Idea: Bridging Theoretical Fairness Definitions with Multi-Agent Coordination in the Real World

Promise Ekpo
Cornell Tech, NY, USA
poe6@cornell.edu

Brian La
Cornell Tech, NY, USA
byl8@cornell.edu

Thomas Wiener
Cornell Tech, NY, USA
tfw29@cornell.edu

Saesha Agarwal
Cornell Tech, NY, USA
sa2388@cornell.edu

Arshia Agrawal
Cornell Tech, NY, USA
aa2443@cornell.edu

Gonzalo Gonzalez-Pumariega
Cornell Tech, NY, USA
gg387@cornell.edu

Lekan Molu
Microsoft Research, NY, USA
patlekano@gmail.com

Angelique Taylor
Cornell Tech, NY, USA
amt298@cornell.edu

Abstract

Current multi-agent reinforcement learning (MARL) theory defines fairness primarily as workload balance, assuming homogeneity across agents in real-world domains. Meanwhile, in fast-paced collaborative settings such as healthcare, there is a need to reduce cognitive overload by ensuring that tasks are allocated to agents based on their expertise (i.e., skill level), which correlates with their efficiency at performing specialized tasks. Hence, fairness requires skill-task alignment. Our findings reveal that (1) optimal fairness emerges from balancing multiple objectives rather than optimizing single metrics, and (2) stronger fairness penalties favor simpler algorithms unless properly tuned. We propose research directions to develop theoretical foundations that bridge current fairness concepts with the real-world coordination challenges of heterogeneous multi-agent systems.

1 Motivation and Problem Statement

Reinforcement learning theory has made substantial progress in formulating fairness mathematically. Established theoretical frameworks define fairness via reward equality in social dilemmas Hughes et al. [2018], worst-case performance guarantees in network scheduling Yuan et al. [2021], and equitable effort distribution in cooperative navigation tasks Aloor et al. [2024]. While these fairness definitions are suitable for some settings, they are not representative of safety-critical environments where fairness involves additional consideration beyond workload balance. For example, in a real-world application such as the ED Anonymous [2025, 2024, 2019], other metrics such as the skill-task alignment contribute to fairness. Our **multi-objective problem** formulation for skill-task fairness alignment seeks to investigate fairness algorithms and definitions that jointly account for (i) equitable workload distribution and (ii) skill–task alignment in the presence of heterogeneous capabilities. Empirically, we observe that while appropriate weighting between these two metrics improves goal success rates and fairness among agents, excessive penalties hurt agents' abilities to learn. Our goal in this idea paper is to ignite a collaboration between theorists and practitioners to create fairness metrics that not only work in MARL benchmarks, but also captures the consideration of domain experts in real-world teams.

2 MARLHospital

MARLHospital is a simulation framework for studying collaboration among heterogeneous agents in medical environments, modeling how varied expertise levels and shared-subtask structure influence coordination in emergency-room workflows. The environment encodes the procedural steps of "Adult Basic Life Support" from American Red Cross code cards American Red Cross [2025], representing CPR and rescue-breath tasks as sequential subtasks such as placing a board, performing chest compressions, and delivering rescue breaths (Figure 1). Task difficulty is controlled through short-horizon CPR goals and longer-horizon rescue-breath goals, both of which can be procedurally modified in the environment configuration. MARLHospital is compatible with standard MARL baselines including IL and CTDE methods implemented via EPyMARL [Papoudakis et al., 2021] and PyMARL [Samvelyan et al., 2019] to support comparison across decentralized and centralized training setups Foerster et al. [2016], He et al. [2016], Du et al. [2019], Amato [2024]. To reflect realistic clinical team structures, the simulator defines uniform teams (identical capabilities), specialized teams (agents with faster performance on particular subtasks), and interdependent teams that require forced cooperation because each agent can only execute a subset of actions. This structure enables controlled investigation of how skill diversity, task hierarchy, and coordination constraints interact with fairness objectives in multi-agent reinforcement learning.

