Reward-Sharing Relational Networks in Multi-Agent Reinforcement Learning as a Framework for Emergent Behavior

AAMAS 2021, Adaptive and Learning Agents Workshop

Hossein Haeri
PhD Student
Department of Mechanical Engineering
University of Massachusetts Lowell

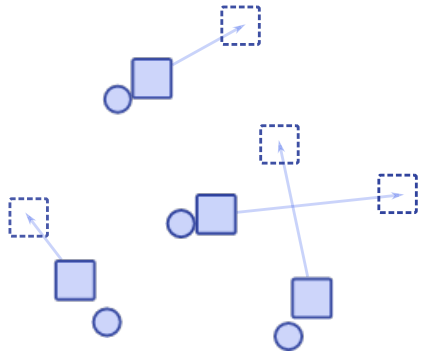
Reza Ahmadzadeh
Assistant Professor
Department of Computer Science
University of Massachusetts Lowell

Kshitij Jerath
Assistant Professor
Department of Mechanical Engineering
University of Massachusetts Lowell

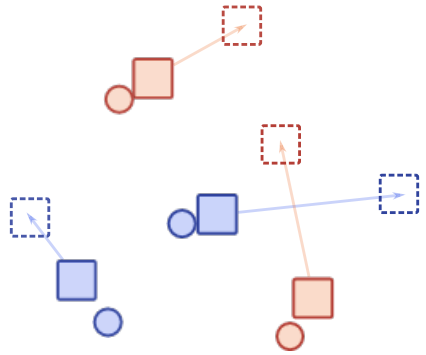
Collaborative vs. Team-based vs. Self-interested MARL

- **Example:** Moving objects from point A to B

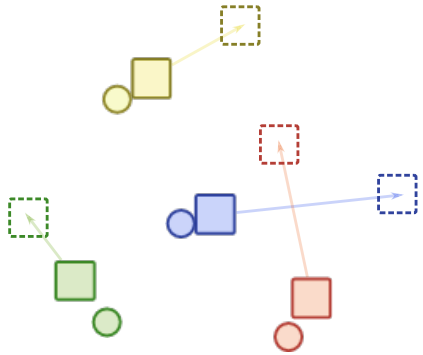
Collaborative (single task)
(single global reward function)



Multiple teams
(single global reward function)



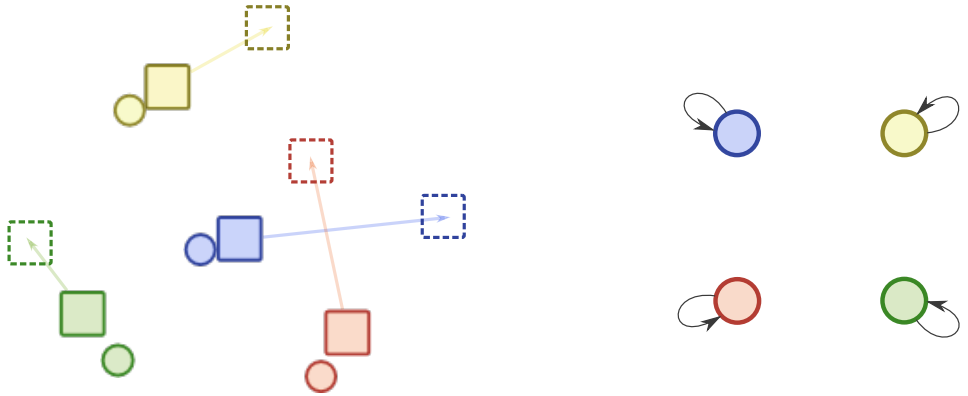
Self-interested
(Single reward function per agent)



What about more complex social networks?

