Introduction

Table of Contents

[Brain Replication 3](#_Toc121296071)

[History 4](#_Toc121296072)

[Modern Applications of AI 5](#_Toc121296073)

[Rational Agents 6](#_Toc121296074)

There are several schools of thought regarding how to define **artificial intelligence**.

1. The first school of thought suggested that it was the science of making machines that **think rationally**. This essentially means using propositions and rules to reach conclusions, referring to traditional programming. This school of thought was quickly disproven because it became evident as systems grew larger that it was simply impossible to define a set of rules that captures all possible scenarios.
2. The second school of thought suggested that it was the science of making machines that **think like people**. This was problematic, since it is difficult to accurately define what that even means. We do not know the process we follow to think in the way that we do.
3. The third school suggested that the machines should **act like people**. This was better, since we did not necessarily need to understand how the machine managed to do this. This refers to the domain of machine learning. The goal was to pass the **Turing test**, a test which basically meant that a human was unable to differentiate between an answer provided by a computer and one provided by another human. Although promising, even this school of thought was wrong because it is possible to fool a human. Just because a machine can tell you its ‘favorite’ novel does not mean it has the same level of intelligence as a human.
4. We are currently on the fourth school of thought, which suggests that the machines must be able to **act rationally**, making rational decisions. We only care about what decisions are made, not how the machine came to those decisions. A rational decision is one which maximally achieves a pre-defined goal. The goal depends on the utility of the outcome. For example, suppose we have to decide whether to go to the barber to get a haircut or to cut our hair ourselves. If our goal is to look good, the machine should decide that the barber is the right choice. If our goal is to save money, the machine should decide that cutting it ourselves is the right choice. In AI, our goal is to **maximize the expected utility**. We are using the word ‘expected’, because most decisions have some uncertainty surrounding them (e.g. there is a possibility of rain, which affects the decision to go to the barber).

## Brain Replication

Since the goal of AI is to behave like the human brain, why not just study the human brain and replicate it? There are multiple issues with this:

* The brain is amazing at rational thinking, but it is still **flawed**. Not every human being makes rational decisions all the time.
* The human brain is not a piece of software that can easily be analyzed. It is absurdly difficult to **reverse engineer**. We basically have no clue how it works step by step to achieve its goals.
* Even if we do manage to replicate a human brain, it will not necessarily mean that we have achieved artificial intelligence. The steps required to produce artificial intelligence might be different from the steps required to produce a human brain without us even realizing there is a difference.

Although we have gotten nowhere close to replicating the human brain, we have learnt a few lessons:

* Any AI will require some **memory**. The human brain remembers things and finetunes its decisions based on it. We learnt that fire damages us at some point in our lives and we stay away from fire because of that memory.
* Any AI must be capable of **simulating** possible outcomes of decisions it does not have a direct memory of. Even if a human has never seen lava before, just by standing near it and feeling the heat coming from it, they will be able to simulate the outcome of touching it and will (hopefully) stay away from doing so.

## History

* 1940-1950: Early days
  + 1943: McCulloch & Pitts: Boolean circuit model of brain
  + 1950: Turing’s "Computing Machinery and Intelligence"
* 1950-70: Excitement
  + 1950s: Early AI programs, including Samuel’s checkers program, Newell & Simon’s Logic Theorist, Gelernter’s Geometry Engine
  + 1956: Dartmouth meeting: "Artificial Intelligence" adopted
  + 1965: Robinson’s complete algorithm for logical reasoning
* 1970-90: Knowledge-based approaches
  + 1969-79: Early development of knowledge-based systems
  + 1980-88: Expert systems industry booms
  + 1988-93: Expert systems industry busts: "AI Winter"
* 1990-2012: Statistical approaches + subfield expertise
  + Resurgence of probability, focus on uncertainty
  + General increase in technical depth
  + Agents and learning systems... "AI Spring"?
* 2012-\_\_\_\_: Excitement?
  + Big data, big compute, neural networks
  + Some re-unification of sub-fields
  + AI used in many industries

