

# Veranstaltungen

## VERANSTALTUNGEN

### Lexikon und Morphologie

2. GLDV-Herbstschule vom 19. bis 24. September 1994 an der Universität Leipzig, Institut für Informatik

Die 2. GLDV-Herbstschule 94 "Lexikon und Morphologie" wendet sich an Student/inn/en und Doktorand/inn/en, insbesondere der Computerlinguistik, Linguistik, Informatik und verwandter Disziplinen, aber auch an Interessierte aus der beruflichen Praxis.

Das Kursangebot beinhaltet Vorlesungen und praktische Übungen (zum Teil am Rechner) zu folgenden Themenbereichen:

Maschinenlesbare Lexika  
Automatische Wortformererkennung  
und morphologische Algorithmen  
MLexD: Ein Standard für die korpusbasierte Lexikographie  
Produkte der Sprachtechnologie  
Transfer-Lexika in der Maschinellen Übersetzung  
Softwareergonomie und natürliche Sprache

### Das Kursangebot

Von den sechs angebotenen Kursen kann jeder Teilnehmer bis zu drei belegen. Da die Veranstaltungen in den Blöcken A, B, C jeweils parallel stattfinden, kann je Block nur höchstens ein Kurs gewählt werden.

Angeboten werden die folgenden Alternativen:

Block A: "Maschinenlesbare Lexika" *oder* "MLexD: Ein Standard für die korpusbasierte Lexikographie"

Block B: "Automatische Wortformererkennung und morphologische Algorithmen" *oder* "Softwareergonomie und natürliche Sprache"

Block C: "Transfer-Lexika in der Maschinellen Übersetzung" *oder* "Produkte der Sprachtechnologie"

### Das Rahmenprogramm

Bei der Eröffnungsveranstaltung am Montagvormittag referiert Prof. Lenders (Bonn) über „20 Jahre GLDV-Absichten, Aufgaben, Positionen" und Frau Dr. Walther (Leipzig) über "Zur Stellung der Morphologie in der generativen Grammatik".

Den ersten Abendvortrag am Montagabend hält Prof. Schnelle (Bochum) über "Wortvernetzungen in Computer und Gehirn", den zweiten am Donnerstagabend hält Prof. Bierwisch (Berlin) über "Universalien und Idiosynkrasien im Lexikon".

Am Dienstagabend findet eine Präsentation der Sieger der ersten Morpholympics 1994 statt.

### Kurse

Maschinenlesbare Lexika (mit einer Einführung in SGML [*Winfried Lenders / M. Volk*])

Der Kurs bietet in seinem ersten Teil eine Einführung in die Struktur, Verwendungsformen und Konstruktionsmethoden maschinenlesbarer und maschinenverwendbarer Wörterbücher (*machine rea-*

*dable* und *machine tractable dictionaries*). Es wird auf folgende Themen auch anhand von Beispielen eingegangen:

- . Wörterbücher, Lexika und lexikalisches Wissen (Definitionen)
- . Gebrauchswörterbücher  
Computerwörterbücher
- . Lexikalische Mikrostrukturen und Makrostrukturen
  - . Wörterbücher in linguistischen Modellen
  - . Wiederverwertbarkeit lexikalischer Ressourcen
- . Das Wörterbuch als Netzwerk
- . Lexikalische Strukturen und Relationen
- . Tools zum Wörterbuchaufbau
- . Lexikalische Datenbanken

Im zweiten Teil des Kurses wird, ausgehend vom Beispiel der Standardisierung von Wörterbuchkodierungen, eine Einführung in die heute international allgemein anerkannte Standardisierungssprache SGML (Standard Generalized Markup Language) gegeben. Es wird gezeigt, wie SGML als Grammatik zur Strukturierung von Dokumenten gesehen werden kann.

