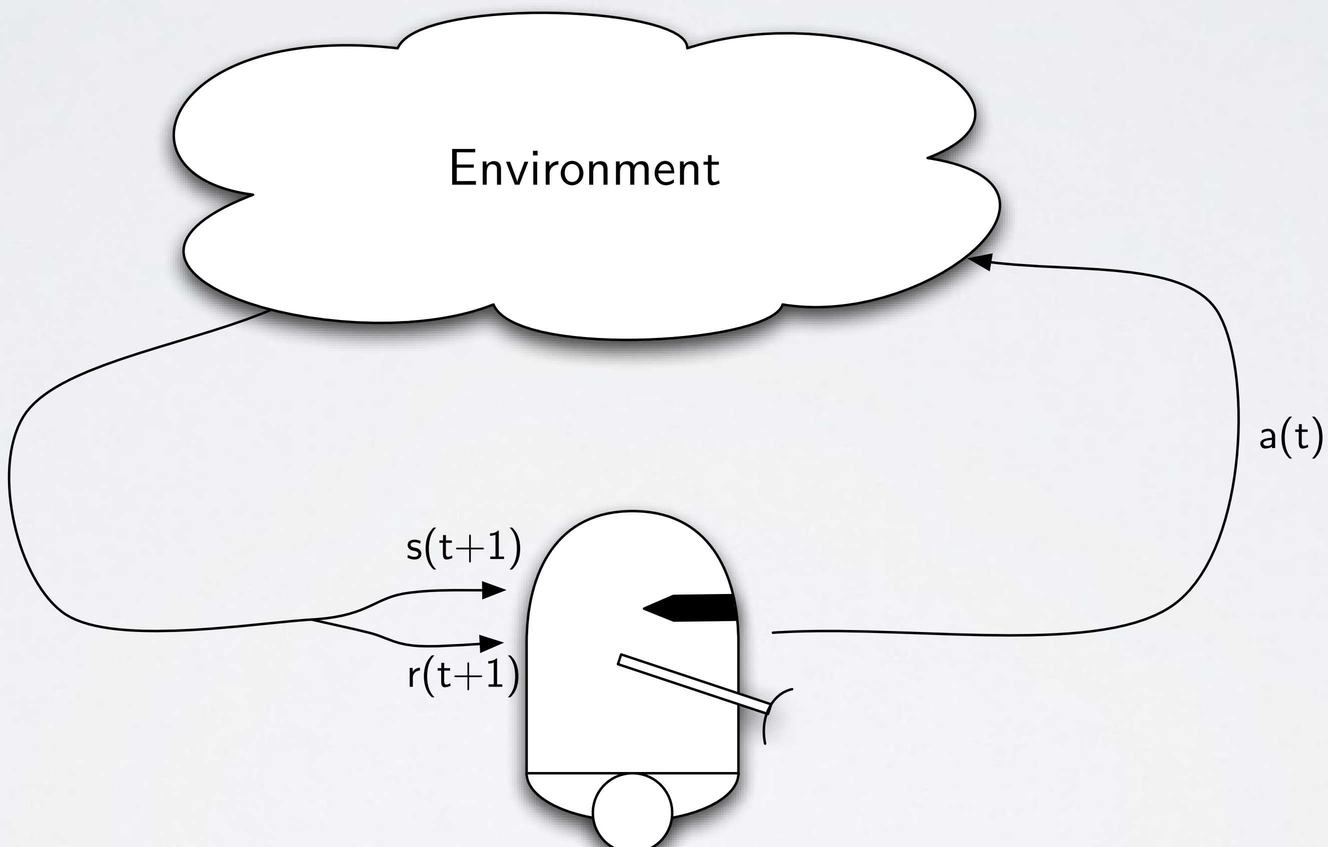
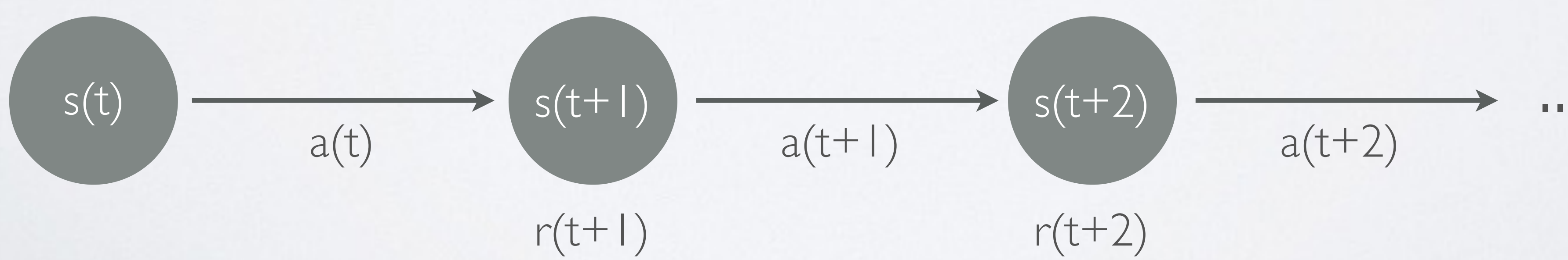
MULTI-AGENT REINFORCEMENT LEARNING

Sparse Interactions

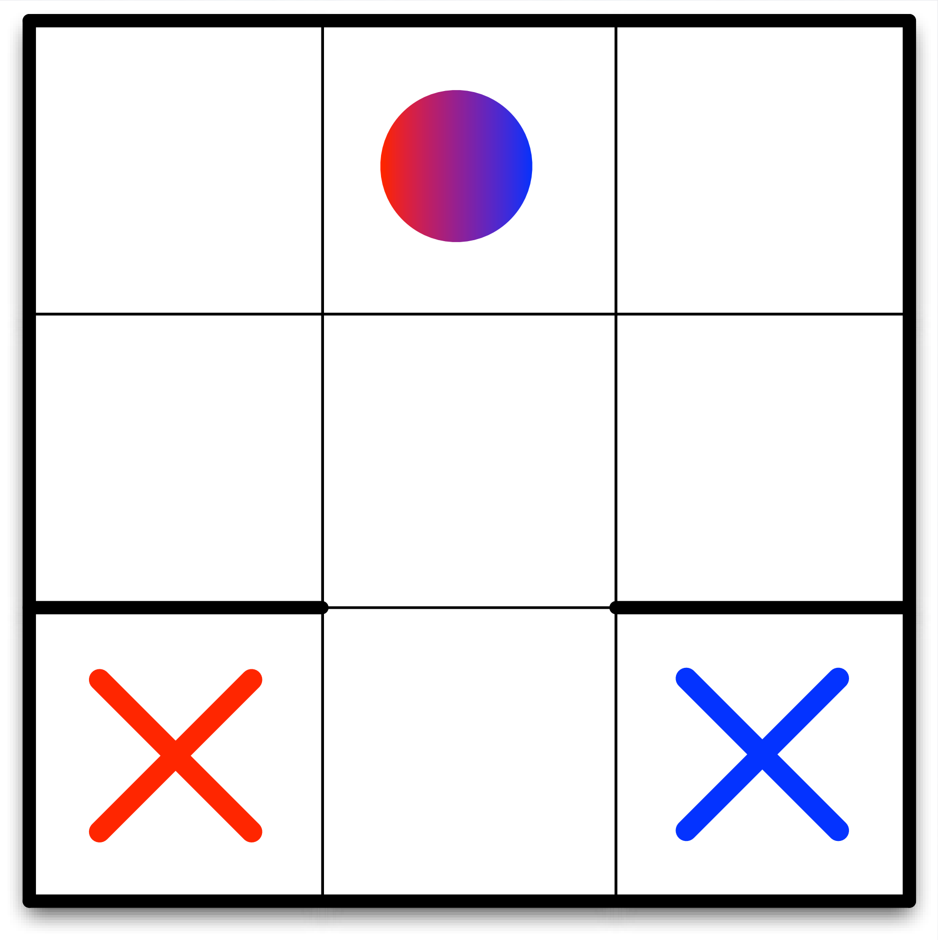
REINFORCEMENT LEARNING

- Agent acting in an unknown environment, learning to maximise a numerical reward signal

