

# Die Grenzen der KI: Können Maschinen denken?



Eine Argumentation von Searle (1986, 1990) und  
Churchland und Churchland (1990)

Proseminar „Denkmaschinen“  
Sommersemester 2005

Vorge stellt von  
Matthias Bonnesen  
und Simon Harder

# Gliederung

---

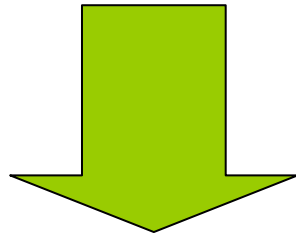
- I. Grundfragen der klassischen KI
- II. Präsentation der Autoren
- III. Das Körper/Geist-Problem (Searle)
- IV. Können Maschinen denken? (Searle)
- V. Churchlands Einwände gegen Searle
- VI. Searles Antwort auf die Churchlands
- VII. Reaktion der Churchlands
- VIII. Zusammenfassung und Schlussteil

# Die Grundfrage der klassischen KI

---

***Anfang der 50er:***

Können Maschinen denken?



Können Maschinen,  
die Symbole anhand von Regeln  
verarbeiten, denken?

# Die Grundfrage der klassischen KI

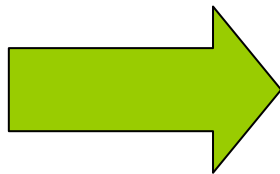
---

*Church:*

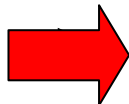
Jede wirklich berechenbare Funktion ist rekursiv berechenbar.

*Turing:*

Jede rekursiv berechenbare Funktion ist in endlicher Zeit von einer maximal einfachen Maschine berechenbar.



Die Formel für das menschliche Verhalten fehlt leider noch.



Ziel der „klassischen“ KI

# Die Anfänge der klassischen KI

---

## *Immer schnellere Computer:*

Arithmetische Probleme

Algebraische Probleme

Taktische Probleme

Einfache Dialoge

Schach

...

### *Ungeklärt:*

- Konstruktion: Computer  $\neq$  Gehirn
- Nonverbales Hintergrundwissen

# Die Anfänge der klassischen KI

---

## Schnellere Computer – größere Probleme



Computer



Gehirn

# John R. Searle

---

- Geboren 1932 in Denver/Colorado
- Studierte in Oxford Philosophie/Linguistik u.a. bei John L. Austin, dem Vater der Sprechakttheorie
- Seit 1959 Professor für Philosophie in Berkeley
- gilt heute selbst neben Austin als der bedeutendste Sprechakttheoretiker
- Weitere Arbeitsgebiete: Semantik, Ethik, Erkenntnistheorie, Philosophiegeschichte



# Churchland

Patricia Smith

Paul M.

---

- 1969 bis 1984 zuerst Assistenz-, später Vollprofessur an der Universität von Manitoba
  - Seit 1984 an der Universität von Kalifornien in San Diego
  - Spezialisiert auf die Philosophie des Geistes und der Wissenschaft, besonders der Neurowissenschaften
- Seit 1966 an verschiedenen Universitäten in den USA und Kanada.
  - Ab 1984 Professor der Philosophie an der Universität von Kalifornien



# Eliminative Materialism (engl.)

---

- Kritik an der „volkstümlichen“ Psychologie
- Bewusstsein eher eine „Nebenwirkung“
- Forderung nach einer „erwachsenen“  
*Neurowissenschaft*
- Erforderliche Konzepte müssen erst  
gefunden werden

*Quelle:*

Dictionary of Philosophers: Eliminative Materialism

# Das Körper/Geist-Problem (Searle)

---

- In welcher Beziehung steht Geist zum Hirn bzw. zum Körper?
- Wo ist im materiellen Universum Platz für Bewusstsein?
- Wie kann aus etwas Geistigem materielle Veränderung hervorgehen?

