Mastering the Game of Go - Can How Did a Computer Program Beat a Human Champion?
In this Lecture:

❖ The game of Go
❖ The match so far
❖ History of man-machine matches
❖ The science
❖ Background
❖ Contributions in AlphaGo
❖ UAlberta and AlphaGo
❖ The future

Source: https://www.bamsoftware.com
The Game of Go
Go

❖ Classic Asian board game
❖ Simple rules, complex strategy
❖ Played by millions
❖ Hundreds of top experts - professional players
❖ Until now, computers weaker than humans
Go Rules

- Start: empty board
- Move: Place one stone of your own color
- Goal: surround
  - Empty points
  - Opponent (capture)
- Win: control more than half the board
- Komi: compensation for first player advantage

The opening of game 1
Final score on a 9x9 board
The Match
## Lee Sedol vs AlphaGo

<table>
<thead>
<tr>
<th>Name</th>
<th>Lee Sedol</th>
<th>AlphaGo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>33</td>
<td>2</td>
</tr>
<tr>
<td><strong>Official Rank</strong></td>
<td>9 Dan professional</td>
<td>none</td>
</tr>
<tr>
<td><strong>World titles</strong></td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td><strong>Processing Power</strong></td>
<td>1 brain</td>
<td>about 1200 CPU, 200 GPU</td>
</tr>
<tr>
<td><strong>Match results</strong></td>
<td>loss, loss, loss, <strong>win</strong>, ?</td>
<td><strong>win, win, win</strong>, loss, ?</td>
</tr>
<tr>
<td><strong>Go Experience</strong></td>
<td>Thousands of games against top humans</td>
<td>Many millions of self-play games</td>
</tr>
</tbody>
</table>
The Match So Far: Game 1

- Black: Lee Sedol
- White: AlphaGo
- Lee Sedol plays all-out at A on move 23. “Testing” the program?
- AlphaGo counterattacks very strongly and gets the advantage
Game 1 continued

- AlphaGo makes a mistake in the lower left. Lee is leading here.
- AlphaGo does not panic and puts relentless pressure on Lee.
- Lee cracks in the late middle game. Move A on the right side may be the losing move.

186 moves. AlphaGo wins by resignation.
Game 1 Reactions

- Shock. Disbelief.
- Huge media interest worldwide
Lee completely changes his style

With white, he plays very safe, solid moves

AlphaGo as Black plays creative, flexible moves and gradually gets ahead

A masterpiece for AlphaGo

211 moves.
AlphaGo wins by resignation
The Match So Far: Game 3

❖ Almost flawless game by AlphaGo
❖ Lee strongly attacks in the first corner, but AlphaGo turns the tables step by step
❖ AlphaGo “relaxes” after that but keeps a safe lead
❖ Professionals:
  ❖ “It played so well that it was almost scary”
  ❖ “Could 31 be the losing move?”
The Match So Far: Game 4

- Lee’s new strategy: take lots of profit, then stake the game on invading the center
- Lee came very close to losing all the center
- Then he produced a fantastic “tesuji”
- AlphaGo needed to compromise here. But it still thought it could get everything, and made things much worse for itself

Lee Sedol wins by resignation
Game 4 Discussion

❖ Why did AlphaGo miss this?
❖ Complex tactical fight
  ❖ Multiple targets
  ❖ Many threats
  ❖ No subset of threats works, but all together they work
❖ Computers lost in combinatorial explosion?
❖ Humans can precisely plan
❖ Human’s only (?) hope: out-calculate the computer (!)
Game 5 Tonight - Watch With Us

• Game 5 is the last game of the match
• It is very important:
  • Was game 4 a “fluke”…
  • …or did Lee figure out how to beat AlphaGo?
• Tonight from **10pm**
• Viewing party on campus, in room CSC 3-33
• Live Youtube feeds with professional commentaries
History of Computer Games Research and Man-Machine Matches
Many pioneers of Computing Science worked on game theory or program designs

Basis for all future work
Chess Man vs Machine

❖ David Levy’s bet - no program can defeat me in 10 years

❖ Easy wins in 1977, much closer in 1978 and 1979 but David Levy wins

❖ 1989: Deep Thought easily defeats Levy, 4-0

❖ 1996, Kasparov wins 4-2 vs Deep Blue

❖ 1996, Kasparov loses 2.5-3.5 vs Deep Blue
Backgammon Man vs Machine

- 1979: Berliner’s BKG wins short exhibition match against Villa
  - Lucky with the dice…

- 1992: Tesauro, TD-gammon
  - Very close to top human experts
- Current programs are almost perfect

*Figure 2.* An illustration of the normal opening position in backgammon. TD-Gammon has sparked a near-universal conversion in the way experts play certain opening rolls. For example, with an opening roll of 4-1, most players have now switched from the traditional move of 13-9, 6-5, to TD-Gammon’s preference, 13-9, 24-23. TD-Gammon’s analysis is given in Table 2.
Checkers Man vs Machine

- Chinook program developed over decades by Jonathan Schaeffer and his group
- 1992+1994: Chinook vs Tinsley
- 2007: Checkers solved
Othello Man vs Machine

- Logistello program developed by Michael Buro
- 1997: Logistello vs Murakami 6 - 0
Poker Man vs Machine

- UAlberta Poker research group, led by Mike Bowling
- 2015: Heads-up limit Texas hold’em is solved
Go Man vs Machine

- 2009: Fuego (open source, mostly UAlberta)
  - First win against top human professional on 9x9 board
- 19x19 board: Many handicap matches (computer starts with an advantage)
- Before AlphaGo, about 3-4 handicap stones
- 2015: AlphaGo beats Fan Hui 2 Dan professional, no handicap
- 2016: AlphaGo - Lee Sedol

White: Fuego
Black: Chou Chun-Hsun 9 Dan
White wins by 2.5 points
The Science
The Science Behind AlphaGo

- AlphaGo builds on decades of research in:
  - Building high performance game playing programs
  - Reinforcement Learning
  - (Deep) neural networks

UAlberta is a world leader
The Science - Background
Making Complex Decisions

- We make decisions every moment of our lives
- What is the process that leads to our decisions?

