1 Was ist Künstliche Intelligenz

Kaum eine Bezeichnung hat so viele widersprüchliche Reaktionen hervorgerufen wie die jenes Faches, in das ich Sie nun einführen will: "künstliche Intelligenz". Es wurde daher schon oft vorgeschlagen, das Fach anders zu benennen, z.B. schlug Wolfgang Bibel die Bezeichnung "Intellektik" vor. Eine Bezeichnung bekommt aber stets nach einiger Zeit den Klang, der dem Bezeichneten anhaftet. Daher ist es wichtiger, sich mit der künstlichen Intelligenz selbst und den Reaktionen auf sie auseinanderzusetzen, als ihr einen anderen Namen zu geben.

Ein Fach konstituiert sich mindestens durch

- ein Erkenntnisinteresse, also ein Thema,
- eine Terminologie,
- einen Stand der Diskussion, also eine Menge von Annahmen, auf die sich ein Kreis von KollegInnen geeinigt hat, und durch
- eine Reihe von Methoden.

Im Verlaufe der Vorlesung werden Sie einige Forschungsthemen, die Terminologie und die dahinter stehenden Annahmen und die wesentlichen Methoden der künstlichen Intelligenz, von jetzt ab kurz "KI" genannt, kennenlernen. Dabei werde ich stets die Entwicklung der fachlichen Diskussion zum Thema skizzieren. Vorab aber soll das globale Thema der KI umrissen werden (1.1). Dann wird die generelle Herangehensweise der KI diskutiert, indem drei verschiedene Ansätze genannt werden (1.2).

1.1 Das Thema der KI

Die am häufigsten angegebene **Bestimmung** der KI ist: Verhalten, das beim Menschen als intelligent betrachtet wird, soll mithilfe von Rechnern modelliert werden, um

- menschliches Verhalten genauer zu verstehen und zu beschreiben
- Rechner menschengerechter zu machen
- menschliche Fähigkeiten auf Rechnern verfügbar zu machen.

Diese Bestimmung hält verschiedene Ansätze der KI zusammen. Sie gibt nämlich nicht an, was als intelligentes Verhalten zählen soll, und sie gibt ebenfalls nicht an, wie wir uns die Modellierung vorzustellen haben. An diesen Stellen verweist die Bestimmung auf den jeweils aktuellen Stand der Diskussion. Was sie genau angibt, ist, daß die Modellierung mithilfe von Rechenanlagen erfolgen soll. Dabei legt sie aber nicht fest, welche Rolle der Rechner spielen soll.

Was als intelligentes Verhalten oder Intelligenz angesehen wird, wandelt sich - nicht zuletzt wegen der Entwicklungen innerhalb der Informatik. So wurde die Fähigkeit zu rechnen früher als ein Zeichen von Intelligenz betrachtet. Heute hingegen, wo es billige Taschenrechner gibt, hat sich das menschliche Selbstbild von dieser Fähigkeit zurückgezogen. Es ist kein we-

sentliches Merkmal menschlicher Intelligenz mehr. Auch Erinnerungsleistungen wurden früher als Zeichen von Intelligenz eingestuft. Heute hingegen, wo das Speichern von Daten Büchern und Rechnern zukommt, wird das Gedächtnis nicht mehr als bedeutendes Merkmal menschlicher Intelligenz betrachtet. Erinnern und Rechnen wird zwar durchaus noch als menschliche kognitive Fähigkeit gesehen, ist aber in seiner Bedeutung herabgestuft worden, weil es Rechnern so leicht fällt.

Andere Fähigkeiten hingegen, die früher als selbstverständlich betrachtet und daher wenig beachtet wurden, sind nun in den Blickpunkt des Interesses geraten, weil sie auf Rechenanlagen kaum oder gar nicht zu reproduzieren sind: eine Schraube eindrehen, einen Gegenstand erkennen, sprechen, lesen, eine Geschichte erzählen, einen Fehler als solchen erkennen und in Zukunft vermeiden, eine Melodie nachsingen. Es sind für Menschen ganz einfache Tätigkeiten. Sie können nur schwer beschrieben werden. Fragen Sie einmal einen Menschen, wie er einen Schuh zubindet. Bitten Sie ihn, das ganz genau zu beschreiben! Es sind Tätigkeiten, die selten erklärt werden. Fragen Sie jemanden, woher er weiß, daß etwas ein Tisch ist. "Ist doch klar" ist die häufigste Antwort. Gerade mit der Beschreibung solcher Tätigkeiten beschäftigt sich die KI. Und darin kann man eine Bedrohung für das Selbstbild des Menschen sehen oder eben auch eine besondere Achtung vor dem Menschen.

Den zu modellierenden Tätigkeiten entsprechen Teilbereiche der KI. Die Liste ist nicht abgeschlossen und besitzt keine prinzipielle Gültigkeit. Sie gibt lediglich den jetzigen Stand der historisch gewachsenen Beschäftigung mit kognitiven Tätigkeiten wieder:

- Problemlösen: Theorembeweisen,Diagnose als heuristische Klassifikation, Spielen: Schach, Dame, Mühle, ...
- Wissensrepräsentation: Schlußfolgern (Inferenzen), Wahrheitserhaltung, Revision
- Planung: Handlungsrepräsentation, Linearisierung von Handlungen
- maschinelles Lernen: induktives Lernen aus Beispielen, induktives Lernen aus Erfahrungen, deduktives Optimieren von Regelmengen
- Bildverarbeitung: Einzelbildverstehen, Szenenanalyse
- Sprachverarbeitung: maschinelle Übersetzung, Textverstehen und Informationsextraktion, Dialogsysteme, gesprochene Sprache analysieren und synthetisieren

Die Gliederung der KI nach solchen Tätigkeit wird zunehmend umgeändert in eine Gliederung nach methodischen Ansätzen. So ist eine eindeutige Zuordnung mancher Arbeiten zu Wissensrepräsentation, Problemlösen, Lernen oder Planung schwierig: zum Beispiel dann, wenn Formalismen zum nicht-monotonen Schließen behandelt werden. Eine solche Arbeit mag ein Phänomen rekonstruieren, das bei verschiedenen Tätigkeiten postuliert wird. Die Sprachverarbeitung ist am einfachsten als Teilbereich herauszustellen, weil sie eine Mutterdisziplin, nämlich die Linguistik hat, auf die sie sich bezieht, und weil sie Phänomene behandelt, die letztlich am menschlichen Sprachverhalten festzumachen sind. Es ist die menschliche Sprachgemeinschaft, die festlegt, was ein wohlgeformter Satz ist, und was nicht. Hingegen legt sie nicht fest, was eine wohlgeformte Formel ist und wie eine

Theorie revidiert werden muß. Dies ist vielmehr in der Diskussion unter Wissenschaftlern begründet, an der die KI-Wissenschaftler teilhaben.

