Can Deep Reinforcement Learning Solve Misere Combinatorial Games?

Oliver, Abraham (School: Brown County High School)

In 2017, Silver et al. designed a reinforcement learning algorithm that learned tabula rasa without any human-generated data. They applied this algorithm to the games of Go, Chess, and Shogi with positive results suggesting that the algorithm is general enough to learn many other games as well. Here, we investigate how well this technique performs on the misère version of a combinatorial game, misère tic-tac-toe, in which the goal is not to complete a row, column, or diagonal. We found that a successful Deep-Q neural network could be trained solely from self-play for the case of a three-by-three board. However, with board sizes of four-by-four and larger, the models failed to converge to a winning strategy and provided no upward trend to suggest that they would converge given more time. Although the computational power of the original tests could not be matched and the hyperparameters could not be as finely tuned, these results suggest that the AlphaZero algorithm may not be as general as previously thought and that misère games might require different specialized algorithms to solve them.

Awards Won: