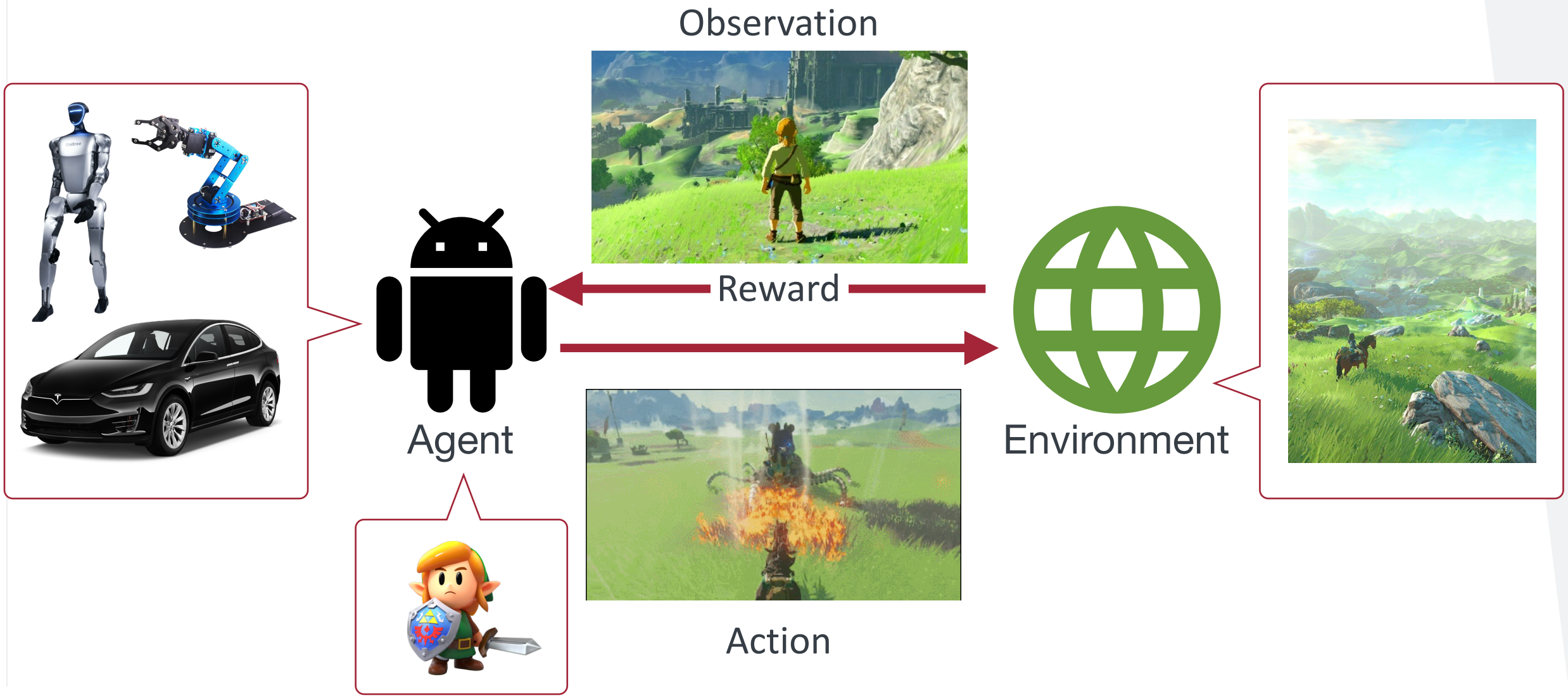
Dynamic Multi-Agent Reinforcement Learning For StarCraft2

Rui Wei, Qingyang Yu, Zixun Xiong

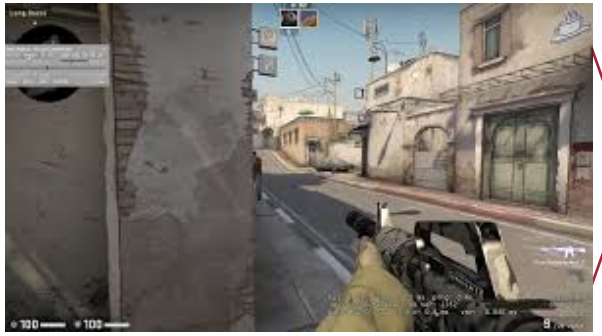
Stevens Institute of Technology

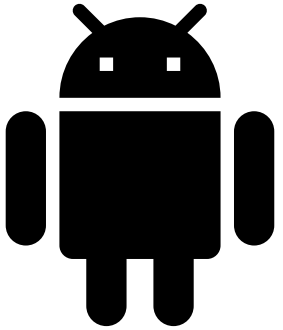
What is Reinforcement Learning (RL)?