1.2 Ansätze innerhalb der KI

In der KI werden - meist implizit - verschiedene Annahmen über die Art der Modellierung und die Rolle von Rechenanlagen bei der Modellierung gemacht. Danach können wir drei Ansätze herausarbeiten:

- Ingenieursansatz:
  Rechenanlagen sollen mit mehr Fähigkeiten als bisher ausgestattet werden, sei es, um Benutzern angemessenere Arbeitsmittel an die Hand zu geben, sei es, um Arbeitskräfte durch Maschinen zu ersetzen. Der Rechner spielt die Rolle eines Werkzeugs im Arbeitsprozeß.

- Beschreibungsansatz:
  Modelle als Beschreibungen intelligenter Tätigkeiten sollen auf dem Rechner dargestellt werden, damit sie besser validiert werden können. Der Rechner spielt die Rolle eines besonderen Buches, das Konsequenzen eines Modells selbst aufzeigen kann.

- Konstruktionsansatz:
  Das Modell eines Wesens soll nachgebaut werden. Der Rechner spielt die Rolle eines künstlichen Wesens.

Der Ingenieursansatz wird als praxis-orientierter Teil von sowohl dem Beschreibungs- wie auch dem Konstruktionsansatz mitverfolgt. Allgemeine Diskussionen zu diesem Ansatz unterscheiden sich nicht von Diskussionen in anderen Bereichen der Informatik (Stichwort: Tailorisierung der Kopfarbeit).

Der Beschreibungs- und der Konstruktionsansatz hingegen haben eine lange Tradition von Auseinandersetzungen: zum einen diskutieren die VertreterInnen der Ansätze untereinander, zum anderen wurden beide von außen kritisiert. Da der Beschreibungsansatz bisher am wenigsten explizit gemacht wurde (und da ich obendrein diesem Ansatz folge), beschreibe ich ihn am gründlichsten (1.2.1). Der Konstruktionsansatz ist durch neuronale Netze, "künstliche Realität" und "künstliches Leben" in jüngster Zeit wieder aktuell geworden (1.2.2).

1.2.1 Beschreibungsansatz

Der Beschreibungsansatz der KI hat zum Ziel, kognitive Tätigkeiten zu beschreiben. Die Beschreibung hat selbst natürlich ganz andere Eigenschaften als das Beschriebene. Das Wort "Hund" bellt nicht, der Satz "Ich falle herunter" fällt nicht. Beschreibung und Beschriebenes befinden sich auf verschiedenen Ebenen, die nur zum Scherz vermischt werden dürfen. Im folgenden soll kurz umrissen werden, was hier unter wissenschaftlichen Beschreibungen oder auch Modellen verstanden wird und welche Eigenschaften sie haben können. Insbesondere werden die operationalen Modelle skizziert und versucht, die häufigsten Mißverständnisse zu vermeiden, mit denen auf sie reagiert wird. Abschließend werden einige Äußerungen von KI-Vertretern zitiert, die den Beschreibungsansatz veranschaulichen.

1.2.1.1 Explizitheit und Bewußtheit

Eine Beschreibung ist **explizit** bezüglich dessen, was sie beschreibt. Sie mag implizit auch Annahmen über andere Sachverhalte einschließen. In jedem Falle aber ist es die Explikation von etwas, die eine Beschreibung von etwas ausmacht. Da Beschreibung und Beschriebenes auf verschiedenen Ebenen sind, läßt sich aus der Explikation nicht folgern, daß das Beschriebene ebenfalls explizit oder bewußt sei. Z.B. messen die Moleküle eines Gases nicht etwa Volumen, Druck und Temperatur, um dem Gasgesetz zu entsprechen. Entsprechend müssen wir auch nicht annehmen, daß Menschen eine bestimmte Grammatik (bewußt) im Kopf haben. Viele kognitive Tätigkeiten werden von Menschen unbewußt ausgeführt und können von ihnen selbst nicht erklärt werden. So kann ein Mensch, der gut im Kopf rechnen kann, nicht unbedingt das Rechnen erklären. Gleichzeitig kann ein Mathematiker, der das Rechnen expliziert und eine Beschreibung von Berechnungen vorlegt, nicht notwendigerweise selbst gut rechnen. Üblicherweise unterschieden wir deshalb Fertigkeiten von Wissen. Dabei sind **Fertigkeiten** unbewußte (kognitive) Vorgänge, während **Wissen** bewußt und explizit ist. Beschreibungen gehören immer zum Wissen. Das Beschriebene kann eine Fertigkeit sein. In der KI beziehen sich die meisten Beschreibungen auf menschliche Fertigkeiten. Damit wird in keiner Weise (es kann gar nicht oft genug betont werden, weil es immer gerade an dieser Stelle zu Mißverständnissen kommt) dem Menschen unterstellt, er übe solche Tätigkeiten aufgrund von Wissen aus.

1.2.1.2 Adäquatheit

Beschreibungen sind dem Beschriebenen in verschiedenem Grade adäquat. In der Sprachwissenschaft haben sich die folgenden zwei Adäquatheitsbegriffe durchgesetzt, die wir für die KI übernehmen können.

