Summary of a Research Paper:
Mastering the game of Go with deep neural networks and tree search by David Silver, et al.

For the first time in history, a game-playing A.I. has surpassed humans in the realm of Go (AlphaGo defeated the 2nd Dan, three-time-consecutive Go-champion of Europe, Fan Hui, 5-nill). Overnight, AlphaGo superseded humans and prior Go-bots by a 99.8% win rate (even defeating Patchi, which performs one hundred thousand simulations per move). This sent ripples across the Asian (Go-playing) world and proved an even greater accomplishment for the A.I. community than DeepBlue's win against Chess grandmaster Gary Kasparov. To understand why, you should appreciate that Go is to Chess as Chess it to Checkers (i.e. Go has an even larger search space and far trickier board positions and moves to evaluate). Furthermore, AlphaGo evaluated "thousands of times fewer positions that Deep Blue did in its chess match against Kasparov"! For the first time, an A.I. has learnt to play Go "at the level of the strongest human players, thereby achieving one of Artificial Intelligence's grand challenges."

The secret to AlphaGo's success lies in the way it plays. Brute-force methods (i.e. exhaustive searches) are infeasible for games like Chess and Go due to the enormous breadth of legal moves and game depths (e.g. Chess has a breadth and depth of 35 and 80, giving 35^{80} variant sequences of moves, Go has a whopping breadth and depth of 250 and 150, resulting in an astronomical 250^{150} possible sequences! These are both numbers higher than the number of atoms in the universe). It direct contrast to this approach, AlphaGo plays Go in a manner far closer to the way humans play. Using no lookahead searches at all, AlphaGo decides its next move using a form of 'artificial intuition'.

AlphaGo uses Convolutional Neural Networks to create a representation of the current position on a 19x19 Go board. It then selects a next move using a non-linear estimation function stored in a supervised learning neural network (trained on 30 million Go game positions on the KGS Go Server and the subsequent moves made by human expert players). Once this policy network learnt to accurately predict expert human moves, the policy network was optimised using self-play (a form of reinforcement learning whereby thousands of random Go games are played against itself. Strategies that win receive a positive reinforcement score; +1, whereas a loss receives a negative reinforcement of -1). Rather than merely learning to predict human expert moves, AlphaGo learnt to select moves which more likely lead to a win (ironically, human players are now studying these moves to learn more advanced Go playing strategies invented by AlphaGo!).

Discussion
Developing a neural network that can approximate an evaluation function for the current position of a game and predict the next best move without using any lookahead is both quicker and more efficient than traditional lookahead strategies. Pre-training the policy network on expert moves and then optimising it using reinforcement learning allows quicker training times, however, I wonder how a learning strategy based on pure reinforcement learning would perform. Perhaps self-play, round-robin, tournament selections as a fitness function for a genetic algorithm could evolve an equally optimal policy network without requiring any expert information apriori (i.e. a neuroevolutionary appraoch).