Artificial Intelligence: Introduction

Slides from Russell & Norvig book, revised by Andrea Roli
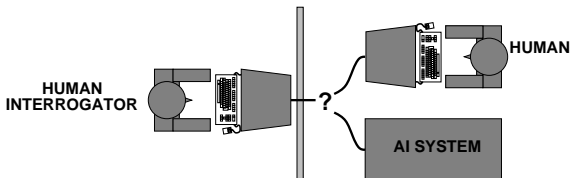
Outline

◇ What is AI?
◇ A brief history
◇ Philosophical foundations
◇ The state of the art

Acting humanly: The Turing test

Turing (1950) "Computing machinery and intelligence":

◇ "Can machines think?" ⟶ "Can machines behave intelligently?"

◇ Operational test for intelligent behavior: the Imitation Game

◇ 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 following 50 years

◇ Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problem: Turing test is not **reproducible**, **constructive**, or amenable to **mathematical analysis**

Thinking humanly: Cognitive Science

- 1960s "cognitive revolution": information-processing psychology replaced prevailing orthodoxy of behaviorism

- Requires scientific theories of internal activities of the brain:

    - What level of abstraction? "Knowledge" or "circuits"?
    - How to validate? Requires (1) Predicting and testing behavior of human subjects (top-down) or (2) Direct identification from neurological data (bottom-up)

- Both approaches (roughly, Cognitive Science and Cognitive Neuroscience) are now distinct from AI

- Both share with AI the following characteristic: **the available theories do not explain (or engender) anything resembling human-level general intelligence**

AI prehistory

Philosophy	logic, methods of reasoning
	mind as physical system
	foundations of learning, language, rationality
Mathematics	formal representation and proof
	algorithms,computation,(un)decidability,(in)tractability
	probability
Psychology	adaptation
	phenomena of perception and motor control
	experimental techniques (psychophysics, etc.)
Economics	formal theory of rational decisions
Linguistics	knowledge representation
	grammar
Neuroscience	plastic physical substrate for mental activity
Control theory	homeostatic systems, stability
	simple optimal agent designs

History of AI

1943	McCulloch & Pitts: Boolean circuit model of brain
1950	Turing's "Computing Machinery and Intelligence"
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
1966–74	AI discovers computational complexity Neural network research almost disappears
1969–79	Early development of knowledge-based systems
1980–88	Expert systems industry booms
1988–93	Expert systems industry busts: "AI Winter"
1985–95	Neural networks return to popularity
1988–	Resurgence of probability; general increase in technical depth "Nouvelle AI": ALife, GAs, soft computing
1995–	Agents, agents, everywhere . . .
2003–	Human-level AI back on the agenda

Philosophical foundations

Two main hypotheses:

- Weak AI: Machine can act *as if* they were intelligent
- Strong AI: Machines *can be* intelligent

Objections to weak AI

- The argument from disability ("A machine can never do X")
- The mathematical objection (computability theory)
- The argument from 'informality'/complexity of behavior

Objections to strong AI

- Consciousness
- Mind-body problem
- Free will

Mental experiments:

- The brain in a vat (The matrix?)
- The brain prosthesis
- The Chinese room

▷ Still ongoing debate. . .

State of the art

- Autonomous planning and scheduling
- Game playing
- Autonomous control
- Diagnosis
- Logistics planning
- Robotics
- Language understanding and problem solving