"Eine Darstellung ist **beobachtungsadäquat**, wenn sie alle im Hinblick auf eine Fragestellung wichtigen Eigenschaften eines Phänomens oder zahlreicher herausgegriffener Phänomene in ihrer jeweiligen Gliederung erfaßt. Dies würde aber zulassen, daß ad hoc für diese gerade herausgegriffenen Phänomene bestimmte Beschreibungsverfahren gewählt würden, die bei anderen Phänomenen vielleicht nicht mehr anwendbar sind. Deshalb wird eine höhere Form der Adäquatheit angestrebt: nämlich die Beschreibungsadäquatheit.

Eine Darstellung ist **beschreibungsadäquat**, wenn sie sich auf ein System von Verfahren stützt, das ein ganzes Feld systematisch zusammengehöriger Phänomene zu erfassen erlaubt. Nicht nur die tatsächlich beobachteten Phänomene, sondern auch alle potentiell beobachtbaren Phänomene innerhalb dieses Feldes sollen erfaßt werden können. ...

Die Berechtigung, derart ein ganzes Feld systematisch zusammengehöriger möglicher Erscheinungen abzustecken und sich damit auf allgemeine Gesetzmäßigkeiten oder allgemeine Regeln zu berufen, ergibt sich nicht allein aus Beobachtungen und Phänomenbeschreibungen,... Sie ergibt sich aus einem bestimmten erklärenden und theoretischen Anspruch und aus der

allgemeinen Problemlage, wie sie von einer Wissenschaft gegenwärtig gesehen wird." (Wunderlich 1974:87f)[1]

Beim maschinellen Lernen werden wir diese Adäquatheitsbegriffe unter anderen Namen wiederfinden. An dieser Stelle geht es darum, Beschreibungen im wissenschaftlichen Sinne etwas genauer zu fassen.

1.2.1.3 Phänomene

Von jetzt an werden wir solche wissenschaftlichen Beschreibungen "Modelle" nennen, um das Wort "Beschreibung" umgangssprachlich lassen zu können. Modelle beziehen sich also auf mögliche und tatsächliche Beobachtungen, die als **Phänomene** festgelegt wurden, und auf einen "theoretischen Anspruch", der festlegt, was wir als Modell eines Phänomens akzeptieren und was wir als "System von Verfahren" begreifen wollen. Das Festlegen von Phänomenen wählt bereits unter Beobachtungen aus und ordnet sie unter einem gewissen Aspekt an. Modelle explizieren also nicht alles, was wir beobachten können, sondern nur jeweils das, worauf es uns unter einem bestimmten Aspekt ankommt. Stellen Sie sich vor, Sie sehen in einem Werbefilm zwei Männer, die Billard spielen und dabei reden. Sie können Aspekte oder Teile dieser Situation als Phänomene betrachten: was die Männer sagen, sind deutsche Texte (syntaktisch und semantisch wohlgeformt), wie die Billardkugeln sich bewegen, ist ein vollkommen elastischer Stoß, der ganze Film ist eine Produktwerbung. Diese Phänomene werden von Modellen unterschiedlicher Wissenschaftsdisziplinen abgedeckt: die Syntax und Semantik des Deutschen beschreibt die Äußerungen, die Impuls- und Energieerhaltung beschreibt den vollkommen elastischen Stoß, und ein Modell der Kommunikationswissenschaft könnte die Werbung beschreiben.

1.2.1.4 Eigenschaften von Sprachen

Ein Modell muß in irgendeiner Sprache geschrieben sein. Diese Sprache kann eine natürliche Sprache sein oder eine formale Sprache. Sie kann als Zeichen Laute, Buchstaben, Ziffern oder Rauchwölkchen verwenden (Alphabet). Die Regeln, nach denen die Zeichen zusammengesetzt werden dürfen (Syntax: Bildung wohlgeformter Ausdrücke oder Wörter), können unbewußt befolgt werden oder explizit bekannt sein. Ebenso können die Regeln, wie Aussagen der Sprache zu interpretieren sind (Semantik: Abbildung auf das Beschriebene oder auf Wahrheitswerte), unbewußt oder bewußt sein. Jede gewählte Sprache für ein Modell hat bestimmte Eigenschaften:

- Sie hat eine bestimmte **Ausdrucksfähigkeit** (expressability). Wenn wir als Beschreibungssprache z.B. die Aussagenlogik wählen, können wir Relationen nicht ordentlich ausdrücken. In natürlicher Sprache hingegen können wir alles ausdrücken.

- Sie ist für bestimmte Menschen leicht oder schwer zu verstehen (**Verständlichkeit**). So gibt es Wissenschaftssprachen, die unter Wissenschaftlern Mißverständnisse vermeiden helfen, für andere Menschen hingegen mißverständlich bis unverständlich sind.

[1] Chomsky (1964) führt für die Sprachwissenschaft noch die Erklärungsadäquatheit ein. Eine Grammatiktheorie ist erklärungsadäquat, wenn sie auf der Basis von primären Daten (Äußerungen) eine beschreibungsadäquate Grammatik auswählen kann.

Manchmal ist eine graphische Sprache leichter zu verstehen, als etwa das Deutsche.

- Es ist (nicht) **entscheidbar**, ob ein beliebiges Wort über dem Alphabet zu der gewählten Sprache gehört. Wenn wir nur eine endliche Menge von Wörtern zugelassen haben, so ist unsere Beschreibungssprache natürlich entscheidbar. Es ist fraglich, ob natürliche Sprachen eine endliche Menge von Wörtern haben und ob wir den Duden für das Deutsche als vollständige Aufzählung akzeptieren.

- Es gibt (k)einen Entscheidungsalgorithmus, der angibt, ob ein beliebiger Ausdruck der Sprache wahr sein kann, also **erfüllbar** ist. Bei der Aussagenlogik ist es einfach, für jeden möglichen Wahrheitswert (wahr oder falsch) jeder Aussagenvariablen (z.B. p, q) festzustellen, ob der Ausdruck (z.B. $\neg p \vee q$) wahr sein kann. Bei der Menge der deutschen Sätze hingegen gibt es wohl keinen Entscheidungsalgorithmus - wir haben ja noch nicht mal die Interpretation, die uns Wahrheitswerte für (welche?) semantischen Entitäten liefern würde!

