Paper review: Mastering the game of Go with deep neural networks and tree search (doi:10.1038/nature16961)

VINCENT ZHA

Background

The ancient game of Go has long been viewed as the most challenging of classic games for artificial intelligence owing to its enormous search space and the difficulty of evaluating board positions and moves. Prior to the paper, various algorithms were applied successfully handling Chess and other games and reached human professional level or even higher. However, the previous algorithms are not good enough to reach human professional player level.

The paper's goals

The authors of the paper developed program called AlphaGo to conquer the Go game with innovative approaches that were based on a combination of deep neural networks and tree search. They have developed, for the first time, effective move selection and position evaluation functions for Go, based on deep neural networks that are trained by a novel combination of supervised and reinforcement learning. They have introduced a new search algorithm that successfully combines neural network evaluations with Monte Carlo Tree Search rollouts.

The goal of the paper is to introduce their new approaches and communicate their inspiring results.

The paper's results

The program AlphaGo made significant accomplishments. It achieved a 99.8% winning rate against other Go programs, and defeated the human European Go champion Fan Hui by 5 games to 0, the first time that a computer program has defeated a human professional player in the full-sized game of Go, a feat previously thought to be at least a decade away. (Update: AlphaGo later defeated the top human player Lee Sedol at 4:1, further sending shock to both the Artificial Intelligence and Go playing world.)

During the match against Fan Hui, AlphaGo evaluated thousands of times fewer positions than Deep Blue did in its chess match against Kasparov; compensating by selecting those positions more intelligently, using the policy network, and evaluating them more precisely, using the value network—an approach that is perhaps closer to how humans play. Furthermore, while Deep Blue relied on a handcrafted evaluation function, the neural networks of AlphaGo are trained directly from gameplay purely through general-purpose supervised and reinforcement learning methods.

The advanced approaches used by AlphaGo can be applied in broader AI domains, such as general game-playing, classical planning, partially observed planning, scheduling, and constraint

satisfaction. The breakthrough provides hope that human-level performance can now be achieved in other seemingly intractable artificial intelligence domains.