CS344M
Autonomous Multiagent Systems

Patrick MacAlpine

Department of Computer Science
The University of Texas at Austin

Good Afternoon, Colleagues

Are there any questions?

Good Afternoon, Colleagues

Are there any questions?

- How is SMDP different from MDP?
- Advantages of tile coding vs other approaches?
- What about SPAR (Strategic Position by Attraction and Repulsion)?

Logistics

- Progress reports due at beginning of class today

Logistics

- Progress reports due at beginning of class today
- Progress report peer review due next Thursday – reports to review will be sent out shortly

Logistics

- Progress reports due at beginning of class today
- Progress report peer review due next Thursday – reports to review will be sent out shortly
- Prize for winning class tournament

Logistics

- Progress reports due at beginning of class today
- Progress report peer review due next Thursday – reports to review will be sent out shortly
- Prize for winning class tournament
- 10+ students went to Undergraduate Writing Center :)

Logistics

- Progress reports due at beginning of class today
- Progress report peer review due next Thursday – reports to review will be sent out shortly
- Prize for winning class tournament
- 10+ students went to Undergraduate Writing Center :)

Reinforcement Learning

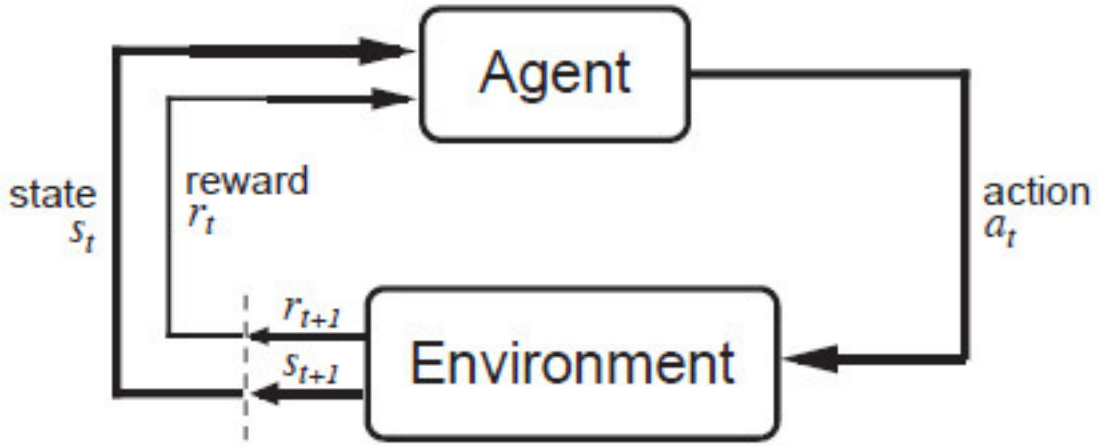

Image from wikipedia

Reinforcement Learning

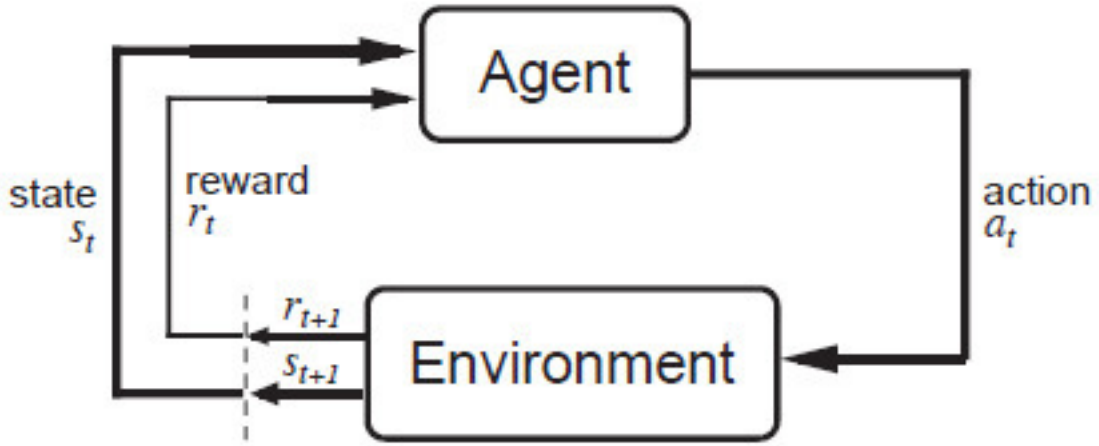

Image from wikipedia

Markov Decision Process (MDP)