- Die Sprache gehört in eine bestimmte **Komplexitätsklasse.** Bei der Aussagenlogik ist schon die Menge aller erfüllbaren Ausdrücke nur in exponentieller Zeit bezüglich der Anzahl von Aussagenvariablen festzustellen, d.h. sie gehört (semantisch) in die Klasse der nichtpolynomialen Sprachen. Natürliche Sprachen werden üblicherweise nach ihrer Syntax in eine Komplexitätsklasse der Chomsky-Hierarchie eingeordnet, wo sie schwach kontextsensitiv sind.

1.2.1.5 Formale Modelle

Die Eigenschaften der Beschreibungssprache sind deshalb wichtig, weil das Modell ja leichter zu handhaben sein soll als das Beschriebene. Die verschiedenen Eigenschaften zeigen unterschiedliche Gesichtspunkte auf, nach denen wir eine Beschreibungssprache auswählen. Die Gesichtspunkte Ausdrucksfähigkeit und Komplexität liefern geradezu entgegengesetzte Sprachen. Nicht unter allen Gesichtspunkten und für alle, die mit dem Modell arbeiten wollen, ist eine formale Sprache günstiger als die natürliche. Auch in der KI wird nicht alles formalisiert. Dennoch ist die Herangehensweise der KI insgesamt eine formale. Es wird eine formale Sprache gesucht, mit der die wichtigen Phänomene zusammenhängend beschrieben werden können. Dabei soll diese Sprache auch möglichst vielen Menschen verständlich sein, da nur Menschen darüber befinden, ob ein Modell akzeptiert wird, oder nicht. Oder, anders ausgedrückt: es werden **formale Modelle** für kognitive Tätigkeiten konstruiert. Die Formalisierung hat den großen Vorteil der leichteren Überprüfbarkeit von Behauptungen am Modell. Am leichtesten läßt sich ein Modell nutzen, wenn man an verschiedenen (syntaktischen) Formen bereits (semantische) Bedeutungsunterschiede festmachen kann, und wenn Prozesse der Veränderung sich auch in einer veränderten Form niederschlagen. Obendrein sollten die Merkmale, die für das Modell wichtig sind, leicht zugänglich sein, während unwichtigere Merkmale ruhig schwieriger abzulesen sein dürfen.

1.2.1.6 Operationale Modelle

Soweit unterscheidet sich der Beschreibungsansatz der KI in seiner Herangehensweise überhaupt nicht von der modernen Sprachwissenschaft, einigen Zweigen der Psychologie und der Informatik. Das KI-Spezifische mag in

der Kombination von Thema und Herangehensweise, in der Auswahl von Kriterien bei der Entwicklung von formalen Sprachen für Modelle und schließlich darin liegen, daß die Modelle nicht nur formal und dennoch verständlich, sondern auch operational sein sollen. Ein **operationales Modell** kann das Verhalten, das es beschreibt, auch hervorbringen. Ein KI-System zur Sprachverarbeitung beschreibt einen Dialog, indem es ihn führt. Ein KI-System zum Planen beschreibt Planung, indem es plant. Ein maschinelles Lernsystem beschreibt Lernverhalten, indem es lernt. Und an dieser Stelle geraten dann viele der eben noch deutlichen Unterschiede aus dem Blickfeld. Jetzt bellt plötzlich das Modell vom Hund, und die Eigenschaften des Modells entsprechen dem Beschriebenen, und damit scheint eine eineindeutige Entsprechung zwischen Menschen und Rechnern möglich. Das ist aber nur bei oberflächlicher Betrachtung so. Tatsächlich gelten alle oben gemachten Aussagen immer noch - ein operationales Modell spielt immer noch die Rolle eines Buches, nur eben die eines besonders praktischen Buches.

Wir dürfen nicht vergessen, daß wir für die Modellbildung Phänomene isoliert haben und schon damit einen Unterschied zwischen dem Beschriebenen und der komplexen Realität hervorgerufen haben. Um im Beispiel zu bleiben: ein operationales Modell von einem Hund bezüglich des Bellens beschreibt das Bellen mit einer formalen Beschreibung vom Bellen, die ausgeführt wird. Damit ist aber keineswegs der Hund in seiner ganzen Komplexität beschrieben. Wie oben schon ausgesagt, abstrahiert ein Modell von den Beobachtungen, die es nicht beschreiben soll. Außerdem entsprechen die Eigenschaften des Modells nicht eineindeutig den Eigenschaften des Beschriebenen. Die Beschreibung vom Bellen mag erfüllbare Sätze enthalten. Aber wie sollte wohl das Bellen eines Hundes erfüllbar sein? Selbst da, wo sich Eigenschaften auf das Modell genauso wie auf das Beschriebene beziehen können, ist die Beziehung ein bißchen komplizierter. So können wir zwar sicher sein, daß wir eine kontextsensitive Sprache nicht mit einem kontextfreien Modell vollständig beschreiben können. Aber nur, weil wir eine einfache Sprache (z.B. Hundebellen) mit einem kontextsensitiven Modell beschrieben haben, muß diese nicht plötzlich komplexer werden; sie bleibt einfach, wie immer wir sie beschreiben.

Genauso wie zu Beginn der Modellbildung unter einem Aspekt wichtige Eigenschaften aus der Realität hervorgehoben werden, so wird auch das Verhalten eines operationalen Modells wieder eingeteilt in **relevante** und **abundante Merkmale**, d.h. lediglich durch die technische Realisierung des Modells bedingte Merkmale (v. Hahn 1978). Bei einer methodisch sauberen Modellbildung müssen wir vorher festlegen, was wir hinterher als relevante Merkmale akzeptieren wollen.

