

Zur Rolle der Linguistik bei der Entwicklung natürlichsprachlicher KI-Systeme

W. Wahlster, FB10-Informatik, Universität des Saarlandes

Die sprachorientierte KI-Forschung hat als Teilgebiet der Informatik zwei Hauptaufgaben:

- (AI) Die komplexen Informationsverarbeitungsprozesse, die dem Verstehen, der Produktion und dem Erwerb natürlicher Sprache zugrundeliegen, sollen mit informatischen Mitteln exakt beschrieben und erklärt werden.
- (A2) An intelligentes Sprachverhalten gebundene Leistungen sollen maschinell verfügbar gemacht werden, und die Mensch-Maschine-Kommunikation soll durch die Entwicklung natürlichsprachlicher Systeme verbessert werden.

In der sprachorientierten KI-Forschung wird das in (AI) formulierte Erkenntnisinteresse wesentlich durch die mit (A2) angesprochene ingenieurwissenschaftliche Zielsetzung der Konstruktion von Anwendungssystemen bestimmt.

Das Paradigma der sprachorientierten KI-Forschung kann durch die Kombination von vier Leitlinien charakterisiert werden (vgl. Wahlster 1982):

- (L1) Angestrebt wird eine vollständig operationalisierte, extreme Form der Referenzsemantik, die bis auf die sensorische Ebene 'durchgeschaltet' wird. Stets wird mit einer expliziten, maschinell handhabbaren und vollständigen Repräsentation der Diskurswelt gearbeitet.
- (L2) Angestrebt wird eine ganzheitliche Modellierung der Interaktion zwischen den dem Sprachverhalten, der Sprachproduktion und dem Spracherwerb zugrundeliegenden Prozessen und anderen für die Sprachverwendung relevanten kognitiven und sozialen Prozessen.
- (L3) Angestrebt wird eine prozeßorientierte Rekonstruktion von sprachlichem Verhalten, wobei Sprachgebrauch prozedural als Ergebnis kommunikativer und kognitiver Prozesse beschrieben wird.
- (L4) Angestrebt wird die Modellierung des instrumentellen Charakters von Sprache, da in KI-Systemen Sprache als Werkzeug in Arbeitsprozessen dienen soll.

Damit unterscheidet sich die sprachorientierte KI als Teilgebiet der Informatik in ihrer Zielsetzung und Vorgehensweise klar von der Linguistik. Ein ernsthafter KI-Forscher wird daher ebensowenig wie ein Sprachphilosoph oder Sprachpsychologe behaupten, daß er Linguistik betreibt (vgl. auch Thompson 1983).

Genauso wie Deduktionssysteme ohne Mathematik und bildverstehende Systeme ohne Physik nicht erfolgreich zu entwickeln sind, führt die Entwicklung natürlichsprachlicher Systeme ohne die Auswertung linguistischer Resultate in Sackgassen. Es ist daher in Projekten zur sprachorientierten KI stets eine interdisziplinäre Zusammenarbeit von Informatikern und Linguisten anzustreben, wie sie z. B. im neugegründeten Center for the Study of Language and Information in großem Stil praktiziert wird. Allerdings setzt eine solche Kooperation auf breiter Basis voraus, daß die unterschiedlichen Erkenntnisinteressen, Methodenrepertoires und Erwartungshaltungen gegenüber dem jeweils anderen Fachgebiet klar artikuliert und gegenseitig anerkannt werden, ohne diese auf verschwommene Grenzfälle zu reduzieren.

In Fig. 1 wird versucht, die sechs wichtigsten Relationen zwischen Linguistik und sprachorientierter KI durch Präpositionen zu charakterisieren und anhand eines typischen Beispiels zu illustrieren. Während in der Vergangenheit einige Ansätze der sprachorientierten KI-Forschung gezielt als Gegenmodelle zu gängigen linguistischen Vorstellungen entworfen wurden (KI contra Linguistik) oder linguistische Forschungsergebnisse ignorierten (KI ohne Linguistik), wird heute die Notwendigkeit eines wechselseitigen Wissenstransfers allgemein anerkannt.

