强化学习原理及应用
Reinforcement Learning (RL): Theories & Applications

DCS6289 Spring 2022

Yucong Zhang （张宇聪）

School of Computer Science and Engineering
Sun Yat-Sen University

Lecture 9： Multi-Agent RL

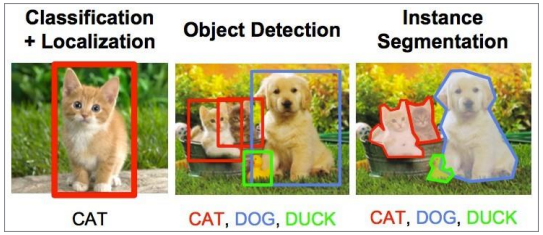
10th May. 2022

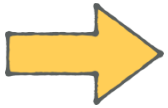
Multi-Agent RL

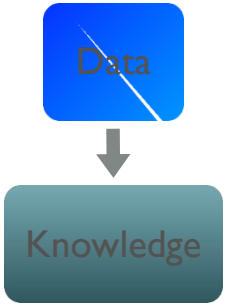
- What is the multi-agent reinforcement learning(MARL)?
- What is the difference between single-agent reinforcement learning and multi-agent reinforcement learning?
- What is the difficulty in the multi-agent reinforcement learning?
- What are the categories of multi-agent reinforcement learning?

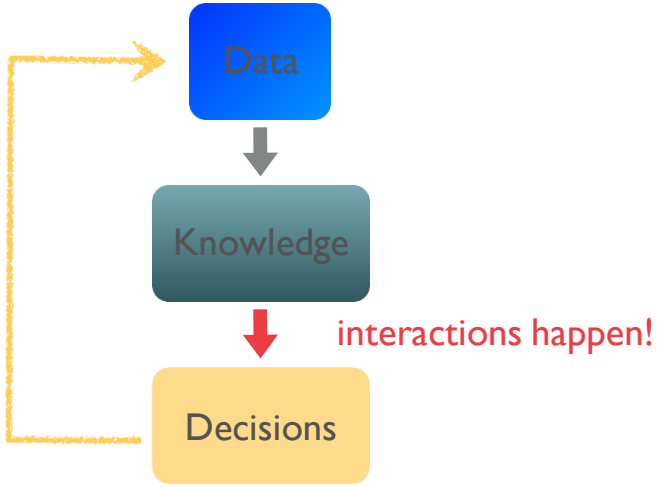
Multi-Agent RL

- What is the multi-agent reinforcement learning(MARL)?
  - Reinforcement learning turns data/knowledge into closed-loop decision making.
  - Multi-agent learning deal with interactions among the learning agents.









Multi-Agent RL

- What is the multi-agent reinforcement learning(MARL)?
  - MARL addresses the sequential decision-making problem of multiple autonomous agents that operate in a common environment, each of which aims to optimize its own long-term return by interacting with the environment and other agents
  - A group of agents work together to optimize team performance
  - Multiagent systems include a set of autonomous entities(agents) that share a common environment and where each agent can independently perceive environment, act acting to its individual objectives and as a consequence, modify the environment
  - In an multiagent system, agents must compete or cooperate to obtain the best overall results.

Multi-Agent RL

- What is the multi-agent reinforcement learning(MARL)?

Robots

Drone Delivery

Games

Autonomous Vehicles

Smart Grids

MALib

Multi-Agent RL

- What is the multi-agent reinforcement learning(MARL)?
- What is the difference between single-agent reinforcement learning and multi-agent reinforcement learning?
- What is the difficulty in the multi-agent reinforcement learning?
- What are the categories of multi-agent reinforcement learning?

