

Ansgar Beckermann Semantische Maschinen

1. Daniel Dennett gehört – anders als etwa John Searle und Hubert Dreyfus – sicher nicht zu den grundsätzlichen Kritikern der KI-Forschung. In seinem Aufsatz »Three Kinds of Intentional Psychology« charakterisiert Dennett jedoch die Situation der kognitiven Psychologie auf eine Weise, die ebenfalls eine große Skepsis im Hinblick auf die Möglichkeiten der KI-Forschung zu implizieren scheint. Dennett zufolge steht die kognitive Psychologie nämlich vor einer prinzipiell unlösbaren Aufgabe – der Aufgabe zu erklären, wie eine rein syntaktische Maschine semantische Leistungen erbringen kann. Denn die Psychologie und die evolutionäre Biologie beschreiben das menschliche Gehirn als eine *semantische Maschine* (*a semantic engine*), deren Aufgabe es ist, die *Bedeutung* der verschiedenen eingehenden Reize zu erfassen, sich ihrem Sinne nach zu unterscheiden und dann entsprechende Handlungen zu veranlassen. Die Physiologie dagegen stellt dieses Organ als eine rein *syntaktische Maschine* (*a syntactic engine*) dar, die alle eingehenden Reize nur aufgrund ihrer strukturellen, zeitlichen und physikalischen Merkmale zu unterscheiden vermag und deren Aktivitäten allein durch diese »syntaktischen« Merkmale bestimmt wird. Nach Dennett stellt sich damit die kritische Frage:

»... how does the brain manage to get semantics from syntax? How could any entity ... get the semantics of a system from nothing but its syntax?« (1982, S. 54)

Und Dennetts Antwort lautet:

»It couldn't. The syntax of a system doesn't determine its semantics. By what alchemy, then, does the brain extract semantically reliable results from syntactically driven operations? It cannot be designed to do an impossible task ...« (ebd.).

Offenbar ist also auch Dennett der Auffassung, daß semantische Maschinen unmöglich sind.¹ Seiner Meinung nach können syntaktische Maschinen – wie zum Beispiel Computer – bestenfalls so konstruiert werden, daß es so scheint, als würden sie seman-

tische Aufgaben lösen, während sie tatsächlich doch immer syntaktische Maschinen bleiben. Diese Auffassung erläutert er am Beispiel einer Maschine, die zur Überwachung der an einem Telegraphen eingehenden Meldungen eingesetzt werden soll und deren Aufgabe es ist, aus diesen Meldungen genau alle Morddrohungen (genauer: alle in Englisch verfaßten Morddrohungen) auszusondern. Keine Maschine kann Dennett zufolge diese Aufgabe vollständig lösen. Denn dafür wäre es nötig, die semantische Kategorie »In Englisch verfaßte Morddrohung« syntaktisch zu definieren. Und das scheint ihm eine Aufgabe zu sein, die prinzipiell nicht gelöst werden kann. Deshalb wird sich die Maschine mit unvollkommenen Hilfsstrategien behelfen müssen, indem sie zum Beispiel alle eingehenden Meldungen daraufhin untersucht, ob sie die Zeichenfolgen

... I will kill you ...

oder

... you ... die ... unless ...

oder ähnliche Zeichenfolgen enthält. Auf diese Weise wird die Maschine zwar nicht alle, aber doch wenigstens eine gewisse Zahl von Morddrohungen aussondern können. Mehr ist nach Dennett jedoch auch nicht möglich. Denn:

»... if you want to get semantics out of syntax ..., variations on this basic strategy are your only hope. You must put together a bag of tricks and hope nature will be kind enough to let your device get by« (a.a.O., S. 54 f.).

2. Wenn man Dennetts System zur Erkennung von Morddrohungen betrachtet, dann fällt auf, daß die simplen Strategien des *pattern matching*, mit denen dieses System arbeitet, den Strategien des Programms ELIZA sehr ähnlich sind, das Mitte der sechziger Jahre von Joseph Weizenbaum als Programm zur Simulierung von Dialogen in einer nicht-direktiven Psychotherapie entwickelt wurde. ELIZA schien vielen ein wichtiger Fortschritt in der Entwicklung natürlichsprachlicher Systeme zu sein. Denn bei der Arbeit mit diesem Programm kamen sehr eindrucksvolle Dialoge zustande, die auf den ersten Blick wirklichen therapeutischen Dialogen zumindest verblüffend ähnlich waren. Der folgende Dialog etwa wird von Weizenbaum selbst zur Veranschaulichung der Leistungen von ELIZA angeführt.² (In diesem Dialog sind die Sätze in Normalschrift die über eine Tastatur in den Computer

eingeegebenen Äußerungen des menschlichen Partners, der sich in der Situation einer Psychotherapie glaubt, und die Sätze in Großbuchstaben die am Bildschirm ausgegebenen Antworten des Computers, der den Part des Therapeuten spielt.)

Men are all alike.

IN WHAT WAY

They're always bugging us about something or other.

CAN YOU THINK OF A SPECIFIC EXAMPLE

Well, my boyfriend made me come here.