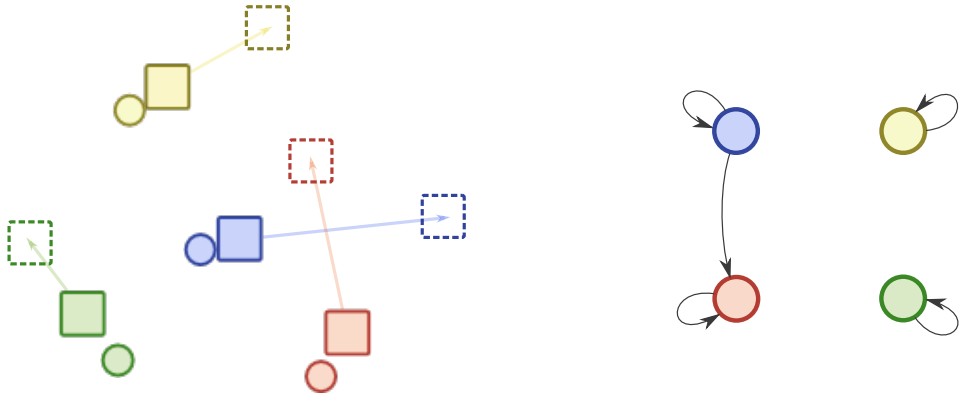
Explicit implementation of networks for multi-agent RL systems

- Building Reward Sharing Relational Network (RSRN) which determines who 'cares' about whom
- If agent i is linked to agent j it means it receives reward according to agent j's individual reward



Explicit implementation of networks for multi-agent RL systems

- Building Reward Sharing Relational Network (RSRN) which determines who 'cares' about whom
- If agent i is linked to agent j it means it receives reward according to agent j's individual reward



Reward Sharing Relational Network (RSRN)

- RSRN can be fully defined by the adjacency matrix $W_{\mathcal{G}}$

$$W_{\mathcal{G}} = \begin{pmatrix} w_{1,1} & w_{1,2} & \cdots & w_{1,N} \\ w_{2,1} & w_{2,2} & \cdots & w_{2,N} \\ \vdots & \vdots & \ddots & \vdots \\ w_{N,1} & w_{N,2} & \cdots & w_{N,N} \end{pmatrix}$$

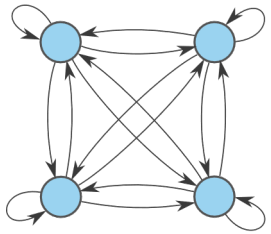
Agent 1's social orientation

Agent 2's social orientation

Agent N's social orientation

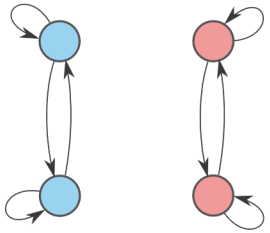
- Weights could be any real number in general (including negative values)

Collaborative vs. Team-based vs. Self-interested vs. Complex MARL



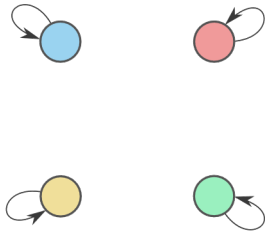
Collaborative

(single global reward function)



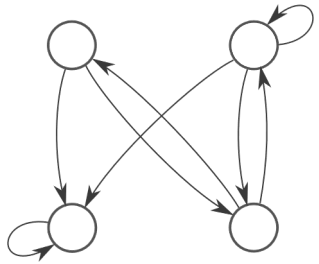
Multiple teams

(single global reward function)



Self-interested

(Single reward function per agent)



Complex Social Network

[2] Baker, Bowen, et al. "Emergent tool use from multi-agent autocurricula." arXiv preprint arXiv:1909.07528 (2019)

