

# Competitive Physical Human-Robot Game Play

Boling Yang  
University of Washington  
bolingy@cs.washington.edu

Golnaz Habibi  
Massachusetts Institute of Technology  
golnaz.habibi@gmail.com

Xiangyu Xie  
University of Washington  
shawnxie@uw.edu

Joshua R. Smith  
University of Washington  
jrs@cs.washington.edu

## ABSTRACT

While competitive games have been studied extensively in the AI community for benchmarking purposes, there has only been limited discussion of human interaction with embodied agents under competitive settings. In this work, we aim to motivate research in competitive human-robot interaction (competitive-HRI) by discussing how human users can benefit from robot competitors. We then examine the concepts from game AI that we can adopt for competitive-HRI. Based on these discussions, we propose a robotic system that is designed to support future competitive-HRI research. A human-robot fencing game is also proposed to evaluate a robot's capability in competitive-HRI scenarios. Finally, we present the initial experimental results and discuss possible future research directions.

## CCS CONCEPTS

• **Human-centered computing** → *Systems and tools for interaction design*; • **Computing methodologies** → *Markov decision processes*; • **Computer systems organization** → *Robotics*.

## KEYWORDS

human-robot interaction, competitive-HRI, reinforcement learning, artificial intelligent for games

### ACM Reference Format:

Boling Yang, Xiangyu Xie, Golnaz Habibi, and Joshua R. Smith. 2021. Competitive Physical Human-Robot Game Play. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21 Companion)*, March 8–11, 2021, Boulder, CO, USA. ACM, New York, NY, USA, 5 pages. <https://doi.org/10.1145/3434074.3447168>

## 1 INTRODUCTION

Competition is one of the most common forms of human interaction, yet *competitive interaction* has rarely been discussed in the context of Human Robot Interaction (HRI). There has indeed been a large focus in HRI on *cooperative interaction*, such as human-aware motion planning, object handover actions, and collaborative manipulation

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).  
*HRI '21 Companion, March 8–11, 2021, Boulder, CO, USA*

© 2021 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
ACM ISBN 978-1-4503-8290-8/21/03...\$15.00  
<https://doi.org/10.1145/3434074.3447168>



**Figure 1:** A robot-based athletic training system based on a PR2 robot and a VIVE VR system. The human and the robot are playing a competitive fencing game. Detailed rules of the game can be found under the System Design & Implementation section. The upper figure shows the gameplay in the VR environment, where the left sub-window shows the game from a third-person perspective, and the right sub-window shows the human's perspective when wearing the VR headset. The lower figure shows the game in actuality. Please refer to this link for example gameplay videos.

[2, 4, 5, 16]. Conversely, the absence of studies in competitive robot interaction may be due to anxieties concerning the actions of robots whose interest do not necessarily align with our own. However, these fears should not prohibit us from considering positive impacts that competitive-HRI can yield, such as providing the participants with motivation, inspiring their potential, and more.

We initiate our research project in competitive-HRI by focusing on how to create a competitive robotic agent that can challenge human users in athletic performance and physical exercise. Physical exercise is essential to our physical and mental health. We hypothesize that a robot with adversarial behaviors can provide athletic practice or exercise sessions that are more personalized, effective, and enjoyable. We believe that athletic practice and physical exercise are scenarios in which nearly anyone can directly benefit from competitive-HRI. In this paper, we first motivate competitive-HRI by discussing psychological studies about how competition can influence participants in a positive manner and how physical exercise is more efficient and motivating when an appropriate amount of competition is introduced. We then survey the literature of reinforcement learning in competitive games and related works in robotic. Next, the design and implementation of our competitive-HRI robotic system are introduced, along with a competitive fencing game we designed to evaluate our system. Finally, we will present initial experimental results and discuss the insights drawn from the results.

## 2 COMPETITION

A number of psychological research studies indicate that competition can effectively increase participants' performance, motivation, and ability to learn in various scenarios. Plass et al. [20] compared how individual, cooperative, and competitive game modes affect learning, performance, and motivation in an educational mathematics video game. They found that players in both competitive and cooperative conditions solve significantly more problems during the game than those in the individual condition. Furthermore, competitive players may have developed better arithmetic skills in the game, such that they were able to complete more problems than those in other conditions. The quality of problem-solving strategies was also affected by different game modes, where competitive and individual players were less likely to employ inefficient problem-solving strategies. The study also found that competitive and collaborative games enhance enjoyment and situational interest. Higher enjoyment refers to a participant having a higher intrinsic desire and tendency to engage in an event [21]. Furthermore, higher situational interest refers to an event that is able to elicit more attention and effective reactions from the participants [10].