Multi-Agent RL

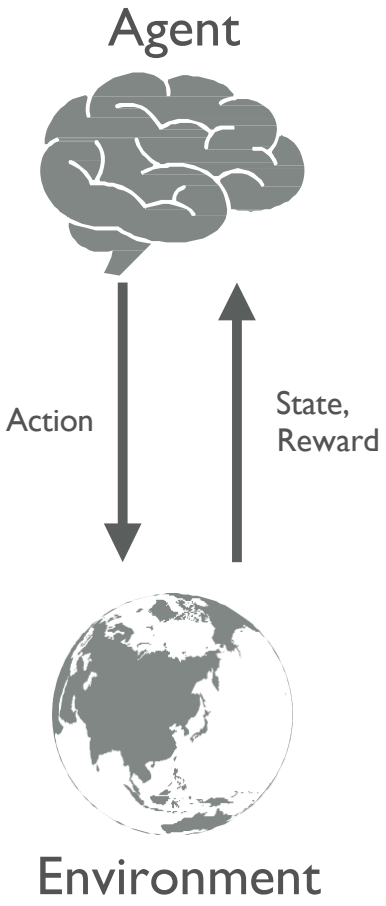
- What is the difference between single-agent reinforcement learning and multi-agent reinforcement learning?
  - Single-agent RL:
    - Only one agent
    - State、local action、single reward
  - Multi-agent RL:
    - At least two agents
    - Local observation、joint action、team reward
    - Agents communicate with each other and interact with environment at the same time.

Multi-Agent RL

- What is the difference between single-agent reinforcement learning and multi-agent reinforcement learning?
  - Problem Formulation: single-agent RL:

    Markov Decision Process(MDP) $(S, A, R, T, P_0, \gamma)$

    - S denotes the state space
    - A is the action space
    - $R = R(s, a)$ is the reward function
    - $T: S \times A \times S \rightarrow [0,1]$ is the state transition function
    - P_0 is the distribution of the initial state
    - γ is a discount factor
    - Goal: find the optimal policy that maximizes expected reward

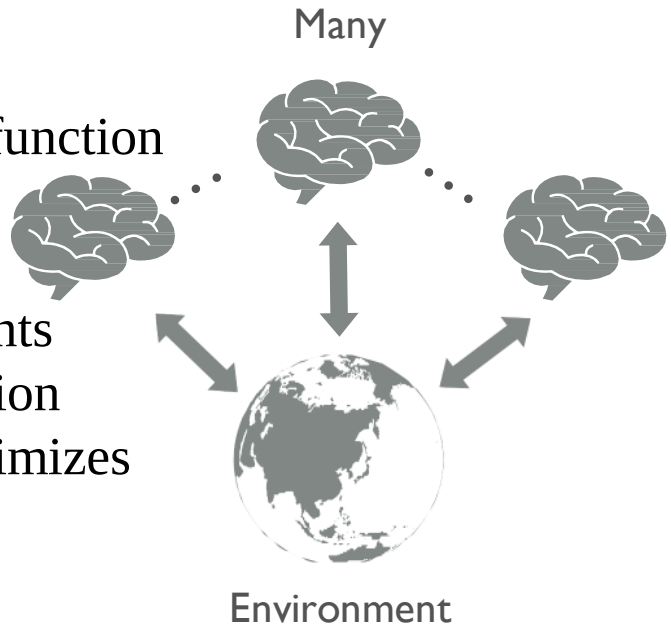


Multi-Agent RL

- What is the difference between single-agent reinforcement learning and multi-agent reinforcement learning?
  - Problem Formulation: multi-agent RL:

    Partially Observable Markov Decision Process(POMDP)$(S, A, R, T, P_0, Z, O, n, \gamma)$

    - n agents in the environment
    - S denotes the state space
    - A is the joint action space $A^1 \times \cdots \times A^n$
    - $R = R(S, A)$ is the share reward function
    - $T: S \times A \times S \to [0,1]$ is the state transition function
    - P_0 is the distribution of the initial state
    - γ is a discount factor
    - Z is the individual observation for each agents
    - $O(s, a): S \times A \to Z$ is the observation function
    - Goal: find the optimal joint policy that maximizes expected team reward



Multi-Agent RL

- What is the multi-agent reinforcement learning(MARL)?
- What is the difference between single-agent reinforcement learning and multi-agent reinforcement learning?
- What is the difficulty in the multi-agent reinforcement learning?
- What are the categories of multi-agent reinforcement learning?

