

Künstliche Intelligenz

Vorlesung 1: Einführung in die Künstliche Intelligenz



ORGANISATORISCHES

Literatur

- W. Ertel, Grundkurs Künstliche Intelligenz
- S. Russel, P. Norvig, Artificial Intelligence, A Modern Approach

Sprechstunden

Nach Vereinbarung



CREDITS

Nach Slides von W. Ertel, D. Sabel, L. Diosan



WAS IST KÜNSTLICHE INTELLIGENZ (KI)

Die **künstliche Intelligenz** ist die Wissenschaft, Maschinen dazu zu bringen, Dinge zu tun, die Intelligenz erfordern, wenn sie von Menschen ausgeführt werden.

Marvin Minsky, 1967

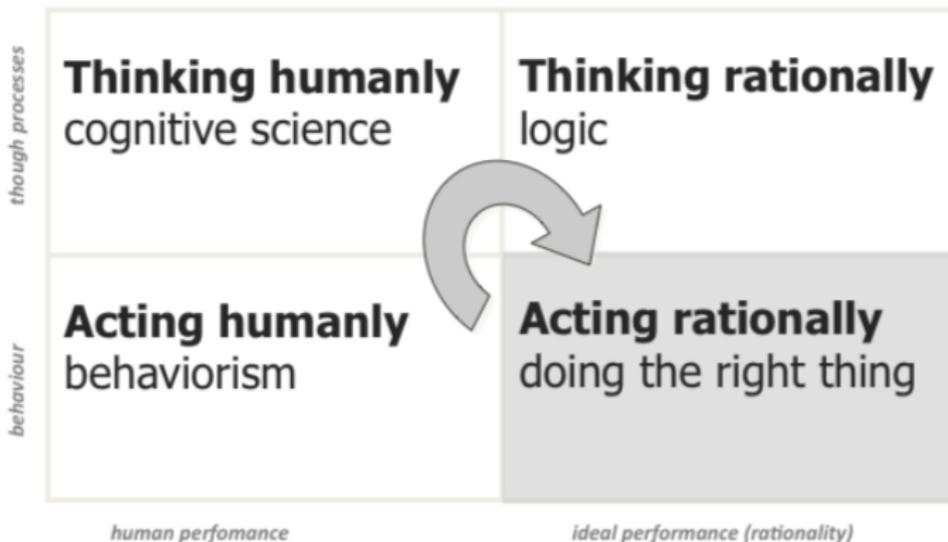


WAS IST KÜNSTLICHE INTELLIGENZ (KI)

- Was ist Intelligenz?
- Wie kann man Intelligenz messen?
- Wie funktioniert unser Gehirn?
- Intelligente Maschine?
- Science Fiction?
- Menschlichen Geist nachbauen?
- Philosophie? z.B. Körper-Seele Problem



VIER PERSPEKTIVEN ÜBER DIE KI



MENSCHLICH HANDELN

ACTING HUMANLY

Alan Turing (1950):

- Können Maschinen denken?
- Können Maschinen wie Menschen handeln?

Turing Test

Ein Computer besteht den Test, wenn ein Mensch, der einige aufgeschriebene Fragen stellt, nicht feststellen kann, ob die geschriebenen Antworten von einem Menschen oder von einem Computer kommen.



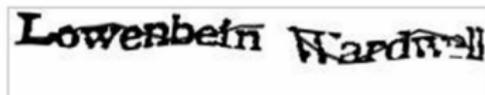
REVERSE TURING TEST

Der Computer versucht zu erkennen, ob er mit einem Computer oder einer Person kommuniziert.

Security Check

Enter **both words** below, **separated by a space**.

Can't read the words below? Try different words or an audio captcha.



Sick of these? [Verify your account](#).

Text in the box:

What's This?

Submit

Cancel



Kognitive Modellierung

- Modellierung des menschlichen Denkens
- Wir müssen bestimmen können, wie Menschen denken
 - Top-down-Ansatz (Psychologie)
 - menschliche Denkschritte folgen (z.B. gefunden durch Introspektion oder durch Beobachten der Handlungen einer Person)
 - GPS: Allgemeiner Problemlöser (General Problem Solver, Newell und Simon, 1957)
 - Bottom-up (Neurowissenschaft)
 - Modellierung des Gehirns (durch Beobachtung)
 - Verbindungsmodelle
 - Intelligentes Verhalten entsteht durch die Verbindung einer großen Anzahl einfacher Einheiten (Neuronale Netze)

RATIONALES DENKEN

THINKING RATIONALLY

Seit der Zeit von Aristoteles (384 - 322 v. Chr.) versuchten die Menschen, **richtiges Denken** zu formalisieren.

■ Syllogismen

- Muster für Argumentationen, die immer korrekt sind falls korrekte Prämissen vorhanden sind.
- Wenn Sokrates ein Mann ist, sind alle Menschen sterblich
⇒ Sokrates ist sterblich

■ Diese Studie initiierte das Gebiet der Logik (und Mathematik).



RATIONALES DENKEN

THINKING RATIONALLY

Haupthindernisse:

- Es ist nicht einfach, informelles Wissen zu nehmen und es in den von der logischen Notation geforderten formalen Bedingungen zu formulieren, insbesondere wenn das Wissen weniger als 100% sicher ist.
- Es besteht ein großer Unterschied, ob man ein Problem **im Prinzip** löst oder, ob man es in der Praxis löst.



RATIONALES HANDELN

ACTING RATIONALLY

Rationales Verhalten = *Richtige Dinge* tun

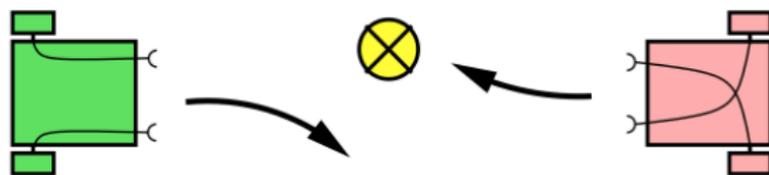
- *Richtig* = das Beste (erwartete) zu erreichen auch wenn es eine gewisse Unsicherheit gibt
- Korrekte Schlüsse zu ziehen (rationales Denken) ist Teil eines **rationalen Agenten**, aber nicht exklusiv.
 - In manchen Situationen gibt es keine beweisbare richtige Handlungsweise, es muss trotzdem gehandelt werden.
 - Es gibt auch rationales Handeln, von dem nicht behauptet werden kann, dass es logische Schlüsse beinhaltet (z. B. Reflexhandlungen).
 - Diese Lehrveranstaltung konzentriert sich auf allgemeine Prinzipien von rationalen Agenten und auf Komponenten für deren Konstruktion.



