

6.034

AlphaGo & AlphaZero

Dr. Howard Shrobe

October 18, 2019



Why we care about Go:

Go played Fan Hui, a professional 2 *dan*, and the winner of the 2013, 2014 and 2015 European Go Championships. Over 5–9 months in late 2015 AlphaGo and Fan Hui competed in a formal five-game match. AlphaGo won the match 5 games to 0.

The first game match between 18-year-old world champion (9 *dan*) Lee Sedol and AlphaGo, played in Seoul, South Korea between the 9th and 15th of March 2016. AlphaGo won all but the fourth game, which was won by a human player.

Michael Redmond, 9-dan Go player, amazed at AlphaGo's unusual move.



Move 37!! Lee Sedol vs AlphaGo Match 2

Some Go Basics:

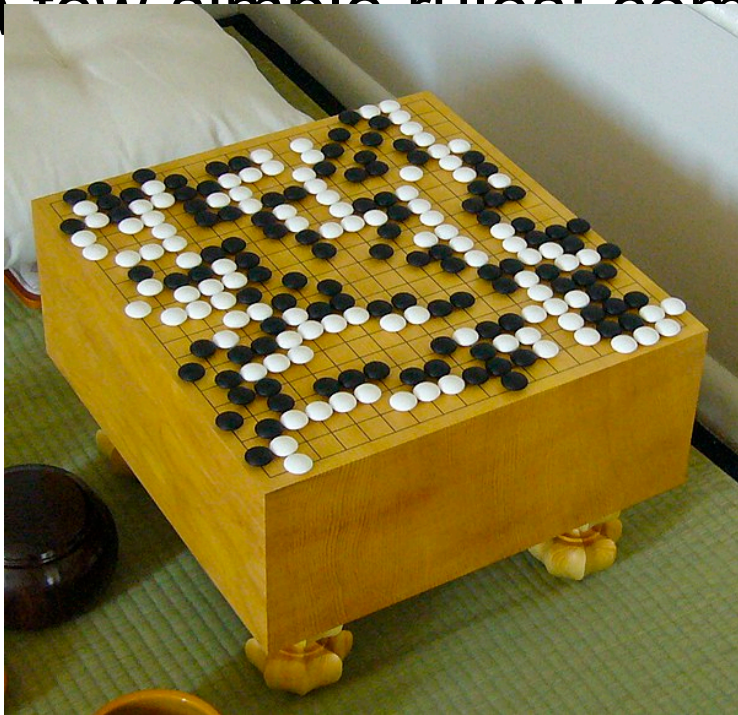
2 Player game – White and Black

Played on a 19 x 19 grid = 361 Intersection points

Playing = Putting a stone on an unoccupied intersection point

Stones don't move after being played

Only a few simple rules: complex emergent properties



Some Go Basics:

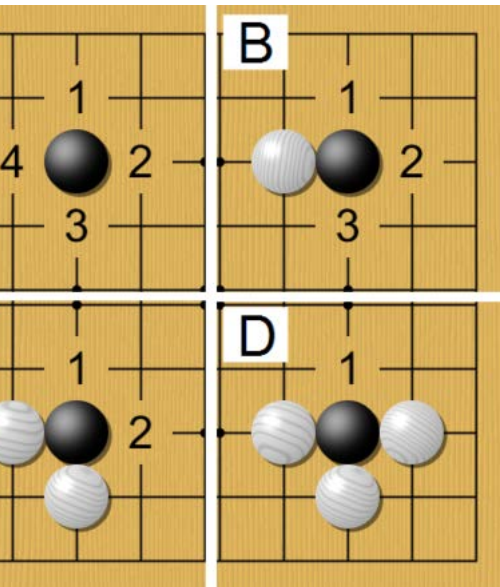
Liberty – A grid line to an unoccupied intersection point

