Individual Contributions as Intrinsic Exploration Scaffolds for Multi-agent Reinforcement Learning
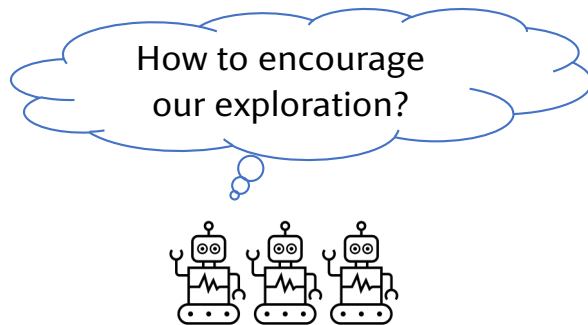
Xinran LI, Zifan LIU, Shibo CHEN, Jun ZHANG

Exploration in Multi-agent Reinforcement Learning

- Exploration is crucial for multi-agent reinforcement learning (MARL) in sparse reward settings
- Intrinsic rewards can often help guide the exploration

How to encourage our exploration?
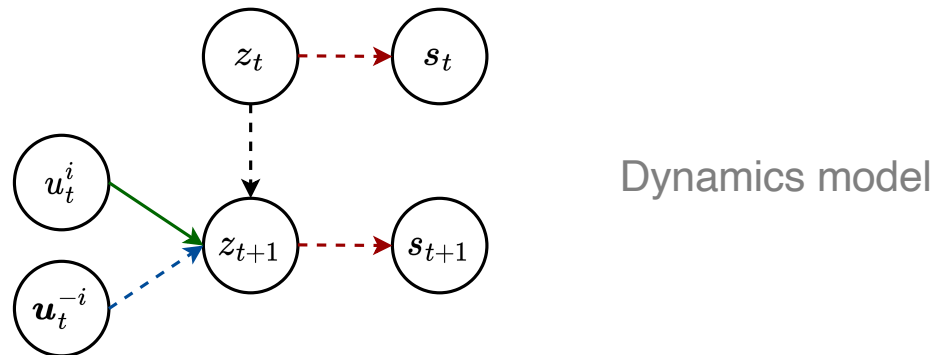
Global
✗ Complicates the credit assignment

Local
✗ Not coordinated exploration

- ICES (Individual Contributions as intrinsic Exploration Scaffolds) motivates exploration by assessing each agent's contribution from a global view

# How to assess individual contributions?

- Bayesian Surprise to Characterize Individual Contributions

Dynamics model

- The contribution $r_{t,\text{int}}^i$ as the mutual information between the latent variable z_{t+1} and the action $u_{t,}^i$, which is given as

$$r_{t,\text{int}}^i = I\left(z_{t+1}; u_t^i \middle| s_t, \boldsymbol{u}_t^{-i}\right) = D_{KL}\left[p(z_{t+1}|s_t, \boldsymbol{u}_t) \parallel p(z_{t+1}|s_t, \boldsymbol{u}_t^{-i})\right]$$