3 Setting

We consider cooperative MARL with N agents and episodic tasks composed of M subtasks. Let $|\tau_i|$ be the number of subtasks completed by agent i; the *workload imbalance* is measured by the Gini index L_1. Let $\mathcal{S}_{i_t}(t) \in [0, 1]$ denote the skill of the executing agent i_t on subtask t, and $\max_j \mathcal{S}_j(t)$ the best available skill; the *skill misalignment* metric L_2 measures deviation from best-skill assignment. We define a composite disparity

$$L_3(\alpha) \; = \; \alpha \, L_1 \; + \; (1 - \alpha) \, L_2, \quad \alpha \in [0, 1], \tag{1}$$

and shaped team reward $r_t = R(s_t, \mathbf{a}_t) - \lambda L_3(\alpha)$ with penalty scale $\lambda > 0$. Intuitively, α controls the fairness *dimension* (workload vs. skill alignment) while λ controls its *strength*. $R(s_t, \mathbf{a}_t)$ is the efficiency reward R_1 while r_t is the fairness shaped reward.

3.1 Theoretical assumptions

Existing MARL theory treats fairness as a constraint satisfaction problem where maintaining reward equality is subject to performance requirements Jiang and Lu [2019]. Thus, assuming equal capabilities across agents (i.e. agents' homogeneity), tasks require similar skill (i.e. task uniformity) and a linear relationship between fairness and efficiency. However , these assumptions rarely hold in the real world as healthcare teams and other applications in human-AI collaboration mostly feature agents with varying expertise (i.e. skill levels) and tasks requiring different skill levels.

3.2 Experiments

To gain insights about this gap between theory and the real world, we conducted experiments comparing the individual contributions of three heterogeneous agents with varying skill levels in a medical-inspired resuscitation task utilising the QMIX MARL algorithm (workload balance in Table 1) with the efficiency reward ($R(s_t, \mathbf{a}_t)$), and then with QMIX FairSkillMARL reward with r_t. For evaluation, we compare this to QMIX Fair Efficient Networks (FEN) Jiang and Lu [2019], an earlier framework for fairness in MARL, which measures agent resource utilization and penalizes agents when they deviate from the average utilization of all agents. In addition, we vary the degree of penalty on fairness and efficiency (λ), as well as the extent to which we care about workload balance and skill-task misalignment(α).

4 Findings

Our systematic evaluation reveals three critical insights that challenge current theoretical assumptions, as seen in Table 1:

Table 1: Comprehensive evaluation of FairSkillMARL and FEN using fairness metrics averaged over last 100 episodes. All fairness metrics (L1) and agent contributions are computed as moving averages to ensure robust measurement.

| Experiment | λ | Method | α | Success Rate ↑ | L1 ↓ | Averaged Agent % | | | Range ↓ |
					L1	A0	A1	A2	
Experiment 1: Impact of Skill-Task Alignment (L2) - Specialized Teams									
$\lambda = 0.0$	0.0	Workload-only	0.0	0.110 ± 0.10	0.270	33.25	36.61	30.14	6.47
		FairSkillMARL	0.7	0.096 ± 0.06	0.323	39.64	39.66	20.69	18.97
		FEN	—	0.086 ± 0.05	0.300	28.49	38.17	33.34	9.68
$\lambda = 1.0$	1.0	Workload-only	0.0	0.104 ± 0.07	0.342	44.17	40.48	15.35	28.82
		FairSkillMARL	0.7	0.067 ± 0.04	0.398	49.94	34.75	15.31	34.63
		FEN	—	0.159 ± 0.18	0.351	40.98	36.84	22.18	18.80
$\lambda = 4.0$	4.0	Workload-only	0.0	0.081 ± 0.07	0.371	46.35	32.65	21.00	25.35
		FairSkillMARL	0.7	0.113 ± 0.08	0.351	37.44	48.74	17.82	30.92
		FEN	—	0.429 ± 0.44	0.254	31.48	41.86	26.65	15.21
Experiment 2: Team Composition Effects (Equal vs Specialized Teams, $\lambda = 1.0$)									
Specialized	1.0	Workload-only	0.0	0.104 ± 0.07	0.342	44.17	40.48	15.35	28.82
		FairSkillMARL	0.7	0.067 ± 0.04	0.398	49.94	34.75	15.31	34.63
		FEN	—	0.159 ± 0.18	0.351	40.98	36.84	22.18	18.80
Equal Skills	1.0	Workload-only	0.0	0.118 ± 0.06	0.401	33.63	39.44	26.93	12.51
		FairSkillMARL	0.7	0.245 ± 0.18	0.342	29.59	39.48	30.92	9.89
		FEN	—	$0.194 \pm$ nan	0.338	45.27	50.40	4.33	46.07
Experiment 3: FairSkillMARL Ablation - Optimal α Selection ($\lambda = 1.0$)									
$\lambda = 1.0$	1.0	Workload-only ($\alpha = 0$)	0.0	0.060 ± 0.04	0.419	46.08	36.75	17.17	28.91
		FairSkillMARL	0.5	0.092 ± 0.02	0.388	45.10	36.63	18.27	26.83
		FairSkillMARL	0.7	0.133 ± 0.06	0.328	44.73	35.28	19.99	24.74
		FairSkillMARL ($\alpha = 1$)	1.0	0.050 ± 0.02	0.316	47.10	39.39	13.71	33.39
		FEN	—	0.146 ± 0.10	0.337	42.26	41.82	15.92	26.34

