Multi-Objective Multi-Agent Decision Making: A Utility-based Analysis and Survey*

JAAMAS Track

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

Patrick Mannion
National University of Ireland Galway, Ireland
patrick.mannion@nuigalway.ie

Diederik M. Roijers
HU Univ. of Appl. Sci. Utrecht, The Netherlands

Ann Nowé
Vrije Universiteit Brussel, Belgium

ABSTRACT

Many real-world decision problems are inherently multi-objective in nature and concern multiple actors, making multi-objective multi-agent systems a key domain to study. We argue that trade-offs between conflicting objective functions should be analysed on the basis of the utility that these trade-offs have for the users of a system. We develop a new taxonomy which classifies multi-objective multi-agent decision making settings, on the basis of the reward structures and utility functions. We analyse which solution concepts apply to the different settings in our taxonomy, which allows us to offer a structured view of the field and identify promising directions for future research.

CCS CONCEPTS

• Computing methodologies → Multi-agent systems; **• Theory of computation → Sequential decision making**;

ACM Reference Format:
Roxana Rădulescu, Patrick Mannion, Diederik M. Roijers, and Ann Nowé. 2020. Multi-Objective Multi-Agent Decision Making: A Utility-based Analysis and Survey. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

The majority of multi-agent systems (MAS) implementations aim to optimise agent's policies with respect to a single objective, despite the fact that many real-world problems are inherently multi-objective in nature. Multi-objective optimisation [5] approaches consider these possibly conflicting objectives explicitly. In multi-objective multi-agent systems (MOMAS) the reward signal for each agent is a vector, where each component represents the performance on 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 [11], and how to measure the quality of solutions [13].

*This paper is an extended abstract of an article published in Autonomous Agents and Multi-Agent Systems [9]

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.

In this work, we focus exclusively on multi-objective approaches to decision making in MAS. We build a taxonomy of what constitutes a solution for a multi-objective multi-agent decision problem based on reward structures and utility functions. We note that many of the different settings we identify are under-explored in the current literature and would merit further investigation.

2 MOMAS TAXONOMY

Reward \ Utility	Team	Social choice	Individual
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.

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 that a multi-objective decision-theoretic algorithm should produce [10]. In multi-agent settings, the situation is more complex, as each individual agent can represent one or more distinct users. In other words, the utility function may vary per agent. That is why we propose a taxonomy based on the *reward* as well as the *utility* functions. We distinguish between two types of reward functions: a *team reward*, in which each agent receives the same value or return vector for executing the policy, and *individual rewards* in which each agent receives a different value/return vector. Furthermore, we make a distinction in three types of *utility*—more or less orthogonally to the types of rewards—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 or are on the same football team; *social choice utility*, when we are interested in optimising the overall social welfare across all agents; and *individual utility*, which is what happens if 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*.

Another factor we identify is the difference between the optimisation criteria: expected scalarised returns (ESR) and scalarised expected returns (SER) [10, 12]. This roughly distinguishes settings where either a single outcome (ESR) or the average outcome over multiple runs (SER) matters.

2.1 Solution Concepts

In the context of MAS, it is difficult to identify what constitutes an optimal behaviour, as the agents' strategies are interrelated. For this reason, we usually try to determine interesting groups of outcomes, i.e., solution concepts.

Coverage sets In single-agent multi-objective decision making, the optimal solution is called a coverage set [10]. A coverage set contains at least one optimal policy for each possible utility function. The team reward and team utility setting is a fully cooperative one where all rewards and the utility derived from that is shared between all agents. Therefore, there is only one true utility function in the execution phase, and the motivation for coverage sets being the right solution concept is the same as for single-agent multi-objective decision making. In a team reward but individual utility setting, coverage sets could be used if all agents will agree (e.g., through negotiation [4, 6]) that they will always execute a policy that is potentially optimal. Furthermore, in an individual reward and utility setting, a coverage set can also be a *set of possible best responses to the behaviours of the other agents*.

Equilibria and stability concepts In the individual utility scenario the utility derived by each agent from the received reward is different, regardless if this reward is the same or not for all the agents. This constitutes the most difficult scenario in our taxonomy. We consider that game theoretic equilibria (e.g., Nash [8] or correlated equilibria [1]) are suitable solution concepts, as we are dealing with decision making between self-interested agents.

Cooperative game theory studies settings where binding agreements among agents are possible. A central problem is therefore that of *coalition formation*, i.e., finding (sub)groups of agents that are willing to make such a binding agreement with each other. In the models in cooperative game theory, the utility for each agent is directly derived from the coalition the agents end up in, however, one can imagine that under the hood, the coalition works together cooperatively (based on their binding agreement) in a sequential decision problem that results in this utility. We further note, that the word cooperative does not imply team utility; typically, the agents will have their own utility functions. Hence, solution concepts from cooperative game theory apply to individual utility settings.

