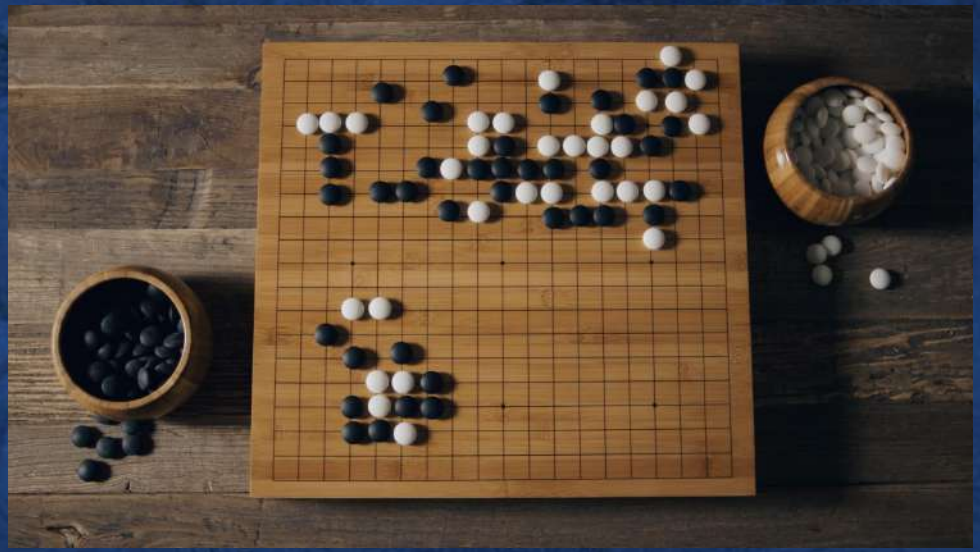
AlphaGo

Why is Go hard for computers to play?

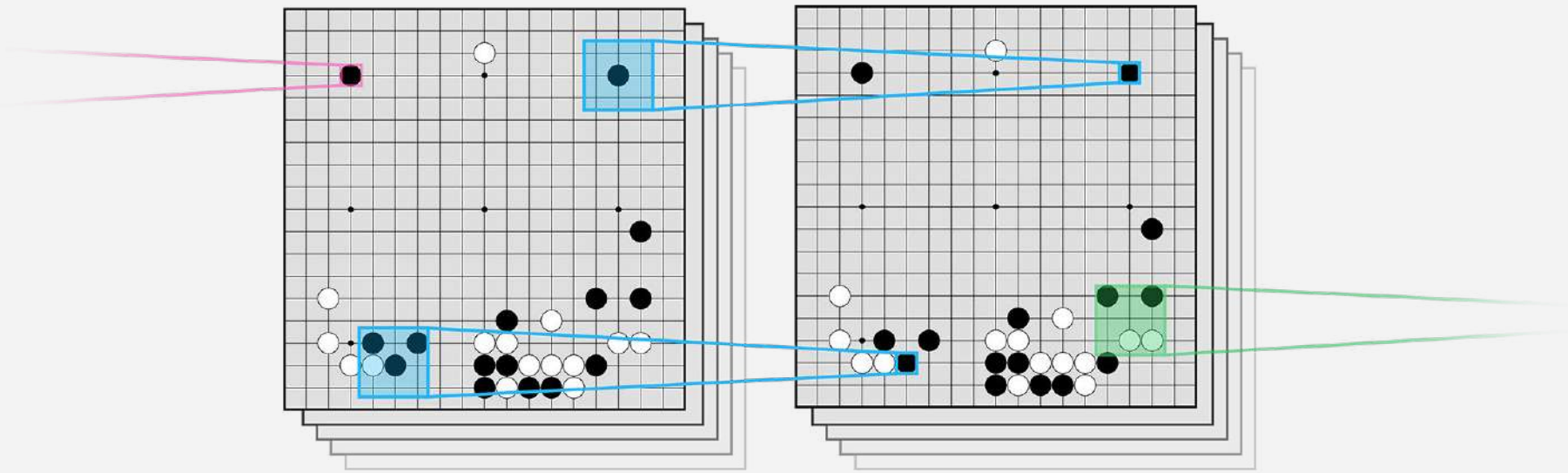
Game tree complexity = b^d

Brute force search intractable:

1. Search space is huge
2. “Impossible” for computers to evaluate who is winning

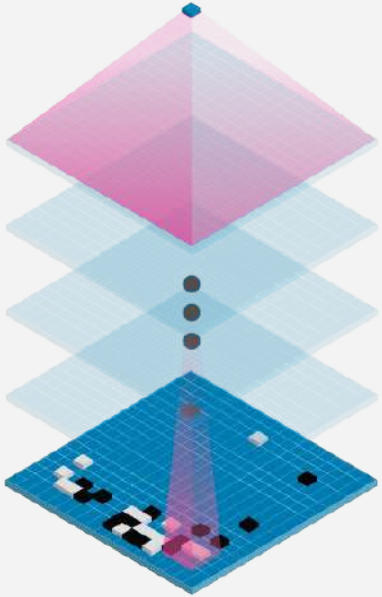


Convolutional neural network

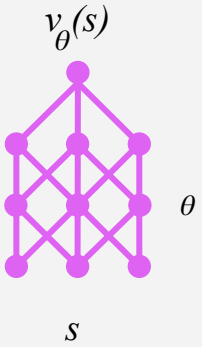


Value network

Evaluation

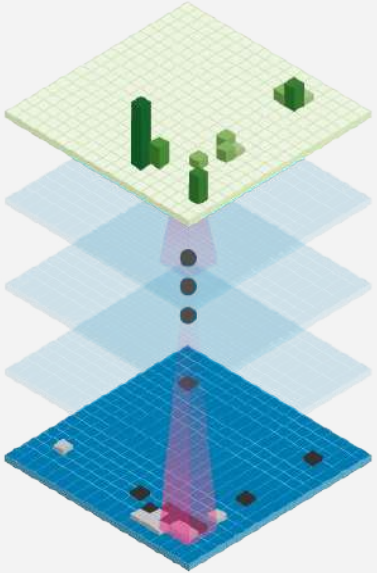


Position



Policy network

Move probabilities



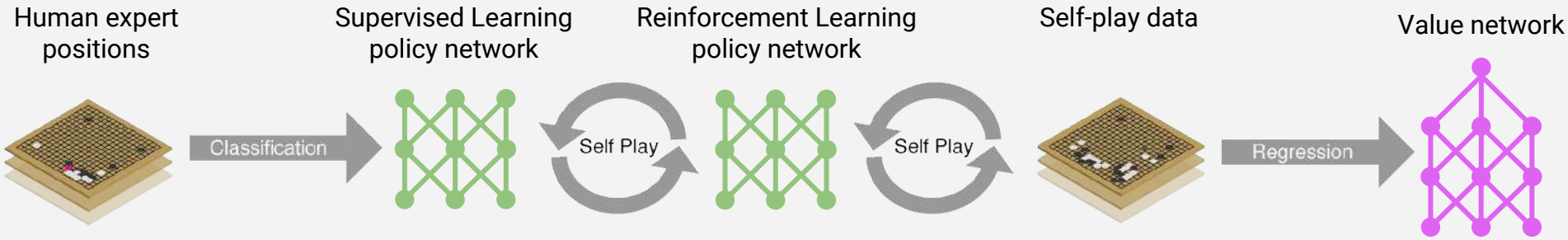
Position

$p_\sigma(a|s)$

σ

s

Neural network training pipeline



Supervised learning of policy networks

Policy network: 12 layer convolutional neural network

Training data: 30M positions from human expert games (KGS 5+ dan)

Training algorithm: maximise likelihood by stochastic gradient descent

$$\Delta\sigma \propto \frac{\partial \log p_\sigma(a|s)}{\partial\sigma}$$

Training time: 4 weeks on 50 GPUs using Google Cloud

Results: 57% accuracy on held out test data (state-of-the art was 44%)

Reinforcement learning of policy networks

Policy network: 12 layer convolutional neural network

Training data: games of self-play between policy network

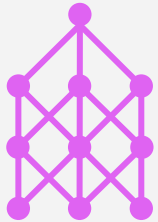
Training algorithm: maximise wins z by policy gradient reinforcement learning

$$\Delta\sigma \propto \frac{\partial \log p_\sigma(a|s)}{\partial \sigma} z$$

Training time: 1 week on 50 GPUs using Google Cloud

Results: 80% vs supervised learning. Raw network ~3 amateur dan.

