

刘 全, 傅启明, 龚声蓉, 伏玉琛, 崔志明. 最小状态变元平均奖赏的强化学习方法[J]. 通信学报, 2011, (1): 66~71

最小状态变元平均奖赏的强化学习方法

DOI:

中文关键词:

英文关键词:

基金项目:

作者

单位

[刘 全](#)

[傅启明](#)

[龚声蓉](#)

[伏玉琛](#)

[崔志明](#)

摘要点击次数: 268

全文下载次数: 179

中文摘要:

针对采用折扣奖赏作为评价目标的Q学习无法体现对后续动作的影响问题, 提出将平均奖赏和Q学习相结合的AR-Q-Learning算法, 并进行收敛性证明。针对学习参数个数随着状态变量维数呈几何级增长的“维数灾”问题, 提出最小状态变元的思想。将最小变元思想和平均奖赏用于积木世界的强化学习中, 试验结果表明, 该方法更具有后效性, 加快算法的收敛速度, 同时在一定程度上解决积木世界中的“维数灾”问题。

英文摘要:

[查看全文](#) [查看/发表评论](#) [下载PDF阅读器](#)

[关闭](#)

版权所有: 通信学报

地址: 北京东城区广渠门内大街80号通正国际大厦6层602室 电话: 010-67110006-869/878/881 电子邮件: xuebao@ptpress.com.cn
技术支持: 北京勤云科技发展有限公司