The effect of competition in physical exercise and athletic training has also been well studied. Feltz et al. [7] created a virtual exercise partner that always slightly outperforms the human participant in cycling or holding a plank exercise. Without knowing it was a manipulated virtual peer, the participants felt less capable than the peer and showed performance improvement during the study. Viru et al. [31] explored the mechanisms of how exercise performance can be enhanced under competitive conditions. They found that, in a treadmill running test, the running duration was prolonged by an average of 4.2% in competition, and was accompanied by a significantly greater peak  $\dot{V}O_2$  (maximum rate of oxygen consumption) response. Inspired by these studies, this project aims to create a robotic exercise partner that is able to challenge a human user.

## 3 RELATED WORK

### 3.1 Reinforcement Learning for Games

Competitive games like Chess, Checkers, and Go have been studied extensively by the AI researchers for the purpose of benchmarking the ability of algorithms to train an agent to learn, reason, and plan [3, 23, 27]. Many algorithms that solve multi-agent Markov games are designed under a multi-agent reinforcement learning scheme, where agents develop emergent and complex behavior through interacting with each other and co-evolving together [8, 9, 11, 26, 28, 30]. The following works have successfully extended game RL to learn complex motor skills. Pinto et al. [19] proposed an adversarial training method that can be applied to most existing RL algorithms. Through the use of PPO and a very large training batch size, Bansal et al. [1] created humanoid and quadrupedal agents that play physically competitive games like soccer, wrestling, and more in simulation. Lowe et al. [15] proposed a centralized action-value function that takes the actions of all agents as input in order to stabilize the DDPG algorithm [14] in multi-agent settings. In this paper, we formulate competitive-HRI tasks as a multi-agent Markov games problem, and we will evaluate two of these approaches and compare their strengths and weaknesses under our competitive-HRI environment.

### 3.2 Competitive and Physical Training Robots

There has been very limited human-robot interaction research that explores competitive interactions. Kshirsagar et al. [12] studied how a human's performance was effected by a robot "co-worker" working in the same workspace when competing for a real monetary prize. Human participants were slightly discouraged when competing against a high performance robot. Another observation was that people would hold a more positive attitude toward a robot with lower performance. Mutlu et al. [18] compared the perceptions of an ASIMO robot when it was playing a video game cooperatively or competitively with human participants. Their results suggested that male participants were more engaged by competitive gameplay, but the cooperative agent was more socially desired. Short et al. [25] found that, when a robot cheated in a "rock-paper-scissors" game, human participants had a greater degree of social engagement and made greater attributions of mental state during the game.

On the other hand, there have been some attempts to use robotic systems to assist in physical exercise. Fasola and Mataric [6] developed a socially assistive robot to provide real-time coaching and encouragement for a seated arm exercise. Most participants did find the robot to be helpful in their exercise and considered it to be an exercise instructor. Süssenbach et al. [29] created an interactive action-based motivation model for an indoor-cycling activity. In order to motivate the user throughout the exercise section, the robot employed communication strategies according to the user's physical state and condition. Their system successfully increased users' workout efficiency and intensity. Sato et al. [22] created a robotic system to assist the training of volleyball players. The system has sufficient mechanical capability to imitate the motion and strategy of top volleyball blockers.

Each of these studies focused on one or a very small set of competitive scenarios. In addition, most of these scenarios only require very limited and simple robot motions. In this work, we examine

the concepts from game AI that we can adopt for competitive-HRI. Furthermore, we design a robotic system that can potentially play various physically competitive games against a human player.