A single stone can have up to 4 liberties

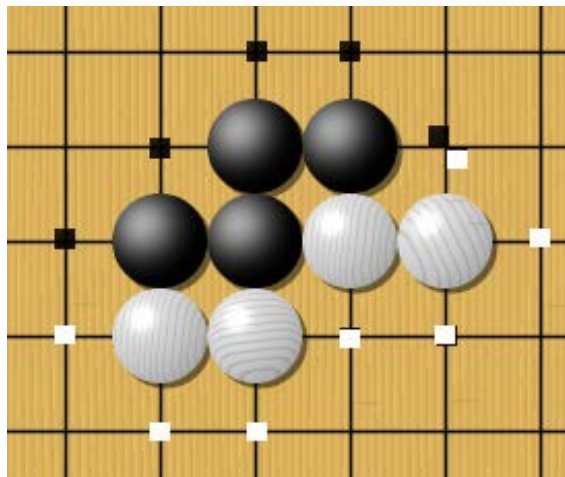
Chain: A set of same colored stones connected by grid lines -- not diagonals

Capturing: A group is captured when it has no liberties

“Liberties”

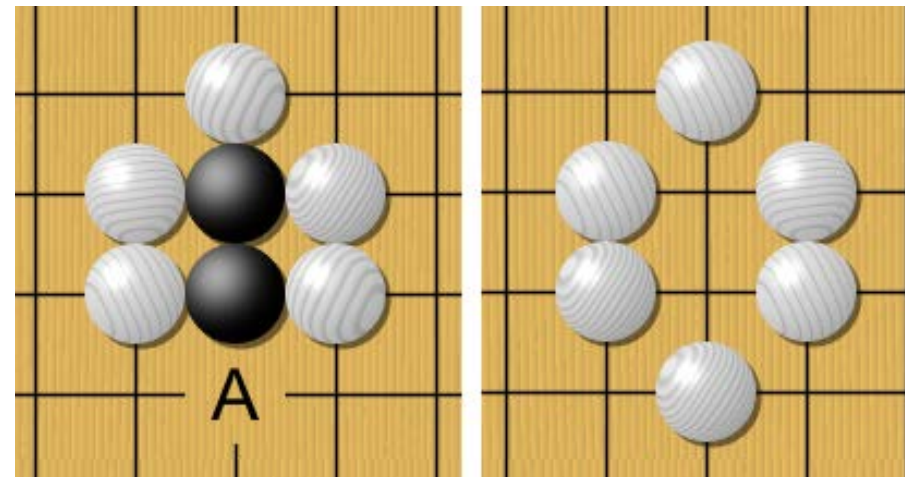


“Chains”



One black chain and two white chains, with their liberties marked with dots. Liberties are shared among all stones of a chain.

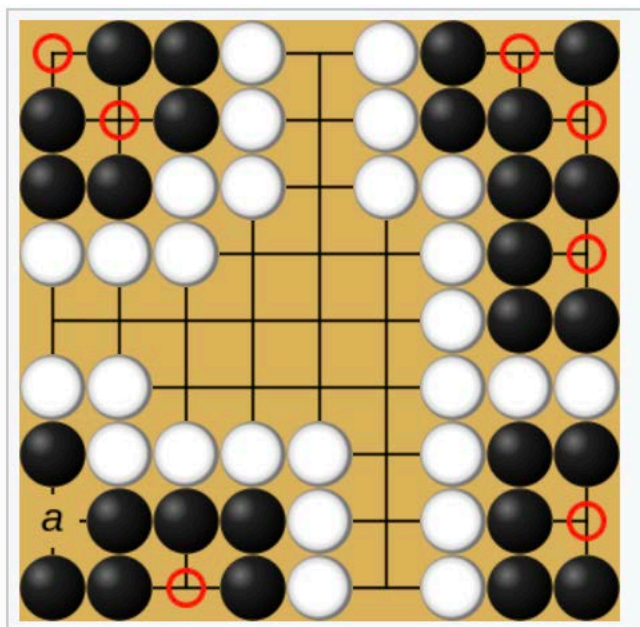
“Capturing”



If White plays at A, the black chain loses its last liberty. It is captured and removed from the board.

