

Course: Introduction to AI

Prof. Ahmed Guessoum

The National Higher School of AI

Chapter 1

Introduction

Outline

- Intelligence
- Definitions of Artificial Intelligence
- The Foundations of AI
- A Brief History of AI
- The State of the Art

What is Intelligence?

Is it the ability to...

- Think?
- Solve problems?
- Sense the environment?
- Plan?
- Take decisions?
- Act?
- Be rational?
- Learn?
- Communicate (in Natural Language)?
- Have feelings; show emotions; show compassion?
- Reason about morality?

The Realm of AI

- AI encompasses a huge variety of subfields, ranging from
 - ♦ the general (learning and perception) to
 - ♦ the specific, such as
 - playing chess, proving mathematical theorems, writing poetry, driving a car on a crowded street, and diagnosing diseases.
- AI is relevant to any intellectual task;
- Truly a universal field.
- Is AI part of CS or CS part of AI? ;-)

Outline

- Intelligence
- Definitions of Artificial Intelligence
- The Foundations of AI
- A Brief History of AI
- The State of the Art

Various definitions of AI

Thinking Humanly

"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense."
(Haugeland, 1985)

"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)

Thinking Rationally

"The study of mental faculties through the use of computational models."
(Charniak and McDermott, 1985)

"The study of the computations that make it possible to perceive, reason, and act."
(Winston, 1992)

Acting Humanly

"The art of creating machines that perform functions that require intelligence when performed by people."
(Kurzweil, 1990)

"The study of how to make computers do things at which, at the moment, people are better."
(Rich and Knight, 1991)

Acting Rationally

"Computational Intelligence is the study of the design of intelligent agents."
(Poole *et al.*, 1998)

"AI . . . is concerned with intelligent behavior in artifacts."
(Nilsson, 1998)

Acting humanly: The Turing Test approach

Turing (1950), "Computing Machinery and Intelligence":

- ♦ "Can Machines think?" → "Can machines behave intelligently?"
- ♦ Operational test for intelligent behavior: the Imitation Game

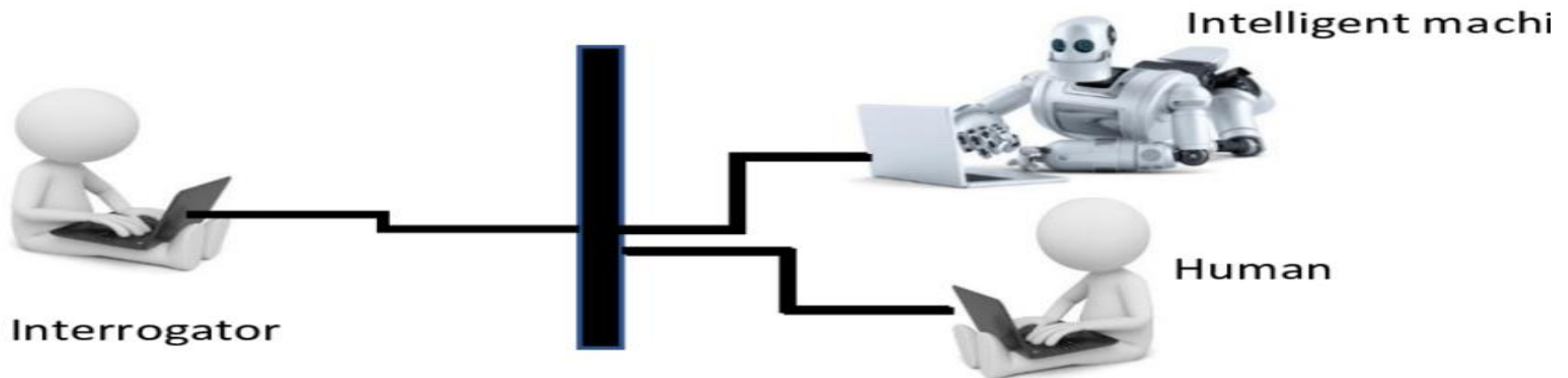


Figure: Pietikäinen, Matti & Silvén, Olli. (2022). Challenges of Artificial Intelligence -- From Machine Learning and Computer Vision to Emotional Intelligence.

The Turing Test

Turing

- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes.
- Anticipated all major arguments against AI in the following 50 years.
- Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problem: the Turing Test is not reproducible, constructive, or amenable to mathematical analysis.

What computers would need to pass the Turing Test

- **Natural language processing** to enable it to communicate successfully in English;
- **Knowledge representation** to store what it knows or hears;
- **Automated reasoning** to use the stored information to answer questions and to draw new conclusions;
- **Machine learning** to adapt to new circumstances and to detect and extrapolate patterns.

The Total Turing Test

- In addition to the previous, it includes
 - ♦ a video signal so that the interrogator can test the subject's perceptual abilities,
 - ♦ the opportunity for the interrogator to pass physical objects through the opening.
- So, additionally, a computer would need:
- **Computer vision** to perceive objects, &
- **Robotics** to manipulate objects and move around.
- NOTE: The 6 capabilities cover most of AI!

