The CoachAI Badminton Environment: Bridging the Gap between a Reinforcement Learning Environment and Real-World Badminton Games

Kuang-Da Wang, Yu-Tse Chen, Yu-Heng Lin, Wei-Yao Wang, Wen-Chih Peng

National Yang Ming Chiao Tung University, Hsinchu, Taiwan
gdwang.cs10@nycu.edu.tw, {s109550094.cs09, clementlin.mg08}@nycu.edu.tw, sf1638.cs05@nctu.edu.tw, wcpeng@cs.nycu.edu.tw

Abstract

We present the *CoachAI Badminton Environment*, a reinforcement learning (RL) environment tailored for AI-driven sports analytics. In contrast to traditional environments using rule-based opponents or simplistic physics-based randomness, our environment integrates authentic opponent AIs and realistic randomness derived from real-world matches data to bridge the performance gap encountered in real-game deployments. This novel feature enables RL agents to seamlessly adapt to genuine scenarios. The *CoachAI Badminton Environment* empowers researchers to validate strategies in intricate real-world settings, offering: i) Realistic opponent simulation for RL training; ii) Visualizations for evaluation; and iii) Performance benchmarks for assessing agent capabilities. By bridging the RL environment with actual badminton games, our environment is able to advance the discovery of winning strategies for players. Our code is available at https://github.com/wywyWang/CoachAI-Projects/tree/main/CoachAI%20Badminton%20Environment.

Introduction

The rising interest in sports analytics has triggered a surge in research (Kao et al. 2022; Wang et al. 2022b), with a focus on leveraging reinforcement learning to enhance player strategies (Won, Gopinath, and Hodgins 2021; Chen et al. 2023). Simulation environments with automatic opponents are critical to swiftly evaluate the designed algorithms; however, previous efforts mainly centered on physics-based interactions or simple rule-based opponents, e.g., Brockman et al. (2016); Kurach et al. (2020). While some prior work attempted to make opponents more realistic through imitation learning (Won, Gopinath, and Hodgins 2021), no existing environments are tailored to turn-based sports that provide authentic opponents. Therefore, we introduce the *CoachAI Badminton Environment*, an RL environment with simulated opponents for badminton, a quintessential turn-based sport. However, two challenges naturally arise in creating an authentic badminton environment: **1) Opponent-Driven Transition Dynamics:** Opponents' actions influence the decisions of agents due to the alternating-return characteristics in turn-based sports. It is important to address such behaviors in terms of transition dynamics. **2) Dynamic Strategies and Randomness:** Effective opponents should adapt their behavior and randomness to prevent being easily seen through, reflecting real-world scenarios for training agents.

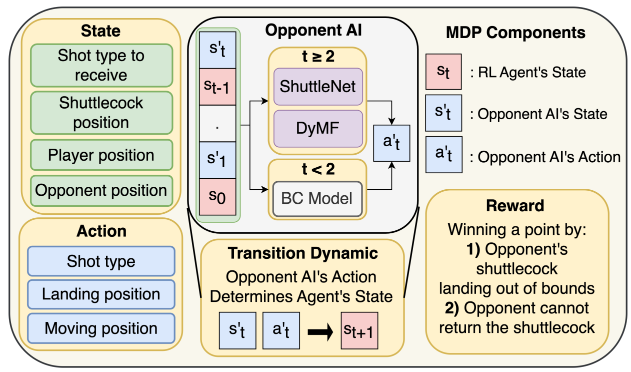

Figure 1: An overview of the *CoachAI Badminton Environment*.

To address these challenges, we integrate two cutting-edge models: **ShuttleNet** (Wang et al. 2022a), a transformer-based stroke forecasting model, and **DyMF** (Chang, Wang, and Peng 2023), a graph-based framework for predicting player movements. This integration reinforces the opponent AI's capability of generating a fine-grained range of actions, encompassing shot types, landing positions, and moving positions. These models are trained on the extensive dataset **ShuttleSet** (Wang et al. 2023), the largest publicly available collection of annotated badminton singles matches, allowing us to capture realistic player strategies from high-ranking players. To enhance the realism, we seamlessly integrate the opponent AI into the only existing badminton environment that simulates shuttlecock trajectories based on the physical parameters of the shuttlecock and players (Huang et al. 2023). To support comprehensive evaluations, we expand the simulation scope from focusing solely on rallies to encompassing entire matches that follow real-world badminton rules for singles matches (Badminton World Federation 2021). These efforts effectively transform our environment into an accurate representation of real-world badminton scenarios. Additionally, we emphasize strategy visualization, shedding light on factors influencing point losses and model disparities, thereby enhancing the understanding

```
import CoachAI.env as badminton_env
# Users can select a specified opponent
env = badminton_env(opponent="Tzu-Ying Tai")
env.reset()
done = False
while not done:
    action = env.action_space.sample()
    next_state, reward, done, info = env.step(action)
```

Figure 2: Users can choose specific opponents for RL.

of strategies. For a more detailed illustration, please refer to our demonstration here[1].

