Adaptive Neural Network Usage in Computer Go

Alexi Kessler, Ian Shusdock
Outline

● The Game of Go
● Computer Go Techniques
● Our Project
● Conclusions
● Future Work
What is Go?

- Two-player alternating stone placing game
- 19x19 board
- **Group**: Connected pieces
- **Liberty**: Empty adjacent position to group
- **Captured**: When a group has no liberties
- **Territory**: Empty locations “controlled” by a player
- No stone sacrifice
- Winner determined by territory and stone captures
What Makes Go Interesting?

- Incredibly complex
- ~$10^{81}$ atoms in the known universe
- Orders of magnitude harder than chess
- Complexity closely resembles real world
- Can lead to advances in artificial intelligence

<table>
<thead>
<tr>
<th></th>
<th>Chess</th>
<th>Go</th>
</tr>
</thead>
<tbody>
<tr>
<td>Possible board states</td>
<td>$10^{47}$</td>
<td>$10^{170}$</td>
</tr>
<tr>
<td>Possible legal move sequences</td>
<td>$10^{123}$</td>
<td>$10^{360}$</td>
</tr>
</tbody>
</table>
Why Study Go AI?

- Functionally infinite states and sequences
- Actions have long term influences
- States are not always as they appear
- In short, **very** hard
- Similar to sequential decision based problems
Previous Techniques
Minimax

- Tree of possible move sequences
- Assumes perfect play
- One player maximizes tree
- One player minimizes tree
- Best move chosen for root player
- Requires the entire tree mapped
- OR
- A heuristic function
Monte Carlo Simulation

- Policy based
- Value estimation
- Simulate games based on policy
- Sensitive to policy choice
- Randomization of policy

Total number of states: 36
Monte Carlo Tree Search

- A combination of game tree search and Monte Carlo simulation
- Limited minimax with heuristic
- Gradually adapt Monte Carlo policy
- Rely on fixed policy for “leaf” nodes
- Works well with Go
Upper Confidence on Trees

- Action selection is treated as separate problem for every node
- Select action $a$ that maximizes following equation
  - $(\text{estimated value of action } a) + \text{(modified bias sequence)}$
- Bias sequence is higher for less explored states/actions
- More likely to choose unexplored nodes
- Handles exploration-exploitation dilemma
Convolutional Neural Networks

- Functions similarly to normal neural network
- Processes overlapping tiles from input
- Great at visual identification
AlphaGo

- Developed by Google
- Two neural networks and MCTS
- Massive computing resources
- Plays moves that humans would not
Last Year’s MQP

- 4 approaches to help move selection
  - Introduce a neural network to Pachi
  - Change the neural network used based on tree depth
  - Train a neural network to inform Pachi search
  - Teach a neural network to use Pachi’s search

- Using a single neural network gave the best result

- Anomalous results
Our Project
Overview

- Investigated anomalous data
- Reinterpreted last year’s results
- Adaptive neural network weighting
- Compared optimized neural network Pachi to default Pachi
Anomalous Data
Reinterpreted Results
Adaptive Neural Network Weighting

- Determine the optimal weighting
- Go is complex, static weighting won’t work
- Based on
  - Board state
  - Game turn
- Trained using Fuego
- Trained two different functions
First Round Performance

![Pachi Theta V1 Performance Chart]

- **Pachi 10k 0.0**
  - Theta V1 Win Rate: X%
  - Opponent Win Rate: X%
- **Pachi 10k 0.5**
  - Theta V1 Win Rate: X%
  - Opponent Win Rate: X%
Second Round Performance

![Graph showing performance against Pachi 0.5 10k](image-url)
V2 vs. Default Pachi

Theta V2 vs Pachi 0.0 10k

Win Rates

- Pachi Theta V2
- Pachi 0.0 10k

Player
V2 vs. Fuego

Performance Against Fuego 12.5k

- Win Rate
- Pachi Theta V2
- Pachi 0.0 10k

- Fuego 12.5k
Depth Based Neural Network
Conclusions

- Adaptive weighting is powerful
- The faster neural network is not very good
- The slower neural network is strong
Future Work

- Use the slower, more accurate neural network
- Train function longer
- Experiment with more parameters
- Revisit the two other approaches from last year
Acknowledgements

- Levente Kocsis, Project Advisor and SZTAKI liaison
- MTA-SZTAKI
- Gabor Sarkozy, MQP Advisor
- Oscar Perez and Joshua Keller, last year’s MQP team
- Worcester Polytechnic Institute
Questions?
Köszönöm!