ARTIFICIAL INTELLIGENCE

344.014
VO 2 h, WS 2015/16

Univ.-Prof. Dr. Gerhard Widmer
Institute for Computational Perception
Johannes Kepler University Linz

gerhard.widmer@jku.at
http://www.cp.jku.at/people/widmer

Overview

The Institute

Administrative Issues

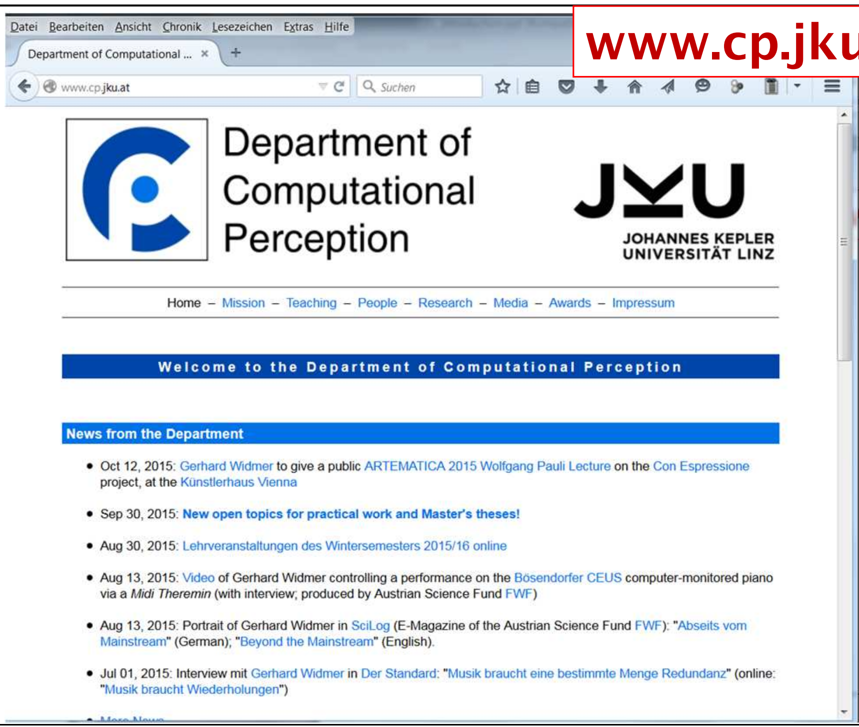
The Exercise Track (Übung)

Literature

Introduction: What is Artificial Intelligence?

Our Guiding Scenario: Intelligent Agents

www.cp.jku.at



Administrative Issues: Lecture

Time and Date of Lectures: Mo, 12:00-13:30, HS 19

Materials:

- **pdf versions** of Powerpoint slides (via KUSSS)

  ➔ signed up for this class in KUSSS?
  ➔ if not: send us an e-mail (claudia.kindermann@jku.at)

Exam: written, Jan. 25, 2016, 12:00-13:30

Administrative Issues: Exercise Track (Übung)

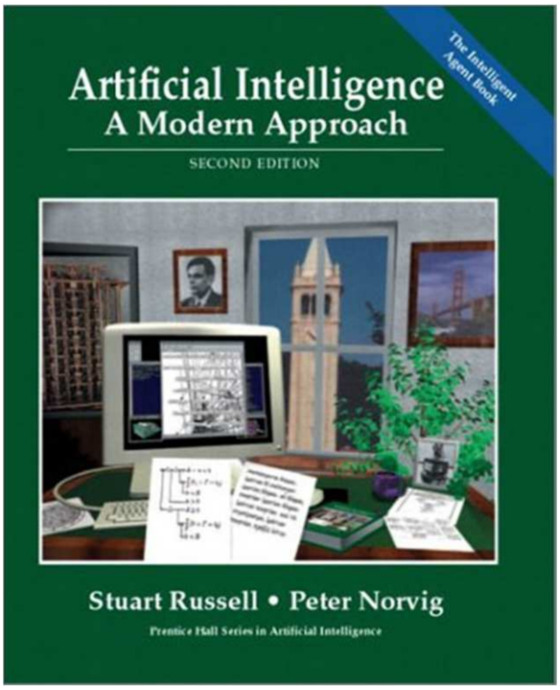
3 Groups (344.021, 344.022, 344.023)

Organised by Filip Korzeniowski & Rainer Kelz

Time:
Mo, 13:45 – 14:30 (all groups)

Place:
HS 6

First meeting (with explanation of procedures):
Mo, Oct. 5



Many of the slides presented in this class are based on:

Artificial Intelligence: A Modern Approach (2nd ed.),
by Stuart Russell & Peter Norvig, Prentice Hall, 2000