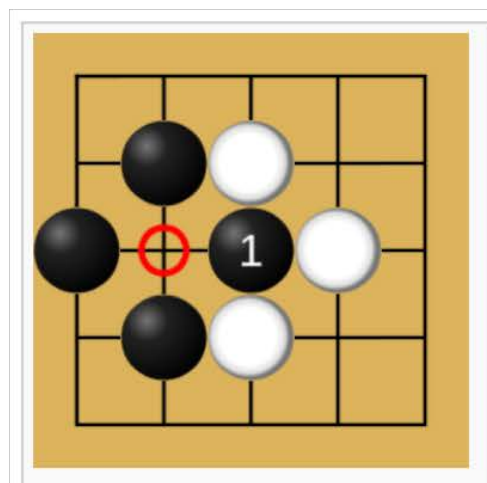
Go basics

2 Eyes = Live



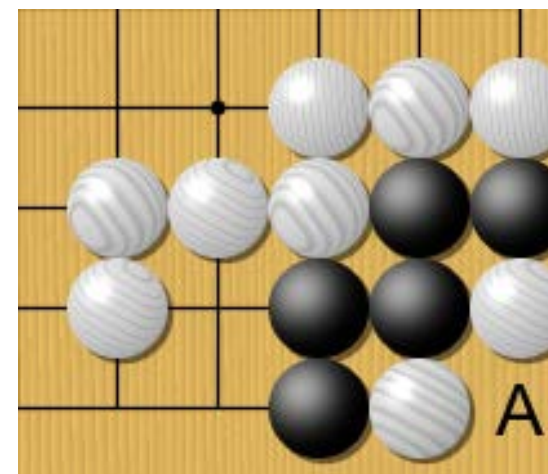
Examples of eyes (marked). The black groups at the top of the board are alive, as they have at least two eyes. The black groups at the bottom are dead as they only have one eye. The point marked a is a false eye.

Ko Rule



Players are not allowed to make a move that returns the game to the previous position. This rule, called the [ko rule](#), prevents unending repetition

Suicide



Under normal rules, White cannot play at A because that point has no liberties.