Note: Fairness metrics (L1) and agent contributions are averaged over the last 100 episodes. Success rates are calculated at convergence. Range = max(Agent%) - min(Agent%).

4.1 Optimal fairness requires multi-objective balance

We observe that neither pure workload balancing nor pure skill alignment achieves optimal coordination. Instead, intermediate balance points consistently outperform both extremes. Theoretically, there is room for application of multi-objective optimisation in fairness in multi-agent RL formulation.

4.2 Different outcomes are obtained for prioritising skill-task misalignment and workload balance

The linear combination of workload balance and skill-task misalignment might not be the most appropriate, as these might not hold or even conflict in some cases. For example, if all tasks require only highly skilled agents, then there might be a conflict that needs to be theoretically bounded, or if all tasks require low-skilled workers, high-skilled workers might get no assignments.

4.3 Fairness definition should be expanded to include the domain expertise preference

Fairness should be context-dependent, as different domains may have additional fairness viewpoints beyond equalised effort. The same fairness mechanism exhibits qualitatively different coordination patterns across team (equal versus varying skilled) and task structures. Also, domain experts should be consulted to determine additional parameters they care about, which would help formulate more practical definitions in theory to align with real-world preferences. One observation from our experiments shows that team heterogeneity should be considered in fairness definitions to design effective fairness approaches. A key insight from this is that there is an opportunity to MARL fairness

definitions to consider additional dependencies that affect fairness, such as the dependency between environmental structure and coordination mechanisms.

5 Recommended Research Directions to Align Fairness Definitions with Real-World Setups

Based on our experiment results, we propose **two research directions** to align theoretical fairness frameworks with real-world coordination. Based on our experiment results, we propose two research directions to align theoretical fairness79frameworks with real-world coordination.80First, there is an opportunity to develop a multi-objective fairness theory treating fairness as a81multi-objective optimisation problem rather than a single-dimensional one. Specifically, we propose82the development of theoretical frameworks that formulate the pareto optimality between fairness83objectives, provide convergence guarantees for composite fairness functions that generalise to different84environment structures. Our preliminary results suggest stable balance points exist, but there is a85need for concrete theoretical formulations to predict them a priori.86Secondly, there is a gap in designing context-dependent fairness frameworks. Theoretical fairness87definitions in MARL assume generalization across team structures and task domains. Hence, there is88an opportunity to develop theoretical tools that characterise how team heterogeneity affects fairness89definitions, design formal methods for adapting fairness definitions to domain structure, determine90edge cases and and theoretically guarantee cases where individual fairness consideration (skill-task91alignment) might conflict with overall group fairness (workload balance).92

6 Conclusion and call for collaboration

Our experimental evidence reveals substantial gaps between theoretical fairness definitions in MARL and multi-agent coordination preferences in the real world. To model real-world team coordination challenges, current theoretical fairness definitions need to be updated. Our main position is that the definition of fairness in MARL is not a solved theoretical problem requiring more engineering. Instead, it presents a challenge that requires theoretical frameworks that consider multi-objective optimisation and adapt to different domain preferences. There is also a need to include insights from domain experts, especially those in safety-critical fields like healthcare, to help expand our understanding and formulation of fairness in multi-agent systems. There are also some concepts that can be included from the fairness in recommendation and information-retrieval systems literature, where attention allocation to content creators is explored either based on everyone receiving equal attention or attention based on some additional parameters or underlying factors. We invite theoretical researchers to engage with our experimental observations and develop frameworks that consider real-world coordination realities in fairness definitions. Our current experiments provide a foundation for more fairness theories to be formulated. In addition, we encourage experimentalists to investigate and identify more of these gaps that exist between theory and real-world multiagent interactions across diverse domains in order to build the experimental evidence base to spark the interest of theoreticians in formulating theories in this field.