# Das Körper/Geist-Problem

---

Zum Körper/Geist-Problem gehören:

- Verursachung
- Bewusstsein
- Intentionalität
- Subjektivität

# Das Körper/Geist-Problem

---

- Jedes geistige Phänomen ist physiologisch, d.h. von Vorgängen im Gehirn verursacht
- Geist und Hirn sind nicht trennbar
- Geist ist eine Eigenschaft des Hirns
- Gedanken sind nicht schwerelos, sondern biochemische Prozesse

# Können Maschinen denken?

---

- Können Maschinen bewusste Gedanken haben wie Menschen?
- Annahme: Maschine=physikalisches System, das gewisse Funktionen ausführt
- Folgerung: Menschen sind Maschinen einer besonderen, biologischen Art

# Können Maschinen denken?

---

Übliche Interpretation der Ausgangsfrage:

- (a) Kann eine Maschine denken, indem ihr ein Computerprogramm implementiert wird?
- (b) Stellt das Programm selbst das Denken bzw. den Verstand/Geist dar?

# Stellt das Programm selbst das Denken dar?

---

Für einige Wissenschaftlern in der KI gilt:

- (1) Ein Programm mit korrektem Input/Output verhalten denkt
- (2) Durch die Programmierung wird Verstand geschaffen („creating minds“)
- (3) Erfolg oder Misserfolg durch Turing-Test nachweisbar

# Der starke KI vs. der schwache KI

## Ansatz nach Searle

---

Starke KI: Programm, das menschliche Kognition simuliert, ist kein Modell, sondern echter Verstand/Geist („mind“)

Schwache KI: Computermodelle sind nützlich, um u.a. menschliche Kognition zu untersuchen und zu simulieren.



# Das chinesische Zimmer

---

In einem Zimmer befinden sich:

- Person ohne Kenntnisse des Chinesischen
- Körbe mit chinesischen Symbolen
- in der Muttersprache der Person geschriebenes Regelbuch, wie diese chinesischen Symbole sinnvoll zusammengefügt werden

# Axiom 1: Computerprogramme sind formal

- Computer verarbeiten/manipulieren Informationen, indem sie:
  - (a) Daten geeignet kodieren
  - (b) auf die Daten präzise definierte Regeln anwenden
- Programme und Symbole sind abstrakt, ohne physikalische Eigenschaften, können im Gegensatz zu unserem Gehirn auf ganz unterschiedlichen Maschinen implementiert werden
- Symbole werden ohne Semantik verarbeitet

## Axiom 2: Menschlicher Verstand hat mentalen Inhalt (Semantik)

- Gedanken, Wahrnehmung, Verständnis haben mentalen Inhalt
- Sprachliches Verstehen benötigt immer auch einen semantischen Rahmen, in dem Wörtern eine Bedeutung zugeordnet wird

# Axiom 3: Syntax allein beinhaltet weder Semantik, noch reicht sie aus, um Semantik zu erschließen

- ❑ Manipulation von Symbolen garantiert nicht deren Verständnis
- ❑ Zu unterscheiden zwischen:
  - (a) formalen Elemente ohne inhärente Bedeutung
  - (b) formalen Elemente mit inhärenter Bedeutung

Schlussfolgerung 1:

Programme beinhalten weder, noch stellen sie Verstand dar

- Schlussfolgerung 1 = starke KI Ansatz ist falsch

# Was Searle nicht zeigen wollte

---

- Ein Computer kann nicht denken
- Nur biologische Systeme können denken
- Menschen werden niemals Systeme erschaffen, die denken können

# Kritik an der Kritik (Churchland)

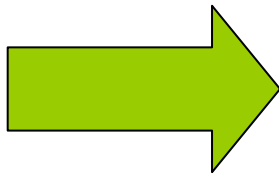
---

*Fragestellung der klassischen KI:*

Die Formel für das menschliche Verhalten finden.

*Searle:*

„Programme sind weder grundlegend, noch ausreichend für Geist.“ (Searle 1990, S.21)



Searle argumentiert nur mit epochenabhängigem „logischen“ Denken!