YOUR BOYFRIEND MADE YOU COME HERE

He says I'm depressed much of the time.

I AM SORRY TO HEAR YOU ARE DEPRESSED

It's true. I am unhappy.

DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY

I need some help, that much seems certain.

WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP

Perhaps I could learn to get along with my mother.

TELL ME MORE ABOUT YOUR FAMILY

My mother takes care of me.

WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU

My father.

YOUR FATHER

You are like my father in some ways.

WHAT RESEMBLANCE DO YOU SEE

You are not very aggressive but I think you don't want me to notice that.

WHAT MAKES YOU THINK THAT I AM NOT VERY AGGRESSIVE

You don't argue with me.

WHY DO YOU THINK I DON'T ARGUE WITH YOU

You are afraid of me.

DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU

My father is afraid of anybody.

WHAT ELSE COMES TO MIND WHEN YOU THINK OF YOUR FATHER

Bullies.

DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE

Erstaunlicher noch als dieser Dialog ist jedoch, mit welchen einfachen Mitteln ELIZA seine Leistungen erzielt.³ Vorab wird dem Programm für die verschiedenen Situationen, zu denen es in ei-

nem Dialog Antworten erzeugen kann, ein *script* eingegeben. Dieses *script* enthält eine Liste von Schlüsselwörtern und für jedes dieser Schlüsselwörter eine Reihe von sprachlichen Mustern, nach denen das Programm sucht, und zu jedem dieser Muster eine Reihe von Transformationsregeln. Zum Schlüsselwort »ME« zum Beispiel gehört das Muster

(* YOU * ME),

und zu diesem Muster die Transformationsregel

(WHAT MAKES YOU THINK THAT I 3 YOU)

Dabei steht »*« für beliebigen Text und »3« in der Transformationsregel für das dritte Element der Eingabe (in diesem Fall also für alles zwischen »YOU« und »ME«). Die eingegebenen Sätze werden von links nach rechts nach Schlüsselwörtern durchsucht. Wenn mehrere Schlüsselwörter gefunden werden, werden sie in der Reihenfolge ihrer Wichtigkeit in einem Schlüsselwort-Speicher abgelegt. Im nächsten Schritt wird jeder eingegebene Satz mit den Mustern verglichen, die mit dem ersten der im Schlüsselwort-Speicher abgelegten Schlüsselwörter verbunden sind. Dabei wird zuerst nach schwierigeren Mustern gesucht. Denn für das zum Schlüsselwort »I« gehörende Muster

(* I * DEPRESSED *)

wird sich schwerer eine Entsprechung finden lassen als für das Muster

(* I *).

Wenn eine Übereinstimmung gefunden ist, generiert ELIZA mit Hilfe einer zugehörigen Transformationsregel eine Antwort, wobei eine zusätzliche Komponente dafür sorgt, daß nicht immer derselbe Antworttyp erscheint. Falls keine Übereinstimmung gefunden wird, wird das Verfahren mit dem nächsten Wort aus dem Schlüsselwort-Speicher wiederholt. Wenn der Schlüsselwort-Speicher leer ist, werden Antworten wie »PLEASE GO ON« oder »VERY INTERESTING« erzeugt. Außerdem verfügt ELIZA bei der Erzeugung von Antworten noch über einige zusätzliche Techniken: So kann das Programm zum Beispiel Schlüsselwörter in der Antwort ersetzen, Schlüsselwörter mit Klassen oder Situationen in Beziehung bringen (»MOTHER« impliziert »FAMILY«) oder im Laufe der »Unterhaltung« auf frühere Schlüsselwörter zurückkommen.

Keine Frage, all das ist sehr geschickt ausgedacht, und es führt auch, wie Weizenbaums Beispiel zeigt, zu sehr eindrucksvollen

Ergebnissen. Dennoch ist ELIZA aber sicher kein System, von dem man zu Recht sagen könnte, daß es die eingegebenen Sätze versteht. Das wird besonders an den folgenden drei Punkten deutlich.

(a) ELIZA kann syntaktisch wohlgeformte nicht von syntaktisch falschen Sätzen unterscheiden. Denn wie das System auf die Eingabe

You are not very aggressive
mit

WHAT MAKES YOU THINK THAT I AM NOT VERY AGGRESSIVE
reagiert, so reagiert es auf die Eingabe

You are not aggressive very
mit

WHAT MAKES YOU THINK THAT I AM NOT AGGRESSIVE VERY.

(b) ELIZA kann semantisch unsinnige Sätze nicht von semantisch sinnvollen Sätzen unterscheiden. Das System murrst nicht, wenn man zum Beispiel den Satz eingibt

Please ask the table to pick up a brick for the baby.⁴
Und auf die Eingabe

You always aboabo me
antwortet ELIZA schlicht mit

WHY DO YOU THINK THAT I ALWAYS ABOABO YOU.

