Artificial Intelligence
Information

AI Slides 5e, 2018 (人工智能讲义)
林作铨 (Lin Zuoquan)
Information Science Department
Peking University
linzuoquan@pku.edu.cn

Course home page
http://www.math.pku.edu.cn/teachers/linzq/ai

The chapter-by-chapter list is syllabus which is subject to lecture-per-week-per as scheduled

Homework (separate file) is required to submit to TA (online, writing in english, by LaTeX) next to lecture per week on time, and has up to 20-30% proportion of total evaluation (only final examination)
Reference Book

Stuart Russell and Peter Norvig

Artificial Intelligence: A Modern Approach (AIMA)
Prentice Hall, 2011 (3e)
Tsinghua University Press, 2011 (3e reprint), 2013 (3e Chinese ed.)

The book web site: http://aima.cs.berkeley.edu/
including implementations for algorithms

Courtesy some sources and slides from the web sites

More references are included in the slides which would be required to reading as the course progresses, and it is encouraged to look for the supplemental materials from else books and papers to expand knowledge.
Overview

1. Introduction

2. Intelligent Agents

3. Search

4. Constraint Satisfaction

5. Logical Agents

6. First-Order Logic

7. Reasoning
8. Planning

9. Knowledge Representation

10. Uncertainty

11. Decision

12. Learning

13. Language

14. Robotics

15. Philosophy
1 Introduction

1.1 AI

1.2 Foundations

1.3 History

1.4 The state of the art

1.5 Debates
What is AI??

What is Intelligence?

Can a machine think?

(Can a machine behave like a thinking person?)

thinking is some process that people engage in every day

intelligence is an intuitive concept

e.g., people engage in every day

There is not a precise definition of intelligence or thinking

Artificial Intelligence (AI) attempts to understand intelligence entities, strives to building intelligent agents that perceive and act in an environment, and makes computer smarter in human-level intelligence

• understanding the principle of intelligence

• making computers in human-level intelligence replace human work
Intelligence and computation

Computation (or computable by algorithm) is an intuitive concept – explicit effective set of instructions to find the answers to any of a given class of problems in finite steps can be precisely defined by the computational models (computability) Turing machine, recursive functions, automata etc. all these computational models are identical the class of problems computable by algorithm is identical with the class of problems solved by the computational models

Computation is typically carried out by an electronic computer, but might also be carried out by a person or by a mechanical device of some sort (machine)

It was fail to precisely define intelligence something like computation by some mathematical models
Big puzzle: brain → mind (conscious, thinking, understanding) → intelligence

The brain is an existence reference of intelligent machines to imitate

E.g., birds were a reference of heavier-than-air flight
   – shouldn’t just copy it, like kite and earlier airplane
   – airplanes were inspired by birds
   – they use the same basic principles for flight
      aerodynamics and compressible fluid dynamics
   – but airplane don’t flap wings and have feathers

AI needs to understand the principle of intelligence

What is the equivalent of aerodynamics for understanding intelligence??
Views of AI

Weak vs. strong AI

Weak AI: a special purpose computer system can solve a problem in some respect of human-level intelligence

Strong AI: a general purpose computer system can solve a class of problems in almost all respects of human-level intelligence

Views of AI fall into four categories

<table>
<thead>
<tr>
<th>Thinking humanly</th>
<th>Thinking rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acting humanly</td>
<td>Acting rationally</td>
</tr>
</tbody>
</table>
Acting humanly: The Turing test

- **Can a machine think??**
- Operational test for intelligent behavior: 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 decades
- Suggested major components of AI: knowledge, reasoning, language understanding, learning etc.
Turing test

The following interaction from Turing’s paper

Q: Please write me a sonnet on the topic of the Forth Bridge.
A: Count me out on this one. I never could write poetry.
Q: Add 34957 to 70764.
A: (Pause about 30 seconds and then give answer as) 105621.

Given the fact that you can fool some of the people all the time it is not clear how rigorous this particular standard is

Note: language plays a special role in human behavior, not seen in other animals

– much of how we deal with new situations involves using what we have read or been told earlier using language

Reading: Turing. A, Computing machinery and intelligence, 1950
Some Turing test programs

- ELIZA, MegaHAL, TIPS, A.L.I.C.E etc.
- Internet chatbots: MGONZ, NATACHATA, CyberLover etc.
- There is the Loebner Prize for Turing-test-like competition since 1991, but have not been won yet
- chatbots.org

Related tests:
- Microsoft Windows 10 Cortana (so called Xiao Na in Chinese)
- Apple Siri
- Google Assistant
- IBM Waston
- Amazon Alexa
- Facebook Messenger etc.

