Artificial Intelligence Application Development

AIMA Chapter 1 (after Russell and Norvig)

Outline

◇ Administration

◇ What is AI?

• the understanding and building of intelligent entitites

◇ Foundations

◇ A brief history

◇ The state of the art

What is AI?

"[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models" (Charniak+McDermott, 1985)
"The study of how to make computers do things at which, at the moment, people are better" (Rich+Knight, 1991)	"The branch of computer science that is concerned with the automation of intelligent behavior" (Luger+Stubblefield, 1993)

Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

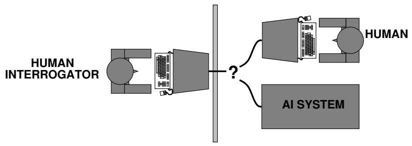
Examining these, we will plump for acting rationally (sort of)

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, reasoning, language understanding, learning

Problem: Turing test is not reproducible, constructive, or amenable to mathematical analysis

Did any of the demos pass the Turing test?

Loebner Prize

"The first formal instantiation of a Turing test"

- http://www.loebner.net/Prizef/loebner-prize.html
- 2001 winner (http://alice.sunlitsurf.com/)

July 7, 2002: Approximating Life. The New York Times Magazine. "Each morning, he wakes before dawn and watches conversations stream by on his screen. Thousands of people flock to his Web site every day from all over the world to talk to his creation, a robot called Alice. It is the best artificial-intelligence program on the planet, a program so eerily human that some mistake it for a real person. As [Richard] Wallace listens in, they confess intimate details about their lives, their dreams; they talk to Wallace's computer about God, their jobs, Britney Spears. It is a strange kind of success: Wallace has created an artificial life form that gets along with people better than he does. ...Is she intelligent? If so, how? In 1950, the pioneering British mathematician Alan Turing grappled with this question in the journal Mind, where he first posed the 'Turing Test' – the gold standard for artificial thought. 'Can machines think?' he asked – and immediately noted that the question hinges, of course, on what 'thinking' is."

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 and thus direction: *the available theroeis 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?
3) Computationally intractable
4) Expressive inadequacies

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

Aristotle (Nicomachean Ethics):
Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good

Beyond programs: autonomy, perception, persistence, adaptive, goal adoption

Rational agents

An agent is an entity that perceives and acts

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
$\rightarrow$ design best program for given machine resources

Course: Rational agents

This course is about designing rational agents

Thinking rationally is only one way of acting rationally

Rationality is more clearly defined and general than human-centered approaches

AI Prehistory and History

Many disciplines (philosophy, mathematics, economics, psychology, linguistics, computer engineering, control theory, neuroscience, and more) have contributed ideas, viewpoints, and techniques to AI

The history of AI has had cycles of success, misplaced optimism, and resulting retrenchments; cycles of new creativity and systematic refinement of best approaches

Philosophy (428 BC-present)

Philosphers made AI conceivable (the mind is on some ways like a machine, operates on knowledge encoded in an internal language, thought can be used to choose action)

Questions

- Can formal rules be used to draw valid conclusions?
- How does the mental mind arise from a physical brain?
- Where does knowledge come from?
- How does knowledge lead to action?

Contributions

- logic, methods of reasoning
- mind as physical system
- foundations of learning, language, rationality

Mathematics (800-present)

Mathematics provided the tools to manipulate logical and probabilistic statments, and to understand computation and reasoning about algorithms

Questions

- What are the formal rules to draw valid conclusions?
- What can be computed?
- How do we reason with uncertain information?

Contributions

- formal representation and proof
- algorithms
- computation, (un)decidability, (in)tractability, (in)completeness
- probability

Economics (1776-present)

Economists formalized the problem of making decisions that maximize the expected outcome to the decision-maker (rationality)

Questions

- How should we make decisions so as to maximize payoff?
- How should we do this when others may not go along?
- How should we do this when the payoff may be far in the future?

Contributions

- utility
- decision theory
- game theory
- operations research / Markov decision processes
- satisficing

Neuroscience (1861-present)

Questions

- How do brains process information?

Contributions

- physical substrate for mental activity

Psychology (1879-present)

Psychologists adopted the idea that humans (and animals) can be considered information processing machines

Questions

- How do humans and animals think and act?

Contributions

- from behavirism to cognitive psychology to cognitive science
- adaptation
- phenomena of perception and motor control
- experimental techniques

Computer Engineering (1940-present)

Computer engineers provided the artifacts that make AI applicaitons possible

Questions

- How can we build an efficient computer?

Contributions

- two way contributions, also software engineering

Control theory and cybernetics (1948-present)

Control theory deals with designing devices that act optimally based on environmental feedback. The mathematical tools here and in AI are coming closer together

Questions

- How can artifacts operate under their own control?