Our future work involves developing comprehensive experimental analyses across additional on-policy and off-policy MARL algorithms in both centralised and decentralised execution settings. We have built a simulator with healthcare-inspired tasks, differently-skilled predefined teams and energy-switching tasks. We have compared our QMIX algorithms against MAPPO using the fairness reward. We have also compared our algorithms to Fair-efficient Networks(FEN) as done in Aloor et al. [2024] and potentially Inequity Aversion (IA) Hughes et al. [2018] and will compare to additional fairness MARL algorithms, such as FA-MAPPO Malfa et al. [2025]. We are also engaging with domain experts at emergency wards in US hospitals to investigate additional variables they care about apart from skill-task misalignment. Additionally, we can explore adding fairness as an additional constraint, rather than only via reward shaping, in the formulation of the optimization problem. These results will be presented at major conferences, building on the theoretical foundations this workshop aims to establish. The future of more robust fairness definitions in MARL depends on interdisciplinary collaboration between theoreticians, domain experts and experimentalists.

References

Jasmine Jerry Aloor, Siddharth Nayak, Sydney Dolan, and Hamsa Balakrishnan. Cooperation and Fairness in Multi-Agent Reinforcement Learning, October 2024. URL `http://arxiv.org/abs/2410.14916`.

Christopher Amato. A First Introduction to Cooperative Multi-Agent Reinforcement Learning, December 2024. URL `http://arxiv.org/abs/2405.06161`.

American Red Cross. American Red Cross Code Cards, January 2025. URL `https://www.redcrosslearningcenter.org/s/american-red-cross-code-cards`.

Anonymous. Coordinating Clinical Teams: Using Robots to Empower Nurses to Stop the Line. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):1–30, November 2019. ISSN 2573-0142. doi: 10.1145/3359323. URL `https://dl.acm.org/doi/10.1145/3359323`.

Anonymous. Towards Collaborative Crash Cart Robots that Support Clinical Teamwork. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*, HRI '24, pages 715–724, New York, NY, USA, March 2024. Association for Computing Machinery. ISBN 979-8-4007-0322-5. doi: 10.1145/3610977.3634967. URL `https://dl.acm.org/doi/10.1145/3610977.3634967`.

Anonymous. Rapidly Built Medical Crash Cart! Lessons Learned and Impacts on High-Stakes Team Collaboration in the Emergency Room, February 2025. URL `http://arxiv.org/abs/2502.18688`. arXiv:2502.18688 [cs] version: 1.

Yali Du, Lei Han, Meng Fang, Ji Liu, Tianhong Dai, and Dacheng Tao. LIIR: Learning Individual Intrinsic Reward in Multi-Agent Reinforcement Learning. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL `https://proceedings.neurips.cc/paper_files/paper/2019/hash/07a9d3fed4c5ea6b17e80258dee231fa-Abstract.html`.

Jakob Foerster, Ioannis Alexandros Assael, Nando de Freitas, and Shimon Whiteson. Learning to Communicate with Deep Multi-Agent Reinforcement Learning. In *Advances in Neural Information Processing Systems*, volume 29. Curran Associates, Inc., 2016. URL `https://proceedings.neurips.cc/paper_files/paper/2016/hash/c7635bfd99248a2cdef8249ef7bfbef4-Abstract.html`.

He He, Jordan Boyd-Graber, Kevin Kwok, and Hal Daumé III. Opponent Modeling in Deep Reinforcement Learning, September 2016. URL `http://arxiv.org/abs/1609.05559`.

Edward Hughes, Joel Z. Leibo, Matthew G. Phillips, Karl Tuyls, Edgar A. Duéñez-Guzmán, Antonio García Castañeda, Iain Dunning, Tina Zhu, Kevin R. McKee, Raphael Koster, Heather Roff, and Thore Graepel. Inequity aversion improves cooperation in intertemporal social dilemmas, September 2018. URL `http://arxiv.org/abs/1803.08884`.