(c) ELIZA kann auf die verschiedenen Sprechakte des Partners nur stereotyp und oft unangemessen reagieren. Auf Fragen zum Beispiel kann es nur mit

WHY DO YOU ASK
oder mit einem Rückgriff auf frühere Themen antworten. Wirkliche Antworten geben kann das System nicht. Auch auf Bitten kann es nicht eingehen, Befehle nicht ausführen. Es kann keine oder kaum Schlüsse ziehen. Behauptungen sind irgendwann vergessen usw. Ein eindrucksvolles Beispiel für dieses stereotype und meist unsinnige Verhalten ist der folgende Dialog⁵, der von einem Partner ausgelöst wurde, der einfach nur den Computer benutzen wollte.

Do you think that I can use the computer this morning?

WHY DO YOU ASK

I might be able to make some additional telcomp sales.

WHY AREN'T YOU SURE

My prospective customers have never seen a demonstration of the system.

DOES THIS MEAN VERY MUCH TO YOU

Of course it does.

YOU SEEM VERY POSITIVE ABOUT THAT

This is ridiculous.

WHAT DO YOU MEAN THIS IS RIDICULOUS

Is there any reason why I can't use the computer now?

TELL ME MORE ABOUT YOUR PROSPECTIVE CUSTOMERS

Please dial me up on 491-1850.

Natürlich kam ELIZA dieser Aufforderung nicht nach.

ELIZA ist also wirklich nur eine syntaktische Maschine. Und was noch schwerer wiegt: eine syntaktische Maschine, die eine semantische Maschine bestenfalls sehr unvollkommen imitiert. Doch dies ist natürlich noch kein Beweis für die These, daß semantische Maschinen prinzipiell unmöglich sind. Denn die Entwicklung natürlichsprachlicher Systeme (NSS) ist nicht bei ELIZA stehen geblieben. Sie hat vielmehr beträchtliche Fortschritte gemacht, so daß sich heute die Frage stellt, ob nicht wenigstens diese neueren Systeme als echte semantische Systeme betrachtet werden können.

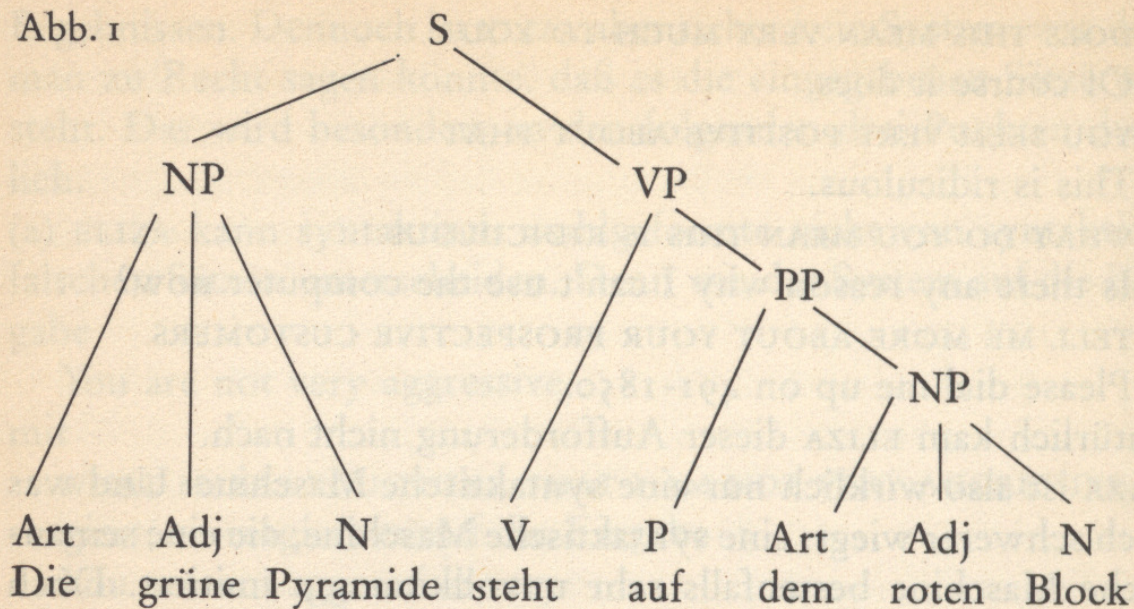
3. Eine Antwort auf diese Frage ist sicher nur möglich, wenn klar ist, wie sich die neueren Systeme von einem System wie ELIZA unterscheiden. Ich will deshalb versuchen, die Arbeitsweise dieser Systeme hier zumindest in ihren Grundzügen kurz darzustellen.⁶

In der Regel beruht die Arbeitsweise aller neueren NSS auf der Tatsache, daß sich natürlichsprachliche Sätze mit Hilfe einer Phasenstrukturgrammatik (PS-Grammatik) analysieren lassen, das heißt, daß man – sofern ein Satz S syntaktisch eindeutig ist (wovon ich im folgenden der Einfachheit halber ausgehen werde) – diesem Satz mit Hilfe eines Parsers in eindeutiger Weise eine strukturelle Beschreibung, zum Beispiel einen PS-Baum zuordnen kann. So entspricht etwa dem Satz

(1) Die grüne Pyramide steht auf dem roten Block
der in Abb. 1 gezeigte Strukturbaum.

