

Artificial Intelligence

For HEDSPI Project

Lecture 1 - Introduction

Lecturers :

- **Le Thanh Huong**
Dept of Information Systems
School of Information and Communication
Technology - HUST

1

Outline

- What is AI?
- Foundations of AI
- Short history of AI
- Philosophical discussions

2

What is AI?

Views of AI fall into four categories:

Think like humans	Thinking rationally
Act like humans	Acting rationally

The textbook advocates "acting rationally"

3

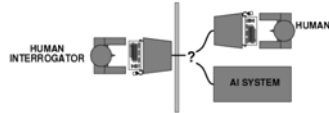
Think like humans

- 1960s "cognitive revolution": information-processing psychology
- Scientific theories of internal activities of the brain
 - What level of abstraction? "Knowledge" or "circuits"?
 - **Cognitive science**: Predicting and testing behavior of human subjects (top-down)
 - This approach now distinct from AI
 - share with AI the following characteristic:
 - The available theories do not explain anything resembling human-level general intelligence

4

Act like humans

- 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

5

Thinking rationally

- The "Laws of Thought" approach
 - What does it mean to "think rationally"?
 - Normative / prescriptive rather than descriptive
- Logician tradition:
 - Logic: notation and rules of derivation for thoughts
 - Aristotle: what are correct arguments/thought processes?
 - E.g.: Socrat is a human, human cannot live forever → Socrat human cannot live forever
 - Direct line through mathematics, philosophy, to modern AI
- Problems:
 - Not all intelligent behavior is mediated by logical deliberation
 - What is the purpose of thinking? What thoughts should I have?
 - Logical systems tend to do the wrong thing in the presence of uncertainty

6

Acting rationally

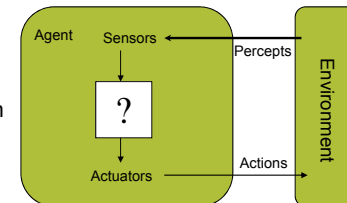
- Rational behavior: doing the "right thing"
 - The right thing: that which is expected to maximize goal achievement, given the available information
 - Doesn't necessarily involve thinking, e.g., blinking
 - Thinking can be in the service of rational action
 - Entirely dependent on goals!
 - Irrational ≠ insane, irrationality is sub-optimal action
 - Rational ≠ successful
- Our focus here: rational agents
 - Systems which make the best possible decisions given goals, evidences, and constraints
 - In the real world, usually lots of uncertainty... and lots of complexity
 - Usually, we're just approximating rationality
- "Computational rationality" a better title for this course

7

Rational agents

- An **agent** is an entity that perceives and acts
- An agent function maps from percept histories to actions:

$$P^* \rightarrow \mathcal{A}$$



- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Computational limitations make perfect rationality unachievable
- So we want the best program for given machine resources

8

Foundations of AI

- 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
- Economics utility, decision theory
- Neuroscience physical substrate for mental activity
- Psychology phenomena of perception and motor control, experimental techniques
- Computer engineering building fast computers
- Control theory design systems that maximize an objective function over time
- Linguistics knowledge representation, grammar

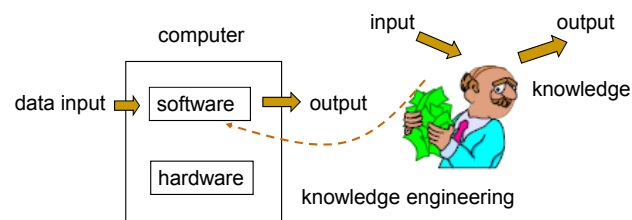
9

Short history of AI

- 1940-1950: Early days
 - 1943: McCulloch & Pitts: Boolean circuit model of brain
 - 1950: Turing's "Computing Machinery and Intelligence"
- 1950—70: Excitement: Look, Ma, no hands!
 - 1950s: Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
 - 1956: Dartmouth meeting: "Artificial Intelligence" adopted
 - 1964: ELIZA
 - 1965: Robinson's complete algorithm for logical reasoning
- 1970—88: 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"
- 1988—: Statistical approaches
 - Resurgence of probability, focus on uncertainty
 - General increase in technical depth
 - Agents, agents, everywhere... "AI Spring"?
- 2000—: Where are we now?

10

Expert system



Expert system = Human Expertise + Inference/Reasoning

Some examples: DENDRAL, MYCIN, PROSPECTOR, MOLGEN, ICAD/ICAM

11

State of the art

- May, '97: Deep Blue vs. Kasparov
 - First match won against world-champion
 - "Intelligent creative" play
 - 200 million board positions per second!
 - Humans understood 99.9 of Deep Blue's moves
 - Can do about the same now with a big PC cluster
- Proved a mathematical conjecture (Robbins conjecture) unsolved for decades
- No hands across America (driving autonomously 98% of the time from Pittsburgh to San Diego)
- During the 1991 Gulf War, US forces deployed an AI logistics planning and scheduling program that involved up to 50,000 vehicles, cargo, and people
- NASA's on-board autonomous planning program controlled the scheduling of operations for a spacecraft
- Proverb solves crossword puzzles better than most humans



12

Philosophical discussions

What Can AI Do?

- Play a decent game of table tennis?
- Drive safely along a curving mountain road?
- Buy a week's worth of groceries on the web?
- Discover and prove a new mathematical theorem?
- Converse successfully with another person for an hour?
- Perform a complex surgical operation?
- Unload a dishwasher and put everything away?
- Translate spoken English into spoken Vietnamese in real time?
- Write an intentionally funny story?

Can machine think?

13

Some problems with AI

- People might lose their jobs to automation.
- People might have too much (or too little) leisure time.
- People might lose their sense of being unique.
- People might lose some of their privacy rights.
- The use of AI systems might result in a loss of accountability.
- The success of AI might mean the end of the human race.

14