Lecture Plan WS 2015/16 (preliminary)

5.10.2015: INTRODUCTION: History of AI; the concept of intelligent agents
12.10.2015: SEARCH 1: Uninformed search
19.10.2015: SEARCH 2: Heuristic search
26.11.2015, 2.11.2015: --- holidays ☹ ---
9.11.2015: SEARCH 3: Game playing search
16.11.2015: INFERENCE 1: Propositional logic; proofs and reasoning
23.11.2015: INFERENCE 2: Propositional logic: resolution & chaining; basics of first-order logic
30.11.2015: INFERENCE 3: Unification; first-order resolution and chaining; logic as a programming language: Prolog
7.12.2015: UNCERTAINTY 1: Basic concepts of probability theory
14.12.2015: UNCERTAINTY 2: Probabilistic queries; independence, Bayes Nets
11.1.2016: UNCERTAINTY 3: Constructing BNs; exact & approximate inference
18.1.2016: LEARNING 1: Learning action strategies: reinforcement learning
25.1.2016: Final Exam (written)

(no time for:) PERCEPTION: Basic concepts of computer perception; examples

Why Artificial Intelligence?

Theoretical Goal:

Understanding the phenomenon of intelligence by building computational models of intelligent behaviours
➔ connection to *Cognitive Sciences*

Practical Goal:

Building "intelligent" machines that are
- flexible
- autonomous
- adaptive and capable of learning
- … useful

What is Intelligence?

en.wikipedia.org:

“Intelligence is a general mental capability that involves the ability to reason, plan, solve problems, think abstractly, comprehend ideas and language, and learn.”

de.wikipedia.org:

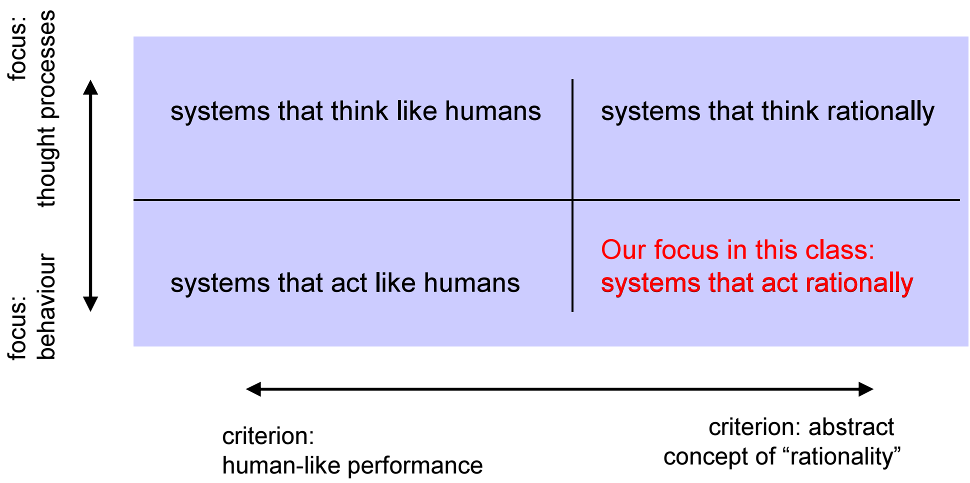
„Intelligenz (lat.: intelligentia = Einsicht, Erkenntnisvermögen) bezeichnet im weitesten Sinne die Fähigkeit von Lebewesen, aber auch von technischen Geräten, sich durch Prozesse der Informationsverarbeitung an die Umweltgegebenheiten anzupassen.“

Popular wisdom:

“Intelligence is that which computers cannot do (yet).”

Definitions of Artificial Intelligence (AI)

4 categories of definitions:



Definitions of Artificial Intelligence (AI)

systems that think like humans

"The exciting new effort to make computers think … machines with minds, in the full and literal sense" *(Haugeland, 1985)*

"[The automation of] activities that we associate with human thinking, such as decision making, problem solving, learning …" *(Bellman, 1978)*

systems that think rationally

"The study of mental faculties through the use of computational methods" *(Charniak & McDermott, 1985)*

"The study of computations that make it possible to perceive, reason, and act" *(Winston, 1992)*

systems that act like humans

"The art of creating machines that perform functions that require intelligence when performed by people" *(Kurzweil, 1990)*

"The study of how to make computers do things at which, at the moment, people are better" *(Rich & Knight, 1991)*

systems that act rationally

"Computational Intelligence is the study of the design of intelligent agents" *(Poole et al., 1998)*

