

<<增强学习与近似动态规划>>

图书基本信息

书名：<<增强学习与近似动态规划>>

13位ISBN编号：9787030275653

10位ISBN编号：7030275659

出版时间：2010-5

出版时间：科学出版社

作者：徐昕

页数：212

版权说明：本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问：<http://www.tushu007.com>

<<增强学习与近似动态规划>>

前言

增强学习 (reinforcement learning, RL) 又称为强化学习或再励学习, 它是近年来机器学习和智能控制领域的前沿和热点, 与监督学习和无监督学习并列三大类机器学习方法之一。

增强学习强调以不确定条件下序贯决策的优化为目标, 是复杂系统自适应优化控制的一类重要方法, 具有与运筹学、控制理论、机器人学等交叉综合的特点。

特别是近十年来, 有关近似动态规划 (approximate dynamic programming, ADP) 的研究成为增强学习、运筹学和优化控制理论等相关领域的关注热点。

例如, 美国国家科学基金会于2006年召开的近似动态规划论坛 (NSF-ADP06), IEEE分别于2007年和2009年召开的近似动态规划与增强学习专题国际研讨会 (IEEE ADPRL 2007、IEEE ADPRL 2009) 等。

另外, IEEE计算智能学会于近年专门成立了近似动态规划与增强学习技术委员会 (IEEE TC on ADPRL)。

在以电梯调度、网络路由控制等为代表的大规模优化决策应用中, 增强学习显示了相对传统监督学习和数学规划方法的优势。

在智能机器人系统、复杂不确定系统的优化控制等领域, 增强学习的应用也正在不断得到推广。

本书是作者多年从事增强学习与近似动态规划理论、算法与应用研究的成果总结, 许多成果是近年来最新取得的研究成果, 是一部系统探讨增强学习与近似动态规划的学术著作。

本书有以下几个特点: (1) 新颖性和前沿性。

本书深入论述了增强学习与近似动态规划的核心与前沿研究课题——大规模连续空间Markov决策过程的价值函数与策略逼近问题, 对近年来取得的研究进展进行了充分讨论。

本书大多数理论、算法与实验结果都是作者近年来在研究工作中取得的成果。

(2) 多学科交叉。

增强学习与近似动态规划的研究涉及机器学习、运筹学、智能控制、机器人学等多个学科领域, 具有较强的学科交叉特点和较宽的学科覆盖面, 对相关领域的学术创新起到了积极的促进作用。

(3) 理论与应用密切结合。

本书在论述增强学习与近似动态规划理论和算法研究进展的同时, 结合智能控制、机器人等领域的应用实例, 在算法研究和理论分析的基础上, 开展了大量的仿真和实验验证, 有利于读者尽快把握理论和应用的结合点。

<<增强学习与近似动态规划>>

内容概要

本书对增强学习与近似动态规划的理论、算法及应用进行了深入研究和论述。

主要内容包括：求解Markov链学习预测问题的时域差值学习算法和理论，求解连续空间Markov决策问题的梯度增强学习算法以及进化一梯度混合增强学习算法，基于核的近似动态规划算法，增强学习在移动机器人导航与控制中的应用等。

本书是作者在多个国家自然科学基金项目资助下取得的研究成果的总结，意在推动增强学习与近似动态规划理论与应用的发展，对于智能科学的前沿研究和智能学习系统的应用具有重要的科学意义。

本书可作为高等院校与科研院所中从事人工智能与智能信息处理、机器人与智能控制、智能决策支持系统等专业领域的研究和教学用书，也可作为自动化、计算机与管理学领域其他相关专业师生及科研人员的参考书。