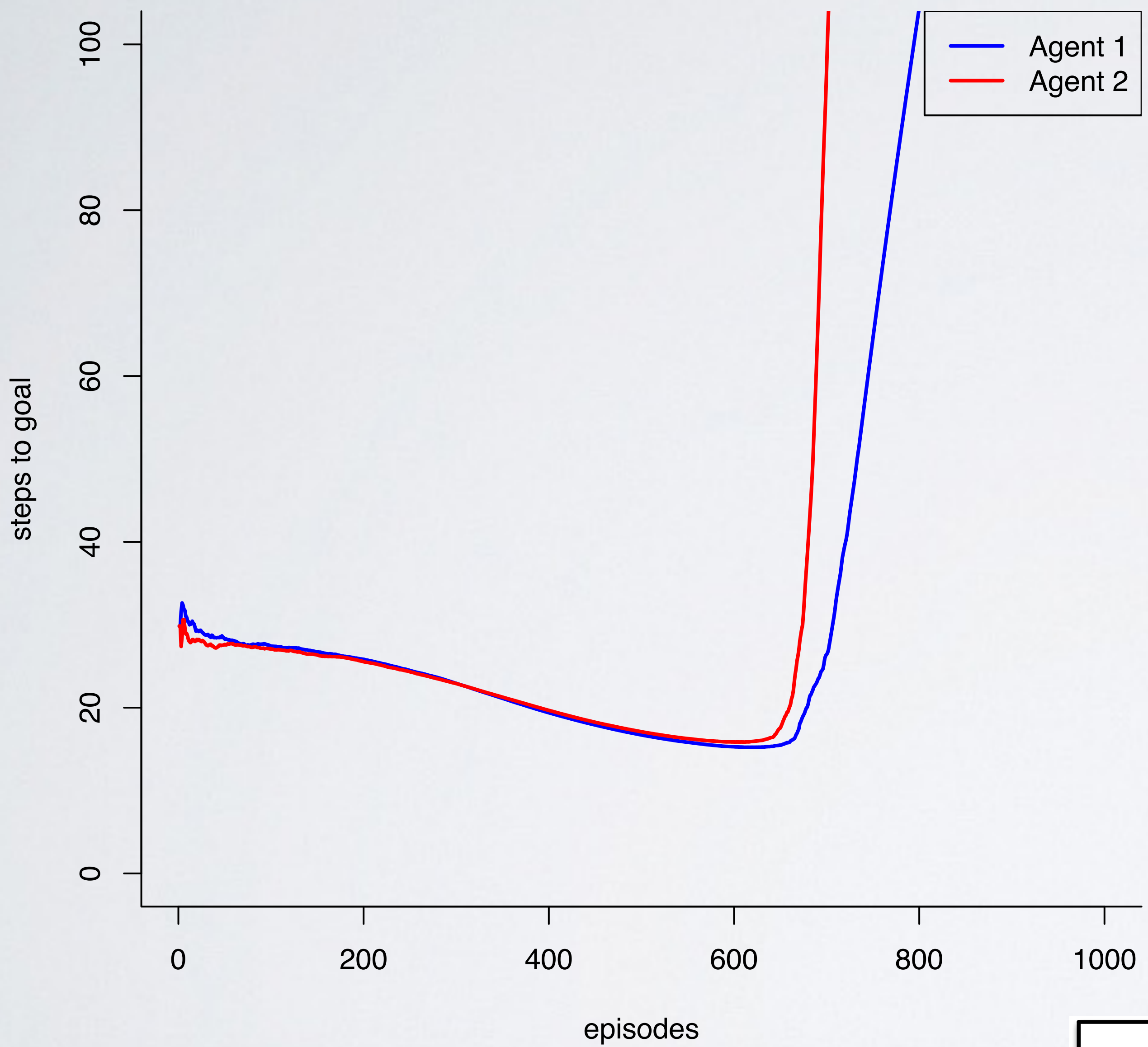


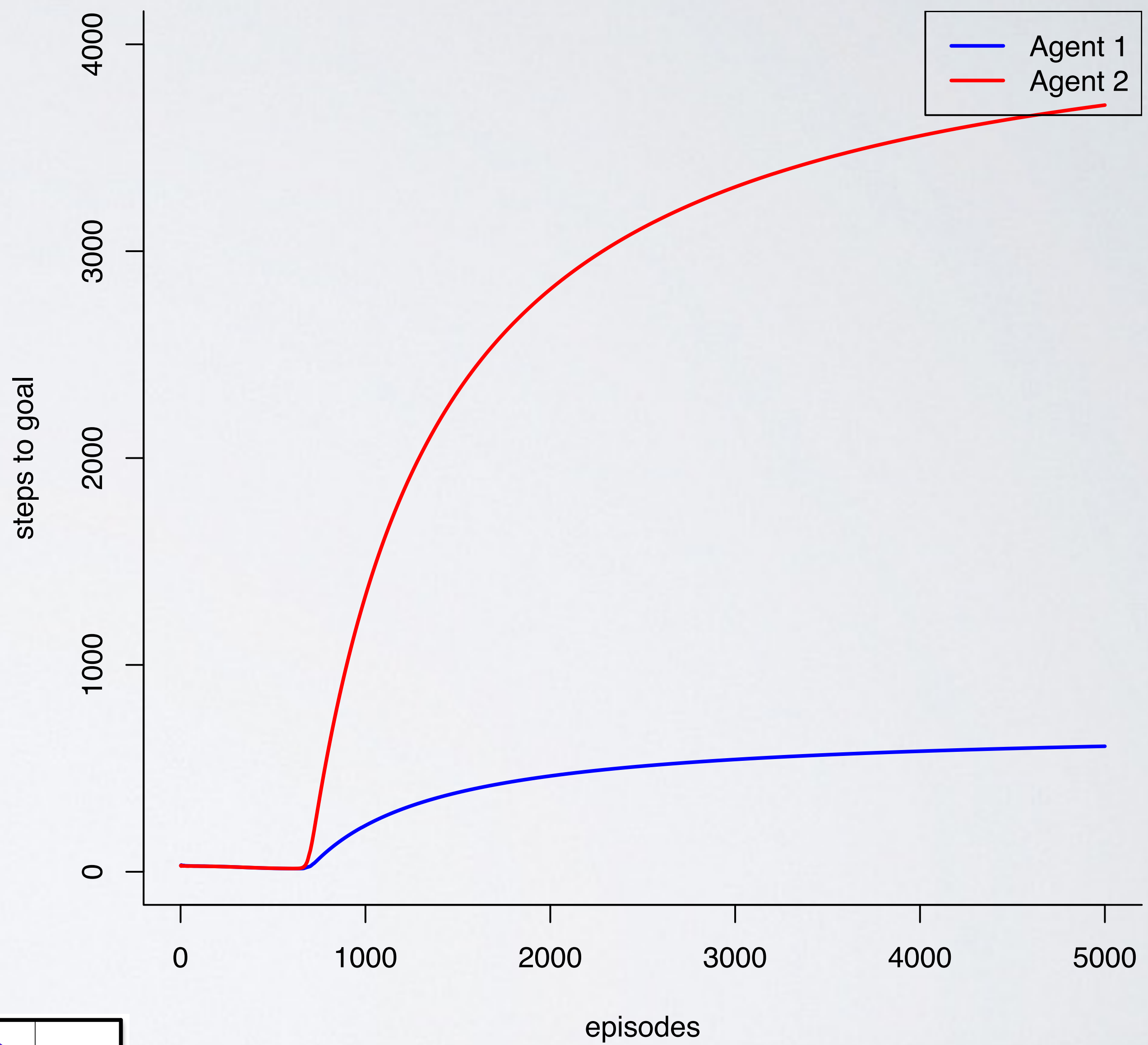


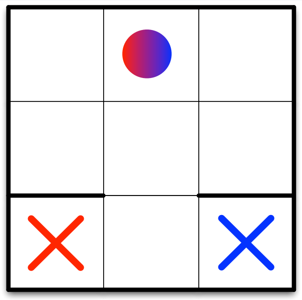
SIMPLE EXAMPLE



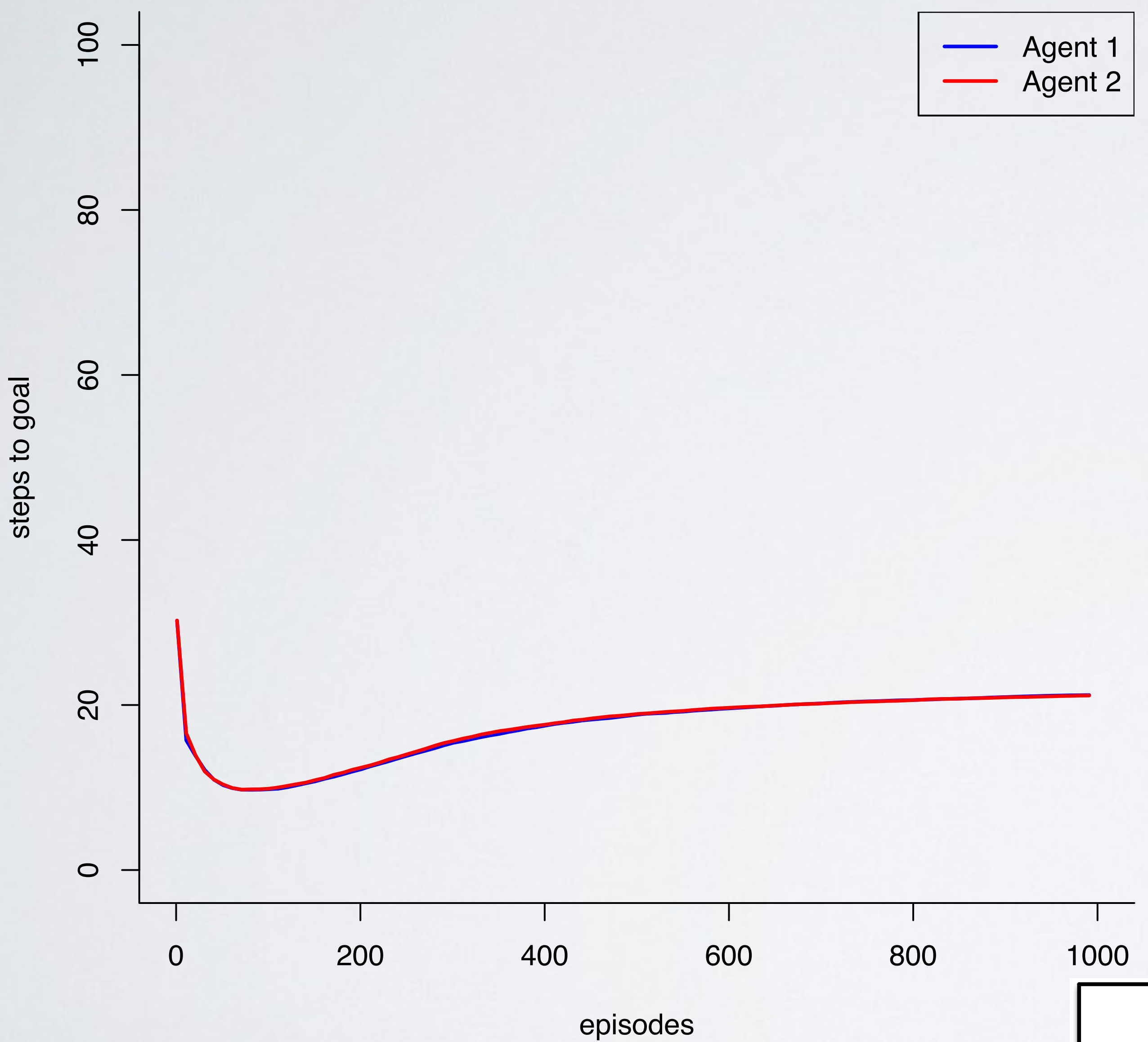
RL WITH BOLTZMANN EXPLORATION

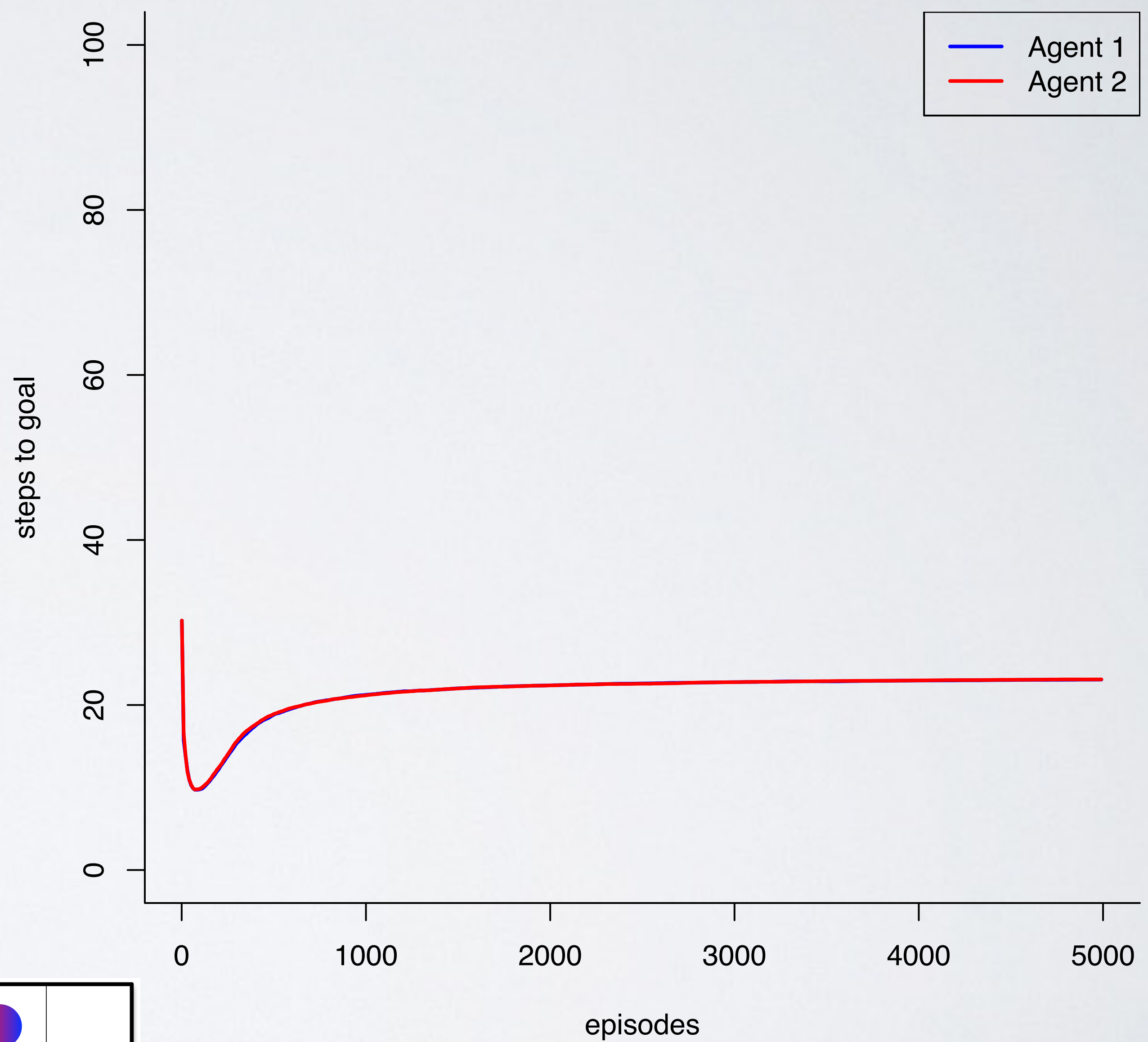


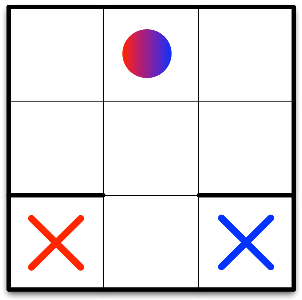




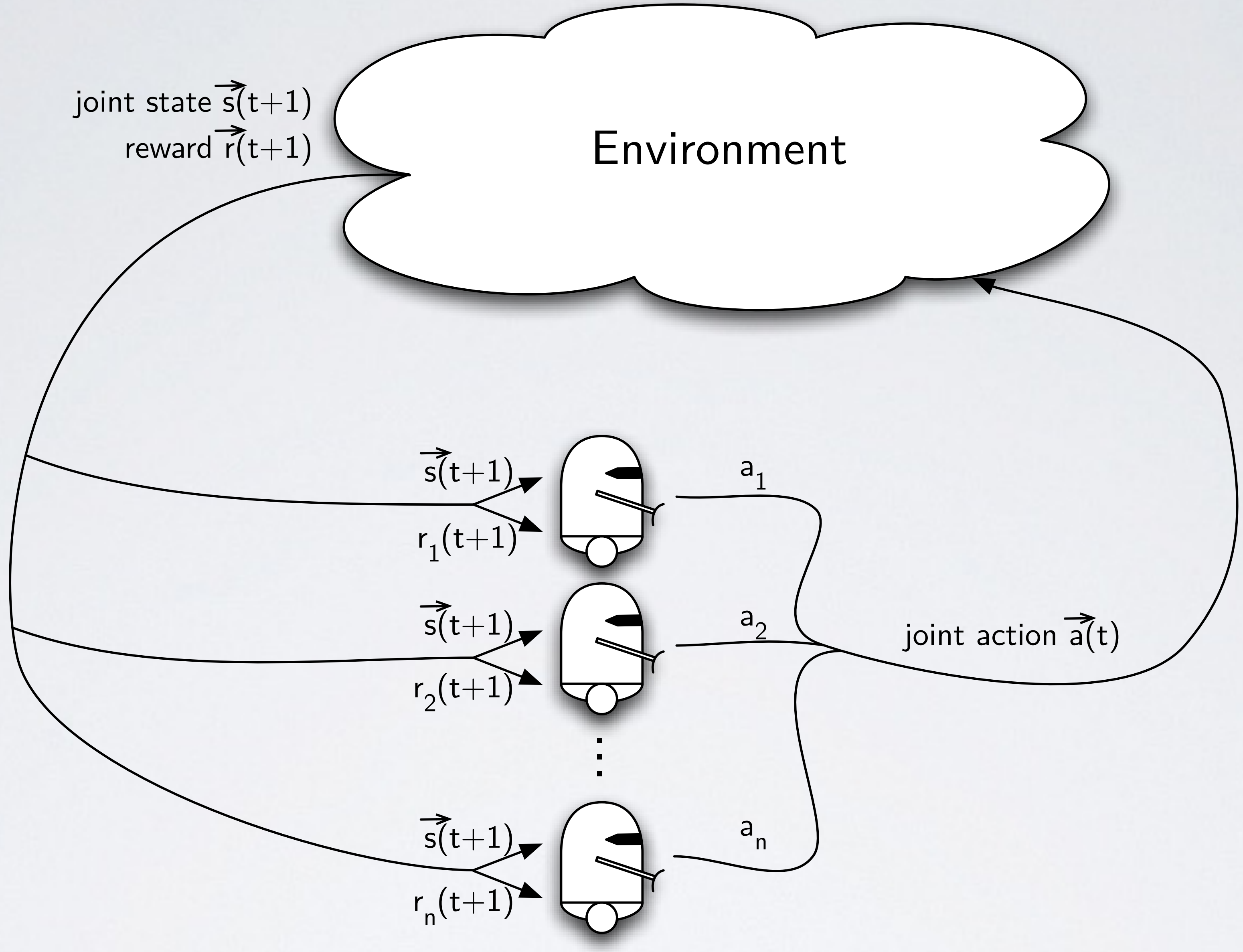
RL WITH **E**-GREEDY (**E** = 0.9)





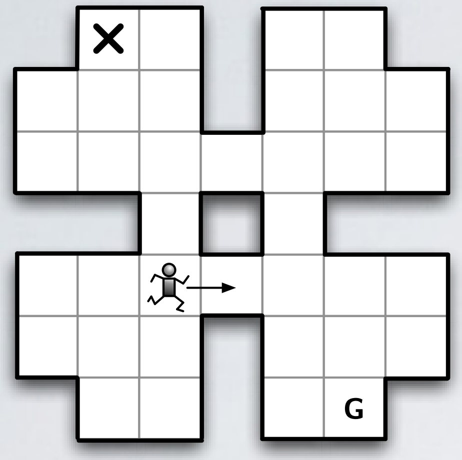


MULTI-AGENT REINFORCEMENT LEARNING



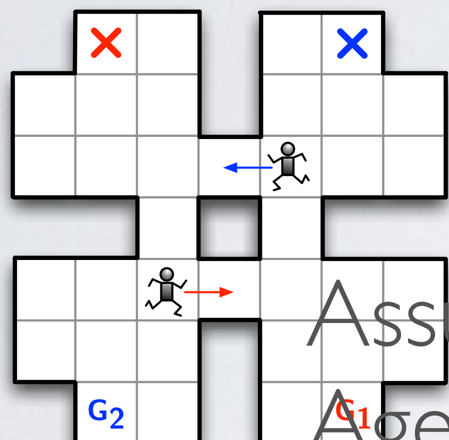
- Agents influence each other
- Possibly conflicting interests
- Observations
- Expensive communication

SPARSE INTERACTIONS



1 agent

Transitions & rewards are only dependent on 1 agent



2 agents

Far away and not interacting with each other

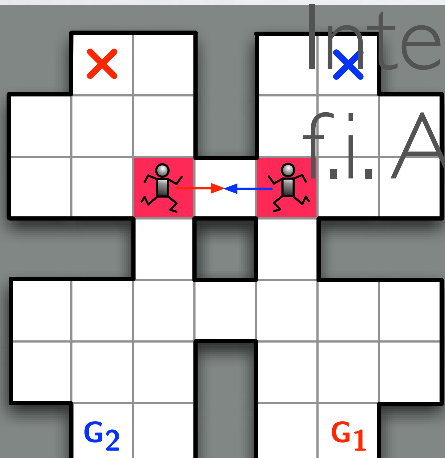
Transitions & rewards are independent of state/action of other agents

Assumptions:

Agents can do something useful alone

Interactions are sparse

f.i. Air traffic control, automated warehouses, ...



2 agents

Close to each other and interacting!!!