**Challenge:** The Turing test is not reproducible or amenable to mathematical analysis
Thinking humanly: Cognitive Science

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

Scientific theories of internal activities of the brain
   – What level of abstraction? “Knowledge” or “circuits”
     1) Predicting and testing behavior of human subjects (top-down)
     2) Direct identification from neurological data (bottom-up)

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

Cognitive Science and AI shares one principal direction

Artificial neural networks: artificial neurons mimick the way biological brain with clusters of biological neurons connected by axons also called connectionism in AI
The Artificial Brain projects

**Artificial Brain**: direct human brain emulation using artificial neural networks on a high-performance computing engine

- IBM Blue Brain project (grant from Pentagon, 2008)
  - Google, Baidu etc.
- BRAIN Initiative (US, 2013)
  - The Human Brain Project (Europ, Japan)
- China Brain Project (China, proposal 2018)

**Challenge**: Artificial Brain is simpler to create general intelligent action directly without the principle of intelligence
Thinking rationally: Logic

Aristotle: what are correct arguments/thought processes?

Originally, logic is study of thought, or intelligence;
but mathematical logic by symbolic method intended to study of inferences in mathematics

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 philosophy, mathematics and logic to AI
so-called logicist in math or symbolism in AI
Knowledge and common sense

Knowledge is power of intelligence, with especially common sense
   – having knowledge
   – solving problems by using knowledge

What is common sense??
   How is having common sense any different from being well trained
   on large amounts of data?

Common sense is not explained, but
   rely on our routines of behavior that we have learned over time
   act in situations that are sufficiently unlike the routines we have
   seen before

AI would bring logic back to original goal

Challenge: Not all intelligent behavior is mediated by mathematical
   logic
**Acting rationally: Rational Agent**

**Rational behavior**: doing the right thing

- **right thing**: that is expected to maximize goal achievement, given the available information

A rational agent is one that acts so as to achieve that best (expected) outcome

View points of rational agent something like engineering

- No matter symbolism vs. connectionism
  - (1) all the skills needed for the Turing test allow an agent to act rationally
  - (2) logical inference is one but not all of possible mechanisms for achieving rationality
  - (3) human behavior is adapted for agent design

**Challenge**: The rational agent doesn’t necessarily involve thinking
Foundations

- Philosophy (428BC-present)
- Mathematics (800BC)
- Economics (1776)
- Neuroscience (1861)
- Psychology (1879)
- Computer engineering (1940)
- Control theory and Cybernetics (1948)
- Linguistics (1957)
## AI as interdiscipline

<table>
<thead>
<tr>
<th>Field</th>
<th>Contributions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Philosophy</td>
<td>logic, methods of reasoning</td>
</tr>
<tr>
<td></td>
<td>mind as physical system</td>
</tr>
<tr>
<td></td>
<td>foundations of learning, language, rationality</td>
</tr>
<tr>
<td>Mathematics</td>
<td>formal representation and proof</td>
</tr>
<tr>
<td></td>
<td>probability</td>
</tr>
<tr>
<td>(Computer science)</td>
<td>algorithms, computation, (un)decidability, (in)tractability</td>
</tr>
<tr>
<td>Psychology</td>
<td>adaptation, perception and motor control</td>
</tr>
<tr>
<td></td>
<td>experimental techniques (psychophysics, etc.)</td>
</tr>
<tr>
<td>Linguistics</td>
<td>knowledge representation</td>
</tr>
<tr>
<td></td>
<td>grammar</td>
</tr>
<tr>
<td>Neuroscience</td>
<td>physical substrate for mental activity</td>
</tr>
<tr>
<td>Control theory</td>
<td>homeostatic systems, stability</td>
</tr>
<tr>
<td></td>
<td>simple optimal agent designs</td>
</tr>
<tr>
<td>Economics</td>
<td>decision and operations (e.g., information processing)</td>
</tr>
</tbody>
</table>

*AI is a discipline of computer science*
# History

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>1943</td>
<td>McCulloch &amp; Pitts: Boolean circuit model of brain</td>
</tr>
<tr>
<td>1950</td>
<td>Turing’s “Computing Machinery and Intelligence”</td>
</tr>
<tr>
<td>1952–69</td>
<td>Early AI (early enthusiasm, great expectations)</td>
</tr>
<tr>
<td>1950s</td>
<td>Early AI programs, including Samuel’s checkers program,</td>
</tr>
<tr>
<td></td>
<td>Newell &amp; Simon’s Logic Theorist, Gelernter’s Geometry Engine</td>
</tr>
<tr>
<td>1956</td>
<td>Dartmouth meeting: “Artificial Intelligence” adopted (AI birth)</td>
</tr>
<tr>
<td>1965</td>
<td>Robinson’s complete algorithm for logical reasoning (resolution)</td>
</tr>
<tr>
<td>1966–73</td>
<td>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</td>
</tr>
<tr>
<td>1980–</td>
<td>AI becomes an industry</td>
</tr>
<tr>
<td>1986–</td>
<td>Neural networks return to popularity (deep learning)</td>
</tr>
<tr>
<td></td>
<td>ALife, GAs, soft computing</td>
</tr>
<tr>
<td>1987–</td>
<td>Rapid increase in technical depth of mainstream AI: AI as a science</td>
</tr>
</tbody>
</table>
Brief history of AI

1995– Intelligent agents
2000– Semantic web and web services
2001  Very large data sets (big data)
2004  Human-level AI
2007  Artificial General Intelligence, strong AI
2010– Smart earth and smart products, Internet of things
       AI embedded in the infrastructure of almost every industry
       (ambient intelligence, human-machine intelligence)
2015– **AI age** coming
       -Present  +AI: from Internet/internet of things+AI to industry+internet+AI

          (ai.stanford.edu/ nilsson/QAI/qai.pdf)
          http://aitopics.org/misc/brief-history
State of the art

Which of the following can be done at present?

