

Werden Rechner jemals denken können?

Beginn einer Diskussion

Hintergrund...

- Vortrag im Rahmen des Prosimares “Denkmaschinen”
 - A.M. Turing - „Computing Machinery and Intelligence” (1950)
 - Beschreibt den Turing-Test
-
- Christian Merkle
 - Georg Zetzsche

Gliederung

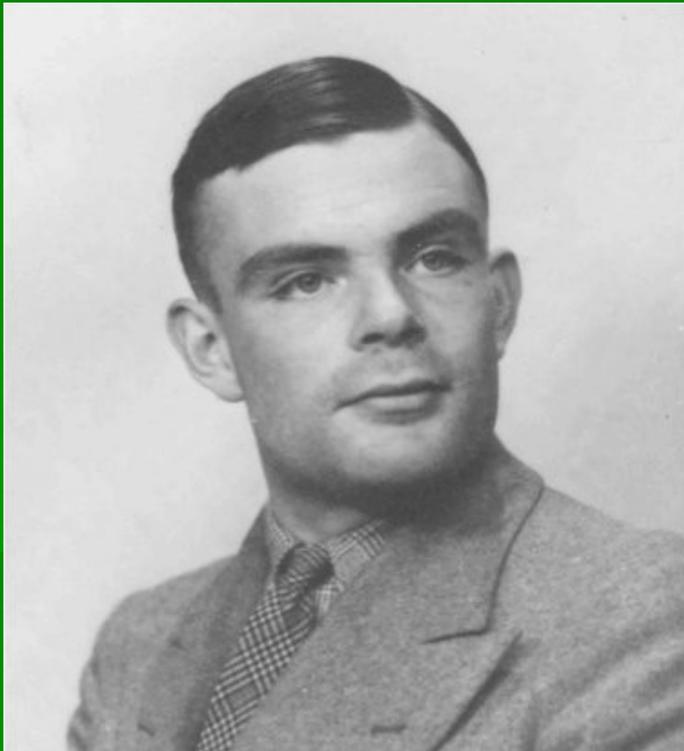
Teil 1

- Alan Turing
- Der Turing-Test
 - Das Imitationsspiel
 - Verwendete Maschinen
- Der Stand 2005
- Eliza
- Zusammenfassung

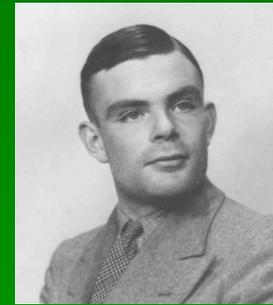
Teil 2

- Konträre Ansichten zur Hauptfrage
- Lernende Maschinen

Alan Turing

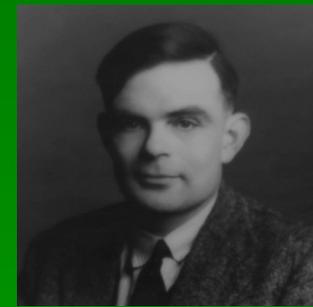
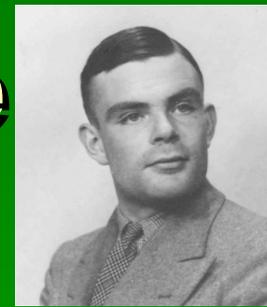


Alan Turing



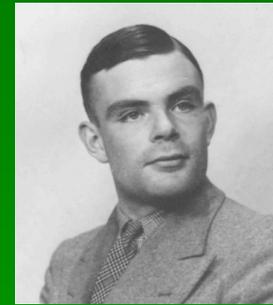
- Mathematiker
- Kryptoanalytiker
- Einer der Urväter des Computers

Wissenschaftliche Verdienste



- Lieferte wichtige Entwurfsideen für die Britische “Bomba” (elektromechanisches Dechiffriergerät für “Enigma” verschlüsselte Nachrichten) (WW2)
- Turing-Maschine (1936)
- ***Turing-Test (1950)***
- Diverse weitere Arbeiten

