Article from

Predictive Analytics and Futurism

July 2016
Issue 13

From Deep Blue to DeepMind: What AlphaGo Tells Us

By Haofeng Yu

Early February in 2016, Demis Hassabis, one of Google DeepMind's founders, tweeted: "Thrilled to officially announce the 5-game challenge match between #AlphaGo and Lee Sedol in Seoul from March 9th-15th for a $1M prize!" While Hassabis was a name I barely knew and AlphaGo sounds like another of Google's toys with a catchy name, growing up playing Go, I knew about Lee very well. The Korean professional Go player had been at the top of the game for almost a decade. The 18 world championships he collected are nothing short of Roger Federer's 17 or Tiger Woods' 14 grand slam titles in their respective fields, tennis and golf. The competition didn't seem to be a good match-up. "I would bet anything that AlphaGo won't go anywhere," I told my friends.

The competition took place in Seoul as scheduled. To the surprise of many fans, including Go professionals, AlphaGo beat Lee four games to one, with the human's sole win coming from the fourth game, merely a consolation that doesn't matter in the best-of-five setting. This is stunning, devastating, yet equally interesting. It inevitably reminds people of the chess match that took place in 1997 between IBM's super computer Deep Blue and Garry Kasparov, the reigning champion at that time. Deep Blue won.

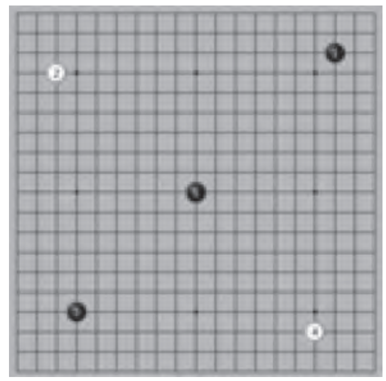
Humanity's intellectual pride continued to be humbled with IBM's Watson beating two champs on the game show "Jeopardy!" in 2011, and now AlphaGo winning at Go, the game many applaud as the final line of defense of human intelligence. Many questions ensue, including: What are DeepMind and AlphaGo? What can AlphaGo tell us, particularly, actuaries? To begin with, let's talk about Go.

ABOUT GO

Have you seen the 2001 movie "A Beautiful Mind"? There is a scene at the beginning where John Nash (Russell Crowe) awkwardly wanders around Princeton's campus, "extracting an algorithm to define [the] movement" of pigeons, while making notes. Very soon, he is dragged into a game of Go, with one semester's free laundry service at stake. Nash loses and claims the game is flawed and his perfect move was ruined.[1] Had he extracted some algorithms for Go instead of for the birds, like AlphaGo did, he could have enjoyed the free laundry service.

The game of Go originated in China more than 2,500 years ago. The rules are simple: Players take turns placing black or white stones on the board, a 19-by-19 square grid, trying to capture the opponent's stones or surround empty space to mark as their own territory. See Figure 1.

Figure 1: The Game of the 20th Century: Go Seigen (Black) vs. Honinbo (White) 1933



Source: Unknown

As simple as the rules are, Go is a game of profound complexity. Its abstract concept of "shi," sensible but indescribable, is unique to the game and often linked to Oriental philosophy, even at national strategic level.[2] Unlike Western chess, which has about 40 moves in a game, Go can last for up to 200.

According to Google DeepMind's site, "There are more possible positions in Go than there are atoms in the observable universe. … Go is played primarily through intuition and feel, and because of its beauty, subtlety and intellectual depth, it has captured the human imagination for centuries."[3]

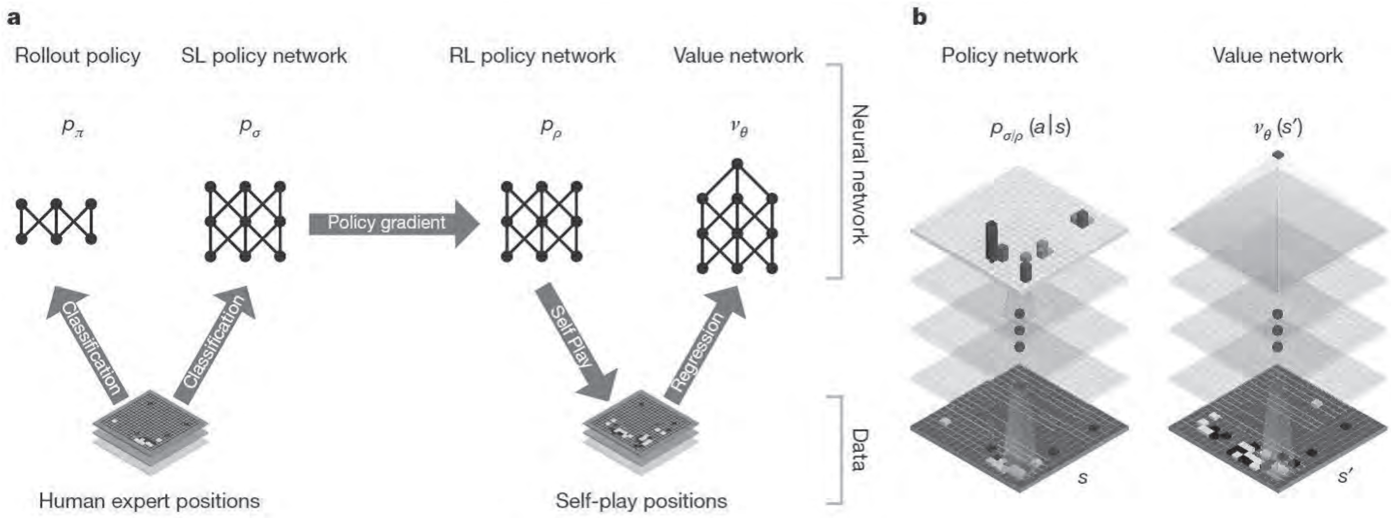
Two quotes from 20th century Chess and Go player Edward Lasker summarize chess and Go this way:

"It has been said that man is distinguished from animal in that he buys more books than he can read. I should like to suggest that the inclusion of a few chess books would help to make the distinction unmistakable." – *The Adventure of Chess*

"While the Baroque rules of Chess could only have been created by humans, the rules of Go are so elegant, organic, and rigorously logical that if intelligent life forms exist elsewhere in the universe, they almost certainly play Go."

No wonder there are so many efforts, including from Facebook, to build a Go application—it simply offers higher levels of—if not the ultimate—challenge. It was thought it would be at least

Figure 2: Neural network training pipeline and architecture



Source: Silver et al., "Mastering the Game of Go."

another 10 years before a machine could beat a human professional in Go; it happened much more quickly.

ABOUT DEEP BLUE

Back in 1997, how did Deep Blue beat Kasparov, the reigning world champion? IBM explains on its Deep Blue website, "The answer lies in its unique combination of innovative software engineering and massive parallel processing power."[4]

To the first point, the keys to IBM's software engineering are tree search with alpha-beta pruning technique and hand-crafted evaluation functions, which do not necessarily represent advanced mathematics or a heavy use of statistics!

To the second, Deep Blue is a massively parallel "32-node IBM RS/6000 SP high-performance computer … capable of evaluating 200 million positions per second"; now we know this kind of computing power can be available in each household.

While Deep Blue attains its strength more or less out of brute force computing power, back in the day, it was a modern marvel.

ABOUT DEEPMIND AND ALPHAGO

Founded in Britain in 2010, the artificial intelligence company Google DeepMind was acquired and renamed by Google in 2014. Google describes AlphaGo as a computer Go that combines Monte Carlo tree search with deep neural networks that have been trained by supervised learning (SL), from human expert games, and by reinforcement learning (RL) from games of self-play.[5] The AlphaGo team also published a paper in Nature in January 2016, which offers comprehensive technical details, for your academic curiosity.[6]

As mentioned earlier, since the search space of future moves of a Go game is so large that no AI can explore every possibility, how did AlphaGo accomplish the mission impossible? Figure 2 tells you where AlphaGo derives its amazing playing strength.

To the best of my understanding, its secret power comes from the following four elements.

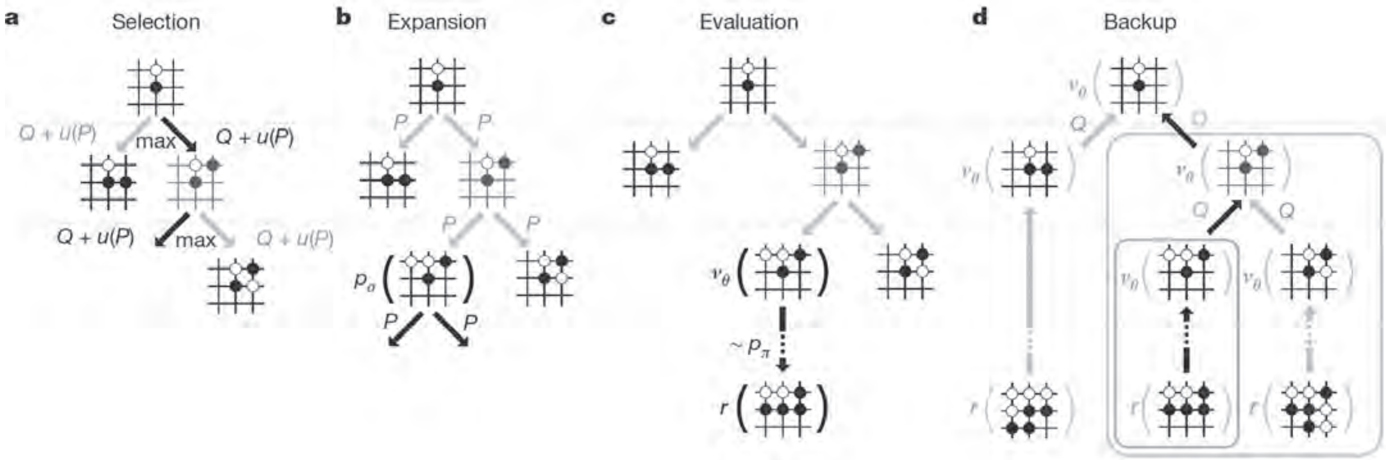
1) **Policy networks** (AlphaGo's left brain). Given the current situation, these networks predict moves that human experts would likely pick. There are two kinds, or phases:

   - Supervised learning. The policy network was trained with numerous information, 30 million positions, from Go games that had been played by human experts; it predicts by maximizing the likelihood of human expert moves.

   - Reinforcement learning. The policy network was training by playing "against itself" millions of times, in a sense teaching itself which moves and strategies worked and which didn't; it predicts by maximizing expected outcomes (of winning).

2) **Rollout policy** (AlphaGo's legs and hands, as it "acts" without thinking/using its "brains"). Given the current situation, this policy predicts moves that human experts would make, similarly to policy networks, but with much

Figure 3: Monte Carlo tree search in AlphaGo



Source: Silver et al., "Mastering the Game of Go."

greater speed. It plays in a way akin to "intuition," so as to achieve a balance between accuracy and speed.

3) **Value network** (AlphaGo's right brain). Given the current situation, it evaluates and spits out the odds of winning or losing. With this function, AlphaGo is able to evaluate its moves quantitatively. Generally speaking, the value function of Go is highly nonsmooth and irregular.

4) **Monte Carlo tree search** (AlphaGo's body). This is the framework that integrates all the parts.

In a training pipeline, the AlphaGo team "pass in the board position as a 19 × 19 image and use convolutional layers to construct a representation" and then "use neural networks to reduce the effective depth and breadth of the (Monte Carlo) search tree (4), evaluating positions using a value network (3), … sampling actions using a policy network (1)," and balancing speed and accuracy with the fast rollout policy (2).

None of the four pieces is utterly new; however, the integration of these concepts, in such a creative and efficient way, is a work of beauty.[7]

WHAT DOES ALPHAGO TELLS US?

The first lesson is that while computing power is still important, its weight has declined. Back in 1997, IBM touted its computing power as one major contributing factor; in 2016, DeepMind seems to intentionally refrain from using super power. The AlphaGo that defeated Lee was a distributed version that uses 1,202 central processing units (CPUs) and 176 graphics processing units (GPUs). Given Google's capacities, it can certainly come up with a stronger AlphaGo if they wish.

The second lesson is that the brute force of Deep Blue has evolved into a whole new form, in the name of machine learning. Unlike Deep Blue, which employed exhaustive tree search with alpha-beta pruning, AlphaGo learns things "brute-force-ly" from scratch. In a sense, the brute force is manifested by its "diligence"—AlphaGo mimics an extremely diligent, but not necessarily genius, student who is willing to learn from millions of human's play and self-play, tediously.

The third lesson we take away here is that data is the key. Deep Blue relied on a huge database of hand-crafted books on openings and endgames to simplify its search; without the daunting 30 million human positions AlphaGo has learned, I doubt the reinforcement learning by self-play can add much value and AlphaGo's strength shall be discounted.

I believe these points, especially the third one, are particularly important for us actuaries. While we have started seeing so-called "disruptive innovations" of machine learning and predictive analytics in our work, without high quality and business specific data, anything that they mean could be misleading. So, for companies who strive to automate their agency, underwriting and claims, or even investment and asset liability management (ALM) processes, they had better invest in data, so as to save for a rainy day.

WHAT HAS ALPHAGO NOT TOLD US YET?

First, the new form of brute force mentioned above may be easily translated into other logic-based territories, not limited to games. True, AlphaGo can only play Go right now. It cannot even move one stone by itself—one of its creators, Ajay Huang, had to sit in front of Lee Sedol and place stones on its behalf.

But its way of learning, guided by minimal hand-crafted rules, is truly inspiring.

Second, there exists another powerful and relevant machine learning tool that has not been mentioned yet—unsupervised learning (UL). Judging from the paper in Nature, AlphaGo doesn't seem to have been trained by UL, or at least, DeepMind didn't make it explicit. But some of AlphaGo's moves are far from we humans' play book. For example, the 37th move in Game 2—no human player would play like this; yet, it was a key play whose importance only was revealed after 20 more exchanges. Its own way of playing! One has to wonder, if DeepMind does train AlphaGo using UL, can it teach humans even more?

