

# ARTIFICIAL INTELLIGENCE & CYBER SECURITY

## REINFORCEMENT LEARNING PASSIVE AND ACTIVE LEARNING

Nacir Bouali

[n.bouali@utwente.nl](mailto:n.bouali@utwente.nl)

UNIVERSITY  
OF TWENTE.

# PASSIVE Vs. ACTIVE LEARNING

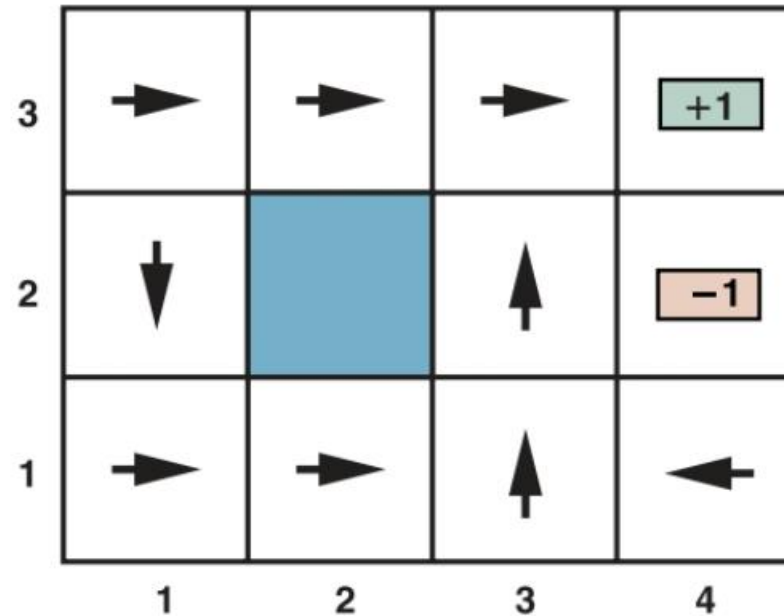
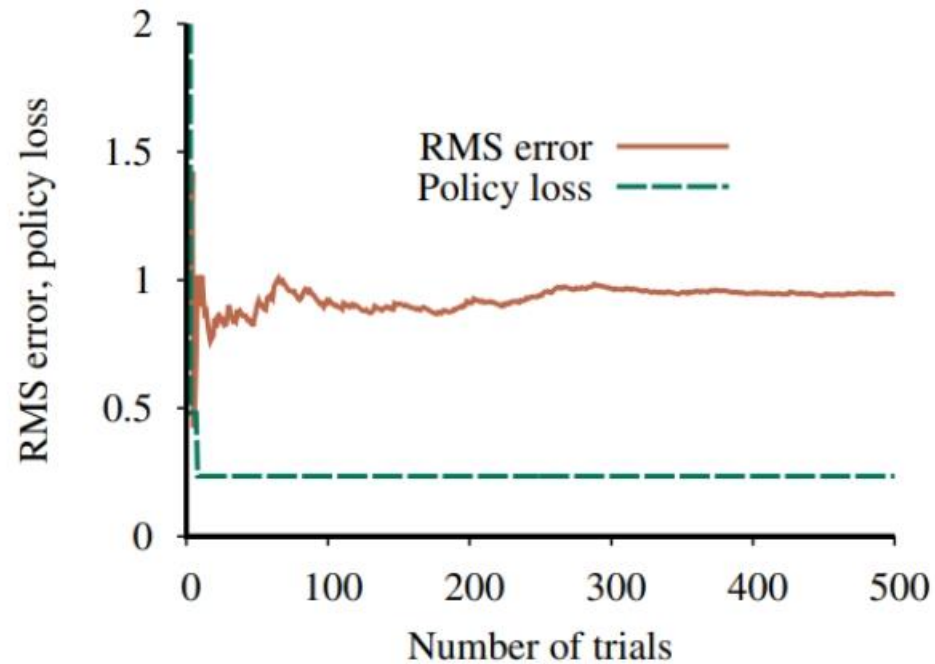
- With passive learning, the agent executes a fixed policy. This limits its freedom.
- Useful if the policy is the optimal policy, less useful otherwise.
- In active learning, the agent needs to also find the optimal policy.
- We can adapt ADP by doing the following;
  - Estimate  $P(s'|s,a)$  for more than one action.
  - With the retrieved transition model we can solve the MDP using value iteration.
  - The action the agent should then choose is;
    - $\pi(s) = \underset{a \in A(s)}{\operatorname{argmax}} \sum_{s' \in S} P(s'|s, a)U(s')$

# PASSIVE Vs. ACTIVE LEARNING

- The approach (updating PDA for active learning) updates the policy taken into account partial information about the environment.
- The agent updates its actions according to its current point of view, not the reality.
- Greedy approach
- Needed ← An approach that allows the agent to explore more the environment.

# EXPLOITATION Vs. EXPLORATION

- In state (2,1) the agent chooses to take the action « down » instead of « up ». It never explored that option.



# EXPLOITATION Vs. EXPLORATION

- Exploitation: Converges to solution faster.
- Exploration: Slows the learning.
- Balancing the two options is a challenging dilemma.

# EXPLOITATION Vs. EXPLORATION

- Introduce an exploration function  $f(u,n)$

$$U^+(s) \leftarrow \max_a f \left( \sum_{s'} P(s' | s, a) [R(s, a, s') + \gamma U^+(s')], N(s, a) \right)$$

$$f(u, n) = \begin{cases} R^+ & \text{if } n < N_e \\ u & \text{otherwise,} \end{cases}$$

- Where  $N(s,a)$  is the number of times action  $a$  has been executed in state  $s$ .
- Enforcing that  $a$  can be executed in  $s$   $N_e$  times.

# CONCLUSION

- We discussed MDPs and the two algorithms which we can use to solve them.
- We discussed active and passive learning.
- We discussed exploitation and exploration.
- There's more to say...

# FURTHER READINGS

- For those wishing to invest more time on RL, refer to this book;
  - Reinforcement Learning: An Introduction