



Plusieurs sortes d'IA ?

Frédéric Alexandre

► To cite this version:

Frédéric Alexandre. Plusieurs sortes d'IA ?. Sébastien Konieczny; Henri Prade. L'intelligence artificielle: De quoi s'agit-il vraiment ?, Cepadues Editions, 2020, 9782364938502. hal-02891487

HAL Id: hal-02891487

<https://inria.hal.science/hal-02891487>

Submitted on 7 Jul 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Plusieurs sortes d'IA ?

Frederic Alexandre

Inria Bordeaux Sud-Ouest; Labri UMR 5800; IMN UMR 5293

146 rue Leo Saignat; 33076 Bordeaux, France

Frederic.Alexandre@inria.fr

Il n'aura pas échappé au lecteur attentif de ce petit livre que l'intelligence artificielle est diverse dans ses préoccupations et ses méthodes, même si tous ses travaux relèvent, d'une manière ou d'une autre, de la conception de systèmes offrant des fonctionnalités en rapport avec la réalisation de tâches réputées complexes. Depuis les débuts de l'intelligence artificielle il y a eu des chercheurs qui ont rêvé de construire des cerveaux artificiels dotés de capacités cognitives riches, et des chercheurs qui se sont attachés à développer des systèmes artificiels capables de réaliser de manière très performante des fonctions particulières, comme la déduction automatique, ou la reconnaissance de formes par apprentissage, pour prendre deux exemples.

Pour mécaniser de telles fonctions, l'intelligence artificielle peut faire appel à des modélisations * très différentes, fondées sur la logique* (on parle d'« IA symbolique »), ou sur des modèles quantitatifs, inspirés en particulier par les statistiques et le traitement du signal (on parle alors d'« IA numérique »). Une autre formulation plus actuelle se référant à des notions proches de cette dichotomie est de parler d'IA basée sur les connaissances * versus d'IA basée sur les données *. S'il est clair que les connaissances* sur le monde sont issues ultimement d'observations, donc de données*, on peut aussi avoir besoin de manipuler les connaissances* de façon explicite (ce qui peut se faire avec des outils logiques* ou numériques comme des réseaux bayésiens * par exemple). Mais il existe aussi des situations où il est intéressant d'utiliser des connaissances* et des données* conjointement.

Même si dès l'origine l'ambition de créer une "intelligence artificielle" en modélisant le cerveau a été affichée, force est de constater que ce Graal n'est pas encore atteint et que des voix s'élèvent régulièrement pour suggérer que comprendre le cerveau de façon unifiée n'est pas à notre portée et qu'il vaut mieux voir l'intelligence artificielle comme un ensemble de techniques pour résoudre différents problèmes. Ceci se réfère à la distinction, quelquefois faite, entre « IA faible » (réalisation de tâches spécifiques, telles que décrites dans ce document) et « IA forte » (avec un modèle complet de cerveau, une machine pourrait rassembler l'ensemble des compétences intelligentes d'un humain), aussi appelée « Intelligence Artificielle Générale » .

Ces questions impactent à l'évidence différents domaines de la philosophie, comme la théorie de l'esprit, l'étude de la conscience et, bien sûr, l'éthique. Débattre de la pertinence d'une forme d'intelligence artificielle plutôt qu'une autre est légitime, tant que l'on reste dans le cadre de démarches scientifiques falsifiables et que l'on ne se cantonne pas au simple énoncé de croyances. Ces débats ne doivent en aucun cas obérer la poursuite de différents programmes scientifiques, tant sur la modélisation du cerveau dans le cadre de l'intelligence artificielle que sur la conception et la réalisation de systèmes dédiés à des tâches spécifiques.