

# Einführung in die Künstliche Intelligenz

Maximilian Stock, Daniel Motz

Fakultät für Mathematik und Informatik, Friedrich-Schiller-University,  
Jena, Germany

`maximilian.stock@uni-jena.de`, `daniel.motz@uni-jena.de`

## 1 Der Gegenstandsbereich

### 1.1 Lernfragen

**Was meint "KI"?** KI ist das Unterfangen, künstliche intelligente Systeme zu bauen.

**Der Turing-Test**

**"Starke" und "schwache" KI**

**Symbolische KI vs. Konnektionismus**

### 1.2 TODOs

- (Seite 10 FS 1) teleologischer Aspekt WB-Systeme verstehen
- (Seite 10 FS 1) "nach generellen Prinzipien strukturiert" verstehen
- (Seite 10 FS 1) Wissenstransfer, Erklärung. Was bedeutet das?
- (Seite 11 FS 1) Wissenstypen: Begriffe abgrenzen

### Done TODOs

- Den "nicht-einfachen" Turing-Test (die Verhörsituation) verstehen
- Was ist eine einbettende Umgebung in einem PSS?
- PSS genauer definieren
- Starke und Schwache KI nach Searle verstehen

### 1.3 Was ist KI?

**1. Definitionsversuch** KI ist das Unterfangen, künstliche intelligente Systeme zu *bauen*. Mit Intelligenz sind Aufgaben, die Menschen gut bewältigen gemeint. Sie demonstrieren Intelligenz:

- Bilderkennung
- Stimmerkennung
- Aufstellen vernünftiger Erwartungen (z.B. wenn etwas "John" heißt, erwarten wir, dass es sich nicht um ein Kino oder ein Eis, sondern wahrscheinlich um eine Person; vielleicht noch ein Haustier oder Ähnliches).

Nicht alles, was der Mensch gut tut, ist interessant für die KI.

Viele Aufgaben, die gemeinhin als schwierig (Intelligenz voraussetzen) gelten, sind zu *speziell*: Das einzig intelligente, was ein Schachprogramm kann, ist Schachspielen!

### Einfacher Turingtest

- Zwei Räume verbunden durch Fernschreiber
  - in einem Raum der *Tester*
  - im anderen: *Mensch* oder *Maschine*
- Tester hat 30 Minuten Zeit über den Fernschreiber herauszufinden, ob es sich um einen Menschen oder eine Maschine handelt.

**2. Definitionsversuch** KI ist das Unterfangen, ein künstliches System zu bauen, das den Turingtest zuverlässig besteht.

**Der Turing-Test** In einer "Verhörsituation" gibt es drei Personen, die sich gegenseitig nicht sehen und nur schriftlich kommunizieren können. Person A ist

männlich, Person B ist weiblich und Person C hat ein beliebiges Geschlecht. Person C leitet das Verhör und soll bestimmen, welche der Personen welches Geschlecht hat. C kann beide mit X und Y ansprechen und muss am Ende A X und B Y oder B X und A Y zuordnen. A hat das Ziel, dass C die Geschlechter falsch bestimmt. Auch hier: Zeitlimit von 30 Minuten.

Nun fragen wir uns: "Was passiert, wenn eine Maschine die Rolle von A übernimmt?"

**Einwände gegen den Turing-Test:** Kann die Maschine als Mensch durchgehen, nur weil der Tester schlecht ist? (... z.B. einschläft ...) Man wiederholt den Test mehrfach. Der Test ist sinnfrei, wenn *B* lügen kann, z.B. vorgibt ein Mann zu sein, weil die "Intelligenz" der Maschine dann irrelevant für den Test ist. Er ist auch sinnfrei, wenn *A* nicht lügen darf. Man geht explizit davon aus, dass *B* keine Agenda hat.

**Searle's Einwand** Der Tester tippt nicht mehr als 10 Zeichen pro Sekunde. Die Maschine muss maximal 18.000 Zeichen beantworten. Da die Anzahl an möglichen Anfragen begrenzt ist, kann man in einer riesigen Tabelle alle möglichen Konfrontationen für diesen Dialog speichern und die Maschine kann mit ihrer Hilfe alle intelligenten Reaktionen nachschlagen. Hier scheiden sich die Geister. Wann ist eine KI eine KI und was muss sie können?

**Definition 1.** *Die "starke" KI-Hypothese*

*... eine funktionale Definition von Intelligenz  $\Rightarrow$  Turing-Test bestehen.*

**Definition 2.** *Die "starke" und die "schwache" KI-Hypothese*

*Eine schwache KI ist ein Werkzeug, das genau eine Klasse von Problemen lösen kann; zum Beispiel ein Schachcomputer. Für andere Probleme, etwa das Steuern eines Fahrzeugs, ist er ungeeignet. Während eine schwache KI ihren Job*

*nur anhand ihres Trainings erledigt, versteht eine starke KI ihre Aufgabe, und ist in der Lage eigenständig eine Lösungsstrategie zu entwickeln. Diese Fähigkeit zu argumentieren und zu abstrahieren, verleiht ihr außerdem die Fähigkeit, sich weiterzuentwickeln, sich zu verbessern und sich neue Fähigkeiten anzueignen. Sie ist autonom.*

**Was heißt hier "künstlich"?** Minimalforderung: künstlich  $\Rightarrow$  anorganisch. Bringt uns aber nicht weiter...

Besser: Ein künstliches System ist ein *physical-symbol-system* (PSS) im Sinne von Newell und Simon. Ein PSS ist eine *symbol-manipulierende Einrichtung* (enthält Symbole, die zusammen Symbolstrukturen bilden, welche erschaffen, bearbeitet, kopiert und gelöscht werden können. Beispiele: **Turingmaschine**, Universalrechner). Es muss in eine externe Umgebung eingebettet sein, damit man sein Verhalten evaluieren kann.