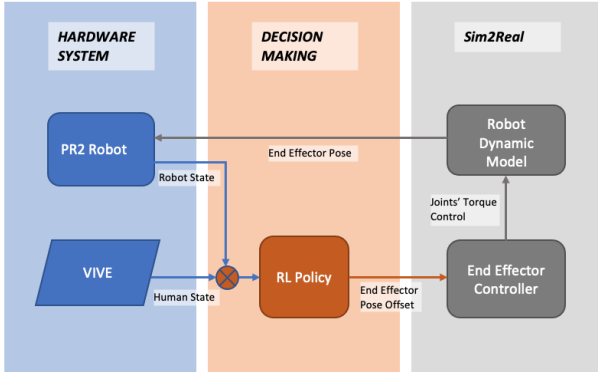


Figure 2: A block diagram demonstrating the pipeline of the proposed robotic system.

## 4 SYSTEM DESIGN & IMPLEMENTATION

We created a robotic system to explore whether a robot can effectively improve human athletic training under competitive conditions. In this section, we will introduce our competitive-HRI robotic system. A preview of the system pipeline is shown in Fig.2.

### 4.1 Evaluation Environment

A zero-sum fencing game environment was created to train and evaluate our agents. It is an attack and defense game where the human is the antagonist and the robot is the protagonist. A screenshot of the fencing game in the Mujoco simulation environment is shown in Fig. 3. Because the overall form-factor of a PR2 robot is similar to a human being, both agents are represented by a PR2 model in the simulation. The antagonist agent on the left scores by placing its bat within the orange spherical(target) area located between the two agents. The antagonist’s score will increase by 1 for every 0.01 seconds that its bat is placed within the target area without contacting the opponent’s bat. However, the antagonist will receive -10 points of score deduction if its bat is placed within the target area and makes contact with the protagonist’s bat simultaneously. Meanwhile, the protagonist agent on the right’s goal is to minimize the antagonist’s score in a game. Moreover, the protagonist will lose the game immediately if its bat is placed within the target area for more than 2 seconds. Each game will last for 20 seconds. The observation space for both agents is represented by the following tuple:

$$O^h = O^r = (p_h, p_r, v_h, v_r, t)$$

Where  $p_h$  and  $p_r$  are the Cartesian pose of the bat frame for the human player and robot player respectively.  $v_h$  and  $v_r$  describe the velocity of humans’ and robots’ bat.  $t$  specifies the game time in seconds.

### 4.2 Hardware Components

The physical system is designed and built based on a PR2 general-purpose robotic research platform. The PR2 robot has two 7 degrees

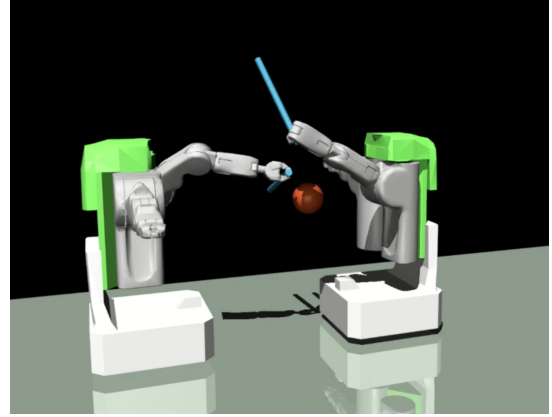


Figure 3: The fencing game in Mujoco simulation environment. The left agent simulates the human antagonist player, and the right agent simulates the robot protagonist player.

of freedom(DoF) arms, and its overall form-factor is similar to a human adult. It is therefore a suitable embodied agent for physical interaction with a human. An HTC VIVE VR headset and two controllers serve two purposes in this system: First, the use of the headset and controllers allows the robot to easily track the pose of the human’s head and hands. Second, VR technology provides an immersive experience to users, and it is commonly used in HRI research that requires a particular interaction scene [13, 17]. By synchronizing the actual robot’s location and joint angles to the PR2 model within the VR environment, human users will be able to see the robot’s behavior in real-time. Consequently, this system has the advantages of both the real and virtual environment, where actual physical contacts, robot motions, and ambient sound can be perceived by the human user, yet, the experimental environment can be easily and quickly modified based on the requirements of different tasks.

### 4.3 Learn to Play Games

We evaluated the performance of two multi-agent actor-critic algorithms trained in the fencing game environment. In the first approach, we combined the iterative training structure from [19] and the classic PPO algorithm[24]. The detail of this approach is presented in Algorithm 1.