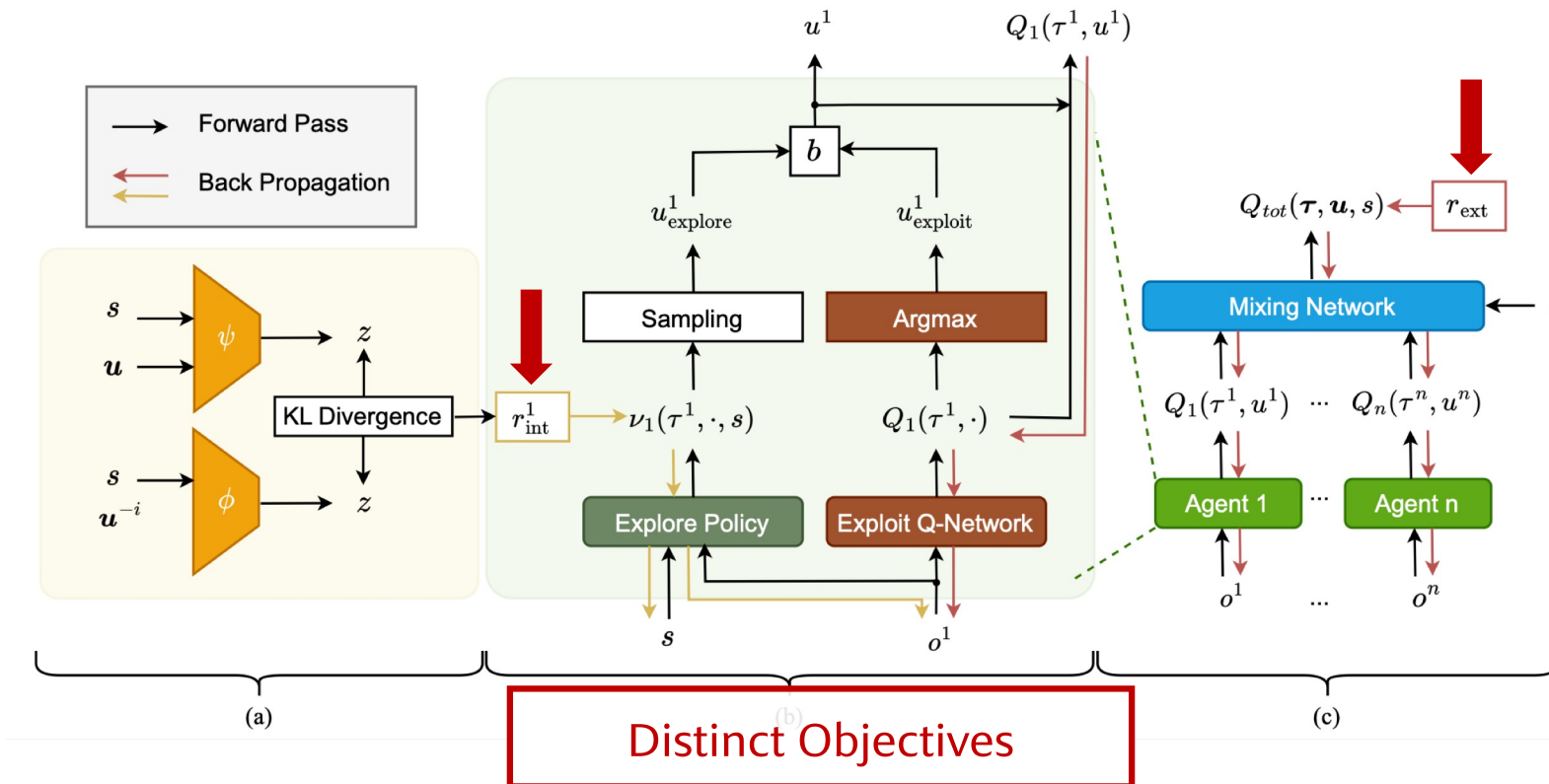# How to assess individual contributions?