Bisher wurden erst wenige in der Kern-Linguistik anerkannte Resultate mithilfe von KI-Methoden erzielt (Linguistik durch KI) und kaum KI-Softwarewerkzeuge für den Linguisten entwickelt (KI für Linguistik). Dagegen hat sich der Einfluß neuerer linguistischer Modelle auf die

sprachorientierte KI-Forschung (KI mit Linguistik) in letzter Zeit verstärkt. Beim gegenwärtigen Stand der KI-Forschung muß eine systematische Aufbereitung linguistischen Wissens für die Entwicklung natürlichsprachlicher Systeme (Linguistik für KI) mit höchster Priorität verfolgt werden.

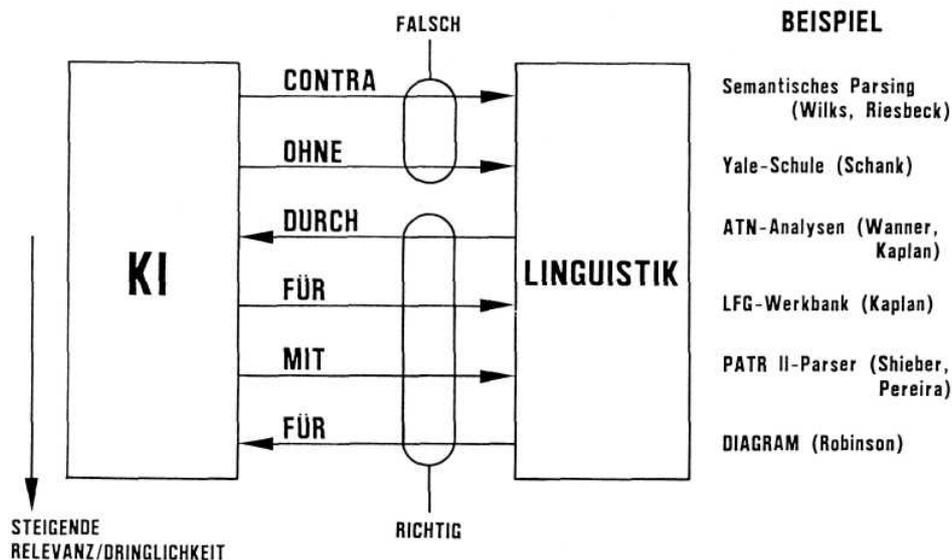


Fig. 1: Relationen zwischen Linguistik und KI

Natürlichsprachliche Systeme gehören zusammen mit den Expertensystemen zur Klasse der wissensbasierten KI-Systeme. Neben sprachunabhängigen Wissensquellen (u.a. Repräsentation der Diskurswelt) umfassen natürlichsprachliche Systeme in der Regel mehrere sprachbezogene Wissensquellen (z.B. Lexikon, Grammatik, Dialogregeln).

Es ist daher verständlich, daß Informatiker - genauso wie sie beim Aufbau der Wissensbasis für ein juristisches Expertensystem Rechtswissenschaftler oder bei der Konstruktion eines Expertensystems für den Finanzbereich Wirtschaftswissenschaftler als Experten zu Rate ziehen - beim Aufbau sprachbezogener Wissensquellen für natürlichsprachliche Systeme von Sprachwissenschaftlern Hilfe erwarten. Hier befragt der Informatiker als KI-Fachmann den Linguisten als 'Experten für Sprache'.

Obwohl größere Teile der Linguistik eher an wichtigen sprachtheoretischen und metatheoretischen Fragestellungen arbeiten und Modelle vorschlagen, die aufgrund ihres geringen Operationalisierungsgrades oder ihrer extremen Speicherplatz- und Rechenzeitkomplexität einer Implementierung nicht zugänglich sind, gibt es Sprachwissenschaftler, die bereit und durch solide, sprachtheoretisch abgesicherte einzelsprachliche Untersuchungen auf der Basis systematischer Feldarbeit auch dazu in der Lage sind, für die sprachorientierte KI relevante Ergebnisse zu liefern. So formuliert v. Hahn als eines der vordringlichen Ziele der Fachsprachenforschung 'kommunikativ orientierte Analysen von fachlich gebundenen Sprachzusammenhängen zu liefern, die die Sprachwirklichkeit in typischen modernen Handlungszusammenhängen nicht nur der Industrie und Technik erheben' (vgl. v. Hahn 1983, S. 11).