Ein operationales Modell ist **praktisch**, weil es Konsequenzen von theoretischen Behauptungen deutlich macht und damit dem Menschen eine Reihe von Arbeiten beim Testen abnimmt. Zum Beispiel kann eine operationale Grammatik Sätze produzieren, die wir dann nur noch daraufhin prüfen müssen, ob sie wohlgeformt sind oder nicht. Die Sätze als Konsequenz der Grammatik werden vom operationalen Modell automatisch geliefert. Hätten wir nur eine formale Grammatik in einem Buch stehen, so müßten wir selbst die Grammatikregeln ausführen, um einen Satz zu konstruieren, den wir dann prüfen könnten. Dabei würden uns Fehler unterlaufen, und wir würden nicht viele Sätze am Tag erzeugen. Operationale Modelle regen zu Erkenntnissen an, weil sie dazu zwingen, alles zu explizieren, was zum Hervorbringen des gewünschten Verhaltens notwendig ist. Da das Modell nach explizit angegebenen Vorschriften ausgewertet wird, liefert es nichts oder Falsches, wenn die Ausführungsvorschriften nicht korrekt oder unvollständig sind. Schreiben wir das Modell nur auf, merken wir vielleicht nicht, daß wir

etwas vergessen haben. Als Leser eines Buches interpretieren wir leicht Ausdrücke so, wie sie gemeint waren, statt sie so zu interpretieren, wie es angegeben ist. Schließlich bieten die operationalen Modelle eine Terminologie für komplexe Vorgänge. Die Beschreibung bleibt damit nicht nur auf einzelne Daten beschränkt (Leistungsmodell), sondern kann ein Modell von (theoretischen) Prozessen bieten (Funktionsmodell). Zum Beispiel können wir bei der Sprachverarbeitung von einer "Suche im Lexikon" sprechen. Wir behaupten damit nicht, daß Menschen im Kopf ein Lexikon haben, das sie durchsuchen. Wir behaupten damit aber, daß wir einen Teil menschlichen Sprachverhaltens als Suche im Lexikon beschreiben können.

1.2.1.7 Zitate

Herbert Simon kann nicht durchgängig als Vertreter des Beschreibungsansatzes gelten. 1957 machte er Prophezeiungen zur Zukunft der KI. Innerhalb von 10 Jahren

- würde ein Rechner Schachweltmeister
- würden Rechner hochwertige Musik komponieren können
- würde das Programm die Standardform einer psychologischen Theorie werden.

Heute, mehr als 40 Jahre später, wissen wir: Die jetzigen Schachprogramme schlagen selbst den Weltmeister. Die Komposition ist nicht so stark modelliert worden, obwohl es einige Ansätze gibt. Programmbeschreibungen, also Beschreibungen von Beschreibungen, haben in der Psychologie einen großen Einfluß. Immerhin mit dieser Modifikation seines letzten Punktes, der im Sinne des Beschreibungsansatzes gelesen werden kann, hat er ein Ziel der KI formuliert. Wir werden bei der Wissensrepräsentation sehen, daß tatsächlich als operationales Modell nicht mehr die unterste Sprachebene eines Programms, sondern seine logische Rekonstruktion angegeben wird.

Von Herbert Simon stammt auch das illustrierende Beispiel der Ameise: Die Spur einer Ameise, die von einem Ort zum anderen geht, ist sehr kompliziert. Versuche, mit mathematischen Funktionen die Spur zu beschreiben, sind kompliziert und erklären nichts - wenn sie möglich sind. Betrachtet man aber die Umgebung der Ameise, ist eine Regel ausreichend:

> immer den untersten Weg in ungefährer Richtung auf das Ziel wählen!

Man braucht dazu auch nicht die Zellen und Moleküle der Ameise zu betrachten. Die Zellstruktur der Ameise ist sehr komplex. Aber sie ist irrelevant für das Verhalten der Ameise in einer komplexen Umwelt. Die mikroskopische Ebene erklärt nicht das Verhalten der Ameise. (Simon 1969 :35f zitiert in Weizenbaum 1978:176)

Ein anderer KI-Vertreter, der in vielen Äußerungen jedoch dem Konstruktionsansatz angehört, ist Roger Schank. Shirley Turkle hat ihn interviewt: "Wir sind in vieler Hinsicht die Philosophen von heute. Wir wenden uns den gleichen Fragen zu, mit denen sich schon Aristoteles und nach ihm alle anderen beschäftigt haben. Wir haben eine andere Methode, dies zu tun. Diese Methode läßt sich in einem Begriff zusammenfassen: Prozeß. Wir sagen, daß der richtige Ansatz zur Analyse von Ereignissen der Welt der prozessuale

Ansatz ist, bei dem Klarheit darüber besteht, wie die einzelnen Schritte aussehen und welches die Eingaben sind, und der die Algorithmen liefert, um von einem Ort zum anderen, von Ort A nach Ort B zu gelangen." (Turkle 1984: 321)

Schließlich ein Beispiel, das illustriert, wie ein operationales Modell es erlaubt, über Prozesse zu reden, ohne daß man dem Funktionsmodell eine physische Realität im Menschen zusprechen müßte. Siegmund Freud hat Versprecher als Fehlleistungen in der Weise gedeutet, daß dort unbewußte Wünsche an die Oberfläche dringen. Donald Norman hingegen hat die von Freud angegeben Fälle untersucht und dabei festgestellt, daß die verwechselten Wörter inhaltlich nah beeinander liegen: "trocken"-"naß", "eröffnet"-"geschlossen", "heiß"-"kalt" sind sich jeweils sehr ähnlich. Wenn wir diese Wörter inhaltlich repräsentieren, haben sie nur ein Bit Unterschied. Aus diesem Grund ist eine Verwechslung wahrscheinlich - insbesondere wenn der Mensch aufgeregt ist. Tatsächlich findet man keine Fälle, wo Menschen inhaltlich fern liegende Wörter fälschlich äußern. Noch weitere Annahmen, wie die von Freud, sind gar nicht nötig, um die Fehlleistung zu beschreiben. Wenn man also an ein Modell die Anforderung stellt, es solle nur so viel annehmen, wie zur Beschreibung notwendig ist, reicht das Informationsmodell aus und ist einfacher als das Freudsche Modell (Turkle 1984: 306).

