Deep Learning Lab – Assignment 4

Theresa Eimer, Maryna Kapitonova, Hakan Yilmaz

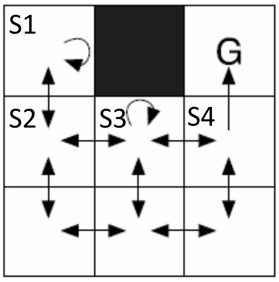
Introduction

The goal of this assignment is the investigation and the implementation of a deep Q-Learning algorithm applied to a grid-world environment similar to that in the third assignment. Before the actual implementation and to strengthen the theoretical background of the Q-learning algorithm, we were challenged to answer the questions in the first part of the exercise sheet.

Q-Learning (on paper)

Prerequisites:

- Deterministic state transitions and absorbing goal state G (agent can never transition out of it)
- Transitions into the wall are not possible, attempts result in same position as before
- All transitions have immediate reward -1, only transitions within the goal-state have reward 0
- Discounting factor $\gamma = 0.5$



1.1:

Update rule of Q-Learning: $q(i,u) = q(i,u) + \alpha\big(r(i,u) + \gamma \cdot max_u q(j,u) - q(i,u)\big)$

Transitions to the goal state can be handled just like other "normal" transitions.

In order to make the goal state absorbing, all transitions within the goal state leave the agent where it is.

1.2:

1. $q(S1, down) = q(S1, down) + \alpha\big(r(S1, down) + \gamma \cdot max_u q(S2, u) - q(S1, down)\big)$
   $= 0 + 1(-1 + 0.5 \cdot 0 - 0) = -1$

2. $q(S2, right) = q(S2, right) + \alpha\big(r(S2, right) + \gamma \cdot max_u q(S3, u) - q(S2, right)\big)$
   $= 0 + 1(-1 + 0.5 \cdot 0 - 0) = -1$

3. $q(S3, up) = q(S3, up) + \alpha\big(r(S3, up) + \gamma \cdot max_u q(S3, u) - q(S3, up)\big)$
   $= 0 + 1(-1 + 0.5 \cdot 0 - 0) = -1$

4. $q(S3, right) = q(S3, right) + \alpha\big(r(S3, right) + \gamma \cdot max_u q(S4, u) - q(S3, right)\big)$
   $= 0 + 1(-1 + 0.5 \cdot 0 - 0) = -1$

5. $q(S4, up) = q(S4, up) + \alpha\big(r(S4, up) + \gamma \cdot max_u q(G, u) - q(S4, up)\big)$
   $= 0 + 1(-1 + 0.5 \cdot 0 - 0) = -1$

All Q-values of state-action pairs that were not visited in this scenario remain unchanged and are therefore 0.

Deep Q-Learning (implementation)

The network architecture ..