- Bayesian Surprise to Characterize Individual Contributions
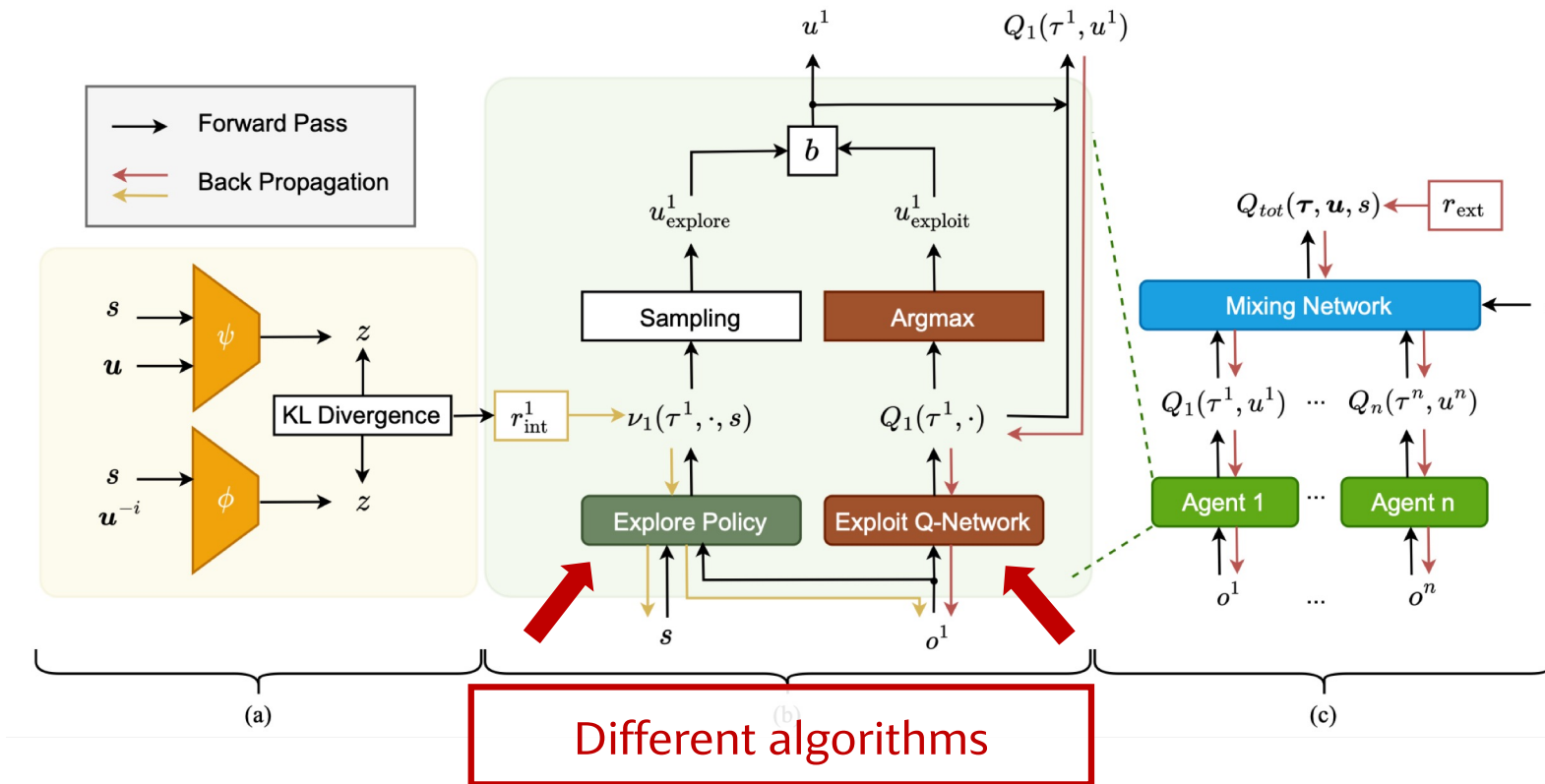- CVAE to Estimate the Bayesian Surprise:
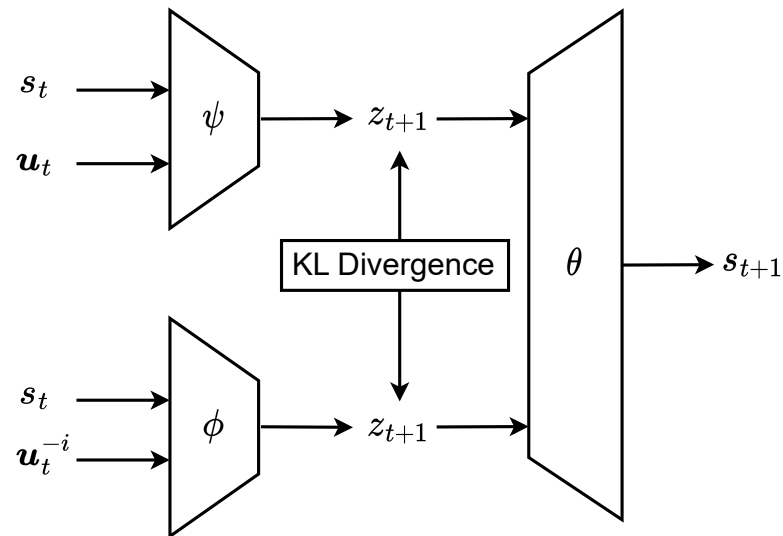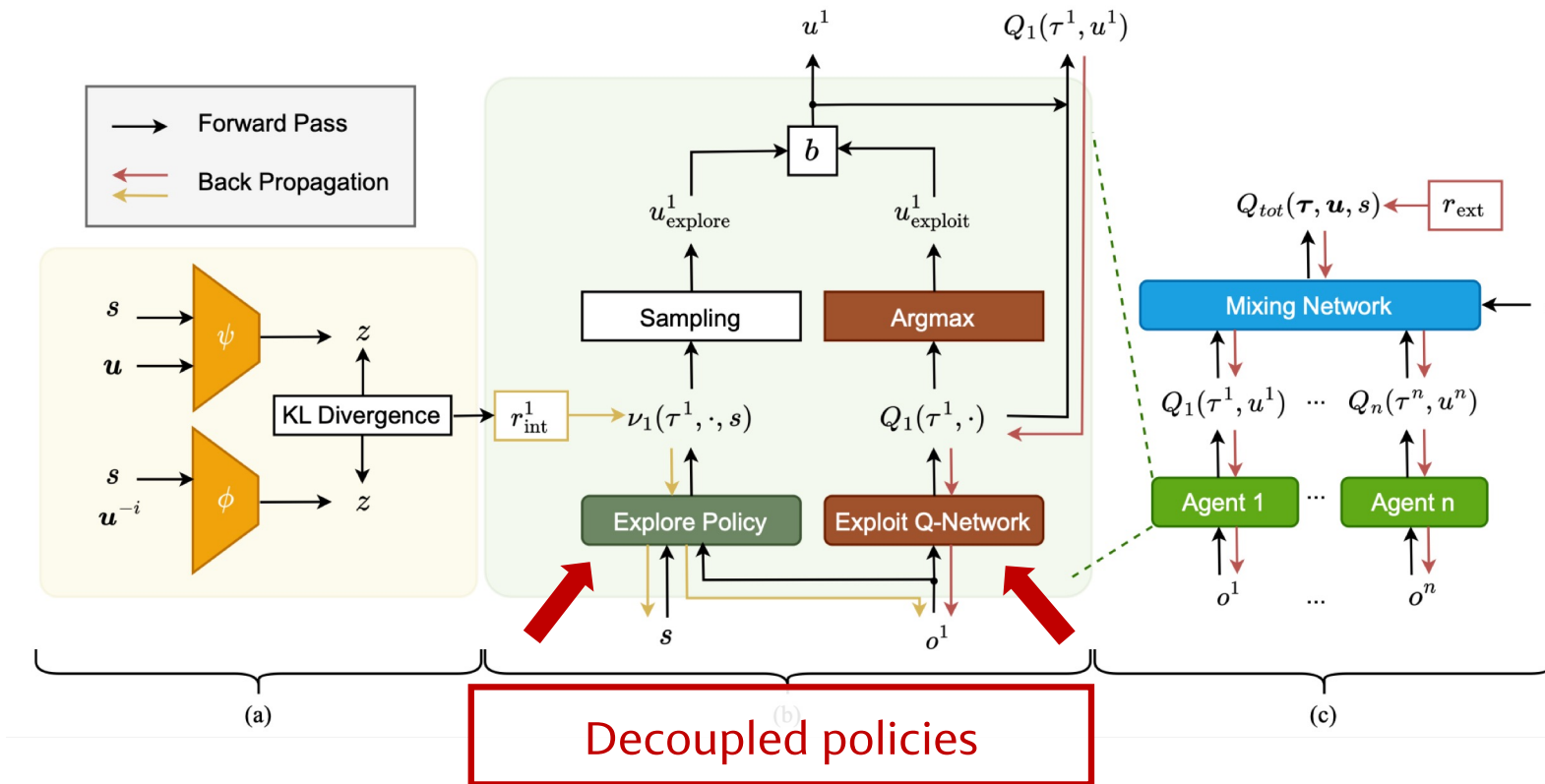
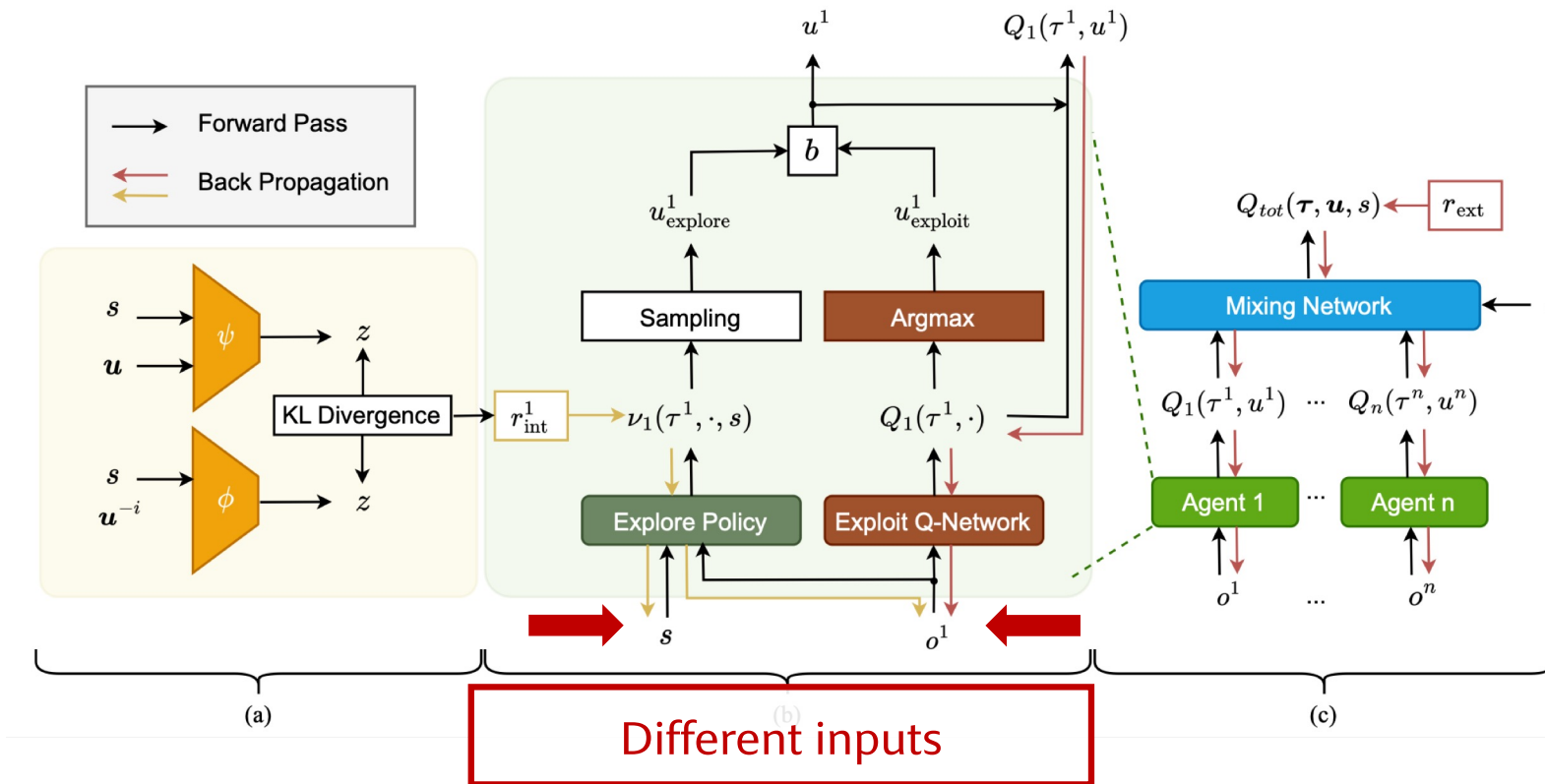# How