

華中科技大學

## 研究生课程报告

姓 名	张昭骏
学 号	M202573989
系、年级	计算机科学与技术 25 级
类 别	学术型硕士
报告科目	强化学习
日 期	2025 年 10 月 25 日

## 目 录

1	引言 .....	1
2	方法原理与设计 .....	2
2.1	问题建模 .....	2
2.2	蒙特卡洛学习方法 .....	2
2.3	Q-Learning 方法 .....	2
2.4	近似 Q-Learning 方法.....	2
3	方法实现 .....	3
3.1	环境实现 .....	3
3.2	蒙特卡洛学习实现 .....	3
3.3	Q-Learning 实现 .....	3
3.4	近似 Q-Learning 实现.....	3
4	实验与分析 .....	4
5	总结 .....	5

## 1 引言

## 2 方法原理与设计

### 2.1 问题建模

#### 2.1.1 Pacman 游戏环境

#### 2.1.2 马尔可夫决策过程建模

### 2.2 蒙特卡洛学习方法

#### 2.2.1 基本原理

#### 2.2.2 算法设计

### 2.3 Q-Learning 方法

#### 2.3.1 基本原理

#### 2.3.2 算法设计

### 2.4 近似 Q-Learning 方法

#### 2.4.1 函数近似原理

#### 2.4.2 特征设计

#### 2.4.3 算法设计

## 3 方法实现

### 3.1 环境实现

### 3.2 蒙特卡洛学习实现

### 3.3 Q-Learning 实现

### 3.4 近似 Q-Learning 实现

## 4 实验与分析

## 5 总结