Beispiel Schach: Unser Schachbrett ist die *planning domain*. In dieser domain kann man sich nur über Züge von Figuren, die Positionen auf dem Brett und ob das Spiel gewonnen oder verloren ist unterhalten. Unsere KI repräsentiert ihre Erkenntnisse als Objekte in dieser domain.

**3. Definitionsversuch** KI ist das Unterfangen, ein PSS zu bauen, das den Turing-Test zuverlässig besteht. Zusatzanforderung mancher KI-ler: Die von dem PSS manipulierten Symbole sollen Objekten in der eingebetteten Umgebung entsprechen.

**Definition 3.** *Annahme des Deklarativismus.*

*Nach der Annahme des Deklarativismus kann ein System, das alle "intelligenten Antworten" in einer Tabelle nachschlägt, nicht intelligent sein.*

*Remark 1.* Ein neuronales Netz ist keine deklarativistische Teildisziplin. Außer wenn man so viele Neuronen anlegt, dass alle Beispiele "gelernt" werden können  $\Rightarrow$  *overfitting*.

*Remark 2.* Kann eine KI stark sein, wenn man auf einem Von-Neumann-Rechner ist? Wahrscheinlich nein, weil er deterministisch ist.

### Modellierungsansatz

1. Algorithmische Theorie (Problemstellung definieren, Beweis der algorithmischen Lösbarkeit)
2. Repräsentation und Algorithmus (Algorithmus abstrakt entwerfen. Mithilfe eines Modells, z.B. Von-Neumann-Rechner, geeignete Datenstrukturen wählen)
3. Implementation (Umsetzung in einer Programmiersprache bzw. Maschinenanweisungen, Optimierung)

Kompetenz: 1. und 2.

Performanz: 2. und 3.

**Wissensbasierte Systeme** Ein wissensbasiertes System kann als eine besondere Art von Programmiersystemen angesehen werden. Die Inferenzmaschine ist dabei ein Berechnungsmechanismus für mit der Wissensbasis gegebene Programme. Durch die Eingabe von Wissen wird die Inferenzmaschine programmiert. Das Wissen wird deklarativ repräsentiert. Es besteht sowohl aus Faktenwissen (ähnlich den Daten in einer herkömmlichen Datenbank) als auch Regelwissen, zum Beispiel in Form von Produktionsregeln (wenn ..., dann ...), die symbolisch vorliegen.

Quelle: Wikipedia: Wissensbasiertes System

- methodischer Aspekt – automatische Lösung mit Fachwissen
  - können mithilfe der Ableitungsregeln und dem Wissen in der WB schlussfolgern
- qualitativer Aspekt – schwieriges Problem
  - WB-Systeme werden in Feldern angewendet, auf denen imperative, algorithmische Lösung schwer realisierbar sind
- teleologischer Aspekt – praktische Bedeutung
  -

Ein WB-System benötigt explizit und deklarativ dargestelltes fachgebietsspezifisches Wissen, das nach generellen Prinzipien strukturiert ist und einen generellen Schlussfolgerungsmechanismus.

## **Entwurf**

- Wissensdarstellung ("Knowledge Representation")
- Wissensausnutzung ("Problem Solving")
- Wissenserwerb ("Knowledge Acquisition") → Erweiterung / Veränderung der WB
- Wissenstransfer, Erklärung ("Explanation")

## **Wissensrepräsentation**

### *Wissenstypen*

- elementare Tatsachen, Schlussweisen
- spezielle Fälle, Situationen
- bestimmte Vorgehensweisen, Algorithmen
- allgemeine Gesetzmäßigkeiten, Alltagswissen

## Logik-basierte Repräsentationsformalismen

- Terminologische Logiken, Beispiel: Diagnosesystem für Fahrzeugdefekte. Symptome als Input, Diagnose als Output)
- Vererbung (von Eigenschaften)
- Nicht-Monotonie (Monotonie bedeutet, dass bei Erweiterung der Wissensbasis alle bisherigen Schlüsse weiterhin gültig sind. Verwendung von Defaults ist bspw. nicht-monoton.)

## Problemlösungstechniken in WB-Systemen

**Klassifikation** ...

**Zugriff** ...

**Synthese** Schließen von neuem Wissen aus der bestehenden WB (und anschließendes Aufnehmen in die WB)

**Deduktion** Ableitung mithilfe von Ableitungsregel

**Simulation** ...

**Zielorientierte Suche** ...

**Planen als zielorientierte Suche auf Metaebene**

**Strukturanalyse** ...

**Transformation** ...

## Einsatzbereiche für WB-Systeme

- Interpretation physikalischer Daten (Messungen interpretieren)
- Diagnose von Systemzuständen (Fehler im Auto, kranker Patient)
- Planung von Handlungen, um Ziele zu erreichen (Marjapussi)
- Entwurf und Konstruktion nach Spezifikation
- Unterricht und Training von Kenntnissen und Fähigkeiten (Deep Learning?)

## **1.4 Wichtige Teilgebiete der KI**

### **Klassische (symbolische) KI**

- Wissensrepräsentation