The protagonist’s parameters  $\theta_i^H$  will first be trained by collecting trajectories that result from playing against the antagonist with a static policy. This continues until the protagonist’s policy achieves good performance against the antagonist’s current policy. The antagonist will then be trained against the protagonist with a static policy in order to find a policy  $v$  with parameters  $\theta_i^V$  that the protagonist’s policy is not robust to. This training sequence is repeated for  $N_{iter}$  iterations. Both agents are updated by optimizing the PPO clipped surrogate objective. Since this method could suffer from stability problems when  $N_{iter}$  is large, we ran the algorithm with  $N_{iter} = 2$ , and receive a robot agent that can play the fencing game sufficiently well.

In the second approach, we directly apply the algorithm in [1] to our environment. This method also trains both agents in an

**Algorithm 1:** Iterative Two-Agent Training

---

**Input:** Environment  $\mathcal{E}$ ; Stochastic policies  $\mu$  and  $\nu$   
**Initialize:** Parameters  $\theta_0^\mu$  for  $\mu$  and  $\theta_0^\nu$  for  $\nu$   
**for**  $i = 1, 2, \dots, N_{iter}$  **do**  
   $\theta_i^\mu \leftarrow \theta_{i-1}^\mu$   
  **for**  $j = 1, 2, \dots, N_\mu$  **do**  
    rollout  $\leftarrow \text{roll}(\mathcal{E}, \mu_{\theta_i^\mu}, \nu_{\theta_{i-1}^\nu}, N_{traj})$   
     $\theta_i^\mu \leftarrow \text{PPO\_Update}(\text{rollout})$   
  **end**  
   $\theta_i^\nu \leftarrow \theta_{i-1}^\nu$   
  **for**  $j = 1, 2, \dots, N_\nu$  **do**  
    rollout  $\leftarrow \text{roll}(\mathcal{E}, \mu_{\theta_i^\mu}, \nu_{\theta_i^\nu}, N_{traj})$   
     $\theta_i^\nu \leftarrow \text{PPO\_Update}(\text{rollout})$   
  **end**  
**end**

---

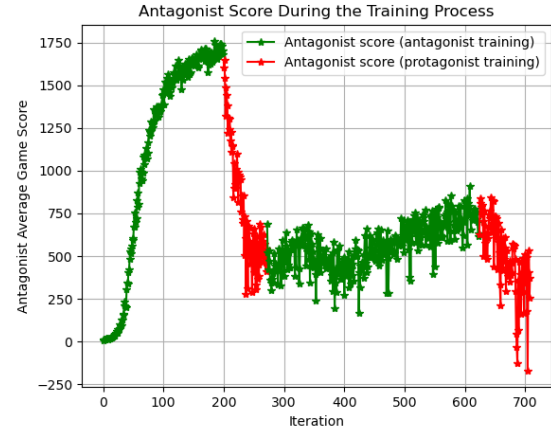
iterative scheme. Nevertheless, when training each of the agents, instead of having such an agent face against the latest opponent’s policy, this algorithm proposed to randomly load a previous version of the opponent’s policy from history. When using a very large batch size, this algorithm demonstrated better stability during a long sequence of training.

## 5 INITIAL RESULT & DISCUSSION

We were able to create a robotic agent with reasonable performance by using Algorithm 1. Fig. 4 shows the antagonist’s score throughout the iterative training process. The inclining green curves indicate the antagonist learns to score higher while avoiding the penalization from contacting the protagonist’s bat. The declining red curves indicate the protagonist learns to minimize the antagonist’s score. Each iteration is stopped when convergence is detected or a timeout occurs. The algorithm from [1] provides more stable training in a longer training sequence, and the resulted policies are seemingly more sophisticated comparing to those resulted from Algorithm 1. Due to the time limitation, we will compare and evaluate the two learning methods in our future works. Although we are still experimenting and improving the RL learning process, the other parts of the robotic system are ready for experiments. We performed an initial experiment with a policy that has been trained with Algorithm 1. In this experiment, one human participant was asked to play five consecutive games with the PR2 robot. Table 1 presents the participant’s score and average heart rate in each game.

The human participant got a negative score in the first game because he made contact with the robot’s bat frequently. However, the participant was able to quickly make adjustments and scored increasingly higher in the consecutive games. Meanwhile, the participant’s heart rate also increased when playing the games with the robot. The participant has a baseline average heart rate of 83 BPM when resting and 111 BPM when walking. The participant’s faster heart rates during the competitive games indicate that he exerted greater physical effort compared to the two baseline scenarios.

