Benchmarking Robustness and Generalization in Multi-Agent Systems: A Case Study on Neural MMO

Extended Abstract

Yangkun Chen[*1,3]
chen-yk21@mails.tsinghua.edu.cn

Joseph Suarez[*†2]
jsuarez@mit.edu

Junjie Zhang[*1,3]
zhangjj21@mails.tsinghua.edu.cn

Chenghui Yu[1,3]
ych20@mails.tsinghua.edu.cn

Bo Wu[3]
bowu@chaocanshu.ai

Hanmo Chen[1,3]
chm20@mails.tsinghua.edu.cn

Hengman Zhu[3]
henryzhu@chaocanshu.ai

Rui Du[4]
durui@bilibili.com

Shanliang Qian[4]
qianshanliang@bilibili.com

Shuai Liu[4]
liushuai01@bilibili.com

Weijun Hong[5]
hongweijun@corp.netease.com

Jinke He[6]
J.He-4@tudelft.nl

Yibing Zhang[7]
554011619@qq.com

Liang Zhao[8]
zhaoliang@idea.edu.cn

Clare Zhu[9]
czhu529@gmail.com

Julian Togelius[10]
julian.togelius@nyu.edu

Sharada Mohanty[11]
mohanty@aicrowd.com

Jiaxin Chen[†3]
jiaxinchen@chaocanshu.ai

Xiu Li[†1]
li.xiu@sz.tsinghua.edu.cn

Xiaolong Zhu[†3]
xiaolongzhu@chaocanshu.ai

Phillip Isola[2]
phillipi@mit.edu

ABSTRACT

We present the results of the second Neural MMO challenge, hosted at IJCAI 2022, which received 1600+ submissions. This competition targets robustness and generalization in multi-agent systems: participants train teams of agents to complete a multi-task objective against opponents not seen during training. We summarize the competition design and results and suggest that, considering our work as a case study, competitions are an effective approach to solving hard problems and establishing a solid benchmark for algorithms. We will open-source our benchmark including the environment wrapper, baselines, a visualization tool, and selected policies for further research.

KEYWORDS

Multi-agent Reinforcement Learning, Benchmark, Competition

ACM Reference Format:

Yangkun Chen[*1,3], Joseph Suarez[*†2], Junjie Zhang[*1,3], Chenghui Yu[1,3], Bo Wu[3], Hanmo Chen[1,3], Hengman Zhu[3], Rui Du[4], Shanliang Qian[4], Shuai Liu[4], Weijun Hong[5], Jinke He[6], Yibing Zhang[7], Liang Zhao[8], Clare Zhu[9], Julian Togelius[10], Sharada Mohanty[11], Jiaxin Chen[†3], Xiu Li[†1], Xiaolong Zhu[†3], and Phillip Isola[2]. 2023. Benchmarking Robustness and Generalization in Multi-Agent Systems: A Case Study on Neural MMO: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

*: Equal contribution. †: Corresponding author. 1: Shenzhen International Graduate School, Tsinghua University. 2: Massachusetts Institute of Technology. 3: Parametrix.ai. 4: bilibili. 5: NetEase Games AI Lab. 6: Delft University of Technology. 7: Chengdu Goldwin Electronics Technology Co., Ltd. 8: International Digital Economy Academy. 9: Stanford University. 10: New York University. 11: AICrowd. Correspondence to Jiaxin Chen (Parametrix.ai), Xiu Li(Shenzhen International Graduate School, Tsinghua University), Joseph Suarez (MIT), and Sharada Mohanty (AICrowd).

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

1 INTRODUCTION

Real-world applications of reinforcement learning (RL) require robust algorithms [2] that can adapt to dynamic environments. While substantially studied in single-agent RL [1, 5], this subject has been less explored in multi-agent systems. There is a distinct scarcity of multi-agent environments and supporting infrastructure.

This paper summarizes the IJCAI 2022 Neural MMO challenge and offers a solution to these problems. Neural MMO is a good environment to start with because it supports large-scale populations, is computationally efficient, and is actively maintained. On top of the environment, we built a large-scale parallel evaluation tool and a TrueSkill[3] rating system on the AICrowd platform. We hope that our methodology can serve as a stepping stone towards establishing more general benchmarks and promoting future research in Neural MMO and other multi-agent systems.

