The NeurIPS 2023 Neural MMO Challenge
Multi-Task Reinforcement Learning and Curriculum Generation

Joseph Suárez JSUAREZ@MIT.EDU
Phillip Isola PHILLIPI@MIT.EDU
Massachusetts Institute of Technology

David Bloomin DAVEEY@GMAIL.COM
Kyoung Whan Choe CHOE.KYOUNG@GMAIL.COM
Hao Xiang Li HXL23@CAM.AC.UK
Ryan Sullivan RSULLI@UMD.EDU
Nishaanth Kanna NISHAANTHKANNA@GMAIL.COM
Daniel Scott DSCOTT45@GATECH.EDU
Rose S. Shuman ROSE.SHUMAN@ALUMNI.BROWN.EDU
Herbie Bradley HB574@CAM.AC.UK
Louis Castricato LOUIS_CASTRICATO@BROWN.EDU
CarperAI

Kirsty You KIRSTYYOU@CHAOCANSHU.AI
Bo Wu BOWU@CHAOCANSHU.AI
Yuhao Jiang YUHAOJIANG@CHAOCANSHU.AI
Qimai Li QIMAILI@CHAOCANSHU.AI
Jiaxin Chen JIAXINCHEN@CHAOCANSHU.AI
Xiaolong Zhu XIAOLONGZHU@CHAOCANSHU.AI
Parametrix.AI

Dipam Chakraborty DIPAM@AICROWD.COM
Sharada Mohanty MOHANTY@AICROWD.COM
AICrowd

Abstract

In this competition, we challenge participants to create teams of 8 agents to complete a variety of tasks involving foraging, combat, tool acquisition and usage, and item trading in Neural MMO 2.0, a simulated environment featuring 128 players and procedurally generated maps. The competition includes two tracks focused on multi-agent reinforcement learning and curriculum generation, respectively. This is the fourth challenge on Neural MMO, and the previous competitions have all yielded state-of-the-art performance on earlier versions of this environment as well as more general improvements to learning methods. We will provide the full source code of the environment, an easy to use starter kit, baselines for both tracks, and 200,000 A100 hours of GPU time for training and evaluating participants' submissions. Success in this competition would produce robust, task-conditional learning methods and establish a new benchmark for the same.

Preprint. Under review.

Keywords

game, multi-agent, reinforcement learning, tasks, curriculum

0.1 Background and impact

This is the 4th competition in a series of challenges on Neural MMO Suarez et al. [2021]. While previous competitions on Neural MMO have yielded important advancements to single-task learning, this competition will feature open-ended curriculum generation and generalization to new, language-grounded tasks. Accomplishments from previous competitions include architectures for processing structured observations of the environment, attention mechanisms for combining information across agents, and improvements to policy learning via variations on self-play. The second and third competitions were among the most popular on AICrowd, with a combined total of 3520 submissions from 1289 registered participants in 35 countries

This year, we have tripled our development team's size and have additionally partnered with Stability AI to provide training compute to participants. While the objective of the competition is to advance state-of-the-art reinforcement learning methods on a complex, multi-objective, many-agent environment, we have worked to simplify participation in the reinforcement learning track, making it comparable to conducting research on Atari Bellemare et al. [2012]. The curriculum generation track is new this year and does not require an AI background to participate.

The competition is hosted entirely within the simulated game environment of Neural MMO. Participants complete mixed cooperative and competitive tasks involving exploring the map, foraging for resources, engaging in strategic combat, training skills, collecting and trading items, and combinations thereof. The results of the competition will be applicable generally to task-conditional learning, including language-grounded learning, as well as multi-agent reinforcement learning and curriculum learning.

Scientific Questions: The reinforcement learning track investigates whether it is possible to learn robust, task-conditional policies Team et al. [2021] without a high-quality curriculum over the presented tasks Kanitscheider et al. [2021]. The curriculum generation track investigates whether it is possible to learn robust, task-conditional policies with standard PPO Schulman et al. [2017] purely through quality generation and ordering of tasks. The no-holds-barred track evaluates how much more can be accomplished with both simultaneously.