are these scaffolds utilized effectively?
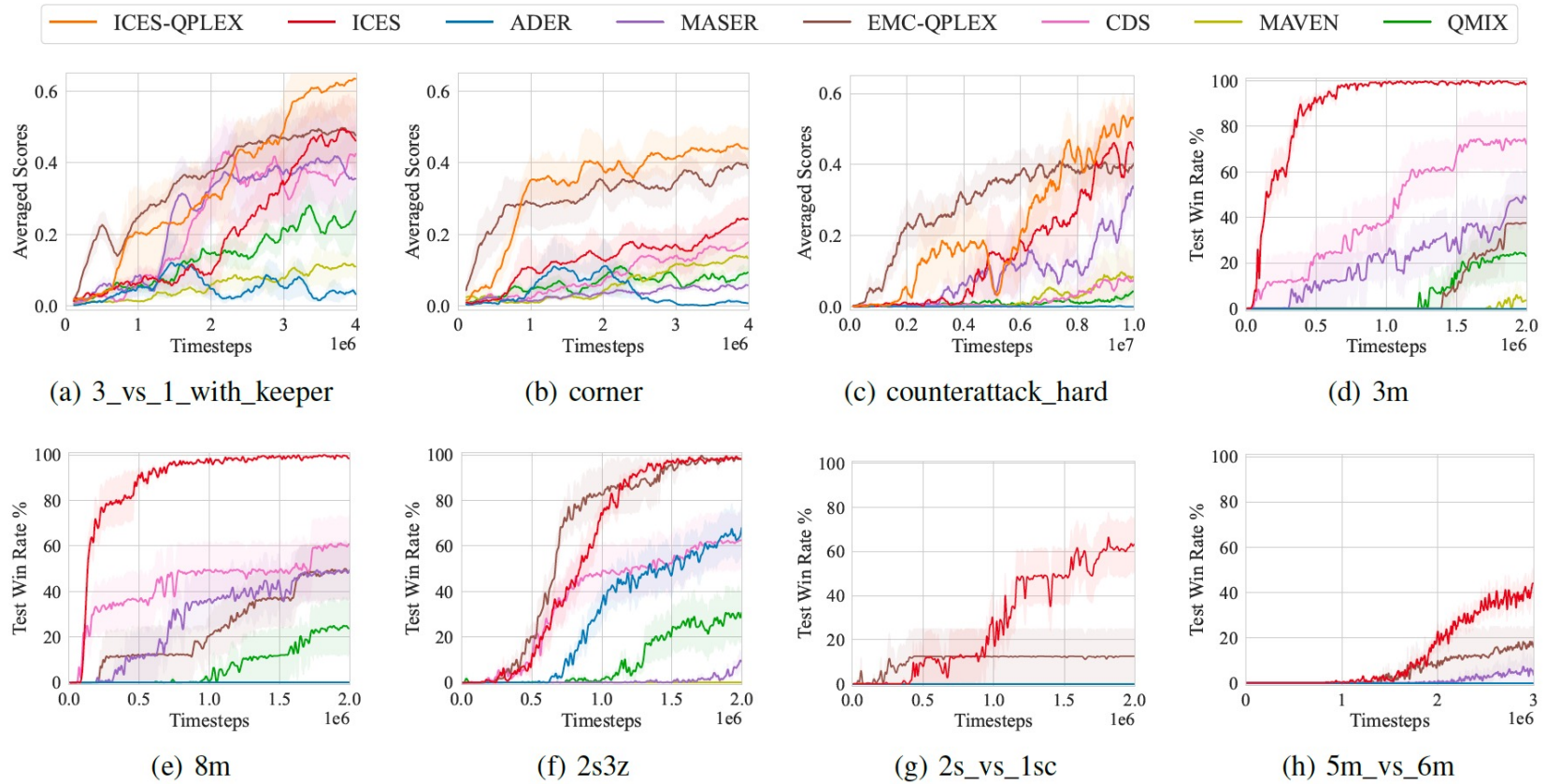
How are these scaffolds utilized effectively?

