Understanding Individual Agent Importance in Multi-Agent System via Counterfactual Reasoning

基于反事实推理解释多智能体系统中的个体重要性

陈建明，王亚文，王俊杰，谢肖飞，胡军，王青，徐帆江

Proceedings of the AAAI Conference on Artificial Intelligence (AAAI, CCF-A), 39, 15, 2025.

联系人：陈建明，王亚文，王俊杰，徐帆江

联系方式：{jianming2023, yawen2018, junjie, fanjiang}@iscas.ac.cn

Background

- **Multi-Agent Systems**
  - Become increasingly prevalent.
  - Complexity of team strategies.
  - Lack of explainability.
- **Existing RL Explanation**
  - In-training explanation: inaccurate explanation; cannot work for the black-box agents.
  - Post-training explanation: cannot assess the importance of an agent.

Approach

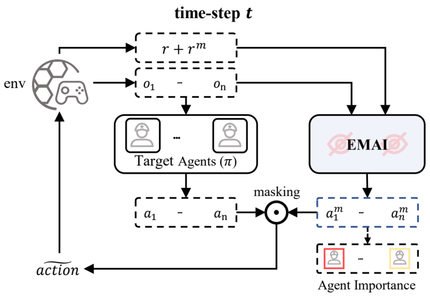
- **Problem Definition**
  - Importance of every agents, at each time-step.
  - The black-box setting where only each agent's observation and action decision can be queried.
- **Problem Modeling**
  - Learn the masking probabilities to denote importance.
  - Optimize the reward differences due to masking and encourage exploration.
- **Architecture and Training**
  - MARL with the CTDE paradigm.

Evaluation

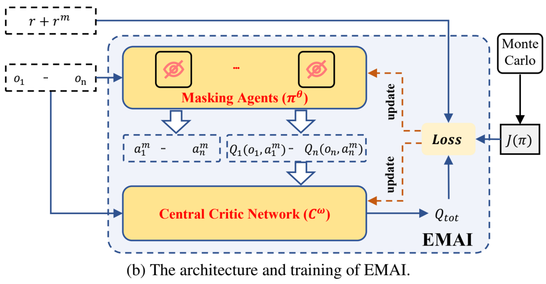
- **Fidelity Evaluation**
  - Average fidelity on 7 tasks: RRD=2.40.
  - Improvement compared to the best baseline: 38% Higher.
- **Usefulness Evaluation**
  - Understanding Policies: more aligned with human intuition; helpful in identifying strategy highlights.
  - Launching Attacks: 14%~289% higher compared to baselines.
  - Patching Policies: the patches guided by EMAI achieve the greatest improvement; in some cases, several baselines could not guide patches correctly, even leading to a decrease in rewards.

Motivation

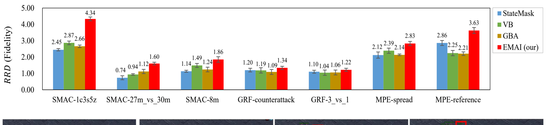
- **Counterfactual Reasoning**
  - Modify the most important elements will exert the greatest impact on the outcome.
- **Key Challenge**
  - Capture the dependencies between agents and across time-steps in MAS.

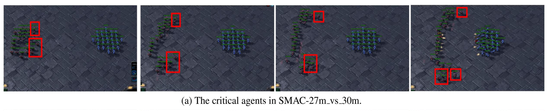


(a) The workflow of EMAI.



(b) The architecture and training of EMAI.





(a) The critical agents in SMAC-27m_vs_30m.

(b) The critical agent in GRF-counter_attack.

Tasks	StateMask	VB	GBA	EMAI (ours)
SMAC-1c3s5z	-0.25 (0.08)	-0.21 (0.16)	-0.19 (0.09)	**-0.68 (0.22)**
SMAC-27m_vs_30m	-0.48 (0.27)	-0.35 (0.17)	-0.45 (0.29)	**-1.41 (0.18)**
SMAC-8m	-0.46 (0.16)	-0.53 (0.25)	-0.41 (0.26)	**-1.43 (0.33)**
GRF-counter_attack	-3.45 (0.61)	-3.19 (0.66)	-3.16 (0.45)	**-4.45 (0.31)**
GRF-3vs1_with_keeper	-1.89 (0.20)	-1.63 (0.38)	-1.34 (0.36)	**-2.30 (0.28)**
MPE-spread	-17.81 (8.49)	-18.14 (7.78)	-17.54 (6.60)	**-23.57 (7.07)**
MPE-reference	-5.40 (0.55)	-6.24 (0.88)	-5.09 (0.51)	**-7.18 (0.92)**

Attack: the changes of episode team rewards before and after the attacks.

Tasks	StateMask	VB	GBA	EMAI (ours)
SMAC-1c3s5z	+0.19 (0.17)	+0.37 (0.13)	+0.22 (0.16)	**+0.75 (0.24)**
SMAC-27m_vs_30m	+0.89 (0.16)	+0.84 (0.18)	+0.90 (0.22)	**+1.11 (0.12)**
SMAC-8m	+0.71 (0.41)	+0.26 (0.51)	+0.51 (0.59)	**+0.92 (0.56)**
GRF-counter_attack	+0.07 (0.64)	-0.63 (0.64)	+0.01 (0.55)	**+1.44 (0.50)**
GRF-3vs1_with_keeper	+0.03 (0.42)	-0.06 (0.58)	-0.09 (0.46)	**+0.33 (0.41)**
MPE-spread	+10.56 (1.39)	+10.01 (0.92)	+8.03 (1.24)	**+12.57 (0.77)**
MPE-reference	+0.24 (1.04)	+0.14 (0.80)	+0.13 (1.06)	**+0.72 (1.11)**

Patching: the changes of episode team rewards before and after the patching.