Multi-Agent RL

- What is the difficulty in the multi-agent reinforcement learning?
  - Non-stationarity:
    - an agent observes not only the outcomes of its own action but also the behavior of other agents
    - Learning among the agents is complex because all agents potentially interact with each other and learn concurrently
  - Partial observability:
    - The agents only capture partial information about the environment before making decision
  - Dimension catastrophic:
    - Joint action space and Joint state space
    - Large-scale multi-agent decision-making、
  - Credit assignment:
    - Lazy agent
  - Sample efficiency、Exploration and Exploitation、complex mixed environment, etc.

Multi-Agent RL

- What is the difficulty in the multi-agent reinforcement learning?

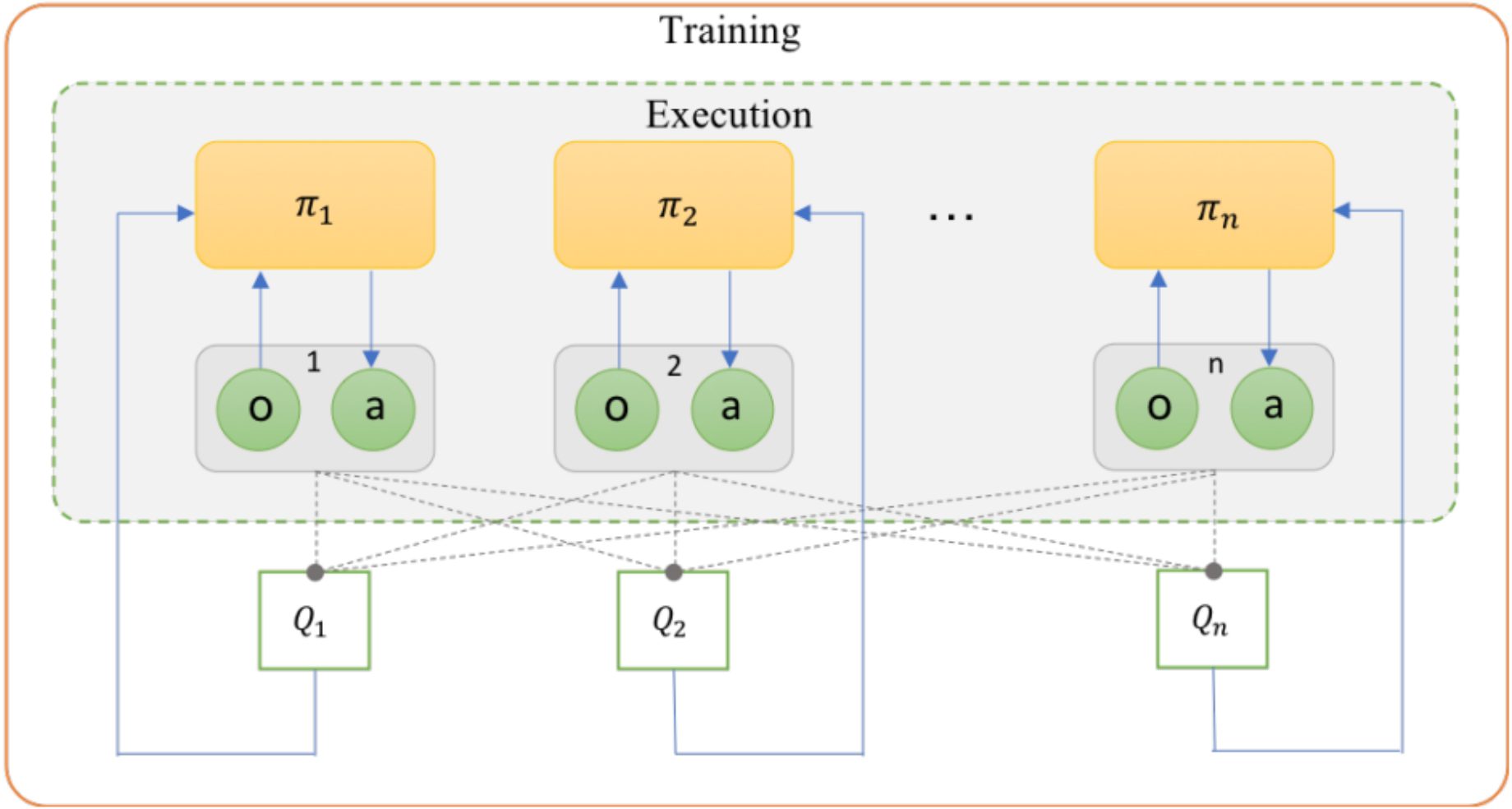
  Solution

  - Non-stationarity:
    - Centralized Train and Decentralized execution(CTDE), e.g. MADDPG
    - Communication, e.g. CommNet
  - Partial observability:
    - RNN、GRU、LSTM, e.g. DRQN
  - Dimension catastrophic:
    - CTDE, e.g. VDN、QMIX
    - Mean-Field, e.g. MFAC
  - Credit assignment:
    - Counterfactual mechanisms, e.g. COMA
  - Exploration and Exploitation:
    - Reward shaping(intrinsic reward、novelty)
    - UCB
    - Influence(mutual information)

Multi-Agent RL

- What is the difficulty in the multi-agent reinforcement learning?

  Centralized Train and Decentralized execution(CTDE)



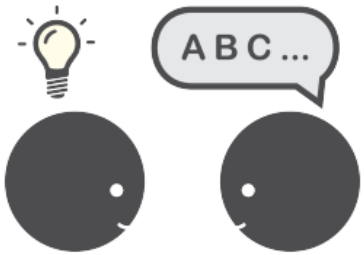
Multi-Agent RL

- What is the multi-agent reinforcement learning(MARL)?
- What is the difference between single-agent reinforcement learning and multi-agent reinforcement learning?
- What is the difficulty in the multi-agent reinforcement learning?
- What are the categories of multi-agent reinforcement learning?

Multi-Agent RL

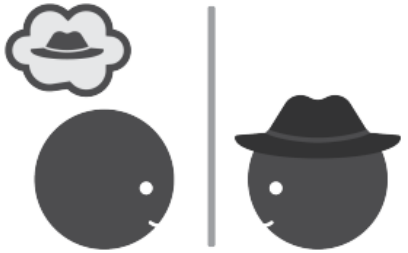
- What are the categories of multi-agent reinforcement learning?
  - Analysis of emergent learning:
    - Simply using Single-agent RL algorithm in multi-agent scenarios
  - Learning communication:
    - Learning communication protocols among agents
  - Learning cooperation:
    - Learning to cooperate using only actions and local observation
  - Agents modeling agents:
    - Reasoning about others