Die Frage ist nun, wie ein Programm von einem solchen Struktur-Baum ausgehend in einem Prozeß der *semantischen Analyse* eine interne Repräsentation aufbauen kann, die den Sinn dieses Satzes wiedergibt. Ich will anhand eines einfachen Beispiels versuchen, eine Antwort auf diese Frage wenigstens zu skizzieren.

Abb. 1



Gehen wir aus von einem System S , das über eine interne Repräsentation der in der Abb. 2 gezeigten Blockwelt-Szene verfügt. Eine solche Repräsentation könnte zum Beispiel die Form propositionaler Listen haben, das heißt in dem System S könnten die für diese Szene wesentlichen Fakten in der folgenden Datenbasis gespeichert sein:

- (IST-EIN OBJEKT-1 BLOCK)
- (FARBE OBJEKT-1 ROT)
- (IST-EIN OBJEKT-2 PYRAMIDE)
- (FARBE OBJEKT-2 GRÜN)
- (STEHT-AUF OBJEKT-2 OBJEKT-1)
- (IST-EIN OBJEKT-3 BLOCK)
- (FARBE OBJEKT-3 BLAU)

Wie kann für dieses System eine interne Repräsentation des Satzes (1) aussehen, von der man sinnvollerweise sagen kann, daß sie den Sinn dieses Satzes wiedergibt? Wenn wir selbst diesen Satz

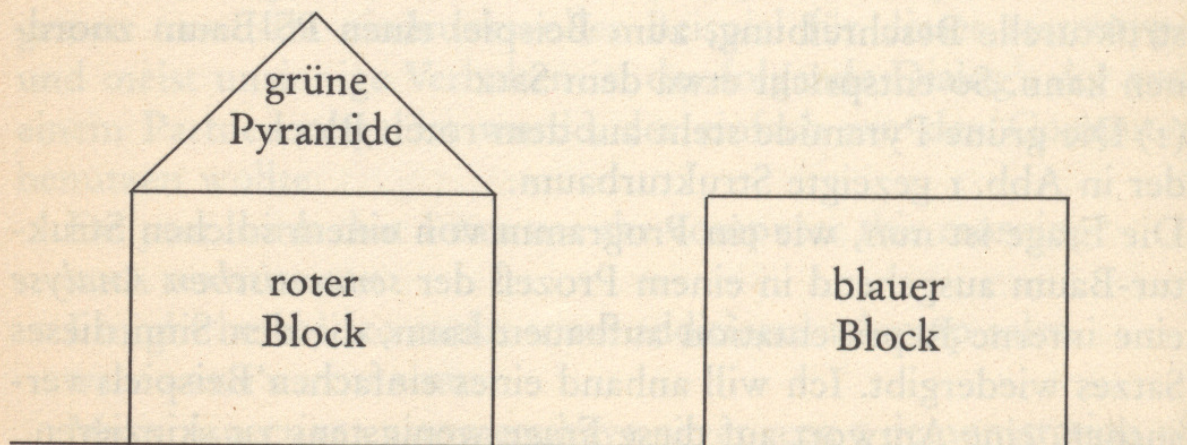


Abb. 2: Blockwelt-Szene

interpretieren, dann ist uns klar, daß in ihm von zwei Objekten die Rede ist – den Objekten, auf die sich die N-Phrasen »die grüne Pyramide« und »dem roten Block« beziehen – und daß der Satz besagt, daß diese beiden Objekte in einer bestimmten Relation zueinander stehen. Für die Interpretation des Satzes ist es also zunächst einmal wichtig, den Bezug der in ihm vorkommenden N-Phrasen zu klären. Fragen wir also zuerst, welche internen Repräsentationen im System *S* diese Aufgabe erfüllen können. Am Aufbau der Datenbasis, in der die für die in der Abb. 2 gezeigte Szene wesentlichen Fakten im System *S* repräsentiert sind, wird deutlich, daß *S* für alle an dieser Szene beteiligten Objekte interne Namen verwendet: für die Pyramide den Namen »OBJEKT-1«, für den roten Block den Namen »OBJEKT-2« und für den blauen Block den Namen »OBJEKT-3«. Um den Referenten der Phrase »die grüne Pyramide« zu finden, benötigt *S* daher eine Funktion, die als Wert den Namen des Objekts liefert, das zugleich grün und eine Pyramide ist. Diese Aufgabe kann aber zum Beispiel durch die folgende Funktion erfüllt werden:

(2) (FINDE-WERT '?X '(UND (IST-EIN ?X PYRAMIDE)
(FARBE ?X GRÜN)))

Man erhält also eine vernünftige Interpretation, wenn das System *S* bei der semantischen Analyse der N-Phrase »die grüne Pyramide« diese Funktion als interne Repräsentation zuordnet. Denn wenn *S* diese Funktion aufruft, ergibt sich als Wert der Name »OBJEKT-2«, also der interne Name des Objekts, auf das sich die interpretierte N-Phrase bezieht.

