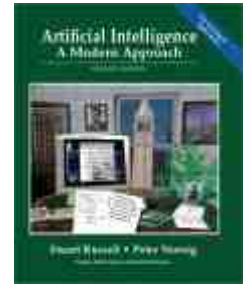


# Einführung in die Künstliche Intelligenz

- Homepage
  - <http://www.ke.informatik.tu-darmstadt.de/lehre/ss07/ki/>
- Termine:
  - Dienstag 12:35-14:15 S202/C120 oder C205
  - Donnerstag 11:40-13:20 S103/123
- 3 VO + 1 UE
  - Vorlesungen und Übungen werden in Doppelstunden abgehalten
  - Terminplan wird auf der Web-Seite aktualisiert
- Übungsbetreuung
  - Dr. Gunter Grieser
  - Tafelübungen
- Schriftliche Klausur
  - Termin wird noch bekannt gegeben
- Anrechenbar für Bereich 3 (zumindest bei mir)

# Text Book

- The course will mostly follow
  - Stuart Russell und Peter Norvig: *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2nd edition, 2003.
- Deutsche Ausgabe:
  - Stuart Russell und Peter Norvig: *Künstliche Intelligenz: Ein Moderner Ansatz*. Pearson-Studium, 2004. ISBN: 978-3-8273-7089-1.
- Home-page for the book:
  - <http://aima.cs.berkeley.edu/>
- Course slides
  - in English (lecture is in German)
  - will be available from Home-page



# What is Artificial Intelligence

- Different definitions due to different criteria
- Two dimensions:
  - Thought processes/reasoning vs. behavior/action
  - Success according to human standards vs. success according to an ideal concept of intelligence: rationality.

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

# Definitions of Artificial Intelligence

<p><b>Systems that think like humans</b></p>	<p><b>Systems that think rationally</b></p>
<p>“The exciting new effort to make computers think . . . <i>machines with minds</i>, in the full and literal sense.” (Haugeland, 1985)</p> <p>“[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . .” (Bellman, 1978)</p>	<p>“The study of mental faculties through the use of computational models.” (Charniak and McDermott, 1985)</p> <p>“The study of the computations that make it possible to perceive, reason, and act.” (Winston, 1992)</p>
<p><b>Systems that act like humans</b></p>	<p><b>Systems that act rationally</b></p>
<p>“The art of creating machines that perform functions that require intelligence when performed by people.” (Kurzweil, 1990)</p> <p>“The study of how to make computers do things at which, at the moment, people are better.” (Rich and Knight, 1991)</p>	<p>“Computational Intelligence is the study of the design of intelligent agents.” (Poole <i>et al.</i>, 1998)</p> <p>“AI . . . is concerned with intelligent behavior in artifacts.” (Nilsson, 1998)</p>
<p><b>Figure 1.1</b> Some definitions of artificial intelligence, organized into four categories.</p>	

# Systems that think like humans

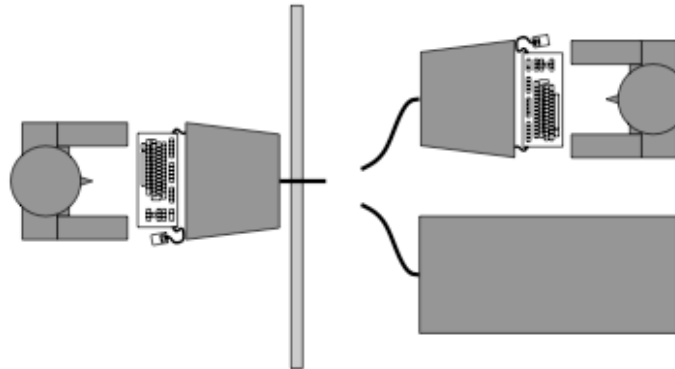
- How do humans think?
  - Requires scientific theories of internal brain activities (cognitive model):
    - Level of abstraction? (knowledge or circuitry?)
    - Validation?
      - Predicting and testing human behavior
      - Identification from neurological data
- Cognitive Science
  - brings together computational models from AI
  - and experimental techniques from psychology
  - to construct precise and testable theories of the mind
  - Cognitive Science is now distinct from AI
- Cognitive Neuroscience
  - How does the brain work at the neuronal level?

