AI/ADM Case study： Go AI Technology

Wenqian Zhang(z5286111), Ziyi Liang(z5325156),

Zhangpeng Chen(z5271769)

University of New South Wales

1 Introduction

Go is a strategic board game played with black and white pieces. It originated in China and is one of the world's oldest chess sports. It is also considered to be the pinnacle of human board games. As measured by state-space complexity, the variations of Go can reach ten to the 172nd power. According to the complexity of the game tree, it is ten to the 300th power, which is much larger than the number of atoms in the universe, which is ten to the 80th power. The fact that no game has been repeated from the 23rd century B.C. to the present makes it so fascinating. There are still tens of millions of Go enthusiasts worldwide, and in China alone, about two to three million new children and young people are introduced to the game every year (Xinhua 2020).

Go is more accurately defined as a game of complete information in which each player has accurate information about all other players' characteristics, strategies, and benefits functions. As of 2015, of the more than 6,000 complete information games played worldwide, only the Go program cannot compete against professionals in its field. However, in October 2015, AlphaGo became the first computer Go program to defeat a professional Go player on a 19-way board without a handicap. Therefore the AI technology we will investigate is this AlphaGo Go AI that, relies on machine learning and neural network technology developed in recent years to give birth.

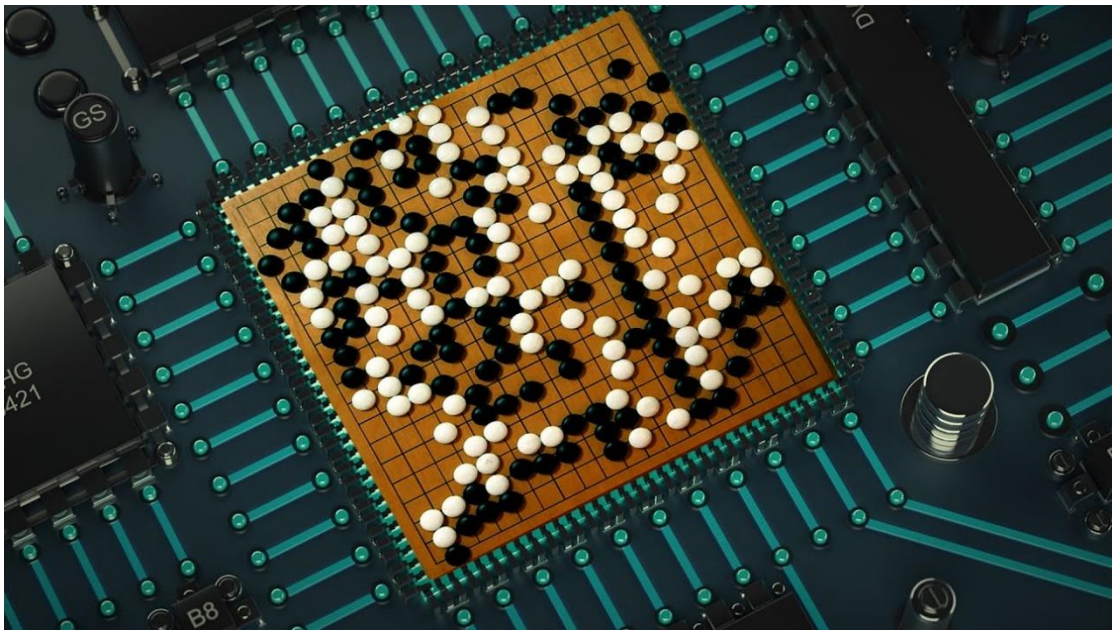

Figure 1. Abstract image of the fusion of Go and computer technology

2 Purpose, Benefits

AlphaGo is an artificial intelligence Go software developed by the Google DeepMind team in London, UK. It can learn to train spontaneously like the human brain, improving its playing strength with trillions of self-playing trainings.

