

Künstliche Intelligenz: Wo stehen wir auf dem Weg zur „Superintelligenz“?

Gerhard Friedrich

Universität Klagenfurt

Fakultät für Technische Wissenschaften

Institut für Artificial Intelligence and Cybersecurity

Österreich

Artificial Intelligence (AI) ist wieder in den Medien

The State of Artificial Intelligence, World Economic Forum in Davos 2016



Andrew Moore:

- CMU, statistical learning

Was Computer können:

- Computers understand what is in the document

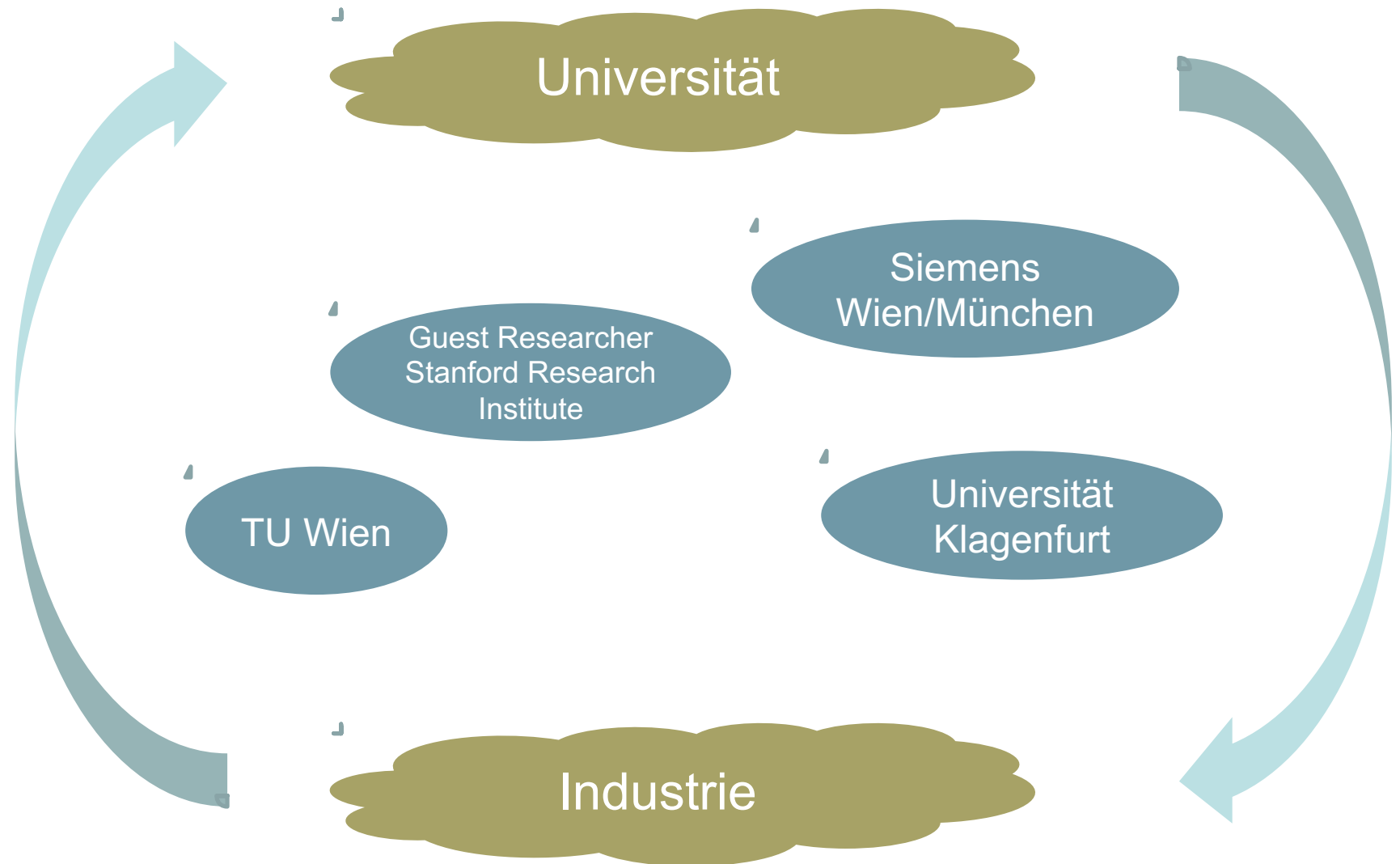
Stuart J. Russell:

- UC Berkeley, AI

Was Computer NICHT können:

- Computers understand little or nothing of the content of the document

Fachlicher Hintergrund



Weit verbreitete große Unsicherheit

Manche (AI) Wissenschaftler sagen:

- „Nur wenn Firmen Künstliche Intelligenz verwenden, können sie wettbewerbsfähig bleiben.“
- „Die Menschheit wird durch Künstliche (Super)Intelligenz bedroht.“
- „Auf Basis von Neuronalen Netzwerken wurde die allgemeine Problemlösungsmaschine gefunden.“

Manche AI
Anwendungen sind
sehr erfolgreich



Wo stehen wir?
Was bedeutet das für
meine Organisation?

Ziel: Klarheit schaffen

Artificial Intelligence

Machine Learning

Neural Networks

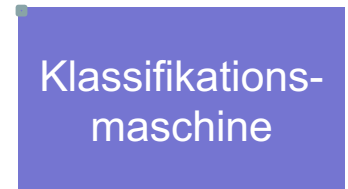
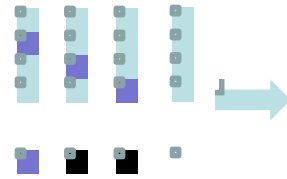
Knowledge
Representation and
Reasoning

Supervised Learning

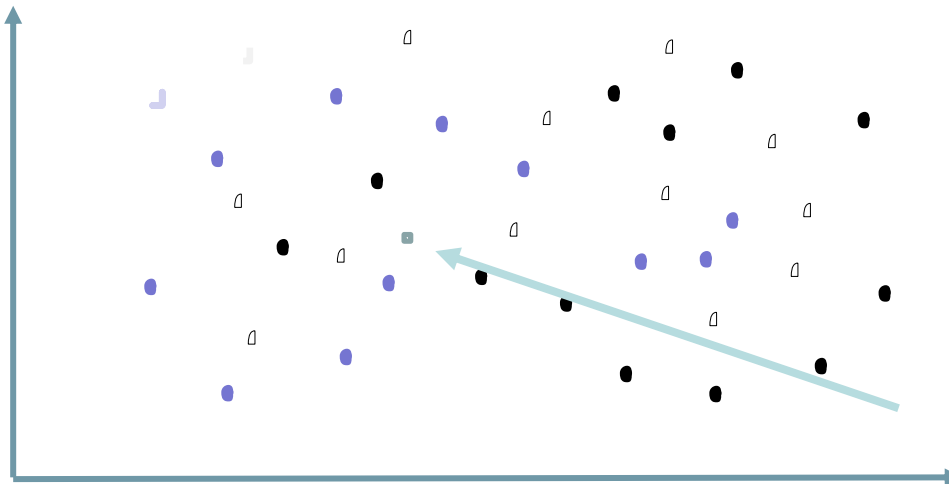
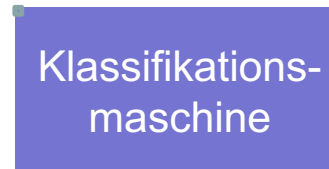
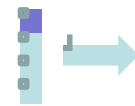
Trainingsbeispiele:

Input-Vektoren

Gewünschte Klassifikation

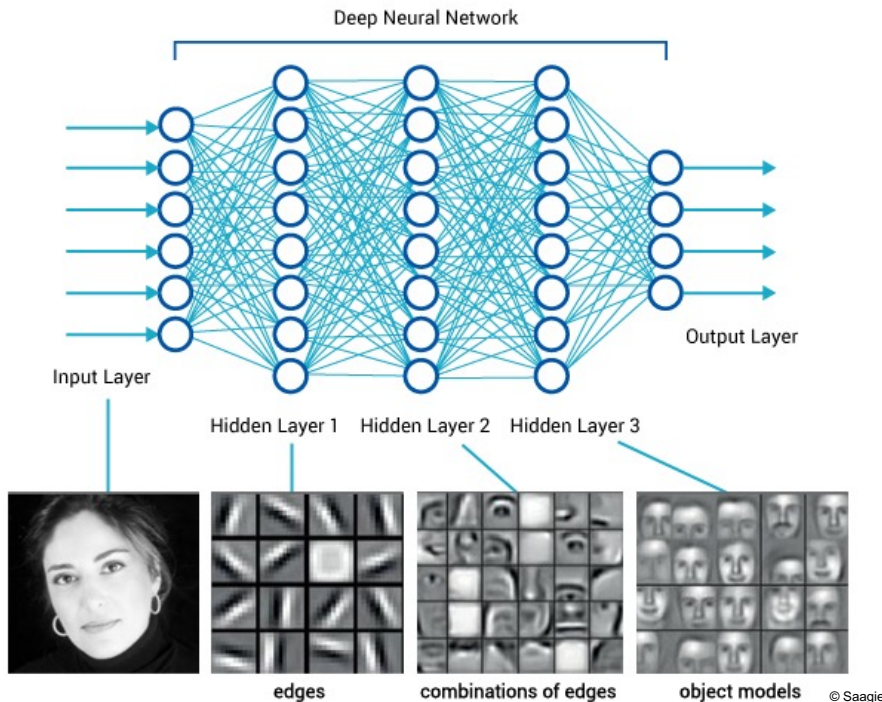
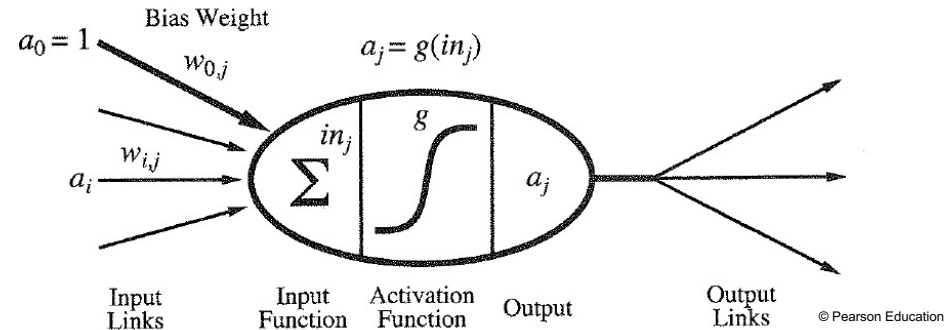
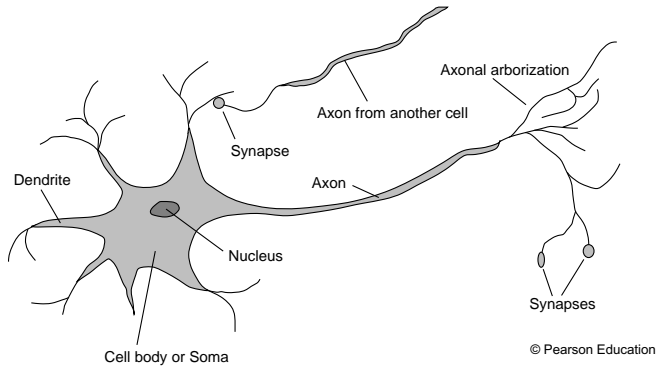


Anwendungsdaten:



Neuer Input-Vektor:
Farbe?

Deep Learning / Deep Neural Networks



Lernen = Anpassung der Gewichte

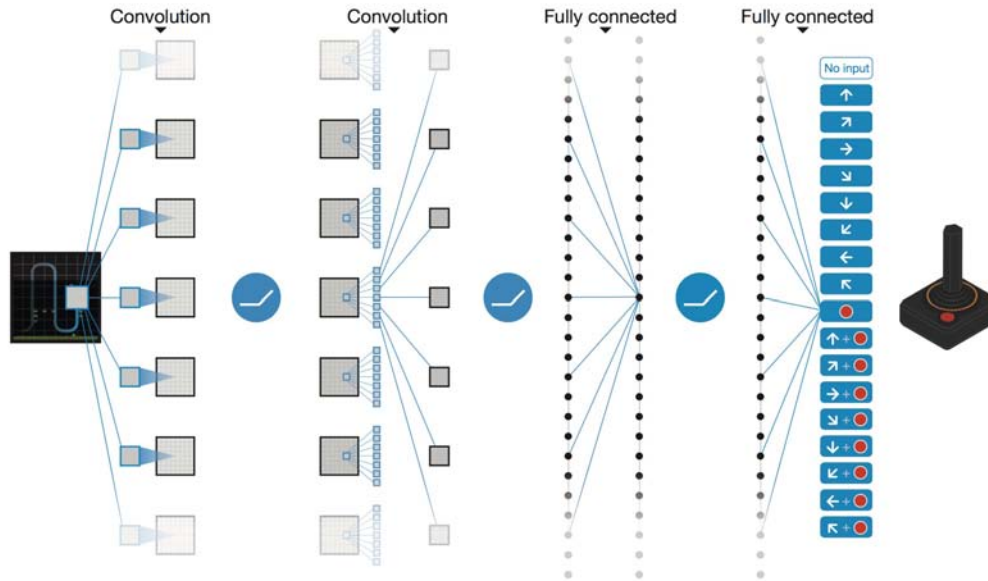
Erfolgsfaktoren:

- Neue Architekturen und Algorithmen
- Große Mengen an Daten
- Rechnerleistung und Parallelisierung

Maschinelles Lernen von Atari Spielen

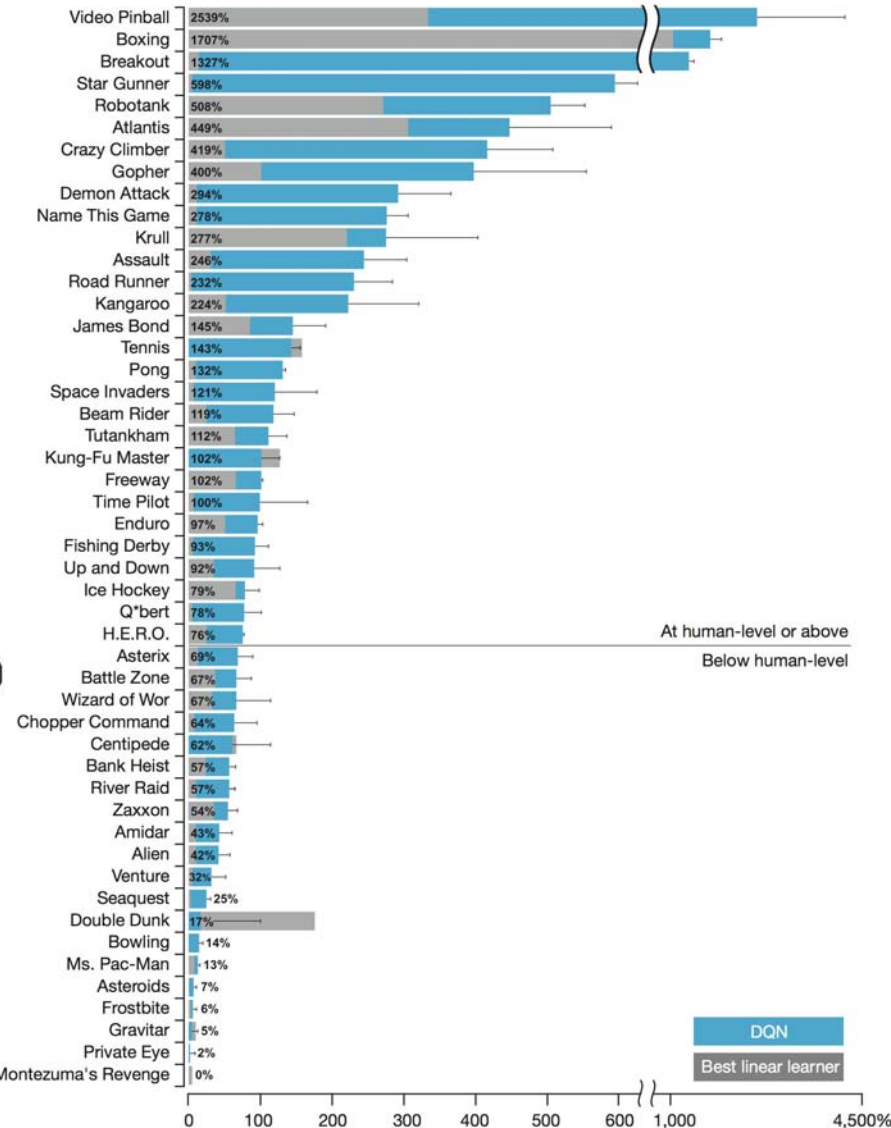
Human-level control through deep reinforcement learning

Volodymyr Mnih et al., NATURE Vol 518, 2015



© Macmillan Publishers Limited

Planungsstrategie über einen längeren Zeitraum notwendig



© Macmillan Publishers Limited

Erfolge der neuronalen Netze (Ausschnitt)

