A Utility-Based Perspective on Multi-Objective Multi-Agent Decision Making

Doctoral Consortium

Roxana Rădulescu
Artificial Intelligence Lab
Vrije Universiteit Brussel, Belgium
roxana.radulescu@vub.be

ABSTRACT

Numerous real-world problems involve multiple interacting entities and are inherently multi-objective in nature. Multi-objective multi-agent systems are a suitable paradigm to model such settings Despite the rising interest in this field, it has become difficult to compare or categorise approaches and identify the state-of-the-art solutions. Therefore, our first contribution is to develop a new taxonomy on the basis of the reward structures and utility functions, to offer a more structured view of the field. We note that utility functions are usually modelled as weights that define preferences over objectives, despite the fact that in many problems this assumption is not valid. We analyse the effect of non-linear utility functions on the set of equilibria in general multi-objective normal form games, under different optimisation criteria and look at how opponent modelling can aid the learning process in this setting. For future work, we are interested in how sequential settings can be approached under these considerations, to get a step closer to creating hybrid, artificial-and-human, multi-agent collectives that can deal with the different preferences w.r.t. the objectives of the different agents in the collective.

CCS CONCEPTS

• Theory of computation → Multi-agent reinforcement learning; **Convergence and learning in games**; **• Computing methodologies → Multi-agent systems**;

ACM Reference Format:
Roxana Rădulescu. 2020. A Utility-Based Perspective on Multi-Objective Multi-Agent Decision Making. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 2 pages.

1 INTRODUCTION

In multi-objective multi-agent systems (MOMAS) the reward signal for each agent is a vector, where each component represents the performance w.r.t. a different objective. Compromises between competing objectives should be made on the basis of the utility that these compromises have for the users. In other words, if we can define a utility function that maps the vector value of a compromise solution to a scalar utility, then we can derive what to optimise [3]. However, in many situations, knowing or applying the utility function in advance is not feasible. In such cases, we should search for a set of solutions that are optimal with respect to the space of all possible utility functions (i.e., a coverage set). Furthermore, in MAS, the structure of the reward function will also determine the nature of the system (e.g., collaborative, competitive), adding an additional complexity layer to how a problem should be approached.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

		Utility		
		Team	Social choice	Individual
Reward	Team	Coverage sets	Mechanism design	Coverage sets (+ Negotiation) Equilibria and stability concepts
	Individual		Mechanism design	Equilibria and stability concepts Coverage sets as best responses

Figure 1: Multi-objective multi-agent decision making taxonomy and mapping of solution concepts.

2 A STRUCTURED VIEW ON MOMAS

In single-agent multi-objective problems, the shape of the utility function, in conjunction with the allowed policy space, can be used to derive the optimal solution set [3]. In multi-agent settings, the situation is more complex, as the utility function may vary per agent. That is why we propose a taxonomy based on the *reward* as well as the *utility* functions [6]. We distinguish between two types of reward functions: a *team reward*, in which each agent receives the same reward vector, and *individual rewards* in which each agent receives a different reward vector. Furthermore, we make a distinction in three types of *utility*, i.e., *team utility*, which is what happens when all the agents serve the same interest, e.g., when they all work for a single company; *social choice utility*, when we are interested in optimising the overall social welfare across all agents; and *individual utility*, when each agent serves a different agenda and just tries to optimise for that. This results in the taxonomy provided in Figure 1. Furthermore, we note that the individual rewards with a team utility setting is not realistic; even if the utility function of all the individual agents would be the same, that would still lead to different individual utilities due to different rewards. Hence, we also treat this situation as *individual utilities*.

3 LEARNING IN MULTI-OBJECTIVE NORMAL FORM GAMES

In single-objective learning scenarios, agents are interested in optimising a return, i.e., a sum of reward signals received over the entire duration of the task. In the multi-objective case, when dealing with non-linear utility functions that should be applied to this return, we are faced with two options: i) to apply the utility function on the vectorial reward after each time step – expected scalarised return (ESR); or ii) to apply the utility function at the end of the task – scalarised expected return (SER) [3]. The choice between these two optimisation criteria depends on what is important to the user: the outcome after each interaction or over multiple interactions.

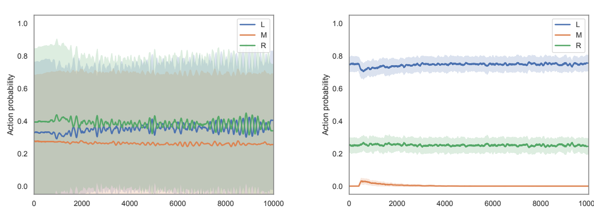
As a model of MOMAS we have studied multi-objective normal form games (MONFG) under the SER optimisation criterion with non-linear utility functions [5]. We introduce a new MONFG (Table 1) and show that while Nash equilibria (NE) [2] need not exist, correlated equilibria [1] can still be present when optimising with respect to a given signal (i.e., single-signal CE).