Reinforcement learning of value networks



Value network: 12 layer convolutional neural network

Training data: 30 million games of self-play

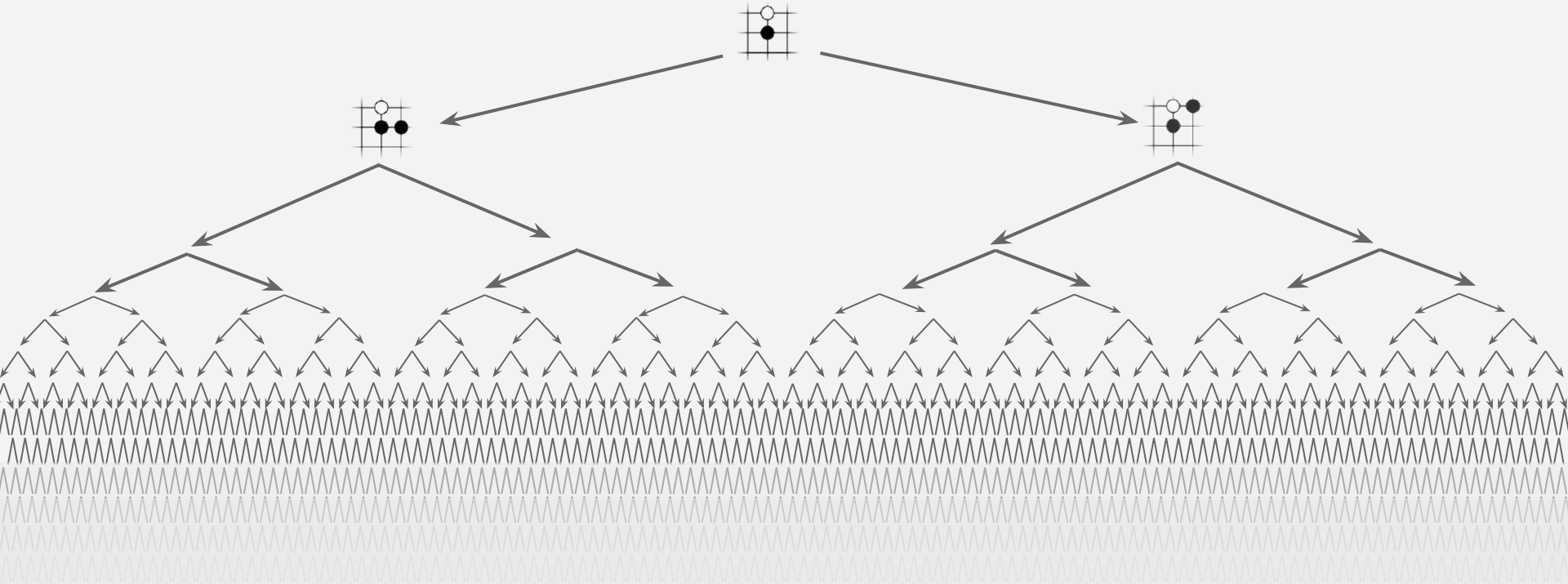
Training algorithm: minimise MSE by stochastic gradient descent

$$\Delta\theta \propto \frac{\partial v_\theta(s)}{\partial\theta}(z - v_\theta(s))$$