Mechanism design In game theory, the field of mechanism design takes the system's perspective for multi-agent decision problems: taking an original decision problem where the agents have individual reward functions that are unknown to the other agents and the "owner" of the game, as well as a social welfare function as input, the aim is to design a system of additional payments that would a) force the agents to be truthful about their individual utilities, and b) leads to solutions that are (approximately) optimal under the social welfare function. In multi-objective settings, the situation is more complex, as the individually received rewards determine the individual utilities via individual private utility functions. In general, it might even be very hard, or even impossible to articulate these functions, so being "truthful" about their utilities might be infeasible from the get-go. Nevertheless, it is possible to design mechanisms for some multi-objective multi-agent problems if the individual utilities can be articulated.

3 CONCLUSION AND NEW HORIZONS

In this paper, we analysed multi-objective multi-agent decision problems from a utility-based perspective. We hope that the taxonomy we build, together with the solution concept mapping, helps to place existing research papers in the larger multi-objective multi-agent decision problem context, and informs and helps to inspire further research.

In future work, it would be worthwhile to further explore the link between the multi-objective optimisation criteria ESR and SER, and solution concepts for MOMAS with non-linear utility functions.

In single-objective reinforcement learning (RL), an agent often aims to learn a model of the other agents' behaviours and uses this model when selecting or learning best responses. In multi-objective multi-agent settings, a good and possibly even sufficient predictor for this behaviour would be the utility function of the other agents. Therefore, explicitly estimating the utility functions of the other agents in a MOMAS is likely to be important in future research.

Interactive querying approaches, in which more information about the utility functions is actively pursued by querying the agents while planning or learning to limit the set of viable alternatives, could be helpful for mechanism design in settings where individual reward vectors are common knowledge, but agent preferences are (partially) unknown. For challenging real-world applications of MOMAS, it will be necessary to develop methods which consider continuous or high-dimensional state and action spaces. An important next step is therefore to extend existing deep RL methods for multi-objective multi-agent decision making settings.

Now that we have identified the different settings and solution concepts which are relevant to MOMAS, significant opportunities exist to revisit problems initially modelled as single-objective multi-agent decision problems using a multi-objective perspective. This could provide a richer set of potential solutions for cooperative MAS using the concept of coverage sets, or potentially improve performance by considering additional synthetic objectives which represent sub-tasks explicitly (i.e., multi-objectivisation [3]). One promising direction for future work is to use multi-objectivisation to improve team behaviour through social welfare. The possibility also exists to use MORL techniques to develop agents which may be tuned to adopt a range of different behaviours during deployment [7] in MAS (e.g. cooperative vs. competitive), or even creating populations of agents that develop effective behaviours against a large range of opponents [2].

REFERENCES

[1] Robert J Aumann. 1974. Subjectivity and correlation in randomized strategies. *Journal of Mathematical Economics* 1, 1 (1974), 67–96.

[2] David Balduzzi, Marta Garnelo, Yoram Bachrach, Wojciech Czarnecki, Julien Perolat, Max Jaderberg, and Thore Graepel. 2019. Open-ended learning in symmetric zero-sum games. In *International Conference on Machine Learning (ICML 2019)*. 434–443.

[3] Tim Brys, Anna Harutyunyan, Peter Vrancx, Ann Nowé, and Matthew E. Taylor. 2017. Multi-objectivization and ensembles of shapings in reinforcement learning. *Neurocomputing* 263 (2017), 48 – 59. Multiobjective Reinforcement Learning: Theory and Applications.

[4] Eugenio de Oliveira, Jose Manuel Fonseca, and Adolfo Steiger-Garção. 1999. Multi-criteria negotiation in multi-agent systems. ≪*KSSESE* (1999), 190.

[5] Kalyanmoy Deb. 2014. *Multi-objective Optimization*. Springer US, Boston, MA, 403–449. https://doi.org/10.1007/978-1-4614-6940-7_15

[6] Catholijn M Jonker, Reyhan Aydoğan, Tim Baarslag, Katsuhide Fujita, Takayuki Ito, and Koen Hindriks. 2017. Automated negotiating agents competition (anac). In *Thirty-First AAAI Conference on Artificial Intelligence*.

[7] Johan Källström and Fredrik Heintz. 2019. Tunable Dynamics in Agent-Based Simulation using Multi-Objective Reinforcement Learning. In *Proceedings of the Adaptive and Learning Agents Workshop (ALA-19) at AAMAS*.

[8] John Nash. 1951. Non-Cooperative Games. *Annals of Mathematics* 54, 2 (1951), 286–295.

[9] 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

[10] 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.

[11] Diederik M Roijers and Shimon Whiteson. 2017. Multi-objective decision making. *Synthesis Lectures on Artificial Intelligence and Machine Learning* 11, 1 (2017), 1–129.

[12] Yijie Zhang, Roxana Rădulescu, Patrick Mannion, Diederik 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).

[13] Luisa M Zintgraf, Timon V Kanters, Diederik M Roijers, Frans A Oliehoek, and Philipp Beau. 2015. Quality assessment of MORL algorithms: A utility-based approach. In *Benelearn 2015: Proceedings of the Twenty-Fourth Belgian-Dutch Conference on Machine Learning*.