(a) Analysis of emergent behaviors



(b) Learning communication

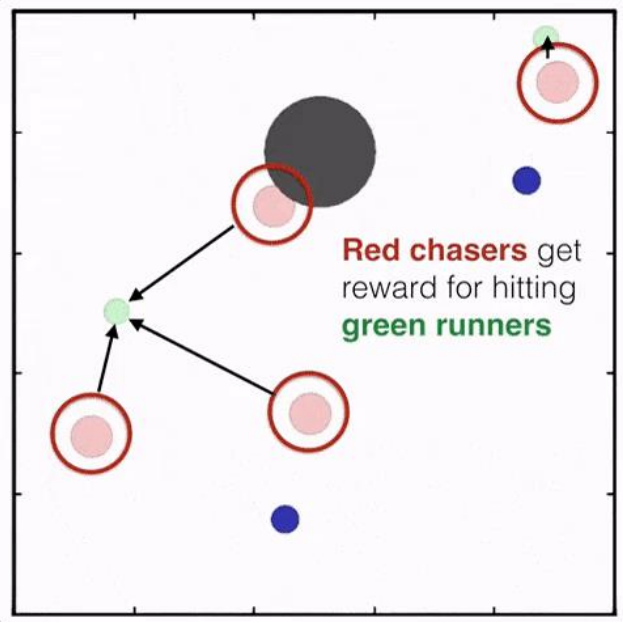
(c) Learning cooperation

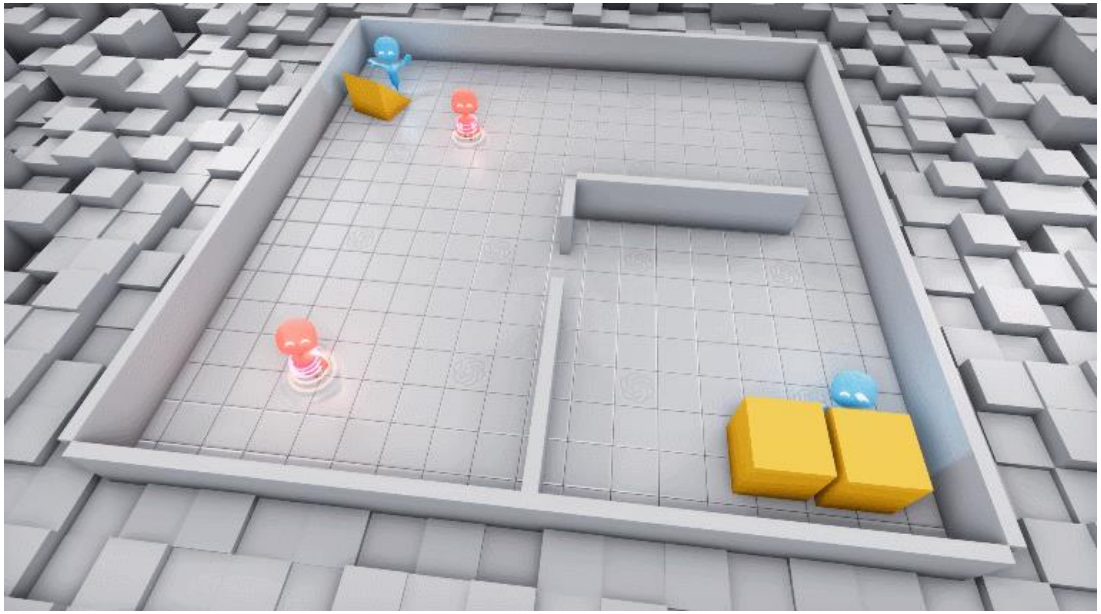


(d) Agents modeling agents

Multi-Agent RL

- What are the types of multi-agent reinforcement learning?
  - Analysis of emergent learning:
    - Three major settings: cooperative, competitive, mixed scenarios
      - Cooperative： working together and coordinating their actions、maximizing a shared team reward
      - Competitive： self-interested(maximizing an individual reward)、opposite rewards、zero-sum games
      - Mixed scenarios： general-sum games