1951: erste Maschine auf Basis von künstlichen neuronalen Netzen

1969: Der lange Marsch durch die Wüste der Geringschätzung beginnt

Anfang der 2000er industrieller Einsatz von Deep Learning, um Handschrift auf Schecks zu erkennen

2012: Deep Learning gewinnt den „Large Scale Visual Recognition Challenge“ und dominiert den Wettbewerb

2014: Learning of Human-Level Control in Atari 2600 Games

2015: dramatische Verbesserung der Speech Recognition durch Google

2015: AlphaGo (Deep Learning) gewinnt gegen europäischen Go Champion

2016: Signifikante Verbesserung der maschinellen Übersetzung von natürlicher Sprache durch Google

XXX: ähnliche Erfolgsmeldungen

Derzeitige Einschränkungen

Fehlleitende Eingabe



© Goodfellow et al., Google, Explaining and Harnessing Adversarial Examples

„Panda“ 57,7%
Konfidenz

Erzeugte
Störung

„Gibbon“ 99.3%
Konfidenz

- Fehlklassifikationen:
keine Garantie bzgl. Korrektheit und Vollständigkeit
- Fehlgeleitetes Lernen über die Zeit
z.B. Verherrlichung von Gewalt durch Chatbots

Common Sense Reasoning / Naive Physics

Text verstehen (Winograd Schema questions)

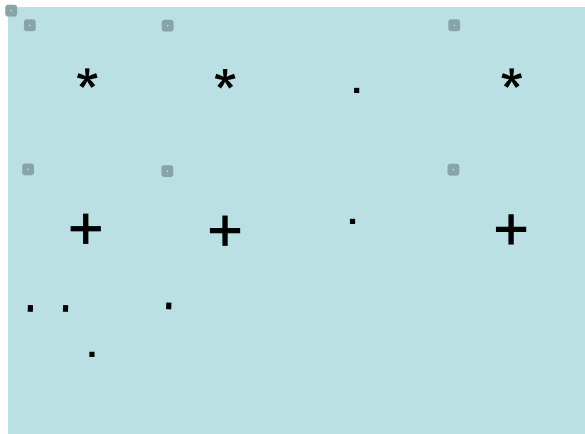
Ein alter Mann ging mit seinem jungen Jagdhund auf einem Feldweg. Plötzlich lief ein Hase über den Feldweg. Er hetzte ihm hinterher.

Frage: Wer hetzte wem hinterher?

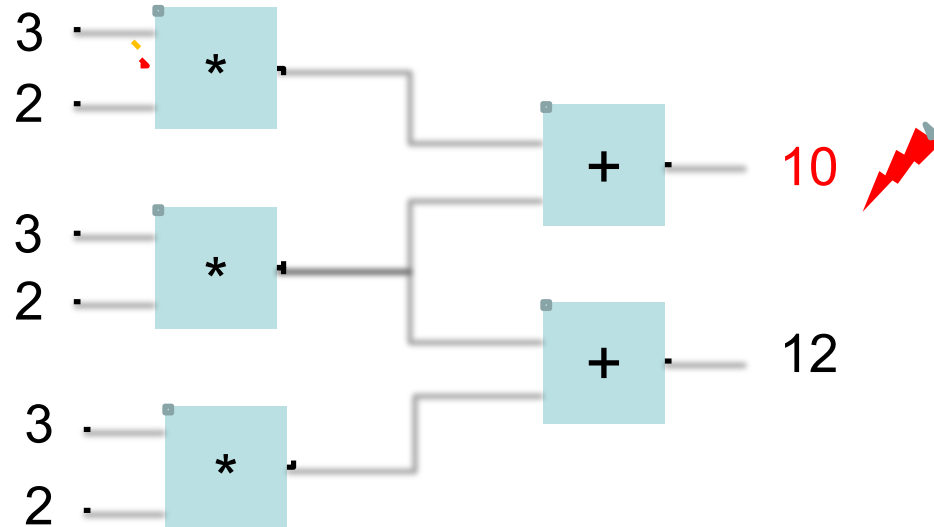
Intelligente, breit einsetzbare Agenten benötigen
Verständnis über die Welt.