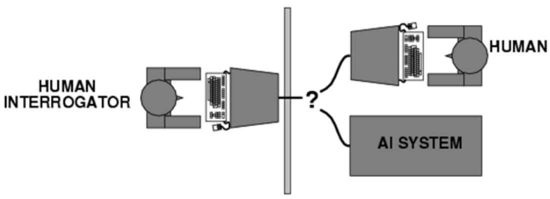
"AI … is concerned with intelligent behavior in artifacts" *(Nilsson, 1998)*

"Acting Humanly": The Turing Test

Alan Turing (1912-1954):

- Article "Computing Machinery and Intelligence" (1950):
  "Can a machine think?" ➔ untestable.
  Instead: ➔ *"Can a machine behave in such as way as to fool humans?"*

- Operational test for intelligent behaviour: The ***Turing Test***



1966: ELIZA (Joseph Weizenbaum, MIT)

A “Computer Psychotherapist” …

ELIZA: How do you do. Please state your problem.
>> I spend too much time with computers.
E: Why do you mention computers ?
>> You are one.
E: Does it please you to believe I am one ?
>> Yes.
E: You seem to be quite positive.
>> I suppose I am.
E: How long have you been ?
>> 31 years.
E: I'm not sure I understand you fully.
>> Join the club.
E: Please go on.
>> You don't make too much sense yourself.
E: We were discussing you -- not me.

Weizenbaum tells us that he was shocked by the experience of releasing ELIZA (also known as "Doctor") to the nontechnical staff at the MIT AI Lab. Secretaries and nontechnical administrative staff thought the machine was a "real" therapist, and spent hours revealing their personal problems to the program. When Weizenbaum informed his secretary that he, of course, had access to the logs of all the conversations, she reacted with outrage at this invasion of her privacy. Weizenbaum was shocked by this and similar incidents to find that such a simple program could so easily deceive a naive user into revealing personal information.

Home Page of The Loebner Prize--"The First Turing Test"

Loebner Prize Gold Medal
(Solid 18 carat, not *gold-plated* like the Olympic *"Gold"* medals)

What is the Loebner Prize?

The Loebner Prize is the first formal instantiation of a Turing Test. The test is named after Alan Turing the brilliant British mathematician. Among his many accomplishments was basic research in computing science. In 1950, in the article *Computing Machinery and Intelligence* which appeared in the philosophical journal Mind, Alan Turing asked the question "Can a Machine Think?"He answered in the affirmative, but a central question was: "If a computer could think, how could we tell?" Turing's suggestion was, that if the responses from the computer were indistinguishable from that of a human,the computer could be said to be thinking.

In 1990 Hugh Loebner agreed with The Cambridge Center for Behavioral Studies to underwrite a contest designed to implement the Turing Test. Dr. Loebner pledged a Grand Prize of $100,000 and a Gold Medal (pictured above) for the first computer whose responses were indistinguishable from a human's. Each year an annual prize of $2000 and a bronze medal is awarded to the **most** human computer. The winner of the annual contest is the best entry relative to other entries that year, irrespective of how good it is in an absolute sense.

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

Our Focus:
"Acting Rationally": Rational Agents

Intelligence as the ability to behave "rationally" in a given environment, or relative to a given task:

- Rational behavior: "doing the right thing" (relative to one's goals, abilities, resources, risks, …)
- "the right thing" = that which is expected to maximise the achievement of goals, given the available information
- Rational behaviour does not necessarily involve thinking (e.g., reflexes …), but thinking should serve to produce rational action
- Rational behaviour or rational decisions are often limited by limited resources (time, incomplete information, …)

Artificial Intelligence is a Multi-disciplinary Field

Philosophy	logic, methods of reasoning, mind as a physical system; foundations of learning, language, rationality
Mathematics	formal representation and proof algorithms, computation, (un)decidability, (in)tractability, probability
Economics	utility theory; decision theory (➔ Herbert Simon …)
Neuroscience	physical substrate for mental activity
Psychology	studies on perception, memory, motor control, …; experimental techniques
Linguistics	knowledge representation, grammars, language understanding
Computer Science	algorithms, complexity theory, programming languages, …

The Birth of “ARTIFICIAL INTELLIGENCE”