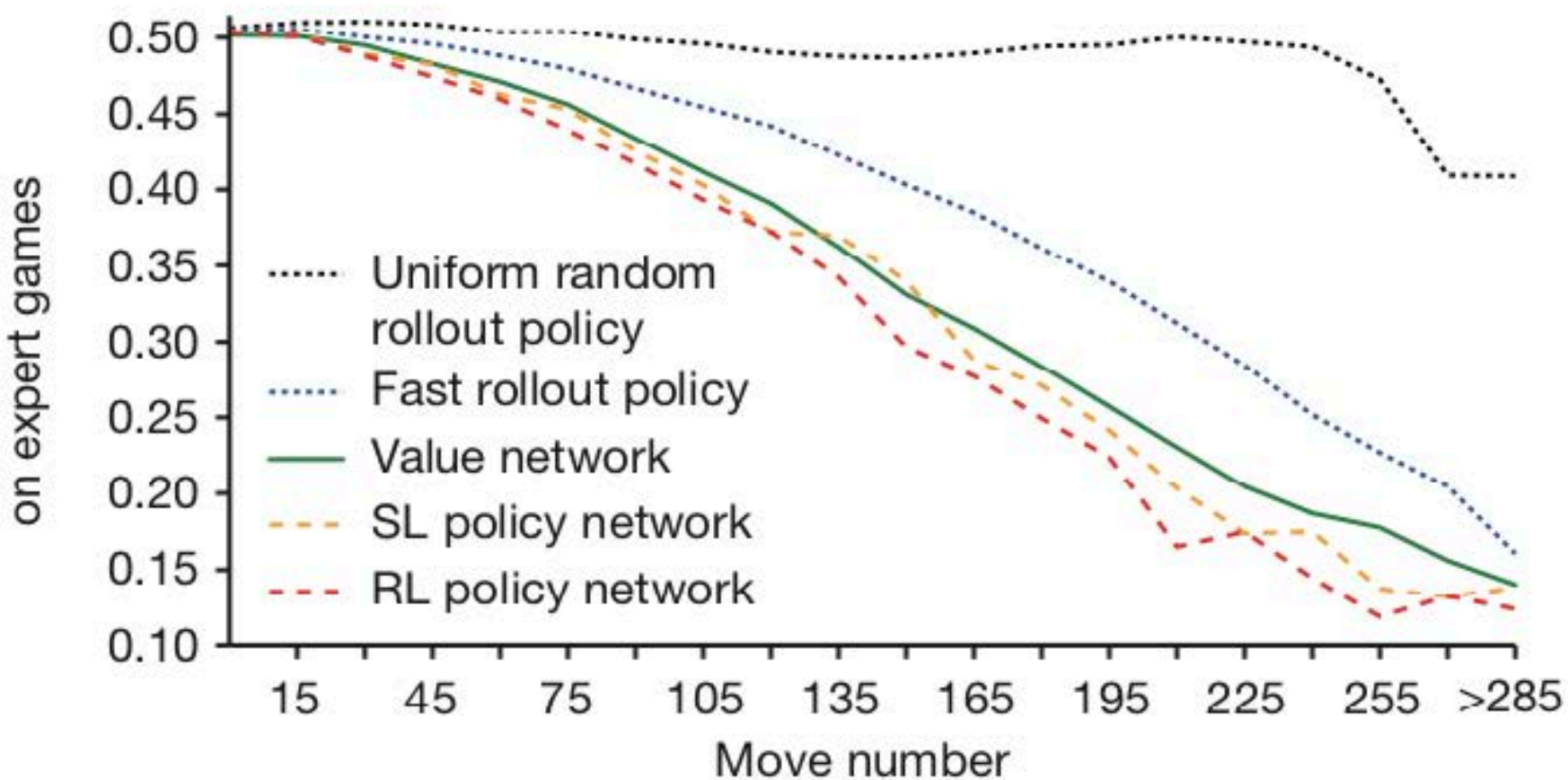
Features used in Policy Network

Extended Data Table 2 | Input features for neural networks

Feature	# of planes	Description
Stone colour	3	Player stone / opponent stone / empty
Ones	1	A constant plane filled with 1
Turns since	8	How many turns since a move was played
Liberties	8	Number of liberties (empty adjacent points)
Capture size	8	How many opponent stones would be captured
Self-atari size	8	How many of own stones would be captured
Liberties after move	8	Number of liberties after this move is played
Ladder capture	1	Whether a move at this point is a successful ladder capture
Ladder escape	1	Whether a move at this point is a successful ladder escape
Sensibleness	1	Whether a move is legal and does not fill its own eyes
Zeros	1	A constant plane filled with 0
Player color	1	Whether current player is black

Feature planes used by the policy network (all but last feature) and value network (all features).

