



国际电气工程先进技术译丛

WILEY

决策用强化与 系统性机器学习

**Reinforcement and Systemic Machine
Learning for Decision Making**

[印度] 巴拉格·库尔卡尼 (Parag Kulkarni) 著

李宁 吴健 刘凯 等译

 **机械工业出版社**
CHINA MACHINE PRESS



国际电气工程先进技术译丛

决策用强化与系统性 机器学习

[印度] 巴拉格·库尔卡尼 (Parag Kulkarni) 著
李 宁 吴 健 刘 凯 等译



机械工业出版社

Copyright © 2012 by the Institute of Electrical and Electronics Engineers, Inc.

All Right Reserved. This translation published under license. Authorized translation from English language edition, entitled < Reinforcement and Systemic Machine Learning for Decision Making >, ISBN: 978 - 0 - 470 - 91999 - 6, by Parag Kulkarni, Published by John Wiley & Sons. No part of this book may be reproduced in any form without the written permission of the original copy-rights holder.

本书中文简体字版由机械工业出版社出版, 未经出版者书面允许, 本书的任何部分不得以任何方式复制或抄袭。版权所有, 翻印必究。

北京市版权局著作权合同登记 图字: 01-2013-7863 号。

图书在版编目 (CIP) 数据

决策用强化与系统性机器学习/ (印) 库尔卡尼著; 李宁等译. —北京: 机械工业出版社, 2015. 7

(国际电气工程先进技术译丛)

书名原文: Reinforcement and Systemic Machine Learning for Decision Making
ISBN 978-7-111-50241-8

I. ①决… II. ①库…②李… III. ①机器学习—研究 IV. ①TP181

中国版本图书馆 CIP 数据核字 (2015) 第 101156 号

机械工业出版社 (北京市百万庄大街 22 号 邮政编码 100037)

策划编辑: 顾 谦 责任编辑: 顾 谦

责任校对: 陈 越 封面设计: 马精明

责任印制: 刘 岚

北京富生印刷厂印刷

2015 年 7 月第 1 版第 1 次印刷

169mm×239mm • 15.75 印张 • 303 千字

0001—2600 册

标准书号: ISBN 978-7-111-50241-8

定价: 79.00 元

凡购本书, 如有缺页、倒页、脱页, 由本社发行部调换

电话服务

网络服务

服务咨询热线: 010-88361066

机工官网: www.cmpbook.com

读者购书热线: 010-68326294

机工官博: weibo.com/cmp1952

010-88379203

金书网: www.golden-book.com

封面无防伪标均为盗版

教育服务网: www.cmpedu.com

机器学习是人工智能领域中一个极其重要的研究方向。强化学习是机器学习中的一个重要分支。作为解决序贯优化决策的有效方法，强化学习有效地应用于计算科学、自动控制、机器人技术等各个领域。

当前，强化学习的核心任务是提高学习效率，本书就是针对此问题展开的。第1章介绍系统概念和增强机器学习，它建立了一个突出的相同的机器学习系统范例；第2章将更多关注机器学习的基本原理和多视角学习；第3章关于强化学习；第4章处理机器学习系统和模型建立的问题；决策推理等重要的部分将在第5章展开；第6章讨论了自适应机器学习；第7章讨论了多视角和全局系统性机器学习；第8章讨论了增量学习的需要和知识表示；第9章处理了知识增长方面的问题；第10章讨论了学习系统的建立。

本书适合于机器学习、自动化技术、人工智能等方面的相关专业教师与研究生阅读，也可供自然科学和工程领域相关研究人员参考。

译者序

进入 21 世纪以来，人们对机器学习（machine learning）的期望与日俱增。十年前所想的智能系统如今只是被认为是个普通系统而已。人们都希望机器变得更加智能，能自主地学习并具备高效的解决日常问题的复杂行为的能力。机器学习的应用不局限于一种特定的领域，而是分布在所有领域。

强化学习（reinforcement learning，又称再励学习、评价学习）是一种重要的机器学习方法，在智能控制机器人及分析预测等领域有许多应用。但在传统的机器学习分类中没有提到过强化学习，而在连接主义学习中，把学习算法分为三种类型，即非监督学习（unsupervised learning）、监督学习（supervised learning）和强化学习。