• Play a decent game of ping-pong
State of the art

Which of the following can be done at present?

- Play a decent game of ping-pong
- 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 ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
State of the art

Which of the following can be done at present?

- Play a decent game of ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
State of the art

Which of the following can be done at present?

- Play a decent game of ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
State of the art

Which of the following can be done at present?

- Play a decent game of ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
State of the art

Which of the following can be done at present?

- Play a decent game of ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
- 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 ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
- 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 ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
- Discover and prove a new mathematical theorem
- Design and execute a research program in molecular biology
- Write an intentionally funny story
Which of the following can be done at present?

- Play a decent game of ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
- Discover and prove a new mathematical theorem
- Design and execute a research program in molecular biology
- Write an intentionally funny story
- 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 ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
- Discover and prove a new mathematical theorem
- Design and execute a research program in molecular biology
- Write an intentionally funny story
- Give competent legal advice in a specialized area of law
- Translate spoken English into spoken Chinese in real time
Which of the following can be done at present?

- Play a decent game of ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
- Discover and prove a new mathematical theorem
- Design and execute a research program in molecular biology
- Write an intentionally funny story
- Give competent legal advice in a specialized area of law
- Translate spoken English into spoken Chinese 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 ping-pong
- Drive safely along a curving mountain road
- Drive safely along in downtown
- Buy a week’s worth of groceries on the web
- Buy a week’s worth of groceries at supermarket
- Win the national championship Chinese chess
- Discover and prove a new mathematical theorem
- Design and execute a research program in molecular biology
- Write an intentionally funny story
- Give competent legal advice in a specialized area of law
- Translate spoken English into spoken Chinese 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 ping-pong
• Drive safely along a curving mountain road
• Drive safely along in downtown
• Buy a week’s worth of groceries on the web
• Buy a week’s worth of groceries at supermarket
• Win the national championship Chinese chess
• Discover and prove a new mathematical theorem
• Design and execute a research program in molecular biology
• Write an intentionally funny story
• Give competent legal advice in a specialized area of law
• Translate spoken English into spoken Chinese in real time
• Converse successfully with another person for an hour
• Perform a complex surgical operation
• Walk with robot secretary in downtown
Debates

AI debates
- Outside: AI vs. philosophy
- Inside: symbolism vs. connectionism
  - with knowledge vs. without knowledge (data)
    - Within symbolism
    - logic vs. probability
    - technical development, e.g., probabilistic logics
    - Within connectionism
    - with reason, e.g., artificial brain vs deep network
    - without reason, e.g., controllers
Symbolism vs. connectionism

Symbolism vs. connectionism
  – two different approaches
  – challenging to each other
  – integration of symbolic and connection methods

All share one principal direction
  the available theories do not explain anything resembling human-level general intelligence

What are principles of intelligence??
Readings

More reference books

- Readings in AI series (various areas of AI, source papers)
  - E.g., Readings in Knowledge Representation

AI philosophical debate

- Hubert Dreyfus, What Computers Can’t do, 1972; What Computers Still Can’t Do, 1992
- Hubert and Stuart Dreyfus, Mind Over Machine, 1986
- Ray Kurzwell, The Singularity Is Near, 2005
- Nick Bostrom, Superintelligence: Paths, Dangers, Strategies, Oxford University Press, 2014
- etc.
AI fiction

- Mary Shelley, *Frankenstein* or Modern Prometheus, 1818
- Samuel Butler, *Darwin among the Machines*, 1863
- Karel Capek, *R.U.R* (Rossum’s Universal Robots), 1921
- Terry Bisson, *They’re Made out of Meat*, 1990
- too much
AI Movie

- Future world, 1976
- The terminator, 1984
- The matrix, 1999
- AI, 2001
- Persons of interest, 2014-15
- West world, 2016
- too much (appear every year)
Asimov’s Three Laws of Robotics:

1. A robot must not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where those orders would conflict with the First Law.
3. A robot must protect its own existence, except where such protection would conflict with the First or Second Law.

- Three Laws were clear, direct, and logical. Asimov’s stories, on the other hand, told how easily they could fail
- The contradictions in Asimov’s laws encouraged others to propose new rules
- any set of rules will always have conflicts and grey areas
AI News

• You can find AI almost everyday from news
• Check it, from now on

• At the time of AI, work hard and enjoy