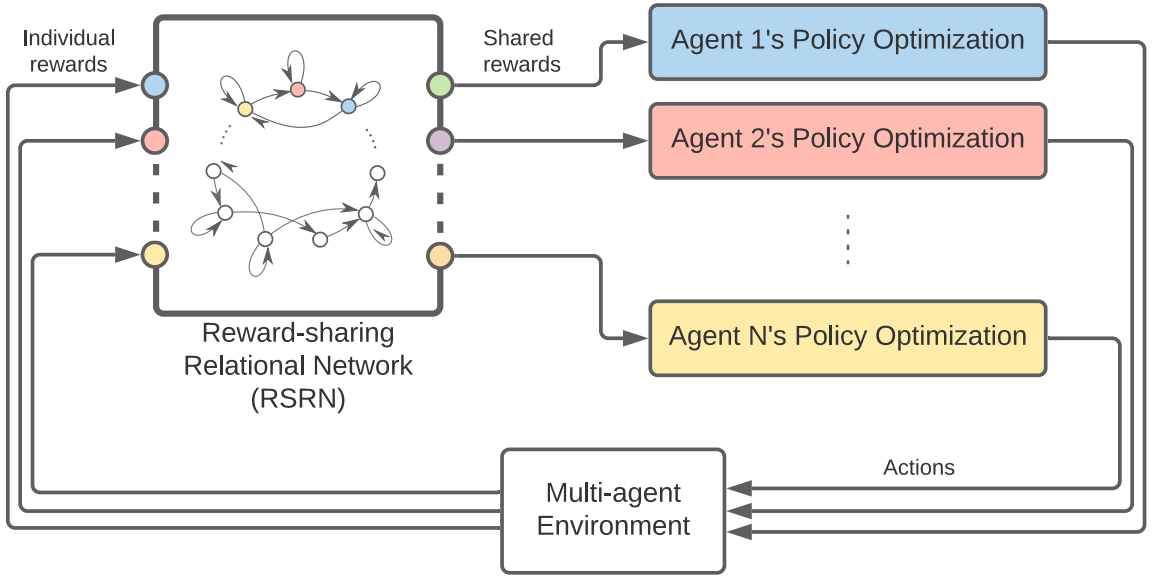
Building composite rewards by scalarization

- Agents try to maximize their shared return $\bar{R}_i$ by accumulating the discounted shared rewards:

$$\bar{R}_i = \mathbb{E}\left(\sum_{k=0}^{T} \gamma^k \bar{r}_{i,k}\right)$$

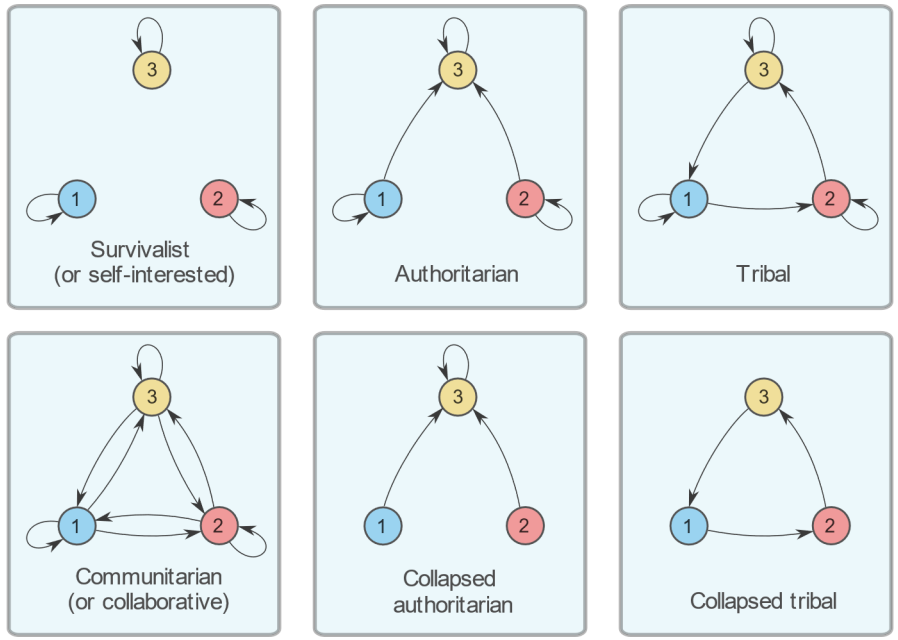
- Reward scalarization using Weighted Product Model (WPM)