The aims of this technology are:

1. To enhance AI technology to solve complex AI problems for broad societal benefits. Especially in applications where AI technology is uniquely suited, such as advancing science and addressing climate and sustainability issues.
2. Since Go can produce almost infinite variations and branches, it is a huge arena where we can train, study, practice, and validate machine learning and neural network techniques in depth.
3. Guide human Go players one step further, enhancing human potential.
4. It is a good thing in itself to explore the unknown and thus open our horizons. AI is like opening a 'skylight' that allows us to explore the vast world in Go further.

The scope of what this technology, AlphaGo, involves and can affect is:

1. For professionals and beginners.
2. It can be used for introductory Go instruction and advanced Go-level reinforcement.
3. Anywhere at any time; the advantage is that users can open the software at any time to learn, with no need to hire a personal Go tutor.
4. It can help professionals or players who want to take a step closer to reviewing the game. It gives the mistakes and wins curves for the entire sentence of the game. It helps players understand what kind of branches and results different plays will produce.

The demand for Go AI as a technology comes from the desire to have a better way to learn and understand how to play better. For example, this technology can help detect if some players can play better than they did before, including what mistakes they made before. This provides them with data to improve their strategy to become better go players. In addition, AlphaGo was originally made to beat humans in Go tournaments. Now, Go AI technology is becoming more collaborative and engaging for the masses. Allowing for this symbiotic relationship between humans and machines, the emotional and technical experience that comes with the game of Go can be reinvigorated.

3 Stakeholders

The population that can be affected by Go AI technology can be broadly divided into the following categories:

1. Direct stakeholders include the companies developing Go AI and the technical staff involved.
2. Go professionals who are directly affected by this technology, as well as tournament organizers and Go associations.
3. The vast number of Go enthusiasts is also a direct stakeholder in this technology.

3.1. The first category of stakeholders

This category is highly technical experts, highly receptive to AI technology, and creators and improvers of Go AI technology. Their needs and goals are to make Go AI more accessible and interactive to more people. They want to get more data to improve their technology and publish papers about it for the development of the academic community. As can be expected, they will be positively impacted by the deployment of the technology, both financially in terms of revenue and the benefits to the industry as a whole from the expansion of practitioners. By relying on technological developments, this technology can be transplanted, or transferred, to other complex human task domains in the future. Provide more possibilities for the future of AI technology as a whole.

3.2. The second category of stakeholders

If the first category is the producers, the second category is the most direct consumers or users. Go professionals, in particular, survive on the Go game, and many have been learning Go since childhood. Many enter the National Go Academy directly after junior high school, dedicating their lives to Go and motivated to become professional players. Their potential need for this technology lies in the ability to help them understand their Go strategy flaws more deeply and accompany them in their training day and night.

This group has a common understanding of technology and a mixed acceptance of it. For example, some players are very resistant to the technology, believing that Go AI could ruin the fun of the game. They believe that AI has taken away the future of this historic game and that everything is cold data that can be calculated. Most players expressed their concerns about the technology: they were afraid it would lead to the loss of their jobs; after all, no human effort can beat a computer that computes hundreds of millions of times per second.

On the other hand, the opponents believe that this technology will allow humans to enter a new era of Go and that in normal games, sometimes the winning rate increases after humans play a move that the AI does not recommend. This proves that the AI has a "blind spot" in its calculations and that humans can see good moves that the AI cannot see. Players are now trying to find the blind spots that the AI cannot find. Many people have experienced the joy of challenging AI. To summarize, deploying AI Go technology will have both positive and negative effects on such people.

3.3. The third category of stakeholders

This categorized population is large, and the age span can be from 6 to 90 years old. Moreover, their views on this technology are also good and bad. Which can be roughly divided into two groups according to the acceptance of AI technology:

1. People who have experienced the benefits of AI are also users of this technology. This group's goals and needs are more or less the same as those of the second category. They all believe that AI can help them further their Go skills. For them, the role of Go AI as a "teacher" is more prominent.
2. The resisters of this technology, that is, people who refuse to use Go AI technology. This group is more likely to be older and less receptive to not only Go AI but also to other Internet technologies available today. They tend to believe that Go AI will cause Go to lose "the joy of human play." The deployment of Go AI will have a negative impact on them because they will not be able to use the new technology.

