

60 Jahre Künstliche Intelligenz: Bilanz und Perspektiven

Bernhard Nebel

Technische Fakultät, Institut für Informatik
Grundlagen der Künstlichen Intelligenz

Albert-Ludwigs-Universität Freiburg



**UNI
FREIBURG**

Intelligente Maschinen?



- Könn(t)en Maschinen denken?



- Können Maschinen so agieren, als ob sie menschlich wären?
- Können Sie kognitive Fertigkeiten besitzen?

- Turing-Test
- Eine kurze Geschichte der KI
- Teilgebiete der KI
- KI-Forschung in Freiburg
- Aktuelle Entwicklungen
- Ausblicke

- Alan Turing in *Computing Machinery and Intelligence* (1950):
 - I propose to consider the question: 'Can machines think?'
 - Da unklar ist, was „Denken“ ist, soll die Frage durch „Können Maschinen das **Imitations-Spiel** erfolgreich spielen?“ ersetzt werden.



- Der Turing-Test ist ein **Gedankenexperiment**, kein ernst gemeinter Test, um Intelligenz von Maschinen zu **messen**!
- **Chatbots** sind heute tatsächlich Realität.
- Es wurden andere Tests vorgeschlagen, z.B. Aufnahmeprüfungen für Schulen zu bestehen ...
- Tatsächlich interessiert uns ja aber, ob wir bestimmte **kognitive Fertigkeiten** auf Maschinen realisieren können (Bildverstehen, Textverstehen, Übersetzung, ...)

- Turing-Test
- Eine kurze Geschichte der KI
- Teilgebiete der KI
- KI-Forschung in Freiburg
- Aktuelle Entwicklungen
- Ausblicke

- **Vorgeschichte:** Philosophie, Psychologie, Linguistik und Mathematik haben seit ihren Anfängen für die KI relevante Fragestellungen bearbeitet.
- **1941-1956:** Mit der Entwicklung der ersten Rechenanlagen entstand auch die Frage: Kann man menschliche Intelligenz auf dem Computer nachbilden?

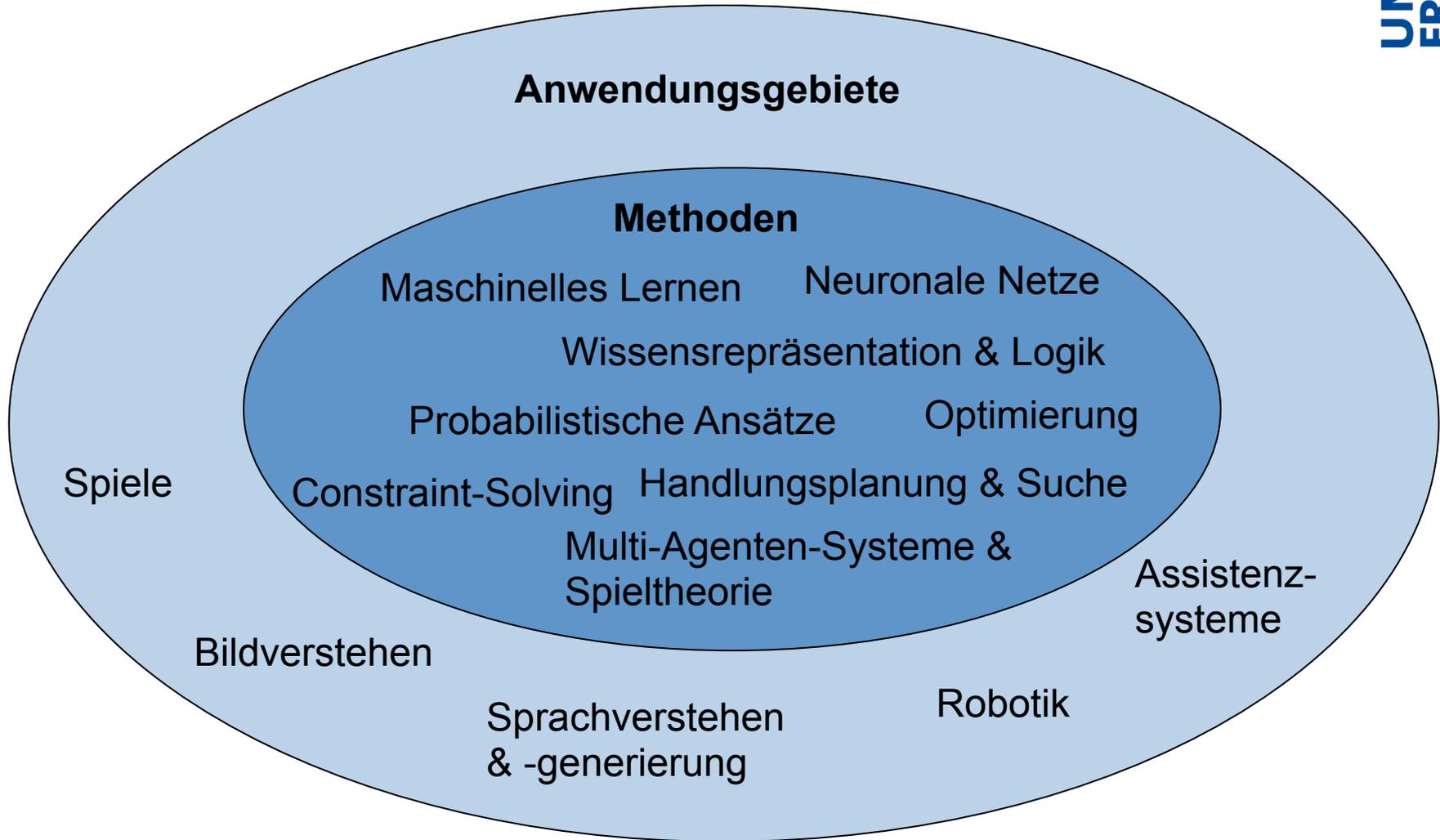
- **1956:** John McCarthy organisiert den Workshop *Dartmouth Summer Research Project on Artificial Intelligence* (Teilnehmer u.a.: Marvin Minsky, Allen Newell, Herbert Simon, Claude Shannon)
 - We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 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. [...]

