AI-Olympics: Exploring the Generalization of Agents through Open Competitions

Chen Wang[1] , **Yan Song**[1] , **Shuai Wu**[2] , **Sa Wu**[1] , **Ruizhi Zhang**[1] ,
Shu Lin[1] and **Haifeng Zhang**[1,2,3,*]

[1]Institute of Automation, Chinese Academy of Sciences

[2]Nanjing Artificial Intelligence Research of IA

[3]School of Artificial Intelligence, University of Chinese Academy of Sciences

{chen.wang, yan.song, sa.wu, ruizhi.zhang, chen.wang, shu.lin, haifeng.zhang}@ia.ac.cn,
wushuai@airia.cn

Abstract

Between 2021 and 2023, *AI-Olympics*—a series of online AI competitions—was hosted by the online evaluation platform Jidi in collaboration with the IJCAI committee. In these competitions, an agent is required to accomplish diverse sports tasks in a two-dimensional continuous world, while competing against an opponent. This paper provides a brief overview of the competition series and highlights notable findings. We aim to contribute insights to the field of multi-agent decision-making and explore the generalization of agents through engineering efforts.

1 Introduction

In recent years, AI-based decision-making has gained much attention in both academia and industries. Particularly, Reinforcement Learning (RL) has matured rapidly yielding high-level performance in various tasks in games [Schrittwieser *et al.*, 2019], robotics [Kober *et al.*, 2013], advertising [Cai *et al.*, 2017], etc. However, many of these carefully trained agents have been criticized for their poor generalization abilities when applied to a slightly difference task [Leibo *et al.*, 2021; Gorsane *et al.*, 2022; Agapiou *et al.*, 2022]. Implementing an adaptive agent that can perform equally well on a series of tasks is still an open and challenging problem. On the other hand, assessing an agent's generalization skills normally relies on benchmark suits encompassing various environments [Packer *et al.*, 2018; Kirk *et al.*, 2023], tasks [Leibo *et al.*, 2021; Oh *et al.*, 2017] and opponents [Leibo *et al.*, 2021; Agapiou *et al.*, 2022], which are often static and may prove inflexible amid the rapid evolution of agents.

To advance the research in this domain, we have developed a Python-based two-dimensional physical game engine *AI-Olympics* environment, accompanied by a variety of scenarios. Meanwhile, we innovatively conducted a series of AI competitions utilizing diverse scenarios created within the framework for a more generalized evaluation. In this paper, we provide a brief introduction to *AI-Olympics* environment and the accompanying scenarios, as well as the corresponding competition series and the evaluation of different
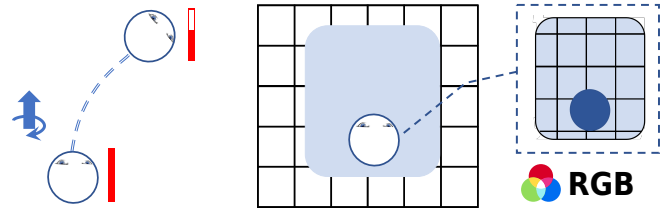
---

[*]Corresponding Author

Figure 1: An illustration of agent's mobility and vision. An agent consumes stored energy and moves, while only observing partially of its surroundings.

generalization skills. A demonstration video is available at https://youtu.be/SFXRe1JI6C8.

2 *AI-Olympics* Environment

AI-Olympics environment is a two-dimensional physical simulator built from scratch with minimal dependencies, utilizing Pygame [Shinde, 2021] for visualization. The design ensures flexible deployment and extensibility. The underlying engine replicates the physical dynamics of agents and elastic collisions between geometric shapes, including straight lines, curves, and circles. Leveraging this game engine, various scenarios and tasks can be generated, spanning different configurations. Three essential characteristics define the environment: (1) **continuous control with partial observation**, (2) **zero-sum**, and (3) **multi-tasking**.