强化学习由统计学、控制理论、心理学等相关学科发展而来，经过多年的发展，已经成为解决序贯优化决策的一种有效方法。所谓强化学习就是智能系统从环境到行为映射的学习，以使奖励信号（强化信号）函数值最大。强化学习中由环境提供的奖励信号是对产生动作的好坏做一种评价，并非直接告诉强化学习系统（reinforcement learning system）如何去执行正确的动作。由于外部环境提供的信息很少，强化学习系统必须靠自身的经验知识进行学习。通过这种方式，强化学习系统在动作—评价的环境中获得知识，改进行动方案以适应环境。

这是一本全面讨论强化学习与系统性学习的书籍，适合于机器学习、自动化技术、人工智能等方面的相关专业教师与研究生阅读，也可供自然科学和工程领域相关研究人员参考。本书包括强化的不同方面，通过机器学习来建立知识库。本书有助于计划通过智能学习和实验做出智能机器的人并尝试新的方式，打开一种相同的新范例。本书第 1 章主要介绍系统概念，如机器学习、强化学习、系统学习、系统性机器学习等；第 2 章主要介绍系统性和多视角的机器学习；第 3~9 章主要介绍本书的主要内容——决策用强化与系统性学习的各个方面内容，有强化学习、系统性机器学习、推理和信息集成、自适应学习、全局系统性学习、增量学习及表示和知识增长。第 10 章列举了一些例子来说明如何构建一个学习系统。

本书主要由李宁、吴健和刘凯翻译。参与翻译本书的人员还有李娜、刘建强、伍宏芳、费盛、聂凤梅、陈蓉、王天伟、葛桃桃、杨雪峰等人，感谢他们的辛苦工作。感谢机械工业出版社给予我们这个难得的机会。

由于个人水平有限、经验不足，书中翻译不足之处在所难免，敬请读者指正批评。

译者

原 书 前 言

人们研究人工智能已经很多年，甚至早于计算机时代。在现代，基于事件的人工智能被广泛应用于部件设备或者是设备整体中。人工智能起了很大程度上的引导作用，但人工干预是强制性的。甚至反馈控制系统也是人工智能系统的一种初步形式。之后自适应控制系统和混合控制系统在系统中增加智能的鉴别能力。随着计算机技术的发展，人工智能技术受到了更多的关注。基于计算机学习简单的事件很快成为诸多智能系统的一部分，人们对智能系统的期望在持续增长，这就致使一种广受欢迎的学习范例，其是以学习为基础的模式。这使得系统在很多实际方案下表现得智能化，其中包括天气模式、入住率模式以及其他可以帮助决策的不同模式。这种模式发展成为一个行为模式学习的范例。这与其说是一种行为模式，倒不如说是一种特定测量参数的简单模式。行为模式试图给出一个更好的描绘和洞察力，这有助于学习和在网络及业务方案下进行决策，这将智能系统提升到了另一个水平。学习是智能的表现，使机器进行学习是使得机器智能化行为一个主要的部分。

决策方案的复杂度和复杂方案中的机器学习在机器智能方面提出了很多问题。孤立的学习是永远不会完成的。人类聚居在一起学习，开发聚居地并通过互动去创造智慧。聚集和合作学习让人类取得了统治地位。此外，人类的学习与所处环境相关联。他们与环境互动，并获得两种形式的反馈——奖励或惩罚。人类的协作学习方式给了他们探索式学习的力量，利用已经了解到的事实以及参照发生的行动去探索。强化学习的范例上升到了一个新的层面，并可以覆盖所需动态方案学习的很多新的方面的问题。

正如 Rutherford D. Roger 所说：“我们淹没在信息的海洋中并渴求着知识的养分。”越来越多的信息可供我们支配，这些信息的存在形式多样化，且有很多的信息来源和众多的学习机会。学习时的实践假设能够制约学习。实际上系统的不同部分之间都是有联系的，系统思维状态的基本原则之一就是在时间和空间上因果是分开的。可以感受到决定和行动的影响超越了可察觉的极限。当学习时，如果不考虑系统性方面的关联，会导致很多的局限性。因此传统的学习范例会遭受现实生活中高度动态和复杂性的问题。对相互依赖关系的整体把握和理解能够帮助人们学到很多新的方面的知识，并用更现实的方式理解、分析和解释信息。根据现有的资料学习、构建新的信息并映射其到知识面和理解不同的观点能够提高人们的学习效率。学习不仅仅是获得更多的数据和整理这些数据，甚至不是建立