- **1956-1960: Früher Enthusiasmus**
 - It is not my aim to surprise or shock you—but the simplest way I can summarize is to say that there are now [...] machines that **think**, **learn** and that **create**. Moreover, their ability to do these things is going to increase rapidly until—in **the visible future**—the range of problems they can handle will be coextensive with the range to which human mind has been applied [H. Simon, 1957]
- **1960-1970: Intelligentes Verhalten wird in vielen Mikrowelten demonstriert**
 - Lösen von geometrischen Analogie-Problemen, Sprachverstehen in eingeschränkten Domänen, ...

- 1970-1980: Wissensbasierte Systeme werden das wesentliche Paradigma:
 - Nicht alleine die Methode ist wichtig, sondern das Wissen, das für die Problemlösung genutzt wird
- 1980-1990: Das Jahrzehnt der Expertensysteme
 - Basierend auf Erfolgen von regelbasierten Systemen (für Konfiguration und Diagnose) hoffte man, Expertenwissen formal darstellen zu können, um so Experten zu ersetzen
- Ende der 80er: Der KI Winter
 - Die Expertensysteme können die Erwartungen nicht erfüllen

- Seit den 90ern
 - neue Methoden (Mathematisierung, z.B. probabilistische Robotik)
 - neue Sichtweisen (rationale Agenten)
- Sichtbare Erfolge:
 - *Chinook*: Dame-Weltmeistertitel (1994)
 - *Deep-Blue*: Schachweltmeister besiegt (1996)
 - *Stanley*: Selbstfahrendes Auto (DARPA Challenge) (2005)
 - *Watson*: Jeopardy!-Quiz gewonnen (2011)
 - *AlphaGO*: Go-Großmeister besiegt (2016)

- Turing-Test
- Eine kurze Geschichte der KI
- Teilgebiete der KI
- KI-Forschung in Freiburg
- Aktuelle Entwicklungen
- Ausblicke



- Turing-Test
- Eine kurze Geschichte der KI
- Teilgebiete der KI
- **KI-Forschung in Freiburg**
- Aktuelle Entwicklungen
- Ausblicke

