

Deep Q-Learning

Modifications & Improvements

导师：Alex



Deep Reinforcement Learning with Double Q-Learning

带有Double-Q Learning的深度强化学习

作者：Hado van Hasselt等

单位：Google Deepmind

发表会议及时间：AAAI 2016



Prioritized Experience Replay

优先化经验回放

作者：Tom Schaul等

单位：Google Deepmind

发表会议及时间：ICLR 2016



Dueling Network Architectures for Deep Reinforcement Learning

使用Duel网络结构的深度强化学习

作者：Ziyu Wang等

单位：Google Deepmind

发表会议及时间：arxiv 2016

前期知识储备

Pre-knowledge reserve



Double Q Learning

充分了解从Q Learning 发展出的 Double Q-Learning的思想

Sum Tree

了解“线段树”这一个数据结构，包括原理和基本的时间复杂度

Heap

了解“堆”这一个数据结构，包括原理和基本的时间复杂度

Importance Sampling

Importance Sampling 的原理，计算方法和应用

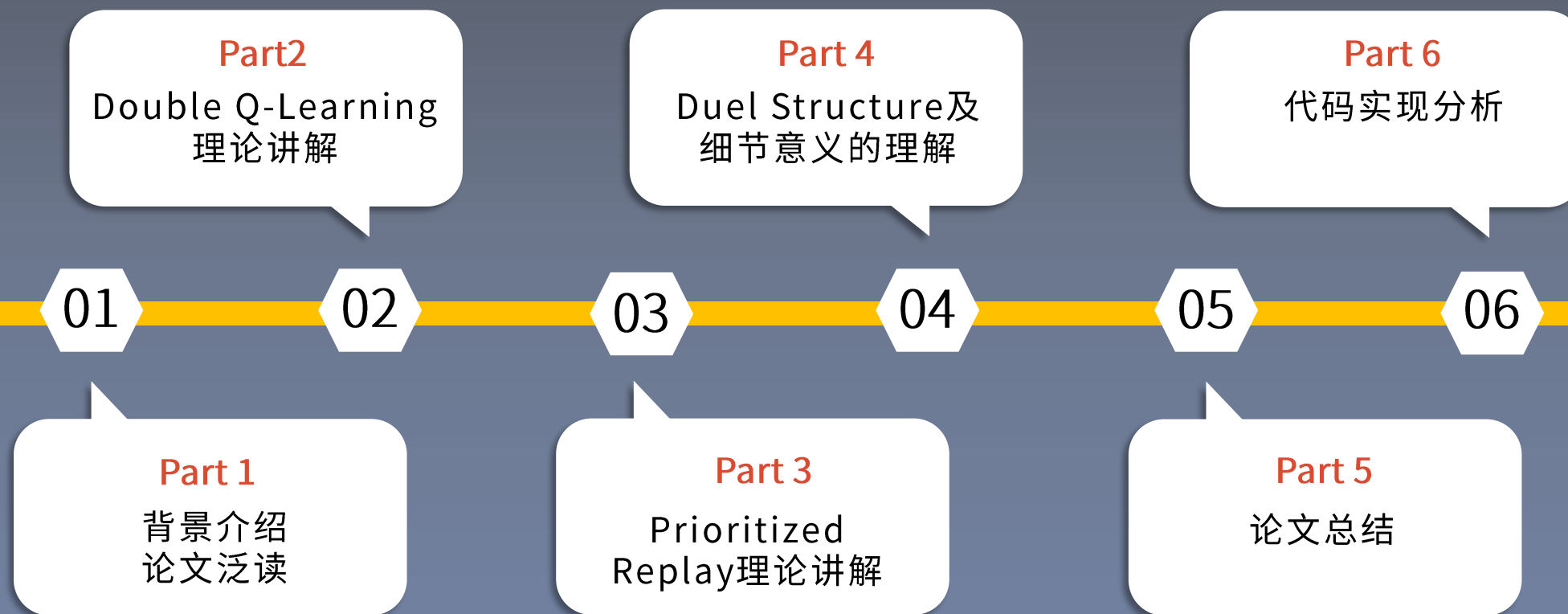
学习目标

Learning objectives

1. 学会一种将Double Q-Learning应用于DRL的算法
2. 知道Duel Network Structure并清楚实践细节和在DRL中的意义
3. 学会一种调整Replay Buffer中采样概率的方法
4. 学会用较为高效的算法进行不等概率抽样
5. 将新的内容与旧的内容结合，拓展原先的算法