<<增强学习与近似动态规划>>

书籍目录

《智能科学技术著作丛书》序前言第1章 绪论 1.1 引言 1.2 增强学习与近似动态规划的研究概况
1.2.1 增强学习研究的相关学科背景 1.2.2 增强学习算法的研究进展 1.2.3 增强学习的泛化方法与近似动态规划 1.2.4 增强学习相关理论研究 1.2.5 增强学习应用的研究进展 1.3 移动机器人导航控制方法的研究现状和发展趋势 1.3.1 移动机器人体系结构的研究进展 1.3.2 移动机器人反应式导航方法的研究概况 1.3.3 移动机器人路径跟踪控制的研究概况 1.4 全书的组织结构 参考文献第2章 线性时域差值学习理论与算法 2.1 Markov链与多步学习预测问题 2.1.1 Markov链的基础理论 2.1.2 基于Markov链的多步学习预测问题 2.2 TD()学习算法 2.2.1 表格型TD()学习算法 2.2.2 基于值函数逼近的TD()学习算法 2.3 多步递推最小二乘TD学习算法及其收敛性理论 2.3.1 多步递推最小二乘TD(RLS-TD()学习算法 2.3.2 RLS-TD()学习算法的一致收敛性分析 2.4 多步学习预测的仿真研究 2.4.1 HlopWorld问题学习预测仿真 2.4.2 连续状态随机行走问题的学习预测仿真 2.5 小结 参考文献第3章 基于核的时域差值学习算法 3.1 核方法与基于核的学习机器 3.1.1 核函数的概念与性质 3.1.2 再生核Hilbert空间与核函数方法 3.2 核最小二乘时域差值学习算法 3.2.1 线性TD()学习算法 3.2.2 KLS-TD()学习算法 3.2.3 学习预测实验与比较 3.3 小结 参考文献第4章 求解Markov决策问题的梯度增强学习算法 4.1 Markov决策过程与表格型增强学习算法 4.1.1 Markov决策过程及其最优值函数 4.1.2 表格型增强学习算法及其收敛性理论 4.2 基于改进CMAC的直接梯度增强学习算法 4.2.1 CMAC的结构 4.2.2 基于C2MAC的直接梯度增强学习算法 4.2.3 两种改进的C2MAC编码结构及其应用实例 4.3 基于值函数逼近的残差梯度增强学习算法 4.3.1 多层前馈神经网络函数逼近器与已有的梯度增强学习算法 4.3.2 非平稳策略残差梯度(RGNP)增强学习算法 4.3.3 RGNP学习算法的收敛性和近似最优策略性能的理论分析 4.3.4 Mountain-Car问题的仿真研究 4.3.5 Acrobot学习控制的仿真研究 4.4 求解连续行为空间Markov决策问题的快速AHC学习算法 4.4.1 AHC学习算法与Actor-Critic学习控制结构 4.4.2 Fast-AHC学习算法 4.4.3 连续控制量条件下的倒立摆学习控制仿真研究 4.4.4 连续控制量条件下Acrobot系统的学习控制 4.5 小结 参考文献第5章 求解Markov决策问题的进化-梯度混合增强学习算法 5.1 进化计算的基本原理和方法 5.1.1 进化计算的基本原理和算法框架 5.1.2 进化算法的基本要素 5.1.3 进化算法的控制参数和性能评估 5.2 求解离散行为空间MDP的进化-梯度混合算法 5.2.1 HERG算法的设计要点 5.2.2 HERG算法的流程 5.2.3 HtERG算法的应用实例: Mountain-Car学习控制问题 5.2.4 Acrobot系统的进化增强学习仿真 5.3 求解连续行为空间MDP的进化-梯度混合增强学习算法 5.3.1 进化AHC算法 5.3.2 连续控制量条件下Acrobot系统的进化增强学习仿真 5.4 小结 参考文献第6章 基于核的近似动态规划算法与理论 6.1 增强学习与近似动态规划的若干核心问题 6.2 基于核的近似策略迭代算法与收敛性理论 6.2.1 策略迭代与TD学习算法 6.2.2 核策略迭代算法KLSPI的基本框架 6.2.3 采用核稀疏化技术的KLSTD-Q时域差值算法 6.2.4 KLSPI算法的收敛性分析 6.3 核策略迭代算法的性能测试实验研究 6.3.1 具有20个状态的随机Markov链问题 6.3.2 具有50个状态的随机Markov决策问题 6.3.3 随机倒立摆学习控制问题 6.4 小结 参考文献第7章 基于增强学习的移动机器人反应式导航方法 7.1 基于分层学习的移动机器人混合式体系结构 7.2 基于增强学习的移动机器人反应式导航体系结构与算法 7.2.1 未知环境中移动机器人导航混合式体系结构的具体设计 7.2.2 基于神经网络增强学习的反应式导航算法 7.3 移动机器人增强学习导航的仿真和实验研究 7.3.1 CIT-AVT-VI移动机器人平台的传感器系统与仿真实验环境 7.3.2 增强学习导航的仿真研究 7.3.3 CIT-AVT-VI移动机器人的实时学习导航控制实验 7.4 小结 参考文献第8章 RL与ADP在移动机器人运动控制中的应用 8.1 基于增强学习的自适应PID控制器 8.2 自动驾驶汽车的侧向增强学习控制 8.2.1 自动驾驶汽车的动力学模型 8.2.2 用于自动驾驶汽车侧向控制的增强学习PID控制器设计 8.2.3 自动驾驶汽车直线路径跟踪仿真 8.3 基于在线增强学习的室内移动机器人路径跟踪控制 8.3.1 一类室内移动机器人系统的运动学和动力学模型 8.3.2 增强学习路径跟踪控制器设计 8.3.3 参考路径为直线时的仿真研究 8.3.4 参考路径为圆弧时的仿真研究 8.3.5 CIT-AVT-VI移动机器人实时在线学习路径跟踪实验 8.4 采用近似策略迭代的移动机器人学习控制方法研究 8.4.1 基于近似策略迭代的学习控制方法与仿真研究 8.4.2 基于P3-AT平台的学习控制器设计 8.4.3 直线跟随实验 8.4.4 曲线跟随实验 8.5 小结 参考文献第9章 总结与展望