Beim Aufbau einer solchen Interpretation kann *S* zum Beispiel so vorgehen: Im internen Lexikon von *S* sind die Wörter »grün« und »Pyramide« mit den Klauseln

(3) (FARBE ?X GRÜN)

und

(4) (IST-EIN ?X PYRAMIDE)

assoziiert. Da »grüne Pyramide« im PS-Baum als Adj-N Komponente analysiert wurde, deutet *S* diesen Ausdruck als Konjunktion, das heißt: ausgehend von diesen Einträgen baut die semantische Komponente von *S* als interne Repräsentation für diese beiden Wörter den Ausdruck

(5) (UND (IST-EIN ?X PYRAMIDE) (FARBE ?X GRÜN))

auf. Dies ist natürlich noch nicht die Interpretation der gesamten N-Phrase »die grüne Pyramide«. Denn es fehlt die Interpretation

des bestimmten Artikels. Mit diesem Artikel ist als Bedeutung im internen Lexikon von *S* zum Beispiel der Ausdruck

- (6) (FINDE-WERT (REFERENT \$NP)
(SINN \$NP))

verbunden, der bei der Gesamtinterpretation der N-Phrase »die grüne Pyramide« den Ausgangspunkt bildet. Dabei geht die semantische Komponente von *S* so vor, daß sie im Ausdruck (6) zunächst den Teilausdruck »(REFERENT \$NP)« durch die Variable »?x« ersetzt und dann den Teilausdruck »(SINN \$NP)« durch das Ergebnis der Analyse des Teilausdrucks »grüne Pyramide«, das heißt also durch den Ausdruck (5). Auf diese Weise entsteht als Repräsentation der gesamten NP der Ausdruck (2).

In ähnlicher Weise wird auch die Repräsentation für den ganzen Satz (1) aufgebaut. Dabei beginnt die semantische Komponente von *S* mit dem Verb »steht«, für das sich im Lexikon zum Beispiel der Ausdruck

- (7) (STEHT-AUF (REFERENT \$SUBJ-NP)
(REFERENT \$»AUF«-PP-NP))

findet. Beim Aufbau der Repräsentation für den ganzen Satz (1) wird – ausgehend von diesem Ausdruck – im ersten Schritt die Subj-NP analysiert, wobei sich als Ergebnis der Ausdruck (2) ergibt. Durch diesen Ausdruck wird in (7) der Teilausdruck »(REFERENT \$SUBJ-NP)« ersetzt. Im zweiten Schritt wird die in die PP eingebettete NP analysiert, wobei sich – analog zum Ausdruck (2) – als Ergebnis der Ausdruck ergibt:

- (8) (FINDE-WERT ' ?X '(UND (IST-EIN ?X BLOCK)
(FARBE ?X ROT)))

Durch diesen Ausdruck wird in (7) der Teilausdruck »(REFERENT \$»AUF«-PP-NP)« ersetzt, so daß sich als Ergebnis der semantischen Analyse des Satzes (1) der Ausdruck ergibt:

- (9) (STEHT-AUF (FINDE-WERT ' ?X '(UND (IST-EIN ?X PYRAMIDE)
(FARBE ?X GRÜN)))
(FINDE-WERT ' ?X '(UND (IST-EIN ?X BLOCK)
(FARBE ?X ROT))))

Das heißt, da die semantische Komponente von *S* auch noch berücksichtigt, daß es sich hier um einen Behauptungssatz handelt, wird dieser Repräsentation noch ein »BEHAUPT« vorangestellt, so daß letzten Endes die folgende Repräsentation entsteht:

- (1*) (BEHAUPT '(STEHT-AUF
(FINDE-WERT ' ?X '(UND (IST-EIN ?X PYRAMIDE)

(FARBE ?X GRÜN)))
(FINDE-WERT ' ?X '(UND (IST-EIN ?X BLOCK)
(FARBE ?X ROT))))))

4. Ich hatte im letzten Abschnitt den Aufbau entsprechender interner Repräsentationen als semantische Analyse bezeichnet. Aber es ist natürlich eine legitime Frage, ob dieser Sprachgebrauch angemessen ist, das heißt, ob es sich bei dem im letzten Abschnitt geschilderten Verfahren tatsächlich um eine Analyse der Bedeutung von Sätzen handelt. Und die Antwort auf diese Frage ist natürlich für unsere Ausgangsfrage von entscheidender Bedeutung, ob neuere NSS – im Gegensatz zu Systemen wie ELIZA – tatsächlich semantische Maschinen sind oder ob es sich auch bei diesen Systemen nur um bloß syntaktische Maschinen im Sinne Dennetts handelt.