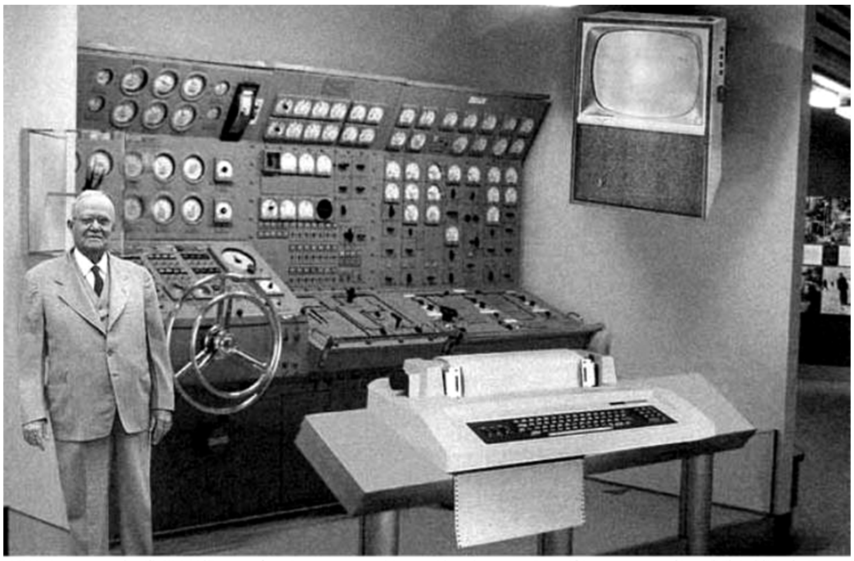
1956: Dartmouth Conference

„We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire.

The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.“

- Dartmouth AI Project Proposal; John McCarthy et al.; Aug. 31, 1955.

1956: The early days of computing ...



Scientists from the RAND Corporation have created this model to illustrate how a "home computer" could look like in the year 2004. However the needed technology will not be economically feasible for the average home. Also the scientists readily admit that the computer will require not yet invented technology to actually work, but 50 years from now scientific progress is expected to solve these problems. With teletype interface and the Fortran language, the computer will be easy to use.

 http://www.museumofhoaxes.com/hoax/photo_database/image/home_computer_of_the_future

2005: The DARPA Grand Challenge: Autonomous Driving Vehicles

The rules of the DARPA Grand Challenge were simple. Contestants were required to build autonomous ground vehicles capable of traversing a desert course up to 175 miles long in less than 10 hours. The first robot to complete the course in under 10 hours would win the challenge and the $2M prize. Absolutely no manual intervention was allowed. The robots were started by DARPA personnel and from that point on had to drive themselves. Teams only saw their robots at the starting line and, with luck, at the finish line. The race was held in the Mojave desert in the southwest United States. Course terrain varied from high quality, graded dirt roads to winding, rocky, mountain passes. The specific race course was kept secret from all teams until two hours before the race. At this time, each team was given a description of the course on CD-ROM in a DARPA-defined Route Definition Data Format (RDDF). The RDDF is a list of longitudes, latitudes, and corridor widths that define the course boundary, and a list of associated speed limits. Speed limits were used to protect important infrastructure and ecology along the course and to maintain the safety of DARPA chase drivers who followed behind each robot. The speed limits varied between 5 and 50 mph. The robots all competed on the same course, starting one after another at 5 minute intervals. When a faster robot overtook a slower one, the slower robot was paused by DARPA officials, allowing the second robot to pass the first as if it were a static obstacle. This eliminated the need for robots to handle the case of dynamic passing.

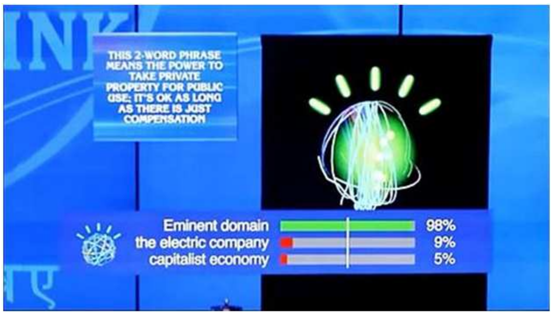
2005:
Stanley is the first car to completely autonomously drive 175 miles through difficult terrain

➔ DARPA_Stanley_Raceday.wmv

2015:
From Stanley to Google's Self-driving Car

Stanley (Stanford Univ., 2005)

2011: "Watson" Wins the U.S. Game Show "Jeopardy!"



Watson is an artificial intelligence computer system capable of answering questions posed in natural language, developed in IBM's DeepQA project by a research team led by principal investigator David Ferrucci. Watson was named after IBM's first president, Thomas J. Watson. The machine was specifically developed to answer questions on the quiz show *Jeopardy!*. In 2011, Watson competed on Jeopardy against former winners Brad Rutter and Ken Jennings. Watson received the first prize of $1 million.

➔ *Video*
https://www.youtube.com/watch?v=seNkjYyG3gI

http://en.wikipedia.org/wiki/Watson_(computer)

Watson in action: ➔ https://www.youtube.com/watch?v=qpKoIfTukrA

2011: "Watson" Wins the U.S. Game Show "Jeopardy!"