更多的信息。从根本上说,学习的目的是为了使个人做出更好的决策,并提高其创造价值的能力。在机器学习中,有必要参照不同的信息来源和学习的机会提升机器的能力。在机器学习中,也有必要赋予机器做出更好决策并提高其创造价值的能力。

本书试图参照不同机器学习的各方面提出系统性机器学习和研究机器学习机会的新范例。本书试图依据精心设计的案例研究构建系统性的机器学习基础。机器学习和人工智能在本质上是跨学科的,其涉及统计学、数学、心理学、计算机工程,许多研究者致力于丰富这一领域并获得更好的效果。本书基于这些机器学习领域众多的贡献以及作者的研究,试图探索系统性机器学习的概念。系统机器学习是全面的、多视角的、增量的和系统性的。在学习时可以从同一数据集中学到不同的东西,也可以从已知的事实中学习。本书是建立一个框架使所有的信息来源得到充分利用并参考全局系统体系建立知识的一种尝试。

在许多情况下,这个问题也不是一成不变的,它随着时间的推移而变化且依赖于环境。环境可能不只是局限于几个参数,但一个问题的整体信息建立环境。一个没有环境的通用系统可能不能够处理特定环境的决定。本书不仅讨论学习的不同方面,也讨论参照复杂决策问题案例的需求。本书可作为进行专门研究的参考用书,并可以帮助读者和研究者欣赏机器学习的新模式。

本书的内容结构如图 0.1 所示。

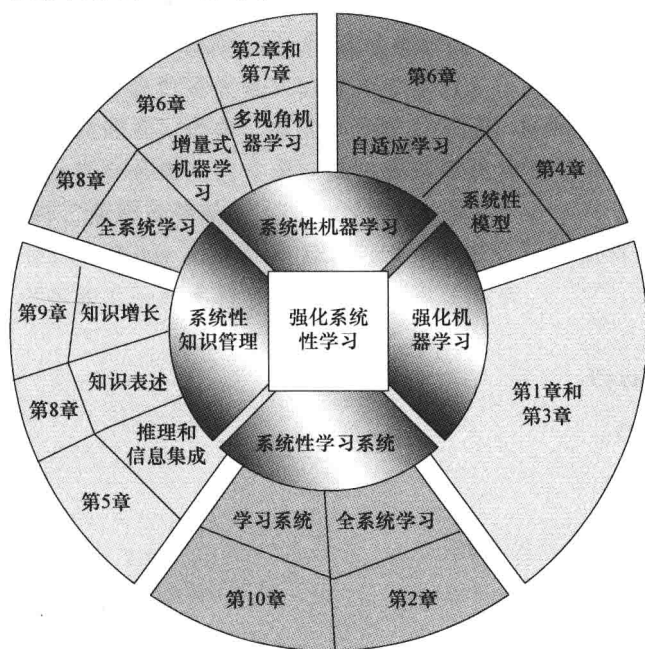


图 0.1 本书的内容结构

VI 决策用强化与系统性机器学习

第1章介绍系统概念和增强机器学习，它建立了一个突出的相同的机器学习系统范例：第2章将更多关注机器学习的基本原理和多视角学习；第3章关于强化学习；第4章处理机器学习系统和模型建立的问题；决策推理等重要的部分将在第5章展开；第6章讨论了自适应机器学习，第7章讨论了范例的多视角机器学习和系统性机器学习；第8章讨论了增量学习的需要，第8章和第9章处理了知识库表示和知识库扩展的问题；第10章讨论了学习系统的建立。

本书试图包括学习的不同方面，同时引入一种新的机器学习范例，通过机器学习来建立知识库。本书有助于计划通过智能学习和实验做出智能机器的人，并尝试新的方式，打开一种相同的新范例。