Social Impact: The Neural MMO team has invested significant resources to broaden the accessibility for participants across a wider range than historically in evidence at NeurIPS competitions. The curriculum generation track is designed in part to integrate underrepresented groups. The only requirement is a competent coding background and a passing interest in AI. The reinforcement learning track is accessible to participants at the level of a first graduate course in reinforcement learning. This represents a substantial milestone in accessibility, given that Neural MMO is arguably the most sophisticated many-agent simulation environment available that still receives active updates.

Industry Impact: The advancements anticipated in this competition are directly applicable to the $200b+ billion gaming market, but they are also more broadly useful in integrating RL with industry simulators, which are typically more complex than the environments popular in academic reinforcement learning research.

Marginalization: Neural MMO is an abstracted game simulator with no concepts of human social divisions. There are no anticipated negative impacts specific to Neural MMO beyond those associated with general technological progress in online reinforcement learning research which, unlike other areas of AI, does not rely on data obtained from or curated by humans.

Inclusiveness Plan: We provide computational resources to ensure fair access to all participants regardless of hardware ownership. The competition's grounding in gaming is likely to make it popular among a wide constituency across social groups. Neural MMO hosts a Discord server with over 800 members. Discord is the main venue for both gaming and open-source development groups. We will advertise the competition on several other such Discord channels, including those devoted to underrepresented affinity groups in AI.

Figure 1: Infographic providing a high level overview of the competition. To be used in promotion.

0.2 Novelty

The proposed competition would occur on Neural MMO 2.0, a new release of the environment comprising a full rewrite of the 6-year-old codebase yielding 3x efficiency improvements, a flexible task system to specify competition objectives, a performant and simple reinforcement learning baseline built using CleanRL [Huang et al., 2021] and supported by PufferLib, and dozens of bug fixes. The previous three competitions on Neural MMO were hosted on independently, at IJCAI 2022, and at NeurIPS 2022, respectively. Whereas the previous competition was on a multimodal but otherwise fixed task, the additions in Neural MMO 2.0 enable this competition to address task-conditional modeling as well as curriculum learning. The new track focuses specifically upon the generation and curation of tasks for efficient training and would not be computationally feasible on previous versions of the environment. Tasks are grounded by the text of their implementations Lehman et al. [2022], and a large language model will be included with the baselines.

0.3 Data

Neural MMO is a platform that simulates populations of agents in procedurally generated virtual worlds inspired by Massively Multiplayer Online games (MMOs). Agents need to forage for resources to survive and can train various skills, engage in strategic combat, acquire tools, equipment, and supplies, and trade with each other on a global market. See our documentation page at `https://neuralmmo.github.io` for full details on the environment.

Neural MMO, the associated baselines, and all tooling used are free and open-source software (FOSS). The only data released for the competition is generated from Neural MMO by a trained neural network and is likewise FOSS. The Neural MMO environment can simulate agents at roughly 5000x real-time per CPU core, producing decades of training data in days of compute. Participants will be given access to sufficient hardware to simulate the environment and train their models. Our evaluations consist of multiple stages; the final test sets will be generated specifically for this competition and have not been made available anywhere else.

0.4 Tasks and application scenarios

In this competition, participants will train a team of 8 agents. The objective is to generalize to maps, opponents, and even tasks not seen during training. While this objective remains the same across the competition, we offer three separate tracks, each focusing on a unique approach:

- **Reinforcement Learning Track**: Participants accomplish the objective by improving the reinforcement learning algorithm, also including the environment reward structure, observation featurization, and the neural network architecture. The presentation and sampling of tasks are provided by the baseline and are treated as constants. Each submission is trained for up to 8 A100 hours on our servers.
- **Curriculum Generation Track**: Participants accomplish the objective by improving the generation and sampling of tasks, also including the reward function assigned to the completion of each task. This track is fully new and will include a baseline that uses a large language model to ground task generation. The reinforcement learning algorithm, observation featurization, and neural network architecture are provided by the baseline and treated as constants. Each submission is trained for up to 8 A100 hours on our servers.
- **No Holds Barred Track**: Participants accomplish the objective however they like, with the exception of hacking our servers or modifying the game and other submissions in an unauthorized manner. We do not provide compute for this track, as it is mostly intended for larger organizations seeking to push the limits of Neural MMO, but we do not restrict compute usage either.

