

Künstliche Intelligenz (Sommersemester 2024)

Kapitel 01: Einführung

Prof. Dr. Adrian Ulges

Künstliche Intelligenz (KI)

Ziel: Implementierung von **intelligentem Verhalten** in künstlichen Systemen

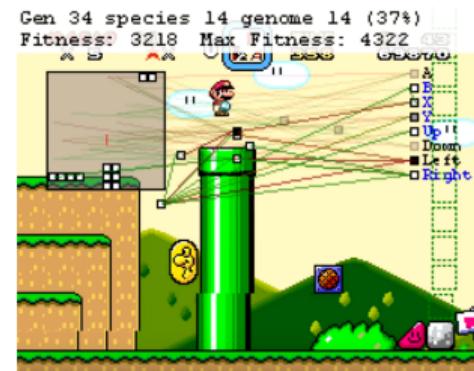
"For thousands of years, we have tried to understand how we think. [...] The field of A.I. goes further still: It attempts not just to understand but also to build intelligent entities."

(Russel/Norvig)

Was sind Beispiele für "intelligentes Verhalten"?

- ▶ Schach / Go / Super Mario spielen?
- ▶ Routen planen? Buchstaben in Scans erkennen?
- ▶ Text automatisch "verstehen"?

Die **Grenzen der KI** sind unscharf und ständig im Wandel! ("AI is whatever hasn't been done yet.")



5 0 4 1



Typische Aspekte intelligenten Verhaltens

- ▶ Schlussfolgern, Problemlösen, Planen
- ▶ Wahrnehmung, Kommunikation, Interaktion
- ▶ (robustes) Lernen

KI als Ingenieurwissenschaft

- ▶ Systeme entwickeln, die bei der **Lösung praktischer Probleme** unterstützen
- ▶ Bsp.: Suchmaschinen, autonomes Fahren, ...

KI als Erkenntniswissenschaft

- ▶ Die **Grundprinzipien** von Intelligenz erforschen
- ▶ Bsp.: Wie wird Wissen repräsentiert? Wie funktioniert unser Gehirn? Ist Artificial General Intelligence (AGI) möglich? ...



Atlas, The Next Generation



Künstliche Intelligenz hat verschiedene Zielsetzungen

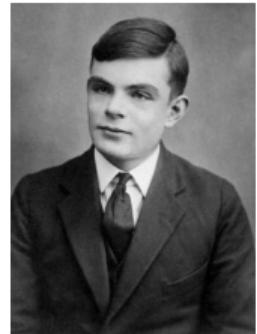
image: [1]

White-Box-Ansatz
Black-Box-Ansatz *

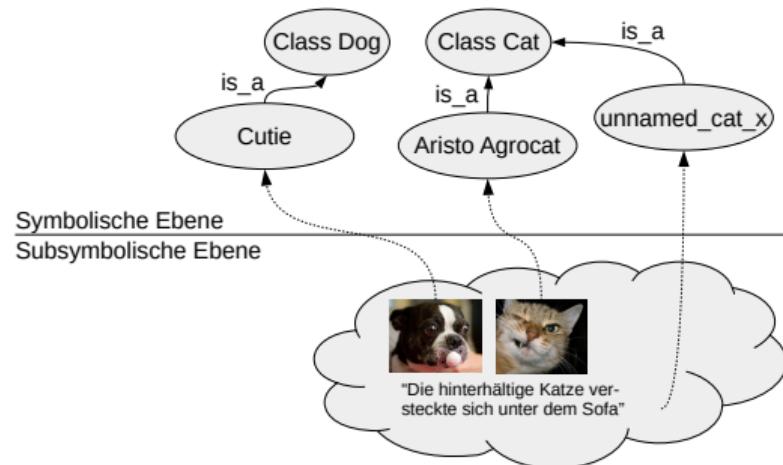
Menschliches Denken (kognitive Wissenschaft) Automatisierung von [...] Aktivitäten, die wir mit menschlichem Denken in Verbindung bringen, wie Entscheidungsfindung, [...] und Lernen (Bellmann 1978)	Rational Denken (Logik, Deduktion) Die Untersuchung geistiger Fähigkeiten mithilfe von Rechenmodellen (Charniak & McDermott, 1985)
Menschliches Handeln (Bsp. Turing-Test) ChatGPT Die Kunst, Maschinen zu erschaffen, die Funktionen ausführen, die Intelligenz erfordern, wenn sie von Menschen ausgeführt werden (Kurzweil, 1990)	Rational Handeln (Agenten, Roboter) Die Nachahmung intelligenten Verhaltens in Form von Rechenprozessen (Schalkoff, 1990)

Wo ordnen wir Large Language Models (LLMs) wie ChatGPT ein?

- ▶ LLMs = Reproduktion menschlichen Handelns
(*keine Nachahmung menschlichen Denkens!*)
- ▶ Der **Turing-Test**: Ein Tester chattet mit einer Maschine und mit einem Menschen. Wenn der Tester nicht identifizieren kann welcher Partner die Maschine ist, hat diese den Test bestanden.
- ▶ Ob moderne LLMs den Test bestehen wird derzeit debattiert [7].