Falls die Vermittlung grundlegender Methoden und Techniken der KI künftig vermehrt in das Ausbildungsprogramm für Linguisten übernommen wird, sehe ich die Möglichkeit, den bisher noch von Informatikern mit KI-Spezialkenntnissen vorgenommenen Wissenstransfer zwischen 'Sprachexperten' und der Wissensbasis von KI-Systemen durch die direkte Interaktion des Linguisten mit geeigneten Wissenserwerbskomponenten von KI-Systemen zu ersetzen und somit aufgrund der wachsenden wirtschaftlichen Bedeutung natürlichsprachlicher Systeme auch mittelfristig gesicherte, neue Beschäftigungsmöglichkeiten für Linguisten zu schaffen.

Das sogenannte Knowledge Engineering für die sprachbezogenen Wissensquellen von KI-Systemen kann dann eine wichtige Teilaufgabe der Angewandten Linguistik werden. Die Domäne des Linguisten ist in diesem Zusammenhang der Aufbau von Wissensquellen mithilfe vorgegebener Wissensrepräsentationssprachen (z.B. Erstellung eines größeren Verblexikons im KL-ONE-Formalismus, Erstellung einer Grammatik für Nominalphrasen im FUG-Formalismus).

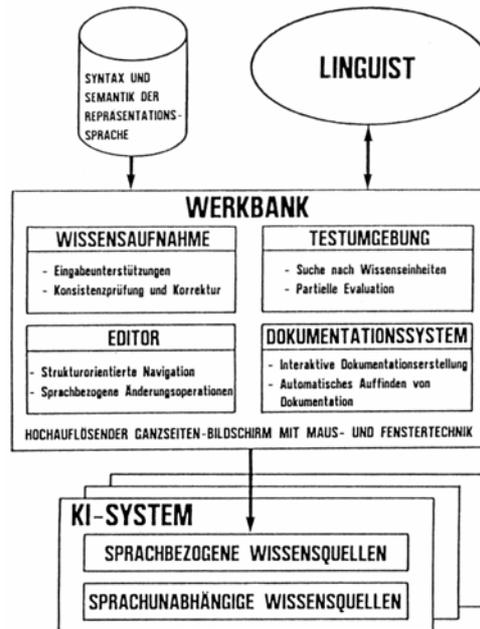


Fig. 2: Das Konzept einer Werkbank für Linguisten

Zur Domäne des Informatikers gehört dagegen der Entwurf und die Weiterentwicklung von Wissensrepräsentationssprachen (z.B. Entwurf von Repräsentationskonstruktionen für die Semantik von Reflexivpronomen), sowie die Bereitstellung intelligenter Werkzeuge für am Aufbau von Wissensbasen arbeitende Linguisten (z.B. die LFG-Werkbank von Kaplan, die auf einer LISP-Maschine von XEROX realisiert ist). Das in Fig. 2 skizzierte Konzept einer Werkbank für Linguisten, das mit modernen Symbolverarbeitungssystemen effizient realisiert werden kann, bietet dem in KI-Projekten mitwirkenden Sprachwissenschaftler eine Arbeitsumgebung, in der er zunächst problemorientiert eine der vorgesehenen Wissensrepräsentationssprachen auswählen kann, um dann linguistisches Wissen systemgestützt zu erfassen, in einer anspruchsvollen Testumgebung experimentell zu erproben, ggf. zu korrigieren und schließlich zu dokumentieren. Beim gegenwärtigen Entwicklungsstand kann es kaum das Ziel eines solchen Vorgehens sein, jeweils in Bezug auf eine bestimmte Gesamtsprache vollständige Wissensquellen zu erstellen, sondern es können zunächst lediglich die Wissensbereiche abgedeckt werden, die notwendig sind, um die durch die jeweiligen Forschungsziele oder Anwendungsbereiche festgelegten Sprachfragmente oder Subsprachen zu erfassen.

Literatur:

- v. Hahn, Walther (1983): Fachkommunikation: Entwicklung, linguistische Konzepte, betriebliche Beispiele. Berlin, New York : de Gruyter.
- Thompson, Henry (1983): Natural language processing: a critical analysis of the structure of the field, with some implications for parsing. In: Sparck Jones, K., Wilks, Y. (eds.): Automatic natural language parsing. Chichester: Ellis Horwood, S. 11-31.
- Wahlster, Wolfgang (1982): Aufgaben, Standards und Perspektiven sprachorientierter KI-Forschung: Einige Überlegungen aus informatischer Sicht. In: Batori, I., Krause, J., Lutz, H.D. (eds.): Linguistische Datenverarbeitung : Versuch einer Standortbestimmung im Umfeld von Informationslinguistik und Künstlicher Intelligenz. Tübingen : Niemeyer, S. 13-24.