Artificial Intelligence Notes

Marco Natali

INDICE

ELENCO DELLE FIGURE

1 INTRODUCTION

This course will provide an introduction to AI techniques and approach analyzed nowadays and to understand the current state of art we have to provide an Timeline to see progress and discover done during the time, so in figure 1 we will see all important events related with AI.

The major discover happened on 2020 are the following:

GPT3 (GENERATIVE PRE-TRAINED TRANSFORMER): produced by OpenAI in May 2020, is a larger and richer language model consisting in 175 billion machine learning parameters used for automatic text generation, translation, user interface synthesis

DARPA CHALLENGE (ALPHADOGFIGHTS) with simulated F-16 Air Fighters where on $18-20$ August 2020 there was the final Event, where AI system was against each other and the winner was a system by Heron system, that was also able to defeated a human expert top gun fighter $5-0$.

On [**?**] Andrew NG says that AI will transform many industries, but it's not magic and almost all of AI's recent progress is based on one type of AI, in which some input data (A) is used to quickly generate some simple response (B) $[A \to B]$.
Also Andrew Ng says that if a typical person can do a mental task with less than one second of thought, we can probably automate it using AI either now or in the near future.
Choosing A and B creatively has already revolutionized many industries, it is poised to revolutionize many more.

ML systems are not (yet?) able to justify in human terms their results, so for some application it is essential the human knowledge to be able to generate explanations, infact some regulations requires the right to an explanation in decision-making, and seek to prevent discrimination based on race, opinions, health, sex and so on, like GPDR.
ML systems learn what's in the data, without understanding what's true or false, real or imaginary, fair or unfair and so it is possible to develop bad/unfair models.

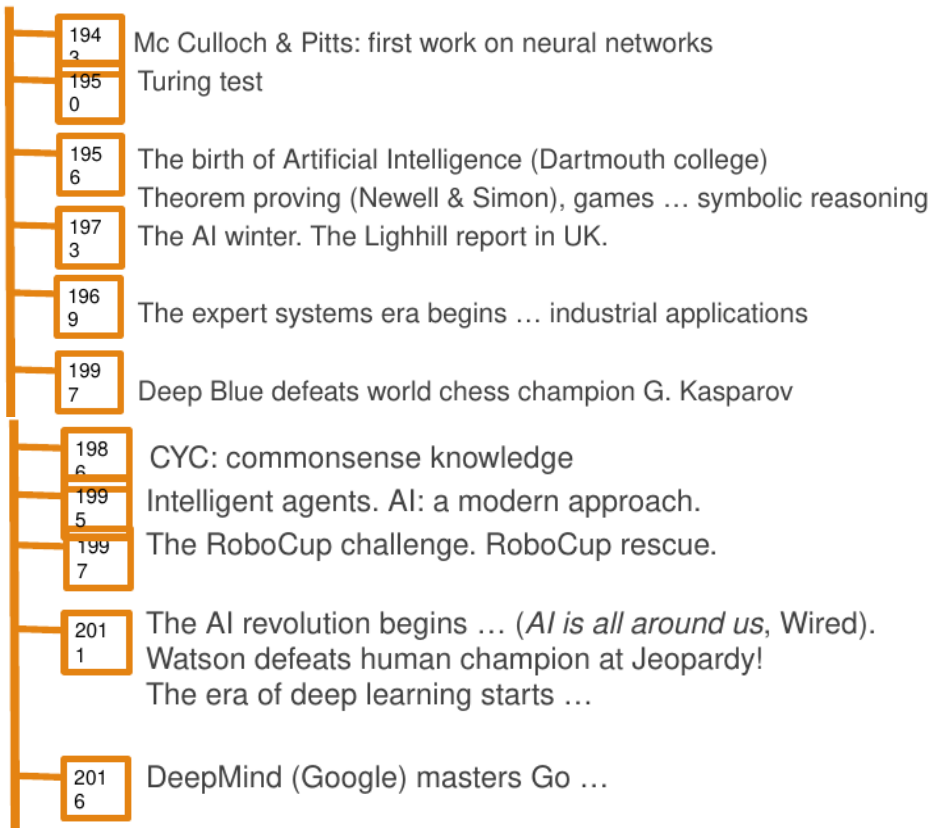
The goal of building AI systems is far from being solved and is still quite challenging in its own. Building complex AI systems requires the combination of several techniques and approaches, not only ML.
One of the most challenging tasks ahead of us is integration of perception and reasoning in AI systems.

AI fundamentals is mostly about "Slow thinking" or "Reasoning" and AI fundamentals has the role, within the AI curriculum, of teaching you about the foundations of a discipline which is now 60 year old.
We will cover different approaches, also some coming of the "Good Old- Fashioned Artificial Intelligence" (GOFAI) or symbolic AI.

Def. Symbolic AI is an high-level "symbolic" (human-readable) representations of problems, the general paradigm of searching for a solution, knowledge representation and reasoning, planning.
Symbolic AI was the dominant paradigm of AI research from the mid 1950s until the late 1980s and central to the building of AI systems is the *Physical symbol systems hypothesis*, formulated by Newell and Simon.

The approach is based on the assumption that many aspects of intelligence can be achieved by the manipulation of symbols (the physical symbol system hypothesis):

Figura 1: AI Timeline evolution



Def. A physical symbol system has the necessary and sufficient means for general intelligent action

Human thinking is a kind of symbol manipulation system (a symbol system is necessary for intelligence) and machines can be intelligent (a symbol system is sufficient for intelligence).
The hypothesis cannot be proven, we can only collect empirical evidence and observations and experiments on human behavior in tasks requiring intelligence.

We have two different typologies of AI, that was introduced and considered:

STRONG AI: relies on the strong assumption that human intelligence can be reproduced in all its aspects (general A.I.).
It includes adaptivity, learning, consciousness and not only pre-programmed behavior.

WEAK AI: simulation of human-like behavior, without effective thinking/understanding and no claim that it works like human mind; it is the dominant approach today.

A problem of AI is that computer can't have needs, cravings or desires and Abraham Maslow's define a hierarchy of human needs:

1. Biological needs (food, sleep, sex, ...)
2. Safety, protection from environment
3. Love and belonging, friendship
4. Self esteem and respect from others
5. Self-actualization