Through these three tracks, the competition addresses task-conditional reinforcement learning, grounded language, generalization to new maps, opponents, and tasks, efficient training curricula, and other such academic reinforcement learning problems. Neural MMO is much closer to an industry-scale simulation than typical arcade environments used in research, and this competition addresses the limited applicability of reinforcement learning outside of simple simulators. We also expect this competition to produce tooling advancements that will make learning easier in more complex settings.

The problem posed is a natural extension of previous competitions, which were simple enough for multiple participants to convincingly solve the task but challenging enough to leave room for improvement. In particular, the reinforcement learning track is substantially similar to the previous competition while the curriculum generation track is fully new and offers another axis of improvement. Similar competitions on other simulated environments are possible, but Neural MMO is unique in its combination of computation efficiency and problem complexity. This makes it well suited to a competition format with limited compute allocation to each participant.

0.5 Metrics

Participants' policies are evaluated on three separate rounds of held-out tasks:

1. **Random**: A set of tasks randomly sampled from the open-source, baseline task generator with a secret random seed. This is an in-distribution test.
2. **Hard Negative**: A set of tasks randomly sampled from the open-source, baseline task generator that the baseline policy failed to complete. This is a hard in-distribution test.
3. **Hand-crafted**: A set of tasks that we wrote to present interesting and high-level interpretable challenges. This is a hard and at least partially out of distribution test.

In all cases, evaluation takes place on randomly generated maps not seen during training. We will provide a fixed set of pretrained opponents for initial evaluations, but later in the competition, participants' policies will be evaluated against each other in the same shared environments. In the final evaluation, the top 16 submissions will compete to complete the most tasks against each other over at least 1000 randomized trials. The scoreboard will be ranked by number of completed tasks.

0.6 Baselines, code, and material provided

We will provide participants with the following materials:

- Comprehensive documentation on Neural MMO at neuralmmo.github.io written by a professional technical writer with consultation and correctness review by engineers on the team
- The full source code of Neural MMO, rewritten from the ground up for simplicity and speed.
- An open-source web app for rendering the environment, including the ability to view replays of past games and rewind/pause/fast-forward as well as pan, zoom, and view relevant statistics.
- A reinforcement learning baseline adapted from CleanRL Huang et al. [2022], a simple, widely used, rigorously tested library that provides a single-file PPO implementation with easily modifiable algorithmic details. Our version includes async sampling, hyperparameter and scale optimization for Neural MMO, historical self-play, and integration with the curriculum baseline below.
- A curriculum generation baseline integrated with the Syllabus library that uses a simple heuristic derived from the OpenAI multi-task curriculum paper on Minecraft to sample training tasks from a fixed set for the reinforcement learning baseline
- A curriculum generation baseline using Evolution through Large Models (ELM) Lehman et al. [2022], which leverages a language model to directly generate the code implementations of new tasks, coupled with the aforementioned heuristic to control the sampling of these tasks.
- A starter kit repository with a simple local set-up script, a prebuilt Docker, and a submission script to upload the participant's code to our evaluation servers

We will release the Neural MMO environment, visualizer, and baselines on GitHub. The environment is available as a pip package, and the visualizer is hosted on neuralmmo.github.io. Participants can choose between a local installation, which includes typical deep learning GPU driver setup and a pip installable requirements.txt, and an all-in-one Docker setup. We will release the starter kit and turn on our submission servers upon the launch of the competition. All other materials will be in open-source live beta testing in partnership with CarperAI and Parametrix.ai from now until launch.

0.7 Website, tutorial and documentation

The following documentation sites are relevant to participants:

- `https://neuralmmo.github.io` - the documentation page for Neural MMO, which has been professionally rewritten for the competition.
- `https://www.aicrowd.com/search?q=neural%20mmo` - the competition page on AICrowd, intended as a starting point for participants with links to all other relevant sites, the starter kit, support channel, and more. New competition page not yet live.
- `https://pufferai.github.io` - documentation for PufferLib, a middleware layer that substantially simplifies the usage of Neural MMO with existing reinforcement learning tools including CleanRL.
- `https://discord.gg/BkMmFUC` - the support and discussion server where participants can get help from the organizers

FAQ and Tutorial sections will be included in the documentation page. Several contact emails will be provided, but Discord will serve as our main support hub. Various references to Neural MMO publications, talks, and presentations are available on `https://neuralmmo.github.io`.

