Artificial Intelligence

What is AI?

Ch1.1

- *"We call ourselves Homo sapiens—man the wise—because our **intelligence** is so important to us.*
- *For thousands of years, we have tried to understand how we think; that is, how a mere **handful of matter** can perceive, understand, predict, and manipulate a world far larger and more complicated than itself.*
- *The field of artificial intelligence, or AI, goes further still: it attempts not just to **understand** but also to **build intelligent entities**."*

What Is AI?

- What is intelligence?
  - The short answer is that there is no consensus.
  - Legg and Hutter [1] summarized and organized over 70 different published definitions of "intelligence", most oriented toward general intelligence, emanating from researchers in a variety of disciplines.

What Is AI?

- **Narrow AI** [2, 3]: systems that carry out specific "intelligent" behaviors in specific contexts.
  - Ex.: classification of breeds of dogs from pictures, playing chess, detecting traffic signs from images, diagnosing diseases from medical records, recognizing songs, grouping consumers, etc.
- For a narrow AI system, if one changes the context or the behavior specification even a little bit, some level of human reprogramming or reconfiguration is generally necessary to enable the system to retain its level of intelligence.

What Is AI?

- **Narrow AI** [2, 3]: systems that carry out specific "intelligent" behaviors in specific contexts.
  - Ex.: classification of breeds of dogs from pictures, playing chess, detecting traffic signs from images, diagnosing diseases from medical records, recognizing songs, grouping consumers, etc.

For a narrow AI system, if one changes the context or the behavior specification even a little bit, some level of human reprogramming or reconfiguration is generally necessary to enable the system to retain its level of intelligence.

https://theinformr.com/cell-plans/p/how-much-data-do-i-need-290/

What Is AI?

- **Artificial General Intelligence (AGI)** [3]:
  - General intelligence involves the ability to achieve a variety of goals, and carry out a variety of tasks, in a variety of different contexts and environments.
  - A generally intelligent system should be able to handle problems and situations quite different from those anticipated by its creators.
  - A generally intelligent system should be good at generalizing the knowledge it's gained, so as to transfer this knowledge from one problem or context to others.

Most of the current research on AI is focused on Narrow AI.
This is mostly because no one knows how to get closer to AGI.

What Is AI?

- **Artificial General Intelligence** [3]: Note the capacity of reacting creatively to unexpected situations different from day-by-day life:
  - Attempt to reduce the dog eating speed fails:
    - https://www.youtube.com/watch?v=9r1xqw4NP9Y
  - Crow solving puzzles to receive food:
    - https://www.youtube.com/watch?v=cbSu2PXOTOc
  - Octopus scaping maze:
    - https://www.youtube.com/watch?v=303S9kOxd0c
  - Drivers react fast to avoid accident:
    - https://www.youtube.com/watch?v=B5LxAlaHV8c

What Is AI?

- Weak AI hypotheses and Strong AI hypotheses
  - **Weak AI hypothesis**: machines can act as if they were intelligent
  - **Strong AI hypothesis**: machines that do so are actually thinking (not just simulating thinking)
  - Most AI researchers take the weak AI hypothesis for granted, and don't care about the strong AI hypothesis

OBS: Several people awkwardly use the term strong AI to refer to general AI. Do not confuse that with the strong AI hypothesis.

What Is AI? The answer of Russell and Norvig [4]

thought processes and reasoning

behavior

Thinking Humanly	Thinking Rationally
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985)
"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998)
"The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Figure 1.1 Some definitions of artificial intelligence, organized into four categories.

success measured by fidelity to human performance

success defined by an ideal performance measure, called **rationality**

Acting Humanly: The Turing Test Approach

- The Turing Test, proposed by Alan Turing (1950), was designed to provide a satisfactory operational definition of intelligence.
- A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer.

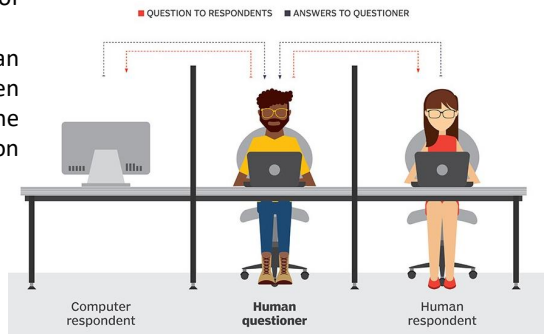


Figure extracted from: https://medium.com/thinkmobiles/evaluating-artificial-intelligence-from-turing-test-to-now-b64a8fced070

Acting Humanly: The Turing Test Approach

The computer would need to possess the following capabilities:

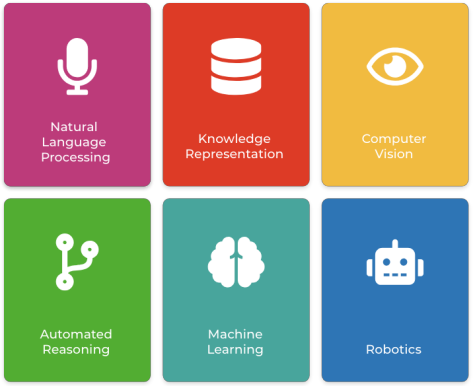
- **natural language processing** to enable it to communicate successfully;
- **knowledge representation** to store what it knows or hears;
- **automated reasoning** to use the stored information to answer questions and to draw new conclusions;
- **machine learning** to adapt to new circumstances and to detect and extrapolate patterns.

Acting Humanly: The Turing Test Approach

The **Total Turing Test** includes a video signal so that the interrogator can test the subject's perceptual abilities, as well as the opportunity for the interrogator to pass physical objects "through the hatch.". To pass the total Turing Test, the computer will need:

- **computer vision** to perceive objects, and
- **robotics** to manipulate objects and move about.

These Six Disciplines Compose most of AI



- Turing deserves credit for designing a test that remains relevant 60 years later. Yet AI researchers have devoted little effort to passing the Turing Test.
- A reason is the belief that it is more important to study the underlying principles of intelligence than to duplicate an exemplar.
- The quest for "artificial flight" succeeded when the Wright brothers and others stopped imitating birds and started using wind tunnels and learning about aerodynamics. Aeronautical engineering texts do not define the goal of their field as making "machines that fly so exactly like pigeons that they can fool even other pigeons".

Thinking Humanly: The Cognitive Modelling Approach

- If we are going to say that a given program thinks like a human, we must have some way of determining how humans think.
  - **Thought Experiments:** observing our thoughts as they fly by;
  - **Psychological experiments:** observing a person in action; and
  - **Brain imaging:** observing the brain in action
- The interdisciplinary field of **cognitive science** brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind. Once we have a sufficiently precise theory of the mind, it becomes possible to express the theory as a computer program.

Thinking Rationally: The "Laws of Thought" Approach

- The Greek philosopher Aristotle was one of the first to attempt to codify "**right thinking**," that is, irrefutable reasoning processes.
- These **laws of thought** were supposed to govern the operation of the mind; their study initiated the field called **logic**.
- There are two main obstacles to this approach:
  - It is not easy to **take informal knowledge and state it in the formal terms** required by logical notation (particularly in face of uncertainty).
  - There is a big difference between solving a problem "in principle" and solving it in practice. Even problems with just a few hundred facts can exhaust the computational resources of any computer unless it has some guidance as to which reasoning steps to try first.

Acting Rationally: The Rational Agent Approach

This will be the focus of our course!

- An **agent** is just something that **acts**.
- All computer programs do something, but 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**.

Acting Rationally: The Rational Agent Approach

- In the "laws of thought" approach to AI, the emphasis was on **correct inferences**. Making correct inferences is sometimes part of being a rational agent, because one way to act rationally is to reason logically to the conclusion that a given action will achieve one's goals and then to act on that conclusion.
- On the other hand, **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 also ways of acting rationally that cannot be said to involve inference.
  - Ex.: recoiling from a hot stove is a reflex action that is usually more successful than a slower action taken after careful deliberation.

Acting Rationally: The Rational Agent Approach

- The skills needed for the Turing Test also allow an agent to act rationally.
  - Knowledge representation and reasoning enable agents to reach good decisions.
  - We need to be able to generate comprehensible sentences in natural language to get by in a complex society.
  - We need learning not only for erudition, but also because it improves our ability to generate effective behavior.

Acting Rationally: The Rational Agent Approach

- The rational-agent approach has two advantages over the other approaches.
  - First, **it is more general than the "laws of thought"** approach because correct inference is just one of several possible mechanisms for achieving rationality.
  - Second, **it is more amenable to scientific development** than are approaches based on human behavior or human thought.
    - The standard of rationality is mathematically well defined and completely general, and can be "unpacked" to generate agent designs that provably achieve it.
    - Human behavior, on the other hand, is well adapted for one specific environment and is defined by, well, the sum total of all the things that humans do.

References

[1] Legg, S., and Hutter, M. 2007a. A collection of definitions of intelligence. Frontiers in Artificial Intelligence and Applications 157:17.

[2] Goertzel, B. (2014). Artificial general intelligence: concept, state of the art, and future prospects. Journal of Artificial General Intelligence, 5(1), 1-48.

[3] Kurzweil, R. 2005. The singularity is near: When humans transcend biology. Penguin.

[4] Russell, S., & Norvig, P. (2002). Artificial intelligence: a modern approach.