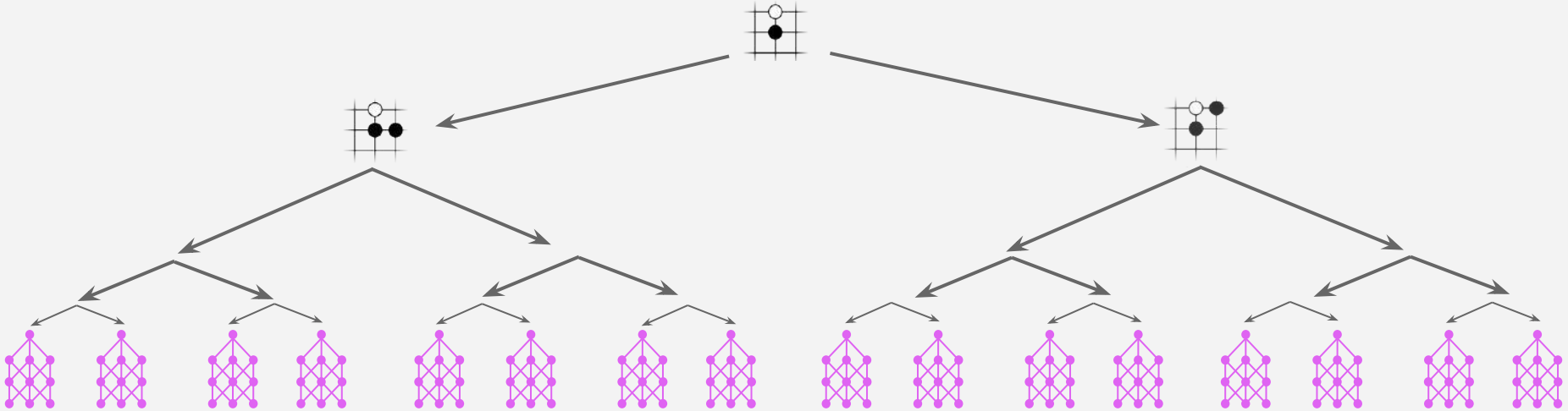
Training time: 1 week on 50 GPUs using Google Cloud

Results: First strong position evaluation function - previously thought impossible

