**Un exemple d'IA** : AlphaGo Zero

Nous allons commencer par le début en parlant d’une autre IA qui est la version antérieur d'AlphaGo Zero : AlphaGo

AlphaGo c'est l'intelligence artificielle développé par la société britannique DeepMind pour affronter les meilleurs joueurs humains de go, elle est très complexe.

Il faut savoir aussi que go est un jeu qui est plus complexe que les échecs, c’était le seul jeu de ce type pour lequel les humains se montraient encore supérieurs aux machines.

Selon Demis Hassabis , l'un des fondateurs de DeepMind, ce jeu de réflexion offre un nombre de combinaisons « supérieur au nombre d'atomes que compte l'univers ».

Mais tout a changé en janvier 2016, lorsque DeepMind a annoncé qu'AlphaGo avait battu Fan Hui, triple champion européen en titre du jeu de go. Quelques mois plus tard, l'IA défiait Lee Sedol, l'un des meilleurs joueurs professionnels du monde. AlphaGo l'a battu par quatre victoires à une.

En mai 2017, une nouvelle version du programme s'est attaquée au champion du monde en titre, Ke Jie, qu'il a vaincu en trois manches.

AlphaGo utilise une méthode de calcul probabiliste dite méthode Monte-Carlo à laquelle sont associés deux réseaux neuronaux d'apprentissage profond ayant chacun une tâche différente :

* Un premier appelé policy network (réseau d'objectifs) travaille à prédire le prochain coup ;
* Un second appelé value network (réseau de valeur) estime l'issue favorable d'un mouvement à partir de la configuration du plateau de jeu.

Pour que l'intelligence artificielle développe ses propres stratégies de jeu, DeepMind lui a fait jouer des milliers de parties contre elle-même.

AlphaGo était considéré ainsi le meilleur IA au monde au jeu de go et pourtant elle a été battue par une nouvelle version c'est à dire AlphaGo Zero par 100 à 0.

AlphaGo Zero est une version améliorée d’AlphaGo produite aussi par l'entreprise DeepMind . Il s'agit d'une version développée sans l'aide de données provenant de parties jouées entre humains, ni de connaissances autres que les règles du jeu. Cette version est plus forte que n'importe quelle version précédente. En jouant contre lui-même, AlphaGo Zero a dépassé la force de la version d’AlphaGo du match contre Lee Se-dol en trois jours en gagnant par 100 jeux à 0, a atteint le niveau d’AlphaGo Master en 21 jours, et a dépassé toutes les anciennes versions en 40 jours.

Le réseau neural d'AlphaGo Zero a été formé en utilisant TensorFlow, en faisant travailler 64 processeurs graphiques (GPU) et 19 processeurs (CPU) paramétrés comme serveurs. Seulement quatre processeurs dédiés à l'apprentissage automatique (TPU, tensor processing unit) ont été utilisés pour les inférences. Initialement, le réseau de neurones ne savait rien d'autre du go que les règles du jeu. Contrairement aux versions antérieures d’AlphaGo, AlphaGo Zero, plutôt que de recevoir quelques situations programmées pour reconnaître quelques positions inhabituelles, n'appréhende que le plateau de jeu. L'IA est engagé dans l'apprentissage par renforcement, en jouant contre elle-même jusqu'à ce qu'elle puisse anticiper ses propres mouvements et l'incidence de ceux-ci sur le résultat du jeu. Dans les trois premiers jours, AlphaGo Zero a joué 4,9 millions de jeux contre lui-même en succession rapide. Il est apparu qu'il a développé les compétences requises pour battre les meilleurs joueurs humains en quelques jours, alors que les versions précédentes d'AlphaGo ont nécessité des mois de formation pour atteindre le même niveau.

Pour résumer, AlphaGoZero joue contre lui-même, en lançant des simulations de jeu qui se basent sur l’algorithme de Monte-Carlo pour prendre des décisions et évaluer chaque état. Grâce à toute ces simulations, le réseau de neurones est ensuite entraîné à prendre les meilleures initiatives et à mieux évaluer chaque partie. Pour créer une meilleure représentation du jeu, et entraîner un seul réseau au lieu de deux, le Policy Network, et le Value network sont à présent combinés en un seul et même réseau. Enfin, AlphaGoZero est basé sur un réseau Résiduel qui permet un meilleur entrainement des poids du réseau.

Pour mieux comprendre comment marche l’algorithme AlphaGo : 👇

<https://www.youtube.com/watch?v=P6eBraOUBLw>