Parag Kulkarni

原书致谢

在过去的 20 年中，我做了很多关于决策和基于人工智能的 IT 产品公司的工作。在这段时期，我用不同的机器学习算法将其应用于不同的方面。这项工作让我意识到需要一个机器学习的新范式和思维变化的需要，这建立了本书的基础，并开始构建系统性机器学习的思维过程。我非常感谢曾经工作过的地方，包括西门子和艾蒂尔（IDeaS）公司，也同样感谢这两个公司的同事。我还想感谢我的朋友和同事的支持。

我要感谢我的博士和我的研究生——Prachi、Yashodhara、Vinod、Sunita、Pramod、Nitin、Deepak、Preeti、Anagha、Shankar、Shweta、Basawraj、Shashikanth 和其他人，感谢他们直接和间接的贡献。他们总是准备用新思路来工作，通过集体学习起到了推波助澜的作用。特别感谢 Prachi 在制图和文本格式方面的帮助。

我要感谢 Chande 教授、Ramani 教授、Sinha 博士、Bhanu Prasad 博士、Warnekar 教授和 Navetia 教授为本书做的注释和评论。我也要感谢那些帮助我的机构，有 COEP、PICT、GHRIET、PCCOE、DYP COE、IIM、马萨里克（Masaryk）大学等，感谢他们容许我在学生面前进行交互并展示我的思想。我也要感谢 IASTED、IACSIT 和 IEEE，让我通过技术交流会的平台来让我展示我的研究成果。我也要感谢我的研究论文的评审专家。

我感谢指导过我的人，老师、祖父和已故的 D. B. Joshi，他们激励我不同的思维。我也想借此机会感谢我的母亲。最主要感谢我的妻子 Murdula 和我的儿子 Hrishikesh 的支持、激励和帮助。

我也感谢 IEEE /Wiley 出版社和 IEEE /Wiley 出版社的编辑团队，感谢他们对我的研究、思想和实验的支持和帮助，并出版了本书。

Parag Kulkarni

关于作者

Parag Kulkarni 博士是普纳埃拉特研究所 (EKLaT Research, Pune) 的 CEO 和首席科学家。他在知识管理、电子商务、智能系统和机器学习咨询、研究和产品建设等领域有超过 20 年的经验。印度理工学院和加尔各答印度管理研究院的校友，Kulkarni 博士是 IIM 的兼职教授、捷克马萨里克大学访问研究员和普纳工程学院兼职教授。他领导的公司、研究实验室和团体，其中包括很多 IT 公司，有艾蒂尔公司、西门子信息系统有限公司、普纳的卡皮森公司和新加坡的 ReasonEdge 公司。他通过战略创新和研究引领了很多公司成功创业。瑞士的 UGSM 皇家商业学校授予 Kulkarni 荣誉博士学位。他是三个专利的共同发明人，并合著了超过 100 篇研究论文并有著作若干本。

目 录

译者序

原书前言

原书致谢

关于作者

第1章 强化与系统性机器学习 1

1.1 简介	1
1.2 监督学习、无监督学习、半监督学习和强化学习	1
1.3 传统机器学习方法和机器学习发展历史	3
1.4 什么是机器学习?	6
1.5 机器学习问题	6
1.5.1 学习的目标	6
1.6 学习模式	7
1.7 机器学习技术和范例	9
1.8 什么是强化学习?	11
1.9 强化函数和环境函数	12
1.10 强化学习的需求	13
1.11 强化学习和机器智能	14
1.12 什么是系统学习?	14
1.13 什么是系统性机器学习?	15
1.14 系统性机器学习的重点	15
1.15 强化性机器学习和系统性机器学习	16
1.16 车辆检测问题的案例研究	16
1.17 小结	16
参考文献	17