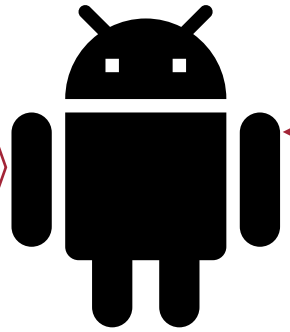
Multi-Agent Reinforcement Learning (MARL)

More complex!

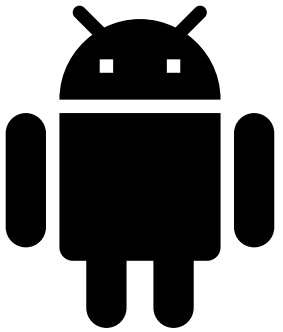




Agent



Agent



Agent

Environment

Training Efficiency of MARL

Significantly Increasing Number of Agents and Resource Costs!

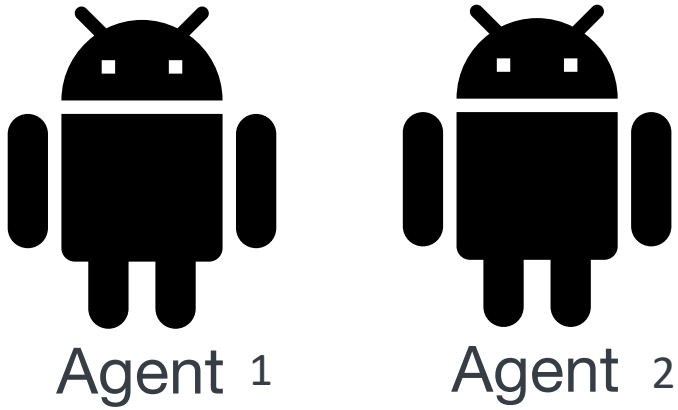
Key Aspects

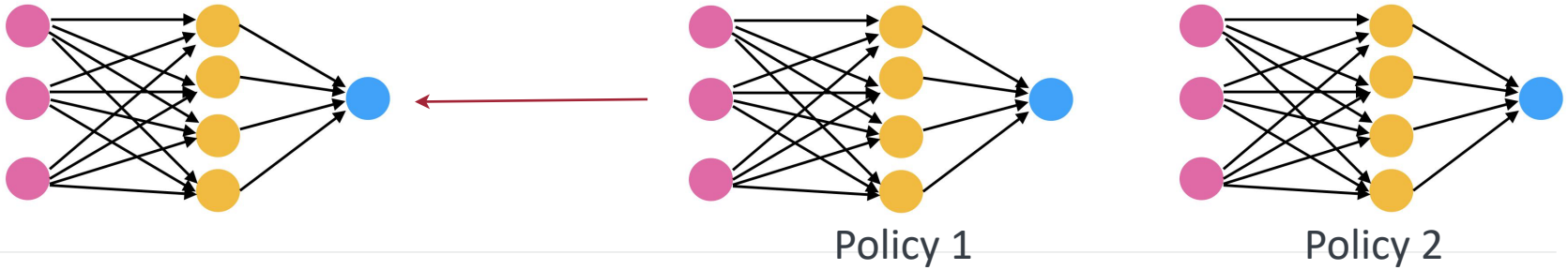
Similar Policies

Could it be possible that the policies of Agent 1 and 2 will become more and more similar during training?