# Systems that think rationally

- Capturing the laws of thought
  - Aristotle: What are ‘correct’ argument and thought processes?
    - Correctness depends on irrefutability of reasoning processes.
    - Syllogisms:
      - first patterns of correct formal reasoning
      - “Socrates is a man, all man are mortal → Socrates is mortal”
  - This study initiated the field of logic.
    - The logicist tradition in AI hopes to create intelligent systems using logic programming.
- Problems:
  - It is hard to formalize knowledge exactly
    - Feigenbaum Bottleneck in expert systems
  - Practical constraints
    - it is easy to write a logically optimal chess player...
  - Not all intelligence is mediated by logic behavior

# Systems that act like humans

- When does a system behave intelligently?
  - Turing (1950) Computing Machinery and Intelligence  
Operational test of intelligence: imitation game



- Test still relevant now, yet might be the wrong question.
  - Requires the collaboration of major components of AI: knowledge, reasoning, language understanding, learning, ...
- Problem with Turing test:
  - not reproducible, constructive or amenable to mathematical analysis.

# Systems that act rationally

- Rational behavior: “doing the right thing”
  - The “Right thing” is that what is expected to maximize goal achievement given the available information.
- Can include thinking, yet in service of rational action
  - Action without thinking: e.g. reflexes.
- Two advantages over previous approaches
  - More general than law of thoughts approach
    - in many situations, a provably correct action does not exist
  - More amenable to scientific development
    - rationality can be defined and optimized
- On the other hand
  - perfect rationality is only feasible in ideal environments.
  - rationality is often not a very good model of reality.
    - humans are, e.g., very bad in estimating probabilities...



# Foundations of AI

Different fields have contributed to AI in the form of ideas, viewpoints and techniques.

- *Philosophy*: Logic, reasoning, mind as a physical system, foundations of learning, language and rationality.
- *Mathematics*: Formal representation and proof algorithms, computation, (un)decidability, (in)tractability, probability.
- *Psychology*: adaptation, phenomena of perception and motor control.
- *Economics*: formal theory of rational decisions, game theory.
- *Linguistics*: knowledge representation, grammar.
- *Neuroscience*: physical substrate for mental activities.
- *Control theory*: homeostatic systems, stability, optimal agent design.

# Subdisciplines of AI

- Natural Language Processing
- Knowledge Representation
- Automated Reasoning
- Machine Learning
- Computer Vision
- Robotics
- ...

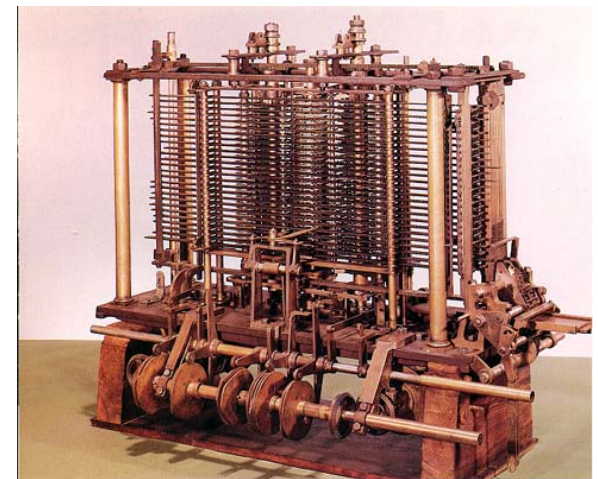
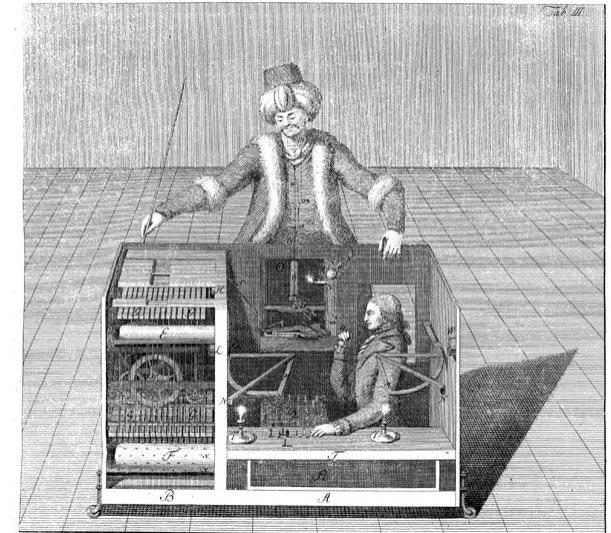
# A Brief History of AI