课程安排

The schedule of course



第一课：论文导读

The first lesson: the paper guide

目录

1/ 论文研究背景、成果及意义

2/ Double Q Learning论文泛读

3/ Prioritized Experience Replay论文泛读

4/ Duel Network Structure论文泛读

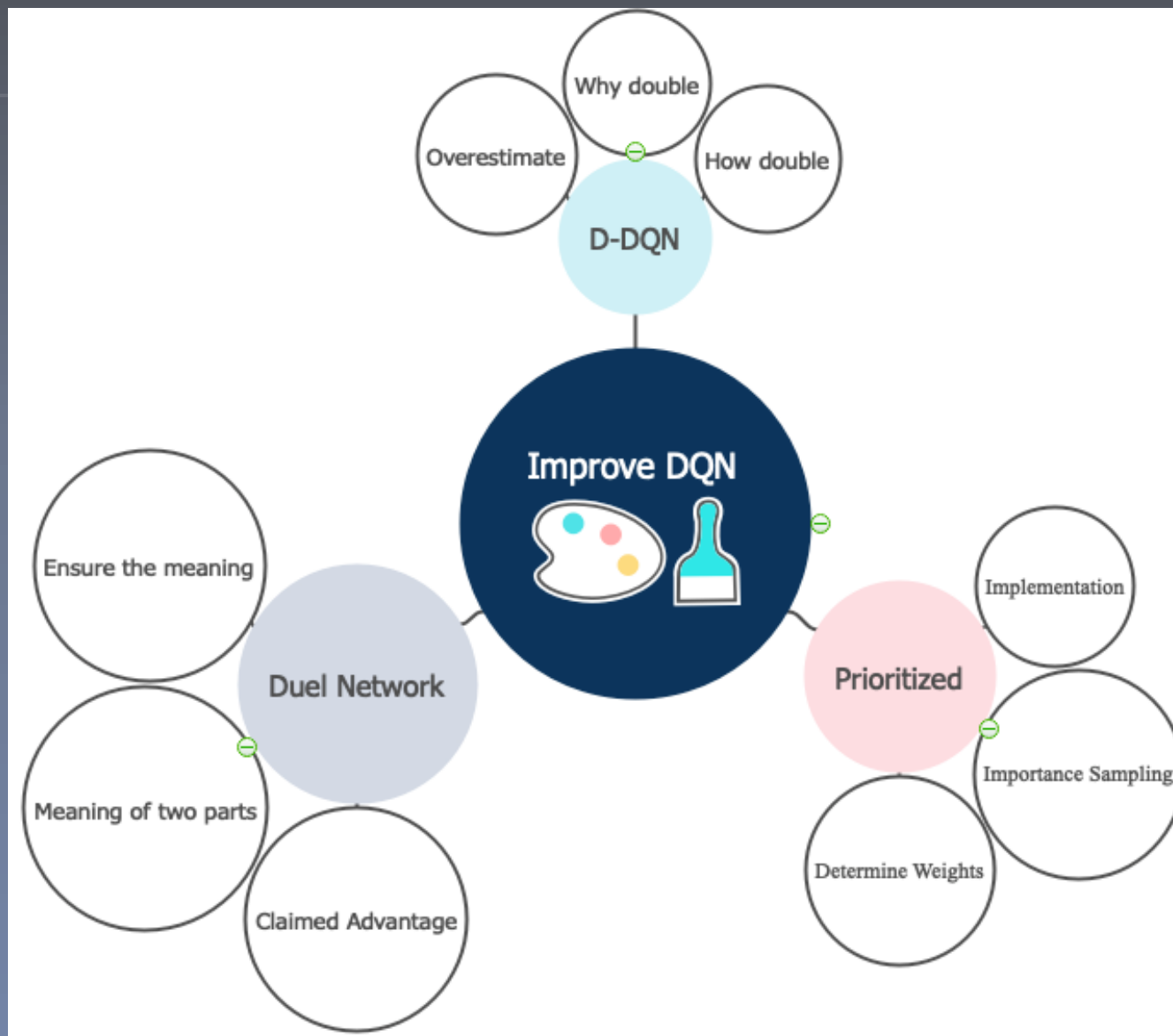
5/ 本课回顾及下节预告

知识树

knowledge tree



深度之眼
deepshare.net



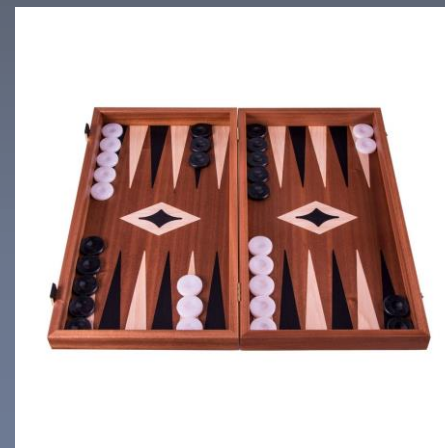
论文研究背景、成果及意义

研究背景

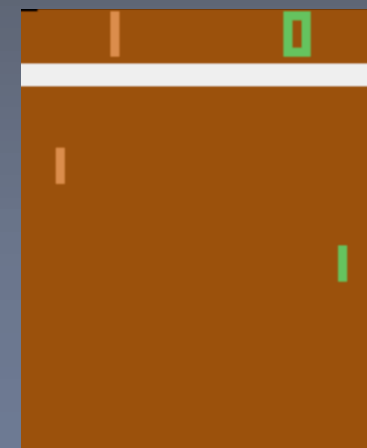
Research background

人们研究如何使用**强化学习**玩游戏，有传统的人工构建特征使用线性控制器实现的，也有结合深度学习的重要成果Deep Q-Learning

这鼓励了大量研究人员进一步推动这个领域的发展



Backgammon



Pong

研究意义

Research Meaning

Double DQN

- 展示了overestimation在使用神经网络时的存在和大小
- 充分验证了Double Q-Learning对于DQN的帮助

通过大量的示例，很直观的反映了overestimation的存在，并验证了其来源并不是某种简单的设定，而是普遍存在的现象



重点 重点来了！

研究意义

Research Meaning

Prioritized Experience Replay

- 提出了一种很符合直觉的加速训练的方法
- 针对该方法，很好的补全了工程上的细节，确保了方法和实用性

根据每个transition的误差判断重要性，再根据重要性进行不等概率采样，并通过importance sampling保证gradient方向没有引入过多的误差



