文敏

高等教育出版社·北京

## 内容简介

本书对人工智能学科及其研究领域进行抽象,梳理出一个人工智能的研究体系,然后按照这个体系讲述其原理,再穿插一些实例。全书分为五篇,共十四章。

这五篇分别是人工智能的体系、求解、规划、学习以及推理。

体系篇中有两章,即结论与体系论。求解篇分成四章,讲述搜索问题、优化问题、博弈问题,以及约束问题求解。规划篇中有两章,分别是时空关联规划和决策理论规划。学习篇先用一章论述机器学习的三个视角,然后再用三章分别讲述学习的任务、范式以及框架。推理篇用两章分别介绍知识表示和推理机制。

本书可作为人工智能、计算机及其相关专业的高年级本科生或研究生的教材,也可作为人工智能的参考书,供有关教学和科研人员使用。

## 图书在版编目(CIP)数据

人工智能原理/王文敏主编. -- 北京:高等教育出版社,2019.8

ISBN 978-7-04-052188-7

I. ①人… II. ①王… III. ①人工智能-高等学校-教材 IV. ①TP18

中国版本图书馆 CIP 数据核字(2019)第 133935 号

策划编辑 张江漫  
插图绘制 于博

责任编辑 张江漫  
责任校对 刘娟娟

封面设计 赵阳  
责任印制 赵义民

版式设计 于婕

出版发行 高等教育出版社  
社址 北京市西城区德外大街4号  
邮政编码 100120  
印刷 北京中科印刷有限公司  
开本 787mm×960mm 1/16  
印张 30.5  
字数 550千字  
购书热线 010-58581118  
咨询电话 400-810-0598

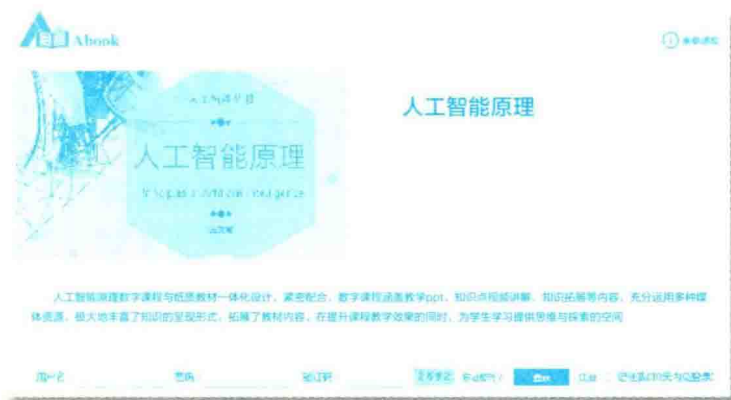
网 址 <http://www.hep.edu.cn>  
<http://www.hep.com.cn>  
网上订购 <http://www.hepmall.com.cn>  
<http://www.hepmall.com>  
<http://www.hepmall.cn>  
版次 2019年8月第1版  
印次 2019年8月第1次印刷  
定 价 63.00元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换  
版权所有 侵权必究  
物料号 52188-00

# 人工智能原理

王文敏

- 1 计算机访问<http://abook.hep.com.cn/1256271>, 或手机扫描二维码、下载并安装 Abook 应用。
- 2 注册并登录, 进入“我的课程”。
- 3 输入封底数字课程账号(20位密码, 刮开涂层可见), 或通过 Abook 应用扫描封底数字课程账号二维码, 完成课程绑定。
- 4 单击“进入课程”按钮, 开始本数字课程的学习。



