

Introduction à l'Intelligence Artificielle

Notes de cours 1

Nadjib Lazaar
Université Paris-Saclay
lazaar@lisn.fr

1 Qu'est-ce que l'Intelligence Artificielle ?

L'Intelligence Artificielle (IA) est le domaine scientifique et technique visant à simuler les fonctions cognitives humaines. Une IA doit créer l'impression d'une intelligence humaine pour un observateur.

Définition

« L'IA est la science de faire faire aux machines des choses qui nécessiteraient de l'intelligence si elles étaient faites par des humains » - Marvin Minsky.

2 L'Intelligence Humaine et son Lien avec l'IA

L'intelligence humaine sert de modèle pour l'IA, mais certaines de ses caractéristiques restent difficiles à reproduire. Voici quelques éléments clés :

- **Perception et Sensation** : L'intelligence humaine repose sur la perception sensorielle, qui permet de comprendre l'environnement. L'IA utilise des capteurs (ex. : vision par ordinateur) mais reste limitée par la complexité humaine.
- **Cognition et Compréhension** : L'humain apprend, comprend et s'adapte aux situations nouvelles de manière flexible. L'IA tente de reproduire cela via l'apprentissage machine, mais elle reste limitée par les données et les algorithmes.
- **Mémoire et Représentation** : La mémoire humaine est flexible et hiérarchique, permettant une gestion efficace des connaissances. L'IA utilise des bases de données et des réseaux de neurones mais manque de la profondeur et de l'intégration de la mémoire humaine.
- **Émotions et Motivation** : Les émotions humaines influencent la prise de décision et l'interaction. L'IA simule des réponses émotionnelles, mais n'a pas la capacité de ressentir ou comprendre les émotions.

- **Raisonnement Abstrait et Créativité** : L'intelligence humaine excelle dans le raisonnement abstrait et la créativité. L'IA, bien qu'elle puisse simuler certains raisonnements, est limitée par ses algorithmes et ses données.
- **Apprentissage Autonome** : L'humain apprend en permanence de manière informelle et adaptative, contrairement à l'IA qui dépend de données préalablement définies.

2.1 Curiosité : L'Essence de l'Intelligence Humaine

La curiosité est un moteur fondamental de l'intelligence humaine. Elle pousse les individus à explorer, à expérimenter et à chercher des réponses à des questions nouvelles, souvent sans but immédiat. C'est cette capacité à poser des questions et à vouloir comprendre l'inconnu qui permet à l'être humain d'innover, d'évoluer et de s'adapter constamment.

Bien que l'IA puisse être programmée pour résoudre des problèmes ou optimiser des processus en explorant des solutions au sein de paramètres définis, elle n'a pas cette capacité à s'interroger d'elle-même ni le désir intrinsèque d'explorer l'inconnu. L'IA fonctionne selon des algorithmes et des données qui lui sont fournies, et ne développe pas de curiosité autonome.

C'est cette différence fondamentale qui fait que l'intelligence artificielle restera toujours "artificielle" : elle manque de la motivation intrinsèque à explorer le monde, à poser des questions nouvelles et à rechercher des réponses au-delà de ses algorithmes. Ainsi, la curiosité humaine, ce besoin perpétuel de savoir et de comprendre, constitue une caractéristique que l'IA ne pourra jamais imiter, et c'est ce qui marque la frontière entre l'intelligence humaine et l'intelligence artificielle.

3 Origines et Histoire de l'IA

- **1950** : Alan Turing publie "Computing Machinery and Intelligence", introduisant le test de Turing pour évaluer une machine capable d'imiter l'intelligence humaine.
- **1956** : Lors de la conférence de Dartmouth, John McCarthy propose le terme "intelligence artificielle" pour désigner les machines capables de réaliser des tâches associées à l'intelligence humaine.
- **1959** : Arthur Samuel développe un programme d'apprentissage machine capable de jouer aux dames.
- **1973** : Le rapport Lighthill critique les progrès de l'IA, entraînant une réduction importante du financement, marquant le début du premier hiver de l'IA.
- **1996** : Deep Blue, une IA basée sur une approche symbolique, bat Garry Kasparov, champion du monde d'échecs.
 - Deep Blue utilise des règles expertes codées manuellement et des algorithmes d'exploration exhaustive (brute force).

- L'évaluation des positions repose sur des heuristiques prédéfinies et une énorme base de données d'ouvertures et de fins de parties.
- Aucune capacité d'apprentissage n'était présente dans ce système.
- **2011** : Watson, développé par IBM, remporte le jeu télévisé Jeopardy!, démontrant les capacités des systèmes d'IA avancés pour le traitement du langage naturel (youtube).
- **2014** : Google rachète DeepMind, une entreprise spécialisée dans l'apprentissage profond (deep learning).
- **2015** : AlphaGo, conçu par DeepMind, utilise des réseaux neuronaux profonds et l'apprentissage par renforcement pour battre pour la première fois un champion humain, Fan Hui, au jeu de Go.
 - Aspect connexionniste : AlphaGo repose sur des architectures de réseaux neuronaux profonds combinées à des techniques d'apprentissage supervisé et par renforcement. Ces réseaux permettent à la machine d'analyser des positions complexes sur le plateau de Go, une tâche auparavant considérée comme inaccessible pour les ordinateurs.
 - Cette victoire est une démonstration majeure des capacités des modèles connexionnistes, qui s'appuient sur des représentations numériques apprises automatiquement à partir de vastes ensembles de données.
- **2018** : OpenAI introduit GPT-2, un modèle avancé de traitement du langage naturel basé sur des réseaux neuronaux.
- **2022** : GPT-3, la version suivante, est mis gratuitement à disposition en ligne, illustrant les avancées rapides de l'IA générative.

4 Les Deux Écoles de l'IA

L'intelligence artificielle (IA) se divise en deux grandes écoles de pensée qui abordent différemment la manière de modéliser et de simuler l'intelligence humaine : l'IA symbolique et l'IA connexionniste. Chacune de ces approches repose sur des fondements théoriques distincts et met l'accent sur des aspects différents de l'intelligence.

IA Symbolique

L'IA symbolique, également appelée IA "classique", repose sur la représentation explicite des connaissances sous forme de symboles et de règles. Cette approche considère que la résolution de problèmes peut être modélisée par la manipulation de symboles représentant des objets, des concepts ou des relations dans le monde réel. Les systèmes basés sur cette approche utilisent des structures telles que les logiques formelles et les bases de connaissances.

Par exemple, les *systèmes experts* sont un type d'IA symbolique qui imite les compétences humaines dans des domaines spécialisés en utilisant des règles

et des raisonnements déductifs. Cette approche a été largement utilisée dans les premières décennies de l'IA et s'est avérée efficace dans des environnements bien structurés et clairement définis. Toutefois, elle rencontre des difficultés dans des situations complexes ou mal définies où les connaissances doivent être apprises à partir de données.

IA Connectionniste

L'IA connectionniste, en revanche, s'inspire du fonctionnement du cerveau humain et des réseaux neuronaux biologiques. Elle se concentre sur la simulation de l'apprentissage automatique à travers des réseaux de neurones artificiels. Contrairement à l'IA symbolique, qui repose sur des règles explicites et des connaissances préalablement codées, l'IA connectionniste met l'accent sur l'expérimentation, l'apprentissage et l'adaptation des systèmes à partir de données.

Les réseaux neuronaux, en particulier les *réseaux de neurones profonds* (ou deep learning), ont permis des avancées spectaculaires dans des domaines tels que la reconnaissance d'images, la traduction automatique et la reconnaissance vocale. Cette approche se distingue par sa capacité à traiter de grandes quantités de données non structurées, à découvrir des patterns cachés dans les données, et à s'adapter en fonction des informations qu'elle reçoit. Cependant, bien que les réseaux neuronaux aient montré des résultats impressionnants, l'interprétabilité de ces modèles reste un défi majeur.

4.1 IA Hybride et Neurosymbolique

L'IA hybride représente une convergence des approches symbolique et connectionniste. Alors que l'IA symbolique excelle dans des environnements bien définis et structurés, et que l'IA connectionniste est particulièrement performante dans l'apprentissage à partir de grandes quantités de données non structurées, l'IA hybride cherche à combiner le meilleur de ces deux mondes.

L'un des domaines les plus prometteurs dans cette direction est l'IA **neurosymbolique**, qui vise à intégrer les réseaux neuronaux avec des techniques symboliques pour surmonter leurs limitations respectives. L'objectif de l'IA neurosymbolique est de tirer parti des capacités d'apprentissage profond des réseaux neuronaux tout en conservant la structure explicite, l'interprétabilité et la manipulation de connaissances propre à l'IA symbolique.

Un exemple d'application de l'IA neurosymbolique pourrait être la résolution de problèmes complexes où les connaissances doivent être combinées avec des capacités d'apprentissage, comme dans la compréhension du langage naturel. Par exemple, un système neurosymbolique pourrait utiliser un réseau neuronal pour apprendre des relations complexes dans des données textuelles, puis utiliser un raisonnement symbolique pour effectuer des inférences logiques à partir de

ces données.

L'intégration des approches symbolique et connectionniste dans un cadre hybride présente encore plusieurs défis, notamment la gestion de la complexité des systèmes, l'interopérabilité entre les différents paradigmes, et la recherche de méthodes pour garantir l'efficacité et l'interprétabilité des modèles.

5 Principaux Domaines d'Application de l'IA

L'IA est un domaine multidimensionnel qui regroupe divers sous-domaines répondant à des défis spécifiques, souvent interconnectés :

- **Représentation des Connaissances et Raisonnement Automatique (KRR)** : Ce domaine traite de la structuration des connaissances à l'aide de graphes et d'ontologies pour permettre un raisonnement logique. Il est essentiel dans des applications complexes comme la planification, la configuration de systèmes, et la conception de solutions. Les compétences acquises dans nos masters permettent de modéliser et de résoudre ces problèmes efficacement.
- **Apprentissage Automatique (ML)** : L'apprentissage automatique est au cœur de l'IA, permettant aux machines de prédire, classifier ou découvrir des patterns à partir des données. Ce domaine est omniprésent dans des applications telles que la détection de fraude, les systèmes de recommandation, et la prédiction dans la bio-informatique. Les étudiants de nos masters acquièrent une expertise pratique grâce à des projets liés à ces domaines.
- **Traitement du Langage Naturel (NLP)** : Le NLP permet aux machines de comprendre et générer du langage humain, avec des applications dans les assistants vocaux, les chatbots, et la traduction automatique. Nos étudiants développent des compétences avancées pour concevoir des systèmes NLP adaptés à divers contextes.
- **Vision par Ordinateur (CV)** : La vision par ordinateur permet aux machines d'analyser et d'interpréter des données visuelles. Ce domaine est crucial pour des applications telles que la conduite autonome, l'analyse d'images médicales, et le contrôle qualité dans l'industrie. Les masters IA et SD offrent une formation complète pour relever ces défis grâce à des algorithmes d'apprentissage profond.
- **Éthique de l'IA et IA Explicable (XAI)** : L'éthique de l'IA englobe des enjeux majeurs tels que la transparence, la responsabilité et les impacts sociaux des technologies intelligentes, notamment dans des secteurs sensibles comme la santé et la justice. Parallèlement, l'IA explicable (XAI) vise à rendre les modèles d'IA compréhensibles et interprétables pour les utilisateurs, renforçant ainsi leur confiance et favorisant une adoption responsable. Dans nos parcours de master IA et SD, les étudiants sont formés à aborder ces problématiques complexes à travers des cours dédiés et des projets appliqués, les préparant à concevoir des systèmes à la fois performants, transparents et éthiques.

- **Systèmes Multi-agents (MAS)** : Les systèmes multi-agents gèrent des interactions complexes entre agents autonomes. Ces systèmes sont utilisés pour optimiser les réseaux logistiques, simuler des comportements humains et coordonner des véhicules autonomes.
- **Robots et Systèmes Autonomes (RA)** : Les robots intelligents et autonomes accomplissent des tâches complexes en s'adaptant à leur environnement, avec des applications en logistique, en santé (robotique chirurgicale) et dans l'automobile (robots industriels, conduite autonome).

Nos parcours de Master IA et SD

Nos deux parcours de master, **Intelligence Artificielle (IA)** et **Science des Données (SD)** à la faculté de sciences d'Orsay, offrent une expertise approfondie dans les domaines KRR, ML, NLP, CV et XAI, préparant les étudiants à relever ces défis technologiques.

5.1 Applications Interdisciplinaires

L'IA se déploie souvent en synergie avec d'autres technologies. Par exemple, elle transforme :

- **La bio-informatique** avec des algorithmes pour la découverte de médicaments ou la médecine personnalisée.
- **L'automobile** grâce aux véhicules autonomes et aux systèmes d'assistance à la conduite.
- **La santé** avec l'analyse prédictive, l'imagerie médicale, et les systèmes d'aide au diagnostic.

Messages Clés à Retenir

- **Tout n'est pas de l'IA** : Tous les algorithmes ou programmes informatiques ne relèvent pas de l'intelligence artificielle. L'IA se distingue par sa capacité à apprendre, raisonner et s'adapter à des situations variées, souvent grâce à des données.
- **L'IA n'est pas magique** : Derrière les performances impressionnantes de l'IA se trouvent des principes mathématiques et des algorithmes rigoureusement établis. Son efficacité dépend largement de la qualité des données et des objectifs définis.
- **L'IA stimule la curiosité humaine plutôt que de la surpasser** : Si l'IA est capable d'exceller dans des tâches spécifiques, elle reste un outil conçu pour nous assister. En nous délestant de certaines tâches complexes ou répétitives, elle nous libère du temps et des ressources pour explorer de nouvelles idées, poser des questions inédites, et repousser les frontières de notre compréhension. C'est un catalyseur qui nourrit notre quête de connaissance et stimule notre créativité.

6 Ressources Supplémentaires

- **Histoire de l'IA** : https://en.wikipedia.org/wiki/History_of_artificial_intelligence
- **Livre : *Machine Learning* de Tom Mitchell** : Un ouvrage clé pour les bases de l'IA, couvrant des techniques classiques de machine learning et des exemples pratiques.
- **Livre : *Artificial Intelligence : A Modern Approach* de Stuart Russell et Peter Norvig** : Un ouvrage de référence pour les étudiants et chercheurs, abordant l'IA de manière complète et approfondie, avec un focus sur la résolution de problèmes complexes.
- **Livre : *Principles of Artificial Intelligence* de Nils Nilsson** : Un classique qui explore les fondements théoriques de l'IA et les algorithmes de base.
- **Livre : *The Society of Mind* de Marvin Minsky** : Ce livre fondamental aborde les théories cognitives et l'intelligence artificielle sous un angle psychologique et philosophique.
- **Livre : *Deep Learning* de Ian Goodfellow, Yoshua Bengio, et Aaron Courville** : Une référence essentielle pour comprendre l'apprentissage profond, son fonctionnement et ses applications.
- **Livret du GDR-RADIA** : Un document qui offre une vue d'ensemble des recherches actuelles en IA, incluant des travaux et des projets collaboratifs en France.