

Waterfall Gridworld

Alessandro Tenaglia

Machine and Reinforcement Learning in Control Applications

May 9, 2022

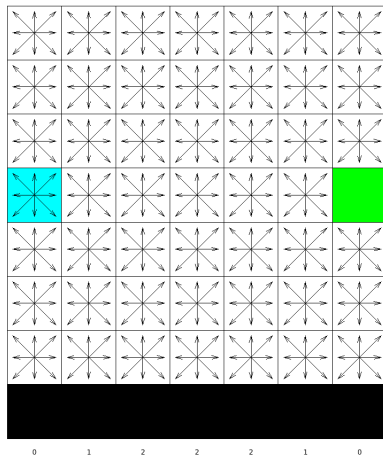
Problem



Learn to move in an unknown map with external disturbances

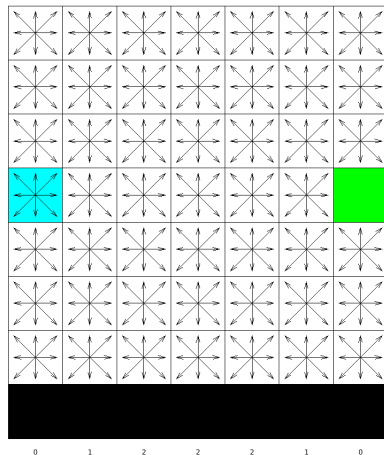
Problem formulation

- Consider the grid world on the right;
- The goal is to reach the green box;
- A waterfall pushes the agent toward the bottom of the grid



Problem formulation

- 8 possible directions:
N, S, E, W, NE, NW, SE, SW;
- Reward:
 - -1 for each step



Model

- The **state** is the position in the Gridworld
 - we have $X \cdot Y$ states.
- The **action** is the direction of the movement
 - we have 8 actions.

TD(λ)

TD(0)



TD(λ)

TD(0)



MC / TD(1)



TD(λ)

TD(0)



TD(λ)

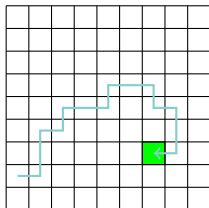


MC / TD(1)

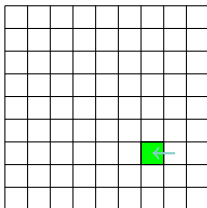


Advantages of $TD(\lambda)$

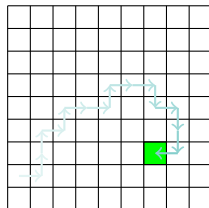
Path taken



Learnt from
 $TD(0)$



Learnt from
 $TD(\lambda)$



Eligibility traces

- The eligibility traces keep memory of the visited state;
- At each step, the eligibility trace of visited state is updated:

- **Accumulating traces:**

$$E_{t+1}(s, a) = E_t(s, a) + 1$$

- **Replacing traces:**

$$E_{t+1}(s, a) = 1$$

- **Dutch traces:**

$$E_{t+1}(s, a) = (1 - \alpha)E_t(s, a) + 1$$

- At each step, the eligibility trace of non-visited states decays

$$E_{t+1}(s, a) = \gamma \lambda E_t(s, a)$$