课程绑定后一年为数字课程使用有效期。受硬件限制, 部分内容无法在手机端显示, 请按提示通过计算机访问学习。

如有使用问题, 请发邮件至[abook@hep.com.cn](mailto:abook@hep.com.cn)。



扫描二维码  
下载 Abook 应用



人工智能简介



人工智能原理  
教学计划

<http://abook.hep.com.cn/1256271>

# 人工智能丛书编委会

<b>主任:</b>	谭铁牛	院士	中国科学院自动化研究所
<b>委员:</b>	李德毅	院士	总参第 61 研究所
	张 钹	院士	清华大学
	徐扬生	院士	香港中文大学(深圳)
	郑南宁	院士	西安交通大学
	陆汝钤	院士	中国科学院数学与系统科学研究院
	柴天佑	院士	东北大学
	李衍达	院士	清华大学
	钟义信	教授	北京邮电大学
	史忠植	研究员	中国科学院计算技术研究所
	何华灿	教授	西北工业大学
	孙富春	教授	清华大学
	刘成林	研究员	中国科学院自动化研究所
	王海峰	教授	百度公司
	焦李成	教授	西安电子科技大学
	沈晓卫	院长	IBM 中国研究院
	周志华	教授	南京大学
	胡 郁	院长	科大讯飞研究院
	周 明	研究员	微软亚洲研究院
	孙哲南	研究员	中国科学院自动化研究所

# 前 言

人工智能被认为是一项对人类社会产生颠覆式影响的科学与技术,其发展之迅猛以及影响面之大,远远超过人们的预料。

人工智能的时代已经到来,准备好去拥抱这个时代了吗?

人工智能于1956年诞生,经历过两次低潮之后,计算能力的迅速提升为其提供了强大的引擎,多媒体数据的爆发性增长为其提供了充足的原料,互联网的迅猛发展为其提供了无垠的空间。人工智能先后战胜了人类的国际象棋、围棋以及德州扑克的顶级选手,图像的识别与分类能力已经超过人类,语音、指纹等识别方式正在改变人机交互的手段,无人驾驶汽车已经在真实环境下试运行数百万公里,双足机器人已经实现了华丽的后空翻。人工智能的学术研究越来越深入,人工智能的创业者越来越多。

套用著名作家狄更斯的一句名言:这是人工智能的“最好时代”,这是人工智能“最有争议的时代”。

世界主要发达国家都把人工智能作为目前最大的发展战略,力图在新一轮国际竞争中掌握主导权。中国国务院于2017年7月8日印发了《新一代人工智能发展规划》,明确提出了中国新一代人工智能实行“三步走”的发展战略,即到2020年人工智能总体技术和应用与世界先进水平同步,到2025年人工智能基础理论实现重大突破,在2030年成为世界主要的人工智能创新中心。这里可以用六个字来概括“三步走”的战略:即同步、突破、领先。所以说,这是人工智能的“最好时代”。

可是,有人担心人工智能会造成大批人失业,有人认为人工智能是人类的威胁,有人游说人工智能可能引发第三次世界大战,更有人惧怕人工智能最终会毁灭人类。所以说,这是人工智能的“最有争议的时代”。

最好也罢,最有争议也罢,人工智能究竟是什么?可以说,人工智能是学术研究,人工智能是工程技术,人工智能是人类科学发展的智慧结晶。

最好也罢,最有争议也罢,人工智能究竟研究什么?可以认为,人工智能研究如何用硬件或软件实现理性的智能,包括智能感知、认知与行为。人类研究人工智能,人类掌握人工智能,人工智能必将造福于人类。

当你在学习人工智能的过程中,你对人工智能的理解会更全面;当你在探索

人工智能的过程中,你对人工智能的认识会更清晰;当你在实践人工智能的过程中,你对人工智能的感受会更深刻。

本书做了一个尝试,即对人工智能的学科及其研究领域进行抽象,梳理出一个人工智能的研究体系,然后按照这个体系讲述其原理,再穿插一些实例分析。

全书分为五篇,共十四章。这五篇分别是:人工智能的“体系篇”“求解篇”“规划篇”“学习篇”以及“推理篇”。

首先,体系篇中有两章。第1章是绪论,介绍人工智能的基础、发展史、突破性进展、图灵测试和塞尔的思想实验以及人工智能的三个层级。第2章是体系论,在探讨人工智能的本源、内涵与外延的基础上,去粗取精,提纲挈领,给出一个人工智能的研究体系,再将人工智能的诸多研究分支囊括在这个研究体系之中。人工智能首先要研究如何用硬件或软件去实现理性思维,而思维的主要表现形式可以归结为求解、规划、学习以及推理,这是后四篇名称的由来。