Interestingly enough, but I bet that DeepMind won't be satisfied by producing merely top video game or Go players. We have reason to believe that DeepMind and its competitors are aiming for more, especially in this era when big data, machine learning, cloud computing, Internet of things (IoT), augmented reality (AR) and virtual reality (VR) bring our physical world closer than ever to virtual worlds. While AlphaGo-like "narrow" AIs (as described by Hassabis) are still far away from their ultimate form, artificial general intelligence (AGI), they are marching in that direction.

NOT JUST FOR APRIL FOOLS'

In Davos-Klosters, Switzerland, this January, the fourth industrial revolution, or Industry 4.0, driven by rising usage of big data and artificial intelligence in all aspects of the economy, emerged as one of the main themes at the 46th World Economic Forum. One report predicts "7.1 million redundancies by 2021, mainly in the fields of management and administration, particularly in the healthcare sector." About the same time, McKinsey released its outlook that automated systems may take over up to 25 percent of insurance jobs.[8]

On the other end, DeepMind just announced on their website that it had struck a deal collaborating with the U.K.'s National Health Service; IBM revealed on their website plans to move into telehealth and telecare five years after IBM Watson toppled the game show "Jeopardy!" Granted, the health and insurance industry is not the only space where actuaries live, but it has been our natural habitat!

Coincidentally, or intentionally in light of the AlphaGo hype, a friend shared with me a news item with the headline "First Robot Run Insurance Agency Opens for Business"—what a "classic" teaser by a "classic" name, Lirpa Loof, on an April Fools' Day! Somehow, it appears not just for April Fools'.

STILL A LONG WAY TO GO

Not all AIs succeeded in challenging humans. Claudico, an AI from Carnegie Mellon University (CMU), lost the Brains vs. Artificial Intelligence challenge in a type of Texas hold 'em poker game in 2015. Interestingly, CMU is also the birthplace of Deep Blue.

In summary, here are two takeaway messages for AI.

- Be humble; even a sophisticated game like Go may represent only a limited and partial perspective of the human unpredictable nature.
- Get more training, supervised or unsupervised, on bluffing.

"Right now, humans are doing OK,"[10] said Doug Polk, a former World Series of Poker champion, who just "defeated" Claudico.

The author would like to thank Aolin Zhang for sharing the April Fools' news and helpful discussion. ■

Haofeng Yu, FSA, Ph.D., is actuary and director, of Inforce Management at AIG. He also serves as webcast and research coordinator of the Predictive Analytics and Futurism Section Council. He can be reached by *Haofeng.Yu@aig.com.*

REFERENCES

1 "The Challenge," Beautiful Mind, directed by Ron Howard, 2001, YouTube clip, 2:26, posted by "andStack," Oct. 8, 2010, *https://www.youtube.com/watch?v=GmlSSSN7C78&nohtml5=False.*

2 David Lai, "Learning from the Stones: A Go Approach to Mastering China's Strategic Concept, Shi" (Advancing Strategic Thought Series, Strategic Studies Institute, May 2004), *http://www.strategicstudiesinstitute.army.mil/pubs/display.cfm?pubID=378.*

3 "The Game of Go," Google DeepMind, accessed May 16. 2016, https://www.deepmind.com/alpha-go.

4 "How Deep Blue Works: Under the Hood of IBM's Chess-Playing Supercomputer," IBM, accessed May 16, 2016, *https://www.research.ibm.com/deepblue/meet/html/d.3.2.html.*

5 "AlphaGo: Using Machine Learning to Master the Ancient Game of Go," Google (blog), Jan. 27, 2016, *https://googleblog.blogspot.nl/2016/01/alphago-machine-learning-game-go.html.*

6 David Silver, et al., "Mastering the Game of Go With Deep Neural Networks and Tree Search," Nature 529 (Jan. 28, 2016), *http://www.nature.com/nature/journal/v529/n7587/full/nature16961.html.*

7 While the fully armed AlphaGo deserves all compliments, let us not underappreciate the fineness of each part—they are refined by DeepMind to a level such that each standalone piece can challenge other competitor Go programs in the world.

8 Unknown,"Robots put five million jobs at risk," SwissInfo, accessed May 24, 2016, *http://www.swissinfo.ch/eng/world-economic-forum_digital-revolution-puts-five-million-jobs-at-risk/41900634.*

9 "How Automation Will Whack Up to 25% of Insurance Jobs: McKinsey", Insurance Journal, accessed May 24, 2016, *http://www.insurancejournal.com/news/national/2016/02/01/397026.htm*.

10 Noah Bierman, "Artificial Intelligence Bot vs The Poker Pros," LATimes, accessed May 24, 2016, http://www.latimes.com/nation/great-reads/la-na-c1-claudico-poker-20150521-story.html.