# Introduction to Artificial Intelligence

## Introduction

**Bernhard Beckert**

**UNIVERSITÄT KOBLENZ-LANDAU**

**Summer Term 2003**

# What is Artificial Intelligence (AI)?

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

# What is Artificial Intelligence (AI)?

**Views of AI fall into four categories**

| Thinking humanly | Thinking rationally |
|---|---|
| Acting humanly | Acting rationally |

**Most AI researchers in Computer Science go for acting rationally**
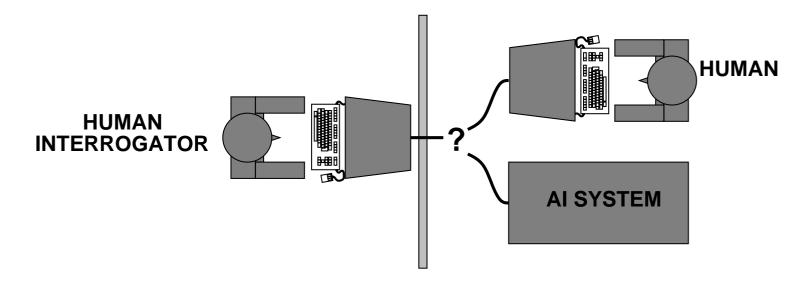
# Acting humanly: The Turing test

**Turing (1950):** *Computing machinery and intelligence*

- **"Can machines think?'**

  **"Can machines behave intelligently?"**

- **Operational test for intelligent behavior: the Imitation Game**

**Classical Turing test**

HUMAN
INTERROGATOR

?

HUMAN

AI SYSTEM

# Acting humanly: The Turing test

**Total Turing test**

**Includes physical interactions with environment**

– **speech recognition**
– **computer vision**
– **robotics**

**Problem of Turing test**

**Turing test is**

– **not reproducible**
– **not constructive**
– **not amenable to mathematical analysis**

# Acting humanly: The Turing test

**Turing's predictions**

- 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 representation, reasoning, language understanding, learning

**Turing's paper online available at**

```
http://www.abelard.org/turpap/turpap.htm
```

# The Turing Test and Subfields of AI

- Knowledge Representation

- Searching

- Automated Reasoning (Deduction)

- Machine Learning

- Natural Language Processing

- Computer Vision

- Robotics

# Turing's and other Tests

## Loebner Prize

**A restricted Turing test, held annually in the form of a competition**

> The Loebner Prize is awarded annually for the computer program that best emulates natural human behavior. During the contest, a panel of independent judges attempts to determine whether the responses on a computer terminal are being produced by a computer or a person, along the lines of the Turing Test. The designers of the best program each year win a cash award and a medal. If a program passes the test in all its particulars, then the entire fund will be paid to the program's designer and the fund abolished.

`http://www.loebner.net/Prizef/loebner-prize.html`

# Turing's and other Tests

## Robot World Cup Initiative (RoboCup)

> Uses playing a soccer game as a standard problem, where a wide range of technologies can be integrated and examined. Carried out for various classes of robots and software agents.
>
> Goal: By the year 2050, develop a team of fully autonomous humanoid robots that can win against the human world soccer champions.

`http://www.robocup.org`

# Thinking humanly: Cognitive Science

**Cognitive revolution (1960s)**

**Information-processing psychology replaced
prevailing orthodoxy of behaviorism**

**Requires scientific theories of internal activities of the brain . . .**

– **What level of abstraction?**
– **"Knowledge" or "circuits"?**

**and Validation**

– **Predicting and testing behavior of human subjects (top-down)**
  $\Rightarrow$ **Cognitive Science**
– **Direct identification from neurological data (bottom-up)**
  $\Rightarrow$ **Cognitive Neuroscience**

**Second-order / Epistemological knowledge**

**"We know what we know and what we don't know"**

# Thinking rationally: Laws of Thought

**Normative (prescriptive) rather than descriptive**

Aristotle: What are correct arguments / thought processes?

Several Greek schools developed various forms of **logic**:

– notation
– rules of derivation (syllogisms)

Direct line through mathematics and philosophy to modern AI

**Problems**

- Not all intelligent behavior is mediated by logical deliberation

- What is the purpose of thinking? What thoughts should I have?

- What is the logic of human reasoning?