其次,求解篇分成四章。本书在人工智能的问题求解中挑选出四类代表性问题,即搜索问题、优化问题、博弈问题以及约束问题,分别在第3章至第6章论述这四类问题的求解方法。

再次,规划篇中有两章。分别是第7章的时空关联规划以及第8章的决策理论规划。前者讲述与时间和空间有关的规划方法,而后者是建立在马尔科夫决策过程和动态规划之上的决策理论规划方法。

接着,是学习篇,这里的学习指的是机器学习。本书又将其进行梳理,分成了三个视角,并在第9章中对这三个视角进行了论述,然后再用三章的篇幅分别讲述这三个视角,即:机器学习的任务、范式以及框架,机器学习的一些典型算法将归纳到这三个视角之中。这是本书的另一个尝试,即从三个角度和层面来诠释机器学习,而不是仅仅讲述若干个机器学习的代表性算法。

最后,是推理篇。这一篇包含第13章知识表示和第14章推理机制。在知识表示中,除了介绍一些代表性的知识表示之外,还讲述了经典逻辑、模态逻辑、描述逻辑,以及本体、知识图谱等。在推理中,不仅论述了逻辑推理、定性推理、语义推理以及概率推理,还涉猎了近几年提出的机器推理等。

本书中每个篇章的组织与结构如图1所示。

综上,本书提出了一个人工智能的研究体系、三个机器学习的研读视角,并围绕一个体系、三个视角来撰写。这是本书的一个尝试,几点愚见,不当之处,敬请批评指正,共同商榷。借用陶渊明诗中的一句话,“疑义相与析”。

本书可作为人工智能、计算机及其相关专业的高年级本科生或研究生的教材,也可作为人工智能的参考书,供有关科研人员使用。



图1 本书篇章的组织与结构

阅读本书后您会发现,书中用到的英文较多。毋庸讳言,这是为了尊重引用到的原文,为了避免翻译上的差异,也为了读者们查阅英文资料时的方便。

不仅是专业术语,对于一些非第一手信息和数据,我都尽量找到原文加以核对,避免道听途说。本书中涉猎到的外国人姓名,我查阅有关资料选择常用的中文翻译,再将其英文放在后面便于对照。对于原文中出现的外国人名的缩写,我也尽量找到其全名。

尽管学习和研究人工智能多年,又参考了大量的书籍、论文以及英文版的维基百科等,但书中错误难免,欢迎专家和读者给予批评和指正。来函请发至 wang.wenmin@qq.com。如果本书能成为读者们学习和研究人工智能过程中的一块敲门砖、抑或是一块铺路石,就颇感欣慰了。

王文敏  
2019年3月

## 致 谢

我于1986年在哈尔滨工业大学计算机系攻读博士学位期间,正好赶上了参加国家863计划智能计算机主题中的项目,使我有机会涉足人工智能这个研究领域。我于1988年12月写就的博士论文,题目就是《支持AI问题求解的成员系统模型研究》。1989年又以“支持协同式问题求解的成员系统语言及其并行模型研究”为题获得了国家青年自然科学基金资助,是当时计算机领域获得资助的仅有的几个青年学者之一。弹指间三十余载,我由衷地感谢祖国和母校的培养、感谢导师陈光照教授和胡明曾教授。

我要感谢北京大学和北京大学深圳研究生院,秉承“兼容并包”的传统,使我能够于2009年底回国后在这个平台上开展科研与学术工作,专心致志、身体力行、知行合一、教学相长。我特别开心的是和研究生们在一起,北大的莘莘学子,求知若渴、思维活跃、风华正茂。