KI-Forschung in Freiburg



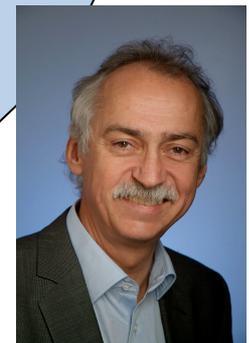
W. Burgard



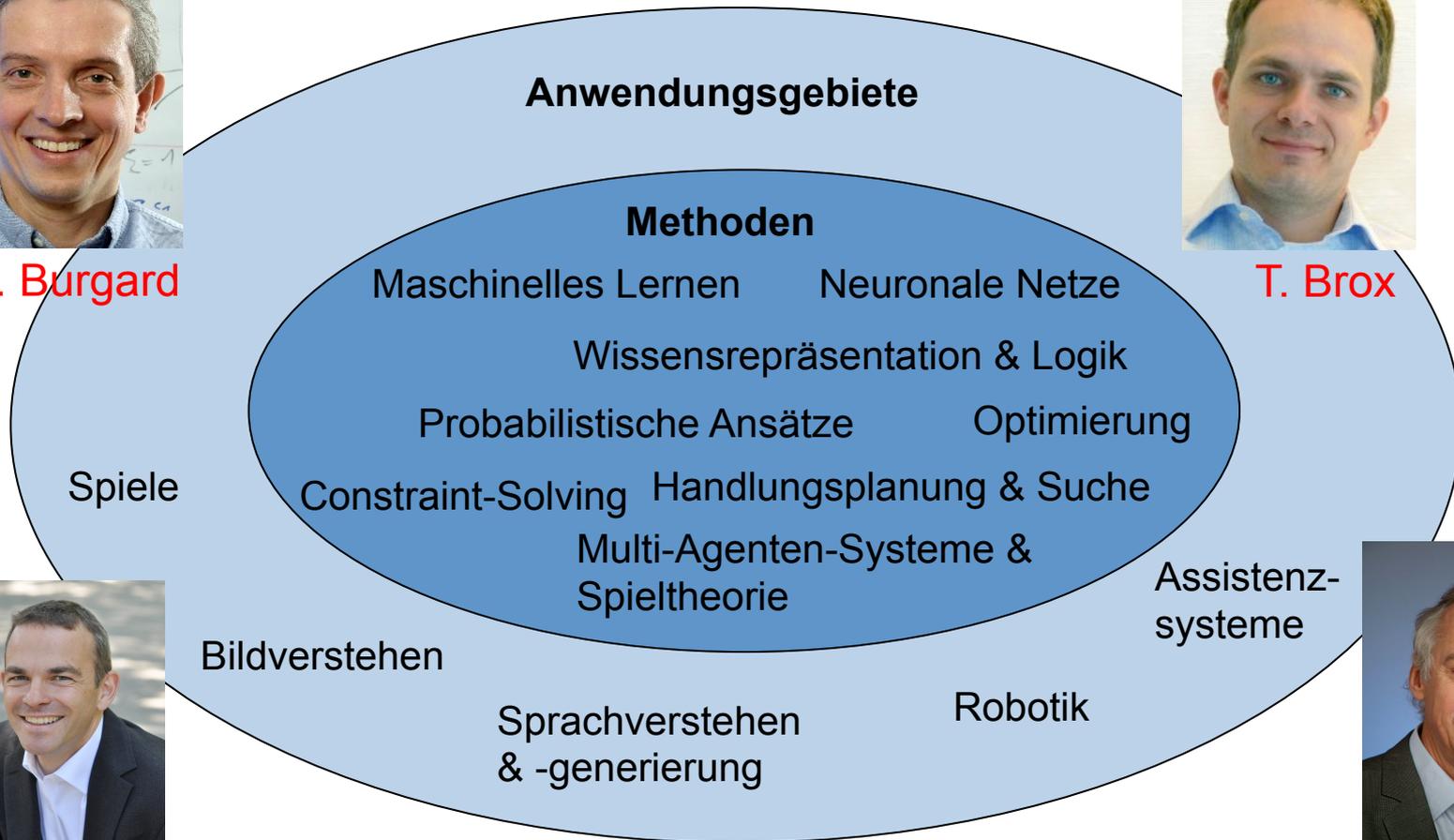
T. Brox



M. Riedmiller (jetzt DeepMind)



B. Nebel



Handlungsplanung: Was ist das?



- „Planning is the art and practice of thinking before acting.“

Patrick Haslum

- Intelligente Entscheidungsfindung: Welche Aktionen sind auszuführen?

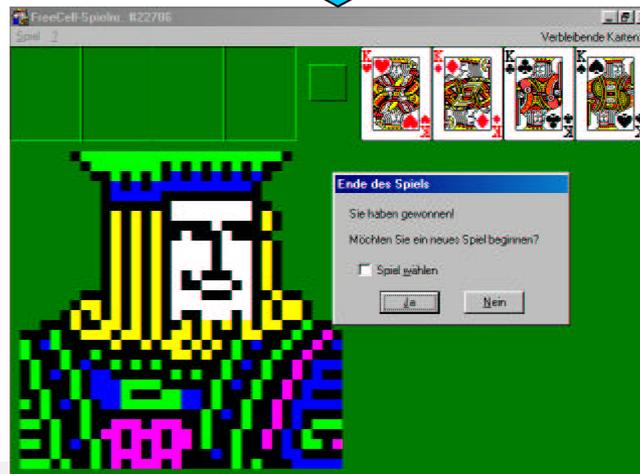
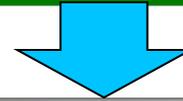
- Vielfältige Anwendungsgebiete:

- Robotik-Szenarien
- Autonome Systeme
- Fahrstuhlsteuerung
- Puzzles (z.B. Sokoban, Rubik's Cube, ...)