Performance of differently trained Policy Nets



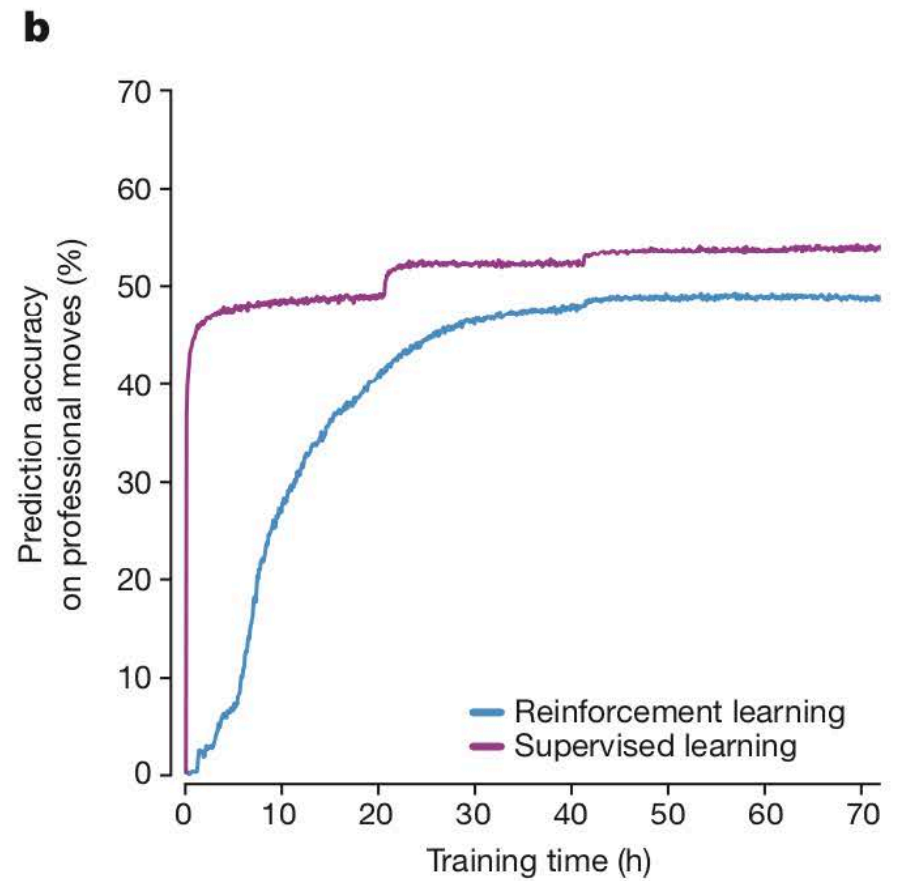
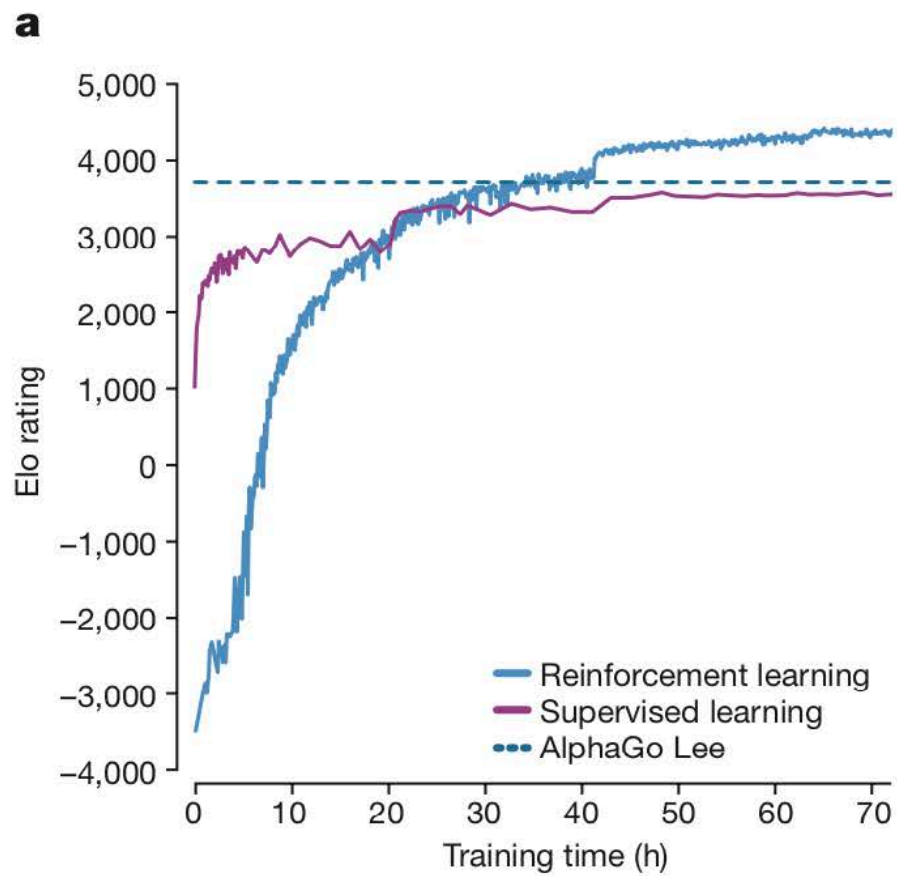
Configuration and performance

Configuration ▼	Search threads ◆	No. of CPU ◆	No. of GPU ◆	Elo rating ◆
Single ^[11] p. 10–11	40	48	1	2,181
Single	40	48	2	2,738
Single	40	48	4	2,850
Single	40	48	8	2,890
Distributed	12	428	64	2,937
Distributed	24	764	112	3,079
Distributed	40	1,202	176	3,140
Distributed	64	1,920	280	3,168

Configuration and strength^[64]

Versions ↕	Hardware ↕	Elo rating ↕	Matches
AlphaGo Fan	176 GPUs, ^[53] distributed	3,144 ^[52]	5:0 against Fan Hui
AlphaGo Lee	48 TPUs, ^[53] distributed	3,739 ^[52]	4:1 against Lee Sedol
AlphaGo Master	4 TPUs, ^[53] single machine	4,858 ^[52]	60:0 against professional players Future of Go Summit
AlphaGo Zero	4 TPUs, ^[53] single machine	5,185 ^[52]	100:0 against AlphaGo Lee 89:11 against AlphaGo Master
AlphaZero	4 TPUs, single machine	N/A	60:40 against AlphaGo Zero

AlphaGo Zero Performance



AlphaGo Zero performance

