# Reinforcement Learning

Exercise 1

1. Consider the following "treasure-hunting" game. A player moves through a two-dimensional maze searching for treasures. There are two treasures in this maze (T1 worth 50, and T2 worth 100). In the figure below the maze is shown together with the states (S1 through S7) and possible actions (South, North, East, West, Stay). Note that not all states allow all actions to be performed. The treasure T1 is located in S5, and T2 is located in S6. The reward function is defined as follows:

State s	Action a	Reward r(s,a)
S1	West	0
S2	West	0
S2	North	0
S2	East	0
S3	North	50
S4	North	100
S5	North	0
S5	East	0
S6	Stay	0
S7	East	100

Picture of the Maze:
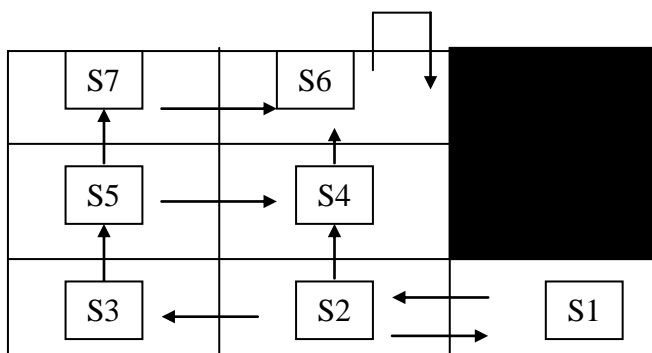
a) Compute the V* values for each state with discount factor $\gamma = 0.8$.
b) What is the optimal policy when $\gamma = 0.8$?
c) Does the optimal policy change if γ is set to 0.5 instead? If yes, give the new policy. If not, explain.
d) Compute the Q(s,a) values for the following state action pairs: (S2,West), (S6,Stay), (S3, North). Let $\gamma = 0.8$ and $\alpha = 1$.

e) Consider applying the Q-learning algorithm to the "treasure-hunting" game. Let Q' be the estimate of Q. Initially all Q' values are set to 0, and $\gamma = 0.8$ and $\alpha = 1$. Assume that the agent moves from state S1, via states S2, S3, S5, and S7, to state S6. Show how the Q' values are updated during this episode. Repeat the same episode twice more and show how the Q' values are revised during each episode.

2. Consider a real-time strategy (RTS) game such as Age of Empires. Each army has two types of units, bowmen and swordsmen. Swordsmen are stronger, but have only a short attack range. Bowmen are weaker, but are able to perform a long-range attack. Units have health-points that are reduced when they are hit. Formulate the problem of creating an AI for the units as a reinforcement learning problem. Specifically, define a:

   a. reward function.
   b. state representation.
   c. action representation.

   Discuss the suitability of your solution. Also, feel free to add more complexity to the RTS game, and discuss the difficulties this added complexity creates for reinforcement learning.

3. [Practical] Work through the OpenAI Gym tutorial: https://gym.openai.com/docs