

Künstliche Intelligenz (Wissensrepräsentation und -Verarbeitung)

Prof. Dr. Sibylle Schwarz
HTWK Leipzig, Fakultät IM
Gustav-Freytag-Str. 42a, 04277 Leipzig
Zimmer Z 411 (Zuse-Bau)
<https://informatik.htwk-leipzig.de/schwarz>
sibylle.schwarz@htwk-leipzig.de

Sommersemester 2019

Was ist Künstliche Intelligenz?

EU-Factsheet on Artificial Intelligence:

Artificial intelligence (AI) refers to systems that show intelligent behaviour: by analysing their environment they can perform various tasks with some degree of autonomy to achieve specific goals.

*Mobile phones, e-commerce tools, navigation systems and many other different sensors constantly gather **data** or **images**. AI, particularly **machine-learning** technologies, can learn from this torrent of data to **make predictions** and **create useful insights**.*

Aussage über das derzeitige (beschränkte) Verständnis von KI

Können Maschinen denken?

Alan Turing 1950

Konkretisierung der Frage:
Können Maschinen **denken**?

zur überprüfbareren Frage:
Können Maschinen konstruiert werden, die einen
speziellen Test bestehen?

Imitation Game

Imitation Game (Alan Turing 1950):

- ▶ zwei verschlossene Räume,
in einem befindet sich **Herr A**, im anderen **Frau B**
- ▶ eine Person C (Frager) stellt Fragen, A und B antworten
- ▶ Kommunikation über neutrales Medium,
an welchem das Geschlecht nicht erkennbar ist,
- ▶ C soll herausfinden, in welchem der Räume Frau B ist
- ▶ Herr A versucht, C irrezuführen
- ▶ Frau B kooperiert mit C

Herr A besteht den Test, wenn ihn C für Frau B hält.

Wie erkennt man Intelligenz: Turing-Test

Turing-Test 1950: verschiedene Versionen des Imitation Game

- ▶ A ist Machine statt Mann (B Person beliebigen Geschlechts)
- ▶ verschiedene Kooperationsverhalten von A und B

Vorschlag zur Bewertung natürlichsprachlicher
Kommunikationsfähigkeiten

Beginn koordinierter Forschung zur Künstlichen Intelligenz

John McCarthy

Programmiersprachen

Marvin Minsky

Kognitionswissenschaft

Claude Shannon

Informationstheorie

stellten 1955 die Vermutung auf, dass

„jeder Aspekt des Lernens oder jedes anderen Ausdrucks von Intelligenz prinzipiell so präzise beschrieben werden kann, dass sich eine Maschine konstruieren lässt, die ihn simuliert.“

Begriff Künstliche Intelligenz

McCarthy formulierte das Ziel,

„herauszufinden, wie man Maschinen konstruiert, die

- ▶ natürliche Sprache benutzen,
- ▶ Abstraktionen und Begriffe entwickeln,
- ▶ Aufgaben lösen, die (bis dahin) nur Menschen lösen konnten,
- ▶ sich selbst verbessern.“

und prägte dafür den Begriff **Künstliche Intelligenz**.

Beginn koordinierter Forschung zur Künstlichen Intelligenz

1956: erste Konferenz zur Künstlichen Intelligenz

Dartmouth Summer Research Project on Artificial Intelligence

Themen:

- ▶ Berechnungsmodelle in Computern
- ▶ Kommunikation mit Computern in natürlicher Sprache
- ▶ Neuronale Netzwerke
- ▶ Berechenbarkeitstheorie
- ▶ Selbst-Verbesserung
- ▶ Abstraktionen
- ▶ Zufälligkeit und Kreativität

Forschung zur Künstlichen Intelligenz

Momentaufnahme 2006:

Dartmouth Artificial Intelligence Conference: The Next Fifty Years

Themen:

- ▶ Modelle des (menschlichen) Denkens
- ▶ Neuronale Netzwerke
- ▶ (Maschinelles) Lernen und Suchen
- ▶ Maschinelles Sehen
- ▶ Logisches Schließen
- ▶ Sprache und Kognition
- ▶ KI und Spiele
- ▶ Interaktion mit intelligenten Maschinen
- ▶ Ethische Fragen und zukünftige Möglichkeiten der KI

Ansätze intelligenter Systeme

- ▶ Simulation menschlichen **Verhaltens**
(Verständnis und eigenes Denken nicht notwendig)
Modellierung von Kognition,
statistische Verfahren, Training mit vielen Fällen
Getroffene Entscheidungen werden nicht begründet.
schwache künstliche Intelligenz

- ▶ Simulation des menschlichen **Denkens**
(Verständnis und eigenes Denken notwendig)
Modellierung des Denkens
logisches Schließen, Abstraktion
Jede Entscheidungen kann nachvollziehbar begründet werden.
starke künstliche Intelligenz

Kritik am Turing-Test

Kritik:

schwache KI genügt, um den Turing-Test zu bestehen

1966: Maschinelle Psychotherapeutin Eliza besteht Turing-Test

Searle (1980) Chinese-Room-Argument:

eine (nicht chinesisch verstehende) Person B in einem verschlossenen Raum mit einem (riesigen) Regelbuch mit chinesischen Fragen und passenden Antworten.

- ▶ A stellt Fragen, B antwortet.
- ▶ B antwortet mit Hilfe des Buches immer passend, ohne die Frage verstanden zu haben.

These: (anscheinend) intelligentes Verhalten ist noch

keine Intelligenz, wenn Verständnis fehlt (Ansatz der starken KI)

außerdem: praktisch nicht umsetzbar

Aktuelle Entwicklung

starker Fortschritt einiger KI-Methoden in den letzten 10 Jahren aufgrund der Entwicklung bei

- ▶ Computertechnik: Parallelrechner, GPU (70% Einfluss)
- ▶ Speichermöglichkeit großer Datenmengen, Verfügbarkeit großer strukturierter und annotierter Datenmengen (20%)
- ▶ neue Typen künstlicher neuronaler Netze, bessere Algorithmen (10%)

sowie starkes Medieninteresse an bestimmten Erfolgen, z.B.

- ▶ 1997 Deep Blue gewinnt gegen amtierenden Weltmeister
- ▶ 2011 Watson schlägt zwei Meister in Quizshow Jeopardy!
- ▶ 2012 erste Zulassung eines autonomen Fahrzeugs für den Test auf öffentlichen Straßen
- ▶ 2016 AlphaGo schlägt Go-Meister
- ▶ ...

führte zum aktuellen Aufblühen der KI-Euphorie

Leistung aktueller (statistischer) KI-Systeme

nahe und teilweise über den menschlichen Fähigkeiten z.B. bei

- ▶ Erkennung von Objekten in Bildern
- ▶ Einordnung / Klassifikation von Objekten und Situationen
- ▶ Reaktion auf klar erkannte Situationen
- ▶ strategischen Spielen mit endlichem Zustandsraum
z.B. Schach, Go

prinzipielle Herausforderungen:

- ▶ Zuverlässigkeit, Sicherheit
- ▶ Begründung, Erklärung

Schwächen aktueller (statistischer) KI-Systeme

KI derzeit noch weit von menschlichen Fähigkeiten entfernt bzgl.

- ▶ Erkennung der eigenen Grenzen
- ▶ Intuition
- ▶ Aufstellen und Überprüfen sinnvoller Annahmen bei unvollständig vorhandener Information
- ▶ Lernen ohne vorheriges Training mit großen Mengen (manuell) annotierter Daten
- ▶ Übertragen von Wissen zwischen verschiedenen Anwendungsbereichen
- ▶ Kombination verschiedener Methoden
- ▶ Schließen bzgl. rechtlicher und moralischer Bezugssysteme, mentaler Modelle

Von Daten zur Intelligenz

Umwelt		Reize, Eindrücke
Agent	Wahrnehmen, Beobachten	Daten
	Erkennen, Verstehen	Information
	Anwenden, Können	Wissen
	Lernen	Wissenserwerb (Intelligenz?)
	Reflektieren, Begründen, Erkennen der Grenzen, Verstehen	Intelligenz