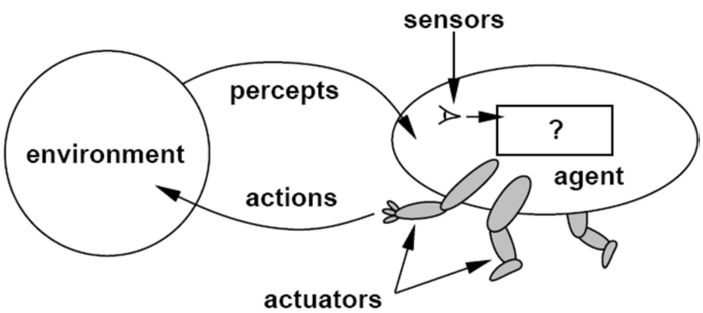
Software:
Watson's software was written in various languages, including at least Java, C++, and Prolog and uses Apache Hadoop framework for distributed computing, Apache UIMA (Unstructured Information Management Architecture) framework, IBM's DeepQA software and SUSE Linux Enterprise Server 11 operating system.[8][17][18] "[...] more than 100 different techniques are used to analyze natural language, identify sources, find and generate hypotheses, find and score evidence, and merge and rank hypotheses."

Hardware:
Watson is made up of a cluster of 90 IBM Power 750 servers (plus additional I/O, network and cluster controller nodes in 10 racks) with a total of 2880 POWER7 processor cores and 16 Terabytes of RAM. Each Power 750 server uses a 3.5 GHz POWER7 eight core processor, with four threads per core.

Information Sources:
The IBM team provided Watson with millions of documents, including dictionaries, encyclopedias, and other reference material that it could use to build its knowledge. Although Watson was not connected to the Internet during the game, it contained 200 million pages of structured and unstructured content consuming four terabytes of disk storage, including the full text of Wikipedia.

http://en.wikipedia.org/wiki/Watson_(computer)

Our Guiding Scenario: Intelligent Agents



Intelligent Agents

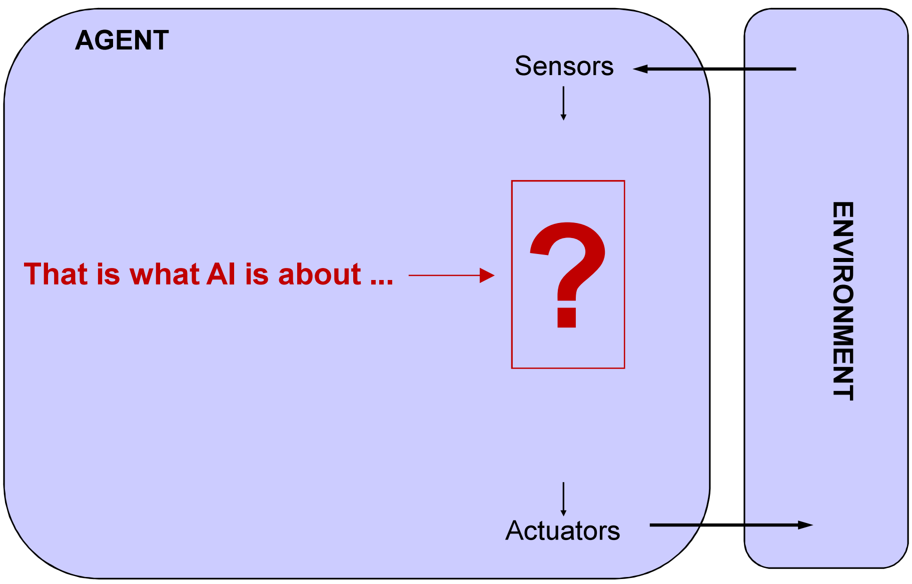
An **agent** is any system that can be viewed as
- perceiving its environment through sensors and
- acting upon that environment through actuators

Human agent:
- eyes, ears, and other organs for sensors;
- hands, legs, mouth, and other body parts for actuators