Multi-Agent RL

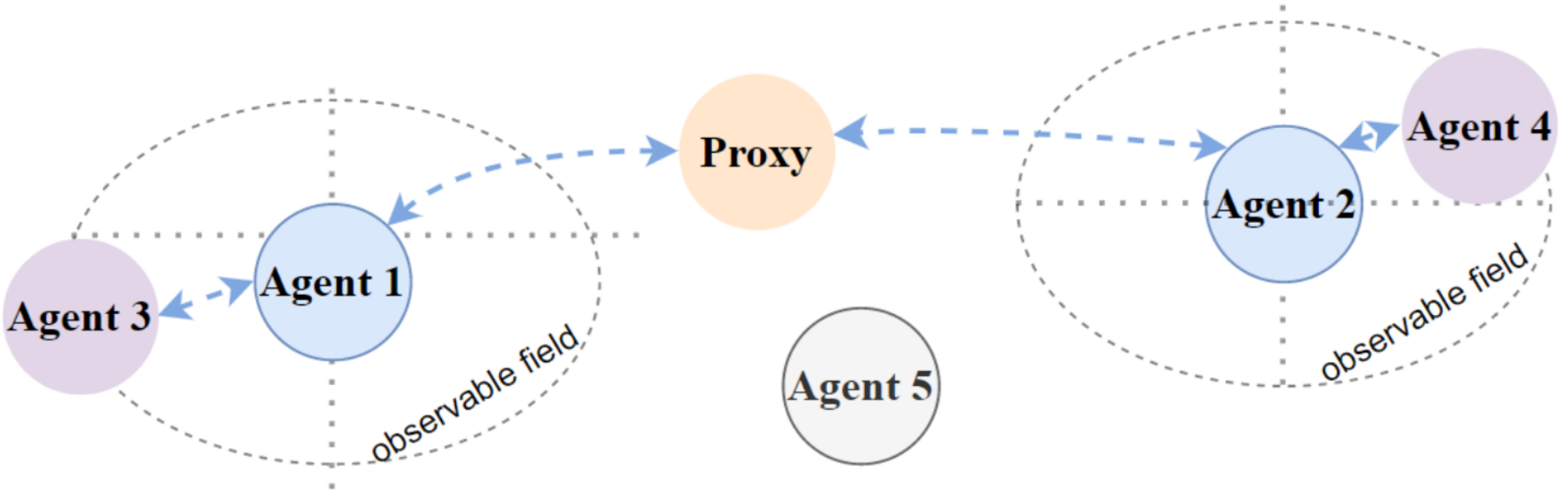
- What are the types of multi-agent reinforcement learning?
  - Analysis of emergent learning:
    - These works, in general, do not propose learning algorithms—their main focus is to analyze and evaluate DRL algorithms, e.g., DQN, PPO and others, in a multiagent environment

Summary	Learning	Setting
Train DQN agents to play Pong.	VB	CO&CMP
Train DQN agents to play sequential social dilemmas.	VB	Mixed
Propose DRL agents able to cooperate in social dilemmas.	VB	Mixed
Propose Malthusian reinforcement learning which extends self-play to population dynamics.	VB	Mixed
Train PPO agents in competitive MuJoCo scenarios.	PG	CMP
Train PPO, A3C, and DQN agents in attacker-defender games.	VB, PG	CMP
Train agents represented with NN to learn a communication language.	PG	CO
Learn communication with an end-to-end differentiable model to train with backpropagation.	PG	CO

Multi-Agent RL

What are the types of multi-agent reinforcement learning?

- Learning communication:
  - These works explore a sub-area in which agents can share information with communication protocols， for example through direct message or via a shared memory
  - This setting usually considers a set of cooperative agents in partially observable environment



Multi-Agent RL

- What are the types of multi-agent reinforcement learning?
  - Learning communication—Dimensions and Targeted Problems

Dimensions	Targeted Problems
Communication Type	Which type of agents to communication with?
Communication Policy	When and how to build communication links among agent?
Communicated Messages	Which piece of information to share?
Message Combination	How to combine received messages?
Inner Integration	How to integrate combined message into learning models?
Communication Constraints	How to fulfill realistic requirements?
Communication learning	How to train and improve communication?
Training Scheme	How to utilize collected experience from agents?
Controlled Goals	What kind of behaviours are desired to emerge with communication?

