Artificial Intelligence

Chapter 1
Outline

♦ What is AI?
♦ A brief history
♦ The state of the art
### What is AI?

The four approaches of AI:

<table>
<thead>
<tr>
<th>Systems that think like humans</th>
<th>Systems that think rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systems that act like humans</td>
<td>Systems that act rationally</td>
</tr>
</tbody>
</table>

Can you guess the trend today?
Turing (1950) “Computing machinery and intelligence”:
♦ “Can machines think?” → “Can machines behave intelligently?”
♦ Operational test for intelligent behavior: the Imitation Game

◇ Suggested major components of AI:
  Natural language processing, knowledge representation, automated reasoning, and 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) Introspection - trying to catch our thoughts as they go by
    2) Psychological experiments - observing a person in action
    3) Direct identification from neurological data (brain image like fMRI)

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
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?
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

Aristotle (Nicomachean Ethics):

Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good
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:

\[ f : \mathcal{P}^* \rightarrow A \]

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**

\[ \rightarrow \] design best **program** for given machine resources
<table>
<thead>
<tr>
<th>Field</th>
<th>Concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Philosophy</td>
<td>logic, methods of reasoning, mind as physical system, foundations of learning, language, rationality</td>
</tr>
<tr>
<td>Mathematics</td>
<td>formal representation and proof, algorithms, computation, (un)decidability, (in)tractability, probability</td>
</tr>
<tr>
<td>Psychology</td>
<td>adaptation, phenomena of perception and motor control, experimental techniques (psychophysics, etc.)</td>
</tr>
<tr>
<td>Economics</td>
<td>formal theory of rational decisions</td>
</tr>
<tr>
<td>Linguistics</td>
<td>knowledge representation, grammar</td>
</tr>
<tr>
<td>Neuroscience</td>
<td>plastic physical substrate for mental activity</td>
</tr>
<tr>
<td>Control theory</td>
<td>homeostatic systems, stability, simple optimal agent designs</td>
</tr>
</tbody>
</table>
## Potted history of AI

<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1943–55</td>
<td>The gestation of AI: first of neural network</td>
</tr>
<tr>
<td>1956</td>
<td>Dartmouth meeting: “Artificial Intelligence” adopted</td>
</tr>
<tr>
<td>1952-1969</td>
<td>Early enthusiasm, great expectations: ”Look, Ma, no Hands” era</td>
</tr>
<tr>
<td>1965</td>
<td>Robinson’s complete algorithm for logical reasoning</td>
</tr>
<tr>
<td>1966–74</td>
<td>The reality: AI discovers computational complexity</td>
</tr>
<tr>
<td></td>
<td>Neural network research almost disappears</td>
</tr>
<tr>
<td>1969–79</td>
<td>Early development of knowledge-based systems: weak methods</td>
</tr>
<tr>
<td>1980–</td>
<td>Expert systems in industry</td>
</tr>
<tr>
<td>1985–95</td>
<td>Neural networks return to popularity</td>
</tr>
<tr>
<td>1987–</td>
<td>AI adopts scientific methods: HMM, data mining, machine learning</td>
</tr>
<tr>
<td>1995–</td>
<td>Agents, agents, everywhere . . .</td>
</tr>
<tr>
<td>2003–</td>
<td>Human-level AI back on the agenda</td>
</tr>
<tr>
<td>2001–</td>
<td>Availability of very large data</td>
</tr>
</tbody>
</table>
State of the art

♦ Robotic vehicles
♦ Speech recognition
♦ Autonomous planning and scheduling
♦ Game playing
♦ Spam filtering
♦ Machine translations