# Das leuchtende Zimmer

---

## *Searle*

Axiom 1:

Computerprogramme  
sind formal (syntaktisch).

Axiom 2: Menschlicher  
Verstand hat mentalen  
Inhalt (semantisch).

Axiom 3: Syntax allein ist  
weder grundlegend, noch  
ausreichend für Semantik.



Folgerung:

Computerprogramme sind  
weder grundlegend, noch  
ausreichend für Geist.

## *Die Churchlands*

Axiom 1: Elektrizität und  
Magnetismus sind Kräfte.

Axiom 2: Die wesentliche  
Eigenschaft von Licht ist  
Leuchtkraft.

Axiom 3: Kräfte an sich sind  
weder grundlegend, noch  
ausreichend für Leuchtkraft.



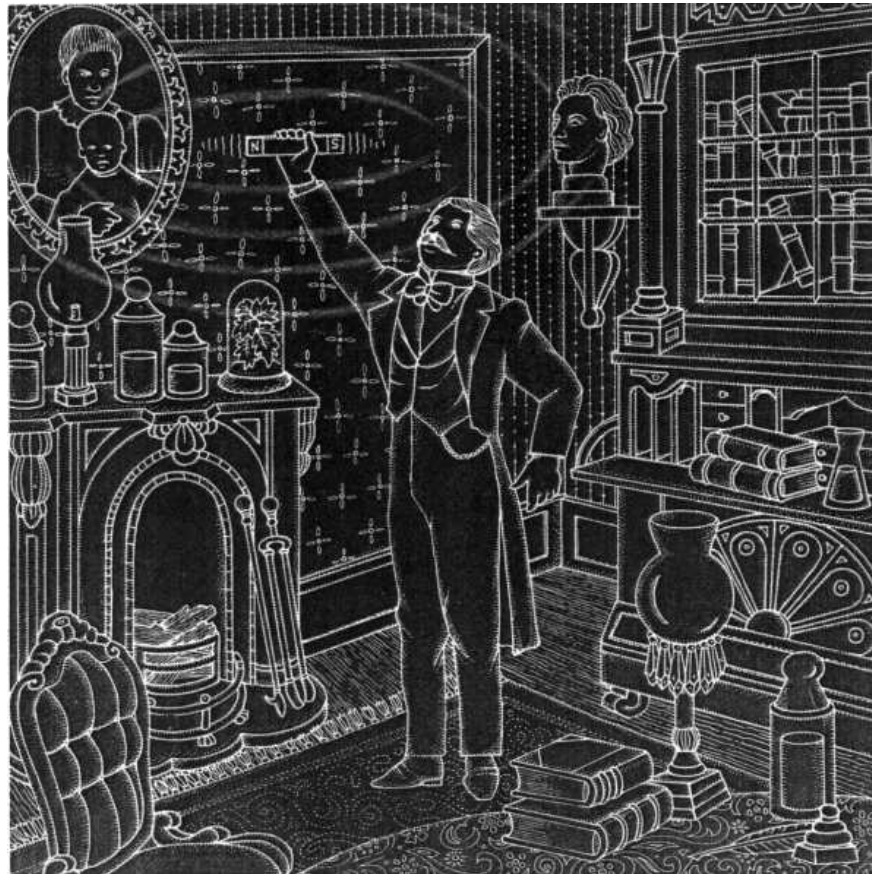
Folgerung: Kräfte an sich sind  
weder konstitutiv, noch  
ausreichend für Licht.



