

# **Un résultat spectaculaire**

AlphaGo

---

AlphaGo un bon exemple de développement industriel en deep learning.

- **deep artificial neural networks**
- **Monte Carlo tree search**
- **Reinforcement Learning**
- **Très gros moyens !**



### Go Vs Echecs

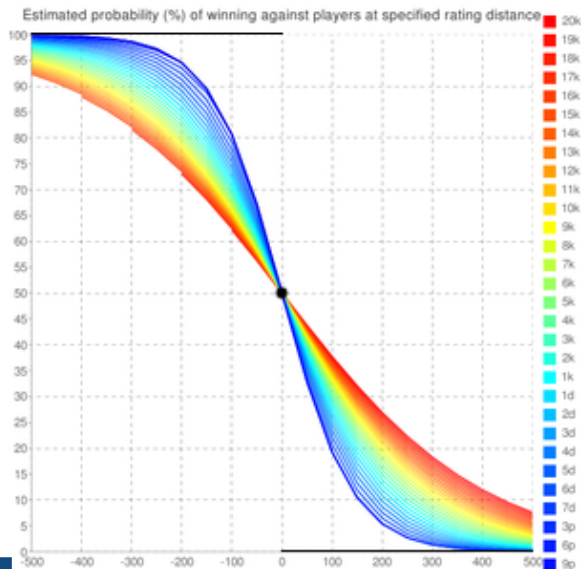
**Le nombre de coups**

$\approx 200$  contre  $\approx 35$

**Le nombre de tours**

$\approx 300$  contre  $\approx 40$

$\Rightarrow \approx 10^{120} \lll \approx 10^{690}$



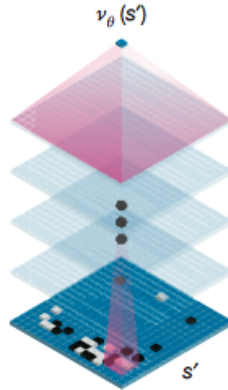
EGF	Classement
2940	9 dan pro
2820	8 dan amateur
2700	7 dan amateur
2600	6 dan amateur
2500	5 dan
2400	4 dan
2300	3 dan
2200	2 dan
2100	1 dan
2000	1 kyu
1900	2 kyu
1800	3 kyu
1500	6 kyu
1000	11 kyu
500	16 kyu
100	20 kyu

Plusieurs versions :

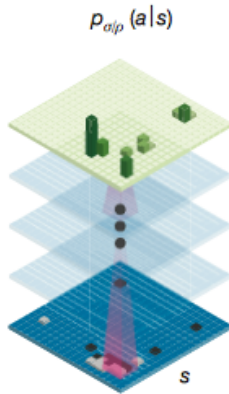
- **AlphaGo Vs Fan Hui 5 à 0** Octobre 2015
- **AlphaGo vs Lee Sedol 4 à 1** Mars 2016 (Prédiction de Rémi Coulom en 2016 : Résultat pas avant 15 ans)
- **AlphaGo-zero** considéré comme ayant au moins 500 EGF au dessus de AlphaGo. 2017

- $\text{softmax}(\text{convolutions}^{13}(19 * 19 * 48\text{bits}))$
- 48bits : input sur la position, la configuration , ...
- Appris sur données supervisées

Value network

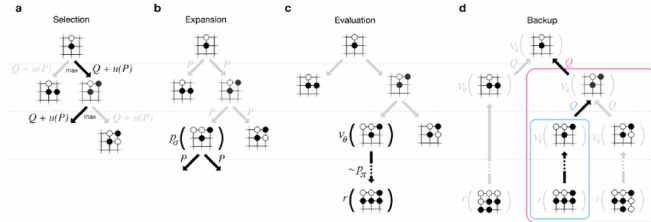


## Policy network



- Appris sur du self-play
- A l'aide d'un MCTS qui apprend à générer des parties efficace en fonction de la value function (fixée) et policy function (que l'on est en train d'apprendre)

## Monte-Carlo tree search



- Selection par maximum de  $(Q\text{-value} + u(P))$  où  $P$  est une proba à priori pour chaque coup.
- Expansion en fonction de la policy qui produit des probabilités qui sont alors stockés pour chaque action



2 phases pour apprendre :

**Value network :**

3 semaines  $\times$  (50  $\times$  TPU)

**1 million de parties APV-MCTS :**

1 journée  $\times$  (5000  $\times$  TPU)

**50 TPU  $\approx$  2000 TFLOPS**

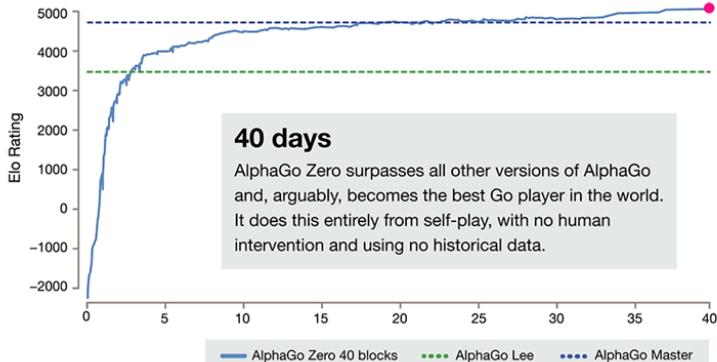
core i7 :  $\approx$  0.07 TFLOPS (donc il en faudrait  $\approx$  30 000)

AlphaGo vs Lee Sedol (ou Fan Hui)

1 202 processeurs et 176 TPU

# AlphaGo Zero

- Plus simple : 4 TPU
- Plus fort : bat alphago 100 :0
- Plus général : self-play uniquement



- Encore plus général
- Bat tous les bot d'échec, go et shogi

## Relativisons

- Un cerveau humain consomme  $\approx 20$  Watts
- AlphaGo  $\approx 440\,000$  Watts (440 grille-pains)
- Un tableur gagnerait n'importe quelle compétition de calcul mental