Multi-Agent RL

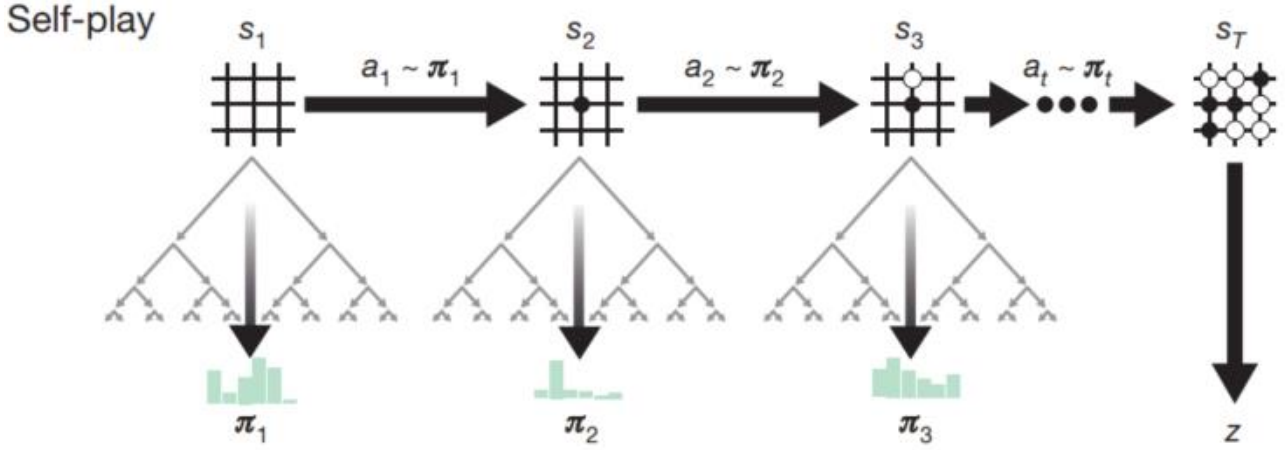
- What are the types of multi-agent reinforcement learning?
  - Learning communication—Communicatee Type

Types	Subtypes	Methods
Agents in the MAS: • communicatees is composed of agents • agent directly communicate with each other	Nearby Agents: Communication is only allowed between neighbors	DGN; MAGNet-SA-GS; Agent-Entity Graph; LSC; NeurComm; FlowComm; GAXNet
	Other Agents: Nearby agents are not identified, the set of communicatees is simply composed of other agents	DIAL; RIAL; CommNet; BiCNet; TarMAC; MADDPG-M; IC3Net; DCC-MD; VBC; ETCNet; TMC
Proxy: A proxy is a visual agent who plays an essential role in communication, but does not have direct effect on the environment		MS-MARL-GCM; ATOC; MD-MADDPG; IMAC; GA-Comm; Gated-ACML; HAMMER; MAGIC

Multi-Agent RL

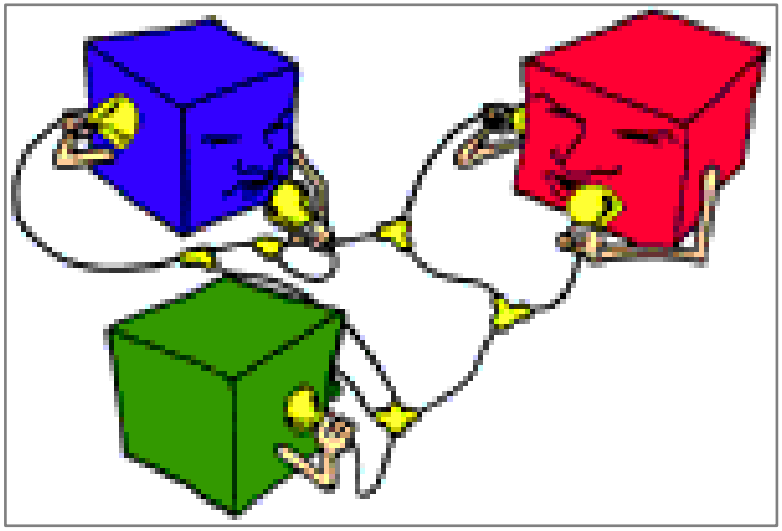
What are the types of multi-agent reinforcement learning?