第2章 全系统原理、系统性和多视角的机器学习 18

2.1 简介	18
2.1.1 什么是系统性学习?	19
2.1.2 历史	20
2.2 什么是系统性机器学习?	21
2.2.1 基于事件的学习	21

X 决策用强化与系统性机器学习

2.3 广义系统性机器学习框架	23
2.3.1 系统定义	24
2.4 多视角决策和多视角学习	26
2.4.1 基于完整信息的表示	32
2.4.2 基于部分信息的表示	32
2.4.3 单视角决策方案图	32
2.4.4 双重视角决策方案图	32
2.4.5 多视角决策方案图	32
2.4.6 定性信念网络和影响图	33
2.5 动态和交互式决策	33
2.5.1 交互决策图	33
2.5.2 决策图和影响图中时间的角色	34
2.5.3 系统性视角的建立	34
2.5.4 信息整合	35
2.5.5 建立典型决策方案图	35
2.5.6 受限信息	35
2.5.7 多决策者系统在系统性学习中的角色	35
2.6 系统性学习框架	39
2.6.1 数学模型	39
2.6.2 系统性学习的方法	39
2.6.3 自适应系统性学习	40
2.6.4 系统性学习框架	41
2.7 系统分析	41
2.8 案例学习：在酒店行业中需要系统性学习	43
2.9 小结	44
参考文献	44

第3章 强化学习

3.1 简介	45
3.2 学习决策者	48
3.3 回报和奖励的计算	50
3.3.1 方案和连续任务	50
3.4 强化学习和自适应控制	51
3.5 动态系统	54
3.5.1 离散事件动态系统	54
3.6 强化学习和控制	55
3.7 马尔科夫性质和决策过程	55
3.8 价值函数	56

3.8.1 行动和价值	56
3.9 学习最优策略(有模型和无模型法)	57
3.10 动态规划	57
3.10.1 动态系统性质	57
3.11 自适应动态规则	58
3.11.1 时间差分学习	59
3.11.2 Q 学习	60
3.11.3 统一的视图	60
3.12 范例——拳击训练器的强化学习	61
3.13 小结	61
参考文献	61
第4章 系统性机器学习和模型	62
4.1 简介	62
4.2 系统学习的框架	63
4.2.1 影响空间	64
4.2.2 交互作用为中心的模型	69
4.2.3 以结果为中心的模型	69
4.3 捕捉系统视图	70
4.4 系统交互的数学表达	73
4.5 影响函数	74
4.6 决策影响分析	74
4.6.1 时空界限	75
4.7 小结	80
第5章 推理和信息集成	82
5.1 简介	82
5.2 推理机制和需要	83
5.2.1 情景推理	85
5.2.2 推理确定影响	85
5.3 情景和推理的集成	88
5.4 统计推理和归纳	91
5.4.1 直接推理	91
5.4.2 间接推理	91
5.4.3 信息推理	91
5.4.4 归纳	92
5.5 纯似然方法	92
5.6 贝叶斯范例推理	93

XII 决策用强化与系统性机器学习

5.6.1 贝叶斯定理	93
5.7 基于时域推理	93
5.8 推理建立系统观点	94
5.8.1 信息集成	94
5.9 小结	96
参考文献	97
第6章 自适应学习	98
6.1 简介	98
6.2 自适应学习和自适应系统	98
6.3 什么是自适应机器学习	101
6.4 基于方案的适应性和学习方法	101
6.4.1 动态适应性和情景感知的学习	102
6.5 系统学习和自适应学习	104
6.5.1 多学习器的使用	105
6.5.2 系统自适应机器学习	108
6.5.3 自适应应用的设计	110
6.5.4 自适应学习的需要和适应的原因	111
6.5.5 适应类型	112
6.5.6 自适应框架	114
6.6 竞争学习和自适应学习	115
6.6.1 适应性函数	116
6.6.2 决策网络	118
6.6.3 自适应学习方案	119
6.7 范例	120
6.7.1 案例研究：基于自适应学习的文本	120
6.7.2 自适应学习的文档挖掘	121
6.8 小结	122
参考文献	122
第7章 多视角和全局系统性的学习	123
7.1 简介	123
7.2 多视角方案构建	124
7.3 多视角决策和多视角学习	126
7.3.1 视角结合	126
7.3.2 影响图和部分方案决策表示图	127
7.3.3 表示决策方案图（RDSD）	130
7.3.4 范例：部分方案决策表示图（PDSRD）表示的不同视角获取的城市信息	131