1 Organizational aspects

1.1 Protocol

AICrowd provides a competition platform with an online leader board. To join the competition, participants will create an AICrowd account, clone the starter project, run the setup to install dependencies, and submit their code to AICrowd via the provided script. Submissions will be evaluated on hardware provided by Stability AI. We will manually verify the code for all prize-eligible submissions at the end of the competition. If a submission is found not to adhere to our rules, the next best scoring submission will be selected as a winner instead. We have allocated the month of June (see above) for integration of our codebases with the competition platform. The AICrowd team and Parametrix are both organizers of previous Neural MMO competitions and have extensive familiarity with the required evaluation protocols.

1.2 Rules and Engagement

Our rules are available on the AICrowd page as a legal document and as summary bullet points. Users are required to agree before making a submission. These rules require that participants do not attempt to circumvent the restrictions put in place by the starter kit, which primarily ensure that advancements to reinforcement learning methods are the sole vector of improvement in the reinforcement learning tracks, advancements to curriculum methods are the sole vectored of improvement in the curriculum learning track, and sufficient freedom for other experimentation is allowed in the no holds barred track. Additionally, it outlines criteria for prize eligibility, including the open-sourcing of model and training code for the reinforcement learning and curriculum tracks.

These rules above enforces basic alignment of the participants' submissions with the scientific goals of each track and of the competition overall. In the first two tracks, we provide training and evaluation compute to participants. This ensures that participants will be minimally disadvantaged by lack of access to specialized hardware resources. The third track is intended for companies and large labs that want to push the bounds of what is possible in the environment without regard to any such restrictions. We have separated the track and prize structure to incentivize all such parties. Participants will have quick access to support via Discord. We will post any updates to the rules or deadlines here as well as on the AICrowd competition site.

1.3 Schedule and readiness

The proposed competition timeline is subject to approval by the NeurIPS organizers. We are flexible as to the start date, end date, and duration of the main competition. Note that the initial launch dates below have been pushed back roughly one week, following a delay in acceptance notification. Our current timeline is as follows:

- **April 27**: Competition submitted. Preliminary build of environment, baselines, and documentation.
- **May 24**: Competition accepted. Neural MMO 2.0 released.
- **June 8**: Baselines released and provided to Parametrix.ai and AICrowd for integration with the evaluation servers. Discord promotional partnerships secured.
- **July 8**: Final held out tasks provided to Parametrix.ai and AICrowd. Competition launches with beta versions of the documentation, environment, and baselines. Promotional material prepared and distributed via social media and partnered AI Discord servers.
- **July 8-31**: Warm-up rounds versus builtin baselines
  - **July 1-10**: Warm up round vs. open-source baseline model
  - **July 17-30** Warm up round vs. larger version of the baseline trained for longer
- **August 1 - October 31**: Main competition with evaluation against other participants' policies in staged rounds on the three sets of held-out tasks
  - **August**: Evaluation on randomly sampled tasks from the baseline distribution
  - **September**: Evaluation on hard negative tasks mined against the baseline
  - **October**: Evaluation on hand-crafted tasks
- **October 31**: Submissions close, final evaluation of top 16 policies
- **November 1-30**: Verification of solutions for integrity; communication with winners; preparation for NeurIPS.
- **December 10-16**: Presentation at NeurIPS
- **May 2024**: Estimated date of submission to the NeurIPS 2024 datasets and benchmarks track. Timeline pending formal announcement of the track.

As of June 8th, we have released a pip package for Neural MMO 2.0 and provided Parametrix.ai with a containerized development setup. The reinforcement learning baseline is implemented in CleanRL and runs but performs poorly. We likewise have initial integrations with OpenELM and Syllabus required for the curriculum baseline. We currently have 7 active developers and a technical writer and are confident in our ability to deliver the final materials well in advance of the deadlines we have set.