WAS IST KÜNSTLICHE INTELLIGENZ (KI)

John McCarthy (1955):

Ziel der KI ist es, Maschinen zu entwickeln, die sich verhalten, als verfügten sie über Intelligenz.



Zwei ganz einfache Braitenberg-Vehikel und deren Reaktion auf eine Lichtquelle.

WAS IST KÜNSTLICHE INTELLIGENZ (KI)

Encyclopedia Britannica:

KI ist die Fähigkeit digitaler Computer oder computergesteuerter Roboter, Aufgaben zu lösen, die normalerweise mit den höheren intellektuellen Verarbeitungsfähigkeiten von Menschen in Verbindung gebracht werden . . .

Nach dieser Definition ist also jeder Computer ein KI-System.



WAS IST KI?

Elaine Rich

Artificial Intelligence is the study of how to make computers do things at which, at the moment, people are better.

- Auch im Jahr 2050 noch aktuell!
- Mensch heute noch besser in vielen Bereichen!
u.a. Bildverstehen, Lernfähigkeit
- Computer heute schon besser in vielen Bereichen!
- Beisp. Schachcomputer/Go



Unterschiedliche Ansätze:

- Wie arbeitet das menschliche Gehirn?
- Problemorientiert: Intelligente Agenten bauen!
- Gemischtwarenladen!

KI-VERFAHREN



WOMIT BESCHÄFTIGT SICH DIE KÜNSTLICHE INTELLIGENZ?

Ziel der Künstlichen Intelligenz:

Herstellung eines intelligenten Agenten

Auch: Herstellung eines möglichst guten

- autonomen
- lernenden
- intelligenten
- automatischen

Informationssysteme



BEISPIELE

- Taschenrechner
- Schachspielende Computer,
z.B. [Deep Blue](#), Deep Thought und Deep Fritz
- Go und KI ([hier](#) und [hier](#))
- Sprachübersetzer wie z.B. GoogleTranslate, Babelfish, etc.
- Wissensbasierte Systeme in vielen Varianten
- Texterkennungssysteme in Kombinationen mit
Hintergrundwissen wie z.B. [IBM Watson](#)
([Jeopardy-Gewinner](#))
- Roboter, z.B. Haushaltsroboter wie Staubsaugerroboter,
Industrieroboter, etc.
- intelligente Informationssysteme



GRUNDLAGEN DER KI

Die KI bezieht sich auf Ideen und Methoden aus vielen Bereichen.

- Philosophie (400 v.C. –) – Wie entstehen die Gedanken im Gehirn? **Logik, Schlussverfahren**
- Mathematik (800 –) – Welche **formale** Regeln werden benutzt um gültige Schlüsse zu ziehen? Was ist berechenbar?
- Wirtschaft (1776 –) – Wie kann man Gewinne maximieren? **Utility Theory, Decision Making**
- Neurowissenschaft (1861 –) – Wie verarbeitet unser Gehirn die Information?
- Psychologie (1879 –) – Wie denken und handeln Menschen? **Behaviourism**



GRUNDLAGEN DER KI

- Computer Engineering (1940 –) – Wie baut man einen Computer?
- Kontroll Theorie (1948 –) – Dynamische Systeme, deren Verhalten durch sogenannte **Eingangsgrößen** von außen beeinflusst werden kann
- Sprachwissenschaften (1957 –) – Welche Beziehungen existieren zwischen Sprache und Denken? **Knowledge Representation**



GESCHICHTE DER KI

- 1931: Der österreichischer Kurt Gödel zeigt, dass in der Prädikatenlogik erster Stufe alle wahren Aussagen herleitbar sind. In Logiken höherer Stufe hingegen gibt es wahre Aussagen, die nicht beweisbar sind .
- 1937: Alan Turing zeigt mit dem Halteproblem Grenzen intelligenter Maschinen auf .
- 1943: McCulloch und Pitts modellieren Neuronale Netze und stellen die Verbindung zur Aussagenlogik her.
- 1950: Alan Turing definiert über den Turingtest Intelligenz von Maschinen und schreibt über lernende Maschinen und genetische Algorithmen
- 1951: Marvin Minsky entwickelt einen Neuronenrechner. Mit 3000 Röhren simuliert er 40 Neuronen.



GESCHICHTE DER KI

- 1955: Arthur Samuel (IBM) baut lernfähige Dame Programme die besser spielen als ihre Entwickler
- 1956 McCarthy organisiert eine Konferenz im Dartmouth College. Hier wird der Name Artificial Intelligence festgelegt.
Newell und Simon von der Carnegie Mellon University (CMU) stellen den Logic Theorist, das erste symbolverarbeitende Programm, vor.
- 1958: McCarthy erfindet am MIT (Massachusetts Institute of Technology) die Hochsprache LISP. Er schreibt Programme, die sich selbst verändern können.
- 1959: Gelernter (IBM) baut den Geometry Theorem Prover.
- 1961 Der General Problem Solver (GPS) von Newell und Simon imitiert menschliches Denken .
- 1963 McCarthy gründet das AI-Lab an der Stanford Universität.



GESCHICHTE DER KI

- 1965: Robinson erfindet den Resolutionskalkül für Prädikatenlogik.
- 1966: Weizenbaum's Eliza-Programm führt Dialoge mit Menschen in natürlicher Sprache.
- 1969: Minsky und Papert zeigen in ihrem Buch Perceptrons auf, dass das Perzeptron, ein sehr einfaches neuronales Netz, nur lineare Zusammenhänge repräsentieren kann.
- 1952 - 1969: Die Zeit der großen Erwartungen...
- J. McCarthy: Look, Ma, no hands! Ära.