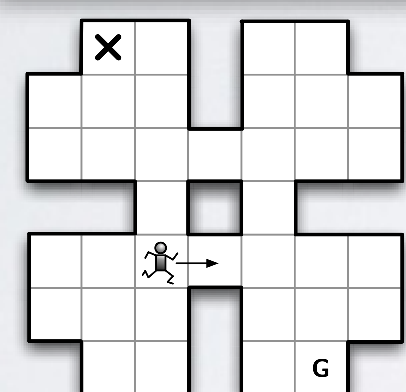
i.e. transitions & rewards are dependent

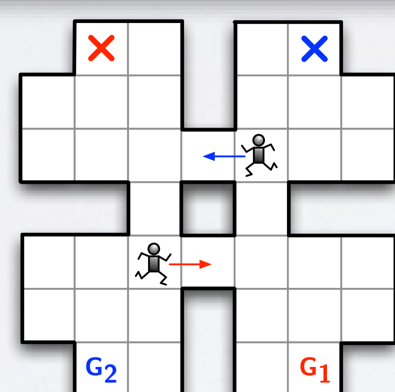
INTUITION OF SPARSE INTERACTIONS

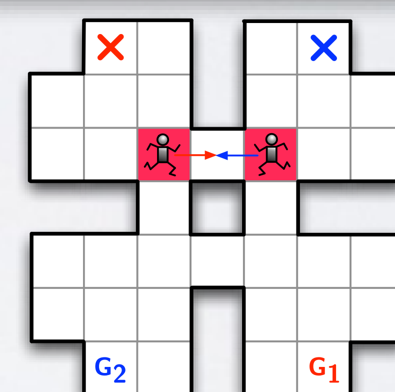
Is there influence from another agent?

No → Act independently, as if single-agent.

Yes → Use a multi-agent technique to coordinate.





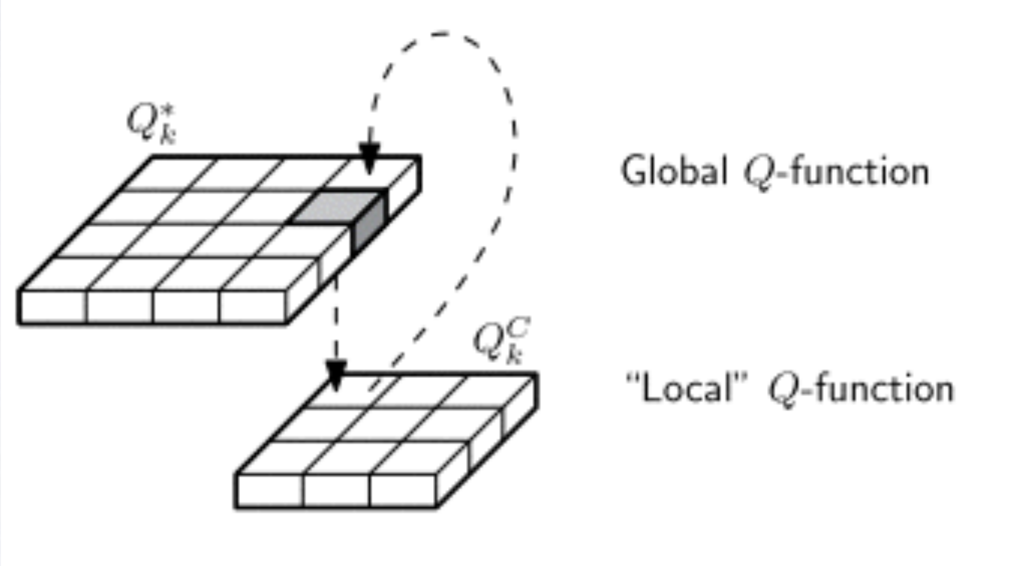


When should agents observe the state information of other agents to avoid coordination problems?

OUTLINE

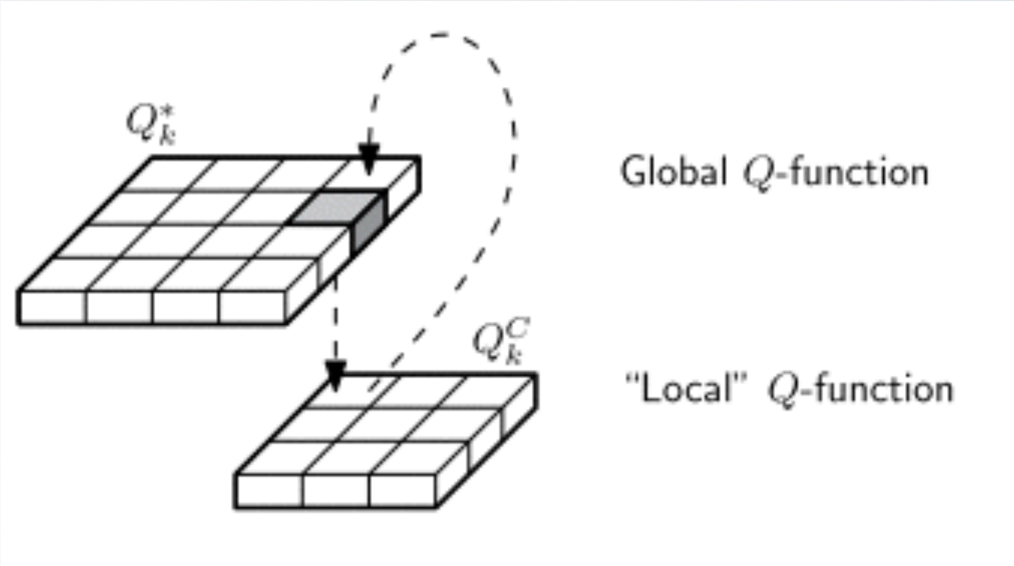
Learning of Coordination
2Observe
CQ-Learning
FCQ-Learning

Learning of Coordination



LEARNING OF COORDINATION

- Add Pseudo COORDINATE action
- External Active Perception
- Cost for coordination



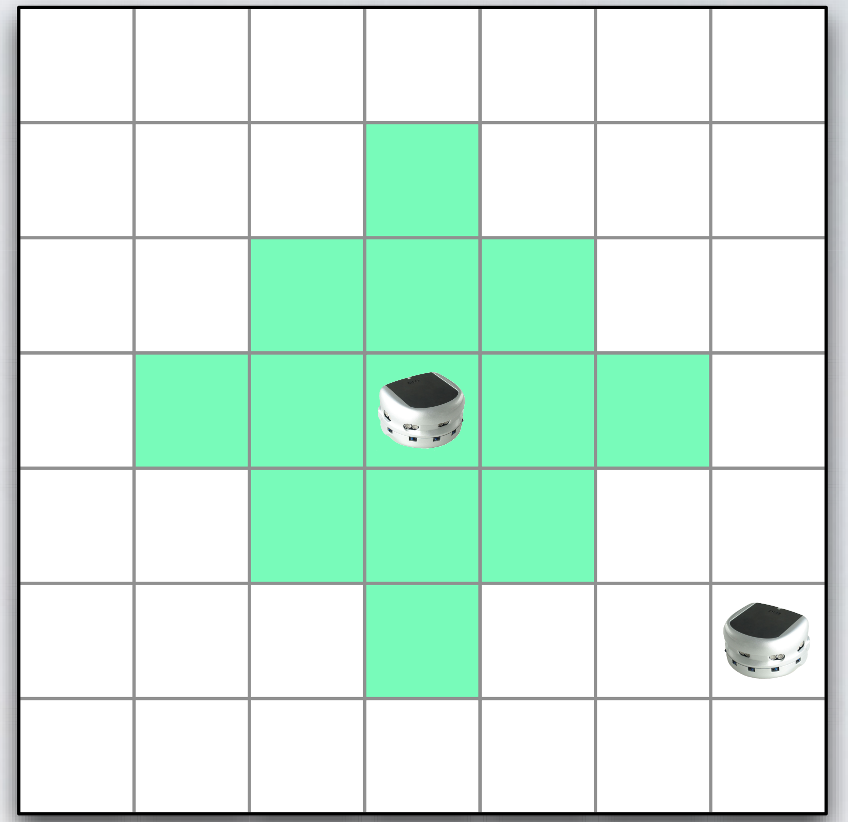
Melo, F. & Veloso, M. (2009). Learning of coordination: Exploiting sparse interactions in multiagent systems.
In: *Proceedings of the 8th International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS)*. International Foundation for Autonomous Agents and Multiagent Systems.

THE ALGORITHM

Algorithm 1 Learning algorithm for agent k

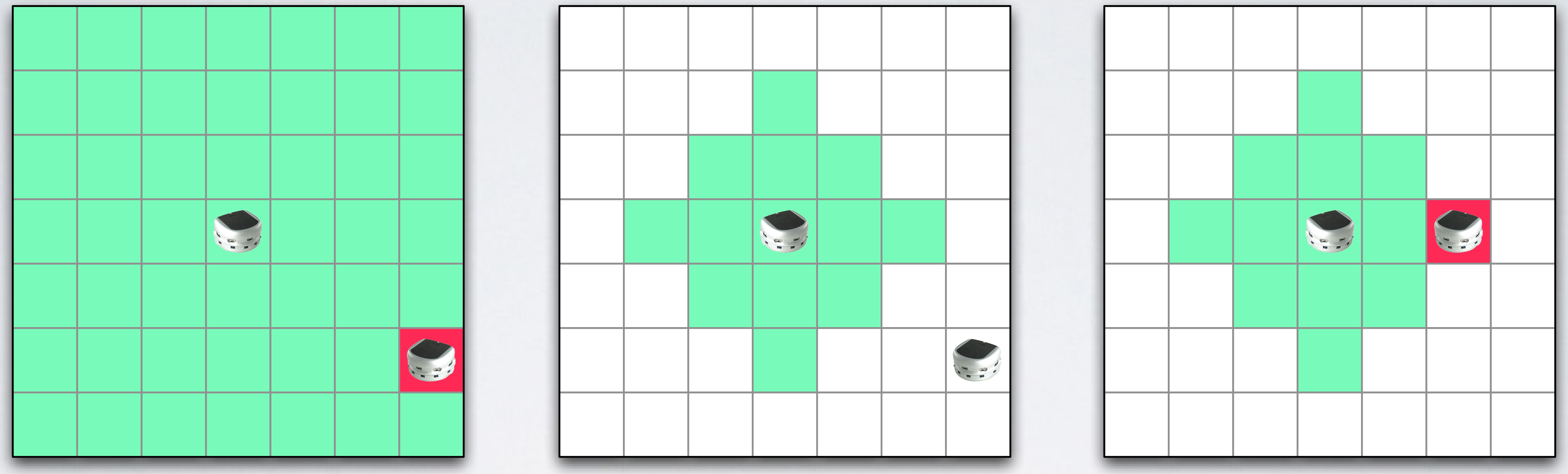
Initialize Q_k^* and Q_k^C;
Set $t = 0$;
while (FOREVER) **do**
  Choose $A_k(t)$ using π_e;
  if $A_k(t) =$ COORDINATE **then**
    if $ActivePercept =$ TRUE **then**
      $\hat{A}_k(t) = \pi_g(Q_k^C, X(t))$;
    else
      $\hat{A}_k(t) = \pi_g(Q_k^*, X_k(t))$;
    end if
    Sample $R_k(t)$ and $X_k(t+1)$;
    if $ActivePercept =$ TRUE **then**
      QLUpdate$\left(Q_k^C; X(t), \hat{A}_k(t), R_k(t), X_k(t+1), Q_k^*\right)$;
    end if
  else
    Sample $R_k(t)$ and $X_k(t+1)$;
  end if
  QLUpdate$\left(Q_k^*; X_k(t), A_k(t), R_k(t), X_k(t+1), Q_k^*\right)$;
  $t = t + 1$;
end while