Vita

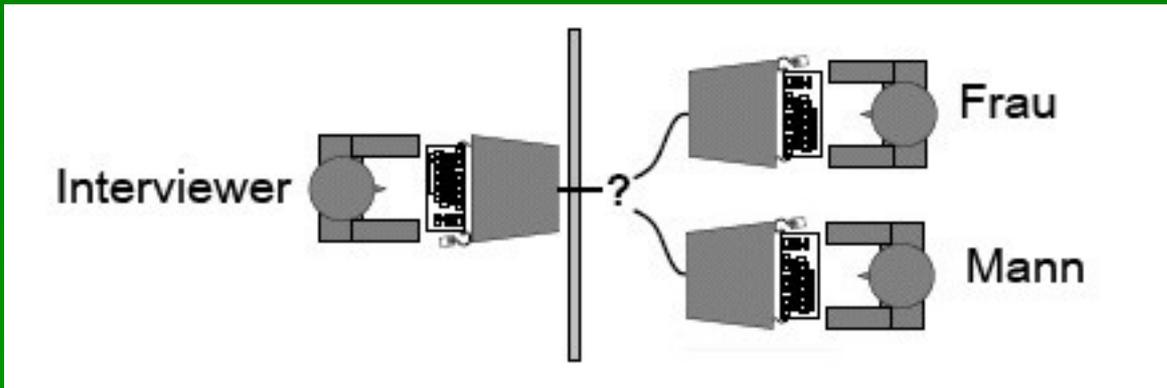


- * 23. Juni 1912 in London
- † 7. Juni 1954 in Wilmslow
- Sohn eines Britischen Beamten
- Studium in Cambridge und Princeton
- 1952 angeklagt wegen "sexueller Perversion"
- Tod durch Cyanid-Vergiftung (vermutlich Suizid)



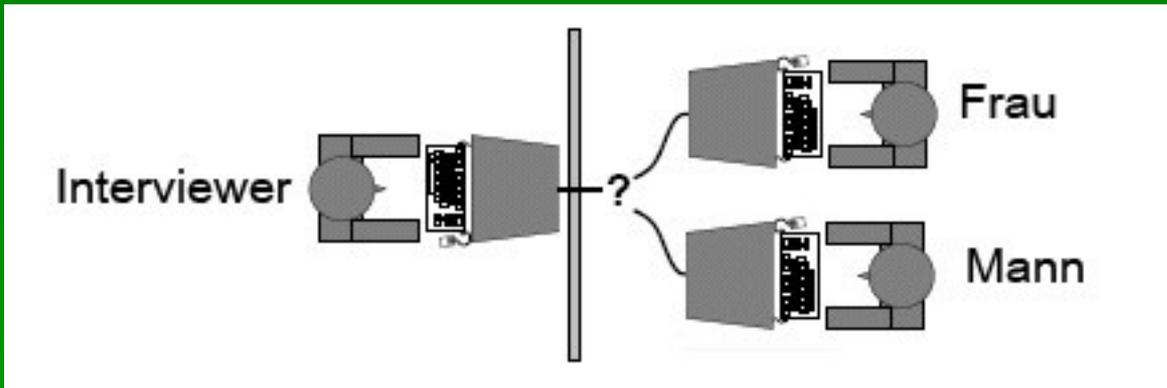
**Können Maschinen
denken?**

Das Imitationsspiel



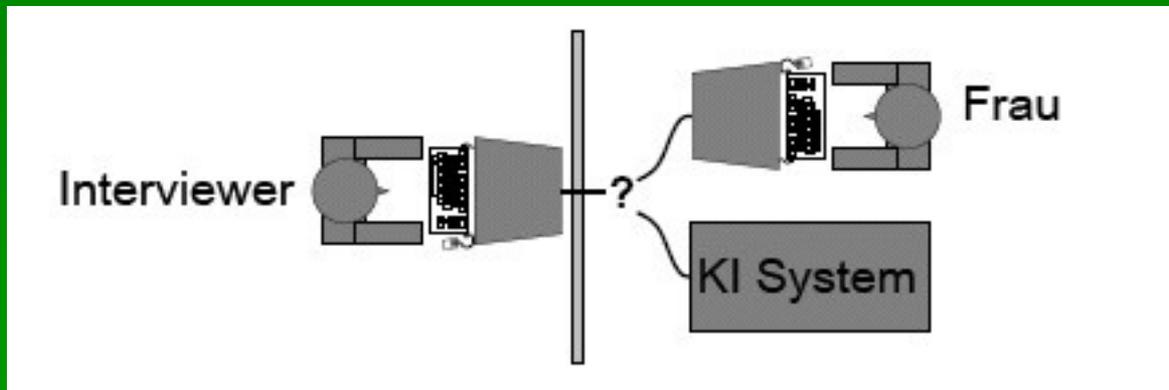
- Teilnehmer: Interviewer, eine Frau, ein Mann
- Teilnehmer sind räumlich getrennt
- Kommunikation natürlichsprachlich, textbasierend
- Frau und Mann müssen sich beide als Frau ausgeben
- Interviewer muss nach einer Befragung die Frau benennen