GESCHICHTE DER KI

EINE DOSIS WIRKLICHKEIT (1966 -1973)

- 1972: Der Franzose Alain Colmerauer erfindet die Logikprogrammiersprache PROLOG.
- Es existieren Maschinen die denken können, lernen können und schaffen können, aber nur einfache Aufgaben meistern. Weshalb?
- Frühere Programme wussten nichts über das Thema, welches sie bearbeitet haben. Der Erfolg entsteht über einfache syntaktische Manipulationen
- Viele KI Probleme sind mit Computer schwer lösbar
- Grundlegende Einschränkungen der Basisstrukturen (z.B. der Perzeptron)



GESCHICHTE DER KI

KNOWLEDGE BASED SYSTEMS (1969 -1979)

- **Domain specific language** wird benutzt Der britische Mediziner de Dombal entwickelt ein Expertensystem zur Diagnose von Bauchkrankheiten . Es blieb in der bis dahin überwiegend amerikanischen KI-Community unbeachtet.
- 1976: Shortliffe und Buchanan entwickeln MYCIN, ein Experten- system zur Diagnose von Infektionskrankheiten, das mit Unsicherheit umgehen kann.
- DENDRAL (Buchanan): inferriert über molekulare Strukturen ausgehend von Informationen, die von einem Masse Spektrometer geliefert werden. Regeln, basiert auf bekannten Mustern werden eingeführt, um die Anzahl möglicher Strukturen zu verringern.
- FRAMES (Minsky, 1957) - Motivation für die heutigen OOP Sprachen



GESCHICHTE DER KI

KI WIRD ZU EINER INDUSTRIE

- 1981: Japan startet mit großem Aufwand das *Fifth Generation Project* mit dem Ziel, leistungsfähige intelligente PROLOG- Maschinen zu bauen.
- Boom der KI Industrie (Ertrag: Mld. Dollars in 1988)
- 1982: Das Expertensystem R1 zur Konfiguration von Computern spart der Digital Equipment Corporation 40 Millionen Dollar pro Jahr .
- 1986: Renaissance der Neuronalen Netze unter anderem durch Rumelhart, Hinton und Sejnowski . Das System Nottalk lernt das Vorlesen von Texten.
- 1990: Pearl, Cheeseman , Whittaker, Spiegelhalter bringen mit den Bayes-Netzen die Wahrscheinlichkeitstheorie in die KI.

Multiagentensysteme werden populär.

- 1993 Weltweite RoboCup Initiative zum Bau Fußball spielender autonomer Roboter.



GESCHICHTE DER KI

- Der Winter... Die Firmen schaffen nicht, die extravaganten Versprechungen einzuhalten.
- Die Rückkehr der neuronalen Netze (1986)
- KI wird wissenschaftlich fundiert (1987), sowie neue Methoden: Hidden Markov Modelle, Bayesische Netzwerke, Data Mining
- Formalisierung und Spezialisierung führt aber zur Fragmentierung



GESCHICHTE DER KI

DIE ENTSTEHUNG INTELLIGENTER AGENTEN

- 1997: Erster internationaler RoboCup Wettkampf in Japan.
- 2003: Die Roboter im RoboCup demonstrieren eindrucksvoll, was KI und Robotik zu leisten imstande sind.
- 2006 Servicerobotik entwickelt sich zu einem dominanten Forschungsgebiet in der KI.
- 2010 Autonome Roboter fangen an, ihr Verhalten zu lernen.
- 2011 Die IBM-Software *Watson* schlägt zwei menschliche Meister in der US-Fernsehshow *Jeopardy!*. Watson kann natürliche Sprache verstehen und sehr schnell beantworten.



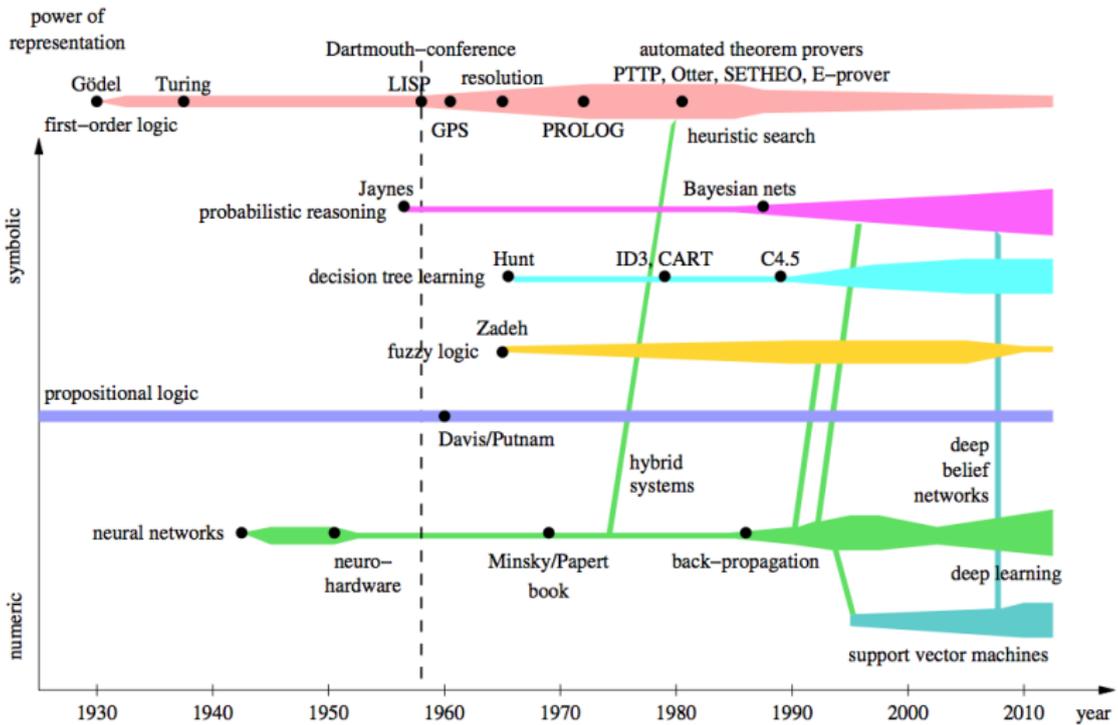
Aktuelle Forschungsrichtungen

- Technologie zum textuellen Sprachverstehen und zur Ermöglichung von Mensch-Computer-Dialogen. (Wahlster: Verbmobil-Projekt)
- Robotik-Ansatz: (Embodied artificial intelligence, R. Pfeifer). Untersuchung und Entwicklung von Sensorik, Motorik, d.h. physikalische Gesetzmäßigkeiten des Roboters und der Steuerung mittels Computer, *intelligentes Insekt*
- Automatische Deduktion Logik
- Robotik
- künstliche neuronale Netze
- Autonomes Fahren

DIE PHASEN DER KI-GESCHICHTE

- Die ersten Anfänge
- Logik löst (fast) alle Probleme
- Der neue Konnektionismus
- Schließen mit Unsicherheit
- Verteilte, autonome und lernende Agenten
- Die KI wird erwachsen





FAZIT DER BISHERIGEN ERFAHRUNGEN

- Fernziel scheint mit aktuellen Methoden, Techniken, Hardware und Software nicht erreichbar
- Motivationen und Visionen der KI sind in der Informatik verbreitet.
- Direkte Programmierung von Systemen (ohne sehr gute Lernalgorithmen) hat Grenzen in der Komplexität des einzugebenden Modells (Regeln, Fakten, ...)