Contributions

- stability
- simple optimal agent designs

Linguistics (1957-present)

Linguistics showed that language use fits into above psychological model

Questions

- How does language relate to thought?

Contributions

- knowledge representation
- grammar
- intersects with AI in computational linguistics / natural language processing

Potted history of AI

1943-1955: Gestation

1956: Birth

1952-1969: Great Expectations

1966-1973: Reality

1969-1979: Knowledge is Power

1980-present: AI and Industry

1986-present: The Return of Neural Networks

1987-present: AI Becomes a Science

1995-present: Intelligent Agents

Gestation

1943 McCulloch & Pitts: Boolean circuit model of brain
1950 Turing's "Computing Machinery and Intelligence"

Birth

1956 Dartmouth meeting: "Artificial Intelligence" adopted

Early enthusiasm, great expectations

1952–69	Look, Ma, no hands!
	Physical symbol hypothesis, lisp, microworlds
1950s	Early AI programs, including Samuel's checkers program,
	Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
1965	Robinson's complete algorithm for logical reasoning

A dose of reality

1966–73	AI discovers computational complexity
	Neural network research almost disappears
	Machine translation research almost disappears

"The spirit is willing but the flesh is weak" translated as "The vodka is good but the meat is rotten"

Knowledge-based systems: The key to power?

1969–79	Early development of knowledge-based systems (e.g., Buchanan et al.)
	From weak methods to expert systems, scripts, frames, etc.

AI becomes an industry

1980–88	Expert systems industry booms
1988–93	Expert systems industry busts: "AI Winter"

The return of neural networks

1986–95	Neural networks return to popularity (connectionism)
1988–	Resurgence of probabilistic and decision-theoretic methods

AI becomes a science

1987– Rapid increase in technical depth
Build on existing theories
Base claims on theorems and/or experments rather than intuition
Real world applications rather than toy examples
Replication of experiments with data and code repositories
Less isolationism

The emergence of intelligent agents

1995– Whole agents rather than fragments
Situated movement
Internet environments ("bots")

2003 View of the field (AIMA)

```
Part I Artificial Intelligence
     1 Introduction
     2 Intelligent Agents
Part II Problem Solving
     3 Solving Problems by Searching
     4 Informed Search and Exploration
     5 Constraint Satisfaction Problems
     6 Adversarial Search
Part III Knowledge and Reasoning
     7 Logical Agents
     8 First-Order Logic
     9 Inference in First-Order Logic
    10 Knowledge Representation
Part IV Planning
    11 Planning
    12 Planning and Acting in the Real World
Part V Uncertain Knowledge and Reasoning
```



```
    13 Uncertainty
    14 Probabilistic Reasoning
    15 Probabilistic Reasoning Over Time
    16 Making Simple Decisions
    17 Making Complex Decisions
Part VI Learning
    18 Learning from Observations
    19 Knowledge in Learning
    20 Statistical Learning Methods
    21 Reinforcement Learning
Part VII Communicating, Perceiving, and Acting
    22 Communication
    23 Probabilistic Language Processing
    24 Perception
    25 Robotics
Part VIII Conclusions
    26 Philosophical Foundations
    27 AI: Present and Future
```

State of the art

Autonomous planning and scheduling (NASA)

Game playing (Deep Blue)

Automomous control (minivan steering)

Diagnosis (medicine)

Logistics planning (Gulf war)

Robotics (surgery)

Language understanding and generation (translation, dialogue)

Problem solving (crossword puzzles)

Discussion Points

There are well-known classes of problems that are intractably difficult or provably undecidable for computers. Does this mean that AI is impossible?

Suppose we extend a classic ANALOGY program so that it can score 200 on a standard IQ test? Would we have a program more intelligent than a human?

Why might the use of introspection (reporting on one's inner thoughts) be a bad methodology?

"Surely computers cannot be intelligent - they can do only what their programmers tell them." Is the latter statement true, and does it imply the former?

What about "Surely animals, humans and computers cannot be intelligent - they can do only what their constituent atoms are told to do by the laws of physics."?

Summary

Different people think of AI differently (thinking or behavior, humans or ideal)

We will adopt the Rational Action view: an Intelligent Agent takes the best possible action in a situation

AI has its roots in many disciplines

The history of AI has had various cycles

Currently: greater use of the scientific method, progress in both theory and practice, integration within subfields and across disciplines

For Next Time

Read Chapter 2 and Empirical Methods Handout

Research and come prepared to report on the latest (or other recent) winner of the Loebner prize. What techniques does it use? How does it advance the state of the art in AI?

HW1 upcoming