

Chapters 1–2: Introduction to AI

DIT410/TIN172 Artificial Intelligence

Peter Ljunglöf
modified from slides by Poole & Mackworth

(Licensed under Creative Commons BY-NC-SA v4.0)

24 March, 2015

Outline

- ➊ *What is AI? (1.1)*
- ➋ *A brief history of AI (1.2)*
- ➌ *Interlude: What is this course, anyway?*
- ➍ *Dimensions of complexity (1.5)*
- ➎ *Solutions (1.4)*

Outline

- 1 *What is AI? (1.1)*
- 2 *A brief history of AI (1.2)*
- 3 *Interlude: What is this course, anyway?*
- 4 *Dimensions of complexity (1.5)*
- 5 *Solutions (1.4)*

What is intelligence?

“It is not my aim to surprise or shock you – but the simplest way I can summarize is to say that there are now in the world machines that can think, that learn, and that create. Moreover, their ability to do these things is going to increase rapidly until – in a visible future – the range of problems they can handle will be coextensive with the range to which human mind has been applied.”

Herbert A Simon

What is intelligence?

“It is not my aim to surprise or shock you – but the simplest way I can summarize is to say that there are now in the world machines that can think, that learn, and that create. Moreover, their ability to do these things is going to increase rapidly until – in a visible future – the range of problems they can handle will be coextensive with the range to which human mind has been applied.”

Herbert A Simon (1957)

What is intelligence?

“It is not my aim to surprise or shock you – but the simplest way I can summarize is to say that there are now in the world machines that can think, that learn, and that create. Moreover, their ability to do these things is going to increase rapidly until – in a visible future – the range of problems they can handle will be coextensive with the range to which human mind has been applied.”

Herbert A Simon (1957)

*“The question of whether machines can think. . .
... is about as relevant as the question whether
submarines can swim.”*

Edsger W Dijkstra (1984)

Strong and Weak AI

One may dream about...

...that computers can be made to think on a level at least equal to humans, that they can be conscious and experience emotions.

Strong AI

This course is about...

...adding “thinking-like” features to computers to make them more useful tools. That is, “not obviously machine like”.

Weak AI

Weak AI

- Weak AI is a category that is flexible, as soon as we understand how an AI-program works, it appears less “intelligent”.
- And as soon as a part of AI is successful, it becomes an own research area! E.g., large parts of advanced search, parts of language understanding, parts of machine learning and probabilistic learning etc.
- And AI is left with the remaining hard-to-solve problems!

Outline

- 1 *What is AI? (1.1)*
- 2 *A brief history of AI (1.2)*
- 3 *Interlude: What is this course, anyway?*
- 4 *Dimensions of complexity (1.5)*
- 5 *Solutions (1.4)*

A brief history of AI

| | |
|----------------|---|
| 1943 | McCulloch & Pitts: Boolean circuit model of brain |
| 1950 | Alan Turing's "Computing Machinery and Intelligence" |
| 1952–69 | Look, Ma, no hands! |
| 1950s | Early AI programs: e.g., Samuel's checkers program, Gelernter's Geometry Engine, Newell & Simon's Logic Theorist and General Problem Solver |
| 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 |
| 1971 | Terry Winograd's Shrdlu dialogue system |

A brief history of AI

- 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; increase in technical depth
“Nouvelle AI”: ALife, GAs, soft computing
- 1995– Agents, agents, everywhere. . .
- 1997 IBM Deep Blue beats the World Chess Champion
- 2001– Very large datasets: Google gigaword corpus, Wikipedia
- 2003– Human-level AI back on the agenda
- 2011 IBM Watson wins Jeopardy
- 2012 US state of Nevada permits driverless cars
- 2014 “Deep learning”: image tagging, recommendation
systems, speech translation, pattern recognition

Outline

- 1 *What is AI? (1.1)*
- 2 *A brief history of AI (1.2)*
- 3 *Interlude: What is this course, anyway?*
- 4 *Dimensions of complexity (1.5)*
- 5 *Solutions (1.4)*

Course overview

- The website:
 - ▶ <http://ai-course-tin172-dit410.github.io/>
- The examinations:
 - ▶ the Shrdlite project (deadlines: 3/4, 23/5)
 - ▶ the written essay (deadlines: 9/5, 23/5)
 - ▶ the written examination (2/6, 18/8)
- The people
 - ▶ the teachers
 - ▶ the student representatives
- The project groups:
 - ▶ forming a group (deadline: 26/3)
- The schedule:
 - ▶ the lectures
 - ▶ the supervision
 - ▶ the exercises

Outline

- 1 *What is AI? (1.1)*
- 2 *A brief history of AI (1.2)*
- 3 *Interlude: What is this course, anyway?*
- 4 *Dimensions of complexity (1.5)*
- 5 *Solutions (1.4)*