- Greek mythology:
  - Hephaestus built Talos, a giant intelligent bronze robot
- 13<sup>th</sup> century:
  - brazen head: oracle in the form of a talking head made of brass
  - supposedly owned by Roger Bacon and Albertus Magnus
- 15<sup>th</sup> century:
  - da Vinci drafted robot design
- 16<sup>th</sup> century:
  - Rabbi Loew made the giant Golem of clay to protect the Jewish community in Prague
  - remains are still supposed to be there...
- 17<sup>th</sup> century:
  - Descartes – “animals are complex machines”



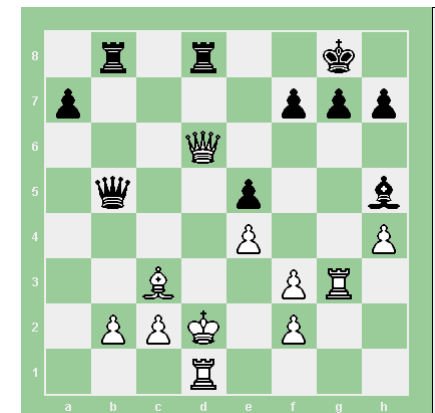
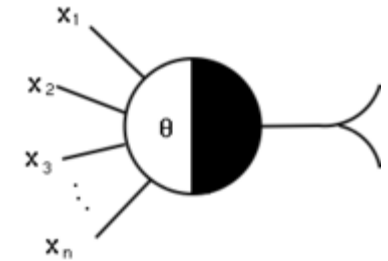
# A Brief History of AI

- 18<sup>th</sup> century:
  - von Kempelen's chess-playing Turk
  - amazing piece of mechanical engineering
  - played and won chess games all over the world (e.g., against Napoleon)
  - unfortunately a hoax...
- 19<sup>th</sup> century:
  - Charles Babbage's Analytical Engine
- 1920:
  - first use of word "robot" in Karel Capek's play *R.U.R (Rossum's Universal Robots)*
- 1940's:
  - Isaac Asimov – "Three Laws of Robotics"

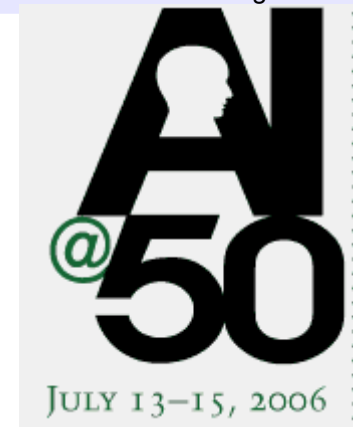


# A Brief History of AI

- 1943:
  - McCulloch and Pitts model artificial neurons
  - 1951: Marvin Minsky and Dann Edmonds constructed the first neural network computer
- 1950
  - Claude Shannon: algorithm for playing Chess
  - foresees strategies that are still used today
    - Shannon Type-A Strategy:
      - brute-force minimax search until a fixed horizon
    - Shannon Type-B Strategy
      - pruning uninteresting lines (as humans do)
      - preferred by Shannon (and contemporaries)
- 1951
  - Turing's chess algorithm
    - computation with paper and pencil
    - first recorded man-machine (chess) game
      - <http://www.chessgames.com/perl/chessgame?gid=1356927>



# The Dartmouth Conference



1956:

- John McCarthy invites 10 scientists with various backgrounds to a 2-week workshop at Dartmouth College
  - bringing together top minds on automata theory, neural nets and the study of intelligence.
  - For the next 20 years the field was dominated by these participants
    - John McCarthy, Herbert Simon, Allan Newell, Marvin Minsky, Arthur Samuel, etc.
  - Allen Newell and Herbert Simon (CMU): The Logic Theorist
    - first nonnumerical thinking program used for theorem proving
    - proved various theorems of Whitehead's *Principia Mathematica*
    - a joint publication by AN, HS, and LT was rejected...

→ Term “Artificial Intelligence” is coined

# Great Expectations (1952-1969)

- Newell and Simon (CMU): the General Problem Solver.
  - Imitation of human problem-solving
  - successfully solved simple puzzles
  - physical symbol system hypothesis
- Arthur Samuel (IBM, 1952-)
  - investigated game playing (checkers) with great success at IBM
    - program beat a regional master
  - pioneered many ideas in game playing and machine learning
    - including alpha-beta search, reinforcement learning, etc.
- John McCarthy (MIT, 1958-)
  - Inventor of Lisp (second-oldest high-level language)
  - Logic-oriented Advice Taker
    - separation between knowledge and reasoning
- Marvin Minsky (MIT)
  - various students working on micro-worlds (e.g., block's world)

# A Dose of Reality (1966-1973)

- Progress was slower than (unrealistic) expectations
  - Simon and Newell's (1958) predictions

1. That within ten years, a digital computer will be the world's chess champion, unless the rules bar it from competition.