FAZIT DER BISHERIGEN ERFAHRUNGEN

- Teilbereiche sind eigenständige Forschungsgebiete: Z.B. Sprachverarbeitung, Robotik, Automatische Deduktion, ...
- Aktuelle Forschungsziele sind eher spezialisierte Aspekte: Logik und Inferenzen, Mensch-Maschine-Kommunikation, Lernverfahren (adaptive Software), Repräsentationsmechanismen; eingebettete KI, nicht-Standard-Logiken und Schlussfolgerungssysteme, ...



ROBOTER UND WEITERE BEGRIFFE

- Roboter = Intelligenter Agent, Computer im Kern, agiert mit der physikalischen Umwelt
- Softbot = Software-Roboter, Umwelt ist i.A. nicht physikalisch, hat Wissensbasis und gibt Antworten und Ratschläge
- Webbot = Web-Roboter, (inter-)agiert im WWW, z.B. um Suchdatenbanken zu erstellen
- Chatbot = Chat-Roboter, interagiert in einem Chat

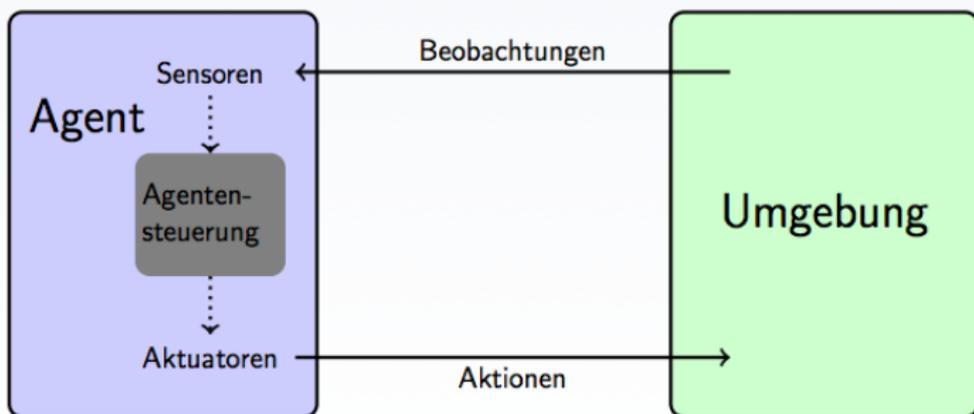


INTELLIGENTE AGENTEN

Agent = Überbegriff für alle KI-Systeme

Ein Agent hat

- *Sensoren* zum Beobachten seiner Umgebung und
- *Aktuatoren* (Aktoren; Effektoren) um die Umgebung zu manipulieren.



INTELLIGENTE AGENTEN

Agent

- macht Beobachtungen (Folge = Beobachtungssequenz)
- Aktion beeinflussen Umgebung und evtl. ihn selbst (z.B. Position)
- Agentenfunktion: Beobachtungsfolgen \rightarrow Aktionen.

Agentfunktion kann durch das Agentenprogramm implementiert werden



BEISPIELE

- Menschliche Agenten. Sensoren: Augen, Nase, Ohren.
Aktuatoren: Hände, Beine, Mund
- Roboter. Sensoren: Kamera, Infrarotsensoren. Aktuatoren:
Greifarme, Räder
- Software Agenten: Sensoren: Tastatur, Datenpakete.
Aktuatoren: Bildschirm, Datenpakete



INTELLIGENTE AGENTEN FORMALISIEREN

- Das Verhalten des Agenten ist beeinflusst von der vollständigen Geschichte der Daten, die der Agent aufgenommen hat.
- Formal, die Auswahl einer Handlung kann durch eine Agentenfunktion (Tabelle) beschrieben werden
 - $V^* \rightarrow A$, V ist eine Menge von Empfindungen, A ist die Menge der Handlungen
 - Die Agentenfunktion kann konstruiert werden, indem das Verhalten des Agenten für alle möglichen Wahrnehmungssequenzen beobachtet wird.
 - Wir brauchen eine Neustart - Fähigkeit für den Agenten und genug Speicherplatz
 - Die Agentenfunktion ist eine abstrakte mathematische Beschreibung.
- Intern wird die Agentenfunktion von einem Agentenprogramm implementiert.



AGENT ALLGEMEIN

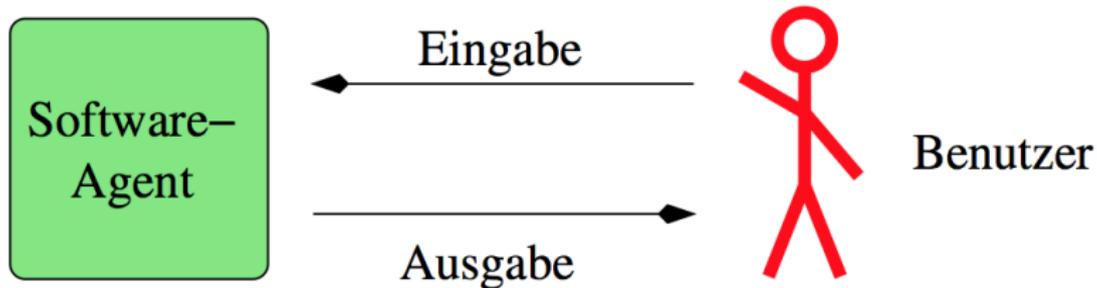


- **Vorwissen**: Wissen über die Umgebung z.B. Karte
- **Erfahrungswissen**: erlerntes Wissen, Testfälle
- **Ziele**: üblicherweise mit Prioritäten und Wichtigkeiten versehen.
- **Beobachtungen**: über die Umgebung und über sich selbst

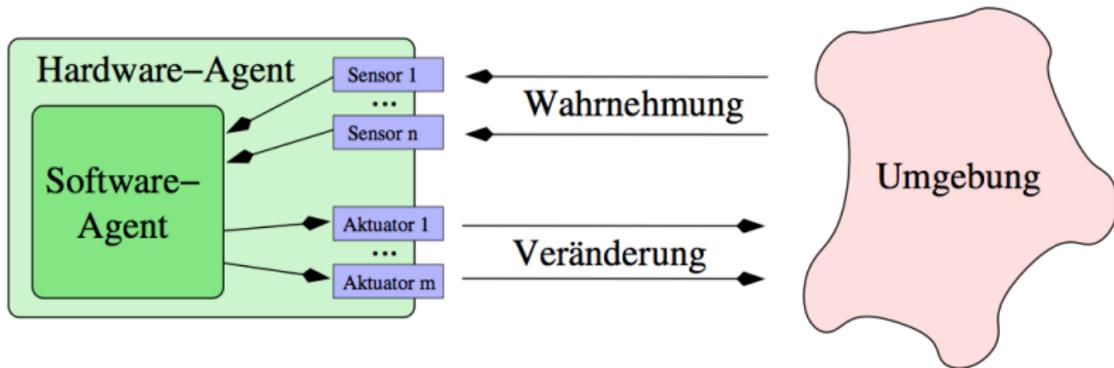
Nächste **Aktion** als Ausgabe (aufgrund von Schlussfolgern, Lernen,...)