Reinforcement Learning

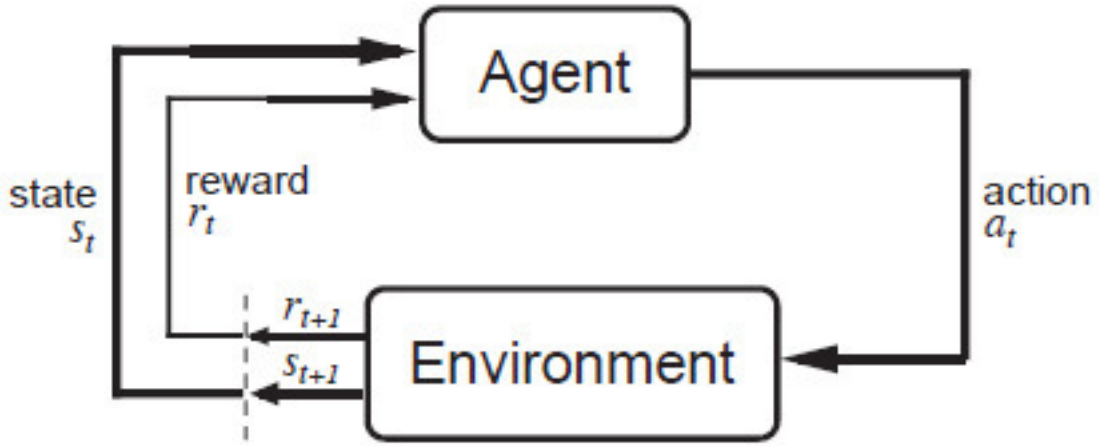

Image from wikipedia

Markov Decision Process (MDP)

Important questions:

- What is your state space?

Reinforcement Learning

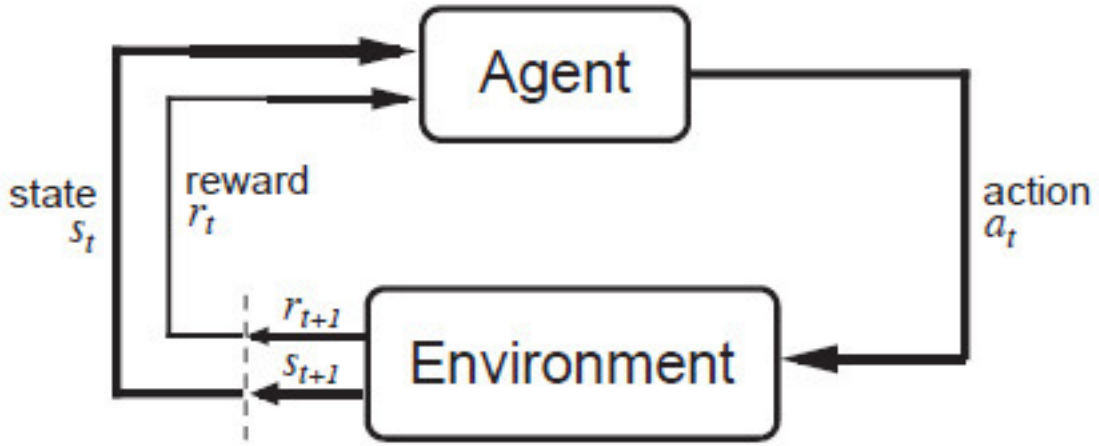

Image from wikipedia

Markov Decision Process (MDP)

Important questions:

- What is your state space?
- What is your action space?

Reinforcement Learning

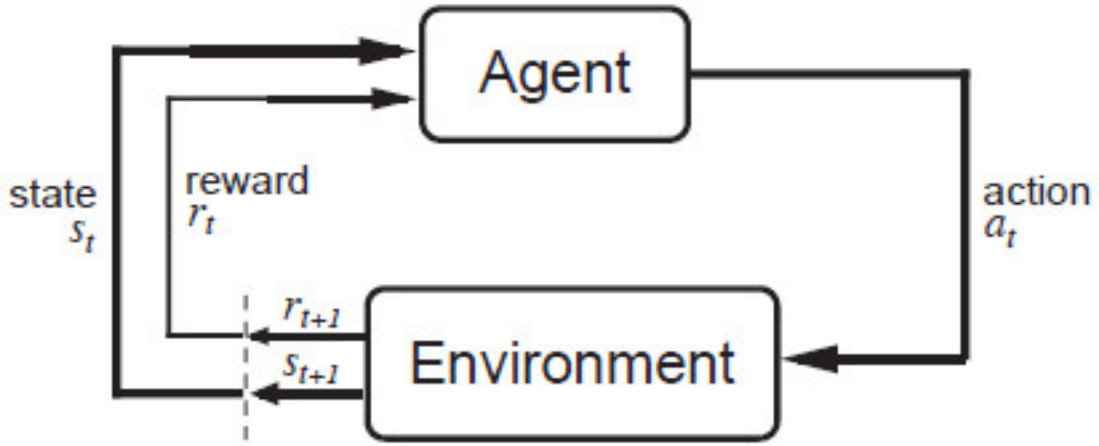

Image from wikipedia

Markov Decision Process (MDP)

Important questions:

- What is your state space?
- What is your action space?
- What is your reward function?

