AlphaGo and Monte Carlo Tree Search

Machine Learning for Big Data
CSE547/STAT548, University of Washington
Sham Kakade

©Sham Kakade 2017

A feat previously thought to be (a decade?) away!!!

2016

AlphaGo deep RL defeats Lee Sedol (4-1)
Chess vs. Alpha Go


Search space: $b^d: b = 35, d = 80$

Search space: $b^d: b = 250, d = 150$

Key Points

• Given current game state,
  – how to decide next action “a”?
  – How to evaluate each legal action?
Monte Carlo Tree Search (MCTS)

• A popular heuristic search algorithm for game play
  – By lots of simulations and select the most visited action.

  ![Diagram of MCTS]

Key point: how to calculate this value

One playout to simulate the game

Monte Carlo Tree Search (MCTS)

• AlphaGo
  – Key point: how to calculate the value for each action (path)

\[
a_t = \arg\max_a (Q(s_t, a) + u(s_t, a))
\]
The visited number for this edge

\[
N(s, a) = \sum_{i=1}^{n} 1(s, a, i)
\]

is estimated in two different ways.

Winning rate is predicted by deep reinforcement learning and predicted by fast rollout playout

Action probability predicted by supervised and reinforcement learning

\[
a_t = \arg\max_a (Q(s_t, a) + u(s_t, a))
\]

\[
Q(s, a) = \frac{1}{N(s, a)} \sum_{i=1}^{n} 1(s, a, i) V(s'_i)
\]

The mean value of all simulations passing through it.

\[
u(s, a) \propto \frac{P(s, a)}{1 + N(s, a)}
\]

(Fast) Methods to evaluate \( V \)

Rollout policy  SL policy network  RL policy network  Value network

\[
\Delta \rho \propto \frac{\partial \log p_{\rho}(a_t | s_t)}{\partial \rho}
\]

\[
\Delta \sigma \propto \frac{\partial \log p_{\sigma}(a | s)}{\partial \sigma}
\]

\[
\Delta \theta \propto \frac{\partial v_{\theta}(s)}{\partial \theta} (z - v_{\theta}(s))
\]

30 million self-play games
Thanks!

- ML for big data:
  - lots of different methods/challenges in the wild.
  - Participate in the ML community.

- Posters:
  - Looking forward to the poster session.

- Have a great summer!

Acknowledgments

- Some slides modified from:
  prlab.tudelft.nl/sites/default/files/deepmind_go_nature.pptx