T-622-ARTI Introduction to AI

- Teacher: Hannes Högni Vilhjálmsson
  (hannes@ru.is)
- Assistant: Angelo Cafaro
  (angelo08@ru.is)

Classes
- Mondays at 13:10 (M.1.02)
- Thursdays at 14:00 (M.1.03)
- Fridays at 10:20 (M.1.02)

Topics Covered
- Agents and Architecture (chapter 2)
- Search (chapters 3-6)
- Logic and reasoning (chapters 7-9)
- Planning (chapter 10-11)
- Bayesian Networks (chapter 14)
- Learning (chapter 18)
- Perception
- Natural Language

Approach
- Lectures (Mondays, Fridays)
  - Introduce theory
- Paper Discussion (Mondays)
  - Your direct participation in topical discussion!
- Labs (Thursdays)
  - Hands-on Practice and Problem Solving
- Assignments and Final Project
**T-622-ARTI Introduction to AI**

- **Final Grade**
  - Discussion 20%
  - Programming Assignments (x2) 10%
  - Problem Sets (x2) 10%
  - Final Project 30%
  - Final Written Exam 30%

- **Attendance**
  - 70% required for taking final exam

---

**T-622-ARTI Introduction to AI**

- **Discussion**
  - Specific short reading is assigned (MON)
  - You post 2 questions online (SUN)
  - We discuss your questions together in class (MON)

  - Your participation here is 20% of grade!

- **Do you know of thought provoking readings?**
  - Let me know and I may schedule them

---

**Hannes’ availability**

- After classes
- "open office policy" 
  - Visit my office anytime (SCS reception, Venus 3. floor)
- Send email or call:
  - hannes@ru.is, 599 6323 (GSM: 824 8814)
- On MSN:
  - skuggavera@hotmail.com
Introduction AI

Russell and Norvig:
Chapter 1

WHAT IS AI?

Chapter 1.1

WHAT IS AI?

What is AI?

Empirical Science

Think like humans

Think rationally

Math / Engineering

Act like humans

Act rationally
Acting Humanly

- The Turing Test
  - Proposed by Alan Turing (1950)
  - Establishes human action as the benchmark
  - AI passes test if written interrogation by human does not unveil it as a computer
  - Provides plenty to work on!
    - Natural Language Processing
    - Knowledge Representation
    - Automated Reasoning
    - Machine Learning

- The Turing Test (cont.)
  - The “Physical” test has also been proposed
  - Involves even more fields including
    - Computer Vision
    - Robotics
  - Seems to cover most of AI!
  - BUT! Does it help us to build intelligence?
    - Human flight came with study of aerodynamics, not by imitating birds.

Thinking Humanly

- Understanding the inner working of the human mind through psychological experiments leading to
  - Precise and testable theories
  - Computational models
  - This is the field of Cognitive Science
    - Computational models may migrate into AI, but in themselves are not enough for Cognitive Science
Thinking Rationally

- What is "right thinking"?
- The Greeks tried to answer this with laws of thought
  - Initiated the field of logic
- Logicist AI tries to describe all kinds of things and problems with a precise logical notation and use that to find "right solutions"
- Problems: (A) Incomplete information; (B) Impractical implementation

Acting Rationally

- Rational Agents try to achieve the best (expected) outcome
- May use logic inference, but also other approaches to rational behavior
  - E.g. Reflexes can produce rational reaction
- Here we choose the Rational Agent perspective because
  - More general than pure logic inference
  - Better defined than human rationality
Philosophy

- Aristotle (384-322 BC)
  - Generating conclusions mechanically given a premise
- Hobbes (1588-1679)
  - Reasoning like numerical computation
- Pascal (1623-1662)
  - Numerical calculating machine - “like thought!”
- Leibniz (1646-1716)
  - Machine operating on concepts, not numbers

So the mind is a machine?

- What about free will?
  - Rocks governed by physics don’t “decide” to fall!

- Explained in terms of “the non-physical side”
  - Dualism
- Explained in terms of a natural choice process
  - Materialism

The mind manipulates knowledge

- Where does the knowledge come from?
- It all starts at the senses, so perception is key!
- And finally, we need action, as part of this picture of the mind

- Aristotle proposed a planning algorithm based on the knowledge of action outcomes
**Mathematics**

- **Logic**: Boolean logic (Boole, 1847)
- **Logic**: First-order logic (Frege, 1879)
- **Computation**: Intractability (1960s)
  - Computation time grows exponentially with instance size
- **Computation**: NP-completeness (Cook, 1971)
  - We can identify the really hard problems
- **Probability**: Cardano, 1501-1576
  - Using new evidence (Bayes, 1702-1761)

**Economics**

- **Rationality leading to preferred outcomes or utility**: Walras, 1834-1910
- **Decision Theory**
  - Combines Probability Theory and Utility Theory (environment and individual)
- **Game Theory**
  - Decision Theory with other rational agents in the environment
- **Operations Research**
  - Sequence of decisions and not immediate payoffs

**Neuroscience**

- The brain seems to “cause minds”!
  - Collection of simple cells leads to thought, action, and consciousness – exactly how is still mystery
  - Areas of the brain seem to map to cognitive functions or body parts, yet this can change
- There are $10^{11}$ neurons in the brain; CPUs will reach that number of gates around 2020 according to Moore’s Law
- But in the brain, all units are active simultaneously!
### Psychology

- **Behaviorism (Watson, 1878-1958)**
  - We can only study the stimulus and response. Knowledge, beliefs, goals and reasoning is "folk psychology".