Modellbasiertes Schließen

Phy. Gesetze & Teile



Modell



- Reasoning from first principles, function-based
 - Realisierung mittels Logik und automatisiertem Schließen
 - Methoden der expliziten Wissensrepräsentation und -verarbeitung
 - vollständig und korrekt
 - Anwendungen: z.B. automatisierte Diagnose, Reparatur, Konfiguration

Automatisches Design, Konfiguration, Planung

Modellbasiertes Schließen / Spezifikation des Problems

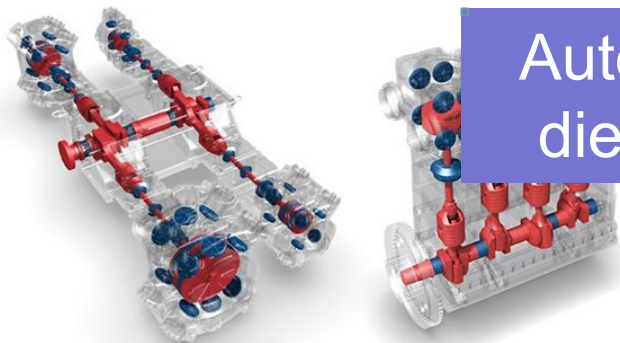


© Dr. G Schmitz CC BY-SA 3.0



© Siemens

Diagnose, Konfiguration, technische
und logistische Planung mit
beweisbarer Qualität



© Siemens

Automatisierung von Tätigkeiten,
die hohe Qualifikation erfordern



© Infineon

Stärken/Schwächen von AI-Methoden



Artificial Neural Networks:

- Stärke: automatische Erzeugung der Funktion durch Beispiele
- Architektur und Gewichte sind die Beschreibung der Funktion
- Rasche Berechnung des Outputs
- genügend viele Beispiele
- genügend Rechenleistung für das Lernen
- „richtige“ Architektur finden
- keine Garantien bzgl. Qualität der Lösung
- schwach bzgl. automatisiertem Schlussfolgern
- Adaption der Funktionalität über Beispiele, Wiederverwendung problematisch
- keine Erklärungen

Modelle / Gesetze:

- Meist händische Spezifikation der Funktion durch Regeln/Gesetze/Einschränkungen
- Symbole und Sätze vom Menschen interpretierbar, Erklärungen möglich
- Stärke: automatisiertes Schlussfolgern
- Garantien bzgl. der Qualität der Lösung
- Spezifikation kann schwierig werden
- unbefriedigendes Lernen der Modelle
- Laufzeit und Speicherverbrauch bei „großen“ Anwendungen
- Adaption der Funktionalität muss in der Spezifikation (hohe Flexibilität) berücksichtigt werden oder durch Änderung der Spezifikation

Design für spezialisiertes Anwendungsgebiet

Warum dieser Widerspruch?

The State of Artificial Intelligence, World Economic Forum in Davos 2016



Andrew Moore:

- CMU, statistical learning
- Computers understand what is in the document

Machine Learning funktioniert teilweise sehr gut: z.B. Übersetzung von Texten

Stuart J. Russell:

- UC Berkeley, AI
- Computers understand little or nothing of the content of the document

AI kann (Fach)texte nicht in Wissen überführen, das für breites automatisierte Problemlösen notwendig ist

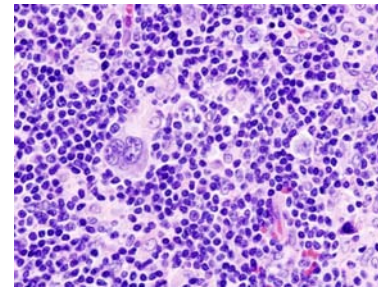
Computersysteme können immer mehr Tätigkeiten verrichten

einfach für den Menschen /
schwierig für die Maschine



© CC BY-SA 2.0

schwierig für den Menschen /
„einfach“ für die Maschine



© CC BY-SA 3.0, <https://commons.wikimedia.org>

Pros:

- Verlängerung des Lebens
- Vermeidung von Unfällen
- mehr Möglichkeiten zur Gestaltung des Lebens
- Produktivitätssteigerung ...

Herausforderungen:

- Analyse und Beeinflussung von Menschen (Gefahr für Demokratie?)
- automatisierte Gewaltanwendung
- Umfangreiche Veränderungen in der Arbeitswelt ...

Der Weg zur Superintelligenz ist noch weit,
aber es geht voran.

Vielen Dank!