인공지능

1차시 : Introduction

서울대학교 컴퓨터공학부
담당 교수: 장병탁
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 What Is AI?</td>
<td>3</td>
</tr>
<tr>
<td>1.2 Foundations of AI</td>
<td>9</td>
</tr>
<tr>
<td>1.3 History of AI</td>
<td>10</td>
</tr>
<tr>
<td>1.4 State of the Art</td>
<td>14</td>
</tr>
<tr>
<td>Summary</td>
<td>15</td>
</tr>
<tr>
<td>Homework</td>
<td>16</td>
</tr>
</tbody>
</table>
### 1.1 What Is AI? (1/6)

#### Some definitions of AI

<table>
<thead>
<tr>
<th>Thinking Humanly</th>
<th>Thinking Rationally</th>
</tr>
</thead>
</table>
| • The new effort to make computers think... *machines with minds.*  
  *Automation of activities such as decision making and problem solving.* | • The study of mental faculties through the use of computational models.  
  *The study of the computation that make it possible to perceive, reason, and act.* |

<table>
<thead>
<tr>
<th>Acting Humanly</th>
<th>Acting Rationally</th>
</tr>
</thead>
</table>
| • The art of creating machines that perform functions that require intelligence when performed by people.  
  The study of how to make computers do things at which, at the moment, people are better. | • Artificial intelligence is the study of the design of intelligent agents.  
  *AI is concerned with intelligent behavior in artifacts.* |
1.1 What Is AI? (2/6)

1) Acting humanly: The Turing test approach

- Build a machine that behaves intelligently like humans
  - Turing test or imitation game (1950): 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 a computer
    - Predicted that by the year 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes

- Key capabilities
  - Natural language processing
  - Knowledge representation
  - Automated reasoning
  - Machine learning

  “Can machines think?” → “Can machines behave intelligently?”

- Turing’s test deliberately avoided direct physical interaction between the interrogator and the computer
- Total Turing Test requires additionally: computer vision and robotics
2) Thinking humanly: The cognitive modeling approach

- Build a human-level general intelligence like human minds
  - How to know the actual workings of human minds: cognitive science

- Once we have a sufficiently precise theory of the human mind, it becomes possible to express the theory as a computer program
  - General problem solver (GPS) (Newell & Simon, 1961)

- AI as science: If the program’s input–output behavior matches human behavior, that is evidence that some of the program’s mechanisms could also be operating in human.

- Cognitive science brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind.
3) Thinking rationally: The “laws of thought” approach

- Build a machine that **thinks right** (or **reasons rationally**)
- **Normative** (or **prescriptive**) rather than descriptive
  - Aristotle: what are correct arguments/thought processes?
  - Syllogism: **deriving correct conclusions** from correct premises
  - **Logic**: notation and rules of derivation for thoughts
- **Logicist tradition in AI**: “good-old-fashioned AI” (GOFAI)

**Obstacles:**
- Not easy to convert informal knowledge to formal terms in logic
- Logical reasoning systems can exhaust computational resources
- Not all intelligent behavior is mediated by logical deliberation
4) Acting rationally: The rational agent approach

- Build a machine that behaves rationally or does the right thing
  - A rational agent is one that acts so as to achieve the best expected outcome
  - 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
  - To act rationally requires to reason logically to the conclusion

- Advantages:
  - More general than the “laws of thought” approach: beyond correct inference
  - Rationality is mathematically well defined and completely general
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:
  \[ f : \mathcal{P}^* \rightarrow \mathcal{A} \]
- 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
  - Design best program for given machine resources
1.2 The Foundations of AI

**AI prehistory**

- **Philosophy**: logic, methods of reasoning
- **Mathematics**: formal representation and proof, algorithms, computation
- **Psychology**: human mind, thinking, acting
- **Economics**: formal theory of rational decisions
- **Linguistics**: knowledge representation, grammar
- **Neuroscience**: plastic physical substrate for mental activity
- **Control theory**: homeostatic systems, stability
1.3 History of AI (1/4)

1943–1956: The birth of artificial intelligence
1956–1969: Early enthusiasm, great expectation
1970–1985: Expert systems become industry
            DENDRAL, MYCIN, X–Con
1986–: Return of neural networks (learning systems)
1987–2005 Al’s Winter & AI adopts scientific methods
            Machine learning (& neural networks)
            Probabilistic inference (& Bayesian networks)
            Intelligent agents (& mobile robots)
2000s: Very large data sets (Web, SNS, mobile)
            Machine learning applications
2012–present: Deep learning becomes industry
1.3 History of AI (2/4)

1943  McCulloch & Pitts: Boolean circuit model of brain
1950  Turing's “Computing Machinery and Intelligence"
1952~69  Early enthusiasm, great expectation
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
Figure 1.4  A scene from the blocks world. SHRDLU (Winograd, 1972) has just completed the command “Find a block which is taller than the one you are holding and put it in the box.”
1.3 History of AI (4/4)

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…
2005~    Human-level AI back on the agenda
2012~    Deep learning becomes industry
1.4 State of the Art

What can AI do today?

» Self-driving cars: Stanley (Stanford)
» Speech recognition: Duplex (Google)
» Face recognition: DeepFace (Facebook)
» Image generation: GAN (Nvidia)
» Personal assistant: Siri (Apple), Alexa (Amazon)
» Planning and scheduling: Sojourner (NASA)
» Game playing: AlphaGo (DeepMind)
» Autonomous robots: Rosie (TUM), Atlas (Boston Dynamics)
» Machine translation: NeuralMT (Google)
Intelligence is concerned mainly with rational action.

Ideally, an intelligent agent takes the best possible action in a situation.

AI is an interdisciplinary effort that has been influenced by and has influenced other disciplines, such as philosophy, mathematics, psychology, and neuroscience.

The history of AI has had cycles of success, misplaced optimism, and resulting cutbacks in enthusiasm and funding.

AI has advanced more rapidly in the past decade because of greater use of the scientific method in experimenting with and comparing approaches.

Deep learning has revolutionized AI by enabling intelligent systems to be developed automatically by data-driven learning rather than knowledge-based programming when big data are available.
Homework

Exercises

- 1.1
- 1.9
- 1.10
- 1.11
- 1.14
출처

사진


