PLAYING NONDETERMINISTIC GAMES THROUGH PLANNING WITH A LEARNED MODEL

Anonymous authors
Paper under double-blind review

ABSTRACT

The MuZero algorithm is known for achieving high-level performance on traditional zero-sum two-player games of perfect information such as chess, Go, and shogi, as well as visual, non-zero sum, single-player environments such as the Atari suite. Despite lacking a perfect simulator and employing a learned model of environmental dynamics, MuZero produces game-playing agents comparable of its predecessor AlphaZero. However, the current implementation of MuZero is restricted only to deterministic environments. This paper presents Nondeterministic MuZero (NDMZ), an extension of MuZero for nondeterministic, two-player, zero-sum games of perfect information. Borrowing from Nondeterministic Monte Carlo Tree Search and the theory of extensive-form games, NDMZ formalizes chance as a player in the game and incorporates it into the MuZero network architecture and tree search. Experiments show that NDMZ is capable of learning effective strategies and an accurate model of the game.