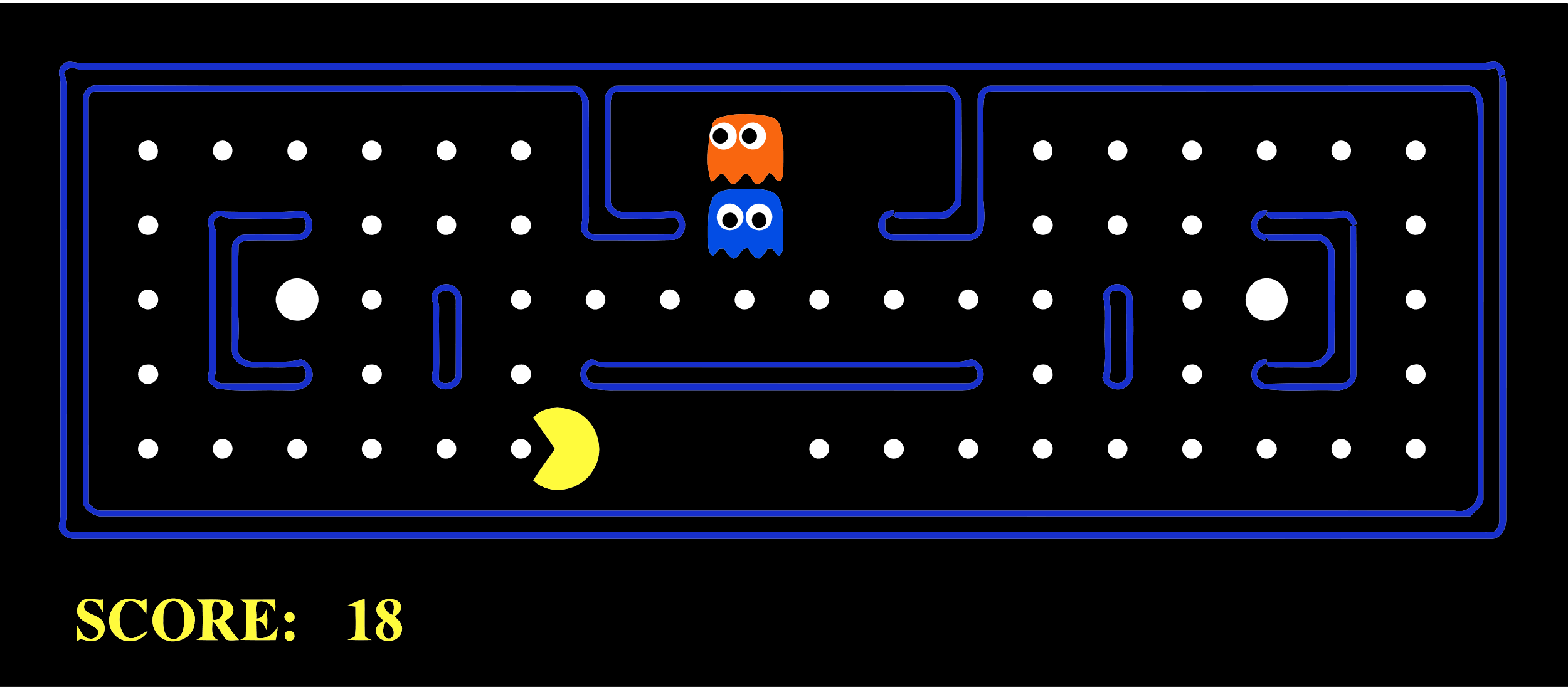
## Modern Applications of AI

* Generated stories
  + Not very good at it yet
* Speech Technologies
  + Automatic Speech Recognition
  + Text-to-Speech Synthesis
  + Dialog Systems
* Language Processing Technologies
  + Question Answering
  + Machine Translation
  + Web Search
  + Text Classification
  + Spam Filtering
* Computer Vision
  + Image Captioning
  + Semantic Scene Segmentation
  + 3D Understanding
  + Deep Fake
* Robotics
  + Rescue
  + Sports
  + Vehicles
  + Automation
* Game Play

## Rational Agents

Our objective in this course will be to design **rational agents**. An agent is anything that can accept **perceptions** and take **actions** based on those perceptions. A rational agent chooses an action from a set of possible actions with the goal of maximizing utility. The techniques for doing this involve using the **perceptions** about the surrounding environment to create an **action space**, from which the optimal one is selected.

For example, consider an automated game of Pac-Man.



The perceived environment contains dots, ghost positions and walls. The available action space consists of moving left, right, up or down.

In the real world, the environment is perceived by the agent using **sensors**. For example, lidar and cameras are used by autonomous vehicles. The actions are executed via **actuators**, which would be the wheels for an autonomous vehicle.