Our contributions are as follows: (1) Orchestration of the competition, which includes the environment, resources, the design of tracks, and the evaluation system. We believe this will be useful to guide future RL competitions. (2) Insights into emergent behaviors and strategies over 1600+ submissions. (3) Policy pool of 20 submitted policies to promote future research on Neural MMO, useful in evaluating policy robustness against a variety of opponents.

You can find links to this and previous AICrowd competitions on Neural MMO, documentation on the environment, source code, and our Discord community server at neuralmmo.github.io.

2 COMPETITION ORCHESTRATION

2.1 Environment

Neural MMO is an open-source research platform that simulates populations of agents in procedurally generated virtual worlds. Unlike other game genres typically used in research, MMOs simulate persistent worlds that support rich player interactions and a wider variety of progression strategies. We refer the reader to the original publication [6] for full information on Neural MMO and its objectives. Our environment is adapted from version 1.5 of Neural MMO with extra configuration to match the competition requirements.

2.2 Competition Structure

Participants are tasked with creating a team of 8 agents to win in our configuration of Neural MMO against 120 other opponents. The competition consists of two tracks: PvE track and PvP. The PvE track pits the participant's policy against 15 built-in opponents. Three increasing stages of difficulty serve as fixed reference points to help participants develop their policies. The main PvP track evaluates submissions against 15 other participants' policies in the same shared environments. This can better test a policy's robustness and generalization to opponents not seen during training.

2.3 Resources

We have created a number of resources for the participants' convenience: (1) A starter kit project containing all required segments to make a successful submission. With this guidance, new participants can make their first submission within 15 minutes. (2) An RL baseline implementation in a single file based on *TorchBeast* [4] to be used as a starting point. (3) Environment documentation and tutorials to help participants to get familiar with Neural MMO. (4) A light web-based replay viewer for our challenge, which allows participants with visual straightforward feedback for their policy development. The effect of our resources is shown in Fig 1.

3 SUMMARY AND ANALYSIS

The competition received over 40k views, 537 individual signups, 110 team signups, and 1679 submissions. This makes it one of the largest RL competitions to date. Of these participants, 48 teams were able to pass our first-round qualifier. 20 teams were able to win at least some games versus better policies that we trained for round 2, with 16 qualifying for round 3. We trained much stronger baselines for these rounds, but 7 teams were still able to win at least some games, and 6 were convincingly better than our best baseline. The best policies fully accomplished the task of the competition.

We categorized all 1600 submissions as rule-based methods (behavior tree, planning-based methods, heuristic methods, etc.) or learning-based methods (reinforcement learning) based on the algorithms employed by the participants. We have observed from

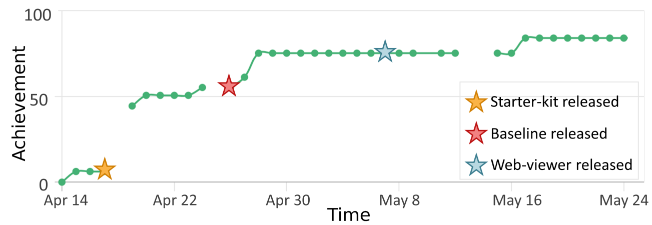

Figure 1: Maximum achievement in PvE stage 1 through time, measured over all participants. The release of the baseline corresponds with a large jump in submission quality.

Fig 2 that the Rule-based or learning-based method both achieve satisfactory performance. Two of the top five were rule-based, two were learning-based, and one was hybrid. Generally, the rule-based methods are quick to get working but do not scale as well against complex opponents. The orange lines of learning-based methods are all climbing, which means there is still room for further research even on this version of Neural MMO. We find that policies that achieve the same score may employ different strategies, and the models of different players can have varied strengths and weaknesses on the sub-tasks of the environment. We find that the performance of participants' policies vary during different stages, which further demonstrates that this environment may facilitate the study of model robustness and generalization by introducing diverse adversaries.

