Introduction

Michael Rose

1.1 | What is AI?

We can split definitions of AI into a set of sorts containing the following subsets:

- Thinking Humanly
- Acting Humanly
- Thinking Rationally
- Acting Rationally

Generally, the humanly definitions are empirical and revolve around observations and hypothesis of human behavior, and the rationality definitions are approached with a combination of mathematics and engineering.

1.1.1 | Acting Humanly: The Turing Test Approach

For a program to pass the **Turing Test**, it would need the following:

- natural language processing: to speak and understand english
- knowledge representation: to store what it knows or hears
- automated reasoning: to use stored information to answer questions or draw new conclusions
- machine learning: to adapt to new circumstances and to detect and extrapolate patterns

There is also the **Total Turing Test** which, in addition, requires:

- computer vision: to see
- robots: to manipulate physical objects

These 6 disciplines compose most of AI.

1.1.2 | Thinking Humanly: The Cognitive Modeling Approach

In order to say that a program thinks like a human, we must understand the workings of human minds. If we were to understand that, it would be reasonable to be able to define a general theory and extrapolate it to computers.

Some ways we try to understand the mind now include:

- introspection: observing ones own thoughts
- psychological experiments: observing others
- brain imaging: observing the brain in action

1.1.3 | Thinking Rationally: The Laws of Thought Approach

There is a **logicist** tradition within AI that hopes to build computers that can handle formal systems and find solutions and relations. This is quite difficult though, as the world contains uncertainty, and there are limitations to the ability of a computer to explore a given problems solution space due to complexity.

1.1.4 | Acting Rationally: The Rational Agent Approach

An **agent** is something that acts. A **rational agent** is one that acts as to achieve the best outcome, or, when there is uncertainty, the best expected outcome. In the laws of thought approach, the emphasis is on correct inferences. This is only part of being a rational agent, as a rational agent may require heuristics to properly manage in its environment. An example would be seeing a predator of sorts. A correct inference may take time to calculate, but a rational inference would be to run.

All the skills needed for the Turing test would allow an agent to act rationally. We will see before long that achieving perfect rationality -always doing the right thing- is not achievable in complicated environments.

1.2 | The Foundations of Artificial Intelligence

1.2.1 | Philosophy

- Can formal rules be used to draw valid conclusions?
- How does the mind arise from a physical brain?
- Where does knowledge come from?
- How does knowledge lead to action?

1.2.2 | Mathematics

- What are the formal rules to draw valid conclusions?
- What can be computed?
- How do we reason with uncertain information?

The first nontrivial algorithm is thought to be Euclid's algorithm for computing greatest common divisors

```
# Euclid's Algorithm
def gcd(a, b):
  if (b == 0):
    return a
  else:
    return gcd(b, a%b)
# results
print(gcd(9999, 888))
```

```
## 3
```

1.2.3 | Economics

- How should we make decisions so as to maximize payoff?
- How should we do this when others may not go along?
- How should we do this when the payoff may be far in the future?

Decision theory (combining probability with utility theory) provides a formal framework for decisions made under uncertainty. This is suitable for cases where probabilistic descriptions appropriately capture the decisiom maker's environment. This is suitable for 'large' economies where each agent need pay no attention to the actions of other agents as individuals.

For small economies, the situation is more like a game (theory), and the actions of one player can significantly affect the utility of another.

For the third question, we must understand how to make rational decisions when payoffs result from actions taken in sequence. This is the field of operations research.

1.2.4 | Neuroscience

- How do brains process information?

1.2.5 | Psychology

- How do humans and animals think and act?

1.2.6 | Computer Engineering

- How can we build an efficient computer?

1.2.7 | Control Theory and Cybernetics

- How can artifacts operate under their own control?

1.2.8 | Linguistics

- How does language relate to thought?