Originalveröffentlichung in: Bielefelder Linguistik. Linguistik. Die Bielefelder Sicht. – Bielefeld: Aisthesis Verlag, 1997. S. 85-88.

Sprache, Wahrnehmung und Handlung

16

Jan-Torsten Milde, Henning Lobin und Dieter Metzing

Als Teil der Linguistik untersucht die Computerlinguistik Sprache mit Hilfe linguistischer und informationstechnischer Verfahren. Die Vorgehensweise bei der Untersuchung des Problems „Sprache" orientiert sich an den in den klassischen Naturwissenschaften gängigen Methoden: Ein auf empirischen Beobachtungen aufbauendes theoretisches Modell wird in eine algorithmische Entsprechung überführt und kann so durch Simulation auf einem Computer validiert werden.

Diese Vorgehensweise hat in der Vergangenheit dazu geführt, daß Sprache als isoliertes Phänomen, ohne weitergehende Einbeziehung des extralinguistischen Kontextes, betrachtet wurde. Dieser Effekt wurde noch verstärkt durch die Nutzung von transkribierten und „bereinigten" Auschnitten natürlicher Sprache als Eingabe in die Computerprogramme.

Will man, im Gegensatz dazu, Sprachverstehensprozesse modellieren, wie sie etwa in einer Instruktionssituation auftreten, so ist die Integration von situativen Informationen in die Sprachanalyse unumgänglich. Hierbei sind

- *Diskurswissen* (sprachlicher Kontext),
- *Handlungskontext* und *Handlungswissen*,
- *aktive sensorische Wahrnehmung* (externer Weltzustand),
- *somatosensorische Wahrnehmung* (interner Zustand)

für ein Verständnis notwendige Teilinformationen. In typischen Gesprächssituationen sind Sprache, Wahrnehmung und Handlung eng verknüpft.

Die Einbeziehung von Diskurswissen wird z.B. notwendig bei der Auflösung von *Referenzen* auf früher eingeführte Objekte. In der Regel wird ein bereits eingeführtes Objekt im weiteren Verlauf nur noch unvollständig (*elliptisch*) sprachlich spezifiziert.

Betrachtet man derartige Äußerungen isoliert, ohne Einbeziehung des Kontextes, so ist eine korrekte Interpretation i.A. nicht mehr zur erreichen. Gleiches gilt für den Handlungskontext. Eingeführte Teilhandlungen werden vom Instruktor sprachlich nur so weit beschrieben, dass der Instruierte in der Lage ist, diese durchzuführen. Nur durch Nutzung von explizitem Handlungswissen wird es möglich, auf nachfolgende natürlichsprachige Referenzen adäquat zu reagieren. Lässt man Bezüge auf Objekte der externen Umwelt zu, so muss das sprachhandelnde System in der Lage sein, die Umwelt aktiv sensorisch zu erfassen und die Ergebnissse dieser Wahrnehmung mit sprachlichen Informationen zu kombinieren. Weiterhin spielt der interne Zustand des Agenten eine Rolle, da dieser Einfluss auf möglich Teilhandlungen haben kann.

16.1 Sprache als Handlungskomplement

Bei der Betrachtung von Sprachverwendung in kooperativen Handlungen sind die Voraussetzungen der Teilnehmer bezüglich dieser Handlungen zu berücksichtigen. Die Kooperation kann zum einen zwischen Partnern erfolgen, die gleiche oder ähnliche Vorstellungen von der Durchführung der Handlung besitzen, oder zwischen einem 'Meister' und seinem 'Lehrling'.

Empirische Untersuchungen zu kooperativen Handlungen zeigen, dass im Falle der Übereinstimmung des Wissens über Planung und Ausführung einer Aufgabe sprachliche Kommunika-

Publikationsserver des Instituts für Deutsche Sprache
URN: http://nbn-resolving.de/urn:nbn:de:bsz:mh39-76262

tion keineswegs Teile der Aufgabenausführung sprachlich substituiert, sondern ihr die Funktion zukommt, den Planungsprozess zu koordinieren und die Integration der Agenten in ihr sensomotorisches Umfeld zu steuern. Es lassen sich dann im Wesentlichen drei Funktionen voneinander unterscheiden: *Exploration, Instruktion* und *Aufmerksamkeitssteuerung*:

Äußerungen mit vorwiegend *explorativer* Funktion koordinieren die Planungen in Situationen, in denen das vorhandene Planungswissen den möglichen Fortgang nicht ausreichend determiniert, sei es dass unvorhergesehene Momente ins Spiel kommen oder sei es dass die Datenlage nicht ausreicht, die Situation einem Planungszustand zu subsumieren („*Versuchen wir es einmal mit dem Schraubenzieher.*").