Section 2

Thinking humanly:

The cognitive modeling approach

- Trying to determine how humans think; to get *inside* the actual workings of human minds.
- Three ways to do this:
 - ♦ through introspection—trying to catch our own thoughts as they go by;
 - ♦ through psychological experiments;
 - ♦ through brain imaging.
- Then develop a sufficiently precise theory of the mind, & express it as a computer program
- Compare the program's I/O behaviour to that of a human.
- **Cognitive science** brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind.

Thinking rationally: The “laws of thought” approach

- Aristotle: one of the first to attempt to codify “right thinking,” i.e. irrefutable reasoning processes.
- His **syllogisms** provided patterns for argument structures that always yielded correct conclusions for correct premises
- These laws of thought initiated the field called **logic**
- 19th century: logicians developed a precise notation for statements about all kinds of objects in the world and the relations among them
- The so-called **logicist** tradition within AI hopes to build on such programs to create intelligent systems

Thinking rationally: The “laws of thought” approach

Two main obstacles to this approach:

1. Not easy to take informal knowledge and state it in the formal terms required by logical notation, particularly when the knowledge is less than 100% certain.
2. Complexity: even problems with just a few hundred facts can exhaust the computational resources of any computer unless guided as to which reasoning steps to try first.

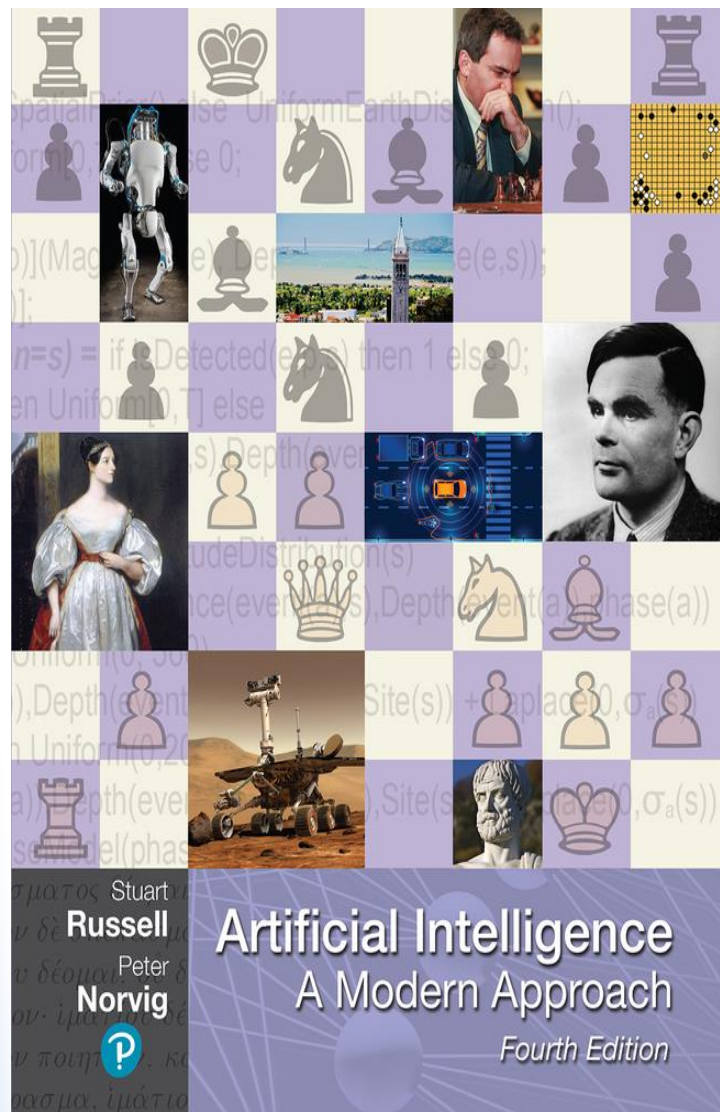
Acting rationally: The rational agent approach

- An **agent** is an entity that acts.
- **Computer agents** are expected to do more: operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change, and create and pursue goals.
- A **rational agent** is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome.
- The best outcome: that which is expected to maximize goal achievement, given the available information.

Acting rationally

- Correct inference is not *all* of rationality; in some situations, there is no provably correct thing to do, but something must still be done.
- There are ways of acting rationally that cannot be said to involve inference: e.g. reflexes; blinking the eye, etc.
- All the skills needed for the Turing Test also allow an agent to act rationally.
- Two advantages of rational-agent approach over the other approaches.
 1. It is more general than the “laws of thought
 2. It is more amenable to scientific development than approaches based on human behavior or thought. The standard of rationality is mathematically well defined and can lead to agent designs that provably achieve it.

Slides based on the textbook



- Russel, S. and Norvig, P. (2020) Artificial Intelligence, A Modern Approach (4th Edition), Pearson Education Limited.