

# Next Steps for AI

*Mengyu Jackson*

Since 1997, when Deep Blue beat Gary Kasparov (then world champion) in a 6 game chess match, Artificial Intelligence has seemed poised to revolutionize the world. Chess was seen as a proving grounds for real, human-comparable intelligence, but transitioning the highly specialized hardware and software that beat Kasparov to more general problems remained extremely difficult.

There are many ways in which chess is "easy" compared to real-world problems: 1.) The game has perfect information for all players (nothing is hidden). It is never required to guess or extrapolate information based on opponent's behavior. 2.) The game is zero-sum (what is good for one player is bad for the other, and vice versa). 3.) It only involves 2 players (no coordination or cooperation is required, nor is anticipating cooperation or coordination between multiple opponents). 4.) Chess is a turn based game. After you observe your opponent's move, the game will not change while you're thinking of which move you'd like to play.

Even within the limited space of 2 player, zero sum, perfect information, turn based games, it was surprisingly difficult to progress to larger games. For nearly 20 years after computers asserted their dominance in chess, computers had difficulty surpassing even amateur players in Go (which is played on a 19x19 board instead of 8x8).

I was in college at VCU when AlphaGo overcame what felt like 20 years of inertia to beat former world champion Lee Sedol. My husband and I stayed up late watching the matches (which were played during night and early morning hours in our timezone). Neither of us knew much about Go, but we had a sense that we were about to witness a historic moment in AI. Since then, advancements in AI have come rapidly, pushing many interesting frontiers.

AlphaZero dethroned AlphaGo using a learning process that only relied on a flexible Neural Network and self-played games (rather than custom evaluation rules written by experts or learning from human games as earlier engines did). For the first time, we saw that an AI could play games at an elite level without input from human experts at any stage. To prove the flexibility of the approach, AlphaZero was also trained to play Shogi and Chess, where it also reached a world class level very quickly. Since experts are expensive and rare in many fields, being able to create an AI expert without requiring a human expert during development or training greatly expands the number of domains AI could realistically be applied in the near future. AlphaStar beat a top human player in Starcraft (a real time game with imperfect information), using a similar training process to AlphaZero.

Pluribus became the first computer program to beat top humans in 6 player poker (a multiplayer game with imperfect information) in 2019, building on some of the same techniques used by AlphaZero.

With all of these recent advancements, AI seems poised to tackle a huge range of problems that seemed frustratingly out of reach only a few years ago.

One of the most interesting frontiers I'm watching is "Explainable AI". Chess Engines have played chess at a grandmaster level for many years. For players looking to get stronger, however, they needed something more than just identification of the best move at each position. They needed to know "why" a move was best. DecodeChess is stepping up to solve this problem for chess: providing human-understandable explanations for why each move is good (or bad) based on very strong engine analysis.

As AI continues to find more uses in solving real-world problems, the need to understand and explain its reasoning will become increasingly vital to maintain confidence in the accuracy and fairness of the decisions they make, which have increasingly serious effects on our daily lives.