Thesis Abstract: Communication In Multi-Objective Games*

Willem Röpke[0000−0001−5045−6127], Diederik M. Roijers[0000−0002−2825−2491], Roxana Rădulescu[0000−0003−1446−5514], and Ann Nowé[0000−0001−6346−4564]

Artificial Intelligence Lab, Vrije Universiteit Brussel, Belgium
{willem.ropke,diederik.roijers,roxana.radulescu,ann.nowe}@vub.be

Keywords: Multi-objective games · Nash equilibria · Reinforcement learning

Multi-objective games model scenarios in which decision makers are confronted with multiple, possibly conflicting, objectives and have many real-world applications, from environmental policy-making [4] to counterterrorism [12]. We study Multi-Objective Normal-Form Games (MONFGs) [1], which are normal-form games with vectorial payoffs. We take a utility-based approach [7], assuming a utility function for every player. Utility for mixed strategies can be defined in two distinct ways [6]. We may scalarise all payoff vectors a priori, which is known as the Expected Scalarised Returns (ESR) criterion. Alternatively, we can apply the Scalarised Expected Returns (SER) criterion by computing the utility from expected payoff vectors [5]. We first contribute two new theorems relating Nash equilibria under both criteria. Second, we contribute four novel communication protocols for learning agents in MONFGs.

1 Theoretical Contributions

We provide insight into the game-theoretic dynamics of multi-objective games, by analysing under what conditions the sets of Nash equilibria [3] under SER and ESR have similar properties or shared elements. We show our first major result in Theorem 1.

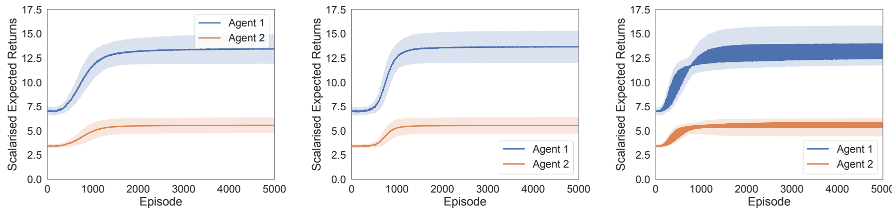
Theorem 1. *For a given MONFG, the sets of Nash equilibria under SER and ESR may have different cardinality and be disjoint.*

Next, we study necessary conditions for which preservation of Nash equilibria under both optimisation criteria can be guaranteed. Our main contribution is stated in Theorem 2.

Theorem 2. *For a given MONFG with convex utility functions, the sets of pure strategy Nash equilibria under SER and ESR are equal.*

We have since generalised this result to quasiconvex utility functions. Moreover, we have built upon these results to provide theoretical contributions in alternative settings and an algorithm which provably computes all pure strategy Nash equilibria in MONFGs given quasiconvex utility functions [10].

* Results from this thesis [8] have been published in [10,11].



(a) None (b) Cooperative action (c) Self-interested action

Fig. 1: Results for the communication protocols on a benchmark game.

2 Learning With Communication

For agents to learn adequate policies in multi-objective games, we contribute four novel communication protocols built upon the actor-critic framework [13]. Protocols were designed for cooperative settings, where agents aim to optimise a joint policy, and self-interested settings, where agents selfishly optimise their policies. Taking inspiration from Stackelberg games [2], we designate one agent as the leader and the other as the follower and switch these roles after every episode. We summarise each protocol below.

Cooperative Action Communication: The leader samples an action from their policy and commits to this publicly. The follower uses this commitment to update their own policy in the direction of a best-response. We find that players converge to their final joint policies faster and show this empirically in Fig. 1b.

Self-Interested Action Communication: The leader commits to an action according to a separate leader policy. The follower learns a best-response policy to each opponent action. This leads to the emergence of cyclic policies and cyclic equilibria [14], resulting in rapidly oscillating utility (Fig. 1c). Recently, we expanded on this both theoretically and algorithmically [9].

Cooperative Policy Communication: The leader commits to their policy, allowing the follower to condition their policy on this commitment. Results were similar to those for protocol 1.