KI operiert auf zwei Ebenen

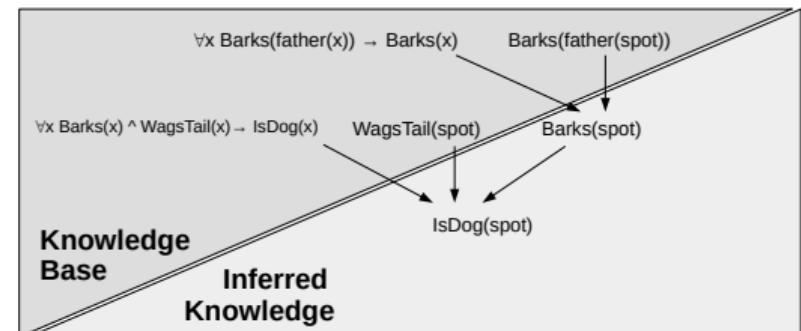
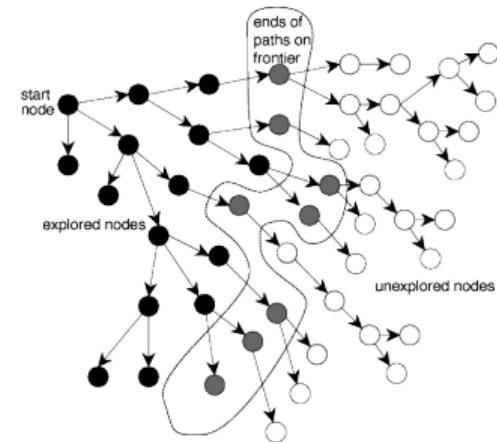


- 1. Subsymbolische Ebene:** Wir verarbeiten zunächst Signale aus der echten Welt. Diese repräsentieren *Konzepte*.
- 2. Symbolische Ebene:** Wir ziehen Schlussfolgerungen über diese Konzepte.

Die symbolische Ebene

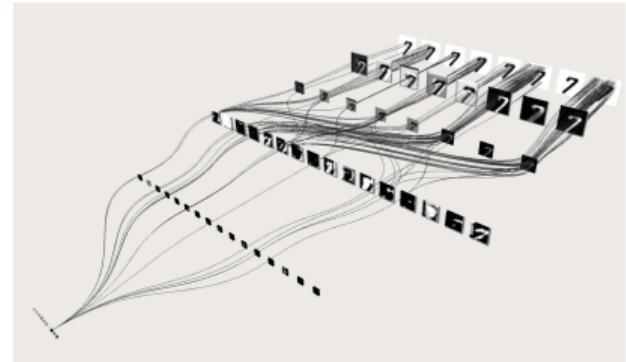
image: [6]

- ▶ Symbolische KI-Ansätze befassen sich mit Konzepten wie z.B. "Katze" und "haarig".
- ▶ Wissen wird explizit repräsentiert (in Form von Regeln, Fakten, Graphen, Zuständen usw.).
- ▶ Verarbeitung diskreter Symbole + Zustände.
- ▶ Symbolische KI-Ansätze verwenden oft Such- und Inferenzalgorithmen.
- ▶ Zentrale Themen: Logik, Wissensgraphen, regelbasierte Systeme.
- ▶ In dieser Vorlesung: Teil II (Krechel).



Die subsymbolische Ebene image: [6]

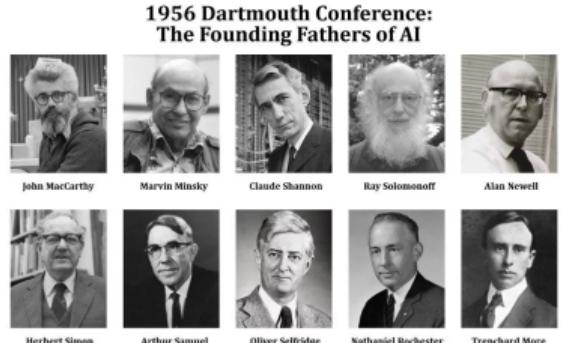
- ▶ Subsymbolische KI-Ansätze verarbeiten häufig Signale, z.B. das Bild einer Katze oder das gesprochene Wort “Katze” als Audiosignal.
- ▶ Verarbeitung numerische Werte / Signale um Muster zu erkennen.
- ▶ Wissen wird **implizit** repräsentiert (z.B., *in numerischen Vektoren/Matrizen*).
- ▶ Verwendung von Optimierungsalgorithmen in \mathbb{R}^d .
- ▶ Kernthemen: Statistik, maschinelles Lernen, neuronale Netze & deep learning.
- ▶ In dieser Vorlesung: Teil I (Ulges).