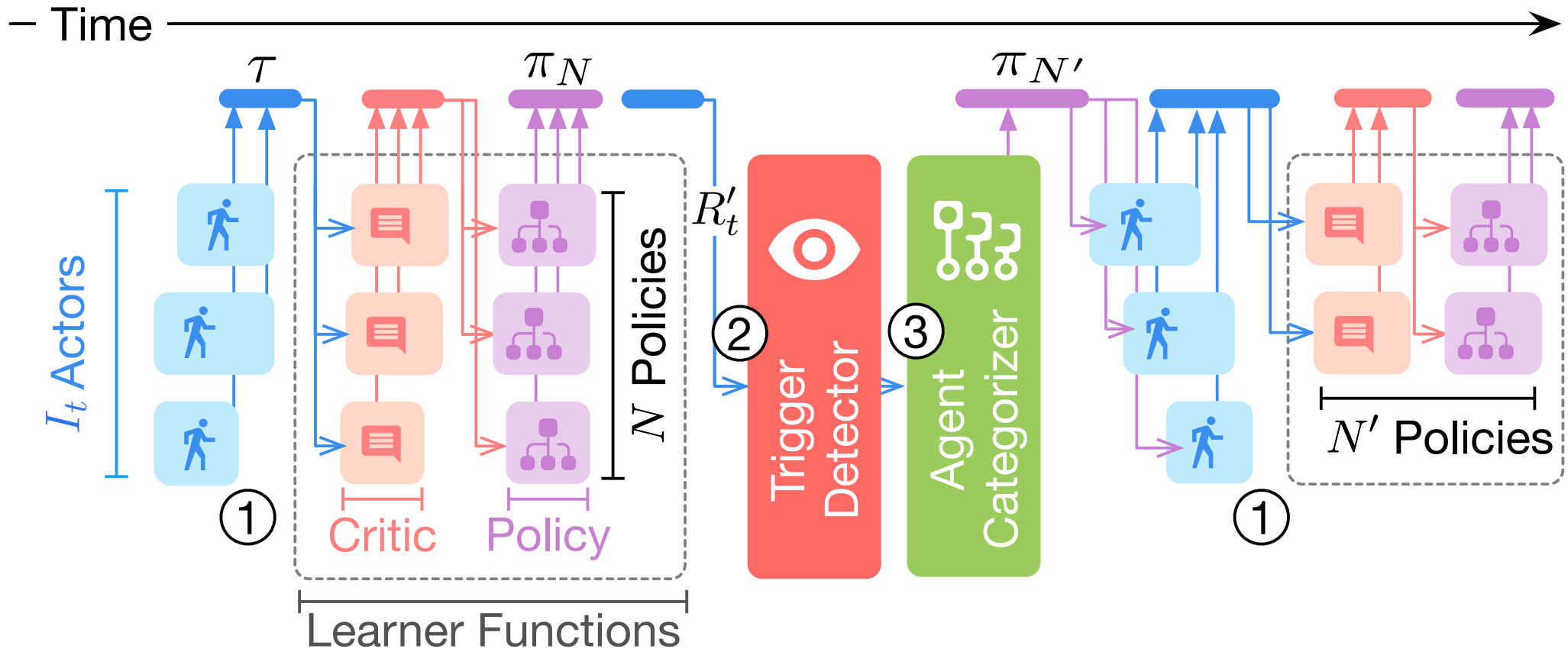
If it's TRUE, we can merge make two agents share the same model

🤔





STEP-WISE DESIGN



① Sampling & training ② Decrease detection
③ Agent categorization for sharing

MAIN TASKS

- *Algorithms Implementation:*
  - **IPPO**, **IQL**, **MAPPO**, **VDN**
- *Environments Integration:*
  - Starcraft Multi-Agent Challenge
    - 8m, 3s5z, …

- *Main Components:*
  - **Detector** - reward, loss, …
  - **Agent Categorizer**
    - Behavior similarity
    - Parameter similarity

EXPECTED TIMELINE & MILESTONES

- *Milestones*
  - **Algorithm Implementation**
    - IPPO, IQL - by 3/10
    - MAPPO, VDN - by 3/15
  - **Agent Categorizer** - by 3/25
  - **Detector** - by 4/10
  - **SMAC Env Integration** - by 4/20
  - **Evaluation & Wrap Up** - by 5/1

- *Workload Distribution*
  - **Rui Wei**
    - Algorithm implementation
    - Environment integration
  - **Qingyang Yu**
    - Agent similarity metric
    - SMAC rendering implementation
  - **Zixun Xiong**
    - Detector design
    - Evaluation design

STEVENS
INSTITUTE OF TECHNOLOGY
1870

THANK YOU

Stevens Institute of Technology
1 Castle Point Terrace, Hoboken, NJ 07030