Äußerungen mit *instruierender* Funktion („*Hol den Schraubenzieher!*") dienen bei gleicher Planungskompetenz beider Agenten der Synchronisierung der Planausführung. Die aufmerksamkeitssteuernde Funktion von Äußerungen schließlich organisiert die Einbettung einer Handlung in ihr sensomotorisches Umfeld. Das sensomotorische Umfeld kooperierender Agenten differiert hinsichtlich der Wahrnehmungsperspektive und der gerade ausgeführten Teilhandlungen. Aufmerksamkeitssteuernde Äußerungen beeinflussen diese Wahrnehmungsperspektive in einer für die Ausführung des globalen Plans relevanten Weise. Dies kann insbesondere durch deiktische Akte geschehen („*Rechts!*"), die die Aufmerksamkeit direkt steuern, oder durch Warnungen, bei der die intendierte Änderung der Wahrnehmungsperspektive entweder vollständig („*Achtung!*") oder teilweise („*Eine Wand!*") zu inferieren ist.

Von diesen aktiven Formen von *Aufmerksamkeitssteuerung* sind passive Formen zu unterscheiden, bei der sich ein Agent seiner defizitären Wahrnehmungsperspektive bewusst ist und deshalb versucht, sie durch Fragen an die Kooperationspartner an die momentanen Erfordernisse anzupassen („*Links oder rechts?*").

Liegt zwischen den Partnern einer Kooperation eine ungleichgewichtige Handlungskompentenz vor, kommen zu den oben genannten Funktionen zwei weitere hinzu: *Planungseinschränkung* und *Generalisierung*:

Die planungseinschränkende Funktion von Sprache korreliert mit der Auffassung, dass Pläne niemals zureichende Handlungsspezifikationen darstellen, sondern stets unterspezifiziert sind (Winograd & Flores, 1986). Viele Handlungen, etwa das Annähen eines Knopfes, können wesentlich besser durch Vormachen vermittelt werden als verbal oder sonstwie mit symbolischen Mitteln.

Durch Vorführen wird demnach der empirisch zugängliche Aktionsteil spezifiziert, durch explizite symbolische Repräsentation der abstrakte, von den empirischen Daten unabhängige. Die demonstrierende Ausführung einer Handlung stellt sensorisches Rohmaterial zur Verfügung, das verbal durch *Verallgemeinerung,* Ausnahmebestimmungen, Toleranzen, Ausführungsrichtwerte usw. zu einer vollständigen Montageanleitung abstrahiert wird, die dadurch unter vielen verschiedenen Rahmenbedingungen anwendbar wird.

16. 3 Intelligente Sprachhandlungssysteme

Von einem übergeordneten Standpunkt betrachtet lassen sich aus den vorangegangenen Bemerkungen einige Anforderungen formulieren, an denen sich *künstliche,* in *Handlungssysteme eingebettete kommunizierende Agenten* messen lassen müssen (vgl. auch Förster & Peters, 1995; Peters & Förster, 1995; Förster et al., 1995; Peters, 1994; Lobin, 1993b; Lobin, 1995; Lobin & Milde, 1994; Milde, 1995; Lobin, Förster, Milde & Müller, 1994).

Gleichwertige Sprach-, Handlungs- und Wahrnehmungskompetenz
Grundsätzlich muss für die Architektur künstlicher Kommunikatoren gelten, dass Sprach-, Handlungs- und Wahrnehmungskompetenz als Ganzes seine Intelligenz konstituieren. Ein einzelner Bereich sollte demzufolge nur so komplex sein, wie es zur Konstituierung der Gesamtintelligenz des Systems erforderlich ist. Dies ist nicht nur ein ökonomisches Desiderat, sondern fußt auf der Überzeugung, dass sich eine Kompetenz nicht ohne die je anderen überhaupt entfalten kann. Die Wahrnehmungsfähigkeit ist um so intelligenter, je integrativer der Sinn des Wahrzunehmenden zu erschließen ist. Der Sinn von Wahrnehmungen ist die Unterstützung von Planungsprozessen und die Kontrolle ihrer Ausführung sowie die Ermöglichung von Kommunikation durch akustische oder visuell gegebene Symbole. Eine Handlungskompetenz ist ohne die Wahrnehmungskompetenz also 'blind', eine intelligente Handlungsplanung bedarf der Einbettung in die Welt durch die Wahrnehmungsfähigkeit. Kommunikation schließlich ermöglicht die Planung und Ausführung von Handlungen durch mehrere kooperierende Agenten. Der Sinn kommunikativer Akte erschließt sich durch die steuernde Funktion in kooperativen Handlungen. Der Komplexitätsgrad von in solche Handlungen integrierten sprachlichen Äußerungen muss also dem Komplexitätsgrad möglicher Handlungen entsprechen.