# 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)

**Aristotle:** *Nicomachean Ethics*

*Every art and every inquiry, and similarly every action
and pursuit, is thought to aim at some good*

# Acting rationally

**A thoroughly pragmatic point of view**

- In practical terms, so far the most fruitful road taken by AI

- Completely misses the perhaps most central aspect
  of being human:

<span style="color:red">**Consciousness**</span>

# Philosophical / theological questions

**Can machines have**

– **minds?**
– **souls?**
– **consciousness?**

**Do sufficiently intelligent machines (automatically) have**

– **minds?**
– **souls?**
– **consciousness?**

**Two theories**

**Dualism:**      **Body and soul/mind are separate things**

**Materialism:**    **There is no immaterial soul/mind
(J. R. Searle: "Brains cause minds")**

# Rational agents

## Agent

- An entity that perceives and acts

- A useful way to think about building AI programs is in terms of designing (and implementing) rational agents

## Abstract definition

An agent is a function from percept histories to actions:

$$f : \mathcal{P}^* \rightarrow \mathcal{A}$$

# Rational agents

## Optimal agent

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

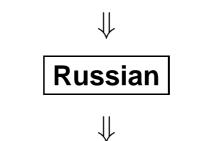$\Rightarrow$ Design best agent for given machine resources

# AI: Historical Roots

| | |
|---|---|
| **Philosophy** | **logic, methods of reasoning** |
| | **mind as physical system** |
| | **foundations of learning, language, rationality** |
| **Mathematics** | **formal representation and proof** |
| | **algorithms** |
| | **computation, (un)decidability, (in)tractability** |
| | **probability** |
| **Psychology** | **adaptation** |
| | **phenomena of perception and motor control** |
| | **experimental techniques (psychophysics, etc.)** |
| **Linguistics** | **knowledge representation** |
| | **grammar** |
| **Neuroscience** | **physical substrate for mental activity** |
| **Control theory** | **homeostatic systems, stability** |
| | **simple optimal agent designs** |

# Potted history of AI

| 1943 | McCulloch & Pitts: Boolean circuit model of brain |
|------|---------------------------------------------------|
| 1950 | Turing's *Computing Machinery and Intelligence* |
| 1952–69 | Look, Ma, no hands! |
| 1950s | Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine |
| 1956 | Dartmouth meeting: "Artificial Intelligence" adopted |
| 1963 | 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 |
| 1980–88 | Expert systems industry booms |
| 1988–93 | Expert systems industry busts: "AI Winter" |
| 1985–95 | Neural networks return to popularity |
| 1988– | Probabilistic methods; enormous increase in technical depth "Nouvelle AI": ALife, GAs, soft computing |
| 1995– | Agents is the new buzzword |

# State of the art

**An early effort in Machine Translation**

"The spirit is willing, but the flesh is weak"

$\Downarrow$

| Russian |

$\Downarrow$

"The vodka is good, but the meat is rotten"

# State of the art, more seriously

**Which of the following can be done by an AI program/robot at present?**

- Play a decent game of table tennis

- Drive along a curving mountain road

- Drive in the center of a big city

- Play a decent game of Bridge or Go

- *Discover* and prove a new mathematical theorem

- Write an intentionally funny story

- Give competent legal advice in a specialized area of law

- Translate spoken English into spoken German in real time

# State of the art

**AI programs . . .**

- **Regularly win a chess game against grandmasters**

  `http://www.lkessler.com/cclinks.shtml`

- *Roughly* **translate a** *text* **from one language into another**

  `http://cslu.cse.ogi.edu/HLTsurvey/HLTsurvey.html`

- **Proved a mathematical problem that was open for 60 years**

  `http://www-unix.mcs.anl.gov/~{}mccune/papers/robbins/`

# AI research challenges

- **Reflective architecture for agents (epistemological reasoning)**

- **Compilation from deliberative reasoning to reflex system (e.g., reinforcement learning)**

- **Make use of massive parallelism in an effective way**

- **Bridge the gap between *human* and *rational* AI**

# Some promising application areas

**Formal software and hardware verification (automated reasoning)**

Intel spends up to 90 % of budget in processor development
for verification

**The Semantic Web (knowledge representation, learning)**

From keyword-based search to content-based search

**Data mining, automatic discovery of structures**

From data to information, *Discovery Science*

Probabilistic methods, learning, fuzzy sets

# Some promising application areas

**Autonomous agents**

– **cleaning robots**
– **military applications**
– **etc.**

**Recognition of speech, gestures, facial expression**

– **handicapped people**
– **cars/planes**
– **surveillance & security**

**Automated translation from/to natural language**

# Summary

- Early success, exaggerated claims, "roller coaster" ride

- Spin-off to mainstream CS
  (e.g., search, knowledge representation, complexity theory)

- Unresolved dichotomy "soft"/human-oriented vs. "hard"/rational AI

- Hard AI gained much in depth and rigour in recent years

- Many impressive tasks can be achieved with AI technology today

- Technological developments

  – WWW
  – computerization of all devices (ubiquitous computing)
  – data explosion

  create highly promising application areas for AI

# Quote: Alan Turing (1950)

We can only see a short distance ahead, but we can see that much remains to be done.

# The Chinese room (J. R. Searle)

**System consists of a room with**

- a human who does not understand Chinese
- a rule book written in English
- blank paper to write on

When the human follows the rules in the book,
he/she can generate intelligent aswers in Chinese
to questions posed in Chinese.

**Questions**

- Does the human now understand Chinese?

- Does the room understand Chinese?

- Is the room intelligent?