AGENTEN

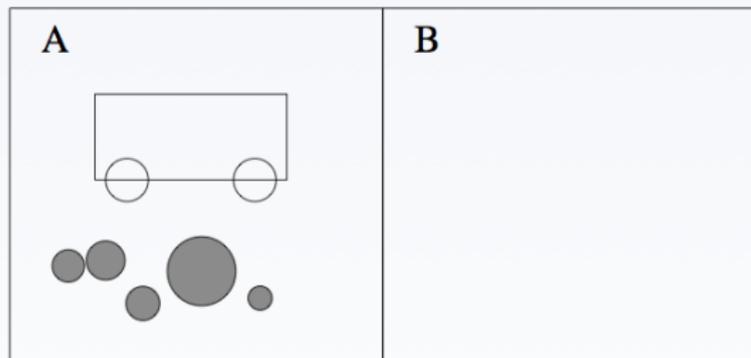
Software-Agent



Hardware-Agent (autonomer Roboter)



BEISPIEL: STAUBSAUGERWELT (RUSSEL & NORVIG)



- Orte: *A* oder *B*
- Jeder Ort: Dreckig / Sauber
- Agent kann nur aktuellen Ort beobachten (Sauber/Dreckig)
- Aktionen: *InsAndereQuadrat*, *Saugen* und *NichtsTun*.

BEISPIEL: STAUBSAUGERWELT (RUSSEL & NORVIG)

Problem:

Wann ist der Agent (das zugehörige Programm) gut / vernünftig bzw. intelligent ?

Notwendig:

Performanzmaß, d.h. eine Leistungsbewertung des Agenten

Z.B.

- Alles immer maximal sauber
- Möglichst sauber, aber wenig Stromverbrauch
- Möglichst sauber, aber wenig störend
- ...

Der optimale agierende Agent ist der intelligente Agent.



INTELLIGENTER AGENT

Definition

Ein vernünftiger (intelligenter, rationaler) Agent ist derjenige, der stets die optimale Aktion bzgl des Performanzmaßes wählt, aufgrund seiner Beobachtungsfolge und seines Vorwissens über die Umgebung.



Guter Agent:

- mittels der Sensoren Wissen über die Umgebung sammeln;
- lernfähig, bzw. sich adaptiv zu verhalten, aufgrund der Beobachtungssequenz

Agent wird als autonom bezeichnet, wenn der Agent eher aus seinen Beobachtungen lernt und nicht auf vorprogrammierte Aktionen angewiesen ist.



AGENTEN

- Reflex-Agent: Funktion von der Menge aller Eingaben auf die Menge aller Ausgaben
- Agent mit Gedächtnis: ist keine Funktion. Warum?
- lernfähiger Agent
- verteilte Agenten
- Markov-Entscheidungsprozess: zur Bestimmung der optimalen Aktion wird nur der aktuelle Zustand benötigt.
- zielorientierter Agent



BEISPIEL

Spamfilter: Sein Ziel ist es, Emails in die richtige Klasse einzuteilen.

Agent 1:

		korrekte Klasse	
		erwünscht	SPAM
Spamfilter entscheidet	erwünscht	189	1
	SPAM	11	799

Agent 2:

		korrekte Klasse	
		erwünscht	SPAM
Spamfilter entscheidet	erwünscht	200	38
	SPAM	0	762

Ist Agent 2 schlechter als Agent 1?



AGENTEN

Definition

Ziel eines kostenorientierten Agenten ist es, die durch Fehlentscheidungen entstehenden Kosten langfristig, das heißt im Mittel, zu minimieren. Die Summe aller gewichteten Fehler ergibt die durch die Fehlentscheidungen entstandenen Kosten.

Beispiel

Blinddarmdiagnosesystem LEXMED. Analog ist das Ziel eines nutzenorientierten Agenten (engl. utility based agent), den durch korrekte Entscheidungen entstehenden Nutzen langfristig, das heißt im Mittel, zu maximieren.



UMGEBUNG

KLASSIFIKATIONEN

- Vollständig beobachtbar vs. teilweise beobachtbar. → Agentensensoren ermöglichen den Zugriff auf den vollständigen Zustand der Umgebung
 - Der Staubsauger kann z.B. nur sein eigenes Quadrat beobachten.
- Deterministisch vs. Stochastisch → Der nächste Zustand der Umgebung wird vollständig durch den aktuellen Zustand bestimmt
 - 8-Puzzle
 - Dreck erscheint zufällig in den Quadraten.
- Nicht-deterministisch: Schachcomputer, Roboter



UMGEBUNG

KLASSIFIKATIONEN

- Episodisch vs. sequentiell → Die Erfahrung des Agenten ist in atomare Episoden unterteilt (die nächste Episode hängt nicht von Aktionen ab, die in vorherigen Episoden ausgeführt wurden)
- Episodisch: Feste unabhängige Zeitabschnitte, in denen beobachtet und nicht agiert wird
- Sequentiell. Es gibt Spätfolgen der Aktionen.



UMGEBUNGEN

KLASSIFIKATIONEN

- Statisch vs. Dynamisch →
 - Die Umwelt ändert sich nicht, während der Agent entscheidet
 - semidynamisch = Umgebung ändert sich nicht, aber der Performance-Score tut es

Dynamisch: die Umgebung kann sich während der Nachdenkzeit des Agenten verändern.

Semi-Dynamisch: Umgebung unverändert, aber Performanzmaß verändert sich

Z.B.: Schachspielen mit Uhr.

- Diskret (Schachcomputer) vs. Stetig. → hängt vom Zustand der Umgebung ab, wie die Zeit behandelt wird und von den Wahrnehmungen und Handlungen des Agenten



UMGEBUNGEN

KLASSIFIKATIONEN

- Ein Agent oder Multiagenten →
 - Welche Entitäten sollten als Agenten betrachtet werden?
 - Wenn ihr Verhalten am besten als maximierendes Leistungsmaß beschrieben wird.
 - Wettbewerbsgegner vs. kooperative Multi-Agenten-Umgebungen
- Bei Multiagenten: Klassifizierung in: Gegner / Wettbewerber / Kooperierenden Agenten.