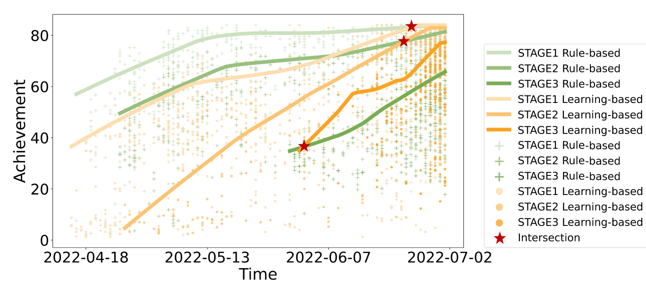

Figure 2: The effectiveness of Rule-Based and Learning-Based methods across three PvE stages. The six lines represent the peak performance of the approach at each stage.

4 CONCLUSION

To benchmark the robustness and generalization of MARL algorithms, we hosted a multi-agent artificial intelligence challenge and received 1600+ policy submissions. The top five submissions all surpassed the best existing baselines while employing strategies ranging from rule-based to full RL. We suggest that a gap in tooling and infrastructure, rather than purely algorithms, is the main short-term bottleneck preventing reinforcement learning from working on complex, multi-agent environments. We argue that the simplest way to realize this result in other environments is to run competitions and open-source the tools built by organizers and participants. We hope that our work will inspire others to adopt the competition model of research and open-source their tooling as we have.

ACKNOWLEDGMENTS

This work was supported in part by the Science and Technology Innovation 2030-Key Project under Grant 2021ZD0201404. We greatly appreciate the series of assistance that the AIcrowd platform has provided for the competition, including event promotion, participation system, and leaderboard. Without them, we would not have been able to complete our work.

REFERENCES

[1] Karl Cobbe, Christopher Hesse, Jacob Hilton, and John Schulman. 2020. Leveraging Procedural Generation to Benchmark Reinforcement Learning. In *Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event (Proceedings of Machine Learning Research, Vol. 119)*. PMLR, 2048–2056. http://proceedings.mlr.press/v119/cobbe20a.html

[2] Karl Cobbe, Oleg Klimov, Chris Hesse, Taehoon Kim, and John Schulman. 2019. Quantifying Generalization in Reinforcement Learning. In *Proceedings of the 36th International Conference on Machine Learning (Proceedings of Machine Learning Research, Vol. 97)*, Kamalika Chaudhuri and Ruslan Salakhutdinov (Eds.). PMLR, 1282–1289.

[3] Ralf Herbrich, Tom Minka, and Thore Graepel. 2006. TrueSkill™: a Bayesian skill rating system. *Advances in neural information processing systems* 19 (2006).

[4] Heinrich Küttler, Nantas Nardelli, Thibaut Lavril, Marco Selvatici, Viswanath Sivakumar, Tim Rocktäschel, and Edward Grefenstette. 2019. TorchBeast: A PyTorch Platform for Distributed RL. *CoRR* abs/1910.03552 (2019). arXiv:1910.03552 http://arxiv.org/abs/1910.03552

[5] Sharada P. Mohanty, Jyotish Poonganam, Adrien Gaidon, Andrey Kolobov, Blake Wulfe, Dipam Chakraborty, Grazvydas Semetulskis, João Schapke, Jonas Kubilius, Jurgis Pasukonis, Linas Klimas, Matthew J. Hausknecht, Patrick MacAlpine, Quang Nhat Tran, Thomas Tumiel, Xiaocheng Tang, Xinwei Chen, Christopher Hesse, Jacob Hilton, William Hebgen Guss, Sahika Genc, John Schulman, and Karl Cobbe. 2020. Measuring Sample Efficiency and Generalization in Reinforcement Learning Benchmarks: NeurIPS 2020 Procgen Benchmark. In *NeurIPS 2020 Competition and Demonstration Track, 6-12 December 2020, Virtual Event / Vancouver, BC, Canada (Proceedings of Machine Learning Research, Vol. 133)*, Hugo Jair Escalante and Katja Hofmann (Eds.). PMLR, 361–395.

[6] Joseph Suarez, Yilun Du, Clare Zhu, Igor Mordatch, and Phillip Isola. 2021. The Neural MMO Platform for Massively Multiagent Research. In *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks 1, NeurIPS Datasets and Benchmarks 2021, December 2021, virtual*, Joaquin Vanschoren and Sai-Kit Yeung (Eds.).