Since this project is still in the development stage, there are still many interesting questions that need to be explored and answered.



**Figure 4:** Antagonist score in the fencing game during the iterative training in Algorithm 1. The antagonist agent is trained first, and  $N_{iter} = 2$ .

	Game 1	Game 2	Game 3	Game 4	Game 5
Human Score	-541	503	329	830	754
Average Heart Rate (BPM)	124.9	129.0	144.6	150.8	144.5

**Table 1:** Game scores and average heart rate of the human participant during five consecutive competitive human-robot fencing games.

From the RL perspective, finding algorithms that can stably create more sophisticated control policies for competitive-HRI problems is essential. On the other hand, our initial experiment data shows that the human participant’s performance was gradually increasing during the games. Creating a robot agent that can learn from small amounts of real world data and quickly improve its policy can help the human’s performance improve faster. Moreover, it will be useful if the robot can help a participant achieve certain training goals. From the user experience perspective, it is important to understand what aspects of competitive-HRI are enjoyable to human users, and what can possibly results in negative emotions.

## ACKNOWLEDGMENTS

This work was supported in part by NSF awards EFMA-1832795 and CNS-1305072. This study (IRB ID: STUDY00012211) has been approved by the University of Washington Human Subjects Division.

## REFERENCES

- [1] Trapit Bansal, Jakub Pachocki, Szymon Sidor, Ilya Sutskever, and Igor Mordatch. 2017. Emergent complexity via multi-agent competition. *arXiv preprint arXiv:1710.03748* (2017).
- [2] Tapomayukh Bhattacharjee, Ethan K Gordon, Rosario Scalise, Maria E Cabrera, Anat Caspi, Maya Cakmak, and Siddhartha S Srinivasa. 2020. Is More Autonomy Always Better? Exploring Preferences of Users with Mobility Impairments in Robot-assisted Feeding. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 181–190.
- [3] Murray Campbell, A Joseph Hoane Jr, and Feng-hsiung Hsu. 2002. Deep blue. *Artificial intelligence* 134, 1-2 (2002), 57–83.
- [4] Min Chen, Stefanos Nikolaidis, Harold Soh, David Hsu, and Siddhartha Srinivasa. 2020. Trust-aware decision making for human-robot collaboration: Model learning and planning. *ACM Transactions on Human-Robot Interaction (THRI)* 9, 2 (2020), 1–23.
- [5] Aaron Edsinger and Charles C Kemp. 2007. Human-robot interaction for cooperative manipulation: Handing objects to one another. In *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 1167–1172.
- [6] Juan Fasola and Maja J Mataric. 2010. Robot exercise instructor: A socially assistive robot system to monitor and encourage physical exercise for the elderly. In *19th International Symposium in Robot and Human Interactive Communication*. IEEE, 416–421.
- [7] Deborah L Feltz, Samuel T Forlenza, Brian Winn, and Norbert L Kerr. 2014. Cyber buddy is better than no buddy: A test of the Köhler motivation effect in exergames. *GAMES FOR HEALTH: Research, Development, and Clinical Applications* 3, 2 (2014), 98–105.
- [8] Jakob N Foerster, Richard Y Chen, Maruan Al-Shedivat, Shimon Whiteson, Pieter Abbeel, and Igor Mordatch. 2017. Learning with opponent-learning awareness. *arXiv preprint arXiv:1709.04326* (2017).
- [9] He He, Jordan Boyd-Graber, Kevin Kwok, and Hal Daumé III. 2016. Opponent modeling in deep reinforcement learning. In *International conference on machine learning*. 1804–1813.
- [10] Suzanne Hidi and K Ann Renninger. 2006. The four-phase model of interest development. *Educational psychologist* 41, 2 (2006), 111–127.
- [11] W Daniel Hillis. 1990. Co-evolving parasites improve simulated evolution as an optimization procedure. *Physica D: Nonlinear Phenomena* 42, 1-3 (1990), 228–234.
- [12] Alap Kshirsagar, Bnaya Dreyfuss, Guy Ishai, Ori Heffetz, and Guy Hoffman. 2019. Monetary-incentive competition between humans and robots: experimental results. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 95–103.