KI: Geschichte

1950er Jahre (sog. "Look ma, no hands" – Ära)

- ▶ Dartmouth-Konferenz (1956): Der **Begriff "KI"** wird geprägt.
- ▶ Erste **Systeme**, u.a.
 - ▶ Logik Theorist (*Beweiser in Aussagenlogik*)
 - ▶ Checker (*schlägt Menschen in "Dame"*)

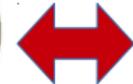


1960er Jahre: Erste Enttäuschungen

- ▶ Systeme **skalieren** (aufgrund kombinatorischer Explosion und limitierter Hardware) nicht auf reale Anwendungsfälle.
- ▶ Probleme sind **komplexer** als angenommen (*Beispiel: maschinelle Übersetzung*).
- ▶ Grundlegende Einschränkungen von **Algorithmen** (*Beispiel: Perzeptron (Minsky & Papert, 1969)*).



Frank Rosenblatt



Marvin Minsky

KI: Geschichte

1980er Jahre (sog. "wissensintensive Phase")

- ▶ **Symbolische Ebene: Expertensysteme**

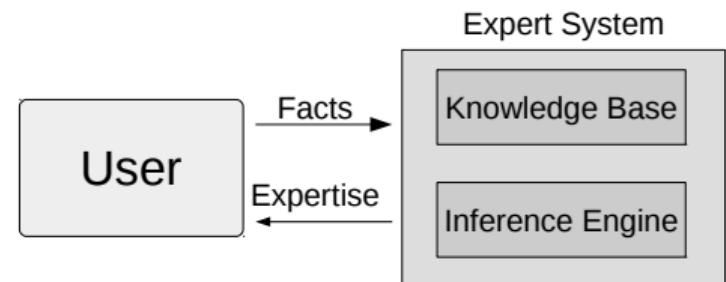
- ▶ Deklarative, dynamische Wissensbasen
- ▶ Problemlösung als Inferenzprozess
- ▶ Entwicklung von Expertensystemen
(*medizinische Diagnose, technische Fehlersuche*)

- ▶ **Subsymbolische Ebene: neuronale Netze**

- ▶ Verbindung von Neuronen zur Lösung allgemeinerer Lernprobleme.

1990er Jahre (KI-Winter)

- ▶ KI-Versprechen werden nicht eingehalten, massive Kürzungen der Finanzierung.
- ▶ Expertensystem: **Wartung** schwierig, **Wissenserfassung** zu teuer (*common-sense knowledge*).
- ▶ **Neuronale Netze**: Training extrem schwierig.

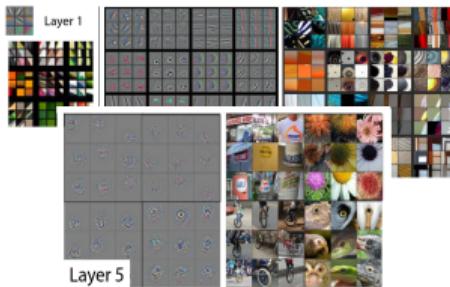




KI: Geschichte images: [3, 8, 5]

2000er/2010er - heute

- ▶ steigende Relevanz von KI-Ansätze
(insbesondere maschinellem Lernen)
- ▶ Mehr Daten
 - leistungsstärkere Modelle
 - Lösung schwierigerer Probleme.
- ▶ “Hype um Deep Learning” (seit 2012).





References |

- [1] Alan Turing Aged 16.
https://upload.wikimedia.org/wikipedia/commons/a/a1/Alan_Turing_Aged_16.jpg (retrieved: Mar 2018).
- [2] Boston Dynamics: Atlas, the Next Generation.
<https://www.youtube.com/watch?v=rVlhMGQgDkY> (retrieved: Oct 2016).
- [3] Google DeepDream robot: 10 weirdest images produced by AI 'inceptionism' and users online (Photo: Reuters).
<http://www.straitstimes.com/asia/east-asia/alphago-wins-4th-victory-over-lee-se-dol-in-final-go-match> (retrieved: Nov 2016).
- [4] UltimateAI Blog.
ChatGPT: The Tech Behind the Hype (and What It Means for Your Support).
<https://www.ultimate.ai/blog/ai-automation/chatgpt-the-tech-behind-the-hype-and-what-it-means-for-your-support> (retrieved: Jan 2023).
- [5] Data never Sleeps X.0.
www.domo.com/blog/ (retrieved: Sep 2020).
- [6] Poole David L. and Mackworth Alan K.
Artificial Intelligence: Foundations of Computational Agents.
Cambridge University Press, 2010.
- [7] Qiaozhu Mei, Yutong Xie, Walter Yuan, and Matthew O. Jackson.
A turing test of whether ai chatbots are behaviorally similar to humans.
Proceedings of the National Academy of Sciences, 121(9):e2313925121, 2024.
- [8] Matthew D. Zeiler and Rob Fergus.
Visualizing and Understanding Convolutional Networks.
CoRR, abs/1311.2901, 2013.