我要感谢澳门科技大学,使我能够延续我的学术生涯。“意诚格物”的校训,勉励我意诚心正、格物致知,激励我继续潜心笃志、脚踏实地。

我要感谢“慕课问道”的布道者——北京大学李晓明教授,使我步入了慕课这个大规模开放式在线课程的平台。从2017年3月起,我的“人工智能原理”慕课课程先后在北京大学“华文慕课”“中国大学MOOC”“北京高校优质课程研究会”以及清华大学“学堂在线”这四个慕课平台上开设。

我要感谢这四个慕课平台,尤其是“中国大学MOOC”,让十几万人工智能的求知者们在网络空间里与这门课程对接。本书就是在我的“人工智能原理”的课堂教学和慕课教学讲义的基础上写成的。当然,讲义和书籍之间有着巨大的差别,我又花了大量的时间重新梳理,许多地方甚至是重写,并充实了大量的内容。

我要感谢 $k$ -modes等算法的发明人——黄哲学教授,抽出时间审阅这本书,并提出了宝贵的意见。

我要感谢高等教育出版社,使得《人工智能原理》实现了从单薄的讲义到厚实的书籍之蜕变,并得以精心的编辑。

我还要感谢其他几家出版社的约稿,并对我的分身乏术而深表歉意。

我要感谢我的妻子和女儿,她们的理解与支持,使我能够“躲进小楼成一统”,从2009年底回国专心从事科研和学术工作。

最后,我还将此书献给含笑九泉的父母,他们是我回报祖国乃至完成此书的力量之源泉。

王文敏

2019年3月

## 郑重声明

高等教育出版社依法对本书享有专有出版权。任何未经许可的复制、销售行为均违反《中华人民共和国著作权法》，其为人将承担相应的民事责任和行政责任；构成犯罪的，将被依法追究刑事责任。为了维护市场秩序，保护读者的合法权益，避免读者误用盗版书造成不良后果，我社将配合行政执法部门和司法机关对违法犯罪的单位和个人进行严厉打击。社会各界人士如发现上述侵权行为，希望及时举报，本社将奖励举报有功人员。

反盗版举报电话 (010)58581999 58582371 58582488

反盗版举报传真 (010)82086060

反盗版举报邮箱 dd@hep.com.cn

通信地址 北京市西城区德外大街4号

高等教育出版社法律事务与版权管理部

邮政编码 100120

### 防伪查询说明

用户购书后刮开封底防伪涂层，利用手机微信等软件扫描二维码，会跳转至防伪查询网页，获得所购图书详细信息。也可将防伪二维码下的20位密码按从左到右、从上到下的顺序发送短信至106695881280，免费查询所购图书真伪。

### 反盗版短信举报

编辑短信“JB,图书名称,出版社,购买地点”发送至10669588128

### 防伪客服电话

(010)58582300

## 网络增值服务使用说明

### 一、注册/登录

访问 <http://abook.hep.com.cn/1256271>，点击“注册”，在注册页面输入用户名、密码及常用的邮箱进行注册。已注册的用户直接输入用户名和密码登录即可进入“我的课程”页面。

### 二、课程绑定

点击“我的课程”页面右上方“绑定课程”，正确输入教材封底防伪标签上的20位密码，点击“确定”完成课程绑定。

### 三、访问课程

在“正在学习”列表中选择已绑定的课程，点击“进入课程”即可浏览或下载与本书配套的课程资源。刚绑定的课程请在“申请学习”列表中选择相应课程并点击“进入课程”。

如有账号问题，请发邮件至：[abook@hep.com.cn](mailto:abook@hep.com.cn)。

# 目 录

## 第一篇 体 系

第 1 章 绪论	3
1.1 引言	3
1.2 关于 agent 的翻译	3
1.3 人工智能的定义	4
1.4 人工智能的基础	5
1.4.1 数学	6
1.4.2 经济学	8
1.4.3 神经科学	9
1.4.4 认知心理学	9
1.4.5 控制理论与控制论	10
1.5 人工智能的历史	11
1.5.1 人工智能的孕育和诞生	12
1.5.2 人工智能的黄金年代	13
1.5.3 人工智能的第一个冬天	14
1.5.4 人工智能的第一次繁荣	15
1.5.5 人工智能的第二个冬天	16
1.5.6 人工智能的第二次繁荣	16
1.6 人工智能的突破	17
1.6.1 计算机游戏	17
1.6.2 计算机视觉	20
1.6.3 语音识别	21
1.6.4 机器翻译	21
1.6.5 自动驾驶	22
1.6.6 人工大脑	22
1.6.7 人工智能芯片	23
1.6.8 医学人工智能	23