MLexD: Ein Standard für die korpusbasierte Lexikographie  
(*Wolf Papprotti*)

Der Kurs gibt einen Überblick über die Architektur ein- und zweisprachiger elektronischer Lexika sowie die morphosyntaktische Information, die ein Lexikon abbilden sollte. Die Adäquatheit des Standards hinsichtlich der Sprachen Englisch, Deutsch und Neugriechisch wird diskutiert. Es ist beabsichtigt, Methoden und Werkzeuge für die Extraktion lexikaler Information aus Korpora vorzustellen.

### **Automatische Wortform-Erkennung und morphologische Algorithmen**

(*Roland Hausser*)

Die linguistische Analyse der Wortformen und Wörter fällt in den Aufgaben-

bereich der Grammatikkomponente 'Morphologie'. Die traditionell sprachwissenschaftlichen Konzepte der morphologischen Analyse sind die theoretische Grundlage für eine effiziente und linguistisch wohlmotivierte Programmierung der automatischen Wortformanalyse im Rahmen der Computerlinguistik. Die Wortformererkennung wiederum ist die Voraussetzung für praktisch alle weiteren computerlinguistischen Anwendungen wie syntaktisches Parsen, semantische Analyse, Indexierung etc.

Der Kurs gliedert sich in einen theoretischen und einen praktischen Teil. Der theoretische Teil vermittelt die Grundbegriffe der morphologischen Analyse, beschreibt unterschiedliche Methoden der automatischen Wortformererkennung und vergleicht ihre programmiertechnische Eignung. Der praktische Teil vermittelt den Umgang mit dem Softwaresystem LA-MORPH, das bei allgemeinen Texten, z.B. Zeitungsartikeln oder Bundestagsreden, etwa 95% der Wortformen erkennt und auf einem PC 486 / 33 unter Linux eine durchschnittliche Geschwindigkeit von etwa 400 Wortformen pro Sekunde entwickelt.

### **Softwareergonomie und natürliche Sprache**

(*Jürgen Krause*)

Bei der Mensch-Maschine Interaktion ist die natürliche Sprache nur eine der beiden Hauptmodi, mit dem Computer "natürlich" zu interagieren. Eine zweite - und in den letzten Jahren wesentlich erfolgreichere Alternative sind die grafischen Benutzeroberflächen. Deshalb befaßt sich der Kurs mit beiden Varianten unter dem gemeinsamen Blickwinkel einer softwareergonomischen Gestaltung von Benutzeroberflächen.

Themen sind:

- a) Was ist Ergonomie und was Softwareergonomie ?
- b) Theoretische Grundlagen "natürlicher" Oberflächen von Computersystemen
- c) Grafische Oberflächen auf der Basis von Direktmanipulation und Schreib-tischmetapher

- d) Natürlichsprachliche Frage- Antwort-Systeme: Grundlagen und kommerzielle Situation heute
- e) Computertalk
- f) Multimodale Mischformen als Überlebensstrategie für den natürlichsprachlichen Modus bei der Mensch-Maschine-Interaktion (WOB-Modell)

Der Kurs hat nichts mit dem Hauptthema "Lexikon und Morphologie" zu tun. Er ist als fakultative, inhaltliche Erweiterung gedacht, für alle diejenigen, die etwas über den engen Zaun einer an der formalen Linguistik orientierten Computerlinguistik hinausschauen wollen.

**Transferlexika in der Maschinellen Übersetzung**

[Johann Haller]

- . Grundbegriffe der Maschinellen Übersetzung: Analyse, Transfer, Synthese
- . Transferkonzepte: direkt, indirekt, Interlingua
- . Transfer in älteren Systemen: SYSTRAN, LOGOS, METAL
- . Transfer in Forschungssystemen (EU ROTRA, CAT2)
- . lexikalischer und struktureller Transfer
- . Beispiele von Einträgen in Transferlexika

Die Teilnehmer erarbeiten und erproben selbständig neue Transfereinträge im CAT2-System.