重点 重点来了！

研究意义

Research Meaning

Prioritized Experience Replay

- 提出了一种很符合直觉的加速训练的方法
- 针对该方法，很好的补全了工程上的细节，确保了方法和实用性

对于实际应用的Prioritized Experience Replay补充了很多细节并解释了相应的原因



重点 重点来了！

研究意义

Research Meaning

Duel DQN

- 无

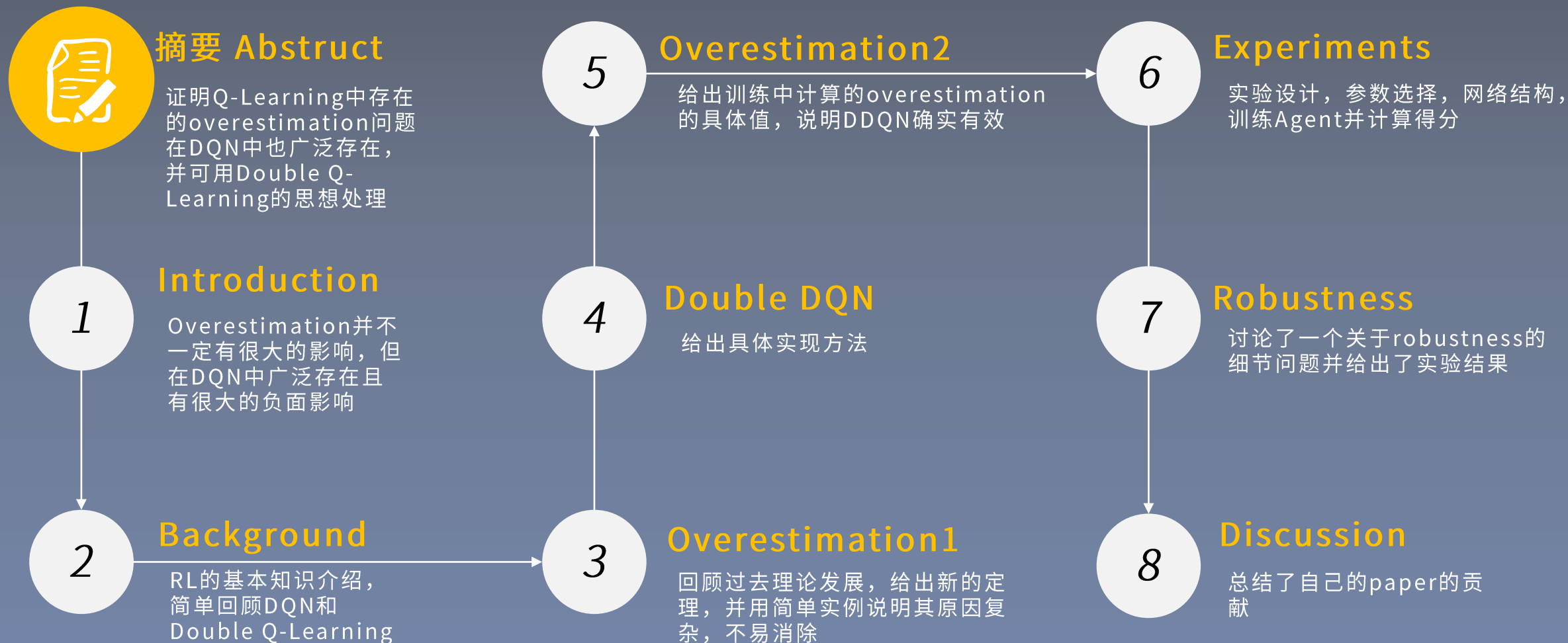
并没有什么很突出的研究意义上的贡献，但是实验结果还是值得了解一下的

论文泛读

Strcuture of Paper

论文结构

Structure of Papers (DDQN)



摘要

abstract (DDQN)

摘要核心

1. 我们都知道Q-Learning中的一个问题是overestimation，但并不清楚实践中这个因素影响多大，是否能避免
2. DQN中存在明显的overestimation
3. Double Q-Learning的思想可以推广到DQN来缓解overestimation

