

Apprentissage par renforcement profond

Nathan Barloy
Telecom Nancy

Universite de Lorraine

Email : nathan.barloy@telecomnancy.eu

Aurélien Delage
Telecom Nancy

Universite de Lorraine

Email: aurelien.delage@telecomnancy.eu

Olivier Buffet
and Vincent Thomas

INRIA

Equipe : Larsen

E-mail : olivier.buffet@loria.fr
and vincent.thomas@loria.fr

Résumé—The abstract goes here.

I. INTRODUCTION

Les débuts de l'intelligence artificielle - ou du moins la perception de l'éventualité de créer, un jour, des machines intelligentes qui surpasseraient les humains - peuvent être attribués aux travaux d'Alan Turing. Il proposa en 1936 une théorie formelle de ce qui sera plus tard un ordinateur. Une Machine de Turing est un concept abstrait, qui consiste en un ruban sur lequel se déplace une tête d'écriture, permettant de réaliser des opérations simples (arithmétiques, par exemple) [2].

14 ans plus tard, Alan Turing présente le concept de Test de Turing. Il concerne directement l'intelligence artificielle en proposant un test qui permettrait de mesurer la qualité de l'intelligence de la machine. L'expérience consiste à faire discuter un humain et une machine, l'humain ne sachant pas s'il discute avec un autre humain ou une machine, et simplement à lui demander à la fin de la discussion, s'il pense avoir parlé avec un humain ou avec une machine. A ce jour, aucune intelligence artificielle n'a *réellement* réussi à passer le Test de Turing [3] [6].

Cela commence à nous amener à notre sujet. Depuis ces définitions formelles d'Alan Turing, la qualité des intelligences artificielle n'a eu de cesse de croître.

En effet, on peut citer la célèbre victoire de l'algorithme *Deep Blue* contre M.Garry Kasparov [4], champion du monde d'échec, en 1996. On peut aussi citer la victoire d'*AlphaGo* battant en 2017 le champion du monde du jeu de Go Lee Sedol. Sa version adaptée au jeu d'échec est aussi particulièrement impressionnante tant dans l'originalité du jeu proposé que dans l'incapacité des maîtres d'échec de comprendre réellement, et d'anticiper les coups de l'algorithme [5].

Et voilà notre sujet: il s'avère que ces deux algorithmes ont été développés par Googletm en utilisant des techniques de *Reinforcement Learning*, ou autrement dit : *Apprentissage par Renforcement Profond*.

Dans ce PIDR, nous nous proposons d'appliquer les techniques de base d'Apprentissage par Renforcement Profond à des jeux de type Atari.

A. Machine Learning, Réseaux de neurones

Le Machine Learning (ou Apprentissage Automatique) consiste, entre autre., à faire apprendre à un ordinateur à réagir, à résoudre un problème, en réaction à un certain jeu de données en entrée. Une des technique principale du Machine Learning est l'utilisation de réseaux de neurones, et c'est ce qui est utilisé pour l'Apprentissage par Renforcement Profond. Nous allons donc nous concentrer dessus pour cette partie.

1) *Réseaux de Neurones*: Qu'est-ce qu'un réseau de neurone?

Commençons par dire - de manière très schématique - que c'est un système, composés de couches et de noeuds, qui permet de répondre quelque chose (une solution) à un jeu de donnée en entrée (un problème). Disons aussi qu'ils réalisent des performances particulièrement impressionnantes sur certains problèmes.

Disons aussi qu'outre les performances sur les jeux, concernant la voiture autonome, ou tout autre application des réseaux de neurones, il est à noter que les algorithmes utilisant des réseaux de neurones tendent à percevoir la *sémantique des mots et des idées*. Par exemple, l'équation roi-homme+femme est résolue par l'algorithme en donnant la solution: reine. Les mots ne sont plus des chaînes de caractères, ils sont des objets sémantiques en espaces de très grande dimension [7].

Nous recommandons très vivement au lecteur de regarder la référence citée ci-dessous, ainsi que la série de vidéos de vulgarisation qui l'accompagne, présentant de manière formelle la perception de la sémantique des mots en tant qu'objets d'espaces de grandes dimension, que l'on cherche à appliquer en espaces de dimension raisonnable, permettant l'analyse algorithmique,..

II. CONCLUSION

The conclusion goes here.

ACKNOWLEDGMENT

The authors would like to thank...

RÉFÉRENCES

- [1] H. Kopka and P. W. Daly, *A Guide to L^AT_EX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.
- [2] <http://zanotti.univ-tln.fr/turing/>
- [3] <https://culture.univ-lille1.fr/fileadmin/lna/lna66/lna66p04.pdf>
- [4] https://fr.wikipedia.org/wiki/Matchs_Deep_Blue_contre_Kasparov
- [5] <https://www.youtube.com/watch?v=bo5plUo86BU>
- [6] <https://www.csee.umbc.edu/courses/471/papers/turing.pdf>
- [7] <https://www.youtube.com/watch?v=r4EtRfHSU7E>