4 Value Sensitive Design

We will use the game "Judgment Call" to help reveal the ethical issues in Go AI technology and predict the social implications. We will role-play each of the following three specific groups of stakeholders who will write product reviews from their perspective to consider the possible harms and impacts of the technology:

1. AI Technicians involved in the development of Go AI.
2. Professional Go players.
3. Go enthusiasts who accept this technology.

Comment format:
[Stakeholders] - [Ethical principles of concern] - [Rating card (1 - 5 stars)]: Specific comments

One star means they do not like the technology, and five stars means they like it.

The following is a summary of the sessions we discussed in three scenarios.

4.1. Scenario I

Go teachers are teaching Go to students on an online Go matchmaking platform. The online platform uses Go AI technology as an aid. Personalized services are provided for different users.

- **AI Technicians - Security and Privacy – 4:** We should design AI systems that are both secure and respectful of privacy. This is because if private matchups, or details of a teacher's teaching with AI, are leaked, it could lead to disadvantages for the teacher's future contests. In, for example, online platforms, failure to focus on security and privacy may also lead to leaks of personal information of teachers and students. The consequences are something we need to be responsible for.
- **Professional Go players – Inclusiveness – 4:** Go AI technology should be able to give different teaching advice for different levels of students, which is beneficial to teachers for targeted teaching, and the advice given by AI should be more diversified to avoid singularity of teaching where professionals may not have the strength to give more solutions. Such AI technology will help us a lot in teaching.
- **Go enthusiasts - Reliability and safety – 5:** Hopefully, when being taught. The impact of each move on the subsequent win rate can be seen more explicitly through Go AI technology, as well as giving students a better understanding of why the move was made by the win rate predicted by the AI for each board position. If this technology can help us avoid learning the wrong ideas and knowledge, then we will be happy.

4.2. Scenario II

An international Go tournament is underway, and the players participating have previously used Go AI as a training tool. The tournament commentators are also using Go AI technology to assist them in explaining the game live to the world.

- **AI Technicians - Reliability and safety – 5:** Our technology should provide real-time win curves while the public watches human games. As well as helping the commentator to analyze the possible subsequent branches of the current game. At this point, our AI technology should be very reliable. Furthermore, we must ensure that our technology works reliably and safely while the tournament is in progress. It can also be personalized for different users. While

watching the tournament online, they can try their own moves to see the subsequent changes in the winning percentage.

- **Professional Go players – Fairness and Accountability – 1:** Our concern is that players may use Go AI technology to cheat. If this happens, it will not only affect the fairness of the tournament but also be detrimental to the spectacle of the tournament, discouraging the public from playing Go. Another situation is that there may be unscrupulous people who use AI to predict the win rate of a tournament for profit. Also, players are concerned about who should be responsible for accidents and whether the user or the inventor should be responsible for mistakes made by AI technology.
- **Go enthusiasts - Transparency – 3:** We want to see in real-time the next highest winning position recommended by AI. We also want to be told how the information is personalized. Are we all getting the same AI-assisted information, and is there any personalization of the user experience that is being done privately because of collecting user behavior? For example, the user has a very low level. Although they may not be able to read complex variation diagrams, they still want the Go AI to give the user information about all the changes in the situation. This would allow the user to be more informed when watching the game.

4.3. Scenario III

A player is playing against the Go AI. This player could be anyone, such as a professional or casual enthusiast.

- **AI Technicians - Inclusiveness – 5:** Our AI Go technology must meet a wide range of human needs and experiences. Specifically it needs to provide different difficulty level options for different users. As well as possible additional optimization and personal customization for people with disabilities. For example, verbal announcements for the visually impaired. Ensuring equitable access to our technology for all populations.
- **Professional Go players – Inclusiveness – 4:** From a professional player's perspective, it is hoped that Go AI technology can be used for personalized training. For example, can this technology collect our own historical games and use them to customize our game strategy, thus helping us to fix our technical flaws? Professional players may not be able to fully identify their opponent's weaknesses in their regular practice. If Go AI technology could mimic the opponent's playing style, professionals would support the deployment of this technology.
- **Go enthusiasts – Inclusiveness – 3:** Enthusiasts may be concerned that AI is all about unbeatable difficulty. Hopefully, they can choose to

train and improve themselves slowly from the basic difficulty and find the AI difficulty that best suits their level. If Go AI technology can accommodate all levels of Go enthusiasts. Then this is an advantage that cannot be matched by real people, exclusively by AI Go technology.

