

DQN

测验, 6 个问题

1
point

1。
DQN

- ☒ Is based on an off-policy method.
 - ☐ Neither off-policy nor on-policy method
 - ☐ Is based on an on-policy method.
-

1
point

2。
How can we speed up the DQN training process?

- ☒ Use smaller model.
 - ☐ Use stacked state and action as input.
 - ☐ Use upsampling.
 - ☒ Use smaller input.
 - ☒ Parallelize learning.
 - ☐ Stack more layers.
 - ☒ Use strides.
 - ☒ Do not use pooling.
 - ☐ Stack 4 last states together.
-

1
point

DQN

测验, 6 个问题

3.

Why do we need to stack four last frames while playing Atari games?

- ☐ At least two stacked frames is the only type of input accepted by a convolutional layer.
 - ☒ It makes Atari game fully observable.
 - ☐ It allows to obtain additional training data.
 - ☐ Without such stacking a neural network will not learn (due to its depth).
 - ☐ It stabilises training.
-

1
point

4.

How can we help an agent to adapt to different scales of a reward signal?

- ☐ Use moving average of last N rewards as a new reward.
 - ☒ Adjust the agent learning rate.
 - ☐ Give the agent more training samples with higher learning rate.
 - ☒ Give the agent more training samples with lower learning rate.
 - ☒ Adjust the reward signal gamma.
 - ☐ Give the agent more training samples with same learning rate.
-

1
point

5.

Why target networks are commonly used?

- ☐ Target networks use smaller amount of parameters.
- ☒ For training stabilization.
- ☒ It is a heuristic against the problem of correlated training data.
- ☐

DQN

测验, 6 个问题

They increase training stability in tasks, where the reward scale is not known beforehand.

☐

They are mandatory when using an experience replay.

☐

Target networks work faster.

1
point

6.

What are the possible ways to improve experience replay (buffer)?

☐

Use smaller buffer (up to 1 thousand).

☐

Add more statistics to it, such as gamma, current learning rate, current step/epoch.

☐

Sample whole sessions from an experience replay instead of single experience (S,A,R,S') tuples.

☒

Use larger buffer (up to 1 million).

☒

Use samples more wisely - taking into account their contribution to the training process.



我（**伟臣 沈**）了解提交不是我自己完成的作业 将永远不会通过此课程或导致我的 Coursera 帐号被关闭。

[了解荣誉准则的更多信息](#)

Submit Quiz