Das Imitationsspiel - Beispiel



- Interviewer sollte Fragen stellen wie: „Wie lang sind Deine Haare?“
- Die Frau sollte möglichst ehrliche Antworten geben: “Ungefähr 20cm”
- Die Frau kann nach hinzufügen: “Ich bin die Frau, bitte glaube mir”.
- Problem: Der Mann könnte die selben Einwürfe machen

Der Turing-Test



- Der Mann wird durch ein KI System/eine Maschine ersetzt!
- Würde sich der Interviewer genauso oft richtig oder falsch entscheiden?
- Letztendlich: Kann der Interviewer zwischen Mensch und Maschine unterscheiden?

Der Turing-Test - Probleme

- Das KI System müßte darauf programmiert sein, wie ein Mensch zu Antworten
 - Wartezeit auf die Beantwortung von Matheaufgaben (z.B. $34957 + 70764$)
 - KI System sollte falsche Antworten geben
 - KI System sollte ebenfalls Rechtschreib- und Flüchtigkeitsfehler machen
- Beste Strategie für das KI System: sich verhalten wie ein Mensch



**Können Maschinen
denken?**



**Können Maschinen
denken?**



**Kann sich eine
Maschine menschlich
verhalten?**

Die Maschine

- Ist eine von Menschen konstruierte Maschine (im technischen Sinn).
- × Auszuschließen sind: Biologische Techniken (z.B. Gentechnik) und Mensch/Maschinen Kreuzungen
- Turing: Die ideale Maschine wäre ein Digitaler Computer

Vorteile eines Digitalen Computers

- x Maschinen arbeiten nach einem festen “Regelbuch”
- Digitale Computer sind universell
- Ein digitaler Computer kann jede Art von Operationen ausführen
 - Kann somit jede Maschine emulieren
- Der Bau einer extra Maschine wäre unnötig
- Ein Digitaler Computer böte die Möglichkeit, freien Willen durch einen Zufallsgenerator zu simulieren.

Aufbau eines Digitalen Computers

- Besteht aus drei Komponenten:
 1. Speicher (im idealfall unendlich)
 - Enthält die zu bearbeitenden Daten und das “Regelbuch”
 2. Ausführende Einheit (CPU/Prozessor)
 - Führt die gewünschten Operationen aus
 3. Kontrolleinheit
 - Das “Regelbuch” im Speicher, quasi der Quellcode/Programmcode

Arbeitsweise eines Digitalen Computers

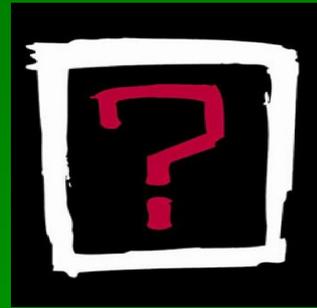
- “addiere die Zahl in Speicher 6809 zu der Zahl in 4302 und schreibe das Ergebnis in den letzteren”
- Maschinencode könnte sein: 6809430217
- “Führe den Programmcode ab Speicher 3309 aus”
- “Falls im Speicher 2110 eine 0 gespeichert ist, führe die Operation in Speicher 4405 aus, ansonsten arbeite einfach weiter”



**Kann sich eine
Maschine menschlich
verhalten?**



**Kann sich eine
Maschine menschlich
verhalten?**



**Gibt es (endliche)
Automaten, die im
“Imitationsspiel” gut
abschneiden würden?**

**Kann man einen
Computer so
programmieren, dass er
im “Imitationsspiel” gut
abschneiden würde?**



**“Im Jahr 2000 hat eine Maschine eine
Chance von 30%, einen Menschen 5
Minuten lang zu täuschen”
(Turing 1950)**

Stand 2005

- Bisher hat es noch kein Programm geschafft, den Turing-Test zu bestehen.
- Das Problem ist die mangelnde Akzeptanz der etablierten KI-Forschung. Der Turing-Test gilt nicht als Intelligenz-Nachweis:

"Die [...] Wettbewerber verstehen den Test als trickreiches Frage-und-Antwort-Spiel, das man mit ein bisschen Lexikon und Grammatik in Gang setzt. Sie machen einfach nur Intelligenzbetrugsprogramme"

(Prof. Dr. Thomas Christaller, Fraunhofer Institute for Autonomous Intelligent Systems (AIS))

Eliza

- Entwickelt in den 60er Jahren von Joseph Weizenbaum
- Programm, das einen Psychotherapeuten simulieren soll
- Simuliert eine Reaktion, indem es Aussagen des menschlichen Gesprächspartners in Fragen umformuliert.
- Reagierte fest auf bestimmte Schlüsselworte
- Beispiel:

Benutzer: *„Ich habe ein Problem mit meinem Vater.“*

Eliza *„Warum, sagen Sie, haben Sie ein Problem mit ihrem Vater?“*

- Praktizierende Psychiater glaubten damals tatsächlich, so zu einer automatisierten Form der Psychotherapie gelangen zu können...

Der Loebner Preis

- Wird seit 1991 im Londoner Science Museum vergeben
- Leicht abgewandelte Form des Turing-Tests
 - Das Programme muss sich einem „Chat“ mit einem menschlichen Prüfer stellen
 - Nach fünf Minuten entscheidet der Prüfer, ob er es mit Mensch oder Maschine zu tun hat.
- Medallien:
 - Gold (\$100.000, Quote 3/10, "audio-visueller Input")
 - Silber (\$25.000, Quote 3/10)
 - Bronze (\$2000, Trostpreis für das "menschlichste" Programm eines Jahrgangs)

Der Loebner Preis











Zusammenfassung

- Computer kann man so programmieren, dass sie für einen Betrachter intelligentes Verhalten zeigen
- Der Turing-Test versucht, Intelligenz präzise zu definieren und zu messen
- Bis heute hat noch kein Programm den Turing-Test bestanden

Quellen

“Computing Machinery and Intelligence” (Alan Turing, 1950)

“Kumpel Computer” (Annette Leßmöllmann, Die Zeit 41/2000)

<http://www.loebner.net/Prizef/loebner-prize.html>

Diverse Wikipedia.de Artikel über Alan Turing, den Turing-Test, Aliza etc.

Ende Teil 1

Können Maschinen denken? (Teil 2)

Christian Merkle

Georg Zetsche

6. Konträre Ansichten zur Hauptfrage

- Die Ansicht Turings zur Hauptfrage,
- dazu entgegengesetzte Einwände und
- Turings Antworten auf diese Einwände.

6.0 Turings eigene Ansicht

- Er glaubt, dass es in fünfzig Jahren möglich sein wird, Computer so zu programmieren, dass sie das „imitation game“ so gut spielen können, dass ein durchschnittlicher Befrager in höchstens 70% der Fälle nach fünf Minuten der Befragung die richtige Entscheidung trifft.
- Er beschränkt sich auf die präzisere Form der Frage, weil die allgemeine Frage „Können Maschinen denken?“ keine genügend klare Bedeutung hat.

6.1 Der theologische Einwand

- Denken ist eine Leistung der Seele.
- Gott hat jedem Menschen eine Seele gegeben, aber den Maschinen nicht.
- Daher können Maschinen nicht denken.
- Turings Antwort:
 - Argumentation schränkt die Allmächtigkeit Gottes ein
 - Gott sollte als allmächtiger auch Maschinen Seelen geben können
 - Konstrukteur erhält dabei nicht die Allmächtigkeit Gottes

6.2 Der „Kopf in den Sand“-Einwand

- Die Folgen denkender Maschinen wären zu schrecklich
- Wir hoffen und glauben daher, dass Maschinen nie denken können werden
- Dies entspringt dem Wunsch, den Menschen als überlegen nachzuweisen
- Keine Antwort Turings

6.3 Der mathematische Einwand

- Verschiedene Ergebnisse aus Mathematik und Logik zeigen Beschränkungen formaler Systeme.
- Prominente Beispiele: der Gödelsche Unvollständigkeitssatz, die Unentscheidbarkeit des Halteproblems
- Ein weiteres unentscheidbares Problem: „Wird die Maschine M auf irgendeine Frage mit Ja antworten?“
- Unentscheidbarkeit bedeutet: jede Maschine für dieses Problem gibt falsche Antworten oder läuft ewig.
- Turings Antwort:

- Es ist keinesfalls klar, dass Menschen diese Beschränkungen nicht auch haben.
- Menschen können zwar jeder Maschine bei je einer Frage überlegen sein, indem sie wissen, dass ein Berechnungsergebnis falsch ist.
- Aber es gibt keine Frage, bei der man so über alle Maschinen siegt.