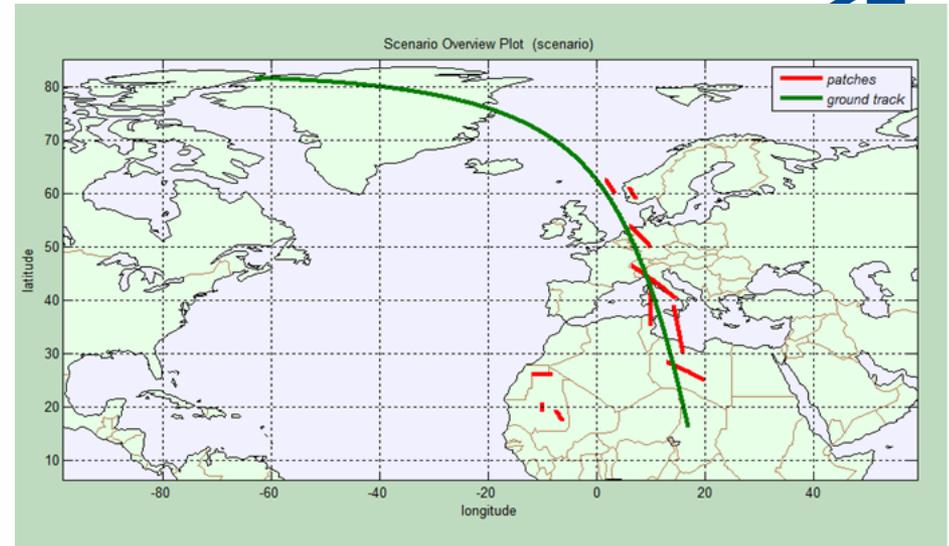
Handlungsplanung



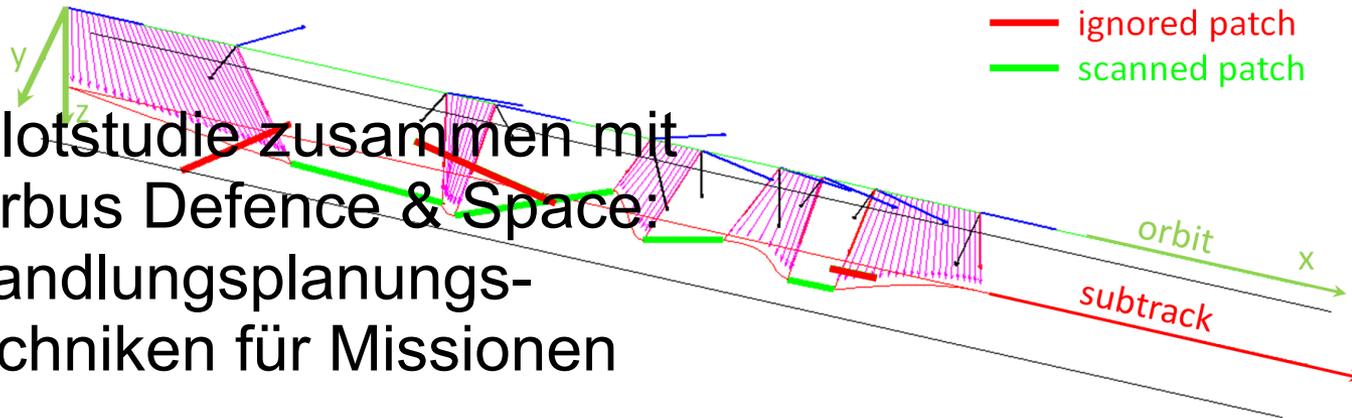
- Logische Formalisierung schwieriger Aufgaben
- Benutze (*automatisch ableitbare*) Heuristiken um die Probleme zu lösen
- ❖ Mit Preisen auf der IPC ausgezeichnete Planungssysteme: IPP, FF, FDD, TFD, PROST
- Erspart der Menschheit viel nutzlos verbrachte Zeit



- Autonome Systeme:
 - Service-Missionen für Satelliten
 - Landungs-Missionen
 - Erdbeobachtungen



- Pilotstudie zusammen mit Airbus Defence & Space: Handlungsplanungstechniken für Missionen



- Handlungsplanung macht den Roboter **intelligenter** und **flexibler**
 - **Ziele** anstatt komplexer Ablaufskripte
 - Dynamisch erzeugte **Pläne**
- **Modellierung** von Anwendungsszenarien als Planungsdomäne
 - Nur Basis-Funktionalitäten modelliert
 - Bedarfsgerechte Kombination durch Planer
- **Schlussfolgern** auf menschlichem Abstraktionsniveau

Planung in der Robotik



- KARIS: **K**leinskaliges **A**utonomes **R**edundantes **I**ntralogistik**S**ystem
- KIT, Univ. Freiburg, Hochschule Ravensburg-Weingarten + 12 Industrieunternehmen
- Intralogistik-Plattform
- Multi-Agenten-Szenario für die Logistik



- Das RoboCup-Projekt:



- 2050 treten Roboter gegen das menschliche Weltmeisterteam an

- Test für die **Integration** verschiedener KI-Techniken



Kooperation: Platzierung & Rollenwechsel

- Für jede **Rolle** (Angreifer, Verteidiger, Unterstützer) gibt es optimale Positionen (abhängig von Ball und Gegner)
- **Kosten** für Neuzuweisungen ergeben sich als Fahrkosten
- **Nutzen** ergibt sich aus Platzierung.
- Gruppe führt gemeinsame Optimierung aus.