2.1 Mobility and Observability

Each agent is depicted as an elastic circle on the map, equipped with the ability to exert impetus and torque, albeit at the cost of internal energy, enabling movement and interaction with the environment. Running out of energy results in a loss of control over the agent. Within the game environment, agents can collide and experience friction, thereby altering their states accordingly. As illustrated in Figure 1, the agent can perceive its surroundings within a limited range, allowing it to observe any other object within the map. Other objects on the map are represented as geometric shapes, each distinguished by its unique color to denote specific characteristics. As a result, we can construct various maps with recognizable items placed, thereby enhancing intelligent AI recognition and generalization across scenarios.
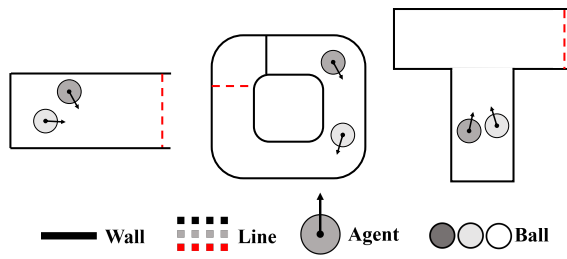
Figure 2: Three example maps in Running game. Agents are represented as circles and the arrow attached to them shows the driving force applied.

2.2 Zero-sum Games

Typically, each game in *AI-Olympics* contains two controllable agents with conflicting interests, rendering it a two-player zero-sum game extensively explored in prior research [Silver *et al.*, 2016; Silver *et al.*, 2017; Song *et al.*, 2024]. Both agents share the same mobility and observability for fair competition, and each can seek to gain an edge by obstructing pathways. Such uncertainty regarding the opponent's strategy during the evaluation phase can pose a significant challenge to the participants, encouraging them to consider generalization skills.

2.3 Multiple Designed Tasks

The game engine is crafted to generate diverse scenarios featuring common items on the map, allowing for an exploration of how agents can generalize their skills across tasks. We have developed six unique scenarios, each comprising a competitive sports task with different objectives and involving two agents with conflicting interests. These scenarios are:

Running game. Running game involves two agents exploring a maze and seeking to reach the goal line (finish line) as fast as possible. Figure 2 gives three example maps with common map elements. For example, the black line represents a sticky wall where the agent loses speed when collisions, and the red dashed line represents the goal line. The first agent that reaches the goal line wins the game.

Other single games. Five additional sports-related games are depicted in Figure 3. Wrestling game has the opposite objective compared to the Running game—each agent should keep itself away from the goal line (border line). In Curling game, players engage in a turn-based format where they strategically "throw" their rocks towards the central goal, aiming to position them as close to the center as possible. Both Table-Hockey and Football share similar map layouts, requiring players to score goals while also playing defensively. However, in Table-Hockey, agents are restricted to movement within their half of the field, while in Football, agents enjoy greater freedom of movement. In Billiard game, the objective of each agent is to pocket balls of the assigned color.

Integrated games. The integrated game combines 4–6 distinct scenarios into one unified task. Two players engage in consecutive games, with the victor of the majority of games
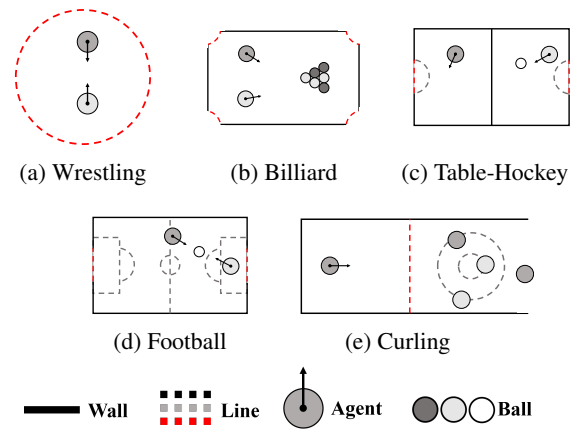

Figure 3: Five additional scenarios built from *AI-Olympics* game engine.

Contest	Scenarios	Participants
RLCN-2021-Winter	Running	145
RLCN-2022-Spring	Curling	103
RLCN-2022-Fall	Wrestling	103
RLCN-2022-Winter	Table-Hockey	89
RLCN-2023-Spring	Billiard	59
RLCN-2022-Summer	Integrated(4)	166
IJCAI-ECAI 2022	Integrated(4)	75
IJCAI 2023	Integrated(6)	58

Table 1: Between 2021 and 2023, *AI-Olympics* Competition Series has held eight online AI contests, among which six of them are preliminary and two are coordinated with the IJCAI committee. The competition series covers a different range of scenarios and has attracted numbers of participants.

as the series winner. While each game features unique objectives, they are underpinned by shared elements including agent movement control and map items. Consequently, a well-trained agent can adeptly identify critical elements within its environment, enabling efficient exploration and navigation to overcome obstacles.