Beispiel: Daten, Information, Wissen, Intelligenz

Daten Darstellungsform (Syntax)
Zeichenketten, Bilder, Ton, ... (z.B 39.7)

Information Bedeutung der Daten (Semantik)
in einem bestimmten Kontext (z.B.
Körpertemperatur= 39.7°)

Wissen Information mit einem Nutzen,
trägt zur Lösung eines Problemes bei,
Nutzen abhängig von vorhandenem Kontextwissen
z.B. Kontext: Körpertemperatur > 39.0° ist Fieber,
bei Fieber ist Fieberbehandlung notwendig,
mögliche Fieberbehandlungen z.B. Wadenwickel,
Medikamente

Wissenserwerb selbständige Informationsgewinnung über Gründe,
Nebensymptome, Therapien für
Körpertemperatur-Unregelmäßigkeiten

Intelligenz Diagnose und Auswahl aus Therapie-Alternativen speziell
für die zu behandelnde Person durch Abwägung der zu
erwartenden Wirkungen, ggf. Überweisung zu Spezialisten

Logische / regelbasierte KI-Methoden

Wissensrepräsentation: formale Beschreibung von Umwelt (Randbedingungen) und Problem

Problemlöseverfahren: zur Lösung vieler Probleme anwendbares Standardverfahren (z.B. logisches Schließen)

Beispiele:

- ▶ Entscheidungsbäume und -tabellen
- ▶ Regelsysteme, Logiken, logisches Schließen
- ▶ Constraint-Systeme und -Löser
- ▶ deklarative Programmierung (logisch, funktional)
- ▶ fallbasiertes Schließen (durch Analogien)
- ▶ Simulation

typische Anwendungen klassischer KI-Methoden:

- ▶ Entscheidungsunterstützung (z.B. Finanzwirtschaft)
- ▶ Diagnosesysteme (z.B. in Medizin, Technik)
- ▶ Bewegungs- und Ablaufplanung

Statistische KI-Methoden

„Soft-Computing“ oft besser geeignet für Probleme

- ▶ die unvollständig beschrieben sind,
- ▶ die keine eindeutige Lösung haben,
- ▶ für die keine effizienten Lösungsverfahren bekannt sind, usw.

einige Ansätze:

- ▶ künstliche neuronale Netze
- ▶ evolutionäre Algorithmen
- ▶ Schwarm-Intelligenz, Ameisen-Algorithmen
- ▶ Fuzzy-Logiken, probabilistische Logiken

Inhalt der LV

- ▶ heuristische Suche
- ▶ Spielbaum-Suche
- ▶ Logisches Schließen
- ▶ Planen
- ▶ Unscharfes / probabilistisches Schließen
- ▶ Bayes-Netze
- ▶ Künstliche Neuronale Netze
- ▶ Kausalität (Zusammenhang von Ursache und Wirkung)
- ▶ Modellierung ethischer Prinzipien (mentale Modelle)

Organisation

6 ECTS

Präsenzzeit 56 h, Vor- und Nachbereitungszeit 124 h

- ▶ wöchentlich eine Vorlesung
- ▶ wöchentlich ein Seminar (Reading group)
zu aktuellen Forschungsbeiträgen
Literatur wird begleitend bekanntgegeben
PVL und Notenbonus
- ▶ Klausur (90 min)
zum Inhalt von Vorlesung und Seminar

Aufgaben für Seminar am 12. April 2019

Lesen Sie die KI-Ethikleitlinien der EU
Ethics guidelines for trustworthy AI

https:

[//ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai](https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai)

- ▶ Grundprinzipien
- ▶ Pläne zur Umsetzung
- ▶ Standards, Überprüfung
- ▶ ...

Informieren Sie sich über die Autoren (52 Expertinnen und Experten)
High-Level Expert Group on Artificial Intelligence

Vergleichen Sie die KI-Ethikleitlinien mit dem Entwurf
https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=57112
(optional)