4.4. Conceptual Investigation Summary

After a conceptual investigation around human values in technology by substituting different stakeholders in three different scenarios, our team came to the following conclusions about the issues or concerns that may arise related to the deployment of Go AI technology:

1. The value principle of greatest concern to the public will probably be the **Inclusiveness** of the technology. They do not want this technology to be the exclusive domain of a small group of people. Not wanting the technology to become a privilege, a tool for the few to exploit the many. It is also hoped that the technology will meet a wide range of human needs. To provide a good personalized experience for different people, including but not limited to people with disabilities, etc.
2. The most immediate stakeholders, the professional Go players, are probably most concerned about the **Fairness** and **Accountability** of this technology. They are afraid that someone will use this technology to gain an undue advantage in a tournament.
3. The **Reliability and safety** of Go AI technology, as well as **Security and privacy**, was also a highly cited principle of value. The technology companies and the technicians involved are concerned about the possible legal risks this technology may pose for them.
4. In general, stakeholders have a positive opinion of the technology. They also scored high in our simulated "Judgment Call" game. Their expectations for Go AI technology are higher than their concerns.

5 Analysis of similar technologies

The year 1997 was a historic year for another game of complete information, International Chess. It was the year that IBM's supercomputer Deep Blue became the world's first artificial intelligence to beat the world chess champion after six rounds. Our report will examine the use of chess AI technology and inform whether Go AI technology should be deployed.

5.1. Case Study I

A serious robot accident occurred at a recent open chess tournament in Moscow: a computer-controlled robot arm operated too fast and crushed the finger of a 7-year-old boy it was playing against into a fracture.

Figure 2. A chess robot broke a 7-year-old boy's finger.
https://www.youtube.com/watch?v=jJxS8GmV5hg&ab_channel=AllNews

According to the Russian news site TASS (2022), the robotic arm performing the chess operation made a movement error, probably due to a programming bug. The boy rushes back before the robot can complete its move. The robot then pinches the boy's finger. The staff rushed to help the boy free his finger from the robot's "vise" and send him to the hospital, but the boy's finger was fractured. The video was uploaded to Twitter and widely distributed, and the boy's parents have contacted the local prosecutor's office to file charges.

This is undoubtedly a violation of the value principle **Reliability and safety**. However, it is impossible to determine the exact cause until we see the detailed application. However, first of all, we can exclude that the robot was artificially designed to "discipline" because it is one of the key principles of the three international principles of robotics that robots cannot harm humans. The child's finger was likely positioned in such a way that the robot calculated the position of the piece that needed to be moved next, and that the child happened to block the piece after the robot formed the trajectory command. This is because the robot does not make any visual judgments after it has formed a move command. In other words, the robot "thinks" it is holding the piece, but it is actually the child's finger. This may just be a coincidence. This example demonstrates that AI technology can deviate from full human control and conventional understanding at this stage. These involve robotics ethics, which also corresponds to the value principle of **Accountability** and requires

strict monitoring and control of the robot's learning evolution. Therefore, people need to remain highly alert to the risks of AI applications.

5.2. Case Study II

On September 4, 2022, in the third round of the Sinquefield Cup of chess in St. Louis, USA, the world's top-ranked Norwegian player Magnus Carlsen unexpectedly lost to 19-year-old American youngster Hans Niemann, ending a 53-game unbeaten streak and blowing up the biggest upset in the world of chess this year. According to Davis (2022), Magnus accused Hans of being involved in cheating using chess AI technology, and rumors were circulating on the Internet that Niemann may have used vibrating wireless anal beads to indicate what moves he should make.

