Homework Assignment

Course: Deep Reinforcement Learning

Sophie Haas, Johanna Linkemeyer, Tjorven Weber (Group 9)

May 1, 2022

1   Review for Group 12

Overall, nice homework! Still, we have some remarks :)

Task 1

- the state space in chess is actually unknown, but most definitely much larger than 768 (estimated complexity of 10^{46})

- the state transition function does not assess the validity of an action (to our knowledge), but, given an action and a state, returns a new state based on a probability distribution; in chess though, the state transition dynamic is deterministic

- action space: might need to consider that the movement of each piece (even same-typed pieces) are distinct actions and cannot not be summarized into only one action $\rightarrow$ that would result in a much larger action space

- rewards: The reinforcement book somewhere mentions, that it is actually better in the case of chess to not assign a positive reward for taking a piece (and negative reward for losing one) $\rightarrow$ that might actually lead to the policy trying to take out as many pieces as possible instead of focusing on winning

Task 2

- Very nice! The rewards could have been at least copy pasted though :p

Task 3

- Again nice answers! The examples would be more interesting though when being different from task 1 and 2

- The explanation on how actions are chosen based on the reward could be a little more extensive (as they are not directly based on the reward of the concurrent state, but on the expected overall return)

2 Review for Group 20

Task 1

- the description of the states and action space could be more precise (e.g. board is 8x8, 32 pieces, different allowed moves for each type of piece etc.)

- reward: the reinforcement book somewhere mentions that it is actually better in the case of chess to not assign a positive reward for taking a piece (and negative reward for losing one) → that might actually lead to the policy trying to take out as many pieces as possible instead of focusing on winning

- policy: as we understood it, this would be a joint categorical distribution (joint probability distribution over all of the categorical distributions for each chess figure)

Task 2

- probabilistic state dynamics: maybe also gravity or wind (generally just some physics simulation like with the self-driving car)

- reward: high reward for landing on the right place with low enough velocity, penalty for not landing on landing pad might be good to add

- policy: good thinking that multiple actions can be taken at once!

Task 3

- minor notation error for the transition function

- otherwise nice :)

3 Review for Group 16

Task 1

- nice that you added additional sources

- detailed and well explained (mentioning that set of legal moves differs for each color), really great job!

- nice that you also explained your thoughts and gave reasons but due to this, it seems like no clear answer is given (e.g. reward)

- policy: your answer is very vague; we probably were supposed to explain that it is a joint categorical distribution (where the categories are the respective actions of each piece)

Task 2

- set of states: simply the pixel space (again, your questions and comments make it difficult to understand what your final answer to the question is)

- your policy is basically the goal state, not a policy

Task 3

- the examples are not really helpful in understanding your definitions of the reward function and state transition function

- otherwise fine :)