UMGEBUNGEN, BEISPIELE (RUSSEL & NORVIG)

Arbeits- umgebung	Beobacht- bar	Deter- min- istisch	Episo- disch	Statisch	Diskret	Agenten
Kreuzwörterzel	vollst.	det.	seq.	statisch	diskret	1
Schach mit Uhr	vollst.	det.	seq.	semi	diskret	n
Poker	teilw.	stoch.	seq.	statisch	diskret	n
Backgammon	vollst.	stoch.	seq.	statisch	diskret	n
Taxifahren	teilw.	stoch.	seq.	dyn.	stetig	n
Medizinische Diagnose	teilw.	stoch.	seq.	dyn.	stetig	1
Bildanalyse	vollst.	det.	episod.	semi	stetig	1
Interaktiver Englischlehrer	teilw.	stoch.	seq.	dyn.	diskret	n



UMGEBUNGEN

- Die einfachste Umgebung
 - vollständig beobachtbar, deterministisch, episodisch, statisch, diskret mit einem Einzelagent
- Die schwierigste Umgebung (real-life)
 - teilweise beobachtbar, stochastisch, sequentiell, dynamisch, kontinuierlich, Multi-Agenten



STRUKTUR DES AGENTEN

Agent besteht aus:

- Architektur des Agenten:
Physikalischer Aufbau inkl. Sensoren und Aktuatoren
- Programm des Agenten



STRUKTUR DES AGENTEN

- Architektur = ein Computergerät mit physikalischen Sensoren und Aktuatoren
- Programm = Implementierung der Agentenfunktion
 - die Abbildung von Wahrnehmungen zu Handlungen
 - Genauer gesagt, nimmt das Agentenprogramm die aktuelle Wahrnehmung als Eingabe (da es von der Umgebung nicht mehr verfügbar ist) und gibt eine Aktion an die Aktuatoren zurück
 - Wenn die Aktionen eines Agenten von der gesamten Wahrnehmungssequenz abhängen, muss sich der Agent an die Wahrnehmungen erinnern
 - Natürlich muss das Programm für die Architektur geeignet sein!



ARTEN VON AGENTEN

- Tabellengesteuerter Agent (endlicher Automat) mit einer Tabelle von Einträgen:
Beobachtungsfolge1 \mapsto Aktion1
Beobachtungsfolge2 \mapsto Aktion2
- leicht abgeändert: Agent mit Zustand implementiert Funktion der Form
(Zustand, Beobachtung) \mapsto Aktion, Zustand



TABELLENGESTEUERTER AGENT

Ein einfacher Agent, der die gesamte Wahrnehmungssequenz im Speicher hält und sie als Index für die Tabelle mit Aktionen verwendet.

```
function TABLE-DRIVEN_AGENT(percept) returns an action
  static: percepts, a sequence initially empty
           table, a table of actions, indexed by percept sequence
  append percept to the end of percepts
  action ← LOOKUP(percepts, table)
  return action
```



TABELLENGESTEUERTER AGENT (2)

Probleme:

- die Tabelle ist zu groß (selbst für Agenten, die mit einer begrenzten Anzahl von Schritten arbeiten)
- Tabelle unvollständig
- Kein Agent könnte jemals aus Erfahrung die richtigen Tabelleneinträge lernen
- Der Entwickler hat keine Anleitung, wie die Tabelleneinträge gefüllt werden sollen

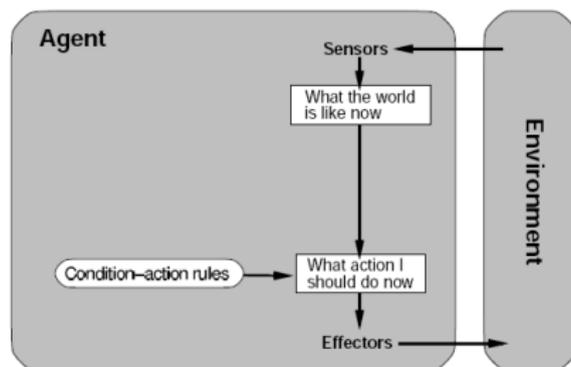
Wir brauchen eine andere Lösung!



ARTEN VON AGENTEN

EINFACHER REFLEXAGENT

- Der Agent wählt eine Aktion auf der Grundlage der aktuellen Wahrnehmung aus
- als Condition-Action-Regeln implementiert (if property = dirty, then clean up)
- signifikante Reduktion der Anzahl der Möglichkeiten (auf die Anzahl der Wahrnehmungen)



ARTEN VON AGENTEN

EINFACHER REFLEXAGENT

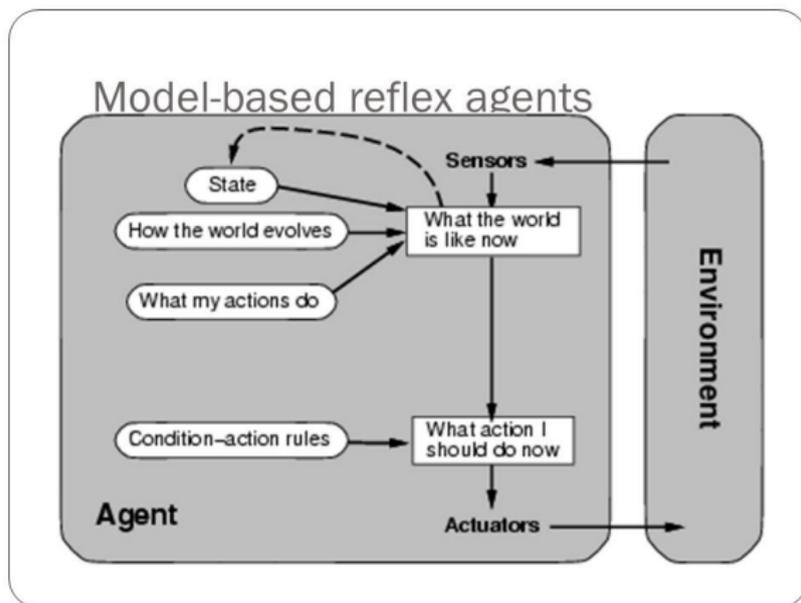
```
function SIMPLE-REFLEX-AGENT(percept) returns an action
  static: rules, a set of condition-action rules
  state ← INTERPRET-INPUT(percept)
  rule ← RULE-MATCH(state, rule)
  action ← RULE-ACTION[rule]
  return action
```

- Funktioniert für vollstndig beobachtbare Umgebungen (andernfalls kann es unendlich viele Schleifen geben).
- Randomisierung von Aktionen kann helfen, aus Endlosschleifen zu entkommen.