# Das leuchtende Zimmer

---

James Clerk Maxwell hätte ein harte Zeit gehabt:



# Das leuchtende Zimmer

---

## ***Maxwells einzig mögliche Antworten:***

- Axiom 3 nimmt die Frage vorweg
- Das Experiment zeigt keine neuen Eigenschaften des Lichts
- Ein Forschungsprogramm zur weiteren Erkundung des Lichts wird benötigt

**= Antwort der klassischen KI**  
(nach den Churchlands)

# Die Meinung der Churchlands

---

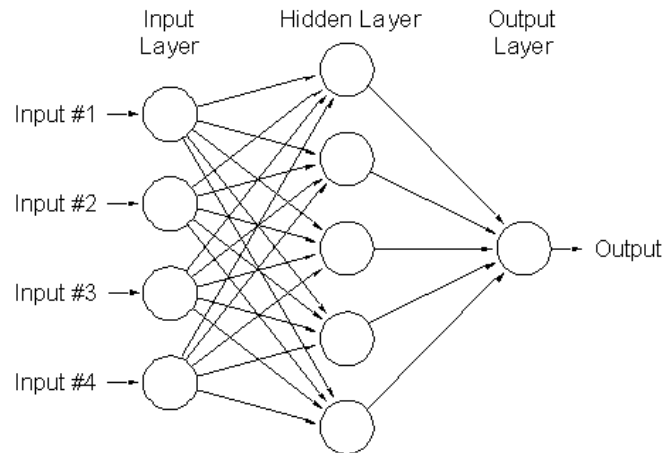
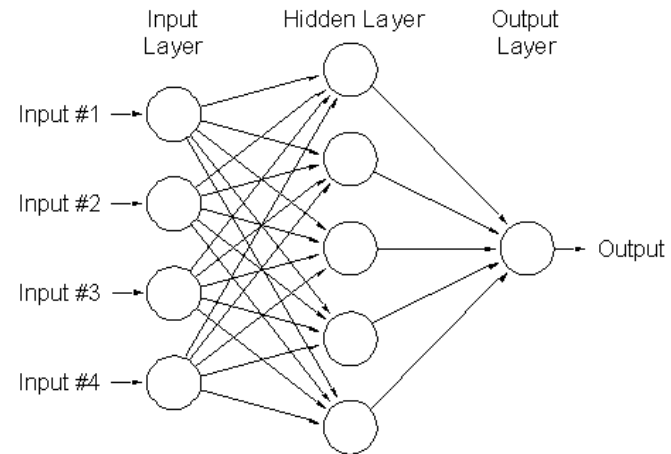
*Also können symbolverarbeitende  
Maschinen bewusste Intelligenz  
hervorbringen?*

**Nach Meinung der  
Churchlands:  
Nein.**

# Die Meinung der Churchlands

## ***Nervensysteme:***

- Massiv parallele Architekturen
- Sehr einfach kleinste Einheit
- Bidirektionale Verbindungen



# Die Meinung der Churchlands

---

## ***Die Vorteile dieser Architektur:***

- 1) Viele kleine Operationen gleichzeitig
- 2) Fehlertoleranz
- 3) Speicherung von Information

## ***Schwäche dieser Architektur:***

Rekursive Funktion auf kleinen Input

# Auch künstliche neuronale Netze/parallele Computerarchitekturen können nicht denken

---

- ❑ Parallele Architekturen können Aufschluss über unsere Gehirnfunktionen liefern
- ❑ Dennoch stellen sie kein Denkprozess dar, da auch hierbei keine Semantik, nur formale Symbolmanipulation
- ❑ Funktionen auf parallelen Architekturen können auch auf traditioneller serieller Maschine laufen

# Axiom 4: Gehirne machen Verstand aus

- ❑ Simulation von Kognition produziert nicht die Effekte von neurobiologischer Kognition
- ❑ Simulation von Gehirnfunktion ist genauso Modell wie Simulation von Entstehung von Autoabgasen oder Simulation von Verdauung

Schlussfolgerung 2: Ein System, das Verstand erzeugt, muss eine verursachende Wirkung haben wie die unseres Gehirns

- ▣ Nur die Systeme können denken, die die gleichen verursachenden Wirkungen wie unser Gehirn haben



Schlussfolgerung 3: Die Art, wie menschliche Gehirne mentale Phänomene produzieren, kann nicht allein dadurch erzielt werden, dass ein Computerprogramm läuft.