	L	M	R
L	$(4,0)$	$(3,1)$	$(2,2)$
M	$(3,1)$	$(2,2)$	$(1,3)$
R	$(2,2)$	$(1,3)$	$(0,4)$

	L	M	R
L	0	0.25	0
M	0	0	0
R	0	0.75	0

Table 1: (Im)balancing act game (left), together with a corresponding correlated equilibrium (right).

The correlated signal used for our experiments is presented in Table 1 (right). Notice how without any action recommendations (Figure 2a) the agent does not manage to converge to a stable strategy, while under the given correlation signal (Figure 2b), the agent follows the recommended actions without any incentive to deviate. We have also shown that using correlated equilibria allows us to reach higher SER for the players, compared to NE, a property also present in the single-objective setting.

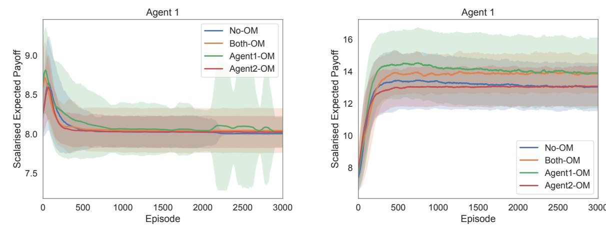


(a) No recommendations **(b) Single-signal CE**

Figure 2: Agent 1 – strategy evolution

3.1 Opponent Modelling

When the same multi-objective reward vector leads to different utilities for each user, it becomes essential for an agent to learn about the behaviour of other agents in the system. In Zhang et al. [7] we present the first study of the effects of opponent modelling (OM) on MONFGs with non-linear utilities, under the SER criteria. We contribute a novel actor-critic formulation to allow reinforcement learning of mixed strategies in this setting, along with an extension that incorporates opponent policy reconstruction using conditional action frequencies. We demonstrate that OM can alter the learning dynamics in this setting: when there are no NE, OM can have adverse effects on utility, or a neutral effect at best (Figure 3a); when equilibria are present, OM can confer significant benefits (Figure 3b).



(a) Game without NE under SER **(b) Game with NE under SER**

Figure 3: Agent 1 – SER, for different OM settings

4 CONCLUSION

We focus on general multi-objective multi-agent systems, without imposing any constraints on the reward or utility functions. Our first contribution is to built a taxonomy of what constitutes a solution for a multi-objective multi-agent decision problem based on reward structures and utility functions. We noted that many of the different settings we identify are under-explored in the current literature and would merit further investigation. We then explore multi-objective normal form games, where agents have distinct non-linear utility functions. We studied the set of equilibria in this setting and how properties of Nash and correlated equilibria translate from single- to multi-objective settings. Furthermore, we investigated how and if opponent modelling can aid agents obtain better outcomes in these cases.

For future work, we plan to expand this discussion to sequential decision making settings (e.g, traffic [4]), where an additional complexity level comes from having to also deal with state information and dependencies. This will be a step closer to achieving hybrid multi-agent collectives, that can deal with the different preferences w.r.t. the objectives of the different agents in the collective.

REFERENCES

[1] Robert J Aumann. 1974. Subjectivity and correlation in randomized strategies. *Journal of mathematical Economics* 1, 1 (1974), 67–96.
[2] John Nash. 1951. Non-Cooperative Games. *Annals of Mathematics* 54, 2 (1951), 286–295.
[3] Diederik M Roijers, Peter Vamplew, Shimon Whiteson, and Richard Dazeley. 2013. A survey of multi-objective sequential decision-making. *Journal of Artificial Intelligence Research* 48 (2013), 67–113.
[4] Roxana Rădulescu, Manon Legrand, Kyriakos Efthymiadis, Diederik M Roijers, and Ann Nowé. 2018. Deep Multi-Agent Reinforcement Learning in a Homogeneous Open Population. In *Proceedings of the 30th Benelux Conference on Artificial Intelligence (BNAIC 2018).* 177–191.
[5] Roxana Rădulescu, Patrick Mannion, Diederik M Roijers, and Ann Nowé. 2019. Equilibria in Multi-Objective Games: A Utility-based Perspective. In *Adaptive and Learning Agents workshop (at AAMAS 2019).*
[6] Roxana Rădulescu, Patrick Mannion, Diederik M Roijers, and Ann Nowé. 2020. Multi-objective multi-agent decision making: a utility-based analysis and survey. *Autonomous Agents and Multi-Agent Systems* 34 (2020). https://doi.org/10.1007/s10458-019-09433-x
[7] Yijie Zhang, Roxana Rădulescu, Patrick Mannion, Diederik M Roijers, and Ann Nowé. 2020. Opponent Modelling for Reinforcement Learning in Multi-Objective Normal Form Games. In *Proceedings of International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS 2020).* (accepted).