1.2.2 Konstruktionsansatz

Der Konstruktionsansatz beherrschte Teile der KI in den frühen 60er Jahren, wurde dann aber zurückgedrängt, da er keine konkreten Ergebnisse liefern konnte oder, wie im Falle des Perzeptrons, seine Ergebnisse sich als zu schwach herausstellten. Der Konstruktionsansatz will mithilfe des Rechners intelligente Wesen konstruieren. Intelligenz sei dadurch zu untersuchen, daß man sie selbst herstellt. Hier wird also nicht mehr einem Modell von Menschen eine Eigenschaft zugeschrieben, sondern die Eigenschaft wird für das geschaffene Objekt behauptet. Damit ist der Rechner mit seinem Programm nicht mehr eine Beschreibung von intelligenten Prozessen, sondern ist selbst ein intelligenter Prozeß. Man sagt, daß der Mensch eine Instanz von intelligenten Wesen ist, und ein KI-System ist eine andere Instanz von intelligenten Wesen. Somit sind die Systeme dem Menschen oder Tieren **analog**. Zu einem intelligenten Wesen gehört natürlich mehr als zu einer Beschreibung eines intelligenten Wesens. Während Beschreibungen in irgendeiner Form symbolisch sind (Gleichungen, logische Formeln, Funktionen, Listen), gehört zu einem Wesen auch die nicht-symbolische Materie. Während es bei einem Buch oder operationalen Modell nicht darauf ankommt, ob es auf Papier, auf dem Bildschirm, über einen CD-Spieler oder wie auch immer lesbar wird, kommt es bei einem Wesen durchaus darauf an, wie es materiell in der Welt ist. Wenn man den Pfad der Ameise beschreiben will, braucht man ihren Körper nicht zu modellieren, wenn man eine künstliche Ameise bauen will, muß man ihre Sinnesorgane (Perzeptoren) und ihre Beine (Effektoren) auch konstruieren.

Es gibt zwei wesentliche Argumente des Konstruktionsansatzes, die ihn von dem Beschreibungsansatz abgrenzen:

- Durch die Abstraktion von der physikalischen Ebene von Lebewesen, verlieren wir die eigentlichen Wurzeln der Intelligenz. Intelligenz ist auf körperliche Erfahrung gegründet und kann nur durch sie erklärt werden.

- Die logischen Beschreibungen sind umständlich, wenn wir Selbstorganisationsprinzipien untersuchen wollen. Andere Verfahren - die ebenfalls symbolisch, aber nicht logisch sind - erlauben es, emergentes (emergent: auftauchend, hervortretend) Verhalten eleganter zu modellieren.

Beide Argumente werden im folgenden kurz skizziert und dann durch Zitate veranschaulicht.

1.2.2.1 Körperlichkeit der Intelligenz

Die meisten Phänomene, die mithilfe des Beschreibungsansatzes untersucht wurden, setzen bereits andere, elementarere Tätigkeiten voraus. Wenn wir eine natürliche Sprache beschreiben, abstrahieren wir von der elementaren menschlichen Fähigkeit, Laute aus einer allgemeinen Geräuschumgebung zu isolieren. Wenn wir Szenenanalyse betreiben, abstrahieren wir von elementaren Vorgängen, die visuelle Eindrücke überhaupt erst ermöglichen. Zwar ist es bekannt, daß die Kenntnis einer Sprache Korrekturen bei der Lautwahrnehmung ermöglicht. Auch werden visuelle Eindrücke von erwartbaren Formen oder Farben beeinflußt. Umgekehrt ist aber noch nicht bekannt, wie die elementaren Fähigkeiten des Menschen zu sehen, zu hören, zu schmecken und sich zu bewegen seine kognitiven Prozesse beeinflussen. Es ist die These des Konstruktionsansatzes, daß Intelligenz durch körperliche Erfahrung nicht nur ermöglicht wird, sondern bedingt und ausgeprägt wird. Es ist deshalb das Programm dieses Ansatzes, bei Wahrnehmung und Reflexen mit der Forschung anzufangen, bevor man Sprachstrukturen, Schachspielen oder das Beweisen von Theoremen untersucht.

Ein Beispiel kann diese Auffassung beleuchten. Die Bedeutung des Wortes "Tasse" oder überhaupt ganz einfacher alltäglicher Dinge läßt sich sehr schwer definieren. Der Verdacht besteht, daß die Bedeutung sich letztlich auf Sinnes- und Handlungserfahrungen gründet. So könnte etwas eine Tasse sein, wenn man eine Flüssigkeit eingießen und diese dann trinken kann. Wenn sich dieser Verdacht bestätigen sollte, so könnte man die Sinneserfahrungen nicht mehr aus dem zu beschreibenden Bereich der Phänomene ausgrenzen, sondern müßte gerade bei ihnen ansetzen. Man müßte dann beschreiben, wie man Flüssigkeiten eingießt und wie man trinkt, um aus solchen Beschreibungen komplexere Definitionen zu formen. Wichtig ist dabei auch, daß das Gelingen der Wahrnehmungen und Handlungen nicht allein von inneren Prozessen abhängt, sondern von außen gesetzt ist. Auch hier begegnet uns wieder Simons Ameise. Die Rückkopplung, ob ein Verhalten zum Ziel führt oder nicht, kommt von der komplexen Umwelt. Gerade deshalb darf der innere Prozeß so einfach sein. Am Beispiel der Tasse bedeutet das, daß die komplexe Umwelt regelt, ob ein Gegenstand eine Flüssigkeit einzugießen zuläßt, oder nicht. Zu beschreiben ist nur der Test, nämlich das Eingießen und das Beobachten, was es bewirkt hat.

Der Konstruktionsansatz der KI hat zum Ziel, Systeme in einer Welt handeln zu lassen, wobei sie aus Erfahrungen dann elementare Bedeutungseinheiten bilden. Diese Bedeutungseinheiten können symbolisch bezeichnet werden. Die Symbole gründen sich dann auf sinnliche Erfahrungen (Stichwort: symbol grounding). Die Annahme dieses Ansatzes ist, daß sich derselbe Prozeß weiter fortschreiben läßt bis hin zu komplexen kognitiven Vorgängen. Es gäbe also ein Prinzip, wie simple Reflexe komplexeres Verhalten hervorbringen, die wiederum komplexeres Verhalten hervorbringen, ... bis hin zum Schachspielen. Dieses Prinzip soll untersucht werden.

