Coordination Problems in Cooperative Multi-agent Reinforcement Learning

Choi, Young Cheol

2009. 4. 8

Contents

- Multi-agent systems(MAS)
- Cooperative MAS
- Coordination Problems

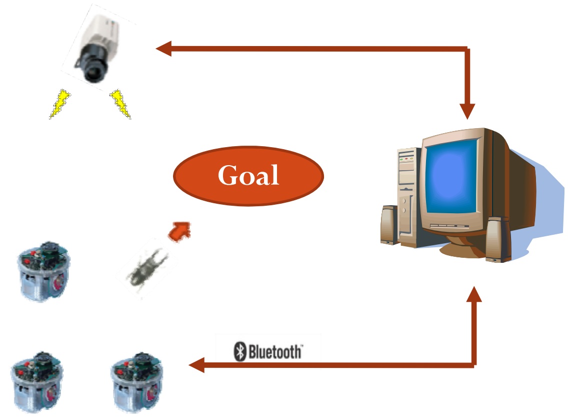
Multi-agent Systems

- Why Multi-agent Systems?
  - Better performance then an individual agent
    - In partially observable environments
    - Distributed control
  - More faster and cheaper

- Multi-agent reinforcement learning
  - Agents learn to map from their states to their actions
  - They learn by rewards or payoffs obtained through interacting with their environment
  - State dynamics are unknown
  - Iterative process

Classification of R.L

- Reinforcement Learning
  - Single-agent
  - Multi-agent
    - Competitive
    - Cooperative
      - Heterogeneous
      - Homogeneous



Cooperative MAS

- All agents can together achieve ideal optimality if and only if they are rational by following a stationary joint deterministic policy

  - Strongly Cooperative MAS : If at least one agent is not rational, then non of the agents in the system can achieve his ideal optimality.

  - Weakly Cooperative MAS : There exists at least one joint policy by which some agents can achieve their ideal optimality while the others can not.

- The learning goal is to maximize the common discounted return.

Cooperative MAS Algorithms

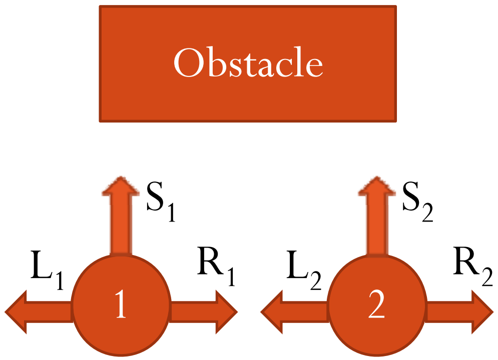
- Team – Q
- Distributed-Q
- Optimal Adaptive Learning(OAL)
- Joint Action Learners(JAL)
- Frequency Maximum Q-value(FMQ)

Resource restriction

- **Communication**: Bandwidth, distance, and other limitations provide limited ability for agents to communicate
- **Bounded reasoning**: Complex domains require abstract representations and anytime reasoning techniques for reaching reasonable conclusions (locally and globally) in near real time
- **Incomplete knowledge**: Agents need models of what is known, not known, or known with limited confidence; active learning techniques for gathering the most critical information (for the team)

Coordination Problems in Cooperative MAS

- In Cooperative MAS, coordination problem arises even if all the agents use the same algorithm.
- Example



Q	L_2	S_2	R_2
L_1	10	-5	0
S_1	-5	-10	-5
R_1	-10	-5	10

Coordination Problems in Cooperative MAS

- Coordination-Free Model

  - Team Q-learning algorithm avoids the coordination problem by assuming that the optimal joint actions are unique

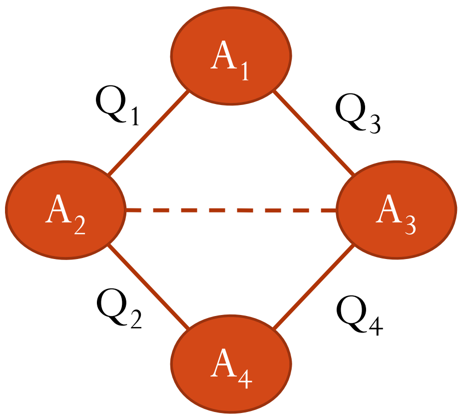
  - Distributed Q-learning algorithm solves the cooperative task without assuming coordination and with limited computation. However, the algorithm only works in deterministic problems.

Coordination Problems in Cooperative MAS

- Coordination-Based Methods

  - Coordination graphs simplify coordination when the global Q-function can be additively decomposed into local Q-functions that only depend on the action of a subset of agents

$$Q = Q_1(a_1,a_2)+Q_2(a_2,a_4)+Q_3(a_1,a_3)+Q_4(a_3,a_4)$$



Node : Agent's action choice

Edge : Q-functions

Coordination Problems in Cooperative MAS

- Indirect Coordination Methods

  - Indirect coordination methods bias action selection toward actions that are likely to result in good rewards or returns.

  - Static tasks

    · Joint Action Learners(JAL) : Employ empirically learned models of the other agents' behavior

    · Frequency Maximum Q-value(FMQ) : heuristic is based on the frequency with which actions yielded good values in the past

  - Dynamic tasks

    · Optimal Adaptive Learning(OAL) : the bias is towards recently chosen Nash equilibria.