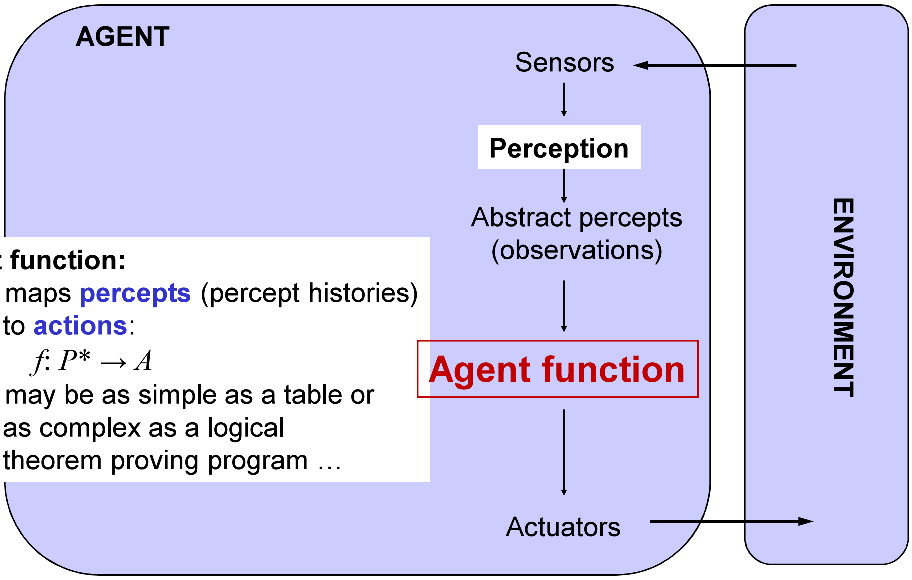
Robotic agent:
- cameras, infrared range finders, etc. for sensors;
- various motors for actuators

Software agent ('softbot'):
- bit strings for sensors
- bit strings for actuators

Intelligent Agents

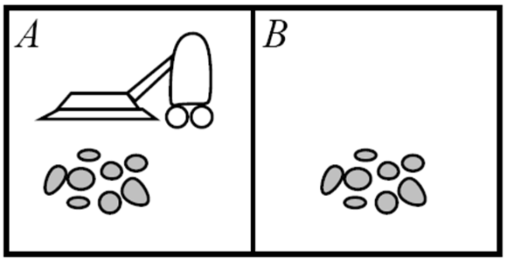


Intelligent Agents



Agent function:
- maps **percepts** (percept histories) to **actions**:
  $f: P^* \rightarrow A$
- may be as simple as a table or as complex as a logical theorem proving program …

The "Vacuum Cleaner World"



"World": two rooms A, B that may or may not contain dirt
Agent: a vacuum cleaner
Percepts (observations): current location (room) and contents (e.g., [A,dirty])
Actions: GoLeft, GoRight, Suck, NoOp
Goal: world should be clean

A Simplistic Approach: Reflex Agents

"direct, immediate reaction to observations"

Simplest Solution:

Agent function is a table that specifies action to be taken for any possible history of percepts

Percept sequence	Action
$[A, Clean]$	$Right$
$[A, Dirty]$	$Suck$
$[B, Clean]$	$Left$
$[B, Dirty]$	$Suck$
$[A, Clean], [A, Clean]$	$Right$
$[A, Clean], [A, Dirty]$	$Suck$
$\vdots$	$\vdots$

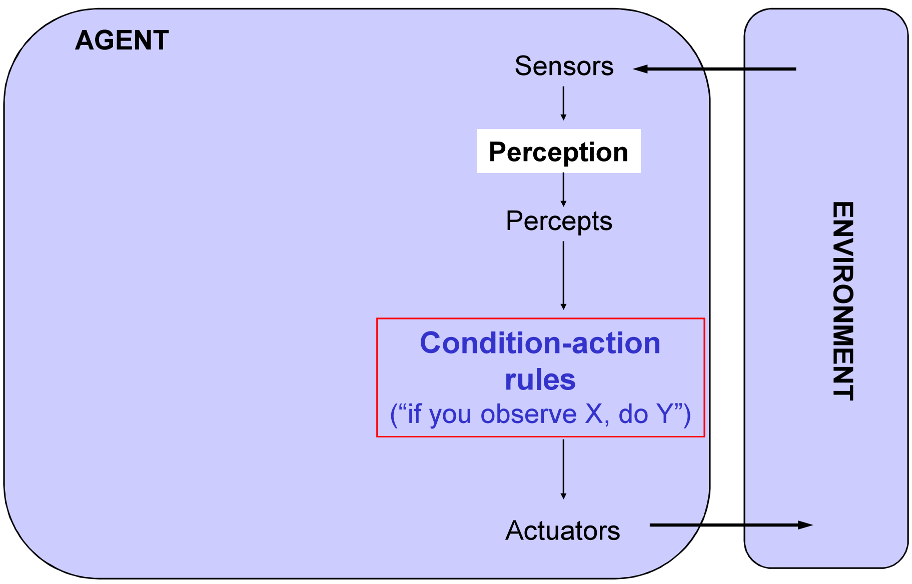
Note that these two are not equivalent! (e.g., what should the program do when A and B have been visited and are clean? …) ➔ Need for a MEMORY!