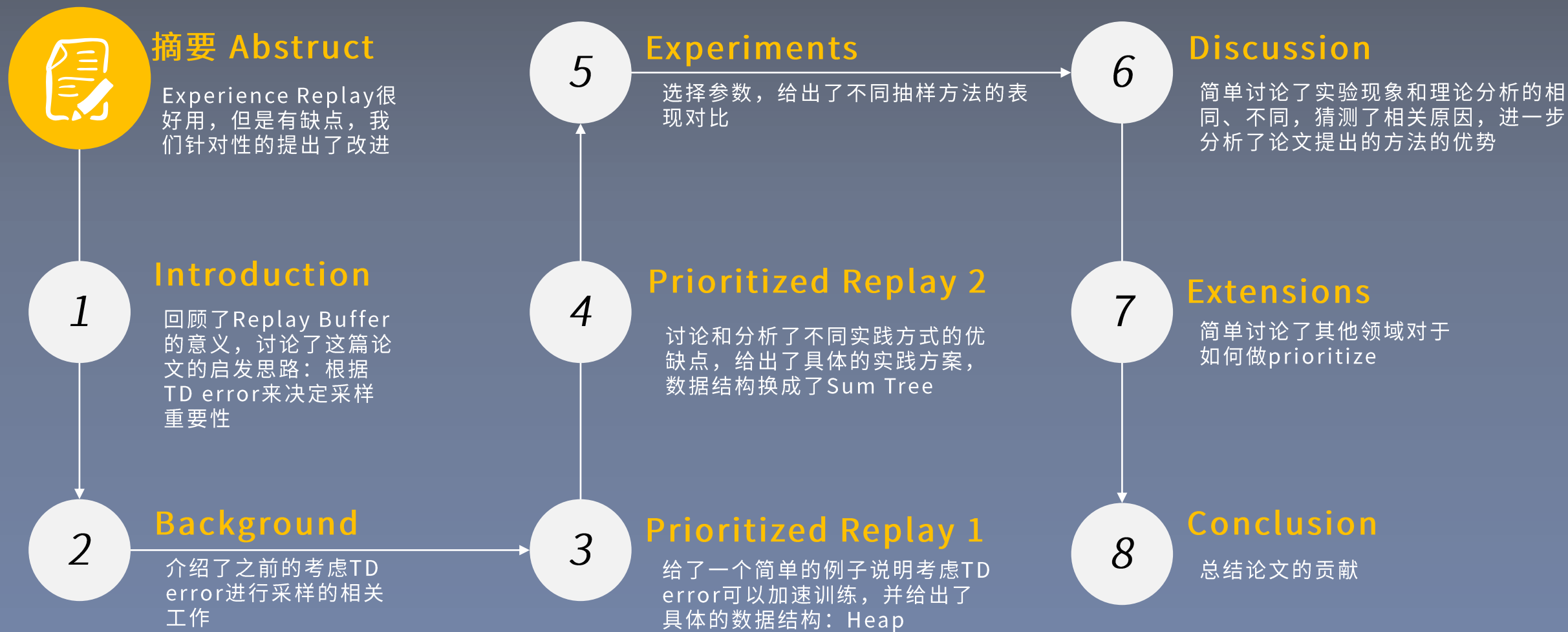
论文小标题

Paper title (DDQN)

1. Introduction
2. Background
 - 2.1 Deep Q Networks
 - 2.2 Double Q-Learning
3. Overoptimism due to estimation errors
4. Double DQN
5. Emperical results
 - 5.1 Results on overoptimism
 - 5.2 Quality of the learned policies
 - 5.3 Robustnestt to Human starts
6. Discussion

论文结构

Structure of Papers (Prioritized Experience Replay)



摘要

abstract (Prioritized Experience Replay)

摘要核心

1. Experience Replay帮助了RL的训练
2. 但是之前的方法对所有transition一视同仁，这不一定合理
3. 本文提出要多采样更”重要“的transition学习以提升效率
4. 我们将改进的Experience Replay用于过去的DQN进行训练，得到了更好的结果

论文小标题

Paper title (Prioritized Experience Replay)