ARTEN VON AGENTEN

MODELLBASIERTE REFLEXAGENTEN



ARTEN VON AGENTEN

MODELLBASIERTE REFLEXAGENTEN

- Partielle Beobachtbarkeit kann gehandhabt werden, indem man den Teil der Welt verfolgt, den der Agent jetzt nicht sehen kann.
- Zwei Arten von Wissen sind notwendig
 - Wie entwickelt sich die Welt? (unabhängig vom Agenten)
 - wie die eigenen Handlungen des Agenten die Welt beeinflussen
- Modell der Welt



ARTEN VON AGENTEN

MODELLBASIERTE REFLEXAGENTEN

function MODEL-BASE-REFLEX-AGENT(*percept*) **returns** an action

static: *rules*, a set of condition-action rules

model, how the next state depends on the current state and action
state, a description of the current world state
action, the most recent action.

state ← UPDATE-STATE(*state*, *action*, *percept*, *model*)

rule ← RULE-MATCH(*state*, *rule*)

action ← RULE-ACTION[*rule*]

return *action*



ARTEN VON AGENTEN

ZWECKBASIERTER REFLEXAGENTEN (GOALBASED, ZIELBASIERT):

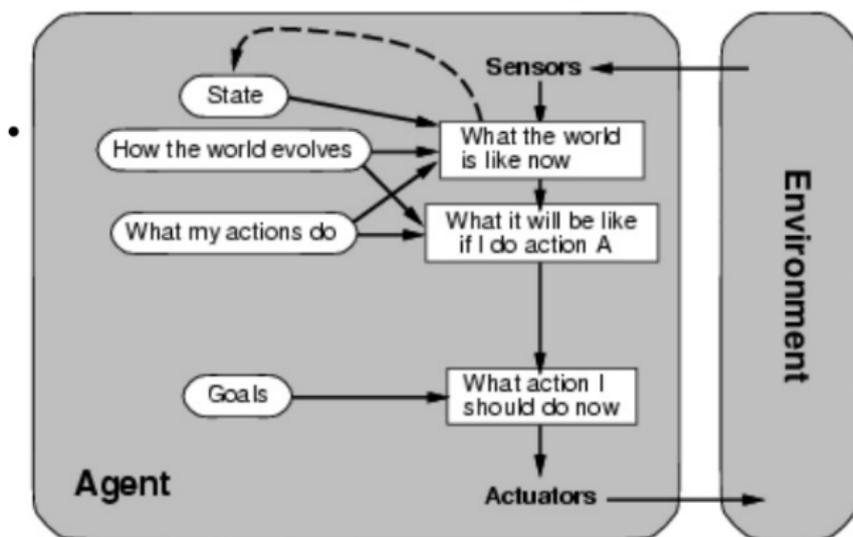
- Aktionen steuern auf das (vorgegebene) Ziel zu.
- Z.B. Taxi kann anhand des Ziels entscheiden ob links oder recht abbiegen
- Agent kann erkennen, ob er das Ziel erreicht hat.



ARTEN VON AGENTEN

ZWECKBASIERTER REFLEXAGENTEN

Goal-based agents



ARTEN VON AGENTEN

ZWECKBASIERTE REFLEXAGENTEN (GOALBASED, ZIELBASIERT):

- Die Aktionsauswahl basiert nicht nur auf dem Status, sondern auch darauf, was der Agent versucht.
- Der Agent benötigt eine Art Zielinformation, die wünschenswerte Situationen beschreibt.
- Die wichtigste Neuerung ist die **Berücksichtigung der Zukunft**.
- **Suche und Planung** sind der Suche nach Handlungssequenzen gewidmet, die die Ziele des Agenten erreichen.
- Der zielgesteuerte Agent scheint **weniger effizient** zu sein als der einfache Reflexagent, ist aber **flexibler**.



ARTEN VON AGENTEN

NUTZENBASIERTER AGENT(UTILITY-BASED, NUTZENBASIERT):

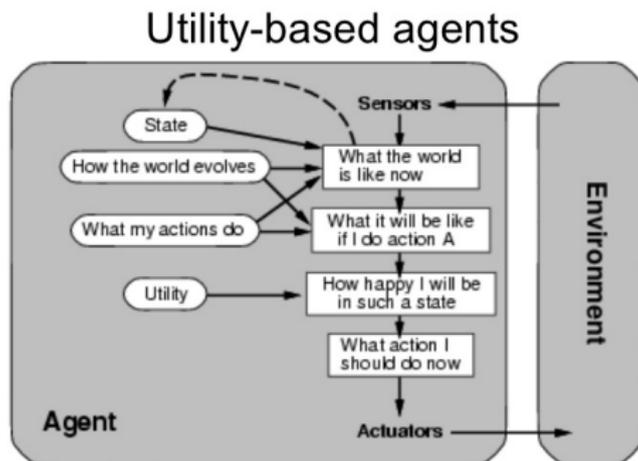
- Nächste Aktion wird anhand einer Nutzenfunktion (internes Gütemaß) berechnet.
- benötigt vorgegebene Bewertungsfunktion (utility function)



ARTEN VON AGENTEN

NUTZENBASIERTE AGENTEN

- Ziele allein reichen nicht aus, um qualitativ hochwertiges Verhalten in den meisten Umgebungen zu erzeugen (Ziel erreicht / nicht erreicht).
- Ein allgemeineres Leistungsmaß ermöglicht den Vergleich zwischen verschiedene Zustände.



ARTEN VON AGENTEN

NUTZENBASIERTE AGENTEN

- Es ist möglich, die Zustände (oder ihre Sequenzen) auf Nützlichkeit abzubilden, um das Leistungsmaß zu beschreiben.
- Die Nützlichkeitesfunktionfunktion ist eine Internalisierung des Leistungsmaßes (der Agent wählt Aktionen aus, um den Nutzen zu maximieren, der rational ist, wenn er dem externen Leistungsmaß entspricht).
- Der Agent kann auch dann erfolgreich sein, wenn widersprüchliche Ziele angegeben sind oder Chancen zum Erreichen unterschiedlicher Ziele nicht gleich sind.



ARTEN VON AGENTEN

LERNENDE AGENTEN

- Bisher haben wir beschrieben, wie Agenten Aktionen auswählen, aber wie finden wir die Programme für die Aktionsauswahl?
- Eine Möglichkeit besteht darin, Lernagenten aufzubauen und sie dann zu unterrichten, anstatt sie zu aufzuweisen.
- Der Lernagent kann in einer zunächst unbekanntem Umgebung arbeiten und kompetenter werden, als es sein anfängliches Wissen zulässt.