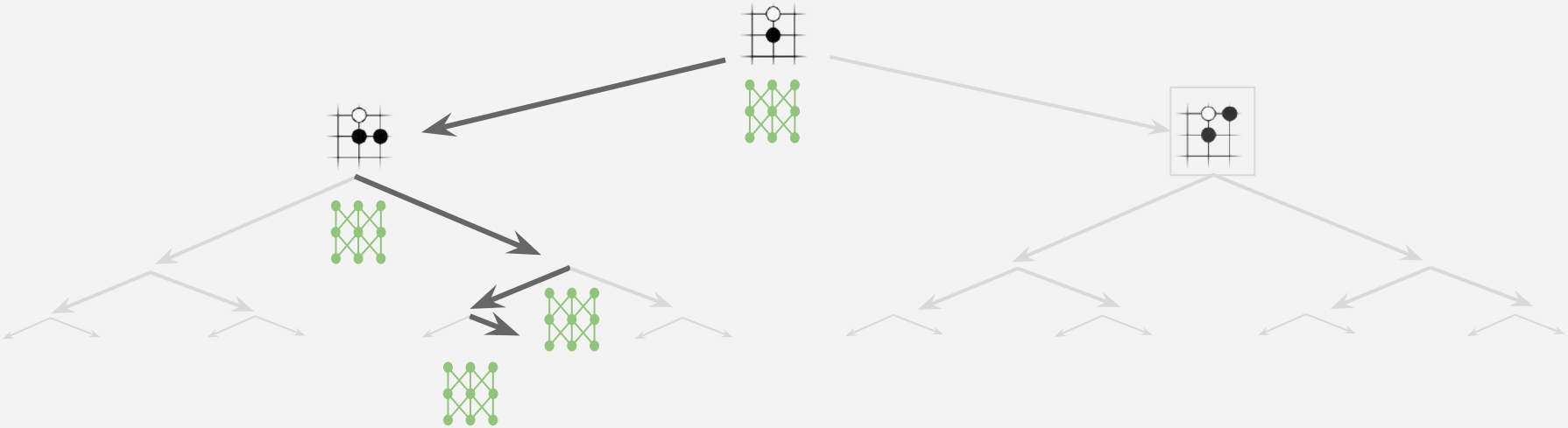
Exhaustive search



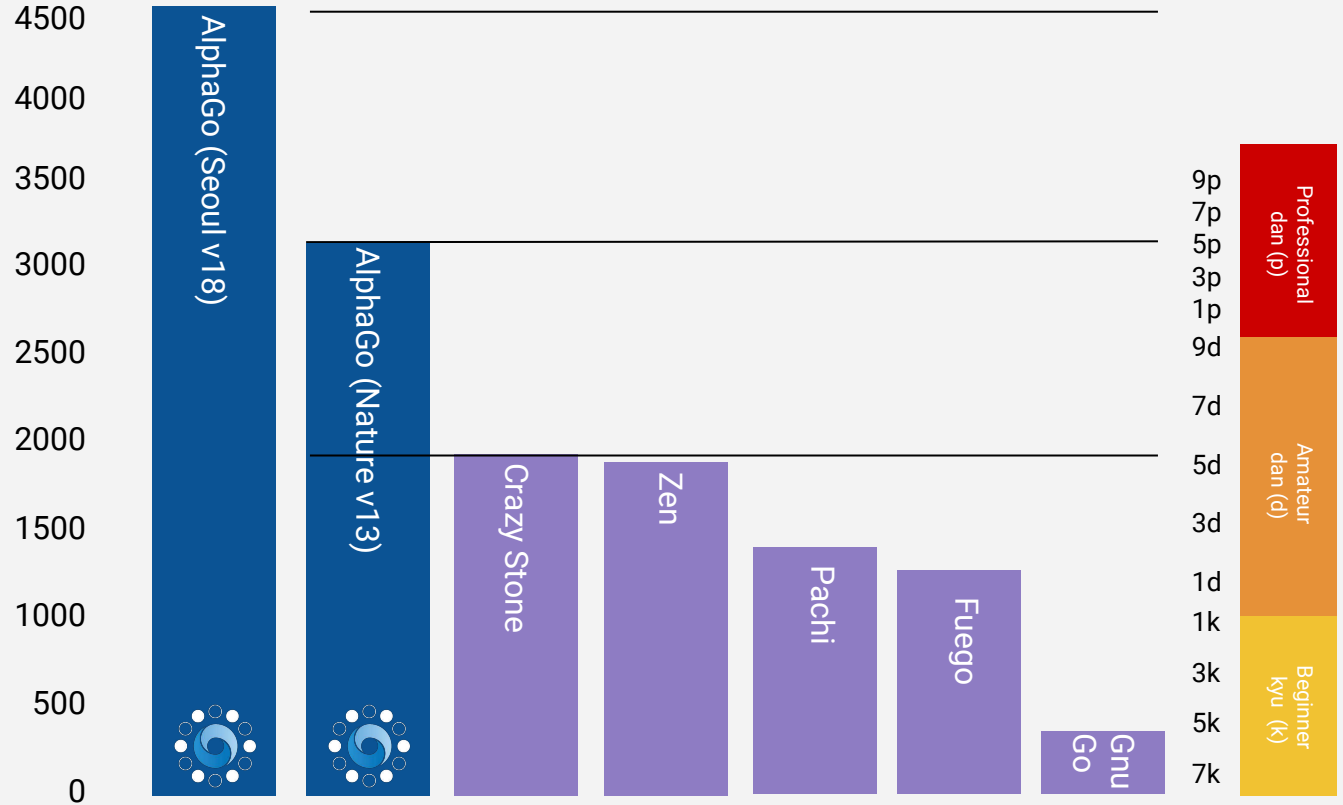
Reducing depth with value network



Reducing breadth with policy network

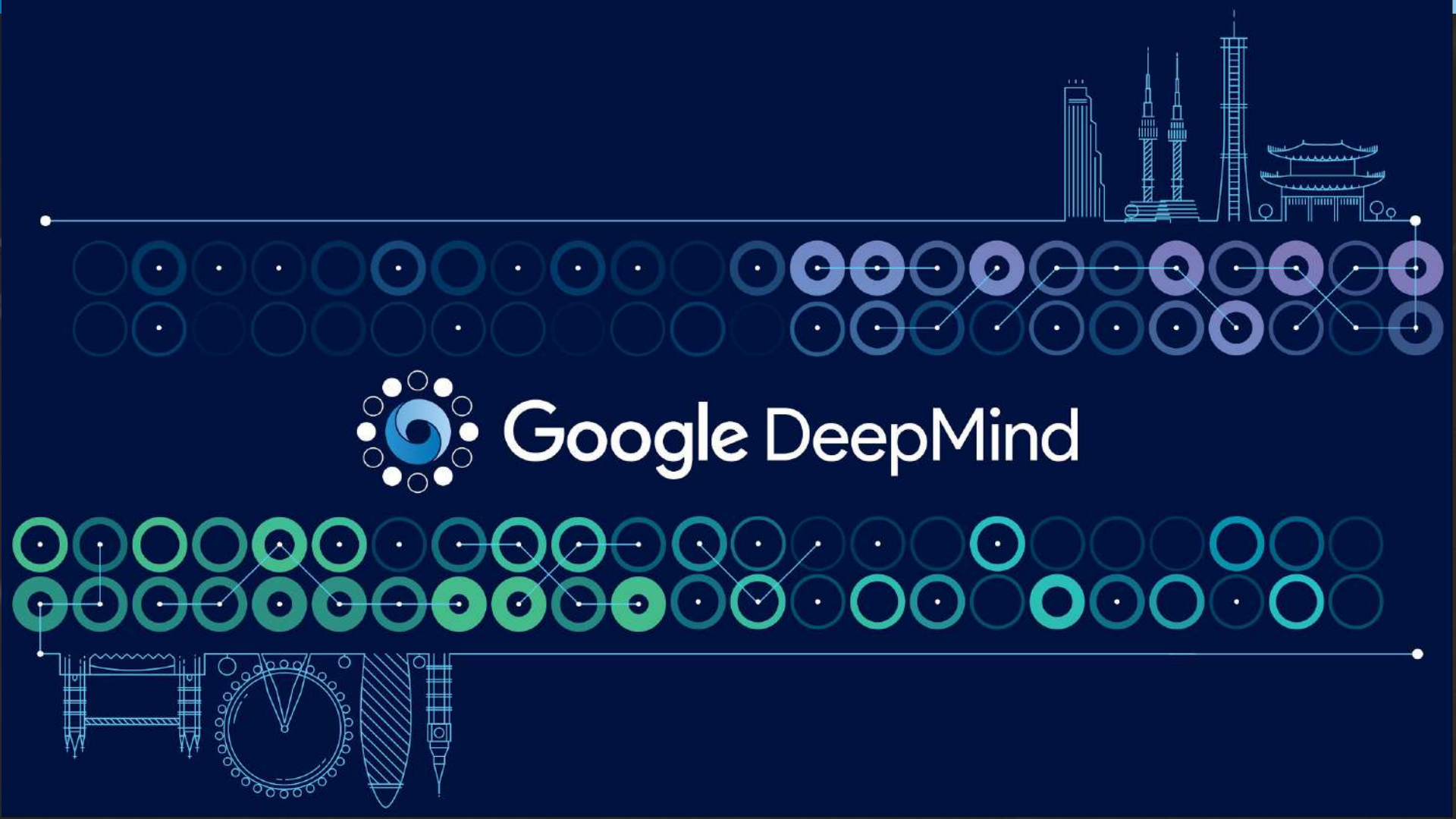


Evaluating AlphaGo against computers



Computer Programs	Calibration	Human Players
AlphaGo (Mar 2016)	*DeepMind challenge match* *4-1*	Lee Sedol (9p) Top player of past decade
Beats		*Beats*
AlphaGo (Oct 2015)	*Nature match* *5-0*	Fan Hui (2p) 3-times reigning Euro Champion
Beats		*Beats*
Crazy Stone and Zen	*KGS*	Amateur humans

What’s Next?


Google DeepMind