1.6.9 药物人工智能 .....	24
1.6.10 法律人工智能 .....	24
1.7 人工智能的判定 .....	25
1.7.1 图灵测试 .....	25
1.7.2 图灵测试的解读 .....	26
1.7.3 图灵测试 2.0 .....	27
1.7.4 中文房间 .....	34
1.8 人工智能的层级 .....	35
1.8.1 人工专用智能 .....	35
1.8.2 人工通用智能 .....	36
1.8.3 人工超级智能 .....	37
1.9 人工智能的安全 .....	38
1.10 人工智能的应用 .....	40
1.11 小结 .....	40
习题 .....	40
<b>第 2 章 体系论</b> .....	<b>42</b>
2.1 引言 .....	42
2.2 人工智能的研究途径 .....	42
2.2.1 控制论与人工智能 .....	42
2.2.2 符号与亚符号 .....	43
2.2.3 整齐与不整齐 .....	44
2.2.4 逻辑与反逻辑 .....	45
2.2.5 符号主义、联结主义与行为主义 .....	46
2.2.6 概率与统计方法 .....	47
2.2.7 计算心理学、计算哲学与机器智能 .....	47
2.2.8 经典与现代人工智能 .....	48
2.3 人工智能的主体论点 .....	48
2.3.1 智能主体 .....	49
2.3.2 计算主体 .....	51
2.3.3 理性主体 .....	52
2.4 智能主体的环境 .....	53
2.5 智能主体的构成 .....	54
2.5.1 主体函数 .....	54
2.5.2 主体算法 .....	54

2.5.3 主体结构 .....	55
2.6 人工智能的研究体系 .....	56
2.6.1 人工智能的本源 .....	56
2.6.2 人工智能的内涵 .....	57
2.6.3 人工智能的外延 .....	57
2.6.4 人工智能的体系 .....	58
2.7 智能主体的思维 .....	59
2.7.1 求解 .....	59
2.7.2 规划 .....	60
2.7.3 学习 .....	60
2.7.4 推理 .....	61
2.8 智能主体的交互 .....	62
2.8.1 感知 .....	62
2.8.2 动作 .....	63
2.9 智能主体的类型 .....	63
2.9.1 主体的层次 .....	63
2.9.2 求解型主体 .....	64
2.9.3 规划型主体 .....	65
2.9.4 学习型主体 .....	66
2.9.5 推理型主体 .....	66
2.10 小结 .....	67
习题 .....	67

## 第二篇 求 解

第3章 搜索问题求解 .....	71
3.1 引言 .....	71
3.2 搜索问题 .....	71
3.2.1 智力游戏问题 .....	71
3.2.2 现实世界问题 .....	74
3.3 搜索问题的要素 .....	74
3.3.1 状态表征 .....	74
3.3.2 状态空间 .....	75
3.3.3 形式化 .....	75
3.3.4 求解的方法 .....	76