7.4 全局系统性学习和多视角途径	134
7.4.1 分散信息整合	135
7.4.2 多视角和全局系统知识表示	135
7.4.3 什么是多视角方案?	135
7.4.4 特定方案	136
7.5 基于多视角途径的案例研究	136
7.5.1 交通控制器用多视角途径	137
7.5.2 情感检测用多视角途径模型	138
7.6 多视角方法的局限性	143
7.7 小结	143
参考文献	144
第8章 增量学习和知识表示	145
8.1 简介	145
8.2 为什么增量学习?	146
8.3 学习已经学会的	147
8.3.1 绝对增量学习	148
8.3.2 选择增量学习	149
8.4 监督增量学习	157
8.5 增量无监督学习和增量聚类	158
8.5.1 增量聚类: 任务	160
8.5.2 增量聚类: 方法	161
8.5.3 阈值	161
8.6 半监督增量学习	162
8.7 增量与系统性学习	163
8.8 增量接近值和学习方法	164
8.8.1 增量学习方法 1	165
8.8.2 增量学习方法 2	166
8.8.3 计算 C 值增量	166
8.9 学习与决策模型	169
8.10 增量分类技术	169
8.11 案例分析: 增量文档分类	170
8.12 小结	171
第9章 知识增长: 机器学习的视角	173
9.1 简介	173
9.2 短暂的历史和相关工作	174
9.3 知识增长和知识启发	178

9.3.1 策略使用进行知识启发	178
9.3.2 基于目标的知识启发	179
9.3.3 基于过程的知识启发	179
9.4 生命周期	180
9.4.1 知识水平	181
9.4.2 直接知识	181
9.4.3 间接知识	182
9.4.4 程序知识	182
9.4.5 问题	182
9.4.6 决策	182
9.4.7 知识生命周期	183
9.5 增量知识表达	184
9.6 案例学习和遗忘学习	186
9.7 知识的扩充: 技术和方法	187
9.7.1 知识增量技术	187
9.7.2 知识增量方法	188
9.7.3 提取知识的机制	189
9.8 启发式学习	190
9.9 系统性机器学习和知识获取	190
9.9.1 全方位知识获取	191
9.9.2 系统知识管理和先进的机器学习	192
9.10 在复杂环境下的知识增量	193
9.11 案例研究	193
9.11.1 银行案例研究	193
9.11.2 软件开发公司	194
9.11.3 杂货集市/零售集市	195
9.12 小结	195
参考文献	196
第 10 章 构建学习系统	197
10.1 简介	197
10.2 系统性学习系统	197
10.2.1 学习单元	199
10.2.2 知识库	200
10.2.3 性能单元	200
10.2.4 反馈单元	200
10.2.5 允许测量的系统	200
10.3 算法选择	201

10.3.1 k 近邻 (k NN)	201
10.3.2 支持向量机 (SVM)	202
10.3.3 质心法	202
10.4 知识表示	203
10.4.1 实用方案和案例研究	203
10.5 学习系统的设计	204
10.6 让系统表现得更智能	204
10.7 案例学习	205
10.8 整体知识框架和强化学习的应用	205
10.8.1 智能算法的选择	207
10.9 智能决策——部署和知识采集以及重用	208
10.10 基于案例的学习：人体情感检测系统	209
10.11 复杂决策问题的整体视角	211
10.12 知识表示和资源查找	213
10.13 组件	215
10.13.1 范例	215
10.14 学习系统和智能系统的未来	216
10.15 小结	217
附录	218
附录 A 统计学习方法	218
A.1 概率	218
A.1.1 互斥事件	218
A.1.2 独立事件	218
A.2 贝叶斯分类	219
A.2.1 朴素贝叶斯分类	220
A.2.2 贝叶斯分类器的优点和缺点	221
A.3 回归	221
A.3.1 线性	222
A.3.2 非线性	222
A.3.3 回归的其他方法	222
A.4 粗糙集	223
A.4.1 不可分辨关系	223
A.4.2 集近似	224
A.4.3 边界区域	224
A.4.4 粗糙集和清晰集	224
A.4.5 约简	224
A.4.6 可有可无和不可缺少的属性	224