

Informatik-Fachberichte 195

Subreihe Künstliche Intelligenz

Herausgegeben von W. Brauer in Zusammenarbeit mit dem
Fachausschuß 1.2 „Künstliche Intelligenz und
Mustererkennung“ der Gesellschaft für Informatik (GI)

I. S. Bátori U. Hahn
M. Pinkal W. Wahlster (Hrsg.)

Computerlinguistik und ihre theoretischen Grundlagen

Symposium, Saarbrücken, 9.-11. März 1988

Proceedings



Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo

Herausgeber

István S. Bátori
Studiengang für Angewandte Informatik, EWH Koblenz
Rheinau 3-4, D-5400 Koblenz

Udo Hahn
Fakultät für Mathematik und Informatik, Universität Passau
Postfach 2540, D-8390 Passau

Manfred Pinkal
Fachbereich Informatik, Universität Hamburg
Bodenstedtstraße 16, D-2000 Hamburg 50

Wolfgang Wahlster
Fachbereich Informatik, Universität des Saarlandes
D-6600 Saarbrücken 11

Veranstalter des Symposiums

Gesellschaft für Linguistische Datenverarbeitung (GLDV)
Deutsche Gesellschaft für Sprachwissenschaft (DGfS) -
Sektion Computerlinguistik
Gesellschaft für Informatik (GI) -
Fachgruppe Natürlichsprachliche Systeme

CR Subject Classifications (1987): I.2.7

ISBN-13: 978-3-540-50554-9 e-ISBN-13: 978-3-642-74282-8
DOI: 10.1007/978-3-642-74282-8

Dieses Werk ist urheberrechtlich geschützt. Die dadurch begründeten Rechte, insbesondere die der Übersetzung, des Nachdrucks, des Vortrags, der Entnahme von Abbildungen und Tabellen, der Funksendung, der Mikroverfilmung oder der Vervielfältigung auf anderen Wegen und der Speicherung in Datenverarbeitungsanlagen, bleiben, auch bei nur auszugsweiser Verwertung, vorbehalten. Eine Vervielfältigung dieses Werkes oder von Teilen dieses Werkes ist auch im Einzelfall nur in den Grenzen der gesetzlichen Bestimmungen des Urheberrechtsgesetzes der Bundesrepublik Deutschland vom 9. September 1965 in der Fassung vom 24. Juni 1985 zulässig. Sie ist grundsätzlich vergütungspflichtig. Zuwiderhandlungen unterliegen den Strafbestimmungen des Urheberrechtsgesetzes.

© by Springer-Verlag Berlin Heidelberg 1988

Geleitwort

Die Frage nach dem Selbstverständnis eines Faches stellt sich am nachdrücklichsten, wenn das Fach und seine Definition sich verändert. Dann reichen die Gewißheiten von gestern nicht mehr aus, um in der Realität von heute erfolgreich zu agieren. Daß die Computerlinguistik sich zur Zeit in einer Expansionsbewegung befindet, kann niemand entgehen, der die akademische Landschaft beobachtet. Im Spiel von Stellenzuwächsen und -streichungen, von neu eingerichteten und eingestellten Studiengängen steht die Computerlinguistik auf der Seite der Gewinner. Sie gewinnt an Volumen und allgemeinem Interesse. Damit steigen jedoch die Anforderungen an das Selbstverständnis des Faches und seine Erklärungsfähigkeit nach außen. Was früher ausreichte, um die Computerlinguistik zu charakterisieren, hält den Ansprüchen nicht mehr stand, die an die Fundierung eines in der Bedeutung gewachsenen Faches gestellt werden. Der vergrößerte Gesamtumfang ermöglicht eine stärkere Binnendifferenzierung. Da fachintern mehr und unterschiedlichere Auffassungen von Computerlinguistik vertreten werden, wird es umso wichtiger, sich klarzumachen, was zwischen den divergierenden Positionen vermittelt. Diese Rolle fällt nicht ausschließlich, aber doch vorrangig einem gemeinsamen theoretischen Grundverständnis zu. Es ist darum in der aktuellen Situation der Computerlinguistik angebracht, sich mit deren theoretischen Grundlagen zu beschäftigen.

Die Frage nach der Rolle der Theorie in der Computerlinguistik läßt sich in die Form einer wissenschaftlichen Veranstaltung kleiden, die die aktuelle Diskussion zu diesem Zeitpunkt darstellen und vorantreiben soll. Dies ist mit dem Symposium "Computerlinguistik und ihre theoretischen Grundlagen" geschehen, das die Gesellschaft für Linguistische Datenverarbeitung (GLDV), die Fachgruppe "Natürlichsprachliche Systeme" in der Gesellschaft für Informatik (GI) und die Sektion Computerlinguistik der Deutschen Gesellschaft für Sprachwissenschaft (DGfS) zusammen veranstaltet haben.

Die Veranstalter hoffen, mit dem Symposium "Computerlinguistik und ihre theoretischen Grundlagen" zur Weiterentwicklung des Theorieverständnisses in der Computerlinguistik beigetragen zu haben.

Brigitte Endres-Niggemeyer
Vorsitzende der Gesellschaft für Linguistische Datenverarbeitung

Vorwort

Der vorliegende Band enthält revidierte Versionen der Beiträge zum Symposium "Computerlinguistik und ihre theoretischen Grundlagen", das in Saarbrücken vom 9. bis 11.3.1988 von der GLDV, DGfS - Sektion Computerlinguistik und von der GI - Fachgruppe Natürlichsprachliche Systeme veranstaltet worden ist. Aufgenommen worden sind die auf der Tagung gehaltenen Vorträge sowie die Positionspapiere der Podiumsdiskussion.

Computerlinguistik (CL) ist ein Fach mit starkem praktischen Bezug, dessen Standort zwischen Linguistik, Informatik und Informationswissenschaft zu suchen ist. Die praktische Arbeit mit natürlichsprachlichen Systemen hat jedoch immer wieder gezeigt, daß Fortschritte in den Anwendungen nur dann erzielt werden können, wenn die Grundlagen ausreichend geklärt worden sind.

Die Einsicht in die Notwendigkeit der linguistischen Grundlagenforschung für Computeranwendungen setzte sich nur langsam durch. Sie führte zu einer Neugliederung der Wissenschaften, mit neuen Nachbarschaften und neuen interdisziplinären Forschungsbereichen, wie Kognitionswissenschaft, Künstliche Intelligenz (und darunter Sprachorientierte Künstliche Intelligenz) und Computerlinguistik.

Das Symposium in Saarbrücken ist in diesem Zusammenhang zu sehen. Die Tagung ging auf einen Vorschlag der Gesellschaft für Linguistische Datenverarbeitung (GLDV) an die KI-Gruppe der GI und an die DGfS zurück, zusammen eine Tagung zu organisieren, die speziell der Thematik der theoretischen Aspekte der CL gewidmet werden sollte.

Als Mitglieder des Programmkomitees für das Symposium sind von den drei Gesellschaften die folgenden Wissenschaftler benannt worden:

- István Bátori, Koblenz,
- Franz Günthner, Tübingen,
- Christopher Habel, Hamburg,
- Siegfried Kanngießer, Osnabrück,
- James Kilbury, Düsseldorf,
- Manfred Pinkal, Hamburg,
- Udo Hahn, Passau,
- Wolfgang Wahlster, Saarbrücken.

Aufgabe des Komitees war die Erstellung des Tagungsprogramms und die Auswahl der geeigneten Beiträge. Es hatte nicht die Aufgabe, die CL und deren theoretische Grundlagen selbst zu definieren.

Natürlich war es unvermeidbar, gewisse Grundprinzipien für die Auswahl der Beiträge festzulegen, die hier aufgelistet werden sollen:

Formale Beschreibung: CL setzt eine beschreibungstechnische Präzision voraus, die ohne formale Mittel nicht zu meistern ist.

Explizite Bezugnahme auf natürlichsprachliche Formulierungen: Mathematische oder formale Modelle müssen an sprachlichem Material konkretisiert werden.