3.4 搜索问题的实例化 .....	77
3.4.1 八数码难题 .....	77
3.4.2 八皇后难题 .....	77
3.4.3 传教士和食人族问题 .....	78
3.4.4 最短路径问题 .....	78
3.5 搜索求解的方式 .....	79
3.5.1 树搜索 .....	79
3.5.2 图搜索 .....	80
3.6 无信息搜索 .....	81
3.6.1 宽度优先搜索 .....	82
3.6.2 深度优先搜索 .....	84
3.6.3 迭代深化搜索 .....	85
3.7 有信息搜索 .....	86
3.7.1 统一代价搜索 .....	87
3.7.2 贪婪最佳优先搜索 .....	88
3.7.3 A* 搜索 .....	90
3.8 小结 .....	91
习题 .....	91
<b>第 4 章 优化问题求解</b> .....	<b>92</b>
4.1 引言 .....	92
4.2 优化问题 .....	92
4.2.1 智力游戏问题 .....	93
4.2.2 现实世界问题 .....	93
4.3 优化问题的求解 .....	94
4.3.1 局部搜索 .....	94
4.3.2 元启发式 .....	95
4.3.3 群体智能 .....	95
4.4 局部搜索方法 .....	95
4.4.1 爬山法 .....	96
4.4.2 局部束搜索 .....	98
4.5 元启发式方法 .....	99
4.5.1 禁忌搜索 .....	100
4.5.2 模拟退火算法 .....	103
4.5.3 遗传算法 .....	105

4.6 群体智能方法 .....	108
4.6.1 蚁群优化算法 .....	108
4.6.2 粒子群优化算法 .....	110
4.7 小结 .....	112
习题 .....	112
<b>第5章 博弈问题求解</b> .....	<b>114</b>
5.1 引言 .....	114
5.2 博弈问题 .....	114
5.2.1 智力游戏问题 .....	114
5.2.2 现实世界问题 .....	115
5.3 博弈问题的类型 .....	116
5.3.1 合作与非合作 .....	116
5.3.2 零和与非零和 .....	117
5.3.3 完美与不完美信息 .....	117
5.3.4 对称与非对称 .....	117
5.3.5 随机与非随机 .....	118
5.3.6 同步与顺序 .....	118
5.4 博弈算法的历史 .....	118
5.5 博弈问题的求解 .....	121
5.5.1 博弈问题的复杂性 .....	121
5.5.2 对抗搜索 .....	122
5.5.3 博弈策略 .....	122
5.6 完美信息博弈 .....	123
5.6.1 最小最大策略 .....	124
5.6.2 Alpha-Beta 剪枝 .....	130
5.6.3 蒙特卡罗树搜索 .....	132
5.6.4 AlphaGo 的博弈算法 .....	137
5.7 不完美信息博弈 .....	139
5.7.1 纳什均衡与相关均衡 .....	139
5.7.2 不完美信息博弈示例 .....	141
5.7.3 遗憾最小化 .....	145
5.7.4 虚拟遗憾最小化 .....	147
5.7.5 DeepStack .....	150
5.7.6 Libratus .....	151

5.8 非对称博弈 .....	154
5.8.1 非对称博弈概述 .....	154
5.8.2 非对称博弈的对称分解 .....	155
5.9 随机博弈 .....	156
5.9.1 随机博弈概述 .....	156
5.9.2 随机博弈模型 .....	156
5.10 小结 .....	158
习题 .....	158
<b>第6章 约束问题求解</b> .....	<b>160</b>
6.1 引言 .....	160
6.2 约束问题 .....	160
6.2.1 智力游戏问题 .....	160
6.2.2 现实世界问题 .....	163
6.3 可能世界及其约束 .....	163
6.3.1 可能世界 .....	163
6.3.2 变量与值域 .....	164
6.3.3 约束 .....	164
6.4 约束满足问题 .....	165
6.4.1 形式化 .....	165
6.4.2 状态的表示 .....	166
6.4.3 求解的方法 .....	167
6.5 约束满足问题的实例化 .....	167
6.5.1 数独 .....	167
6.5.2 算式谜 .....	168
6.5.3 地图着色 .....	169
6.5.4 作业调度 .....	170
6.6 约束传播 .....	171
6.6.1 节点一致性 .....	171
6.6.2 弧一致性 .....	172
6.6.3 路径一致性 .....	173
6.6.4 $k$ -一致性 .....	173
6.7 回溯搜索 .....	173
6.7.1 回溯搜索的实例 .....	174
6.7.2 回溯搜索算法 .....	175