3 *AI-Olympics* Competition Series

AI-Olympics competition series adopts the tasks outlined in Section 2.3 as themes and is evaluated by our online platform, **Jidi** (http://www.jidiai.cn). The platform offers live rankings and hosts AI competitions akin to Kaggle (https://www.kaggle.com/), specifically focusing on decision-making challenges. The objective of the competition series is to provide testbeds and promote practical engineering exploration of agents' generalization abilities. Notably, as shown in Table 1, we have conducted several preliminary contests, two of which were partnered with the IJCAI committee to gain further attention.

3.1 Evaluation Protocol

Platform Jidi offers online evaluation services for submitted agents in various simulated environments. The registered

users submit their code files on the website, then the platform performs back-end evaluation and updates the results on the front-end webpage. Jidi is able to provide real-time ranking which is helpful for participants to test their agents. For *AI-Olympics* competition series, since the games are all zero-sum, Jidi performs a *Swiss-system Tournament* for player pairing during the evaluation process, which has been used in various types of competitions including sports tournaments and academic events [Csató, 2013].

3.2 Generalization Abilities Assessment

During the preliminary competition, we utilize individual scenarios as testbeds and invite participation in each specific task. Additionally, we explore an integrated scenario wherein the agent engages in all games sequentially. Throughout the competition, agents are assessed for their generalization abilities from various perspectives.

Map generalization. In the Running game competition, agents are required to efficiently navigate the track map and reach the goal line (finish line) before their opponent. The competition spans multiple rounds of evaluation, with newly designed maps added to the pool in each round. This ensures that the submitted agents are tested not only on familiar maps but also on unseen layouts. These new maps feature identical elements, as depicted in Figure 2. A proficient agent must adeptly interpret the elements on the map and adjust its movement strategy accordingly, thereby demonstrating its generalization abilities across various maps.

Scenario generalization. In the integrated games competitions, agents are required to compete in multiple sequential sports scenarios, each characterized by a distinct layout and objective described in Section 2.3. To prevent the agent from memorizing the scenario sequence, we employ random shuffling of orders, forcing the agent to identify the current scenario during the evaluation phase. Furthermore, a future extension may involve the addition of new tasks in each evaluation round, placing more emphasis on the agent's ability to generalize across scenarios.

Opponent generalization. Throughout the competition series, the evaluation protocol requires the player not only solve multiple tasks but also compete with various opponents. Playing the same game with different opponents can yield disparate outcomes. Consequently, participants must take into account the strategies employed by other players. This can be achieved by reviewing replays of other players on the website and designing counter-strategies that can be generalized across various opponents.

4 Analysis of the Submitted Agents

During our competition series, some interesting findings echo our motivation as well as provide insights into the field.

Enhancing map generalization through increased data. In the Running game competition, at the beginning there are two maps in the candidate pool, then in subsequent rounds two additional maps are introduced into the pool, each with an increased probability of selection. As the competition progresses, we observe a consistent improvement in the mobility

and completion rates of participants' agents on newly introduced maps. This phenomenon is attributed to participants incorporating the new maps into their training, thereby expanding the breadth of their training data through techniques such as data augmentation [Shorten and Khoshgoftaar, 2019]. Consequently, this broader scope of training data contributes to the enhancement of agents' generalization skills over time.

Improving performance via scenario-specific strategies. Despite the random order of scenarios, many agents still manage to apply a targeted policy to each scenario in the integrated game. In some scenarios such as Curling, which demands more intricate skills, participants demonstrated a notable trend for adapting their strategies to the current game situation. For instance, some agents will stop in front of the goal line (hog line) and observe the situation first and then move backward to leave space for "throwing". This aligns with our motivation to run the competition which is to ask the agent to act adaptively in the game instead of reacting to experience.

Gaining advantages in competitive games through aggressive actions. The presence of opponents introduces varying degrees of challenge. In scenarios emphasizing competition, we observe certain agents exhibiting interference behaviors. For example, in some maps of Running game, agents actively engage in tactics such as (1) intentionally colliding with opponents to slow their progress by pushing them against walls, or (2) strategically maneuvering to obstruct the paths of others during turns.

Employing diverse AI-related methods. After discussing with several participants, we have gained valuable insights into the methodologies they used to tackle the tasks at hand, including:

- reinforcement learning with computer vision techniques;
- learning from demonstration to accelerate training;
- self-play framework for robustness;
- complex reward shaping to guide behavior;
- historical data summarization for long-term planning;
- curriculum learning to ease the training process;
- heuristic methods to fill up the opponent pool.