Hierarchical Communication: Agents learn a policy consisting of two layers. In the top-layer, the leader determines whether to communicate or not. In the bottom-layer, either a no communication protocol or one of the prior communication protocols is followed. The follower responds with the same lower level policy. We find that learned policies depend on the characteristics of the underlying game as well as hyperparameters used while training

We hope that our contributions will further the applicability of multi-objective games in practice, as well as pave the way for further theoretical advancements. We also aim to extend our results to partially observable and sequential settings.

Acknowledgements WR is supported by FWO, grant number 1197622N. This research was supported by funding from the Flemish Government under the "Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen" program.

References

1. Blackwell, D.: An analog of the minimax theorem for vector payoffs. Pacific Journal of Mathematics **6**(1), 1–8 (1954). https://doi.org/10.2140/pjm.1956.6.1
2. Letchford, J., Korzhyk, D., Conitzer, V.: On the Value of Commitment. Autonomous Agents and Multi-Agent Systems **28**(6), 986–1016–986–1016 (Nov 2014). https://doi.org/10.1007/s10458-013-9246-9
3. Leyton-Brown, Yoav, K.a.S.: Essentials of Game Theory: A Concise, Multidisciplinary Introduction. Morgan and Claypool Publishers, 1st edn. (2008). https://doi.org/10.2200/s00108ed1v01y200802aim003
4. Martinez, E., Tazdaït, T., Tovar, E.: Participative democracy and local environmental issues. Ecological Economics **68**(1), 68–79 (Dec 2008). https://doi.org/10.1016/j.ecolecon.2008.01.025
5. Rădulescu, R., Mannion, P., Roijers, D.M., Nowé, A.: Multi-objective multi-agent decision making: A utility-based analysis and survey. Autonomous Agents and Multi-Agent Systems **34**(1), 10–10 (Apr 2020). https://doi.org/10.1007/s10458-019-09433-x
6. Rădulescu, R., Mannion, P., Zhang, Y., Roijers, D.M., Nowé, A.: A utility-based analysis of equilibria in multi-objective normal-form games. The Knowledge Engineering Review **35**, e32–e32 (2020). https://doi.org/10.1017/S0269888920000351
7. Roijers, D.M., Whiteson, S.: Multi-objective decision making. In: Synthesis Lectures on Artificial Intelligence and Machine Learning. vol. 34, pp. 129–129. Morgan and Claypool (2017). https://doi.org/10.2200/S00765ED1V01Y201704AIM034
8. Röpke, W.: Communication In Multi-Objective Games. Master thesis, Vrije Universiteit Brussel, Brussels (June 2021)
9. Röpke, W., Radulescu, R., Nowé, A., Roijers, D.M.: Commitment and Cyclic Strategies in Multi-Objective Games. In: Cruz, F., Hayes, C.F., da Silva, F.L., Santos, F.P. (eds.) Proceedings of the Adaptive and Learning Agents Workshop (ALA 2022). p. 9. Online, https://ala2022.github.io/ (May 2022)
10. Röpke, W., Roijers, D.M., Nowé, A., Rădulescu, R.: On nash equilibria in normal-form games with vectorial payoffs. Autonomous Agents and Multi-Agent Systems **36**(2), 53 (Oct 2022). https://doi.org/10.1007/s10458-022-09582-6
11. Röpke, W., Roijers, D.M., Nowé, A., Rădulescu, R.: Preference communication in multi-objective normal-form games. Neural Computing and Applications (Jul 2022). https://doi.org/10.1007/s00521-022-07533-6
12. Sawant, A., Dickerson, J.P., Hajiaghayi, M.T., Subrahmanian, V.S.: Automated generation of counterterrorism policies using multiexpert input. ACM Transactions on Intelligent Systems and Technology **6**(4) (Jul 2015). https://doi.org/10.1145/2716328
13. Sutton, R.S., Barto, A.G.: Reinforcement Learning: An Introduction. MIT press, Cambridge, MA, second edn. (2018)
14. Zinkevich, M., Greenwald, A., Littman, M.L.: Cyclic Equilibria in Markov Games. In: Proceedings of the 18th International Conference on Neural Information Processing Systems. pp. 1641–1648–1641–1648. MIT Press, Vancouver, British Columbia, Canada (2005)