Following Carlsen's public rebuke of Niemann, Chess.com (2022) also released a 72-page investigative report detailing their findings on Niemann's previous cheating. The investigation report provided extensive evidence based on data analysis, relied on several cheating detection tools to determine whether cheating occurred, including comparing a player's moves to the supercomputer's recommendations, and incorporated input from chess fair play experts. The report concluded that Niemann may have used illegal cheating in more than 100 online tournaments through 2020, including some with win bonuses, and his results were statistically unusual.

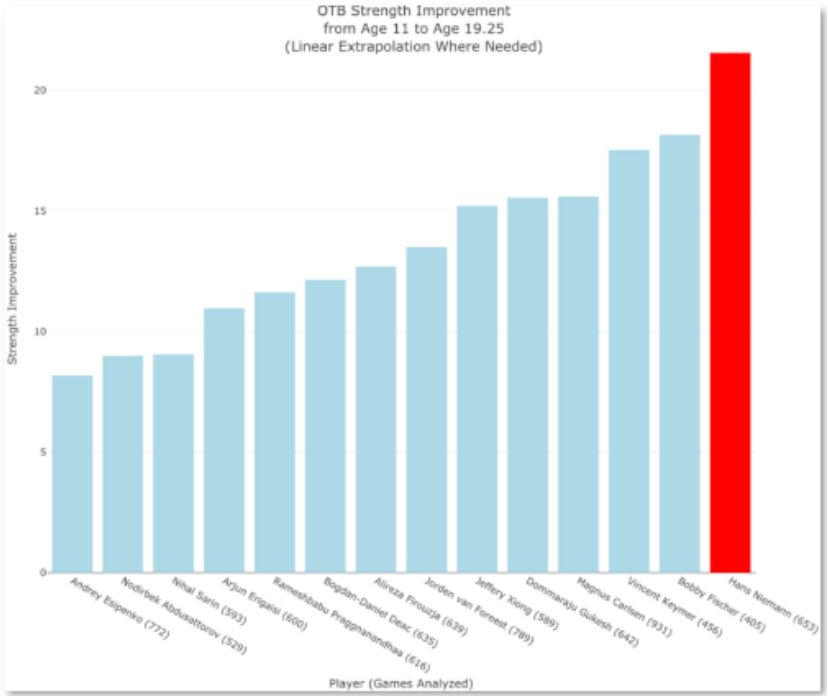

Figure 3. Chess.com(2022) releases heavyweight report accusing Niemann of possibly cheating hundreds of times

This incident undoubtedly involves the principle of **Fairness**, which is the most important value for professional players. The cheating by players using chess AI technology has caused real direct harm to other players and long-term damage to the entire chess game. Future chess tournaments will only be able to maintain the purity of the human intellect of the sport by using ever-escalating high-tech anti-cheating techniques to prevent players from ever-escalating high-tech cheating techniques. Once a player succeeds in cheating using AI technology, the game loses its meaning.

The entire chess community should take the issue of AI cheating seriously. It is an existential threat to the sport. Tournament organizers should seriously consider increasing security measures and cheating detection. We have different moral values. For some people, fairness is a very important value. When someone cheats and creates an uneven playing field, we think it is very unfair. While this is the result of subjective human malice, there is no doubt that chess AI technology is an accomplice to them. So it is all the more important for AI technology companies to take responsibility for this. While ensuring the **Inclusiveness** of their technology, i.e., meeting a wide range of human needs, they also need to restrain the possibility that their technology could be used unethically.