- **Cognitive Psychology (James, 1842-1910)**
  - The brain as an information-processing device
  - Beliefs and goals just as scientific as pressure (Craik, 1943)

- **Cognitive Science (MIT Workshop, 1956)**
  - Computer models addressing psychology

### Computer Engineering

- **Punchcard Loom (Jacquard, 1805)**
  - Programmable machine

- **Difference Engine (Babbage, 1792-1871)**
  - Math tables for engineering (not built, but works)

- **Analytical Engine (also Babbage)**
  - Universal computation
    - memory, programs, jumps
  - Ada Lovelace wrote programs for it
  - Never built: What if?

- **Steampunk Fiction**

- **Heath Robinson (Turing, 1940)**
  - Designed to decipher German messages

- **Colossus (Turing, 1943)**
  - General purpose machine based on vacuum tubes

- **Z-3 (Zuse, 1941)**
  - Programmable

- **ABC (Atanasoff, 1942)**
  - First electronic computer

- **ENIAC (1946)**
Control Theory and Cybernetics

- First, only living things could modify behavior in response to changes in environment!
- Water Clock (Ktesibios, 250 BC)
  - Kept water running at constant pace
- Thermostat (Drebbel, 1572-1633)
- Steam Engine Governor (Watt, 1736-1819)

Control Theory and Cybernetics

- Wiener (1894-1964) looking at control and cognition
  - Mental mechanism trying to minimize error, a challenge to behaviorism

Control Theory and Cybernetics

- Modern Control Theory, especially stochastic optimal control tries to maximize an objective function over time
- Optimal behavior, like the rational agents

- Why not the same field?
- AI breaks out of the math of control theory and considers "softer" things like language, vision and planning

Linguistics

- Behaviorist theory does not address creativity in language
- Chomsky (1957) explains this creativity with syntactic structures, going back to Panini (350 BC), formal enough for programming

Computational Linguistics

- Has to deal with the context of understanding and producing language
- Therefore connected with Knowledge Representation
Chapter 1.3

HISTORY OF AI

Artificial Neuron (1943)
- Warren McCulloch and Walter Pitts
- ON or OFF, depending on enough stimulation by neighboring neurons
- All logical connectives (AND, OR, NOT) could be implemented by simple nets
- Suggested that these could also be made to “learn”

Neural Network Computer (1950)
- Marvin Minsky and Dean Edmonds
- 3000 vacuum tubes simulated 40 neurons
"Complete Vision of AI" (1950)
- Alan Turing articulated this vision in an article called "Computing Machinery and Intelligence"
- That's where he proposed the Turing Test as well as machine learning and genetic algorithms.

Dartmouth Workshop (1956)
- 10 people brought together, who all had shared interest in automata theory, neural nets and study of intelligence
- Newell and Simon showed the Logic Theorist reasoning program
- McCarthy's term "Artificial Intelligence" was adopted for the field.

50s: Exciting Early Years
- General Problem Solver (Newell and Simon, 1957)
- Geometry Theorem Prover (Gelernter, 1959)
- Checkers players (Samuel, 1956-)

---

---

---

---
1958: Good Year for McCarthy
- The Lisp programming language
- Time sharing (multiple users on a computer)
- Describes hypothetical "Advise Taker"
  - General knowledge representation and reasoning

60s: Minsky and students flourish
- Chose series of limited problems that appeared to require intelligence to solve:
  - Microworlds
  - Most famous is the blocks world

But the REAL world is tough!
- The AI lacked general knowledge
  - Russian to English translation programs failed!
- The AI methods didn't scale up
  - Intractable problems out of reach
  - The world is BIG
- Doubts about capabilities of neural nets
  - Could represent less than first expected
**70s: Expert Systems Save the Day**

- **DENDRAL (Buchanan, 1969) - Chemistry**
  - Inferring molecular structure from mass spectrum
  - Intractable to check all structures
  - Instead: Checks structure patterns known by human experts

- **MYCIN (Feigenbaum, 70s) - Medical**
  - Diagnoses blood infections
  - 450 rules and system better than junior doctors

**80s: Industrial AI Boom**

- DEC XCON Expert System saved them $40 million per year in 1986!
- By 1988 DEC’s AI group had 40 expert systems deployed.
- Nearly every major US corporation established their own AI group and was using or looking into expert systems.
- Extravagant promises, but failure to deliver in the end caused a new "AI Winter".
90s: AI Becomes a Science

- Gone back to existing theories to build a strong foundation - comparing methods
- Example
  - Hidden Markov Models (HMMs) based on math
  - Bayesian Networks based on neural nets
- Resurgence of formalization and specialization has lead to isolation of more “cutting edge” work like vision and robotics

Bringing it all together: Agents

- Agents provide an opportunity to work on a complete AI system, across approaches