

目 录

1 研究课题介绍	1
1.1 选题背景及意义.....	1
1.1.1 选题背景	1
插图索引	2
表格索引	3
公式索引	4
参考文献	5
致 谢	6
声 明	7

主要符号对照表

s	状态（任意状态）
a	行动（任意行动）
π	策略
$\pi(s)$	策略函数
$\pi(a s)$	策略在状态 s 下采取行动 a 的概率
r	奖励
$r(s, a)$	在状态 s 下采取行动 a 获得的奖励
G_t	时间 t 之后的总奖励
$v(s)$	状态 s 的价值函数
$q_\pi(s, a)$	在策略 π 下，在状态 s 下采取行动 a 的价值
$V(s)$	状态 s 的价值函数的估计值
$Q(s, a)$	状态 s ，动作 a 的价值的估计值
$p(s' s, a)$	在状态 s 采取动作 a ，状态转移到 s' 的概率
$p(s', r s, a)$	在状态 s 采取动作 a ，状态转移到 s' 并且获得奖励 r 的概率
$\pi(a s, \theta)$	策略函数的参数是 θ ，在状态 s 下采取行动 a 的概率
π_θ	策略函数，其参数为 θ
$\nabla \pi(a s, \theta)$	策略函数的参数是 θ ，在状态 s 下采取行动 a 的概率对 θ 的 导数
γ	总奖励 G_t 计算时，对于后续 R_{t+t_i} 的折扣率
ρ	重要性采样比例（importance-sampling ratio）
δ_t	时序差异误差（TD error）
RL	强化学习（reinforcement learning）
MPC	模型预测控制（model-predictive control）
TD	时序差异（temporal-difference）
MDP	马尔可夫过程 (Markov Decision Process)
POMDP	部分可观测的马尔可夫过程 (Partially Observable Markov De- cision Process)
DQN	深度 Q-学习神经网络（Deep Q Network）
DRQN	深度循环 Q-学习神经网络（Deep Recurrent Q Network）

1 研究课题介绍

1.1 选题背景及意义

1.1.1 选题背景

自动驾驶成为近年来人工智能领域发展最迅猛的技术之一，Alphabet 旗下的子公司 Waymo 为代表的一批自动驾驶企业已经让车辆上路^[1]。不过在复杂场景下，这些车辆的行为仍然不够智能，需要人类驾驶员的干预。

自 Deepmind 的科学家在 *Nature* 发文提出深度强化学习在游戏中的应用后^[2]，强化学习作为人工智能的下一个发展大方向，也在 2015-2018 年迎来了爆发点。然而，尽管强化学习在棋牌等游戏中获得了巨大的成功，目前的强化学习在自动驾驶上仍没有成熟的应用。

把强化学习应用于自动驾驶是一个极其自然的想法。自动驾驶的应用场景本质上是移动机器人的问题，而强化学习应用于机器人更是由来已久^{[3][4][5]}。因此，在这个项目中我们考虑了一个较为复杂的自动驾驶场景——环岛场景，希望应用强化学习框架，设计出适合于自动驾驶在此类场景中决策的算法。

插图索引

表格索引

公式索引

参考文献

- [1] Waymo[EB/OL]. <https://waymo.com>.
- [2] Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518: 529 EP -.
- [3] Kober J, Bagnell J A, Peters J. Reinforcement learning in robotics: A survey[J]. Int. J. Rob. Res., 2013, 32(11): 1238-1274.
- [4] Polydoros A S, Nalpantidis L. Survey of model-based reinforcement learning: Applications on robotics[J]. Journal of Intelligent & Robotic Systems, 2017, 86(2): 153-173.
- [5] Deep reinforcement learning framework for autonomous driving[J]. Electronic Imaging.

致 谢

衷心感谢导师 xxx 教授和物理系 xxx 副教授对本人的精心指导。他们的言传身教将使我终生受益。

在美国麻省理工学院化学系进行九个月的合作研究期间，承蒙 xxx 教授热心指导与帮助，不胜感激。感谢 xx 实验室主任 xx 教授，以及实验室全体老师和同学们的热情帮助和支持！本课题承蒙国家自然科学基金资助，特此致谢。

感谢 L^AT_EX 和 Th_UT_HESIS^[2]，帮我节省了不少时间。

声 明

本人郑重声明：所呈交的学位论文，是本人在导师指导下，独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本学位论文的研究成果不包含任何他人享有著作权的内容。对本论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明。

签 名：_____ 日 期：_____