2Observe



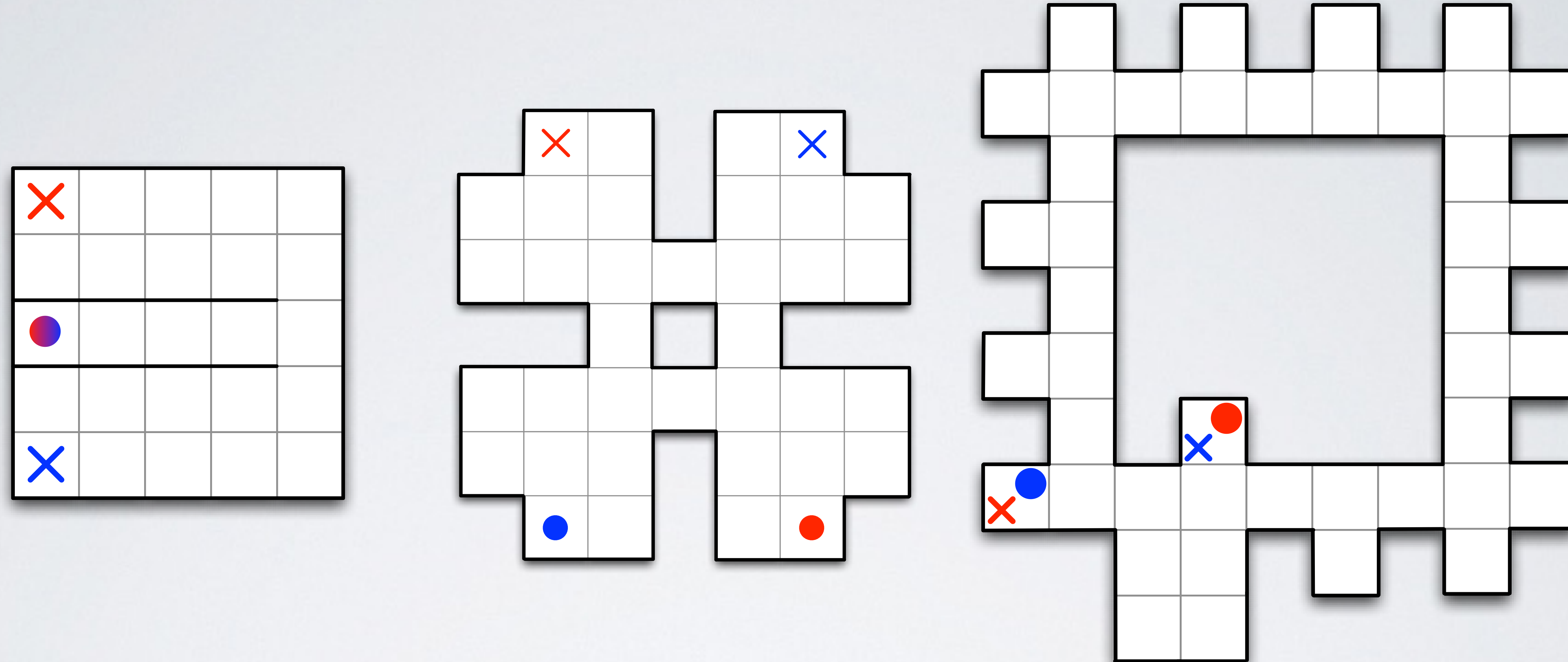
[De Hauwere et al. (2009)a] De Hauwere, Y.-M., Vrancx, P. & Nowé, A. (2009a).
Learning what to observe in multi-agent systems. In: *the 21st Benelux Conference on Artificial Intelligence*. Eindhoven, The Netherlands.

PROBLEM SETTING



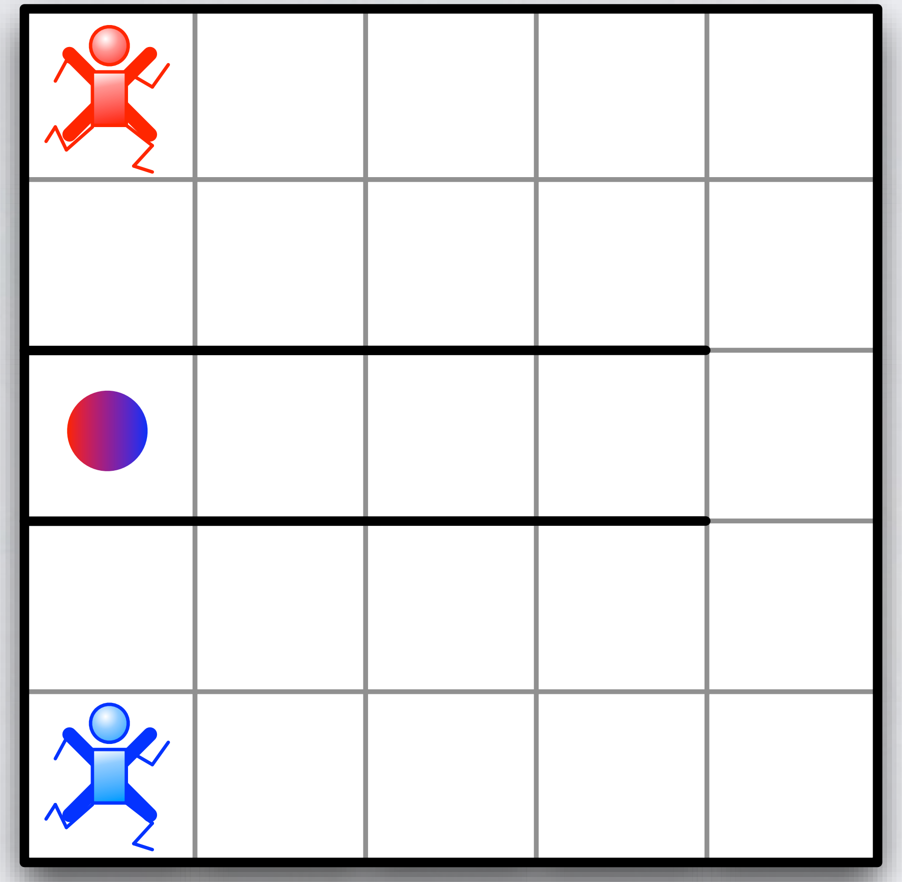
- Learn when to act upon sensory input
- Adaptive obstacle avoidance
- Save energy

EXPERIMENTAL SETTING



- Reach goal
- Avoid collisions

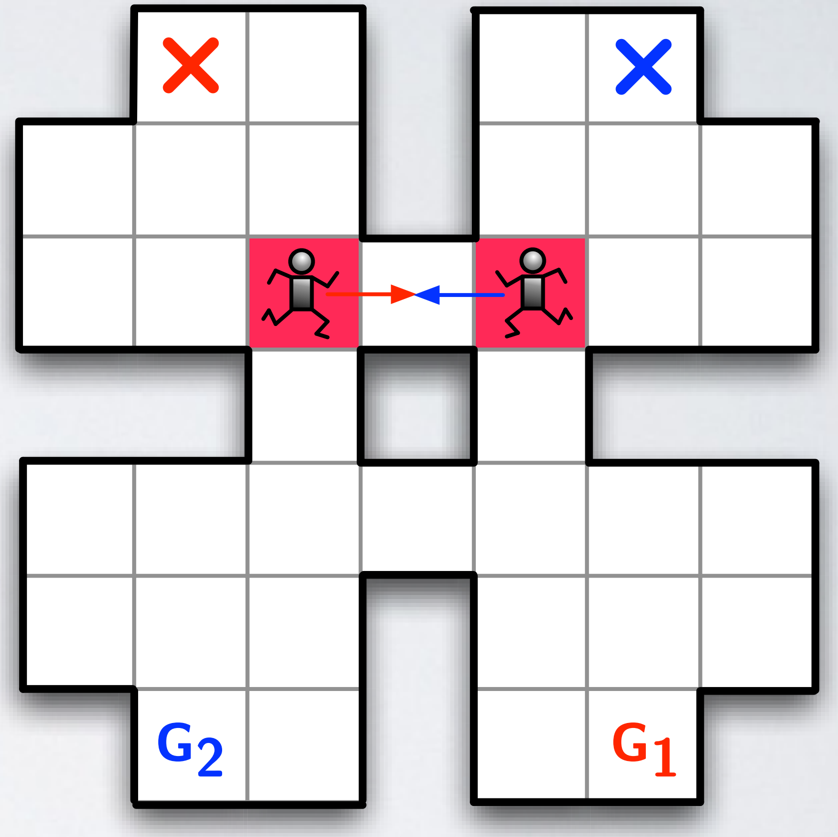
ILLUSTRATION (TUNNELTOGOAL)