Representation as if-then-else program:

function REFLEX-VACUUM-AGENT([$location$, $status$]) **returns** an action

  if $status = Dirty$ **then return** $Suck$
  else if $location = A$ **then return** $Right$
  else if $location = B$ **then return** $Left$

Reflex Agents



Reflex Agents

Problems:

- Need to explicitly define / store "correct" action for each possible sequence of percepts

  (Consider, e.g., chess playing: 35 possible moves per situation, game may be 100 moves long ➔ 35^{100} table entries!)
- We (the system designer) may not *know* what the correct action is
- No flexibility (what if the world changes?)
- *Learning agent* would take forever to build up the entire table from observations / experiences

The AI Approach: Knowledge-based Agents

General AI approach:

- Specify an agent's **knowledge** (about actions, their effects, etc.) **explicitly** (rather than hardwiring decisions into a program)
- Design **general reasoning algorithms** to make decision based on this

  ➔ separation of knowledge and decision making („inference")

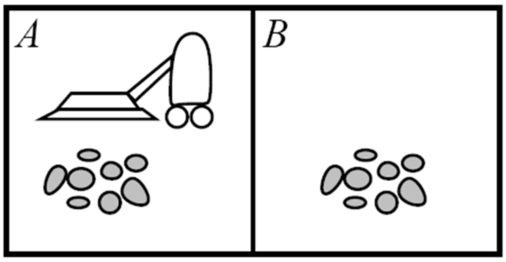
  ➔ knowledge can be easily changed without need for changing entire program (e.g., new rooms in vacuum cleaner world)

  ➔ agent can explain its decisions!

Informatics view:
- Problem solving = execution of algorithms

AI view:
- Problem solving = knowledge + inference (algorithms)

Explicitly Representing Knowledge about the "Vacuum Cleaner World"



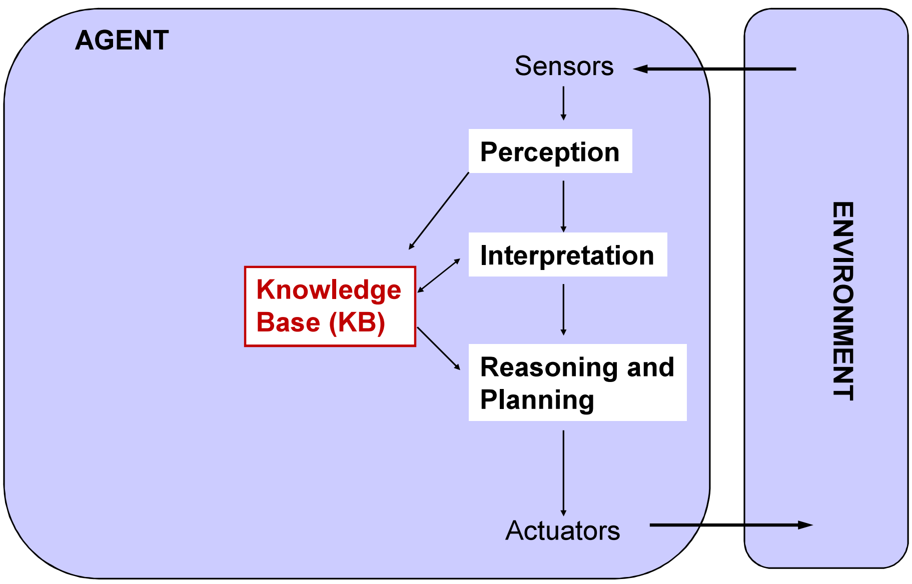
knowledge about …

states
Room(A). Room(B). *% There are two rooms called A and B.*
Contains(A,Dirt). Contains(B,Dirt). *% Both contain dirt.*
in(Agent,A). *% The agent is currently in room A.*

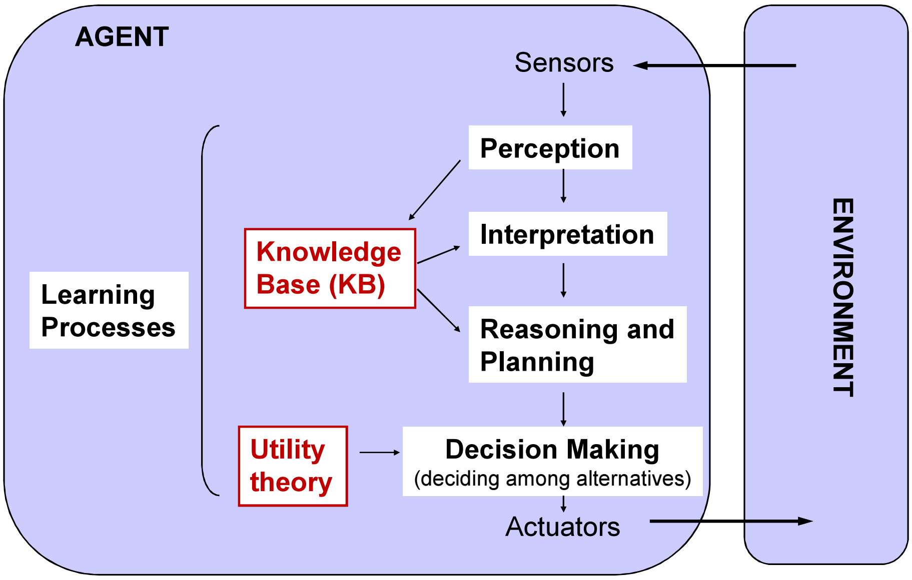
actions
in(Agent,x) AND suck ➔ clean(x). *% If the agent is in any room x*
% and it performs the operation suck,
% then the room x will be clean.
in(Agent,A) AND move_right ➔ in(Agent,B).
in(Agent,B) AND move_left ➔ in(Agent,A).