StarKick: Kick it like Beckham



- Wir wollten nicht bis 2050 warten, um gegen Menschen anzutreten ...
- Wir bauen zusammen mit Gauselmann einen autonomen Tischkicker!



Mensch-Maschine-Interaktion: Implizite Koordinierung



- DESIRE-Roboter übergibt etwas:

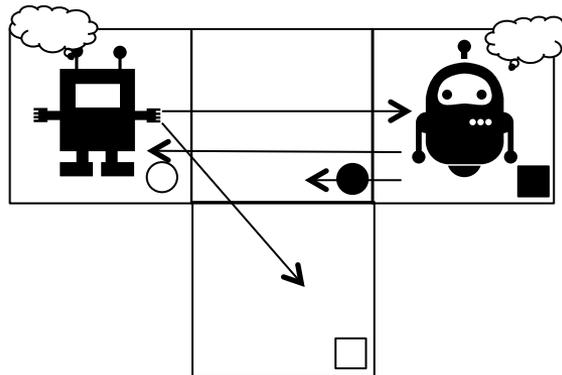


- Annahme: Alle Beteiligten haben ein **gemeinsames Ziel**.
- **Explizite** Koordination: Um einen gemeinsamen Plan zu entwickeln, muss man sich absprechen/verhandeln oder zentrale Vorgaben machen während man plant.
- **Implizite** Koordination: Basierend auf dem vorhandenen Wissen, wird **dezentral ein** Plan (für alle) entworfen, bei dem jeder entsprechend seines Wissens und seiner Fähigkeiten beiträgt, das **gemeinsame Ziel** zu erfüllen.

Implizite Koordinierung unter Zielunsicherheit



- Der rechteckige Roboter **R** will zu dem Feld mit dem schwarzen Quadrat und der runde Roboter **K** will zu dem schwarzen Kreis.
- Beide **glauben gegenseitig**, dass auch das Feld mit dem weißen Quadrat oder Kreis ein Zielfeld für den anderen sein könnte.

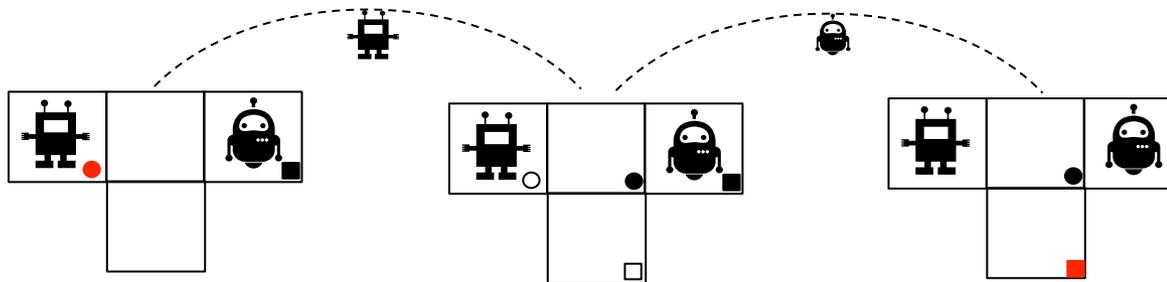


- Beide wollen, dass sie **beide erfolgreich** sind und die jeweiligen Zielfelder erreichen!
- Kann jeder der Roboter einen **implizit koordinierten Plan** erzeugen, der garantiert, dass beide Roboter erfolgreich sind?

Unsicherheit repräsentieren: Mögliche Welten



- Alle möglichen **Weltzustände** müssen betrachtet werden, ebenso wie die **Ununterscheidbarkeitsrelation** zwischen solchen Zuständen:



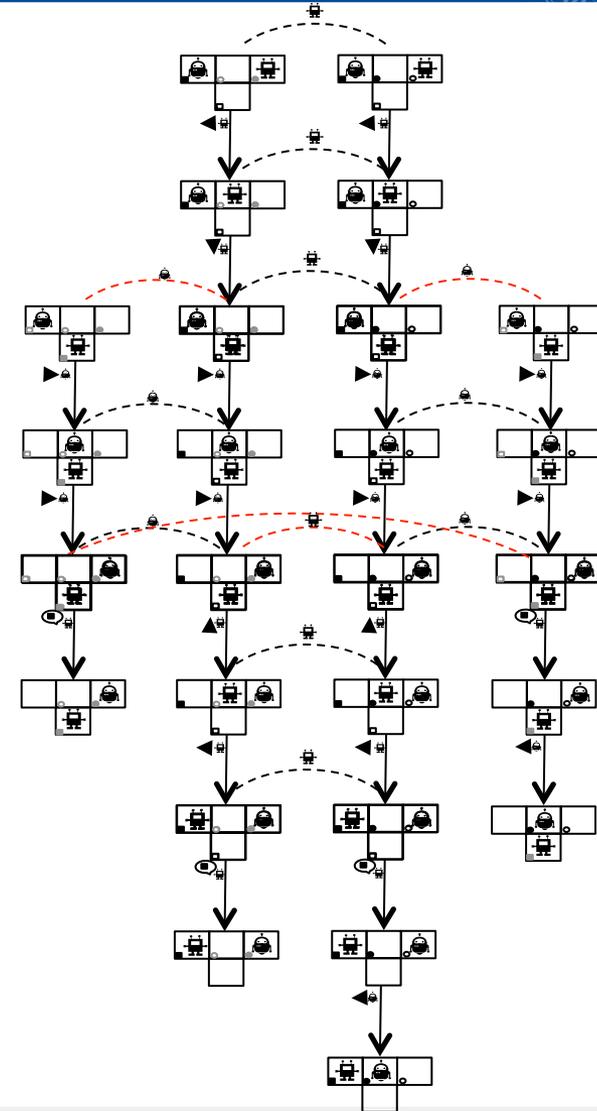
- Wenn Aktionen geplant werden, müssen diese **uniform** für alle ununterscheidbaren Welten sein.
- **Kommunikation** oder **Wahrnehmung** können Welten unterscheidbar machen!

- Während der Planung muss man die **Perspektive wechseln** und aus der Sicht des anderen planen.
- Man muss sich in den Partner versetzen und **seine Unsicherheit** mit einplanen.
- Für alle Welten, die man für **möglich** hält:
 - Betrachte alle Welten, die der andere dann für möglich halten würde
 - Fahre dann mit der Planung fort und führe weitere Perspektivenwechsel durch

Ein implizit koordinierter Plan



- **R** findet einen **implizit koordinierten Plan**, bei dem er beginnt.
- **K** findet nur einen Plan, bei dem er nicht beginnt, aber dann einsteigen kann.
- Die beiden Pläne sind **kompatibel**, aber kann man immer **Erfolg garantieren?**



- Turing-Test
- Eine kurze Geschichte der KI
- Teilgebiete der KI
- KI-Forschung in Freiburg
- Aktuelle Entwicklungen
- Ausblicke

Erfolge jüngerer Zeit I: Watson



- **Watson** und Jeopardy! (2011)



- Ein KI-System wird Meister in der US-Quizsendung *Jeopardy!*
- Information Retrieval, Verstehen natürlicher Sprache und Bewertungsmethoden für mögliche Antworten
- Verblüffend, da es oft im Trickfragen geht
- Anwendungen z.B. in der Medizin

- Autonomes PKW-Fahren (ab 2011)



- DARPA Challenges (2004/5/7) zeigten, dass Autos autonom lange Strecken fahren können
- Basiert auf guter Sensorik, GPS, und guten Verfahren zur Zustandsschätzung
- Für den Stadtverkehr werden Lernansätze genutzt

- Auch im Go-Spiel sind Maschinen mittlerweile besser (2015)



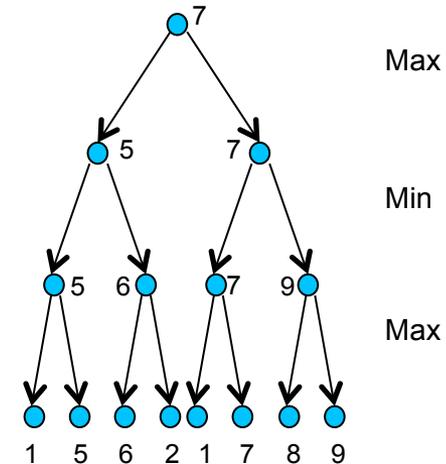
- Um mehrere Größenordnung schwieriger als Schach (361 Felder statt 64, Verzweigungsfaktor rund 200 statt 35, Suchtiefe rund 300 statt 100)
- Durch Kombination von zwei relativ neuen Techniken erfolgreich: **Monte-Carlo-Baumsuche** und **Deep Learning**

Technik-Durchbrüche in den letzten Jahren



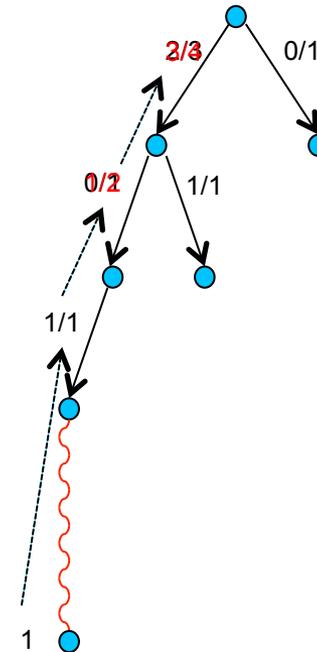
- Monte-Carlo-Baumsuche (speziell in Spielen und im probabilistischen Planen)
- Deep Learning (Künstliche neuronale Netze)
- Kombinatorische Optimierung (Millionen von Klauseln in Booleschen Formeln) und SMT
- Starke probabilistische Verfahren für Zustandsschätzung in der Robotik