Jiechuan Jiang and Zongqing Lu. Learning Fairness in Multi-Agent Systems, October 2019. URL `http://arxiv.org/abs/1910.14472`. arXiv:1910.14472 [cs].

Gabriele Malfa, Jie Zhang, Michael Luck, and Elizabeth Black. *Fairness Aware Reinforcement Learning via Proximal Policy Optimization*. February 2025. doi: 10.48550/arXiv.2502.03953.

Georgios Papoudakis, Filippos Christianos, Lukas Schäfer, and Stefano V. Albrecht. Benchmarking Multi-Agent Deep Reinforcement Learning Algorithms in Cooperative Tasks, November 2021. URL `http://arxiv.org/abs/2006.07869`.

Mikayel Samvelyan, Tabish Rashid, Christian Schroeder de Witt, Gregory Farquhar, Nantas Nardelli, Tim G. J. Rudner, Chia-Man Hung, Philip H. S. Torr, Jakob Foerster, and Shimon Whiteson. The StarCraft Multi-Agent Challenge, December 2019. URL `http://arxiv.org/abs/1902.04043`. arXiv:1902.04043 [cs].

Mingqi Yuan, Qi Cao, Man-On Pun, and Yi Chen. Multi-agent reinforcement learning-based fairness-aware scheduling for bursty traffic. In *2021 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6. IEEE, 2021.
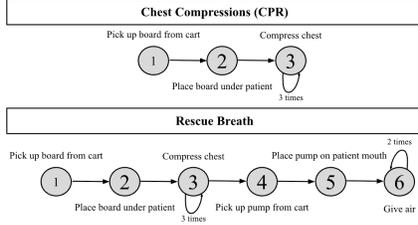
Figure 1: Task flow diagrams for the check compression (CPR) and rescue breath tasks in MARL-Hospital.

A MARLHospital

We present MARLHospital, a framework for modeling multi-agent collaborative tasks in medical environments. We address the challenge of coordinating agents with varying expertise levels and in shared task mode, which is prevalent in multi-agent settings, particularly in ERs.

A.1 Medical Tasks

We model the CPR and AED tasks using the standard procedural steps of "Adult Basic Life Support" from the American Red Cross code cards (see Figure 1). These cards provide visual flow charts guiding resuscitation procedures to capture real-world tasks based on evidence-based practices **?**. These procedural steps are modeled in the action space of MARLHospital.

A.2 Baseline Methods

Our experiments include standard MARL algorithms, including IL and CTDE, trainable on our custom simulation environment as building blocks for more sophisticated algorithms (i.e., EPyMARL library [Papoudakis et al., 2021][1] and PyMARL [Samvelyan et al., 2019][2]) Foerster et al. [2016], He et al. [2016], Du et al. [2019], Amato [2024]. We investigate structured team compositions, including standard baselines that might facilitate different coordination strategies based on their decentralized and centralized training setups.

A.3 Task Difficulty

We define the two goals as Partial (P) and Complete (C) with task difficulties based on the length of the time horizon. For the CPR goal, the HCWs must perform CPR, a short-time-horizon task that consists of picking up and placing a board under the patient and giving N chest compressions. For the rescue breaths, which is a longer horizon task, in addition to the CPR goal, the HCWs must perform the same actions in the CPR task, in addition to picking up the pump and placing it on the patient to give them air. Since we use procedural generation in the environment implementation, these goals can be modified in the environment configuration file.

A.4 Team Compositions

Understanding team composition is essential for modeling clinical settings, as the structure and capabilities of a care team influence task performance, decision speed, and workload distribution. We define three team compositions that vary in expertise levels (EL) of agents: *uniform*, *specialized*, and *interdependent* teams. In *uniform Teams*, all agents possess identical capabilities and can perform every subtask with equal efficiency. This structure allows maximum flexibility as any agent can complete any part of the task independently, and coordination is optional. *Specialized Teams* allows agents to be more efficient in one particular subtask but still retains the ability to perform all others, although more slowly. This specific setting incentivizes but does not necessitate collaboration. In *Interdependent Teams (Forced Cooperation)*, agents are capable of performing only a subset of subtasks and cannot execute at least two others.

---

[1]https://github.com/uoe-agents/epymarl
[2]https://github.com/oxwhirl/pymarl