1.4 Competition promotion and incentives

We will promote our competition through the following vectors:

- Discord servers: Stability AI (177000 members), Learn AI Together (44000 members), Reinforcement Learning (4300 members) Eleuther AI (3600 members), CarperAI (850 members), Neural MMO (800 members), PufferAI (70 members), and several others
- AICrowd's mailing lists
- Professional or personal Twitter accounts of all willing authors
- The LIFE monthly open-source group meeting run by Joseph Suárez
- Lab meetings and university slack channels of student co-authors

Parametrix.AI will sponsor $20,000 US in prizes to be split across tracks. We may decide to offer additional prizes for special categories at our discretion. Winners will be invited to coauthor the summary report to be submitted to the NeurIPS 2024 Datasets and Benchmarks track. All authors who do not request to remain anonymous will be named.

2 Resources

2.1 Resources provided by organizers

Massachusetts Institute of Technology: Lead organizer of the competition and current home of the Neural MMO platform.

Carper AI / Stability AI: 200,000 A100 GPU hours, sufficient for development, training of baselines, and evaluation of up to 20,000 submissions. Our most recent competition received 1663 submissions. Limits will be put in place to prevent malicious waste of resources.

Parametrix.ai: Integrations and deployment engineering, competition orchestration, media, promotion and localization in China, $20k in prizes.

AICrowd: Engineering support for the AICrowd platform.

2.2 Support requested

The 2022 competition received minimal in-person attention due to the competition presentation room's distant location from the main hall. We humbly suggest that, similar to the datasets and benchmarks track, the competitions be given their own section at one of the poster presentation sessions, ideally with tables and outlet access for demos.

2.3 Organizing team

Our team comprises MIT, CarperAI, Parametrix.ai, and AICrowd personel and has representation from a diverse group of contributors spanning four countries, six ethnicities, and it includes female and LGBTQ members.

Massachusetts Institute of Technology is responsible for organizing the competition and coordinating efforts among co-organizers.

- **Joseph Suárez**: 5th year PhD candidate at MIT, the creator and lead developer of Neural MMO and PufferLib, and an organizer of all previous competitions. Joseph is leading the CarperAI development efforts for this competition and managing partnerships with AICrowd and Parametrix.ai.
- **Phillip Isola**: Associate professor in EECS at MIT best known for his work in generative AI. Phillip is Joseph's advisor and will act in an advisory capacity for the competition. Phillip has organized 11+ workshops, tutorials, and competitions at top conferences.

CarperAI is a democratized AI research team, enhancing the performance of preference learning. Carper AI is responsible for the development of Neural MMO 2.0, creation of the baselines and starter kit, and production of updated documentation.

- **David Bloomin**: A software engineer with experience at Google, Facebook, and Asana, where he contributed significantly to the development of Google Drive, Facebook's newsfeed architecture, and large-scale infrastructure systems. David is currently exploring multi-agent reinforcement learning (MARL) in complex social environments.
- **Kyoung Whan Choe**: PhD in cognitive neuroscience, published in journals such as Science Advances and Behavioral Research Methods. Kyoung focuses on the agent/behavioral aspects of Neural MMO.
- **Hao Xiang Li**: CS student at the University of Cambridge. A member of the technical team, Mark is working on building the infrastructure for Neural MMO curriculum design.
- **Ryan Sullivan**: 3rd year PhD student at the University of Maryland focusing on automatic curriculum learning. Ryan developed a curriculum learning library which will be used to create baselines for Neural MMO.
- **Nishaanth Kanna**: Master's Student at the University of New Brunswick. Nishaanth is a member of the technical team, working on integrating OpenELM with Neural MMO and generating diverse tasks.

- **Daniel Scott**: CS Master's student at Georgia Institute of Technology. Daniel is focusing on infrastructure and baselines for the Neural MMO competition curriculum generation track.
- **Rose S. Shuman**: Brown University MA. Rose is a technologist adept at converting complex material into clear concepts. She is developing the technical documentation for Neural MMO.
- **Herbie Bradley**: 2nd year PhD student at the University of Cambridge. Herbie led the development of OpenELM in Carper and helped to supervise the curriculum generation baseline with OpenELM.
- **Louis**: PhD student at Brown University, studying RLHF Mechanistic Interpretability. Louis is also a Co-founder and Team Lead of CarperAI, one of the largest open source RLHF groups.