Die Kybernetik hatte ein ähnliches Ziel und bot als Prinzip den Regelkreis an, der sich jedoch als nicht ausdrucksstark genug erwies. Auch die frühe Zeit der KI bot eine Formalisierung für ein solches Prinzip an, nämlich das Perzeptron. Auch dieses erwies sich als zu ausdrucksarm. Heutige Netze sind weit ausdruckstärker. Auch gibt es erste Versuche, sie bei Robotern einzusetzen. Davon, die Einbindung in die Welt und die Entwicklung von intelligentem Verhalten aus Sinneserfahrungen heraus zu modellieren, ist die KI aber noch weit entfernt.

1.2.2.2 Nicht-logische Modelle

Wie man leicht sieht, wenn man die Argumente explizit aufschreibt, ist das zweite Argument nicht eigentlich gegen die Beschreibung, sondern gegen die logische Form der Beschreibung für bestimmte Phänomene gerichtet. Es werden andere Untersuchungsformen für andere Phänomene ausgesucht. Dafür werden neue Formalismen untersucht.

Die Phänomene, für die eine elegante Untersuchungsmethode gesucht wird, sind vor allem solche der Wahrnehmung, aber auch Reflexe oder evolutionäre Prozesse wie etwa Arterhaltung oder Selektion. Die Phänomene werden auch unter einem besonderen, meist biologischen Aspekt formuliert. So wird eben nicht ein Denkvorgang wie Schachspielen als ein kognitives Phänomen hervorgehoben, sondern es soll der biologische Vorgang im menschlichen oder tierischen Gehirn beschrieben werden. Die Organisation von Gehirnen ist tatsächlich sehr interessant. Einige Eigenschaften, die beschrieben werden sollen, sind die Selbstheilung und die nicht lokalisierbare Leistung. Bei einer geringfügigen Beschädigung von Teilen eines Gehirns wird in gewisser Zeit die Leistungsfähigkeit wieder hergestellt - wie geht das? Die anfänglichen Versuche der Biologen und Mediziner, bestimmten Orten des Gehirns eindeutig bestimmte Leistungen zuzuordnen, haben sich als lediglich grob orientierend erwiesen. Obwohl das Sprachzentrum bei den meisten Menschen überwiegend hinter der linken Schläfe liegt, gibt es durchaus Menschen, bei denen eine Häufung von Sprachverarbeitungsprozessen in der rechten Gehirnhälfte festgestellt wurde. Außerdem gibt es Menschen (z.B. Rudi Dutschke), deren Sprachzentrum zerstört wurde und die dann doch wieder - nach einem langen Re-edukationsprozeß - fließend sprechen und verstehen konnten. Die These, daß man bestimmten kognitiven Leistungen ganz bestimmte "Produktionsstätten" zuordnen könnte, ließ sich nicht aufrecht erhalten.[2] Ein weiteres interessantes Phänomen ist die Bildung der DNA: Proteinketten und ihre Faltung. Dort wirken Prozesse, die mit einer Grammatik für Aminosäurefolgen beschrieben werden können, mit räumlichen Prozessen zusammen (was liegt nah beieinander nach einer bestimmten Faltung der DNA), und sogar zeitliche Prozesse spielen eine Rolle (wie lange dauert es, einen bestimmten chemischen Stoff hervorzubringen, der sich in die Kette eingliedert und dort wie ein "Satzzeichen" das Ende einer bestimmten Abfolge von Aminosäuren markiert oder sogar bewirkt). Diese Phänomene, die von Biologen oder Medizinern gemäß ihrer Disziplin bestimmt werden, sollen erforscht werden mit Hilfe formaler Methoden.

Ungleichungen und cellulare Automaten können als **Simulationen** emergenten Verhaltens genutzt werden. Als emergent gilt ein Muster, in das sich die graphische Ausgabe eines einfachen cellularen Automaten ein-

[2]Beachten Sie, daß dies auch eine Bestätigung für die Annahme des Beschreibungsansatzes ist, daß kognitive Vorgänge nicht auf physiologische/biologische Entitäten abzubilden sind!

schwingt. Für solche Muster gilt das Argument des Beispiels von der Ameise in umgekehrter Weise: hier ist es kompliziert oder gar unmöglich, mit Regeln das Muster zu beschreiben. Dabei ist es einfach, von einem cellularen Automaten Muster erzeugen zu lassen. Die cellularen Automaten müssen selbst wieder beschrieben werden, um sie inspizierbar und diskutierbar zu machen. Welche Erkenntnis gewinnen wir, wenn ein Prozeß ähnliche Ergebnisse liefert, wie der, den wir verstehen wollen, und wir verstehen diesen Prozeß ebenso wenig wie sein Vorbild? Dies ist die Problematik der Simulation. Während ein operationales Modell immer bereits eine Menschen verständliche und von Menschen interpretierte Beschreibung darstellt, muß eine Simulation nicht verständlich sein. Sie kann genauso undurchsichtig sein, wie das, was sie simuliert. Die Simulation ist ein dem Simulierten analoger Prozeß, wobei festgelegt sein muß, welche Aspekte sich analog sein sollen, und welche nicht. Von den verwendeten formalen Methoden und ihrer Inspizierbarkeit hängt es ab, ob die Simulation leichter zu beschreiben ist als das Simulierte.

Die formalen Methoden, die für nicht-logische Modelle genutzt werden, sind unabhängig von ihrer Nutzung zur Simulation interessant[3]. Die Statistik hat neues Interesse gefunden, weil z.B. neuronale Netze formal auf ihr basieren. Dietterich und andere (1990) haben gezeigt, daß unter bestimmten Bedingungen, unter denen die fehlertolerante Verarbeitung von sehr großen Beispielmengen wichtig ist, ein einfaches neuronales Netz (backpropagation) mehr leistet als ein einfaches symbolisches Verfahren (ID3). Dietterichs Versuch behandelte die Zuordnung von Phonemen (Zeichen für Laute) zu Schriftzeichen. Während ein neuronales Netz sozusagen über jede Kombination statistisch Buch führt, gibt es bei ID3 nur eine zentrale statistische Bewertungsfunktion. Reichert man die Statistik von ID3 an, so nähert sich die Performanz der des neuronalen Netzes. Bei statistischen Verfahren ist es umstritten, ob sie eher der Beschreibung oder eher der Simulation zuzurechnen sind. Eine statistische Aussage kann durchaus als Beschreibung eines Sachverhalts interpretiert werden. Die früheren Arbeiten zu Neuronalen Netzen wollten aber bewußt eine Simulation von Gehirnvorgängen darstellen.