In gewisser Weise sind natürlich auch die neueren NSS syntaktische Maschinen. Sie nehmen Zeichenreihen auf und transformieren sie in andere Zeichenreihen – genauso wie Searle dies in seinem Chinese-Room-Beispiel eindrücklich beschrieben hat. Aber folgt aus dieser Tatsache, daß sie nicht auch semantische Maschinen sein können? Schließen sich diese beiden Charakterisierungen gegenseitig aus? Kommen wir, um uns einer Antwort auf diese Fragen anzunähern, zunächst zurück auf Dennetts Problem, ein System zu konstruieren, das aus den eingehenden Telegraphenmeldungen genau alle Morddrohungen aussondert. Dennett war der Meinung, daß dies nicht möglich sei. Und das Hauptargument für seine These war, daß es unmöglich sei, die semantische Kategorie »In Englisch verfaßte Morddrohung« syntaktisch zu definieren. Hinter diesem Argument scheint die generelle Überzeugung zu stehen, daß es grundsätzlich unmöglich sei, ein formales Programm zu entwickeln, das für beliebige Sätze einer natürlichen Sprache entscheidet, ob sie dieselbe Bedeutung haben oder nicht. Wenn man sich die im letzten Abschnitt geschilderte Arbeitsweise neuerer NSS ansieht, gewinnt man jedoch den Eindruck, daß diese Systeme genau dies leisten. Denn die syntaktischen und die semantischen Komponenten dieser Systeme arbeiten gerade so zusammen, daß zum Beispiel dem Satz
(10) John schlägt Jack
dieselbe interne Repräsentation zugeordnet wird wie dem entsprechenden passivischen Satz
(10') Jack wird von John geschlagen.

Nämlich der Ausdruck:

(10*) (BEHAUPT '(UND (IST-EIN EREIGNIS-1 SCHLAGEN)
(AGENT JOHN-1)
(PATIENT JACK-1)))

Analog wird dem der Aussage (1) entsprechenden Fragesatz

(11) Steht die grüne Pyramide auf dem roten Block?

eine Repräsentation zugeordnet, die sich nur im Hinblick auf die Repräsentation der Kraft des Satzes, nicht jedoch im Hinblick auf die Repräsentation des propositionalen Gehalts von der Repräsentation (1*) unterscheidet, nämlich der Ausdruck

(11*) (FRAGE '(STEHT-AUF
(FINDE-WERT '?X '(UND (IST-EIN ?X PYRAMIDE)
(FARBE ?X GRÜN)))
(FINDE-WERT '?X '(UND (IST-EIN ?X BLOCK)
(FARBE ?X ROT))))).

Alle diese Beispiele zeigen, daß neuere NSS im Idealfall in der Lage sind, bedeutungsgleiche von bedeutungsverschiedenen sprachlichen Ausdrücken zu unterscheiden. Denn diese Systeme arbeiten so, daß sie sprachlichen Ausdrücken dann und nur dann dieselbe interne Repräsentation zuordnen, wenn sie dieselbe Bedeutung haben.

An dieser Stelle könnte man jedoch einwenden: Möglicherweise gelingt es neueren NSS, bedeutungsgleiche von bedeutungsverschiedenen sprachlichen Ausdrücken zu unterscheiden, indem sie bedeutungsgleichen Ausdrücken genau dieselbe interne Repräsentation zuordnen; aber daraus folgt noch nicht, daß sie die Bedeutung sprachlicher Ausdrücke auch wirklich verstehen. Denn in der Tat scheint es doch möglich zu sein, zu entscheiden, ob zwei sprachliche Ausdrücke *A* und *B* dieselbe Bedeutung haben, ohne die Bedeutung von *A* und *B* zu kennen.

Gegen diesen Einwand gibt es meiner Meinung nach zwei Erwiderungen: eine defensive und eine offensive. Die defensive knüpft an die Überlegungen Dennetts an. Dennett hatte das Problem der kognitiven Psychologie so geschildert, daß diese Wissenschaft annehme, das Gehirn könne auf die Bedeutungen eingehender Reize reagieren, während es *de facto* doch nur eine syntaktische Maschine sei. Wenn es aber syntaktische Maschinen gibt, die bedeutungsgleiche von bedeutungsverschiedenen Reizen unterscheiden können, dann gibt es hier gar kein Problem. Denn auf die Bedeutung eines Reizes reagieren zu können heißt ja zunächst einmal

nicht mehr, als in der Lage zu sein, auf alle Reize mit einer bestimmten Bedeutung und nur auf diese Reize in gleicher Weise zu reagieren. Und diese Fähigkeit hängt offenbar nur von der Fähigkeit ab, bedeutungsgleiche von bedeutungsverschiedenen Reizen zu unterscheiden.

Ich halte diesen Punkt für äußerst wichtig. Aber vielleicht erscheint er manchen doch nicht ausreichend. Kommen wir also zu der offensiveren Erwiderung. Diese Erwiderung geht von der Gegenfrage aus: Wenn die Unterscheidung bedeutungsgleicher von bedeutungsverschiedenen Ausdrücken nicht ausreicht, was mehr muß ein System können, um die Bedeutung eines Ausdrucks wirklich zu erfassen? Die naheliegende Antwort ist: Es muß wissen, auf was sich die sprachlichen Ausdrücke beziehen, und das heißt: wissen, welcher Gegenstand von einem definiten Gegenstandsbezeichner bezeichnet wird; wissen, auf welche Eigenschaft sich zum Beispiel das Adjektiv »grün« bezieht; wissen, unter welchen Bedingungen Sätze wahr sind; usw. Ich halte diese Antwort nicht nur für naheliegend, sondern auch für richtig. Die Frage ist nur, was hier mit »wissen« gemeint ist.

