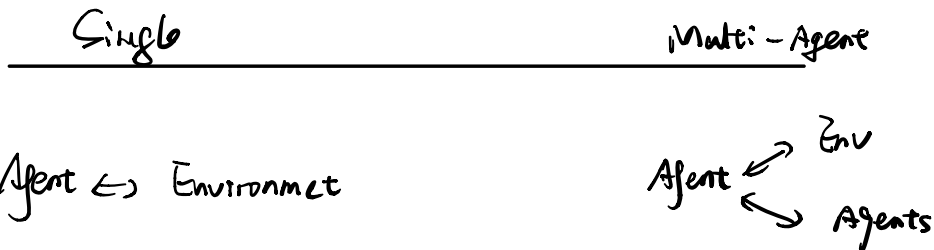


Multi-Agent Reinforcement Learning

1. Single-Agent RL \rightarrow MARL.

环境非静态 (维度爆炸)



① $\langle S_0, A_0, S_1, R_1 \rangle$ MDP 不适用.

② DQN replay Experience 不适用

环境非静态.

2. MARL 方向:

① RL 应用到 MARL

Game 平台

Pong, Sequential social dilemmas.

② Agents Communication

通信 learning

③ Learning cooperation

Agent 之间

- a shared memory
- a single network (parameter sharing)
- Use gradient sharing

共同学习分享多个 Agent 之间

④ Agents Modeling Agents

建模 $\left\{ \begin{array}{l} \text{动作} \\ \text{状态} \end{array} \right.$

(不是 Agent 看成 Env - 部分)

→ MARL Problem:

- ① 维度 $\left\{ \begin{array}{l} \text{States} \\ \text{Actions} \end{array} \right.$ 联合空间

如何降维: Q Mix OxFord Whiteson

- ② Reward (Credit assignment) 如何分配

Counter-factual policy gradient

OpenAI + berkeley

③ RL 收敛、稳定

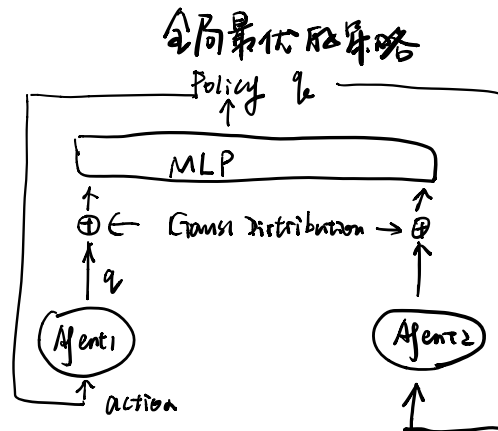
Mean-Field 冯军 伦敦大学

张维楠 上海交大.

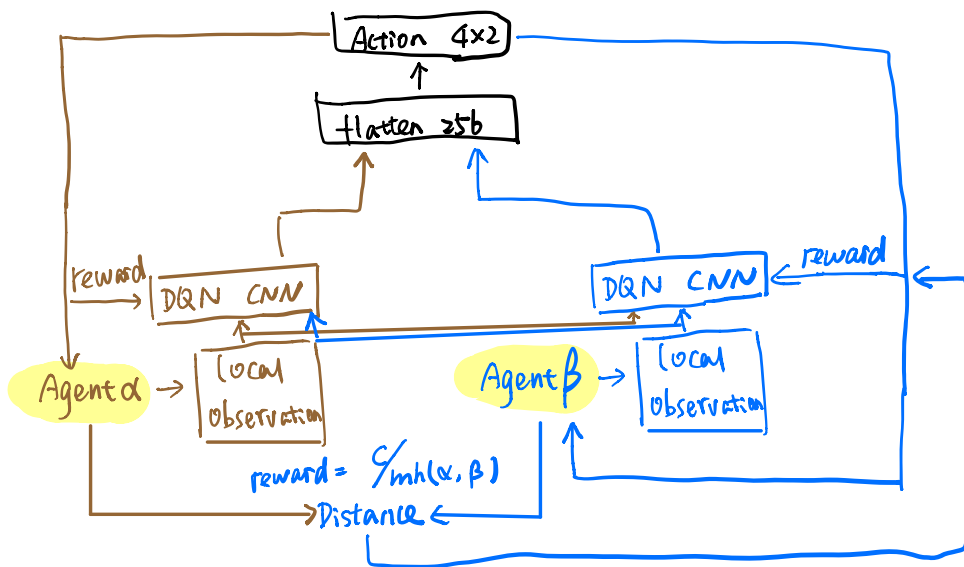
4. 新学期工作:

① Federated learning 杨强

MARL 维度 Federated + RL 框架



比较 CNN 模块比较.



② 工作计划

Task + Agent Path Planning

- 1) 设计框架
- 2) 降低 MARL 维度.
- 3) 全局最优