- Agents modeling agents :
  - An importance ability for agents to have is to reason about the behaviors of other agents by constructing models that make predictions about the modeled agents
  - Opponent model: utilizing network to learn a representation of the opponent's policy or q-value, e.g., DRON; SOM
  - Self-play: agents learn to improve their performance by playing "against themselves", e.g., NFSP; PSRO
  - Minimax: minimizing the worst case scenario (maximum loss), e.g., M3DDPG



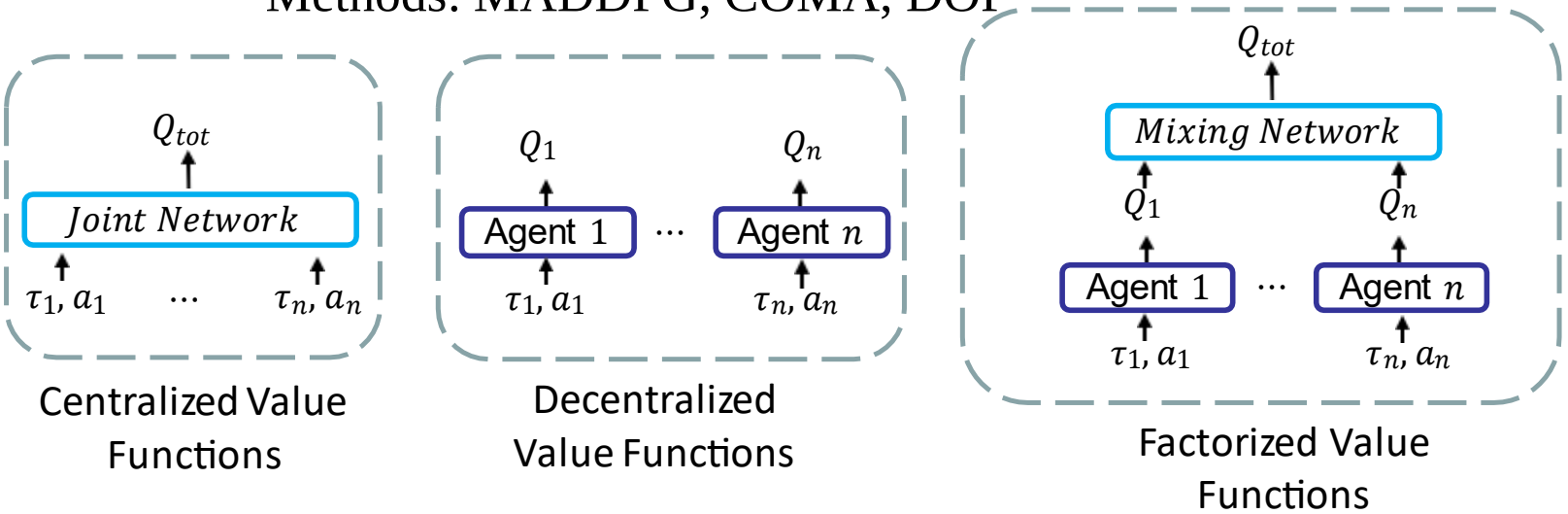
Multi-Agent RL

- What are the types of multi-agent reinforcement learning?
  - Learning cooperation :
    - Open, dynamic, persistent systems
    - Decentralized control
    - Large scale
    - Partial observability
    - No real-time global reward signal
    - Communication delay



Multi-Agent RL

- What are the types of multi-agent reinforcement learning?
  - Learning cooperation :
    - Value-Based Methods
      - Paradigm: CTDE
      - Methods: VDN, QMIX, QPLEX
    - Policy Gradient Methods
      - Paradigm: Centralized Critic and Decentralized Actors
      - Methods: MADDPG, COMA, DOP



Centralized Value Functions

Decentralized Value Functions

Factorized Value Functions

Scalability

Non-stationarity
Credit assignment

Centralized training
Decentralized execution

Multi-Agent RL

- Summary