Nehmen wir als besonders zentralen Punkt das Wissen um die Wahrheitsbedingungen eines Satzes. Unter welchen Bedingungen kann man von einem System zu Recht sagen, daß es die Wahrheitsbedingungen eines Satzes kennt? Was heißt es, zu wissen, unter welchen Bedingungen ein Satz *S* wahr ist?

Meiner Meinung nach sind in diesem Zusammenhang zwei Fähigkeiten entscheidend: nämlich erstens die Fähigkeit, Situationen, in denen *S* wahr ist, von Situationen zu unterscheiden, in denen das nicht der Fall ist; und zweitens die Fähigkeit, den Satz *S* systematisch mit Situationen der ersten Art in Verbindung zu bringen. Wenn das so ist, dann folgt daraus aber – und besonders der erste Punkt macht dies deutlich –, daß für die Fähigkeit, Sprache zu verstehen, nicht nur die Sprachkomponenten, sondern besonders auch die Wahrnehmungskomponenten entscheidend sind. Denn um zum Beispiel Situationen, in denen der Satz »Die grüne Pyramide steht auf dem roten Block« wahr ist, von solchen Situationen unterscheiden zu können, in denen das nicht der Fall ist, muß das System feststellen können, ob es in seiner Umgebung genau eine grüne Pyramide und genau einen roten Block gibt und ob der erste Gegenstand auf dem zweiten steht. Dazu benötigt es jedoch eine Wahrnehmungskomponente, die komplex genug ist,

um folgendes leisten zu können: Sie muß erstens in der Lage sein, Einzelgegenstände in der Umgebung des Systems zu identifizieren (das heißt, sie vom Umgebungshintergrund und von anderen Einzeldingen abzugrenzen); sie muß zweitens in der Lage sein, die identifizierten Gegenstände ihrer Art und ihrer Farbe entsprechend zu klassifizieren (also zum Beispiel Pyramiden von Nicht-Pyramiden und grüne von nicht-grünen Gegenständen zu unterscheiden); und sie muß drittens in der Lage sein, einfache räumliche Relationen, die zwischen den identifizierten Gegenständen in ihrer Umgebung bestehen, zu erkennen (also zum Beispiel, daß ein Gegenstand auf einem anderen steht).

Wenn wir annehmen, daß ein System mit der in Abb. 1 gezeigten Szene konfrontiert ist, reicht es also nicht aus, daß dieses System über eine Kamera verfügt, die auf diese Szene gerichtet ist. Denn durch die Kamera wird im System zunächst nur ein Bild der Szene erzeugt, das zum Beispiel in einer Grauwertmatrix zwischengespeichert wird. Dieser Vorgang selbst kann sicher noch nicht als Wahrnehmung bezeichnet werden. Wahrnehmung beginnt vielmehr erst, wenn durch Verarbeitung der in der Grauwertmatrix gespeicherten Daten ein internes Modell der Szene aufgebaut wird. Dabei muß die Verarbeitung, wie gerade schon erwähnt, folgendes leisten: Sie muß erstens als Ergebnis liefern, daß an der Szene drei verschiedene Gegenstände beteiligt sind (OBJEKT-1, OBJEKT-2 und OBJEKT-3); sie muß zweitens die Gegenstände der Art und der Farbe nach richtig klassifizieren; und sie muß schließlich zu dem Ergebnis führen, daß der erste Gegenstand auf dem zweiten steht (aber zum Beispiel nicht der zweite auf dem dritten). Das heißt, die Verarbeitung muß letzten Endes zum Aufbau eines internen Modells dieser Szene führen, das etwa die Form der im Abschnitt 3 angeführten propositionalen Listen haben könnte.⁷

Wenn unser System über eine Wahrnehmungskomponente verfügt, die in der Lage ist, solch ein internes Modell der Umgebung des Systems aufzubauen, dann sieht man aber sofort, wie – auf dem Wege über dieses Modell – die sprachlichen und die Wahrnehmungskomponenten des Systems so zusammenarbeiten können, wie es für wirkliches Sprachverstehen nötig ist.

Dies wird beispielsweise deutlich, wenn wir noch einmal den Fragesatz

(11) Steht die grüne Pyramide auf dem roten Block?

betrachten, zu dem das System, wie schon gesagt, die interne Repräsentation

(II*) (FRAGE (STEHT-AUF
(FINDE-WERT '?X '(UND (IST-EIN ?X PYRAMIDE)
(FARBE ?X GRÜN)))
(FINDE-WERT '?X '(UND (IST-EIN ?X BLOCK)
(FARBE ?X ROT))))))

aufbaut. Denn diese Repräsentation wird von dem System als Aufforderung interpretiert, auf der Grundlage der gegebenen Datenbasis den Ausdruck