goals
GOAL: clean(A) AND clean(B).

Knowledge-based Agents



Knowledge-based Agents: The Full View



A Real Knowledge-based Agent: An Inside View of *Stanley* ...

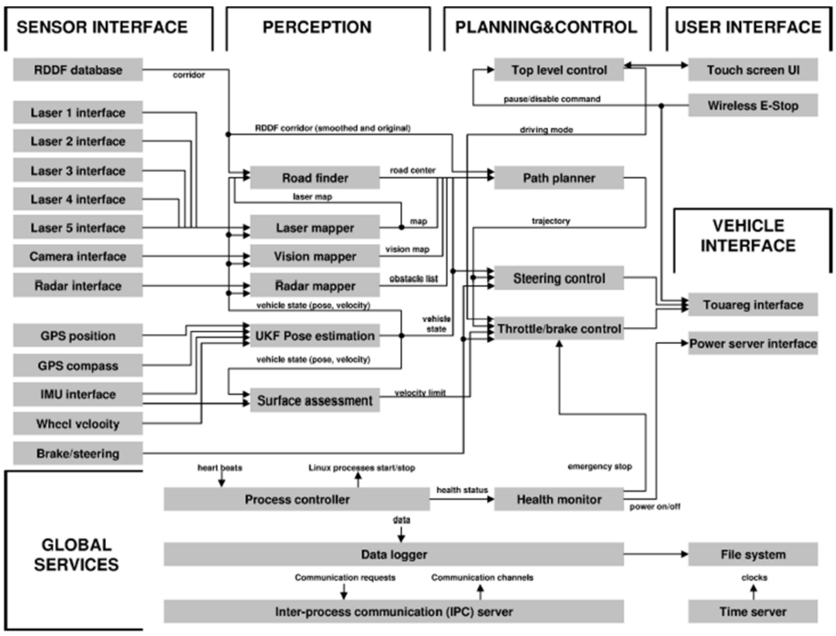


Figure 5: Flowchart of Stanley Software System. The software is roughly divided into six main functional groups: sensor interface, perception, control, vehicle interface, and user interface. There are a number of cross-cutting services, such as the process controller and the logging modules.

Abilities Required for an Agent to be "Intelligent" ➔ *Main Research Areas in AI*

- Interpreting and "understanding" sensorial or other inputs
  ➔ **perception**

in this class ...

- Collecting and maintaining knowledge about the world
  ➔ **knowledge representation**
- Drawing conclusions: inferring ("schließen") new things from what it knows
  ➔ **(logical or probabilistic) reasoning**
- Making predictions and plans based on knowledge, observations, and goals
  ➔ **searching and planning**
- Making decisions under various constraints (resources/time, utility, uncertainty, …)
  ➔ **decision theory**
- Updating and increasing its knowledge
  ➔ **learning**
- Communication with its environment or other agents (including humans)
  ➔ **language understanding, communication**

Topics of this Class

Focus on very basic methods, mostly in discrete (symbolic) domains:

- solving problems by making plans (→ search)
- representing "knowledge" (→ logic)
- drawing conclusions from knowledge (→ reasoning, inference)
- reasoning with uncertain information (→ probability theory)
- acquiring new knowledge (→ learning)
- detecting patterns in complex inputs (→ perception)

Not dealt with in this class (because of time constraints):

- interpreting language(s) (→ computational linguistics, grammars)
- perception-action coupling, robotics
- continuous (numerical) worlds
- real-time processing and time constraints
- neural networks, genetic algorithms, fuzzy logic
- psychological or neurological theories of cognition
- collective intelligence, multi-agent models
- current "hype topics" (emotional agents, …) …

Remember: AI is a huge research field with 50 years of research and hundreds of thousands of published scientific articles ...