# Reinforcement Example Problem

1. The left figure shows a rectangular grid world representation of a simple finite MDP. The cells of the grid correspond to the states of the environment. At each cell, four actions are possible: north, south, east, and west, which deterministically cause the agent to move one cell in the respective direction on the grid. Actions that would take the agent off the grid leave its location unchanged, but also result in a reward of -1. Other actions result in a reward of 0, except those that move the agent out of the special states A and B. From state A, all four actions yield a reward of $+10$ and take the agent to A'. From state B, all actions yield a reward of $+5$ and take the agent to B'.
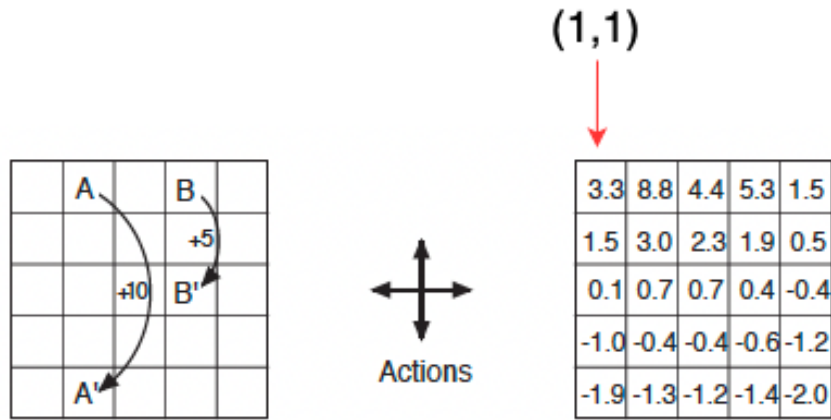
Suppose the right figure shows the optimal value function $v^*(s)$ for $\forall s$ and the discount factor is $\gamma = 0.9$. What is an optimal action for $s = (1, 1)$.

sol) the movement to the **right** direction is optimal action at $s = (1, 1)$.

$$\pi^*((1,1)) = \arg\max_a f(r, s'|a, (1,1))(r + 0.9 \cdot v^*(s'))$$

- a: left $\rightarrow v^*(1, 1) = (-1 + 0.9 \cdot v^*(1, 1)) = -1 + 0.9 \cdot 3.3$

- a: up $\rightarrow v^*(1, 1) = (-1 + 0.9 \cdot v^*(1, 1)) = -1 + 0.9 \cdot 3.3$

- a: down $\rightarrow v^*(1, 1) = (0 + 0.9 \cdot v^*(2, 1)) = 0 + 0.9 \cdot 1.5$

- a: **right** $\rightarrow v^*(1, 1) = (0 + 0.9 \cdot v^*(1, 2)) = 0 + 0.9 \cdot 8.8$