1. Introduction

2. Background

3. Prioritized Replay

3.1 A Motivation Example

3.2 Prioritizing with TD-error

3.3 Stochastic Prioritization

3.4 Annealing the Bias

4. Atari Experiments

5. Discussion

6. Extensions

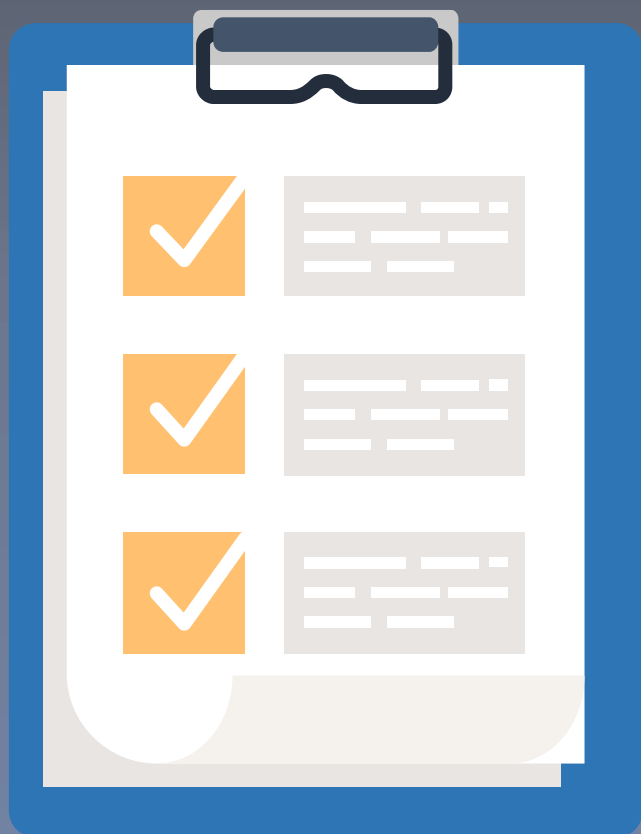
7. Conclusion

本课回顾及下节预告

Review in the lesson and Preview of next lesson

本课回顾

Review in the lesson



01 DQN的几个改进

Double DQN, Prioritized Replay Buffer, Duel Network Structure

02 Double DQN

关于Double DQN的内容的大致概括

03 Prioritized Replay Buffer

关于Prioritized Experience Replay的内容的大致概括

04 Duel Network Structure

几乎一笔带过了Dueling Deep Q-Learning的论文

下节预告

Preview of next lesson



01 Double DQN实验内容

分析从Double DQN论文的实验内容，让大家理解并重视 overestimation 的存在和潜在的问题并直观感受Double DQN的效果

02 Prioritized Experience Replay 思想

简单分析实验，帮助大家理解为什么可以根据TD-error作为重要性依据，并简单介绍Heap这一数据结构，解释为什么需要importance sampling

03 Prioritized Experience Replay 细节

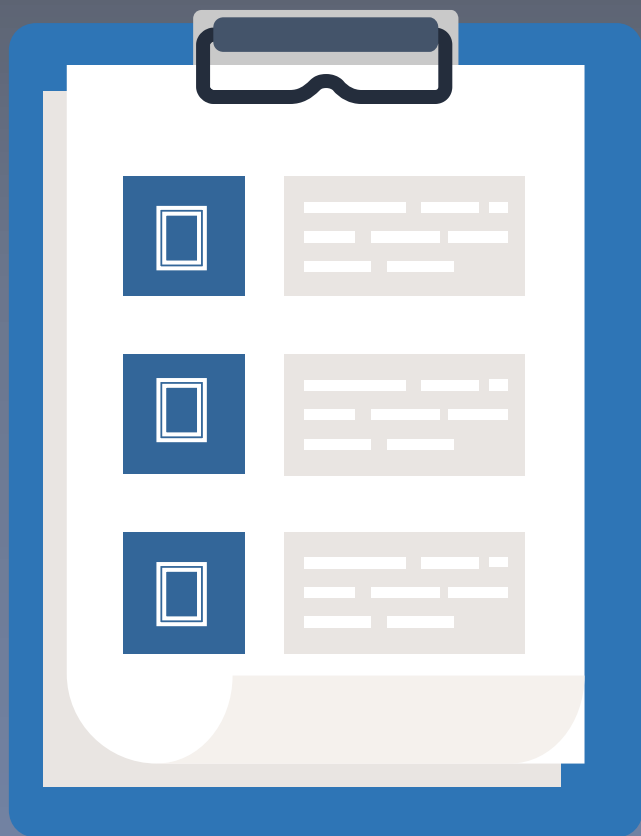
分析Prioritized Experience Replay的其他细节并理解其重要性

04 Duel Structure

介绍Duel Network Structure及其对应的RL的概念，并解释作者如何保证概念得到了对应

下节课前准备

Preview of next lesson



- 下载论文
- 泛读论文
- 可以的话自行预习Heap和Importance Sampling
- 筛选出自己不懂的部分，带着问题进入下一课时

—— 结 语 ——

**Whatever is worth doing is worth
doing well.**





深度之眼
deepshare.net

联系我们：

电话：18001992849

邮箱：service@deepshare.net

QQ：2677693114



公众号



客服微信

