# Artificial Intelligence

## Chapter 1

# Outline

◇ What is AI?

◇ A brief history

◇ The state of the art
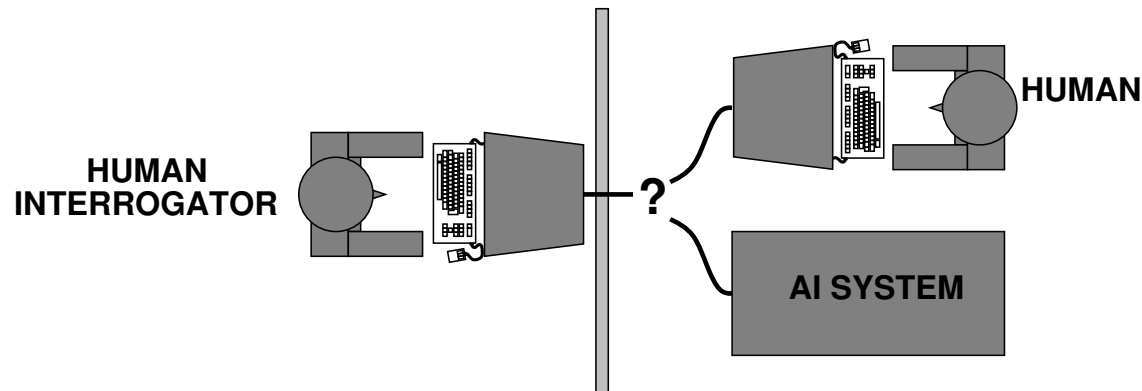
# What is AI?

| Systems that think like humans | Systems that think rationally |
|---|---|
| Systems that act like humans | Systems that act rationally |

# Acting humanly: The Turing test

Turing (1950) "Computing machinery and intelligence":

◇ "Can machines think?" $\longrightarrow$ "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 representation, automated reasoning, natural language understanding, machine 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**

Hence, all three fields share one principal direction!

# Thinking rationally: Laws of Thought

Normative (or prescriptive) rather than descriptive

Aristotle: what are correct arguments/thought processes?

Several Greek schools developed various forms of logic:
    **notation** and **rules of derivation** for thoughts;
may or may not have proceeded to the idea of mechanization

Direct line through mathematics and philosophy to modern AI

Problems:
1) Not all intelligent behavior is mediated by logical deliberation
2) What is the purpose of thinking? What thoughts **should** I have
    out of all the thoughts (logical or otherwise) that I **could** have?

# 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 reflex—but thinking should be in the service of rational action

What counts as the right thing is defined by the objective function we provide to the agent.

Stuart Russell calls this general paradigm the standard model.

# Beneficial Machines

For an artificial task like chess or path planning, the task comes with an obvious objective built in.

But for something like a self-driving car, the objective is difficult to specify and is always going to involve tradeoffs.

The problem of achieving agreement between our true preferences and the objective we put into a machine is the value alignment problem.

Stuart Russell in a recent book, **Human Compatible: AI and the Problem of Control** argues that we should build machines in a way that requires them to learn our objectives and never quite be sure exactly what they are.

In the end, we want agents that are provably beneficial to humans.

# Rational agents

An agent is an entity that perceives and acts

This course is about designing rational agents

Abstractly, an agent is a function from percept histories to actions:

$$f : \mathcal{P}^* \to \mathcal{A}$$

For any given class of environments and tasks, we seek the
agent (or class of agents) with the best performance

Caveat: **computational limitations make
perfect rationality unachievable**
$\to$ design best program for given machine resources

# 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 |

# Potted history of AI

| | |
|---|---|
| 1943 | McCulloch & Pitts: Boolean circuit model of brain |
| 1950 | Turing's "Computing Machinery and Intelligence" |
| 1952–69 | 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 |
| 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 |

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley

# State of the art

Which of the following can be done at present?

◇  Play a decent game of table tennis
◇  Drive safely along a curving mountain road
◇  Drive safely along El Camino Real in Silicon Valley
◇  Drive safely onto a container ship and head back to Japan

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology
◇ Name your next band

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology
◇ Name your next band
◇ Give competent legal advice in a specialized area of law

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology
◇ Name your next band
◇ Give competent legal advice in a specialized area of law
◇ Translate spoken English into spoken Swedish in real time

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology
◇ Name your next band
◇ Give competent legal advice in a specialized area of law
◇ Translate spoken English into spoken Swedish in real time
◇ Converse successfully with another person for an hour

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology
◇ Name your next band
◇ Give competent legal advice in a specialized area of law
◇ Translate spoken English into spoken Swedish in real time
◇ Converse successfully with another person for an hour
◇ Perform a complex surgical operation

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology
◇ Name your next band
◇ Give competent legal advice in a specialized area of law
◇ Translate spoken English into spoken Swedish in real time
◇ Converse successfully with another person for an hour
◇ Perform a complex surgical operation
◇ Unload any dishwasher and put everything away

# State of the art

Which of the following can be done at present?

◇ Play a decent game of table tennis
◇ Drive safely along a curving mountain road
◇ Drive safely along El Camino Real in Silicon Valley
◇ Drive safely onto a container ship and head back to Japan
◇ Buy a week's worth of groceries at HyVee
◇ Play a decent game of bridge, chess, even Go
◇ Discover and prove a new mathematical theorem
◇ Design and execute a research program in molecular biology
◇ Name your next band
◇ Give competent legal advice in a specialized area of law
◇ Translate spoken English into spoken Swedish in real time
◇ Converse successfully with another person for an hour
◇ Perform a complex surgical operation
◇ Unload any dishwasher and put everything away