Parametrix.AI is an AI focused start-up established in January 2019 in Shenzhen, China. Our mission is to create 'living AI beings' and build a virtual world where 1 billion people and 10 billion AI beings can live harmoniously together. In January 2022, the company announced a $100 million Series B investment from Sequoia China, 5Y Capital, and Gaorong Capital, reaching $1 billion valuations as an unicorn. Parametrix.AI was a co-organizer for the previous two Neural MMO Challenges and will be responsible for orchestration of the competition, distributed evaluation infrastructure, promotion in China, and integration with the AICrowd platform.

- **Kirsty You**: Product manager
- **Yuhao Jiang**: Machine learning researcher
- **Bo Wu**: Machine learning researcher
- **Qimai Li**: Senior machine learning researcher
- **Jiaxin Chen**: Senior machine learning researcher. Co-organizer of 3rd and 4th Neural MMO Challenge.
- **Xiaolong Zhu**: Senior R&D Director

AIcrowd is an AI crowdsourcing platform that enables data science experts and enthusiasts to collaboratively solve real-world problems, through challenges. AICrowd was a co-organizer for the previous three Neural MMO challenges and will be providing the platform for the competition.

1. **Dipam Chakraborty**: Engineer at AICrowd, point of contact for previous Neural MMO competitions
2. **Sharada Mohanty**: CEO and co-founder of AICrowd, coordinator of previous Neural MMO competitions

References

Marc G. Bellemare, Yavar Naddaf, Joel Veness, and Michael Bowling. The arcade learning environment: An evaluation platform for general agents. CoRR, abs/1207.4708, 2012. URL http://arxiv.org/abs/1207.4708.

Shengyi Huang, Rousslan Fernand Julien Dossa, Chang Ye, and Jeff Braga. Cleanrl: High-quality single-file implementations of deep reinforcement learning algorithms. CoRR, abs/2111.08819, 2021. URL https://arxiv.org/abs/2111.08819.

Shengyi Huang, Rousslan Fernand Julien Dossa, Chang Ye, Jeff Braga, Dipam Chakraborty, Kinal Mehta, and João G.M. Araújo. Cleanrl: High-quality single-file implementations of deep reinforcement learning algorithms. Journal of Machine Learning Research, 23(274):1–18, 2022. URL http://jmlr.org/papers/v23/21-1342.html.

Ingmar Kanitscheider, Joost Huizinga, David Farhi, William Hebgen Guss, Brandon Houghton, Raul Sampedro, Peter Zhokhov, Bowen Baker, Adrien Ecoffet, Jie Tang, Oleg Klimov, and Jeff Clune. Multi-task curriculum learning in a complex, visual, hard-exploration domain: Minecraft. CoRR, abs/2106.14876, 2021. URL https://arxiv.org/abs/2106.14876.

Joel Lehman, Jonathan Gordon, Shawn Jain, Kamal Ndousse, Cathy Yeh, and Kenneth O. Stanley. Evolution through large models, 2022.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. CoRR, abs/1707.06347, 2017. URL http://arxiv.org/abs/1707.06347.

Joseph Suarez, Yilun Du, Clare Zhu, Igor Mordatch, and Phillip Isola. The neural mmo platform for massively multiagent research. In J. Vanschoren and S. Yeung, editors, Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks, volume 1, 2021. URL https://datasets-benchmarks-proceedings.neurips.cc/paper/2021/file/44f683a84163b3523afe57c2e008bc8c-Paper-round1.pdf.

Open Ended Learning Team, Adam Stooke, Anuj Mahajan, Catarina Barros, Charlie Deck, Jakob Bauer, Jakub Sygnowski, Maja Trebacz, Max Jaderberg, Michaël Mathieu, Nat McAleese, Nathalie Bradley-Schmieg, Nathaniel Wong, Nicolas Porcel, Roberta Raileanu, Steph Hughes-Fitt, Valentin Dalibard, and Wojciech Marian Czarnecki. Open-ended learning leads to generally capable agents. CoRR, abs/2107.12808, 2021. URL https://arxiv.org/abs/2107.12808.