1.2.2.3 Zitate

Das Interesse an biologischen Phänomenen illustriert das Zitat von Varela: "In ähnlicher Weise wird die Fähigkeit des Gehirns, Schädigungen zu verkraften, oder die Anpassungsfähigkeit biologischer kognitiver Prozesse in neuen Umwelten ohne Beeinträchtigung ihrer Gesamtleistungsfähigkeit, von Neurobiologen als Tatsache betrachtet, ist aber im Computerparadigma nirgendwo von Einfluß." (Varela 1990:57)

Zunächst war der ganzheitliche Ansatz der (Re-)Konstruktion gängig in der amerikanischen oder angelsächsischen KI. In England gab es ein großes Robotik-Projekt, das aber 1974 gestoppt wurde, nachdem Lighthill der britischen Regierung über den Stand der KI-Forschung berichtet hatte. Zum Thema Roboter sah er keinerlei Fortschritt. Er wertete das Unterfangen als die Kompensation männlicher Wissenschaftler, daß sie nicht gebären können (Fleck 1982:191).

[3]Zu numerischen Verfahren siehe Schwefel (1981).

Viel weiter zurück reichen Frankenstein, Homunculus oder Golem - Phantasien von der Erzeugung von nicht-menschlichen (Lebe-?)wesen. Im Jahre 1580 kannte Rabbi Loew, so ist es überliefert, den aus 72 Buchstaben zusammengesetzten Namen Gottes und konnte daher eine menschenähnliche Lehmfigur beleben, den *Golem*. Der fand heraus, wann ein Progrom stattfinden würde, so daß der Rabbi die Gemeinde warnen konnte, und half dem Rabbi im Haushalt. In gewissen KI-Kreisen wird behauptet, daß Sussman, Minsky, von Neumann und Wiener von Rabbi Loew abstammen.

Roger Schank äußerte in demselben Interview, das oben zitiert wurde, auch eine stark konstruktivistische Auffassung: "Ich wollte schon immer einen Geist erschaffen. Irgend etwas in der Art. Es ist das Aufregendste, was man machen kann. Das Wichtigste, was man machen kann." (Turkle 1984: 322)

Ähnlich auch Don Norman:"Ich habe einen Traum, mir meinen eigenen Roboter zu bauen. Ihn mit meiner Intelligenz auszustatten. Ihn zu meinem Ebenbild, zu meinem Geist zu machen. Mich selbst in ihm zu sehen." (Turkle 1984: 322)

Auch in etwas subtilerer Weise werden die Maschinen künstlichen Lebewesen gleichgesetzt: "Will man außerdem eine künstliche Intelligenz schaffen, die Maschinen herstellt, die in dem Sinne intelligent sind, daß sie mit den Menschen zusammen (so wie die Tiere) eine gemeinsame Welt des Verstehens und Handelns aufbauen, dann sehe ich keinen anderen Weg dafür, als diese Maschinen in gleicher Weise durch einen Prozeß evolutionärer Transformationen zu entwickeln bzw. zu erziehen, wie es die handlungsbezogene Perspektive nahelegt. Wie ergiebig oder wie schwierig dieser Ansatz sein wird, darüber mag sich jeder seine eigene Meinung bilden." (Varela 1990:121)

1.3 Literatur

Als Leseempfehlung für diejenigen, die sich intensiver mit dem in diesem Kapitel angesprochenen Stoff beschäftigen möchten, gelten nur die mit * markierten Literaturhinweise. Die anderen sind der Überprüfbarkeit wegen und als Verweis auf Spezialwissen zu einzelnen Punkten aufgeführt. Wer genau einen weiteren Artikel lesen möchte, soll sich einen der beiden ** markierten aussuchen.

Chomsky, Noam (1964): Aspects of the Theory of Syntax, deutsche Übersetzung 1969 Frankfurt a.M.: Suhrkamp

Dietterich, Thomas/ Hild, Hermann/ Bakiri, Ghulum (1990): A Comparative Study of ID3 and Backpropagation for English Text-to-Speech Mapping, in: Porter/Mooney (eds): Machine Learning - Procs. of the 7th ICML, San Mateo:Morgan Kaufmann

*Dietz, Peter (1996): Menschengleiche Maschinen, Vorlesungsskript WS 96/97, Universität Dortmund

Fleck, James (1982): Development and Establishment in Artificial Intelligence, in: Elias, Martins, Whitley (eds): Scientific Establishment and Hierarchies. Sociology of Sciences, Vol. VI, Reidel

Opwis, Klaus, Plötzner, Rolf (1996): Kognitive Psychologie mit dem Computer, Heidelberg: Spektrum

Pylyshyn, Zenon (1989): Computing in Cognitive Science, in: Posner (ed) *Foundations of Cognitive Science*, Cambridge (Mass): MIT Press

Simon, H. A. (1969): The Sciences of the artificial, Cambridge

Schwefel, Hans-Paul (1981): Numerical Optimisation of Computer Models, Chichester: Wiley

*Turkle, Sherry (1984): Die Wunschmaschine - Vom Entstehen der Computerkultur, Reinbek: Rowohlt

Varela, Francisco (1990): Kognitionswissenschaft - Kognitionstechnik: Eine Skizze aktueller Perspektiven, Frankfurt a.M.:Suhrkamp

von Hahn, Walther (1978): Überlegungen zum kommunikativen Status und der Testbarkeit in natürlichsprachlichen Artificial-Intelligence-Systemen, in: *Sprache und Datenverarbeitung* 1

Weizenbaum, Joseph (1978): Die Macht der Computer und die Ohnmacht der Vernunft, Frankfurt a.M.: Suhrkamp

Wunderlich, Dieter (1974): Grundlagen der Linguistik, Reinbek: Rowolt