- [13] Rui Li, Marc van Almkerk, Sanne van Waveren, Elizabeth Carter, and Iolanda Leite. 2019. Comparing human-robot proxemics between virtual reality and the real world. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 431–439.
- [14] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. 2015. Continuous control with deep reinforcement learning. *arXiv preprint arXiv:1509.02971* (2015).
- [15] Ryan Lowe, Yi I Wu, Aviv Tamar, Jean Harb, OpenAI Pieter Abbeel, and Igor Mordatch. 2017. Multi-agent actor-critic for mixed cooperative-competitive environments. In *Advances in neural information processing systems*. 6379–6390.
- [16] Ruikun Luo, Rafi Hayne, and Dmitry Berenson. 2018. Unsupervised early prediction of human reaching for human-robot collaboration in shared workspaces. *Autonomous Robots* 42, 3 (2018), 631–648.
- [17] Elias Matsas and George-Christopher Vosniakos. 2017. Design of a virtual reality training system for human-robot collaboration in manufacturing tasks. *International Journal on Interactive Design and Manufacturing (IJDeM)* 11, 2 (2017), 139–153.
- [18] Bilge Mutlu, Steven Osman, Jodi Forlizzi, Jessica Hodgins, and Sara Kiesler. 2006. Perceptions of ASIMO: an exploration on co-operation and competition with humans and humanoid robots. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 351–352.
- [19] Lerrel Pinto, James Davidson, Rahul Sukthankar, and Abhinav Gupta. 2017. Robust adversarial reinforcement learning. *arXiv preprint arXiv:1703.02702* (2017).
- [20] Jan L Plass, Paul A O’Keefe, Bruce D Homer, Jennifer Case, Elizabeth O Hayward, Murphy Stein, and Ken Perlin. 2013. The impact of individual, competitive, and collaborative mathematics game play on learning, performance, and motivation. *Journal of educational psychology* 105, 4 (2013), 1050.
- [21] Katie Salen, Katie Salen Tekinbaş, and Eric Zimmerman. 2004. *Rules of play: Game design fundamentals*. MIT press.
- [22] Kosuke Sato, Keita Watanabe, Shuichi Mizuno, Masayoshi Manabe, Hiroaki Yano, and Hiroo Iwata. 2017. Development of a block machine for volleyball attack training. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 1036–1041.
- [23] Jonathan Schaeffer, Neil Burch, Yngvi Björnsson, Akihiro Kishimoto, Martin Müller, Robert Lake, Paul Lu, and Steve Sutphen. 2007. Checkers is solved. *science* 317, 5844 (2007), 1518–1522.
- [24] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. 2017. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347* (2017).
- [25] Elaine Short, Justin Hart, Michelle Vu, and Brian Scassellati. 2010. No fair!! an interaction with a cheating robot. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 219–226.
- [26] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharmas Kumar, Thore Graepel, et al. 2018. A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science* 362, 6419 (2018), 1140–1144.
- [27] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. 2017. Mastering the game of go without human knowledge. *nature* 550, 7676 (2017), 354–359.
- [28] Kenneth O Stanley and Risto Miikkulainen. 2004. Competitive coevolution through evolutionary complexification. *Journal of artificial intelligence research* 21 (2004), 63–100.
- [29] Luise Süßenbach, Nina Riether, Sebastian Schneider, Ingmar Berger, Franz Kummert, Ingo Lütkebohle, and Karola Pitsch. 2014. A robot as fitness companion: towards an interactive action-based motivation model. In *The 23rd IEEE international symposium on robot and human interactive communication*. IEEE, 286–293.
- [30] Ardi Tampuu, Tambet Matiisen, Dorian Kodelja, Ilya Kuzovkin, Kristjan Korjus, Juhan Aru, Jaan Aru, and Raul Vicente. 2017. Multiagent cooperation and competition with deep reinforcement learning. *PloS one* 12, 4 (2017), e0172395.
- [31] M Viru, Anthony Hackney, K Karelson, T Janson, M Kuus, and A Viru. 2010. Competition effects on physiological responses to exercise: performance, cardiorespiratory and hormonal factors. *Acta Physiologica Hungarica* 97, 1 (2010), 22–30.