(9) (STEHT-AUF (FINDE-WERT '?X '(UND (IST-EIN ?X PYRAMIDE)
(FARBE ?X GRÜN)))
(FINDE-WERT '?X '(UND (IST-EIN ?X BLOCK)
(FARBE ?X ROT))))))

zu beweisen. Dazu wertet es zunächst die beiden Ausdrücke

(2) (FINDE-WERT '?X '(UND (IST-EIN ?X PYRAMIDE)
(FARBE ?X GRÜN)))

und

(8) (FINDE-WERT '?X '(UND (IST-EIN ?X BLOCK)
(FARBE ?X ROT)))

aus, das heißt, das System untersucht seine Datenbasis daraufhin, ob es genau ein Objekt gibt, das sowohl den Ausdruck »(IST-EIN ?X PYRAMIDE)« als auch den Ausdruck »(FARBE ?X GRÜN)« erfüllt, bzw. genau ein Objekt, das sowohl den Ausdruck »(IST-EIN ?X BLOCK)« als auch den Ausdruck »(FARBE ?X ROT)« erfüllt. Im ersten Fall findet es das OBJEKT-2 und im zweiten Fall das OBJEKT-1 und ersetzt daher im Ausdruck (9) die Teilausdrücke (2) und (8) durch die internen Namen dieser Objekte, wodurch der Ausdruck

(9') (STEHT-AUF OBJEKT-2 OBJEKT-1)

entsteht. Im nächsten Schritt versucht das System nun, diesen Ausdruck zu beweisen, was sofort gelingt, da dieser Ausdruck Teil der Datenbasis ist. Das Ergebnis der Gesamtoperation ist also positiv. Und deshalb antwortet das System mit »Ja«. Falls das Ergebnis negativ gewesen wäre, hätte es mit »Nein« geantwortet. Das Zusammenspiel von sprachlichen und Wahrnehmungskomponenten ist somit offensichtlich. Die sprachlichen Komponenten bewirken, daß das System auf die Frage (II) genau dann mit »Ja« antwortet, wenn es den Ausdruck (9') in der Datenbasis findet. Und die Wahrnehmungskomponente bewirkt, daß dieser

Ausdruck genau dann in die Datenbasis aufgenommen wird, wenn in der Szene, auf die die Kamera des Systems gerichtet ist, die grüne Pyramide tatsächlich auf dem roten Block steht. Letzten Endes antwortet das System auf die Frage (11) also genau dann mit »Ja«, wenn die Wahrheitsbedingungen des propositionalen Teils dieses Satzes erfüllt sind. Das heißt, es kann nicht nur Situationen, in denen der entsprechende Aussagesatz (1) wahr ist, von Situationen unterscheiden, in denen das nicht der Fall ist, es bringt diese Situationen auch in der richtigen Weise mit dem Satz (11) in Verbindung. Damit aber sind, denke ich, alle Voraussetzungen erfüllt, um mit Recht sagen zu können, daß es diesen Satz wirklich verstanden hat. Systeme dieser Art sind also, wie mir scheint, wirklich semantische Maschinen.

Anmerkungen

- 1 Dennoch unterscheidet sich die Position Dennetts natürlich erheblich von den Positionen Searles oder Dreyfus'. Denn die referierte Argumentation macht deutlich, daß Dennett auch das menschliche Gehirn für eine rein syntaktische Maschine hält. Dennett geht es also nicht um einen möglichen Unterschied zwischen Menschen und Maschinen, sondern um das ganz anders geartete Problem, welchen Platz Semantik und Intentionalität in einem Universum haben können, in dem es nur syntaktische Mechanismen gibt. Ähnlich wie Searle scheint jedoch auch Dennett anzunehmen, daß zwischen syntaktischen und semantischen Maschinen eine unüberbrückbare Kluft besteht und daß daher syntaktische Maschinen niemals zugleich auch semantische Maschinen sein können.
- 2 Weizenbaum (1976), S. 3 f.
- 3 Die folgende Beschreibung von ELIZA folgt im wesentlichen Barr/Feigenbaum (1981), S. 285-287.
- 4 Dieser Beispielsatz stammt von Margaret Boden; siehe Boden (1977), S. 108.
- 5 Dieser Dialog ist zitiert nach Boden (1977), S. 96.
- 6 Vgl. zum Folgenden Charniak/McDermott (1985), Kap. 4.
- 7 Wahrnehmungskomponenten, die dies leisten, konnten zwar bisher noch nicht so weit entwickelt werden wie die entsprechenden sprachlichen Komponenten. Aber die Grundzüge der Arbeitsweise visueller Systeme sind doch so weit bekannt, daß zumindest im Prinzip klar ist, wie solche Systeme konstruiert werden können. Vgl. etwa Charniak/McDermott (1985), Kap. 3.

Literatur

- Barr, A., und E. A. Feigenbaum (1981), *Handbook of Artificial Intelligence*, 3 Bde., Los Altos, Cal.: William Kaufmann.
- Boden, M. (1977), *Artificial Intelligence and Natural Man*, Hassocks, Sussex: Harvester Press.
- Charniak, E., und D. McDermott (1985), *Introduction to Artificial Intelligence*, Reading, Mass.: Addison-Wesley.
- Dennett, D. (1981), »Three Kinds of Intentional Psychology«, in: R. Healey (Hg.), *Reduction, Time and Reality*, Cambridge: Cambridge University Press, S. 37-61.
- Weizenbaum, J. (1976), *Computer Power and Human Reason*, San Francisco: W. H. Freeman and Company; dt: *Die Macht der Computer und die Ohnmacht der Vernunft*, Frankfurt: Suhrkamp 1977.