COORDINATING

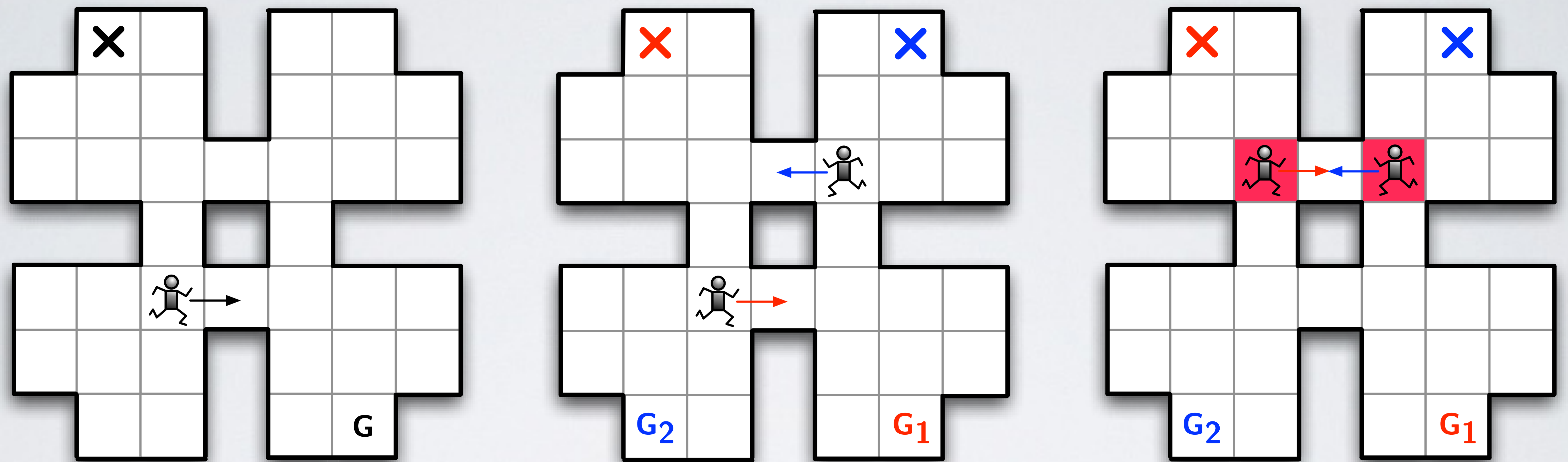
- Interactions are relative to the agent

CQ-Learning



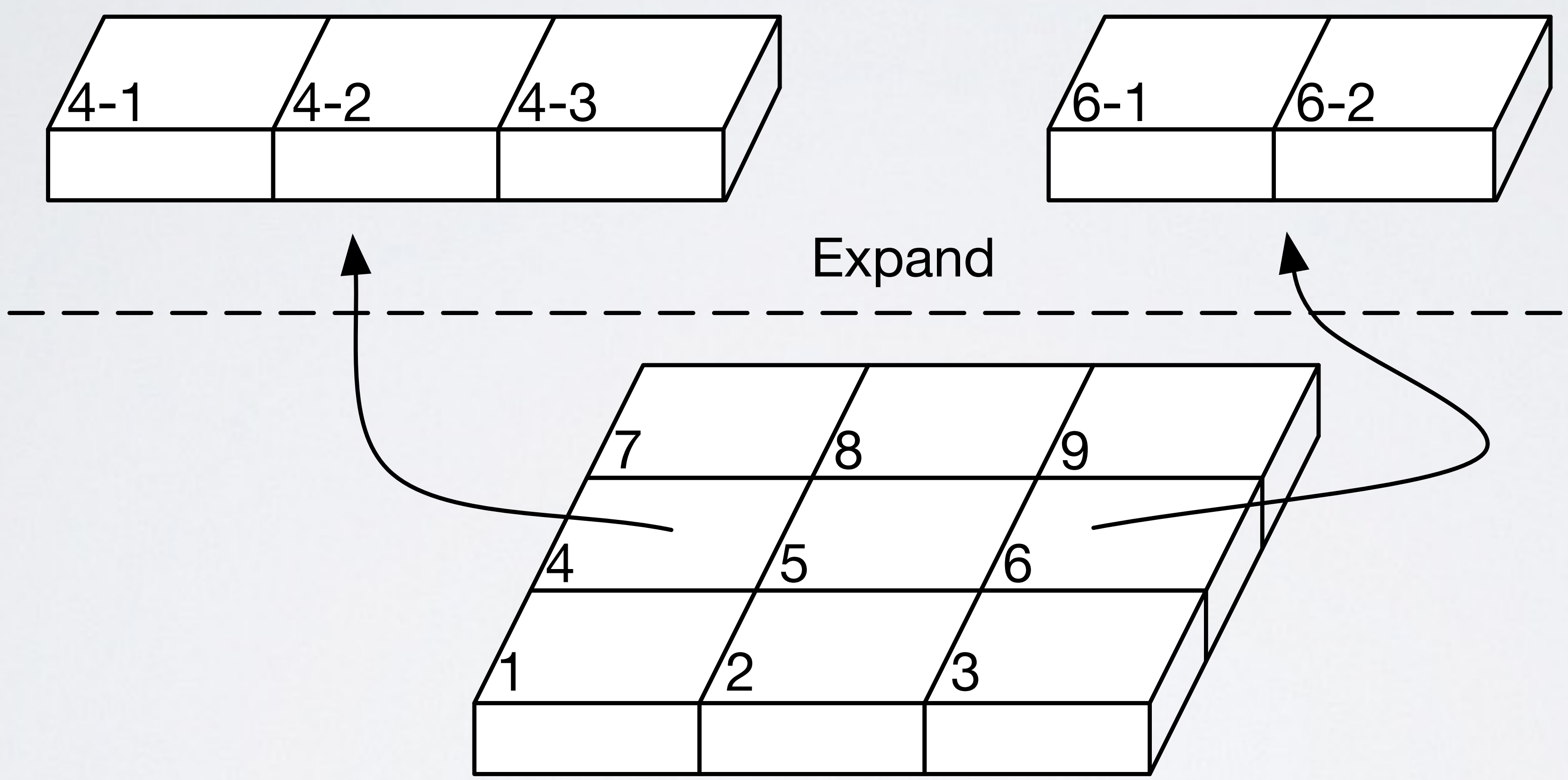
[De Hauwere et al. (2011)a] De Hauwere, Y.-M., Vrancx, P. & Nowé, A. (2011a).
Adaptive state representations for multi-agent reinforcement learn- ing. In: *Proceedings of the 3th International Conference on Agents and Artificial Intelligence*. Rome, Italy.

