Mastering the game of Go with deep neural networks and tree search

Selina Cheng

Authors

David Silver[1]*, Aja Huang[1]*, Chris J. Maddison[1], Arthur Guez[1], Laurent Sifre[1], George van den Driessche[1],Julian Schrittwieser[1], Ioannis Antonoglou[1], Veda Panneershelvam[1], Marc Lanctot[1], Sander Dieleman[1], Dominik Grewe[1],John Nham[2], Nal Kalchbrenner[1], Ilya Sutskever[2], Timothy Lillicrap[1], Madeleine Leach[1], Koray Kavukcuoglu[1],Thore Graepel[1] & Demis Hassabis[1]

[1]Google DeepMind, 5 New Street Square, London EC4A 3TW, UK.

[2]Google, 1600 Amphitheatre Parkway, Mountain View, California 94043, USA.

*These authors contributed equally to this work.

Overview

Problem: Go is a complex board game with 250^{150} possible sequences of moves, brute-force is intractable

Solution: Combine deep neural networks (DNNs) with Monte Carlo Tree Search (MCTS)

Impact: First AI to defeat a world champion (Fan Hui) in Go

Methods

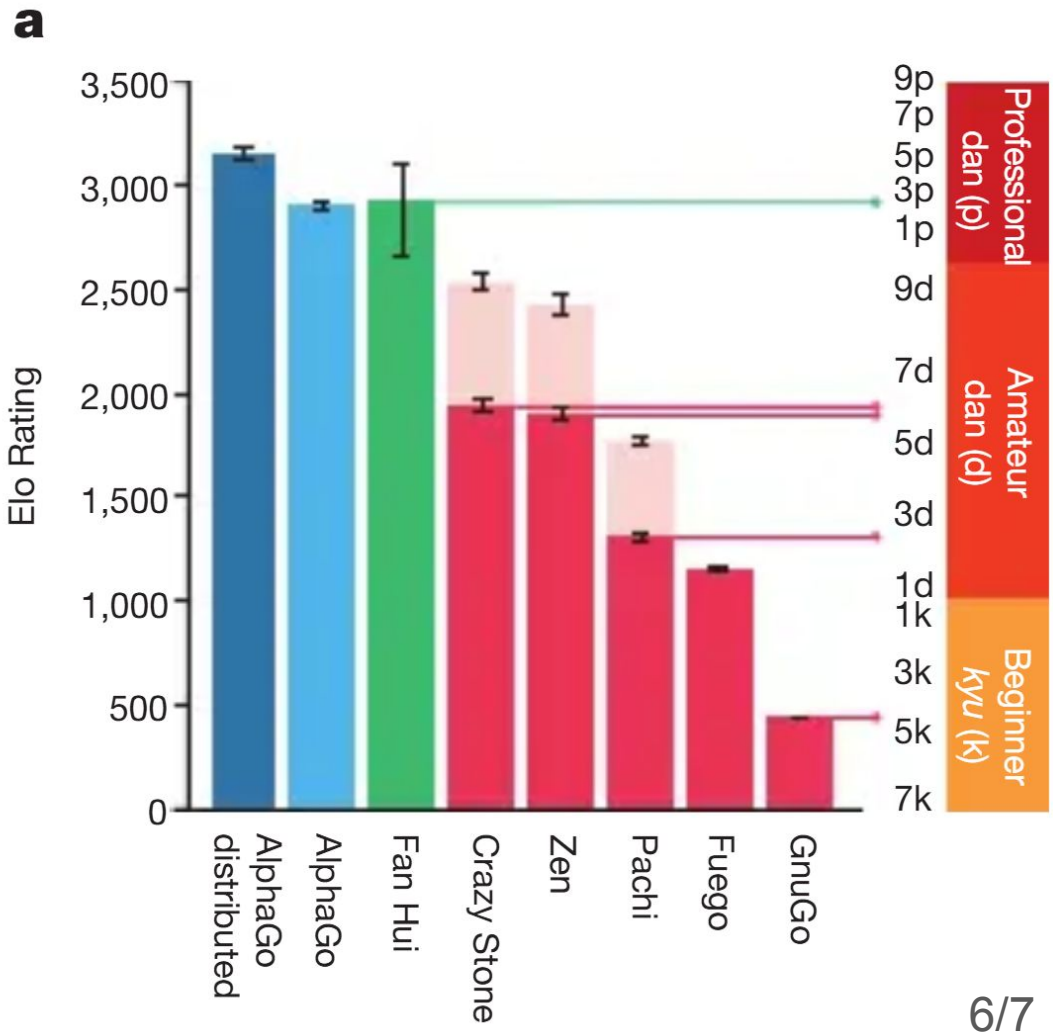
- **Two Neural Networks**
  - Policy Network, prediction and selection
  - Value Network, evaluation
- **Monte Carlo Tree Search** (MCTS)
  - Uses aforementioned networks to search
  - Exploration and exploitation
  - Simulates game trajectories to select optimal moves
- Together, they reduce the depth and breadth of search tree

Experimental Design

- Training
  - Policy Network Training
    - Supervised Learning, 30 million expert moves
    - Reinforcement Learning, self-play
  - Value Network Training, uses self-play positions
- Evaluation

Results

- AlphaGo vs. Competitors
  - 99.8% win rate
  - >77% handicap win rate
  - 2015 European Go: 5-0
- Outperforms pure search or human imitation



Key Takeaways

AlphaGo demonstrated the power behind hybrid AI that combines learning and search for complex decision-making in previously thought to be intractable problems.