Wechselseitige Ergänzung der Einzelkompetenz
Ein Großteil der in realen sprachlichen Äußerungen erscheinenden Phänomene entzieht sich der Beschreibung durch eine normative Grammatik. Dies hat zur Folge, dass zu ihrer Erfassung z.T. äußerst komplizierte Zusatzmechanismen entwickelt werden müssen, wie es z.B. bei elliptischen Konstruktionen der Fall ist. Auf der anderen Seite ist zu beobachten, dass es offensichtlich keinen kommunikativen Nachteil darstellt, derartige Konstruktionen zu verwenden – im Gegenteil: Die Tatsache, dass wir Ellipsen in stark wirkungsbezogenen Kommunikationssituationen (z.B. bei Warnungen oder Befehlen) bevorzugt äußern – also gerade in solchen Situationen, bei denen die Kommunikation eingebettet ist in durch andere kognitive Kompetenzen gesteuerte Aktivitäten –, legt die Vermutung nahe, dass komplexere grammatische Strukturen gerade auf diesem kommunikativen Basisinventar aufbauen und nicht vielmehr diese von jenen abgeleitet sind.

Für die Untersuchung des konkreten Phänomens *elliptischer* Äußerungen bedeutet dies beispielsweise, dass ihre syntaktische Struktur und damit ihre semantische Analyse nicht auf satzsyntaktische Vollformen zurückgeführt werden kann, sondern dass sie im Rahmen eines handelnden und perzipierenden Systems direkt zu interpretieren sind. Ein Ansatzpunkt dafür wäre z.B. die Interpretation von Ellipsen als 'teilversprachlichte' Konzepte oder Perzepte. Ellipsen wären danach als Konstrukte zu verstehen, die auf der Grenze zwischen sensomotorischer und Sprachkompetenz stehen, zu einem Teil versprachlicht, zu einem anderen Teil nur konzeptuell oder perzeptuell analysierbar.

Einbeziehung von Sozialbeziehungen
Eine Instruktion oder eine aufmerksamkeitssteuernde Äußerung erfordern eine soziale Beziehung zwischen den Beteiligten. Die Sozialbeziehung kann permanent bzw. institutionalisiert sein, insbesondere bei Instruktionen (institutionalisiertes Sozialgefälle zwischen Meister und Lehrling). Aufmerksamkeitssteuernde Äußerungen hingegen basieren im Idealfall auf einem allein aus der Situation gegebenen Sozialgefälle, d.h. in relevanter Weise bezüglich des Globalziels differierender Wahrnehmungsperspektiven. So kann z.B. nur ein solcher Agent eine Anzahl von Teilhandlungen koordinieren, der sich zu jeder Zeit über den Stand jeder Teilhandlung informieren kann. Eine Schwierigkeit besteht für die Beteiligten darin, zu entscheiden, wer für den nachfolgenden Handlungsabschnitt die Kooperation sprachlich koordinieren soll

Abbildung 1: Im Rahmen des Forschungsprojektes D1 „Kommunizierende Agenten" wurde ein natürlichsprachlich steuerbarer Roboter entwickelt. Dieser integriert für die korrekte Interpretation von Handlungsanweisungen Information über die aktuell ausgeführte Handlung sowie Wahrnehmungs- und Sprachinformation. Aufgabe des Robotors ist die Konstruktion eines Flugzeugs aus Baufix. Das System verfügt über Teilkompetenzen (z.B. Greifen eines Objekts), ist aber nicht in der Lage, das komplette Flugzeug selbständig zu bauen. Durch sprachliche Äußerungen eines Menschen kann der Roboter angeleitet werden.

und kann. Eine defizitäre Wahrnehmungsperspektive kann auch bedeuten, dass die überlegene Wahrnehmungsperspektive eines Partners nicht als eine solche erkannt werden kann.

Sprache als Mittel zum Zweck

Künstliche Kommunikatoren und die Szenarien, in denen sie agieren und kommunizieren, müssen so beschaffen sein, dass die Sprache nicht als ein Substitut für ein anderes Medium fungiert, das den gestellten Aufgaben möglicherweise sogar besser gerecht wird, sondern dass die Verwendung von Sprache die Bewältigung von Situationen ermöglicht, die ohne Sprache nicht oder nur streng deterministisch bewältigt werden können (s. Abbildung 1).

Die vorangehenden Abschnitte haben gezeigt, dass Sprachverarbeitung, die in einem realistischen Kontext stattfinden soll, sich dem Problem der Integration von Sprache, Wahrnehmung und Handlung stellen muß. Dies gilt insbesondere für die Konstruktion von Sprachverstehensmodulen für komplexe technische Systeme, etwa *Roboter*. Für die Computerlinguistik ergeben sich dabei neue Fragestellungen, die über die bislang verfolgten weit hinausreichen.