PROBLEM SETTING

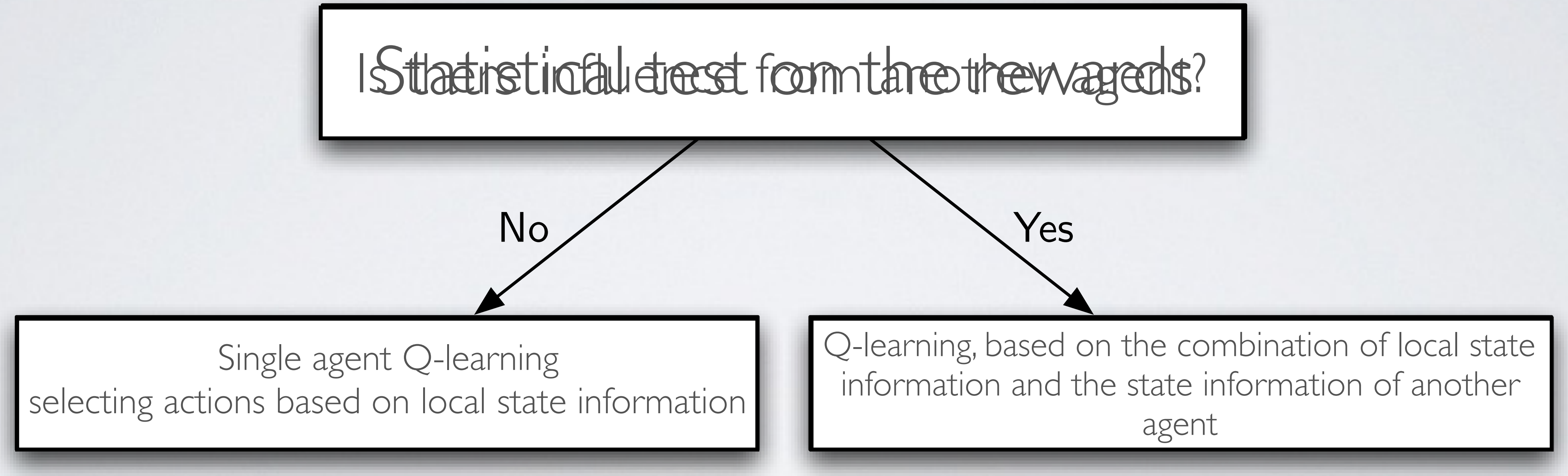


- Agents only interact where their policies interfere
- Locally adapt policy

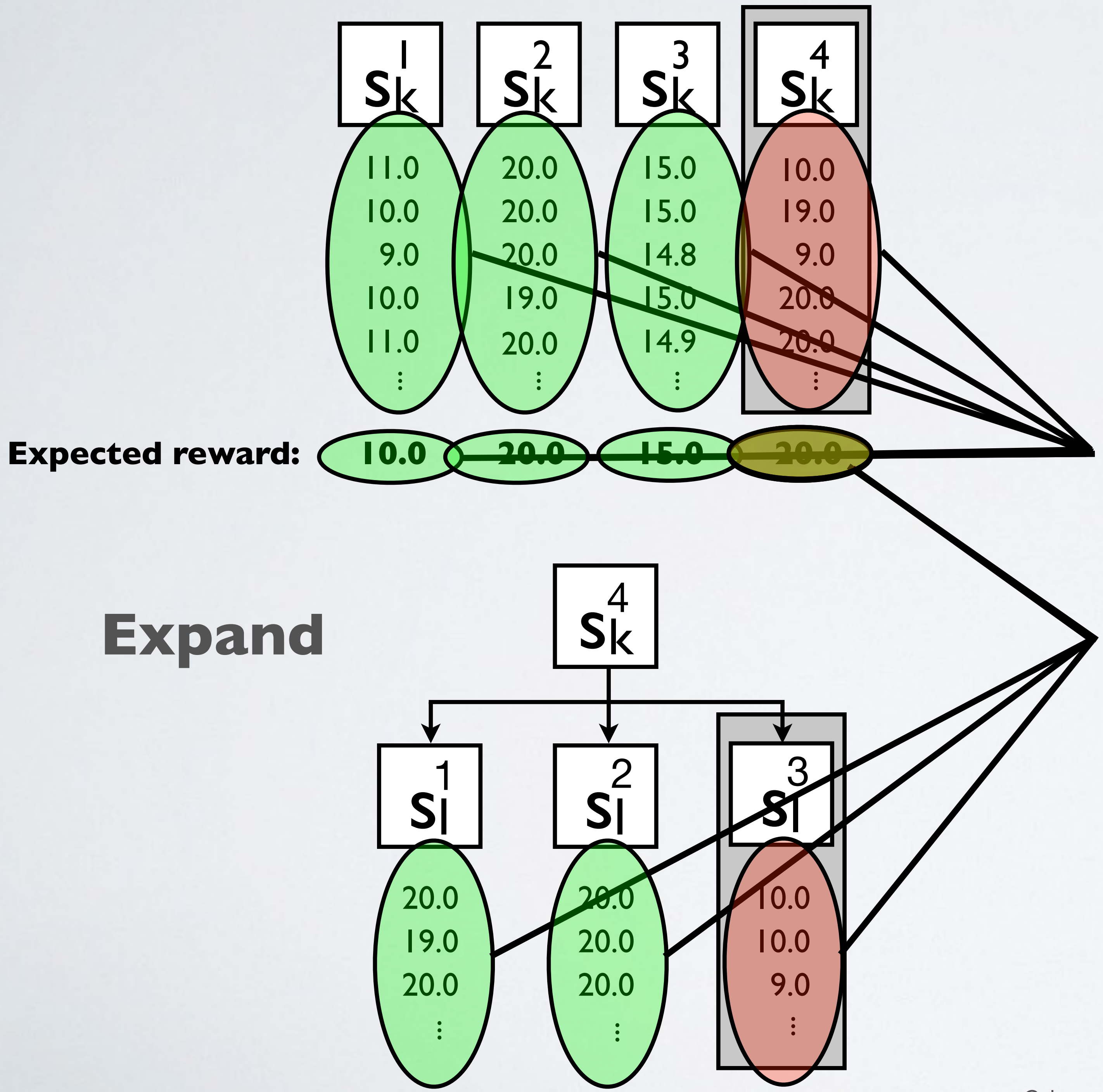
REPRESENTATION IDEA



SOLUTION METHOD: CQ-LEARNING



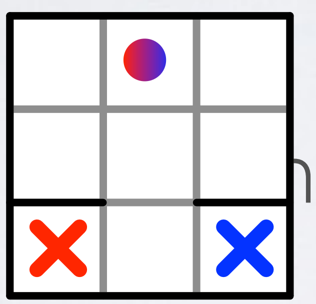
CQ-LEARNING : STATISTICAL TESTS



- Agents have been learning alone in the environment
- Agent k acts independently using only local state information (s_k) in a multi-agent environment
- Perform statistical test against baseline
- Samples its rewards, based on the state information of other agents & performs the same test

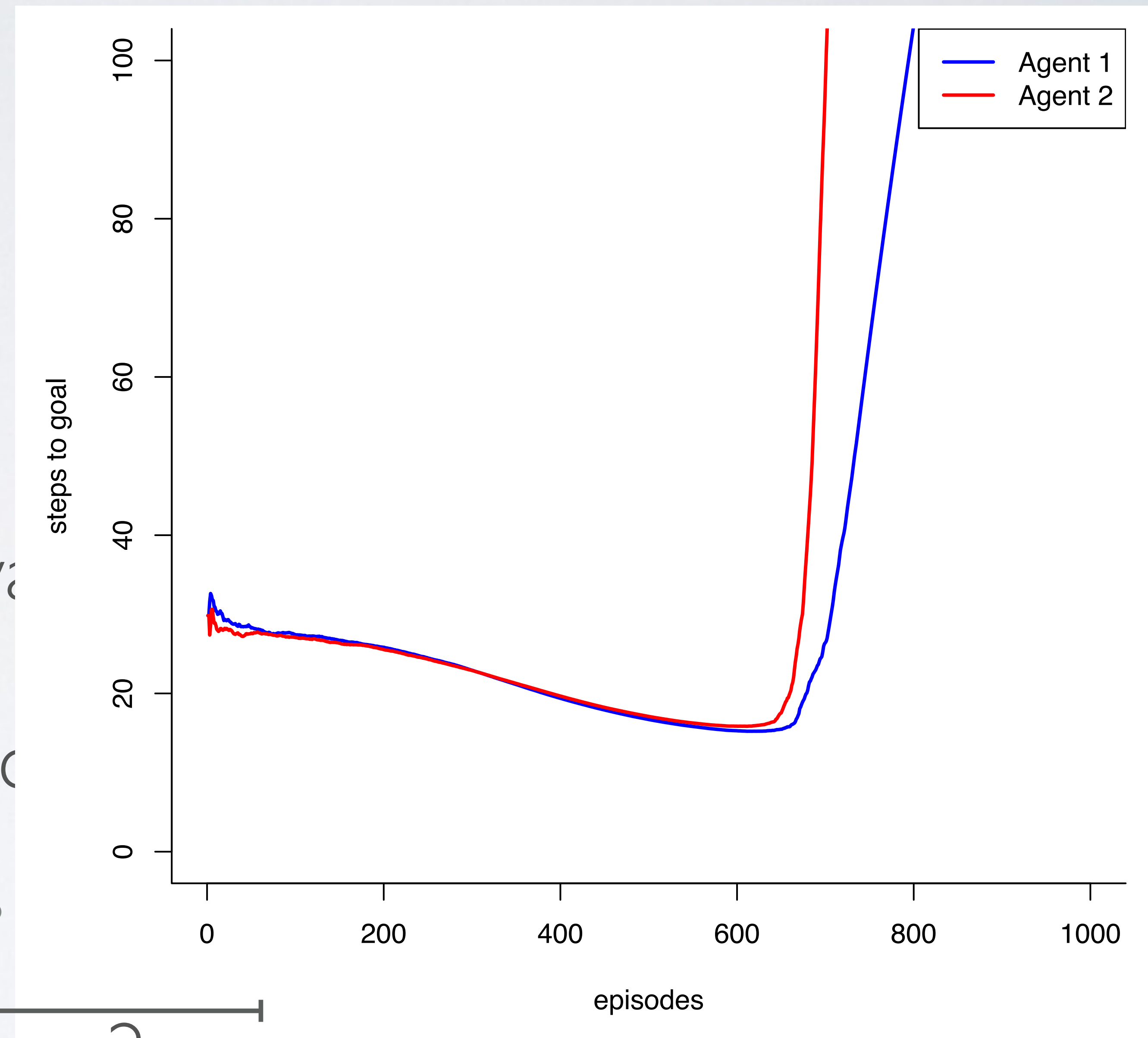
$$s_k^4 \Rightarrow \langle s_k^4 , s_l^3 \rangle$$

CQ-LEARNING   BASELINE FOR STATISTICAL TESTS



itial rewa

for a parti



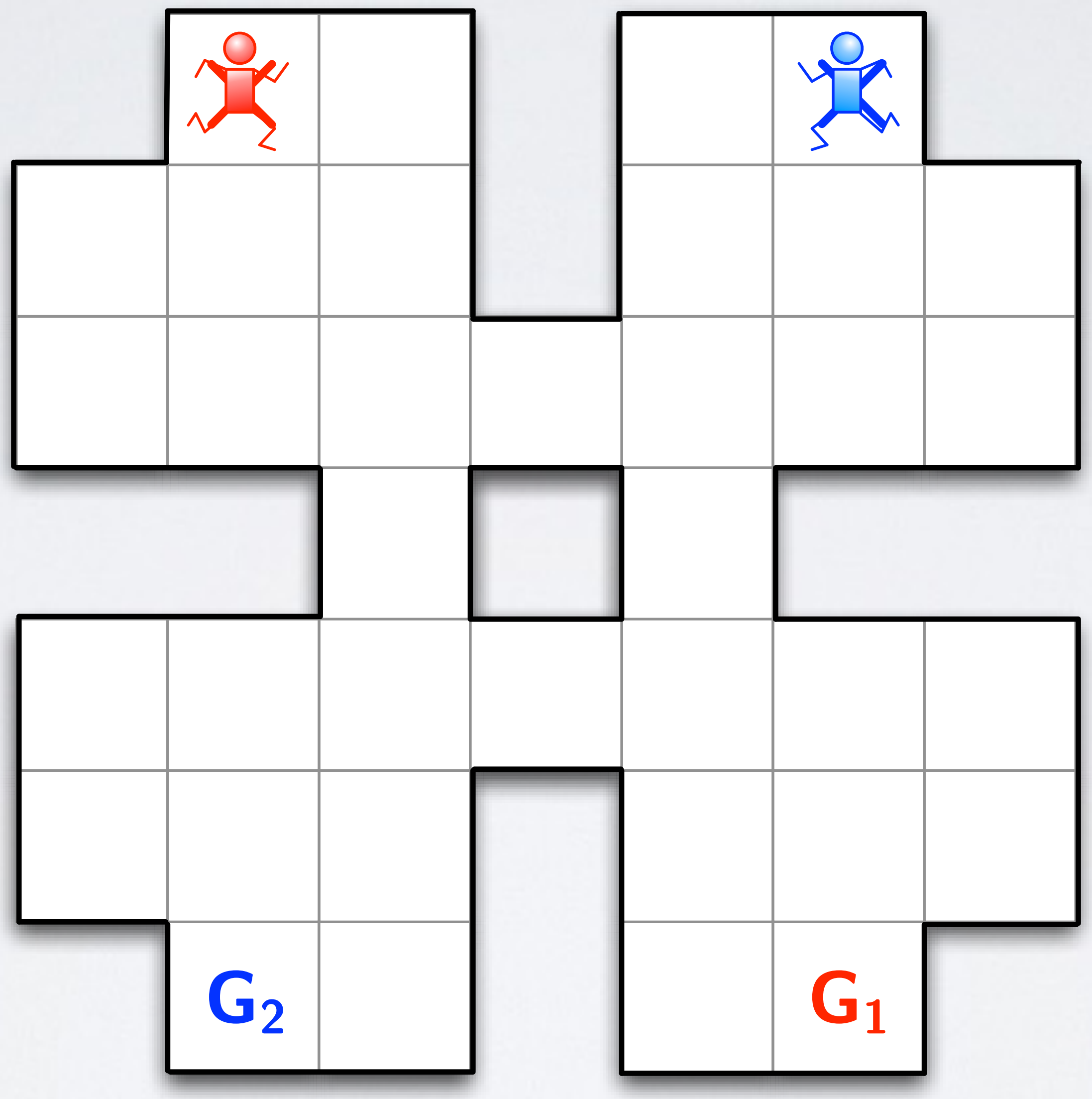
W_k^1

W_k^2

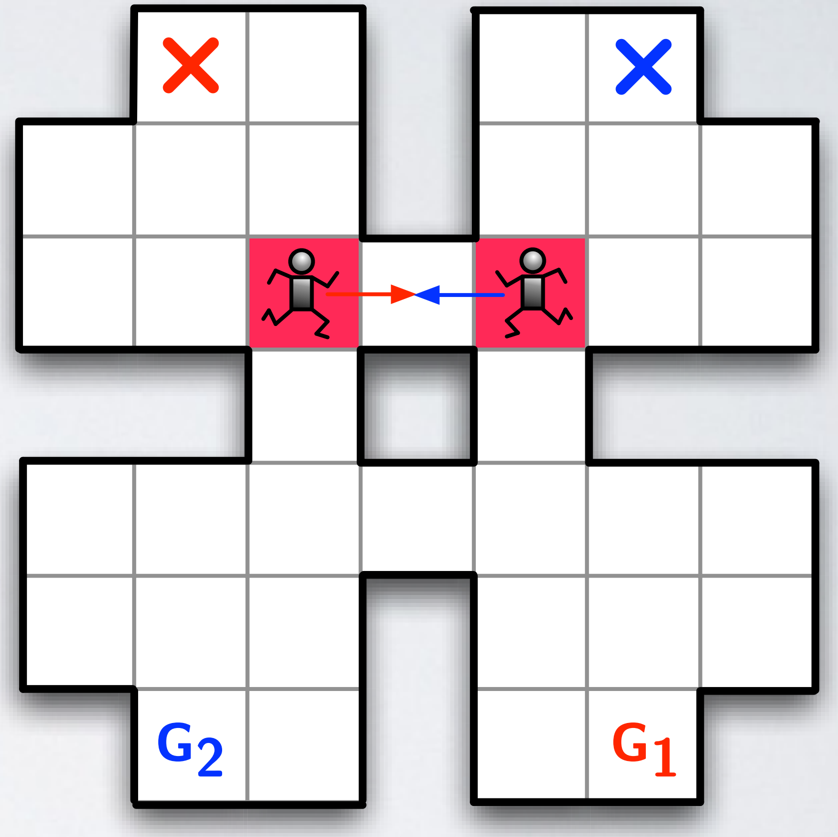
Compare W_k^1 against W_k^2

CQ-LEARNING ILLUSTRATION

- Sample run

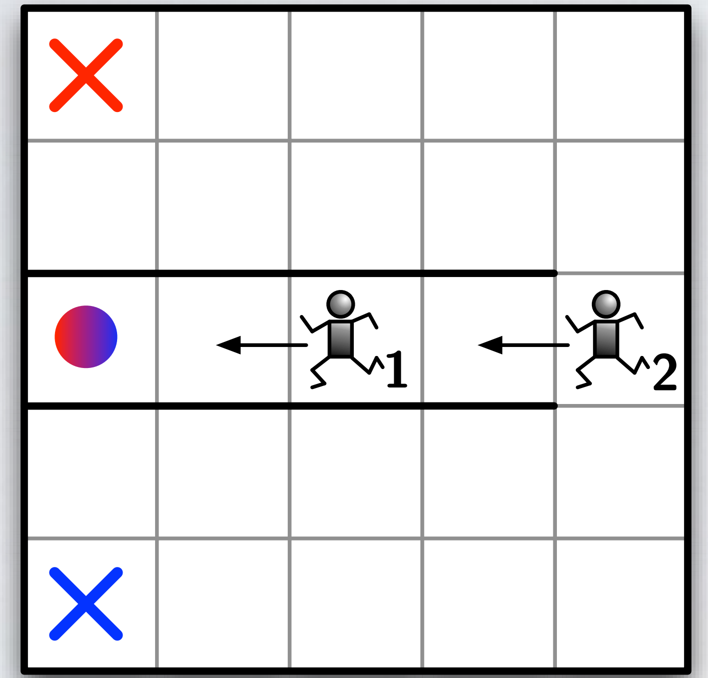


FCQ-Learning

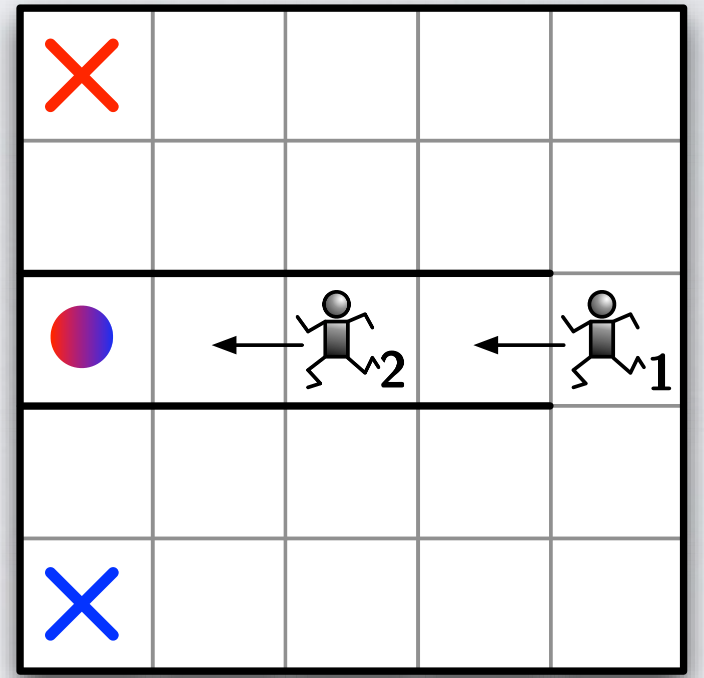


[De Hauwere et al. (2011)b] De Hauwere, Y.-M., Vrancx, P. & Nowé, A. (2011b).
Detecting and solving future multi-agent interactions. In: *Pro- ceedings of the AAMAS Workshop on Adaptive and Learning Agents*. Taipei, Taiwan.

PROBLEM SETTING



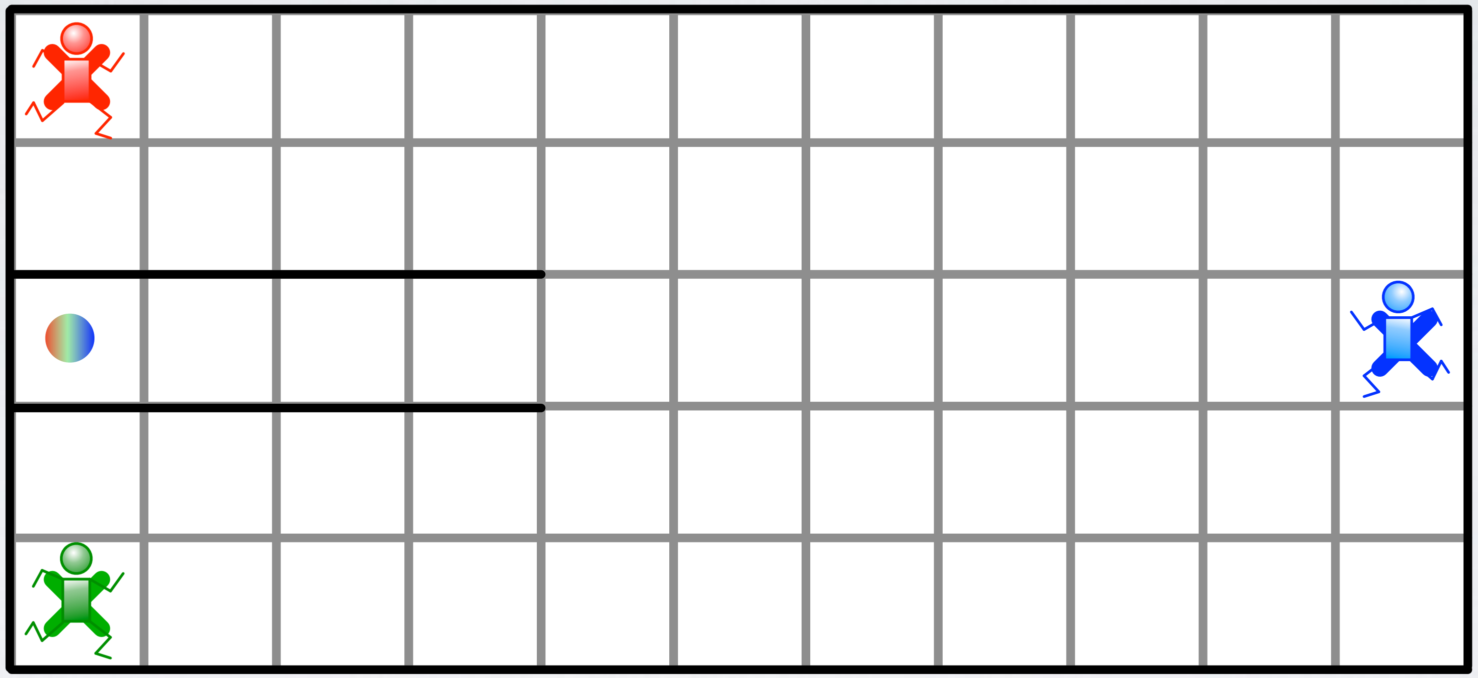
Reward: +20



Reward: +10

- Reflected in immediate reward signal
- Too late to solve the problem

EXPERIMENTAL RESULTS



- Order to reach the goal:
- Red Agent +20
- Blue Agent +20
- Green Agent +20

FURTHER READING

Sparse Interactions in Multi-agent RL

Yann-Michael De Hauwere

available at http://ai.vub.ac.be/publications/904