Distinct Objectives

How are these scaffolds utilized effectively?

Different algorithms

How are these scaffolds utilized effectively?

Is ICES effective in exploration?

(a) 3_vs_1_with_keeper

(b) corner

(c) counterattack_hard
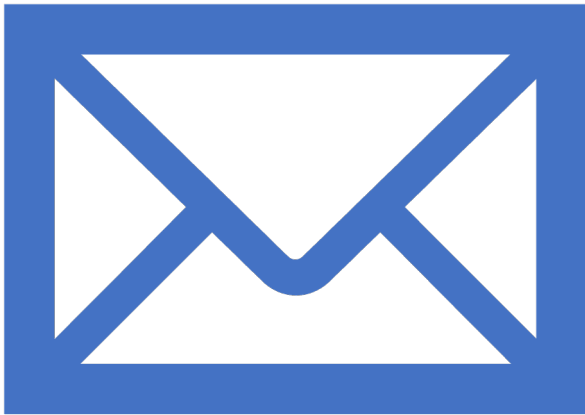
(d) 3m
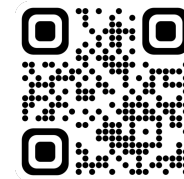
(e) 8m

(f) 2s3z

(g) 2s_vs_1sc

(h) 5m_vs_6m

Take-away

- Key idea:
  - Use individual contributions to encourage agents to have more significant impact on the latent state transition during training time

- Benefits
  - ICES directly assigns the exploration credits to individual agents to bypass global intrinsic reward credit assignment
  - More flexibility brought by decoupled exploration and exploitation

Thank you!

xinran.li@connect.ust.hk

Full paper   Code