Introduction to Artificial Intelligence

Course Notes 1

Nadjib Lazaar
Université Paris-Saclay
lazaar@lisn.fr

---

1 What is Artificial Intelligence?

Artificial Intelligence (AI) is the scientific and technical field aimed at simulating human cognitive functions. An AI should create the impression of human-like intelligence for an observer.

Definition

"AI is the science of making machines do things that would require intelligence if done by humans" - Marvin Minsky.

2 Human Intelligence and its Connection with AI

Human intelligence serves as a model for AI, but some of its characteristics remain difficult to replicate. Here are some key elements:

- **Perception and Sensation:** Human intelligence relies on sensory perception, which helps understand the environment. AI uses sensors (e.g., computer vision) but is limited by human complexity.
- **Cognition and Understanding:** Humans learn, understand, and adapt to new situations flexibly. AI attempts to reproduce this through machine learning but remains limited by data and algorithms.
- **Memory and Representation:** Human memory is flexible and hierarchical, enabling efficient management of knowledge. AI uses databases and neural networks but lacks the depth and integration of human memory.
- **Emotions and Motivation:** Human emotions influence decision-making and interaction. AI simulates emotional responses but cannot feel or understand emotions.

- **Abstract Reasoning and Creativity:** Human intelligence excels in abstract reasoning and creativity. AI, although it can simulate some reasoning, is limited by its algorithms and data.
- **Autonomous Learning:** Humans continuously learn informally and adaptively, unlike AI, which depends on predefined data.

2.1 Curiosity: The Essence of Human Intelligence

Curiosity is a fundamental driver of human intelligence. It pushes individuals to explore, experiment, and seek answers to new questions, often without an immediate goal. It is this ability to ask questions and desire to understand the unknown that allows humans to innovate, evolve, and adapt constantly.

Although AI can be programmed to solve problems or optimize processes by exploring solutions within defined parameters, it lacks the ability to question itself or the intrinsic desire to explore the unknown. AI works based on algorithms and data it is provided, and does not develop autonomous curiosity.

This fundamental difference makes artificial intelligence always "artificial": it lacks the intrinsic motivation to explore the world, ask new questions, and seek answers beyond its algorithms. Thus, human curiosity, the perpetual need to know and understand, is a characteristic that AI can never imitate, marking the boundary between human and artificial intelligence.

3 Origins and History of AI

- **1950:** Alan Turing publishes "Computing Machinery and Intelligence", introducing the Turing Test to evaluate a machine capable of imitating human intelligence.
- **1956:** At the Dartmouth conference, John McCarthy introduces the term "artificial intelligence" to describe machines capable of performing tasks associated with human intelligence.
- **1959:** Arthur Samuel develops a machine learning program capable of playing checkers.
- **1973:** The Lighthill report criticizes AI progress, leading to a significant reduction in funding, marking the beginning of the first AI winter.
- **1996:** Deep Blue, an AI based on a symbolic approach, defeats Garry Kasparov, the world chess champion.
  - Deep Blue uses manually coded expert rules and exhaustive search algorithms (brute force).
  - Position evaluation relies on predefined heuristics and a vast database of opening and endgame positions.
  - No learning capability was present in this system.

- **2011:** Watson, developed by IBM, wins the TV game show Jeopardy!, demonstrating the capabilities of advanced AI systems for natural language processing (youtube).
- **2014:** Google acquires DeepMind, a company specializing in deep learning.
- **2015:** AlphaGo, created by DeepMind, uses deep neural networks and reinforcement learning to defeat a human champion, Fan Hui, at the game of Go.
  - Connectionist aspect: AlphaGo relies on deep neural network architectures combined with supervised and reinforcement learning techniques. These networks allow the machine to analyze complex Go board positions, a task previously considered inaccessible to computers.
  - This victory is a major demonstration of the capabilities of connectionist models, which rely on automatically learned digital representations from vast data sets.
- **2018:** OpenAI introduces GPT-2, an advanced natural language processing model based on neural networks.
- **2022:** GPT-3, the next version, is made freely available online, illustrating the rapid advances in generative AI.

4 The Two Schools of AI

Artificial Intelligence (AI) is divided into two major schools of thought that approach modeling and simulating human intelligence differently: symbolic AI and connectionist AI. Each of these approaches is based on distinct theoretical foundations and focuses on different aspects of intelligence.

Symbolic AI

Symbolic AI, also known as classical AI, is based on the explicit representation of knowledge in the form of symbols and rules. This approach assumes that problem-solving can be modeled by manipulating symbols representing objects, concepts, or relationships in the real world. Systems based on this approach use structures such as formal logics and knowledge bases.

For example, *expert systems* are a type of symbolic AI that imitate human skills in specialized fields using rules and deductive reasoning. This approach was widely used in the early decades of AI and proved effective in well-structured and clearly defined environments. However, it faces difficulties in complex or poorly defined situations where knowledge must be learned from data.

Connectionist AI

Connectionist AI, on the other hand, is inspired by the functioning of the human brain and biological neural networks. It focuses on simulating machine learning through artificial neural networks. Unlike symbolic AI, which relies on explicit rules and pre-coded knowledge, connectionist AI emphasizes experimentation, learning, and adaptation of systems from data.

Neural networks, particularly *deep neural networks* (or deep learning), have enabled spectacular advances in areas such as image recognition, machine translation, and speech recognition. This approach stands out for its ability to process large amounts of unstructured data, discover hidden patterns in data, and adapt based on the information it receives. However, despite the impressive results of neural networks, the interpretability of these models remains a major challenge.

4.1 Hybrid and Neurosymbolic AI

Hybrid AI represents a convergence of the symbolic and connectionist approaches. While symbolic AI excels in well-defined and structured environments, and connectionist AI is particularly strong in learning from large amounts of unstructured data, hybrid AI seeks to combine the best of both worlds.

One of the most promising areas in this direction is **neurosymbolic AI**, which aims to integrate neural networks with symbolic techniques to overcome their respective limitations. The goal of neurosymbolic AI is to leverage the deep learning capabilities of neural networks while maintaining the explicit structure, interpretability, and knowledge manipulation inherent in symbolic AI.

An example of a neurosymbolic AI application could be solving complex problems where knowledge needs to be combined with learning capabilities, such as in natural language understanding. For example, a neurosymbolic system could use a neural network to learn complex relationships in textual data and then use symbolic reasoning to make logical inferences from this data.

The integration of symbolic and connectionist approaches in a hybrid framework still presents several challenges, including managing system complexity, ensuring interoperability between paradigms, and finding methods to guarantee the efficiency and interpretability of the models.

5 Main Application Areas of AI

AI is a multidimensional field that encompasses various subfields addressing specific challenges, often interconnected:

- **Machine Learning (ML)**: Machine learning is at the heart of AI, allowing machines to predict, classify, or discover patterns from data. This field is pervasive in applications such as fraud detection, recommendation systems, and predictions in bioinformatics. Students in our master's programs gain practical expertise through projects related to these areas.

- **Knowledge Representation and Automated Reasoning (KRR)**: This field focuses on structuring knowledge using graphs and ontologies to enable logical reasoning. It is crucial in complex applications such as planning, system configuration, and solution design. The skills gained in our master's programs allow students to model and solve these problems effectively.

- **Natural Language Processing (NLP)**: NLP enables machines to understand and generate human language, with applications in virtual assistants, chatbots, and machine translation. Our students develop advanced skills to design NLP systems tailored to various contexts.

- **Computer Vision (CV)**: Computer vision allows machines to analyze and interpret visual data. This field is essential for applications such as autonomous driving, medical image analysis, and quality control in industry. Our AI and SD master's programs provide comprehensive training to tackle these challenges using deep learning algorithms.

- **AI Ethics and Explainable AI (XAI)**: AI ethics encompasses major issues such as transparency, accountability, and the social impact of intelligent technologies, particularly in sensitive areas like healthcare and justice. At the same time, explainable AI (XAI) aims to make AI models understandable and interpretable for users, building trust and encouraging responsible adoption. In our AI and SD master's programs, students are trained to address these complex issues through dedicated courses and applied projects, preparing them to design systems that are both efficient, transparent, and ethical.

- **Multi-agent Systems (MAS)**: Multi-agent systems manage complex interactions between autonomous agents. These systems are used to optimize logistics networks, simulate human behaviors, and coordinate autonomous vehicles.

- **Robots and Autonomous Systems (RA)**: Intelligent and autonomous robots perform complex tasks by adapting to their environment, with applications in logistics, healthcare (surgical robotics), and the automotive industry (industrial robots, autonomous driving).

Our Master's Programs in AI and SD

Our two master's programs, **Artificial Intelligence (AI)** and **Data Science (SD)** at the University of Orsay, provide in-depth expertise in ML, KRR, NLP, CV, and XAI, preparing students to tackle these technological challenges.

5.1 Interdisciplinary Applications

AI is often deployed in synergy with other technologies. For example, it is transforming:

- **Bioinformatics** with algorithms for drug discovery or personalized medicine.
- **Automotive** through autonomous vehicles and driver assistance systems.
- **Healthcare** with predictive analytics, medical imaging, and diagnostic support systems.

Key Takeaways

- **Not everything is AI**: Not all algorithms or computer programs fall under the category of artificial intelligence. AI is distinguished by its ability to learn, reason, and adapt to diverse situations, often through the use of data.
- **AI is not magic**: Behind AI's impressive performance lie well-established mathematical principles and algorithms. Its effectiveness largely depends on the quality of the data and the defined objectives.
- **AI stimulates human curiosity rather than surpassing it**: While AI excels in specific tasks, it remains a tool designed to assist us. By relieving us of complex or repetitive tasks, it frees up time and resources to explore new ideas, ask novel questions, and push the boundaries of our understanding. It is a catalyst that fuels our quest for knowledge and stimulates our creativity.

6 Additional Resources

- **History of AI**: `https://en.wikipedia.org/wiki/History_of_artificial_intelligence`
- **Book:** ***Machine Learning*** by **Tom Mitchell**: A key resource for the fundamentals of AI, covering classical machine learning techniques and practical examples.
- **Book:** ***Artificial Intelligence: A Modern Approach*** by **Stuart Russell and Peter Norvig**: A reference book for students and researchers, providing a comprehensive and in-depth look at AI, with a focus on solving complex problems.
- **Book:** ***Principles of Artificial Intelligence*** by **Nils Nilsson**: A classic exploring the theoretical foundations of AI and core algorithms.

- **Book:** ***The Society of Mind*** by **Marvin Minsky**: This seminal work discusses cognitive theories and artificial intelligence from a psychological and philosophical perspective.
- **Book:** ***Deep Learning*** by **Ian Goodfellow, Yoshua Bengio, and Aaron Courville**: An essential reference for understanding deep learning, how it works, and its applications.
- **GDR-RADIA Booklet**: A document offering an overview of current AI research, including collaborative works and projects in France.