2. That within ten years a digital computer will discover and prove an important new mathematical theorem.

3. That within ten years a digital computer will write music that will be accepted by critics as possessing considerable aesthetic value.

4. That within ten years most theories in psychology will take the form of computer programs, or of qualitative statements about the characteristics of computer programs.

. . . 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 the human mind has been applied.



# A Dose of Reality (1966-1973)

- Progress was slower than (unrealistic) expectations
  - Simon and Newell's (1958) predictions
  - came more or less true after 40 (instead of 10) years
    - but in very different ways than they had imagined
- Difficulties
  - Difficulty of knowledge representation
    - e.g., attempt for Machine Translation of Russian Scientific papers

the spirit is willing but the flesh is weak  
→ the vodka is good but the meat is rotten

- In 1966 no machine translations were used
- nowadays they are routinely used
  - but still give awful results (e.g., <http://babelfish.altavista.com/>)

# A Dose of Reality (1966-1973)

- Progress was slower than (unrealistic) expectations
  - Simon and Newell's (1958) predictions
  - came more or less true after 40 (instead of 10) years
    - but in very different ways than they had imagined
- Difficulties
  - Difficulty of knowledge representation
  - Lacked of scalability
    - under-estimation of the combinatorial explosion in search
    - things that work well in micro-worlds do not work in real world
      - e.g., theorem proving only worked with very few facts
    - Lighthill report (1973) focused on this issue
      - stopped AI funding in UK in all but two universities
  - Fundamental limitations on techniques and representations
    - Minsky and Papert (1969) noted that perceptrons are only linear separators
      - killed research in neural networks for decades

# Knowledge-Based Systems (1969-1979)

- DENDRAL project (Buchanan et al. 1969)
  - task:
    - infer molecular structure from formula of the molecule and the mass spectrum
  - First successful knowledge-intensive system
    - recognized importance of domain-specific knowledge
- Expert systems
  - MYCIN to diagnose blood infections (Feigenbaum et al.)
    - with ~450 rules, it performed better than junior doctors
    - knowledge had to be tediously acquired from experts (→ Knowledge Engineering Bottleneck)
    - introduction of uncertainty in reasoning
- Increase in knowledge representation research
  - Logic, frames, Schank's scripts, semantic nets, ...
  - CYC project (started by Lenat 1984)
    - attempt to encode common-sense knowledge

# The AI industry (1980-present)

- R1 (McDermott, 1982)
  - expert system for configuring computers at DEC
  - saved about \$40 million a year in 1986
- Fifth generation project in Japan (1981)
  - 10-year plan with strong focus on Logic Programming
  - did not quite live up to its ambitious goals
- Similar programs in US and U put an end to the AI winter
  - in UK, the Alvey report reinstated funding for “Intelligent Knowledge-Based Systems” (to avoid the name AI)
- Neural Networks revival
  - seminal work *Parallel Distributed Processing* by Rumelhart and McClelland (1986)
  - soon became popular in industrial applications
- AI industry grew from a few million dollars in 1980 to billions of dollars in 1988

# Current Trends in AI

- exploit the strengths of the computer
  - fast repetitious computations
  - elementary statistics (counting and probabilities)
  - and do not try to model human thought processes
    - brute-force methods perform much better in many areas
  - aviation was only possible when planes stopped to flap their wings...
- focus on particular tasks and not on solving AI as a whole
  - “Intelligent prostheses”
    - tools that support us in tasks that would otherwise require human intelligence
  - “AI-complete problems”
    - informal notion for very hard problems that cannot be solved unless the system has complete world knowledge
- strong scientific standards developed in the 1980s
  - solid experimentation and scientific verification of hypotheses

# The Science of AI

- Conference Series
  - 1969: Biennial International Joint Conference on AI (IJCAI)
  - 1980: Annual National Conference on AI (AAAI)
  - 1982: Biennial European Conference on AI (ECAI)
- Magazines
  - AI Magazine (published by AAAI)
  - IEEE Intelligent systems
- Journals
  - Artificial Intelligence (Elsevier)
  - Journal of Artificial Intelligence Research
    - pioneered free on-line publication (<http://www.jair.org>)
- Since the 1980s various subfields emerged, joined forces with related fields
  - many journals and annual conferences in subareas

# State of the Art

- Autonomous Planning and Scheduling
  - NASA's on-board autonomous planning program controlled the scheduling of operations for a spacecraft
  - During the 1991 Gulf War, US forces deployed an AI logistics planning and scheduling program
    - DARPA: this application payed back 30 years of investment in AI
- Game Playing
  - TD-Gammon learned an evaluation function for backgammon that led to changes in backgammon theory
  - Deep Blue defeated the reigning world chess champion Garry Kasparov in 1997
    - increased IBM stocks by several billion dollars
- Natural Language Understanding
  - Proverb solves crossword puzzles better than most humans
    - although it does not “understand” the clues
  - translation systems are frequently used (unfortunately...)