ARTEN VON AGENTEN

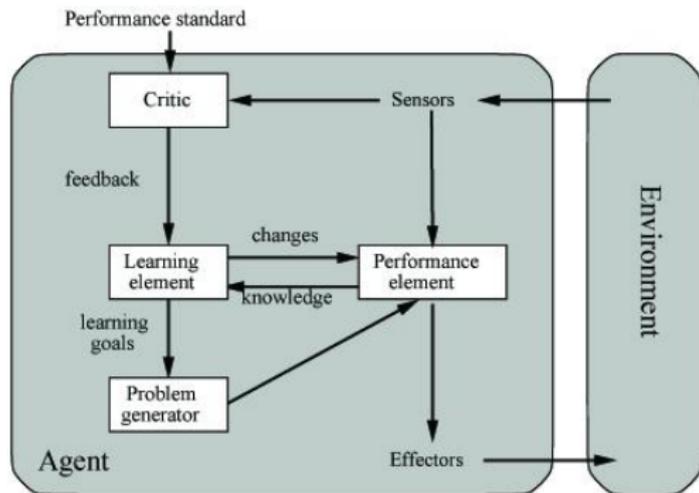
LERNENDE AGENTEN

- Wir können jede Agentenstruktur auf einen lernenden Agenten erweitern, indem wir annehmen:
 - Leistungselement (Performance): die ursprüngliche Agentenstruktur, die verantwortlich für die Auswahl der Aktion ist
 - Lernelement: verantwortlich für Verbesserungen
 - Kritik: Feedback darüber, wie der Agent handelt (sich selbst wahrnimmt)
 - Problem Generator: verantwortlich für Vorschläge von Aktionen, die zu neuen und informativen Erfahrungen führen



ARTEN VON AGENTEN

LERNENDE AGENTEN



ARTEN VON AGENTEN

LERNENDE AGENTEN

- Lernverfahren erhöhen die Autonomie des Agenten.
- Ausführungsmodul: bisheriges Programm
- Lernmodul: verwendet Beobachtungen um Parameter im Programm anzupassen
- Lernmethoden: online, offline (mit Trainingsphase)



ZUSAMMENFASSUNG AGENTEN

- Ein **Agent** ist eine Entität, die in einer Umgebung Wahrnehmungen aufweist und handelt.
- Die **Agentenfunktion** spezifiziert die Handlung eines Agenten als Antwort auf eine Wahrnehmungssequenz.
- Das **Leistungsmaß** bewertet das Verhalten des Agenten in einer Umgebung.
- Ein **rationaler Agent** wirkt, um den erwarteten Wert des Leistungsmaßes zu maximieren.



ZUSAMMENFASSUNG AGENTEN

- Das Agentenprogramm implementiert die Agentenfunktion.
 - Einfache Reflexagenten reagieren direkt auf Wahrnehmungen
 - Modellbasierte Reflexagenten speichern interne Zustände, um Aspekte der Welt zu verfolgen, die in der gegenwärtigen Wahrnehmung nicht sichtbar sind.
 - Zielbasierte Agenten handeln, um ihre Ziele zu erreichen.
 - Nutzenbasierte Agenten versuchen, ihre eigene Glücklichkeit zu maximieren
- Alle Agenten können durch Lernen ihre Leistung verbessern.
- Verschiedene Komponenten der Agentenstruktur beantworten Fragen wie:
 - Wie ist die Welt jetzt?
 - Was soll ich jetzt tun?
 - Was bewirken meine Handlungen?



WISSENSBASIERTE SYSTEME

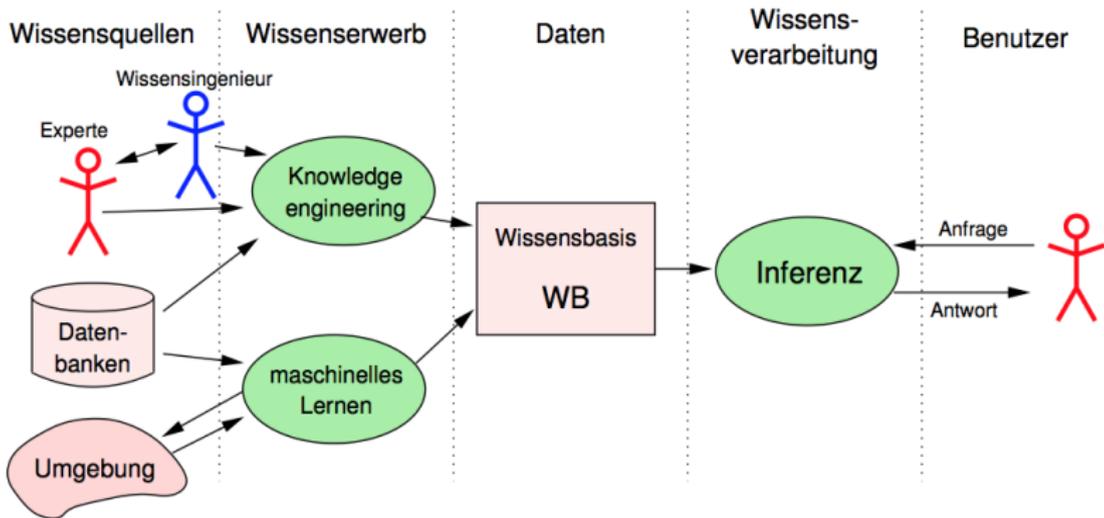
- strenge Trennung:
- Wissen
- Inferenzmechanismus
- Wissensbasis, kurz WB (engl. knowledge base, KB)



KNOWLEDGE ENGINEERING

- Wissensquelle z.B. Datenbanken, menschl. Experten
- Wissensingenieur (engl. knowledge engineer)
- Maschinelles Lernen u.a. aktive Exploration durch Agenten





Struktur eines klassischen wissensverarbeitenden Systems.

TRENNUNG VON WISSEN UND INFERENZ HAT VORTEILE:

- Inferenz ist anwendungsunabhängig (Beisp. med. XPS)
- Wissen kann deklarativ gespeichert werden.



DARSTELLUNG DES WISSENS MIT FORMALER SPRACHE:

- Aussagenlogik
- Prädikatenlogik erster Stufe (kurz: PL1).
- probabilistische Logik,
- Fuzzylogik
- Entscheidungsbäume
- Begriffsverbände