Environment Description

In Figure 1, we show the overview of our proposed environment. Our environment embraces real-world badminton singles match rules, modeling each game as a Markov Decision Process (MDP) (Puterman 2014). This design establishes a badminton environment with three core features:

Opponent AI as Transition Dynamics

Opponent AI serves as the transition dynamic in RL environments for turn-based sports, directly determining the RL agent's next state. To create realistic opponent AIs, we integrate two advanced models: ShuttleNet and DyMF. ShuttleNet predicts shot types and landing positions, while DyMF predicts player movement positions. A limitation of this design is that both models rely on historical strokes, requiring at least two previous strokes. Therefore, we introduce the Behavior Cloning (BC) model, which interacts with the RL agent when historical strokes are insufficient (i.e., fewer than two strokes), facilitating interaction throughout the rally. We integrate these models to enable effective interaction between the opponent AI and the RL agent, resulting in applying trained agents to real-world scenarios.

MDP Modeling & API Integration

Our environment models each badminton game as an MDP, with components depicted in Figure 1. Our environment's API inherits from the widely used OpenAI Gym API (Brockman et al. 2016) to support out-of-the-box compatibility. As shown in Figure 2, we provide example codes for running a random agent against a specified opponent AI.

Evaluation & Visualization

To visualize a clear understanding of the agent's performance, our environment emphasizes two essential functions, including[2]:

- **Animated Simulation:** Obtain a thorough game scenario understanding with our interactive animation. It vividly displays landing and movement position distributions and provides replay functionality via the scroll bar. Real-time updates of shot type probabilities are presented in a bar chart during matches.
- **Error Statistics:** Calculate proportions between two losing action types: hitting the shuttlecock out of bounds and hitting the net, presented as a pie chart. Furthermore, we offer discrete court representations, segmenting it into 9 internal and 1 external regions, creating 10 distinct positional blocks. This discretization aids in identifying the reasons behind point losses. We provide a table of the top four states where losing actions are most likely, and during rallies, our environment displays landing and moving distribution for the most frequent states resulting in point losses. Users can manually select a state to observe its associated distribution.

[1] https://youtu.be/EsIXpBvcvzA

[2] For a detailed demonstration of how these functions operate in the environment, please refer to our demo video.

Demonstration Overview

Two scenarios are shown to showcase the usage of *CoachAI Badminton Environment* to enhance playing strategies:
RL Algorithms Benchmarks. We benchmark two widely used RL algorithms, Advantage Actor-Critic (A2C) (Mnih et al. 2016), and Proximal Policy Optimization (PPO) (Schulman et al. 2017), by competing against two different opponent AIs in 20 badminton sets, reporting average score differences. Figure 3 shows that these algorithms, when facing the same opponents, produce varying performance. This result highlights the impact of algorithm selection on strategy evolution. **RL-Driven Winning Strategy.** We showcase how our RL-trained agent, powered by our opponent AI, identified strategic improvement in a badminton match. Specifically, we observed that for a state prone to point losses, as the PPO agent accumulated more training steps, there was a shift in the moving distribution after taking serves from favoring the left court side to shifting to the right, as illustrated in Figure 4. This adjustment led to a 60% increase in the agent's win rate against opponent A.

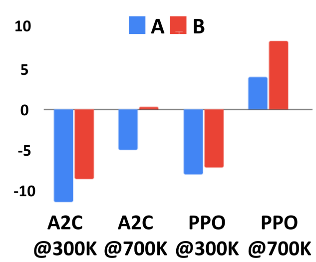

Figure 3: Average score difference for A2C and PPO.

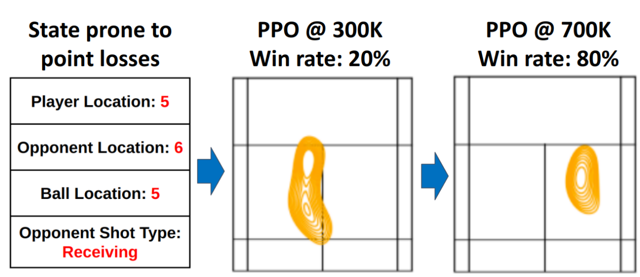

Figure 4: The enhancement of win rate via improvement in moving distribution.