Algorithmische Umsetzbarkeit: Untersuchungen zur theoretischen Linguistik ohne algorithmisch-prozedurale Perspektive konnten nicht berücksichtigt werden.

Mensch-Computer-Interaktion: CL geht in ihrer Zielsetzung tendenziell über die reine sprachliche Strukturbeschreibung hinaus, sie setzt sich mit der Erfassung der sprachlichen Prozesse (Produktion und Verstehen) und mit der Integration der verschiedenen Ebenen der Sprachbeschreibung auseinander.

Fachliches Niveau und Relevanz für das Tagungsthema waren selbstverständliche Voraussetzungen. Die Vorträge sollten einen interessanten (wesentlichen, aktuellen) Aspekt der theoretischen Grundlagen der CL ansprechen und zur Klärung oder Präzisierung dieser Grundlagen beitragen.

Wenn man berücksichtigt, daß in dem Programmkomitee drei Organisationen (GLDV, DGfS und die KI-Gruppe der GI) repräsentiert waren, ist es nicht verwunderlich, daß die theoretischen Grundlagen der CL aus der Perspektive der KI, der (theoretischen) Linguistik, der (linguistischen) Informationswissenschaft und der linguistischen Anwendungen unterschiedlich akzentuiert werden können.

Auch wenn es unterschiedliche Auffassungen über das Wesen der CL innerhalb des Programmkomitees gegeben hat, ist die Auswahl der Beiträge vom Programmkomitee einmütig gefaßt worden.

26 Beiträge wurden eingereicht; 10 konnten aufgrund der oben genannten Kriterien akzeptiert werden.

Die Organisation des Symposiums ist von dem "Institut der Gesellschaft zur Förderung der Angewandten Informationsforschung e.V. (IAI) - Projekt EUROTRA-D" getragen worden. Im Namen des Programmkomitees danken die Herausgeber an dieser Stelle dem IAI und allen seinen Mitarbeitern, die die Tagung nicht nur sorgsam vorbereitet, sondern die Trägervereine auch finanziell entlastet haben. Besonderer Dank gilt Herrn Dr. Johann Haller und Herrn Tom C. Gerhardt, die mit ihrem persönlichen Einsatz den mustergültigen Ablauf des Symposiums ermöglichten.

István S. Bátori

Inhaltsverzeichnis

INPUT WORT. Ansätze der Simulation wortweiser Textverarbeitung Manfred Aulich, Guido Drexel, Gert Rickheit, Hans Strohner	1
An Algebraic Characterization of STUF Christoph Beierle, Udo Pletat, Hans Uszkoreit	15
Ein Konzept zur Komposition der Semantik aus bedeutungstragenden Teilen einer Äußerung und zur Behandlung alternativer Interpretationen Ayşe Erben	33
<i>In diesem Ton lasse ich nicht mit mir reden!</i> – Einige psychologische Überlegungen zu Aufforderungsinteraktionen zwischen Mensch und Computer Joachim Grabowski-Gellert	54
Effiziente Analyse natürlicher Sprache mit TAGs Karin Harbusch	79
A Mathematical Model for the CAT Framework of EUROTRA Theo M.V. Janssen	104
Modellierungen in der Maschinellen Übersetzung Annely Rothkegel	117
PARTIKO. Kontextsensitive, wissensbasierte Schreibfehleranalyse und -korrektur Astrid Scheller	136
TAGDevEnv. Eine Werkbank für TAGs Klaus Schifferer	152
Zur Behandlung von unbounded dependencies im multilingualen Übersetzungssystem EUROTRA Paul Schmidt	172

Panel-Diskussion geleitet von Burghard Rieger

Vorbemerkungen zur Panel-Diskussion Theoretische Grundlagen der Computerlinguistik	192
Computerlinguistik: Eine Positionsbestimmung Burghard Rieger	193
Auspizien einer kognitiven Orientierung der Computerlinguistik Manfred Bierwisch	198
Kognitionswissenschaft als Grundlage der Computerlinguistik Christopher Habel	204
Computerlinguistik, Linguistik und künstliche Intelligenz Hans Uszkoreit	210
Zum Fortschritt in der Computerlinguistik Wolfgang Wahlster	215