Dimensions of complexity

| | |
|-----------------------------|---|
| Modularity | flat / modular / hierarchical |
| Representation | states / features / relations |
| Planning horizon | static / finite / indefinite / infinite |
| Sensing uncertainty | fully / partially observable |
| Effect uncertainty | deterministic / stochastic |
| Preference | goals / complex |
| Number of agents | single / multiple |
| Learning | knowledge is given / learned |
| Computational limits | perfect / bounded rationality |

Modularity

- Model at one level of abstraction: **flat**
- Model with interacting modules that can be understood separately: **modular**
- Model with modules that are (recursively) decomposed into modules: **hierarchical**
 - ▶ *Example:* Planning a trip to see the Mona Lisa in Paris.
- Flat representations are adequate for simple systems
- Complex biological systems, computer systems, organizations are all hierarchical
- A flat description is either continuous or discrete.
- Hierarchical reasoning is often a hybrid of continuous and discrete.

Representation

Much of modern AI is about finding compact representations and exploiting the compactness for computational gains.

- **Explicit states**
 - ▶ a state is one way the world could be
- **Features** or propositions
 - ▶ states can be described using features
 - ▶ 20 binary features can represent $2^{20} = 1,048,576$ states
- **Relations** and individuals
 - ▶ there is a feature for each relationship on each tuple of individuals

Planning horizon

...how far into the future the agent looks.

- **Static:** the world does not change
- **Finite stage:** the agent reasons about a fixed finite number of time steps
- **Indefinite stage:** the agent reasons about a finite, but not predetermined, number of time steps
- **Infinite stage:** the agent plans for going on forever

Uncertainty

There are two dimensions for uncertainty:

- Sensing uncertainty
- Effect uncertainty

In each dimension an agent can have

- No uncertainty: the agent knows which world is true
- Disjunctive uncertainty: there is a set of possible worlds
- Probabilistic uncertainty: a probability distribution over worlds

Why Probability?

- Agents need to act even if they are uncertain.
- Predictions are needed to decide what to do:
 - ▶ definitive predictions: “you will be run over tomorrow“
 - ▶ disjunctions: “be careful or you will be run over”
 - ▶ point probabilities: “probability that you will be run over tomorrow is 0.002 if you are careful and 0.05 if you are not”
 - ▶ probability ranges: “you will be run over with probability in range [0.01,0.34]”
- Acting is gambling: agents who don't use probabilities will lose to those who do.
- Probabilities can be learned from data and prior knowledge.

Sensing uncertainty: Observability

Whether an agent can determine the state from its observations:

- **Fully-observable:** the agent can observe the state of the world
- **Partially-observable:** there can be a number states that are possible given the agent's observations

Effect uncertainty: Determinism

If an agent knew the initial state and its action, could it predict the resulting state?

- **Deterministic:** the resulting state is determined from the action and the state
- **Stochastic:** there is uncertainty about the resulting state

Preferences

- **Achievement goal** is a goal to achieve. This can be a complex logical formula.
 - ▶ *Example:* coffee delivery robot
- **Complex preferences** may involve tradeoffs between various desiderata, perhaps at different times.
 - ▶ **ordinal:** only the order matters
 - ▶ **cardinal:** absolute values also matter
 - ▶ *Example:* medical doctor

Number of agents

- **Single agent** reasoning is where an agent assumes that any other agents are part of the environment.
- **Multiple agent** reasoning is when an agent reasons strategically about the reasoning of other agents.

Agents can have their own goals: cooperative, competitive, or goals can be independent of each other

Learning from experience

Whether the model is fully specified a priori:

- Knowledge is given
- Knowledge is learned from data or past experience

Computational limits

- **Perfect rationality:** the agent can determine the best course of action, without taking into account its limited computational resources.
- **Bounded rationality:** the agent must make good decisions based on its perceptual, computational and memory limitations.

Outline

- 1 *What is AI? (1.1)*
- 2 *A brief history of AI (1.2)*
- 3 *Interlude: What is this course, anyway?*
- 4 *Dimensions of complexity (1.5)*
- 5 *Solutions (1.4)*

Defining a Solution

- Given an informal description of a problem, what is a solution?
- Typically much is left unspecified, but the unspecified parts can't be filled in arbitrarily.
- Much work in AI is motivated by *common-sense reasoning*.
The computer needs to make common-sense conclusions about the unstated assumptions.

Quality of Solutions

- Does it matter if the answer is wrong or answers are missing?

Classes of solution:

- An **optimal solution** is a best solution according some measure of solution quality.
- A **satisficing solution** is one that is good enough, according to some description of which solutions are adequate.
- An **approximately optimal solution** is one whose measure of quality is close to the best theoretically possible.
- A **probable solution** one that is likely to be a solution.