

## Exercise 8

### Task 1 - REINFORCE on the Cart-Pole

a)

$$\pi(a|s, \theta) =$$

b)

$$\begin{aligned}\nabla_{\theta} \log \pi(A_t|S_t, \theta) &= \dots \\ &= x(s, a) - \sum_b (q_{\pi}(s, a) - b(s)) \nabla(a|s, \theta)\end{aligned}$$

c)

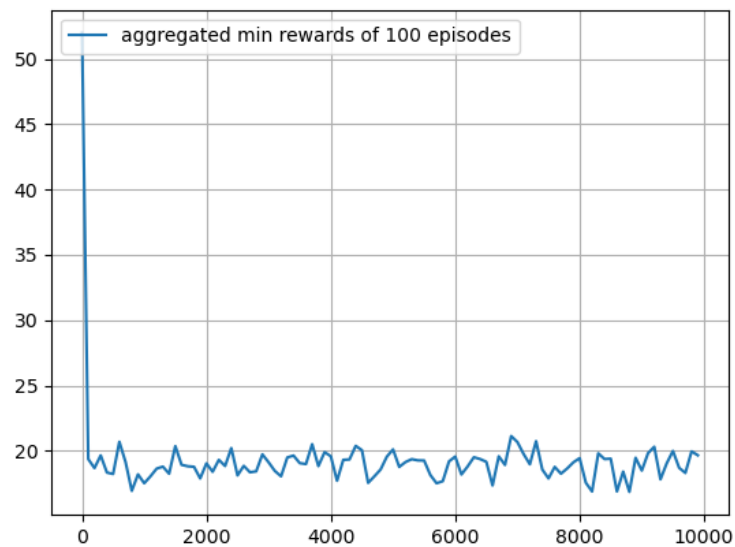


Figure 1: average episode lengths

d)