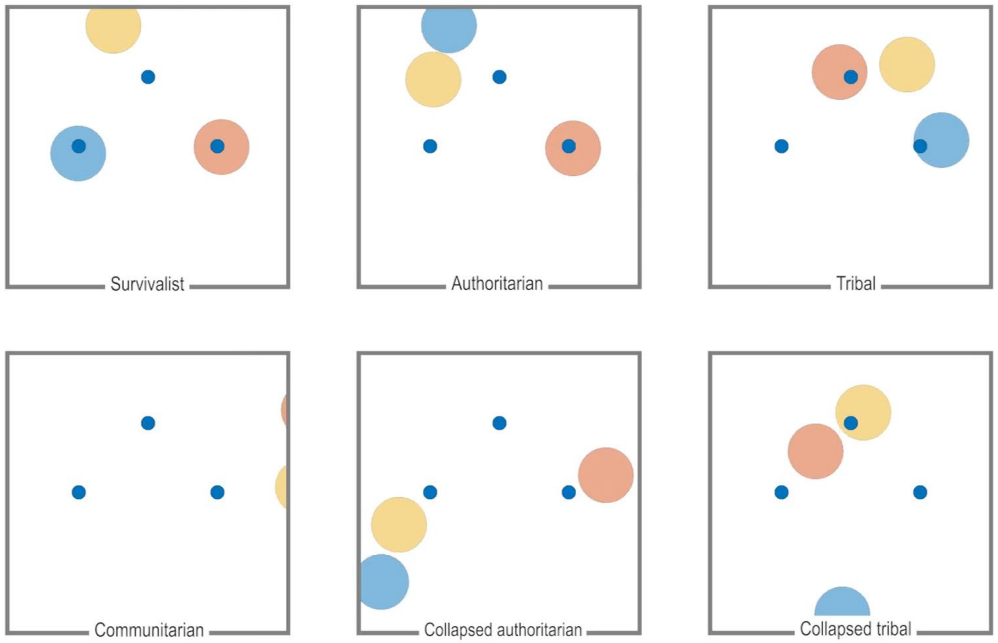
$$\bar{r}_i = \prod_{j=1}^{N} r_j^{w_{i,j}}$$



Evaluating RSRN using a 3-agent MARL environment

- **Scenario:** 3 agents try to get to the closest landmark as fast as possible (agent 3 is slower than others)