- Normalerweise:
Systematische Suche (alle möglichen Züge) bis zu einer gegebenen Tiefe
 - Dann **Stellungsbewertung** durch einfach zu berechnende **Features**
 - Die Werte werden zur Wurzel hoch **propagiert**.
- Funktioniert nur, wenn gute Features bekannt sind und Tiefe und Breite des Baums stark beschränkt sind



- Baum wird **schrittweise** expandiert
(Selektionsstrategie notwendig: **Exploration vs. Exploitation**)
- Bei jeder Expansion wird ein **ganzes Spiel** zufällig bis zum Ende gespielt.
- Die Gewinn/Verlust-Werte werden zurückpropagiert.
- **Stellungsbewertung** ist deshalb nicht notwendig.
- **Große Verzweigungsfaktoren** sind tolerabel.

1. Selektion
2. Expansion
3. Simulation
4. Zurückpropagierung



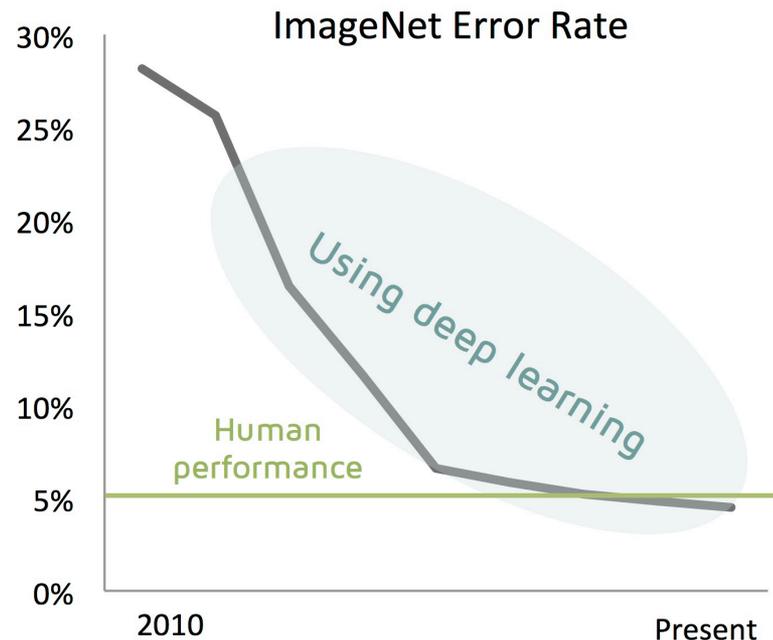
Deep Learning: Erfolge



- Es gibt einen **Objektklassifikations-Wettbewerb**, bei dem Teile von Bildern einer von 200 oder 1000 Klassen zugeordnet werden müssen.
- Beispiel:

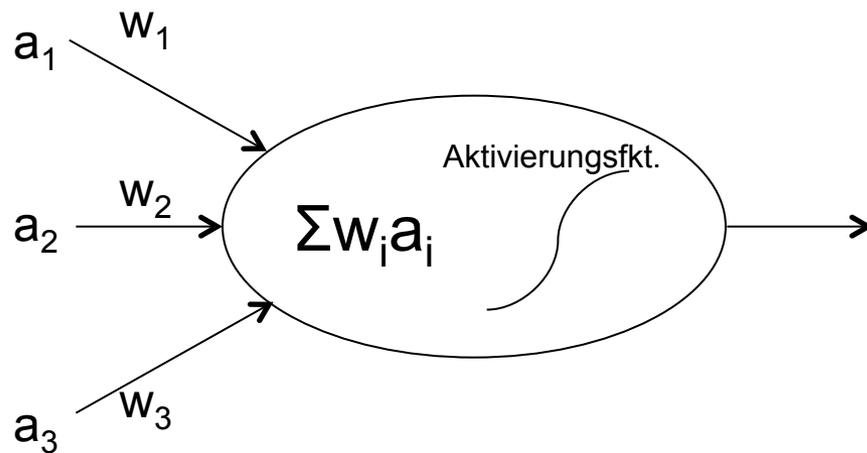


- Steigerung über die Jahre:



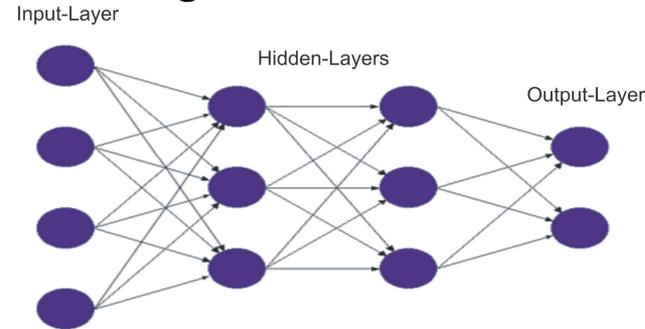
- Erfolgreich u.a. wegen **großer Trainingsmengen!**

- KNNs sind aus einfachen Bausteinen, den **künstlichen Neuronen** aufgebaut:



- Einzelne Neuronen können als **Klassifikator trainiert** werden, sind aber nicht sehr mächtig

- **Mehrlagige KNNs** sind mächtiger



- Lernmethode: **Fehler-Rückpropagierung** (Backpropagation)
- Sehr tiefe Netze mit bestimmter Topologie (**CNNs**), haben sich als sehr erfolgreich herausgestellt.
- Funktioniert nur mit hochparalleler Hardware (GPU-Programmierung)

- Turing-Test
- Eine kurze Geschichte der KI
- Teilgebiete der KI
- KI-Forschung in Freiburg
- Aktuelle Entwicklungen
- **Ausblicke**

- Techniken der KI werden zu massiver Arbeitslosigkeit führen.
 - Ein Argument, das es seit 1970 gibt und eigentlich die **gesamte Informatik** betrifft.
 - Wir arbeiten z.B. z.Zt. in einem Projekt daran, Robotersteuerungen zu entwerfen, die für die Produktion kleiner Stückzahlen effektiv sind.
 - Automatisierung zu stoppen hat relativ wenig Sinn und führt zu Verlust der Wettbewerbsfähigkeit.
 - Die Politik muss diesen **Technologiewandel** begleiten und Rahmenbedingungen setzen.



- Autos werden innerhalb der nächsten 10 Jahre teilautonom und autonom fahren können
 - Die Frage ist nicht ob, sondern wann.
 - Aber wird es auch das sein, was der Kunde will?
 - Ethische und juristische Fragen müssen geklärt werden.
 - Vermutlich wird dies zu einer deutlichen **Reduktion** der Unfallzahlen führen (vgl. Flugverkehr).
 - Es wird u.U. den Verkehr **flüssiger** machen.

- Autonome bewaffnete Roboter
 - Das Militär und auch der Katastrophenschutz entwickeln und setzen sowohl Drohnen als auch landgestützte *ferngesteuerte* Roboter ein
 - Im militärischen Bereich sind diese bewaffnet, z.Zt. aber ferngesteuert
 - Autonomie und Teilautonomie sind nur eine Frage der Zeit
 - Forderung nach einem **Memorandum** für die Entwicklung solcher Waffen!

- Die Forschung im Gebiet der KI könnte (u.U. nicht intendiert) Artefakte erschaffen, die sich selbst bewusst sind,
 - ... und damit tausendfaches Leid bewusster, empfindender „Wesen“, wenn diese abgeschaltet werden.
 - Diese „Wesen“ könnten sich gegen das Abschalten wehren.
 - Ungefähr **so wahrscheinlich** wie dass im CERN **schwarze Löcher** entstehen.
 - Oder vielleicht gibt es bereits ein solches Wesen?
 - Das **WWW** (das ist nicht ernst gemeint!)



- Die KI-Forschung zusammen mit der Hardware-Entwicklung könnte eine Superintelligenz erschaffen
 - Speicherkapazität und Verarbeitungskapazität steigen weiter **exponentiell**.
 - Irgendwann werden Rechner „**intelligenter**“ als der Mensch sein (aber was heißt intelligenter?).
 - Ab dem Zeitpunkt werden sich die Rechner selbst konstruieren, sind nicht mehr auf Menschen angewiesen, und **rottet** diese deshalb (?) aus.
 - Aber dann müssten diese sich **selbst bewusst** sein und **eigene Ziele** verfolgen,
 - oder aber **absolute Macht** und **unklare Ziele** übertragen bekommen haben.

- Ob Maschinen **denken** können, ist schwer zu sagen.
- **Kognitive Fertigkeiten** können aber auf Maschinen realisiert werden.
- Die **KI** hat dazu in den letzten 60 Jahren viele Beiträge geliefert.
- Es gibt eine Menge von **offensichtlichen Erfolgen** und **interessante neue Techniken**.
- Speziell **Deep Learning** ist in letzter Zeit sehr erfolgreich (man hat aber keine gute Theorie)
- Ohne die **Fortschritte bei der Hardware** wären diese Erfolge nicht möglich gewesen.
- Ich denke, wir werden noch viele interessante Dinge aus der KI sehen – die „**Superintelligenz**“ halte ich aber für eher unwahrscheinlich.

Dank an ...

