

强化学习：作业四

黄彦骁 502023370016

December 7, 2023

1 作业内容

在gym Atari环境中实现Dyna-Q算法和Model-based算法并调节参数探究性能。

2 实现过程

实验一Dyna-Q算法所需要的DynaModel实际上只需要实现一个类似table形式的字典对得到的环境进行存储，而相应的Q-agent需要加入是否拿到钥匙的Q-values部分，同时在main1.py函数中进行相关的调用

实验二只需要将相应的改动在main2.py函数中加上即可。

整个过程对于算法收敛性的判断，采取average reward > 90和min reward > 80的条件进行判定。

3 复现方式

运行实验一在code4文件夹下运行 `python main1.py`.

运行实验二算法在code文件夹下运行 `python main2.py`.

4 实验效果

实验一的效果如下，由于实验具有一定的随机性，因此选取了一些值来对n进行大致的估计：

n	time	step
0	15s	9200
100	13s	5000
200	9s	3300
500	8s	2500
1000	5s	1700
1500	4s	1100
2000	3s	1000
2500	5s	1200
3000	5s	1100

可以看出大概n=1500左右算法需要的样本数量不再明显下降
实验二的效果如下所示：

n	m	start_planning	h	time	step
0	0	0	0	17s	9300
50	50	0	1	69s	14100
100	100	0	1	85s	12300
100	200	0	1	86s	8900
100	500	0	1	139s	7700
100	700	0	1	116s	4900
100	900	0	1	172s	5800
200	700	0	1	132s	5300
100	700	5	1	146s	6300
100	700	10	1	185s	8100
100	700	0	2	281s	11400

实验对比得出最优的参数组合为：n=100,m=700, start_planning=0, h=1。
可以看出n,m对实验效果的影响较大，而剩余两个参数并没有特别大的影响。

采取改进一之后的效果如下所示：

n	m	start_planning	h	time	step
100	200	0	1	84s	8000
100	500	0	1	168s	8800
100	700	0	1	137s	5300
100	900	0	1	185s	5800
200	700	0	1	164s	5800

采取改进一之后的最优参数为n=100,m=700, start_planning=0, h=1，并没有发生变化。

采取改进二的效果如下所示：

n	m	start_planning	h	time	step
100	200	0	1	136s	131000
100	500	0	1	131s	6600
100	700	0	1	151s	5900
100	900	0	1	165s	5100
200	700	0	1	165s	6000

采取改进一之后的最优参数为n=100,m=900, start_planning=0, h=1，变化不大，由于算法本身具有一定的随机性，可以得出结论两种改进对于算法效果影响不大。

实验探究三解答

1.对于Dyna-Q算法而言，应该在Model模拟的环境越来越逼近真实环境时采用Model提供的数据进行训练最好，因此我人为可能要采用动态n值能取得较好的效果，前期模型比较随机，与真实环境差异较大，采用较小的n值，当模型不断能够拟合真实环境时，逐渐增大n值，使得agent学习收敛加快。

2.对于采用神经网络的Model-based算法而言，神经网络的训练程度决定了算法的性能，直观来说，想要让神经网络训练的好，转移训练的频率和采样轨迹的条数能够直接影响神经网络的训练过程，但实际上我们也观察到，m和n的大小和最后效果也不是完全正相关的关系，我认为是二者过大时，agent学习初期从环境中采样得到的数据并不能完全贴合环境，而模型则针对这部分数据进行了大量训练，在过拟合的情况下导致模型与环境产生了一定的偏差，从而使得性能变差。因此最关键的应该是不能让神经网络与真实环境偏差过大，否则会耗费一定的训练过程来纠正这种偏差，从而使得学习过程变慢。

3.replay buffer和Dyna-Q是类似的想法，都想重复利用历史数据的价值来进行训练，而不在环境中进行采样。

5 小结

本次实验主要Model-based的相关算法，在这一类算法中，Model的性能变得非常重要，如果模型不能拟合到很好的效果，算法基本会变得无意义，因此，正确高效的使用数据来训练Model是需要进行深度研究的。