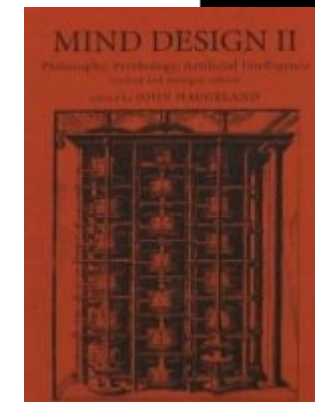
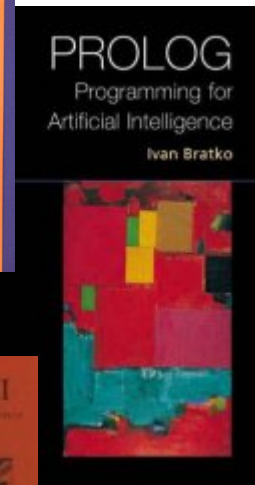
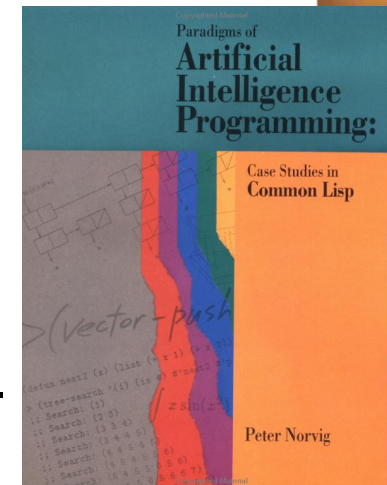
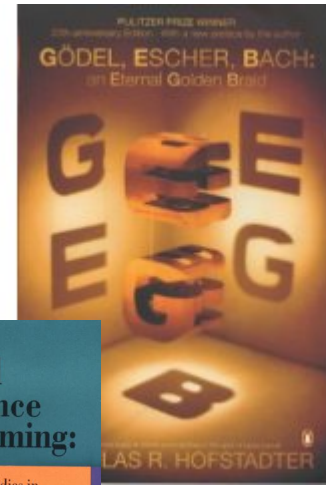
# State of the Art

- Robotics
  - No hands across America
    - CMU's ALVINN drives autonomously 98% of the time from Pittsburgh to San Diego
    - 2% human intervention mostly on exit ramps etc.
  - DARPA Grand Challenge
    - 2005 off-road race for autonomous vehicles
    - 2007 Urban Challenge
  - robot assistants are routinely used in microsurgery
- Scientific Discovery
  - machine learning system helped to discover new quasars
  - automated theorem prover proved Robbins conjecture unsolved for decades
  - Data Mining and Knowledge Discovery is a developing new industry



# Recommended Books

- Douglas R. Hofstadter, *Gödel, Escher, Bach*, Basic Books, New York 1979.  
→ classic, inspiring, readable, original intro into AI
- Peter Norvig, *Paradigms of Artificial Intelligence Programming: Case Studies in Common LISP*, Morgan Kaufmann, 2000.  
→ AI techniques in LISP
- Ivan Bratko, *Prolog Programming for AI*, Addison-Wesley, 3<sup>rd</sup> edition, 2000.  
→ AI techniques in Prolog
- Haugeland, John (ed.), *Mind Design II: Philosophy, Psychology, and Artificial Intelligence*, Cambridge, Mass.: MIT Press, 1981.  
→ collection of classic Philosophical articles
- <http://www.aaai.org/AITopics/>  
→ Web site of the Association for the Advancement of Artificial Intelligence



# Other AI-relevant Course Topics @ TUD

- Machine Learning (Fürnkranz, Schiele)
- Robotics (van Stryk)
- Search and Optimization (Weihe)
- Natural Language Processing (Gurevych)
- Fuzzy-Logic and Genetic Algorithms (Adamy)
- Vision and Perception (Schiele, Fraunhofer IGD)
- Digital Storytelling (Göbel)
- Digitale Spiele (Jantke)
- ...