6 Comparison between Analysis of Similar Technologies and Conceptual Investigation

First of all, it can be found that the problems that may arise from the value principle **Reliability and safety** summarized in Conceptual Investigation have occurred after the deployment of similar technologies (Case Study I). This proves that the concerns we inferred by playing the role of a stakeholder are meaningful. However, the principles involved in the actual situation have emerged in ways that we had not previously anticipated. We were concerned from the perspective of the technical people involved about the possible legal risks if the values of "reliability and security" were not considered. In reality, however, users should probably be more concerned about these values. For example, in Case Study I, the chess AI robot hurt a young boy's finger, causing a serious injury like a broken bone. This is a situation we had not anticipated before. This practical experience can be complementary to our conceptual investigation. That is, the technical staff should not only consider the value of "reliability and security" from their own perspective and should not only think about how the AI system can avoid legal risks to themselves. They should also consider whether AI technology can be used reliably and safely at all times and whether it can be used in a way that is not harmful to humans.

We can also see that in Case Study II, professional players use chess AI technology to cheat, which addresses the value principle of **Fairness** that was

a concern when playing as a professional in Conceptual Investigation. The flawless performance of AI in these industries has given some players a way to cheat to improve their winning percentage. It is conceivable that without AI, it would be nearly impossible for a sport like Go and chess to cheat in a world-class tournament. Nevertheless, the advent of AI has started to raise concerns about the fairness of Go and chess tournaments. At the same time, the overwhelmingly inclusive nature of chess AI technology has created problems that we did not anticipate, embracing aspects that can be exploited unethically. We need to consider these situations before deploying Go AI technology.

Overall, more situations need to be considered before deploying Go AI technology. Suppose some robotic arms or other mechanical products need to be manipulated by Go AI that have direct contact with humans. In that case, attention needs to be paid to human safety in any situation. There also needs to be some restrictions on the use of AI to prevent it from being used in some illegal scenarios to harm society and disrupt order.

7 Deployment Recommendations

After a detailed analysis, we believe that Go AI technology can be deployed, but with some adjustments. First of all, **Reliability and safety** are the most important topics for technicians as well as Go enthusiasts. While the three principles of robotics may sound plausible, many arguments have demonstrated that there are cases where they are inadequate. So it is important to take more of an edge case to avoid them, assuming that AI developers write code that detects whether another player (assumed to be human) has skipped the correct rotation of the robot arm. Assume that the robot arm and board are behind an obstacle, and it needs to detect in time that a human has jumped over the obstacle into the position where the robot arm is. The fixture should have the pressure-sensing capability. This may provide feedback that what is being squeezed is different from the consistency of the pieces. At some threshold, the AI running the gripper should automatically release on the assumption that something other than a piece is captured. This can override all other objects or biological elements that may have entered its grasper incorrectly. We will gradually add more edge cases in the future based on practice to make our Go AI technology more secure and reliable.

Secondly, Go professionals will be more concerned about whether someone will use the AI to cheat improperly, and the **Fairness** of our Go AI is their primary concern. For example, many universities are already using text-based plagiarism detectors such as Turnitin to detect if a student's paper was written with the aid of an AI. In the future, if our Go AI is at risk of being exploited for

cheating. We should add more types of detection to our deployment, comparing moves made to those recommended by the AI engine and comparing players' past performance to their historical ability. This will determine if they are using AI to cheat and protect our Go AI from being used for inappropriate behavior.

Another point that is most relevant to fairness is the law. In deploying our AI technology, legislators should be pushed to enact strict laws governing the various uses of AI. However, the effort to draft and enact such laws is an incremental process that we will slowly refine over the next few practices.

References

Chess.com Team. (CHESScom. (2022, October 4). Hans Niemann Report. Chess.com. https://www.chess.com/blog/CHESScom/hans-niemann-report

Davis, S. (2022, September 26). Magnus Carlsen accuses Hans Niemann of cheating at chess, Sinquefield Cup. Insider; Insider. https://www.insider.com/magnus-carlsen-accuses-hans-niemann-of-cheating-at-chess-sinquefield-cup-2022-9

TASS. (2022, July 28). Moscow Chess Federation calls for strengthening protection after robot breaks boy's finger. TASS; TASS. https://tass.com/sport/1486181

Xinhua. (2020, January 11). Chang Hao: The root of Chinese Go prosperity lies in the masses. Sina.com.cn. https://k.sina.com.cn/article_213815211_0cbe8fab02000tn7w.html#/