# Gegenargument: das gesamte Zimmer versteht Chinesisch

---

Argument nicht plausibel weil:

- (a) Symbolmanipulation erlaubt keinen Zugang zur Bedeutung der Symbole
- (b) Person könnte Symbole und Regelbuch auswendig lernen und im freien die Aufgabe erfüllen. Folge: Kein Buch, kein Zimmer, Keine Körbe mit Symbolen, aber: Verstünde Person nun Chinesisch?

## Die Churchlands: Chinesisches Zimmer entspricht „Licht sei nicht elektromagnetisch“

- Analogie unangemessen, da Licht/Elektromagnetismus verursachende Kräfte hat („causal powers“)
- Formale Symbole haben diese Kräfte nicht

# Churchlands: Searle argumentiert an der eigentlichen Frage vorbei

---

- Warum sollten Symbole nicht mentalen Inhalten entsprechen?
- Weil Symbole/Symbolmanipulationen im Zimmer uminterpretiert werden können, z.B.
  - (a) Für Züge im Schachspiel
  - (b) Als Vorhersage für Aktienkurse
  - (c) ...

# Trugschlüsse in der KI

---

Für alle eindeutig:

- (a) Modell von Wassermolekülen ist nicht nass
- (b) Computersimulation von Verdauungsprozessen verdaut nichts

Warum sollte dann Simulation von Denken tatsächlich denken und somit Geist/Verstand darstellen?

Turing-Test und „residualer“ Behaviorismus betrachten nur Output-Verhalten

# Reaktion der Churchlands

---



= 10.000 Erden oder  $10^{14}$  Menschen

# Schlussstein: Können Maschinen denken?

---

- Antwort hängt von Definition des Denkens/bewussten Denkens ab
- Nach Searles Definition kann eine Maschine in der Tat nicht denken
- Ist man eher Output-orientiert, kann man die Aktivität der Maschine vielleicht als denken oder quasi-denken bezeichnen

# Quellen

---

Brockhaus, Der (2005): Der Brockhaus multimedial 2005 premium. Mannheim: Bibliographisches Institut & F.A. Brockhaus AG.

Churchland, Paul (1990): „Could a Machine Think?“ in: Scientific American, Januar 1990, S. 26 ff.

Dictionary of Philosophers (ohne Jahr): Patricia Churchland. Online-Dokument, verfügbar über [http://www.explore-biography.com/philosophers/P/Patricia\\_Churchland.html](http://www.explore-biography.com/philosophers/P/Patricia_Churchland.html) [Letzter Zugriff 03.06.2005]

Dictionary of Philosophers (ohne Jahr): Paul Churchland. Online-Dokument, verfügbar über [http://www.explore-biography.com/philosophers/P/Paul\\_Churchland.html](http://www.explore-biography.com/philosophers/P/Paul_Churchland.html) [Letzter Zugriff 03.06.2005]

Dictionary of Philosophers (ohne Jahr): Eliminative materialism. Online-Dokument, verfügbar über [http://www.explore-beliefs.com/philosophy/E/Eliminative\\_Materialism.html](http://www.explore-beliefs.com/philosophy/E/Eliminative_Materialism.html) [Letzter Zugriff 03.06.2005]

Searle, John R. (1986): Geist, Hirn und Wissenschaft. Suhrkamp Verlag.

Searle, John R. (1990): „Is the Brain´s Mind a Computer?“ in Scientific American, Januar 1990, S. 20 ff.

Zawidzki, Tadeusz (ohne Jahr): Churchland, Patricia. Online-Dokument, verfügbar über <http://www.artsci.wustl.edu/%7Ephilos/MindDict/churchlandps.html> [Letzter Zugriff: 03.06.2005]

Zawidzki, Tadeusz (ohne Jahr): Churchland, Paul. Online-Dokument, verfügbar über <http://www.artsci.wustl.edu/%7Ephilos/MindDict/churchlandpm.html> [Letzter Zugriff: 03.06.2005]