Acknowledgments

This work was supported by the Ministry of Science and Technology of Taiwan under Grants 112-2425-H-A49-001.

References

Badminton World Federation. 2021. The Laws of Badminton. https://www.worldbadminton.com/rules/. Accessed: 2021-05-23.

Brockman, G.; Cheung, V.; Pettersson, L.; Schneider, J.; Schulman, J.; Tang, J.; and Zaremba, W. 2016. OpenAI Gym. *CoRR*, abs/1606.01540.

Chang, K.; Wang, W.; and Peng, W. 2023. Where Will Players Move Next? Dynamic Graphs and Hierarchical Fusion for Movement Forecasting in Badminton. In *AAAI*, 6998–7005. AAAI Press.

Chen, X.; Wang, W.; Hu, Z.; Chou, C.; Hoang, L.; Jin, K.; Liu, M.; Brantingham, P. J.; and Wang, W. 2023. Professional Basketball Player Behavior Synthesis via Planning with Diffusion. *CoRR*, abs/2306.04090.

Huang, L.; Hseuh, N.; Chien, Y.; Wang, W.; Wang, K.; and Peng, W. 2023. A Reinforcement Learning Badminton Environment for Simulating Player Tactics (Student Abstract). In *AAAI*, 16232–16233. AAAI Press.

Kao, T.-Y.; Pan, T.-Y.; Chen, C.-N.; Tsai, T.-H.; Chu, H.-K.; and Hu, M.-C. 2022. ScoreActuary: Hoop-Centric Trajectory-Aware Network for Fine-Grained Basketball Shot Analysis. In *Proceedings of the 30th ACM International Conference on Multimedia*, MM '22, 6991–6993. New York, NY, USA: Association for Computing Machinery. ISBN 9781450392037.

Kurach, K.; Raichuk, A.; Stanczyk, P.; Zajac, M.; Bachem, O.; Espeholt, L.; Riquelme, C.; Vincent, D.; Michalski, M.; Bousquet, O.; and Gelly, S. 2020. Google Research Football: A Novel Reinforcement Learning Environment. In *AAAI*, 4501–4510. AAAI Press.

Mnih, V.; Badia, A. P.; Mirza, M.; Graves, A.; Lillicrap, T. P.; Harley, T.; Silver, D.; and Kavukcuoglu, K. 2016. Asynchronous Methods for Deep Reinforcement Learning. In Balcan, M.; and Weinberger, K. Q., eds., *Proceedings of the 33nd International Conference on Machine Learning, ICML 2016, New York City, NY, USA, June 19-24, 2016*, volume 48 of *JMLR Workshop and Conference Proceedings*, 1928–1937. JMLR.org.

Puterman, M. L. 2014. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

Schulman, J.; Wolski, F.; Dhariwal, P.; Radford, A.; and Klimov, O. 2017. Proximal Policy Optimization Algorithms. *CoRR*, abs/1707.06347.

Wang, W.; Huang, Y.; Ik, T.; and Peng, W. 2023. ShuttleSet: A Human-Annotated Stroke-Level Singles Dataset for Badminton Tactical Analysis. In Singh, A. K.; Sun, Y.; Akoglu, L.; Gunopulos, D.; Yan, X.; Kumar, R.; Ozcan, F.; and Ye, J., eds., *Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, KDD 2023, Long Beach, CA, USA, August 6-10, 2023*, 5126–5136. ACM.

Wang, W.; Shuai, H.; Chang, K.; and Peng, W. 2022a. ShuttleNet: Position-Aware Fusion of Rally Progress and Player Styles for Stroke Forecasting in Badminton. In *AAAI*, 4219–4227. AAAI Press.

Wang, W.-Y.; Chan, T.-F.; Peng, W.-C.; Yang, H.-K.; Wang, C.-C.; and Fan, Y.-C. 2022b. How Is the Stroke? Inferring Shot Influence in Badminton Matches via Long Short-Term Dependencies. *ACM Trans. Intell. Syst. Technol.*, 14(1).

Won, J.; Gopinath, D.; and Hodgins, J. K. 2021. Control strategies for physically simulated characters performing two-player competitive sports. *ACM Trans. Graph.*, 40(4): 146:1–146:11.