In practice, participants often utilized a combination of these methods when implementing their agent. In contrast to state-of-the-art algorithms in academia, the competition underscored the practicality of machine learning methods and offered valuable engineering insights to the field of game AI.

5 Conclusion

AI-Olympics competition series has provided a unique platform for participants to showcase their expertise in game AI. Through diverse challenges, participants have demonstrated their proficiency in addressing complex scenarios and adapting to dynamic environments. The competition's emphasis on practicality and agent generalization abilities underscores the complexity of game AI research nowadays and brings insights for further research endeavors. In the future, we will keep exploring the generalization skills of agents and continue to gather community support to tackle the problem.

Acknowledgments

References

[Agapiou *et al.*, 2022] John P Agapiou, Alexander Sasha Vezhnevets, Edgar A Duéñez-Guzmán, Jayd Matyas, Yi-ran Mao, Peter Sunehag, Raphael Köster, Udari Madhushani, Kavya Kopparapu, Ramona Comanescu, et al. Melting pot 2.0. *arXiv preprint arXiv:2211.13746*, 2022.

[Cai *et al.*, 2017] Han Cai, Kan Ren, Weinan Zhang, Kleanthis Malialis, Jun Wang, Yong Yu, and Defeng Guo. Real-time bidding by reinforcement learning in display advertising. *Proceedings of the Tenth ACM International Conference on Web Search and Data Mining*, 2017.

[Csató, 2013] László Csató. Ranking by pairwise comparisons for swiss-system tournaments. *Central European Journal of Operations Research*, 21:783–803, 2013.

[Gorsane *et al.*, 2022] Rihab Gorsane, Omayma Mahjoub, Ruan John de Kock, Roland Dubb, Siddarth Singh, and Arnu Pretorius. Towards a standardised performance evaluation protocol for cooperative marl. *Advances in Neural Information Processing Systems*, 35:5510–5521, 2022.

[Kirk *et al.*, 2023] Robert Kirk, Amy Zhang, Edward Grefenstette, and Tim Rocktäschel. A survey of zero-shot generalisation in deep reinforcement learning. *Journal of Artificial Intelligence Research*, 76:201–264, 2023.

[Kober *et al.*, 2013] Jens Kober, J. Andrew Bagnell, and Jan Peters. Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, 32:1238 – 1274, 2013.

[Leibo *et al.*, 2021] Joel Z Leibo, Edgar A Dueñez-Guzman, Alexander Vezhnevets, John P Agapiou, Peter Sunehag, Raphael Koster, Jayd Matyas, Charlie Beattie, Igor Mordatch, and Thore Graepel. Scalable evaluation of multi-agent reinforcement learning with melting pot. In *International conference on machine learning*, pages 6187–6199. PMLR, 2021.

[Oh *et al.*, 2017] Junhyuk Oh, Satinder Singh, Honglak Lee, and Pushmeet Kohli. Zero-shot task generalization with multi-task deep reinforcement learning. In *International Conference on Machine Learning*, pages 2661–2670. PMLR, 2017.

[Packer *et al.*, 2018] Charles Packer, Katelyn Gao, Jernej Kos, Philipp Krähenbühl, Vladlen Koltun, and Dawn Song. Assessing generalization in deep reinforcement learning. *arXiv preprint arXiv:1810.12282*, 2018.

[Schrittwieser *et al.*, 2019] Julian Schrittwieser, Ioannis Antonoglou, Thomas Hubert, Karen Simonyan, L. Sifre, Simon Schmitt, Arthur Guez, Edward Lockhart, Demis Hassabis, Thore Graepel, Timothy P. Lillicrap, and David Silver. Mastering atari, go, chess and shogi by planning with a learned model. *Nature*, 588 7839:604–609, 2019.

[Shinde, 2021] Piyush N Shinde. Pygame: Develop games using python. *International Journal for Research in Applied Science and Engineering Technology*, 2021.

[Shorten and Khoshgoftaar, 2019] Connor Shorten and Taghi M Khoshgoftaar. A survey on image data augmentation for deep learning. *Journal of big data*, 6(1):1–48, 2019.

[Silver *et al.*, 2016] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[Silver *et al.*, 2017] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *nature*, 550(7676):354–359, 2017.

[Song *et al.*, 2024] Yan Song, He Jiang, Zheng Tian, Haifeng Zhang, Yingping Zhang, Jiangcheng Zhu, Zonghong Dai, Weinan Zhang, and Jun Wang. An empirical study on google research football multi-agent scenarios. *Machine Intelligence Research*, pages 1–22, 2024.