How to make good decisions?
- Consider many alternatives
- Consider short-term and long-term consequences
- Evaluate different options and choose the best-looking one

Image Source: https://www.rubegoldberg.com
Making Sequential Decisions

- Loop:
  - Get current state of world
  - Analyze it
  - Select an action
  - Observe the world’s response
  - If not done: go back to start of loop
Heuristic Search

- Heuristic search is a research area in computing science
- It is considered a part of the field of Artificial Intelligence
- It can be used for sequential decision-making problems
- Applications: automated planning, optimization problems, pathfinding, games, puzzles,…
The Three Plus One Pillars of Modern Heuristic Search

❖ Three main ingredients:
  ❖ Search
  ❖ Knowledge
  ❖ Simulations

❖ Plus one:
  ❖ Machine learning to acquire knowledge

❖ We will see all of these used in AlphaGo

❖ Many other modern heuristic search methods also use those

❖ Examples:
  ❖ planning
  ❖ robot motion planning
  ❖ mapping unknown terrain
  ❖ other games
Tree Search

❖ At each step in the loop:
❖ I need to choose one of my actions
❖ The world could react in one of many possible ways
❖ Drawing all possible sequences results in a (huge) tree

Domain Knowledge and Evaluation

- We need to know if a sequence led to a good result
- Exact knowledge: we know the result for sure
- Heuristic knowledge: an estimate of the result
- Evaluation: mapping from a state of the world to a number
- How good or bad is it for us?

Win for X
Loss for X (Win for O)
Draw
What’s your evaluation?
Simulation

- For complex problems, there are far too many possible sequences
- Sometimes, there is no good evaluation
- We can sample long-term consequences by simulating many future trajectories

Image Source: https://upload.wikimedia.org
Computer Go Before AlphaGo

- **Search:**
  - Monte Carlo Tree Search
  - Invented about 10 years ago
  - First successful use of simulations for classical two-player games
  - Scaled up to massively parallel (e.g. Fuego on 2000 cores on Hungabee)

- **Simulation:**
  - Play until end of game
  - Find who wins at end (easy)
  - Moves in simulation: random + simple rules
  - Early rules hand-made, later machine-learned based on simple features
Computer Go Before AlphaGo

- **Knowledge:**
  - Fast, simple knowledge: used for move selection in simulation (“rollout policy”)
  - Slower, better knowledge: used for move ordering in tree search (“SL policy”)
  - Since 2015: even better slow knowledge from deep convolutional neural networks
Computer Go Before AlphaGo

- Summary of state of the art before AlphaGo:
  - Search - quite strong
  - Simulations - OK, but hard to improve
  - Knowledge
    - Good for move selection
    - Considered hopeless for position evaluation

Who is better here?
The Science - AlphaGo’s Contributions
Alpha Go Design

- According to paper in Nature
- Not yet known what changed over the last 5 months, other than much more self-play
- Search: MCTS (normal)
- Simulation (rollout) policy: relatively normal
- Supervised Learning (SL) policy from master games: improved in details, more data
- **New:** Reinforcement Learning (RL) from self-play for value network
- **New:** Reinforcement Learning (RL) from self-play for policy network
Value Network

- Given a Go position
- Computes probability of winning
- No search, no simulation!
- Static evaluation function
- Trained by RL from self-play
- Trains a deep neural network
- Similarly, the policy network is trained to propose stronger moves
Putting it All Together

- A huge engineering effort
- I only showed the tip of the iceberg here
- Many other technical contributions
- Massive amounts of self-play training for the neural networks
- Massive amounts of testing/tuning
- Large hardware: 1202 CPU, 176 GPU used in previous match, “similar hardware” vs Lee Sedol
University of Alberta and AlphaGo
DeepMind and Us

❖ AlphaGo is “big Science”
❖ Dozens of developers, millions of dollars in hardware and Computing costs
❖ What is the role of our university in all of this?
❖ We contributed lots of:
  1. Basic research
  2. Training
UAlberta Research and Training

- Citation list from AlphaGo paper in Nature
- Papers with UofA faculty or UofA trainees in yellow

The Future
Where do we Go from Here?

- Which other problems can be tackled with this approach?
- The methods are quite general, not Go-specific
- We need an internal **model** of the problem in order to learn from self play
- We may be able to use similar approaches when we have lots of data
- Can we build a model from data?
- MCTS started in Go, has found a large number of applications
- Deep learning techniques have revolutionized many fields such as image recognition, speech recognition, natural language processing, drug discovery…
- Go was the first combination of MCTS and deep learning
- Limitless possibilities…
What Should UAlberta Do?

- Keep doing world-leading basic research and training
- Find ways to attract and retain the best students in the field
- Update our computational infrastructure to *not completely* lose touch with industry
- Develop applications beyond games? Big science?
Summary and Outlook

❖ DeepMind’s AlphaGo program is an incredible research breakthrough
❖ Landmark achievement for Computing Science
❖ University of Alberta has played very significant roles on the way there
❖ We must try to stay relevant in the future!

Watch game 5 with us: 10 pm, CSC 3-33