SARSA (s_t,a_t,r_t,s_{t+1},a_{t+1})

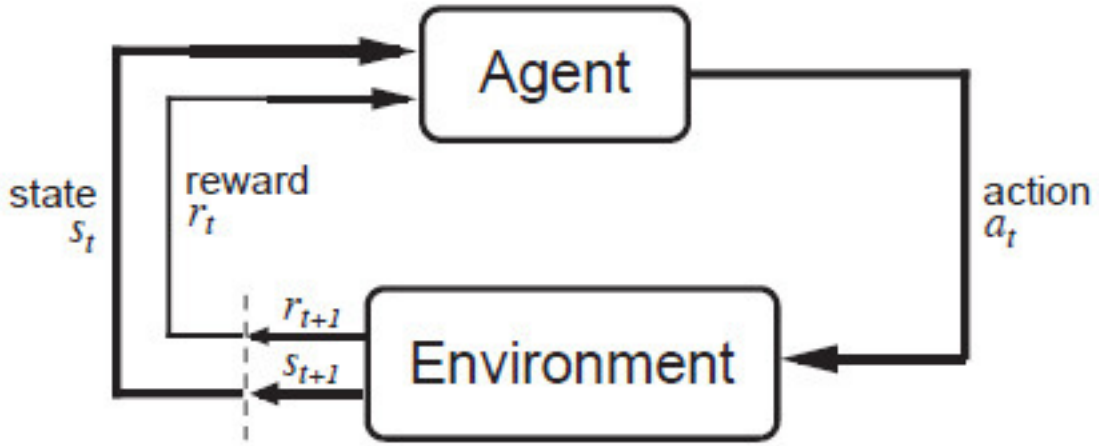

Image from wikipedia

SARSA (s_t,a_t,r_t,s_{t+1},a_{t+1})

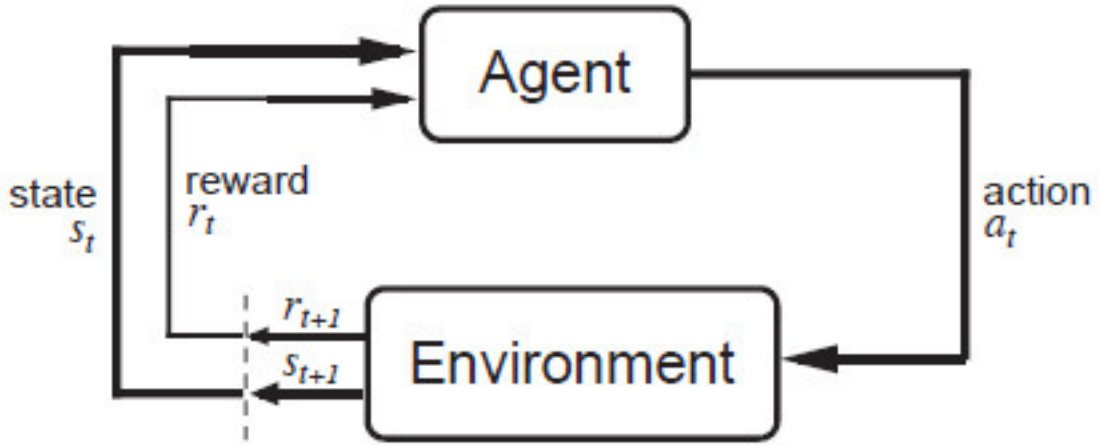

Image from wikipedia

Learn Q table (value function) for state - action pairs

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)]$$

Keepaway

- Keepaway videos

Keepaway

- Keepaway videos
- Slides

Keepaway Discussion

- Could you use learned policies for full soccer game?

Keepaway Discussion

- Could you use learned policies for full soccer game?
- Could we apply competitve co-evolution?

Keepaway Discussion

- Could you use learned policies for full soccer game?
- Could we apply competitve co-evolution?
- Other sub-tasks in soccer that might be learnable?

Keepaway Discussion

- Could you use learned policies for full soccer game?
- Could we apply competitve co-evolution?
- Other sub-tasks in soccer that might be learnable?

Half Field Offense

<Slides>

Policy Search vs Value Function Based RL

	Policy Search	**Value Function Based**
Learn	Policy parameters	Value function
Good For	Tuning parameter values	Learning discrete actions
Evaluation	Fitness function	Reward function
Algorithms	CMA-ES, genetic algorithms, etc.	SARSA, Q-learning, etc.