Summary of the paper: Beating the World's Best at Super Smash Bros. Melee with Deep Reinforcement Learning, by Vlad Firoiu, William R. Whitney and Joshua B. Tenenbaum[1]

Goals and Techniques

The goal is to test the performance of Reinforcement Learning methods on a multi-player console fighting game released by Nintendo and named Super Smash Bros. Melee (SSBM). The team chose to use features read on each frame, from the game's memory, such as: each player's position, velocity, action sate, as input for their agent. One limitation of this choice is that they are not able to take the projectiles thrown during the game into consideration. The game was run at a lower frame per second rate than its native speed and no actions were taken on skipped frames. The team also limited the number of controller inputs that can be used to a number that is both manageable for the agent and sufficient to perform all the important actions in the game. Important events were encoded as well as intermediary stages that inform players on how likely the opponent is to lose. A single character and a single stage were used by all playing agents for the initial experiements.

The team used two main reinforcement learning techniques namely Q-learning with the n-step SARSA approach and policy gradients taking the Actor-Critic approach and performing asynchronous experience generation.

The team ran emulators in parallel due to low frame rates on servers. A central trainer was tasked with collecting experiences from game playing agents working in parallel to maintain a circular queue of recent experiences. The trainer kept snapshots of the agents' neural network weights and at the same time, performed stochastic gradient descent on the set of experiences.

Results

Both approaches, Q-learning and Actor-Critic beat the in-game AI, each using a different strategy to win. The team was unable to duplicate the results obtained one a Gym environment. After playing against older versions of itself for training, the game playing agent also defeated competent human players including professional players competing in tournaments.

Weaknesses were uncovered when human players employed novel strategies unknown to the agent. The agent also performed poorly when there were changes in the character or stage. The same experiment was repeated for five other characters in the game after which characters trained against each other. This seemed to have solved some weaknesses from the single-character agent. Transferring a network from one character to another produced faster results compared to training a new character from scratch.

Q-learning seemed to perform less well than Action-Critic when dealing with evolving opponents. Attempts to train a recurrent agent that would play at a reaction speed similar to humans were unsuccessful.

---

[1]Vlad Firoiu, William F. Whitney, Joshua B. Tenenbaum, Beating the World's Best at Super Smash Bros. Melee with Deep Reinforcement Learning, 2017, https://arxiv.org/pdf/1702.06230.pdf