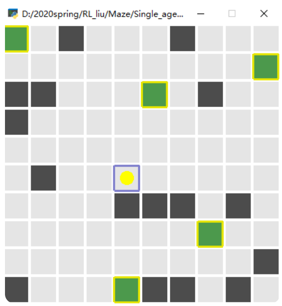
1. Maze 实验

一个Agent 5个Goals 找最短路径。

① environment



Start = (4,4)

ends = (7,2) (4,0) (0,9) (9,8) (5,7)

obstacles = 10 × 10 maze 20%

20个随机域

② DQN Model

256 Dense relu
128 Dense relu
64 Dense relu

'Adam' 'mse'

③ Observation states

1	2	3
4	5	6
7	8	9

\+ distance [current_position, goal]

\+ Start_position

\+ End_position

obstacles : 0 empty : 1

④ activation states

"Up" "down" "Left" "Right"

⑤ reward

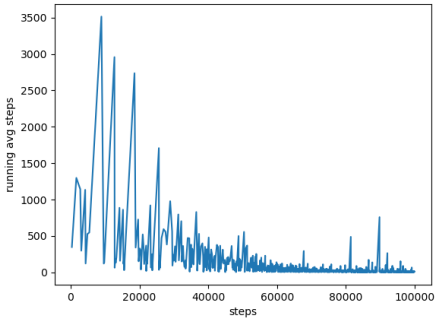
- Obstacles : -10
- moving : -1
- goal : 50

⑥ training :

ε-greedy.

replay experiment

double DQN



2. 多场景下的 Model 训练

场景1下训练好的模型参数，在场景2下继续训练。

训练结果是什么呢?

2个不同 env

本质 obstacle 坐标在变化.

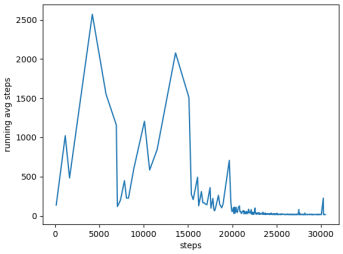
Seed num = 3 14.

一个场景训练迁移到新场景下，提高训练过程.

A2C 框架.


D:/2020spring/RL_liu/Maze/multi_envir...


D:/2020spring/RL_liu/Maze/multi_envir...


running avg steps
steps
0
500
1000
1500
2000
2500
0
5000
10000
15000
20000
25000
30000