6.4 Das Argument vom Bewußtsein

- Erst, wenn eine Maschine sich ihrer selbst bewußt ist, kann sie wirklich denken.
- Turings Antwort:
 - Dies ist eine Kritik an der Prüfmethode, mit der die Denkfähigkeit überprüft wird.
 - Testverfahren, um Bewußtsein nachzuweisen, sind kaum möglich
 - Führt man dieses Argument fort, so kann man auch nicht sicher sein, dass Menschen denken.
 - Deshalb Turings Vorschlag: „imitation game“ als Test

6.5 Argumente von verschiedenen Unzulänglichkeiten

- Argumente der Form „Selbst wenn all diese Dinge funktionieren, wird die Maschine nie X schaffen.“
- Mögliche X : freundlich sein, schön sein, Humor haben, richtig von falsch unterscheiden, Fehler machen, etc.
- Turings Antwort:
 - Für diese Aussagen wird meist keine Begründung angegeben
 - Sie entstammen der „wissenschaftlichen Induktion“: jemand hat viele Maschinen gesehen und schließt dann auf alle anderen Maschinen.

6.6 Der Einwand von Lady Lovelace

- Maschinen sind immer von Menschen geschaffen.
- Also können sie nie etwas wirklich neues tun.
- Turings Antwort:
 - Können wir denn sicher sein, dass Ideen von Menschen wirklich neu sind?
 - Ideen von Menschen können auch als Ergebnis von Lehre oder generellen Vorgehensweisen gesehen werden.
- Variante des Einwands: Rechner werden uns nie überraschen können.

- Turings Antwort:
 - Basiert auf dem Irrglauben, dass Menschen sofort alle Konsequenzen eines Sachverhalts kennen
 - Menschen machen Fehler und können daher von Maschinen überrascht werden

6.7 Das Argument von der Kontinuität im Nervensystem

- Das Nervensystem ist kein (endlicher) Automat.
- Ein kleiner Fehler im Eingangssignal kann das Ausgangssignal stark verändern.
- Also ist eine Nachahmung durch solche Automaten nicht möglich.
- Turings Antwort:
 - Dies wirkt sich beim „imitation game“ nicht aus.
 - Allein durch Befragen lässt sich kein Unterschied feststellen.

6.8 Das Argument vom informellen Verhalten

- Menschliches Verhalten lässt sich nicht in einem Satz von Regeln beschreiben. Es gibt zu viele Eventualitäten.
- „Gäbe es einen Regelsatz für das Verhalten, wären wir Maschinen. Einen solchen gibt es nicht, also sind wir keine Maschinen.“
- Turings Antwort:
 - Logisches Problem: Man muss unterscheiden: „rules of conduct“ und „laws of behaviour“.
 - Abwesenheit von „laws of behaviour“ beim Menschen wohl nicht beweisbar.

6.9 Das Argument von der übersinnlichen Wahrnehmung

- Menschen können beim „imitation game“ Telepathie einsetzen, Maschinen aber nicht.
- Maschinen können höchstens per Zufallszahlen entscheiden.
- Turings Antwort:
 - Dies ist ein starkes Argument.
 - Aber: vielleicht wird der Zufallszahlengenerator von der Telepathie beeinflusst!
 - Lösung: Ein „telepathie-sicherer Raum“.

7. Lernende Maschinen

- Mögliche Vorgehensweise: Konstruktion einer Maschine, die die (Lern)Fähigkeiten eines Kindes aufweist. Dann diese Maschine unterrichten.
- Unterschiedliche Ansichten zur Komplexität der „Kind-Maschine“. Möglich: eingebauter logisches Inferenzsystem.
- Regelvorrang muss festgelegt werden.
- Evtl. sollte ein Zufallselement eingebaut werden
- Offene Frage: Was sollte der Maschine zuerst beigebracht werden? Möglich: abstrakte Fähigkeiten oder Sinnesorgane.

In jedem Fall ist noch viel zutun!

Danke für Eure Aufmerksamkeit!