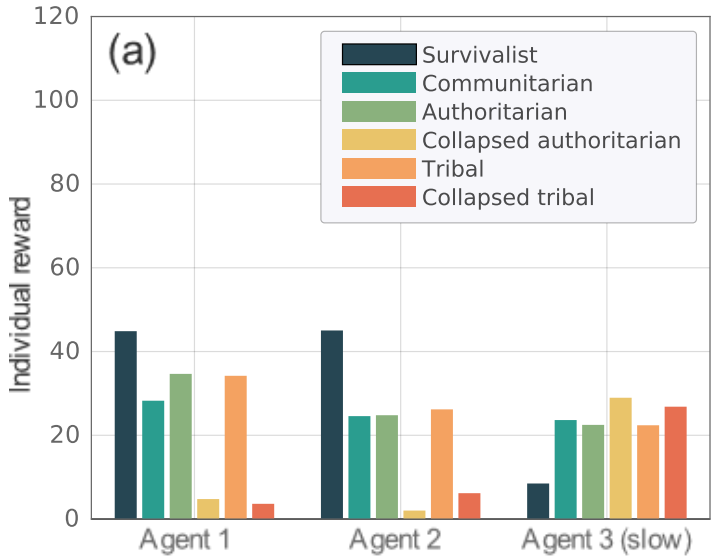


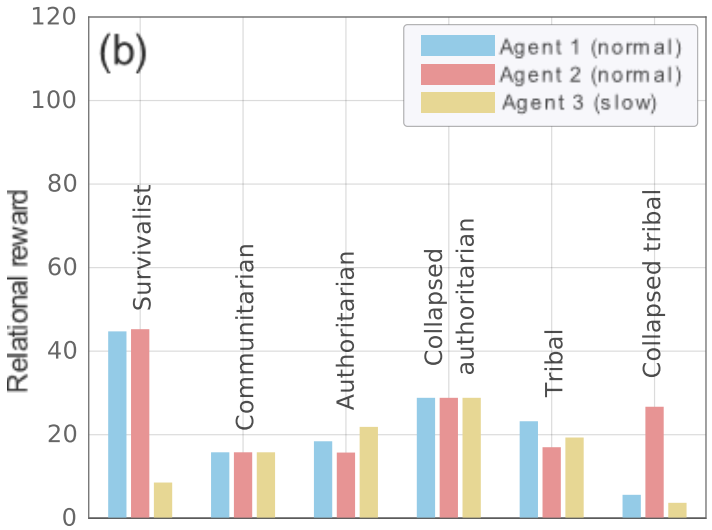


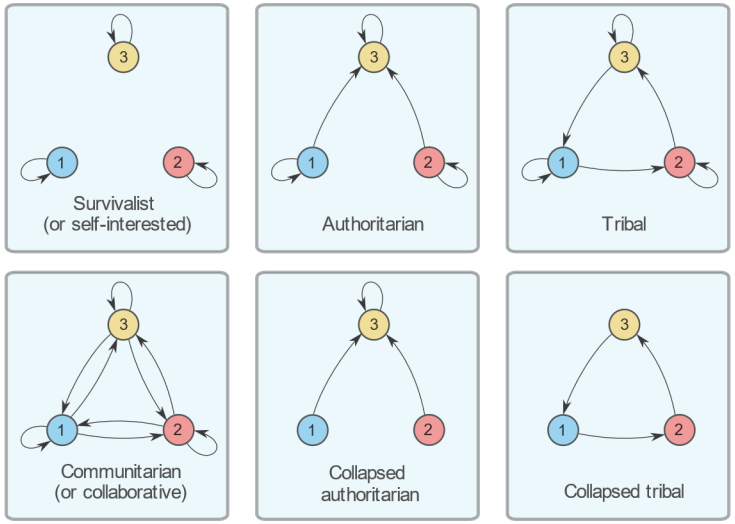
Video is available at sites.google.com/view/marl-rsrn

Balancing agent performances towards RSRN

- Training agents using MADDPG[1] for 500K episodes
- We can measure individual performance of the trained agents using their individual rewards (averaged over 5K test episodes)







[1] Lowe, Ryan, et al. "Multi-agent actor-critic for mixed cooperative-competitive environments." arXiv preprint arXiv:1706.02275 (2017).

Conclusions

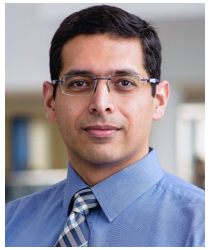
- RSRN can effectively balance A MARL system towards any arbitrary network
- Scalability to large number of agents
- Policy Intractability for agents with several out-going connections
  - Can be addressed by credit assignment approaches
  - Many real-social networks are sparse (small number of out-going connections)
  - We expect to have a sparse RSRN for large scale systems
- Choosing the right scalarization function

Thank you!

Reward-Sharing Relational Networks in Multi-Agent Reinforcement Learning as a Framework for Emergent Behavior

Hossein Haeri
PhD Student
Department of Mechanical Engineering
University of Massachusetts Lowell

Reza Ahmadzadeh
Assistant Professor
Department of Computer Science
University of Massachusetts Lowell



Kshitij Jerath
Assistant Professor
Department of Mechanical Engineering
University of Massachusetts Lowell