

强化学习-2020秋-课程作业三

作业内容

在Atari环境下实现Deep Q-learning Network算法

作业描述

环境描述

本次作业的环境为gym上的Atari Game，默认为Pong。

玩家得到的观测：一个三维数组(12,84,84)，表示4帧彩色图像(3,84,84)的复合。

可执行的动作：离散的动作 {0,1,...,18}，具体数目参考具体的游戏。

奖励：玩家在游戏中每次移动都会得到奖励（可能为0或负数）。

游戏目标：尽可能达到高的累计奖励。

任务描述

请完成：

1. 依据Deep Q-learning算法，实现DQN及其各种变体（包括但不限于Double DQN, DQN with Prioritized Replay Buffer, Dueling DQN等等，至少需要实现一种变体），学习一个游戏策略。
2. 绘制你实现的Q-learning算法的性能图（训练所用的样本与得到的累计奖励的关系图，代码中提供了tensor board接口，可以直接调用）。

代码描述

代码文件夹code由'atari_ddqn.py', 'trainer.py', 'buffer.py', 'model.py'等文件组成，同学们只需了解这四个文件的功能即可。

'atari_ddqn.py': 类CnnDDQNAgent定义了一个基于DQN的agent，其中函数act代表这个agent的策略（即在观测下做出何种动作），函数learning代表这个agent如何学习策略。此处需要同学们在函数learning中利用batch=(s0,s1,a,r,done)来计算loss并更新网络参数。另外此文件中包含了一些超参数，大家可以谨慎地做一些修改以提升性能（对于完成作业来说不是必需的）。

'trainer.py': 类Trainer定义了与环境交互，收集数据并让agent基于数据更新策略的过程。其中self.board_logger定义了tensor board接口，用于绘制性能曲线。line 38提供了绘制gif的功能，大家可以提交一份自己训练出来的gif。此文件夹中没有需要修改的地方。

'buffer.py': 类RolloutStorage定义了Replay Buffer的结构，此处需要调整/重新定一个sample函数来实现**Prioritized Replay Buffer**。

'model.py': 类CnnDQN定义了网络结果，此处需要调整以实现**Dueling DQN**。

提交方式

完成的作业请通过sftp上传提交。上传的格式为一份压缩文件，命名为'学号+姓名'的格式，例如'MG20370001张三.zip'。文件中除原有代码外，还需包含 'performance.png'（性能曲线），'record.gif'（游戏效果展示）和'Document.pdf'（一份pdf格式的说明文档），文档内容至少需要包含：

1. 实验效果说明（如果实现了一种变体，请额外说明）。
2. 如何复现实验效果。
3. 算法的实现说明（如果实现了一种变体，请额外说明）。
4. 如果有相关的改进，也请在其中说明。

文档模板参见'DocumentExample.tex'和'DocumentExample.pdf'。（也可以使用自己的模板。）

Tips:

如何调用tensor board

参考资料

1. 安装tensorboard
2. 在code文件夹下输入tensorboard --logdir out，会收到一个端口号(例如6006)
3. 在浏览器中打开localhost:6006即可（如使用远程服务器，将localhost改成ip地址即可）