**Produkte der Sprachtechnologie**

[Gerhard Heyer / Klemens Waldhör]

Im Unterschied zur Computerlinguistik steht in der Sprachprodukttechnologie weniger die Simulation kognitiver Prozesse im Vordergrund, als vielmehr das Ziel, Verfahren und Methoden für eine ingenieurmäßige Entwicklung von Sprachprodukten bereitzustellen. Der Kurs gibt einen Überblick über die wichtigsten Grundlagen, Probleme und Methoden der Sprachprodukttechnologie. Im einzelnen werden behandelt:

1. Das ingenieurwissenschaftliche Paradigma als Grundlage der Sprachprodukttechnologie
2. Überblick und Klassifikation von Sprachprodukten
3. Softwareengineering für Sprachprodukte (insb. elektronische Wörterbücher)
4. Ergonomische Aspekte von Sprachprodukten (insb. elektronische Wörterbücher)
5. Linguistik für Sprachprodukte

Die Herbstschule findet in den Räumen der Universität Leipzig statt.

Hinweise zur Anmeldung:

Die Gebühren für Kursteilnahme und -unterlagen betragen

für Student/inn/en (mit Immatrikulationsbescheinigung): bis zum 1.7.94. 65,- DM nach dem 1.7.94 90,- DM für sonstige Teilnehmer/innen:

bis zum 1.7.94 130,- DM nach dem 1.7.94 180,- DM

Auskunft und Aumeldungsunterlagen über:

2. GLDV-Herbstschule 1994  
 c/o Prof. Dr. G. Heyer  
 Universität Leipzig  
 Institut für Informatik / Abt. ASV  
 Augustusplatz 10/11  
 04109 Leipzig  
 Tel.: 0341 - 7192188 Fax: 0341 - 7192399 e-mail: heyer@informatik.uni-leipzig.de



**Europäische IR Sommerschule 03.-08  
September 1995, Ilmenau/Thüringen**

Lokale Organisation: Vorsitz      R. Schramm, TH Ilmenau  
Programmkomitee: Initiative:    Norbert Fuhr, Dortmund  
Fachgruppe Information Retrieval der GI  
Vorläufiges Programm ab Nov. 94

**Hypertext-Information Retrieval-Multimedia '95, HIM '95**

05.-07. April, Konstanz

Anforderungen an moderne Informationssysteme lassen aus System- und Nutzersicht eine isolierte Betrachtung von Problemen häufig nicht mehr zu. Die Fachgruppen Hypertext (4.9.1), Information Retrieval (2.5.4 / 4.9.3) und Multimedia (4.9.2) der Gesellschaft für Informatik (GI) führen daher gemeinsam mit der Österreichischen Computer Gesellschaft (ÖCG), der Schweizer Informatik Gesellschaft (SI) und dem Hochschulverband Informationswissenschaft (HI) eine Tagung mit den Schwerpunkten Hypertext, Information Retrieval und Multimedia durch. Die Informationswissenschaft an der Universität Konstanz ist vom 05.04.95 bis 07.04.95 Gastgeber der Tagung. Sowohl fachspezifische als auch fachübergreifende Beiträge in der Form von Vorträgen, Postern und Tutorien sind für alle Teilgebiete und Aspekte von Hypertext, Information Retrieval und Multimedia willkommen. Zu Systemdemonstrationen während der und ergänzend zu den Vorträgen wird ausdrücklich aufgefordert.

Beiträge sind bis spätestens 01.11.94 einzureichen. Tagungssprachen sind Deutsch und Englisch.

Für weitere Informationen oder eventuelle Fragen, wenden sie sich bitte an die Organisationsleitung:

Marc Rittberger

Universität Konstanz

Informationswissenschaft

Postfach 5560

D- 78434 Konstanz

Tel: 49-7531-883595

Fax: 49-7531-882601

EMAIL: ritt@inf-wiss.uni-konstanz.de