<<增强学习与近似动态规划>>

章节摘录

插图：1.3.3 移动机器人路径跟踪控制的研究概况移动机器人的路径跟踪控制是指在完成路径规划的前提下，通过设计反馈控制律实现移动机器人对规划路径的闭环跟踪控制。

路径跟踪控制问题在许多移动机器人的应用场合具有重要的作用，如包括无人驾驶汽车在内的自主地面车辆（ALV）和用于运输环境的自动导引车辆（AGV）等。

在上述应用条件下，全局路径往往可以事先规划完成，移动机器人对规划路径的高性能跟踪成为有效完成导航任务的一个关键。

目前针对移动机器人的动力学建模和路径跟踪控制问题已开展了大量的研究工作。

移动机器人作为一类具有非完整特性的机电系统，难以建立其精确的动力学模型，并且系统的动力学特性还受到运行速度和环境条件变化的影响，特别是对于高速运行的无人驾驶汽车等移动机器人系统，其动力学参数随车速变化明显。

因此移动机器人的路径跟踪控制成为控制理论和工程界的一个研究热点和难点。

已经提出的移动机器人路径跟踪控制方法包括PID控制、滑模控制、非线性状态反馈控制和智能控制方法等。

PID控制是目前在实际系统中应用较多的一种设计方法，该方法对系统模型依赖较少，且具有一定的鲁棒性，但参数优化困难，难以实现对模型变化的自适应调节。

滑模控制和非线性反馈控制能够在一定模型假设下通过离线设计保证系统的稳定性，但存在对模型依赖性大、难以实现在线学习和自适应的缺点。

移动机器人路径跟踪的智能控制方法是近年来得到普遍注意的一个研究领域，有关学者已提出了多种用于移动机器人路径跟踪的智能控制方法，如模糊路径跟踪控制器、基于神经网络的路径跟踪控制和模糊神经网络控制器等。

上述智能控制方法都具有不依赖于系统的动力学模型，具有自适应和学习能力等优点，但仍然需要解决知识自动获取的问题，即监督学习的教师信号设计。

增强学习方法作为一类基于机器学习的自适应最优控制方法，在复杂系统的控制器优化设计方面具有广泛的应用前景，目前已成功地应用于某些非线性系统的学习控制中。

<<增强学习与近似动态规划>>

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:<http://www.tushu007.com>