

NEWS – NACHRICHTEN

Computer Constructs Natural Language Concept System

At the University of Mainz, FRG, a DFG-sponsored research project was started on July 1, 1974, aiming at a computerized display of the concept system behind the ordinary language terms of the German language. Under the guidance of Professor Dr. Gerhard Wahrig, author of the well-known German "Webster", his "Deutsches Wörterbuch", (Bertelsmann 1966, ² 1975 in preparation), a small research team began to input the definitions and explanations of words listed in the "Deutsches Wörterbuch" in such a way that the segments of these definitions may be computer-compared with each other. Polysemous terms as parts of definitia will be checked and – by a stochastic process – be reduced to monosemy by a computer-program developed by Gudrun Guckler: "Semantische Analyse zweisprachiger Wörterbucheinträge mittels EDV" (1974 in prep.). The results of these comparisons will show on the one hand a network formed by the generic, associative, partitive relations between the concepts of the terms and on the other hand it will identify all equivalent terms/concepts. Since one definition is building up on the other, it may finally be shown, which of the terms/concepts will be the highest ones in the hierarchy of predications and thus prove to act as the elementary terms on which language is built. The project is supported by the Deutsche Forschungsgemeinschaft (DFG) (German Research Association) and is estimated to run for about five years. Some 50 000 terms and around 100 000 concepts will have to be investigated. The name of the project: "Semstruktur" (seme-structure).

Ein automatisches Kern-Lexikon der deutschen Sprache

Unter der Projektleitung von Prof. Dr. Janos S. Petöfi, wurde am 1.10.1974 an der Universität Bielefeld mit einem von der Deutschen Forschungsgemeinschaft (DFG) geförderten Projekt begonnen, dessen Aufgabenstellung auf den folgenden Hypothesen basiert:

- 1) Im lexikalischen Bestand jeder natürlichen Sprache kann eine Menge abgegrenzt werden, deren Elemente als nicht zu definierende Grundelemente betrachtet werden können und mit deren Hilfe alle anderen Elemente der Sprache definierbar sind. Eine solche Menge von Grundelementen kann daher ein Definitionswörterbuch und die Menge der Repräsentationen der Grundelemente ein Kern-Lexikon (KL) genannt werden.
- 2) Die optimale Struktur der Repräsentation der Grundelemente ist eine „Funktorkomplex + Argument(e)“-Struktur, wobei die Argumentstellen durch „Funktionskategorien und Sortenspezifikationen“ gekennzeichnet sind.
- 3) Die Menge der Elemente des KL kann in Teilmengen unterteilt werden, wobei die „Labels“ der einzelnen Teilmengen an den Argumentstellen der Elemente des KL als Sortenspezifikationen (Selektions-Charakteristika

hinsichtlich der Besetzung der Argumentstellen) funktionieren.

4) Den Elementen des KL können „Explikationen“ (quasi-Definitionen), die aus anderen Elementen des KL aufgebaut sind, zugeordnet werden. Diese Explikationen de-sambiguieren die Lesarten der Grundelemente.

Dementsprechend ergeben sich die folgenden Aufgaben:

- a) Abgrenzung einer „Menge von Grundelementen“ im lexikalischen Bestand der deutschen Sprache;
- b) Normalisierung der Struktur der Repräsentation der Grundelemente;
- c) Bestimmung der Sortenspezifikationen;
- d) Formulierung der Explikationen.

Folgende Resultate sind aus der Durchführung der oben aufgeführten Aufgaben zu erwarten:

Das herzustellende deutsche Kern-Lexikon (DKL) wird in sich selbst einen wesentlichen Beitrag zur Lexikon-Forschung bedeuten, sowohl im Hinblick auf die Beschreibung der deutschen Sprache als auch in allgemeiner Hinsicht.

Hat man ein DKL zur Verfügung, so kann mit dem Aufbau eines Lexikons (von wenigstens minimalem Umfang) begonnen werden, das auch ein Optimum an Informationen enzyklopädischer Art enthält. Nennen wir dieses Lexikon ein deutsches Minimal-Lexikon (DML).

Hat man ein DML zur Verfügung, so kann mit der Entwicklung einer tatsächlich funktionierenden (und optimal automatisierbaren) Textgrammatik begonnen werden.

Eine funktionierende Textgrammatik ermöglicht die Verwirklichung von solchen Zielsetzungen mit grundlegender Bedeutung wie z. B.:

- Erforschung der Wirksamkeit einer Textgrammatik in der Sprachbeschreibung;
- Erforschung der Wirksamkeit einer Textgrammatik in der nicht-automatisierten und automatisierten Textverarbeitung;
- Erforschung der Wirksamkeit einer Textgrammatik in den verschiedenen, sich mit Texten beschäftigenden Wissenschaftszweigen.

Solange keine funktionierende Textgrammatik existiert, kann die Forschung, die auf die Lösung dieser grundlegend wichtigen Aufgaben ausgerichtet ist, nicht über die ausschließlich theoretische Argumentationsebene hinausgehen, und dies hindert die gesellschaftliche Verwertung der theoretischen Forschung. Zur Entwicklung einer funktionierenden Textgrammatik bedarf es aber zunächst – neben den theoretischen, grammatischen Forschungen – des Aufbaus eines Kern-Lexikons.

(Zur Frage der Lexikon-Problematik siehe den Artikel „Some Aspects of a Multi-purpose Thesaurus“ in diesem Heft.)

First Infoterm Symposium, April 1975

As a first step toward co-ordination in the field of terminology a general view on terminological activities seems to be necessary. Invitations to a "Symposium on International Co-operation in Terminology" have been sent out to be held in Vienna, 9–11 April 1975. Papers

dealing with topics from the following list are invited from organizations and experts active in the field concerned:

1. International co-operation in the field of terminology, especially of terminological documentation (accomplishments, possibilities) in:
 - UDC 002 Documentation and Information
 - UDC 3 Social sciences
 - UDC 53 Physics
 - UDC 54 Chemistry and UDC 66 Chemical industries
 - UDC 56/59 Biology
 - UDC 61 Medical Sciences
 - UDC 62/69 Technology, esp. UDC 621.3 Electrical engineering, electronics
 - UDC 80 Linguistics and philology
 - UDC 91 Geography
 2. International network for terminological documentation (Infoterm network):
 - Decentralization according to subject fields or countries. Reports and discussions on the establishment of such a network
 3. Terminological word banks
 4. Central registration of neologisms
- Summaries of papers should be sent to Infoterm (Secretary: Mr. Helmut Felber, Österreichisches Normungsinstitut, Leopoldsgasse 4, A 1021 Wien 2) as soon as possible.

Conference on Printed Subject Indexes, July 1975

It is proposed to hold an *International Conference on Printed Subject Indexes – Their Design and Evaluation* at the College of Librarianship Wales, Aberystwyth, 14–18 July 1975. The emphasis of the Conference will be with characteristics and performance of indexes; their format, readability, indexing language and overall effectiveness, rather than with the mechanics of their production.

The themes of the various sessions are:

- International aspects of printed subject indexes, their place in the information network, discipline and interdisciplinary-oriented indexes, index surveys;
- Index entry design: term order, structure words, context, access points, rotation methods, index stages;
- Problems of language and terminology: index languages, crossreference structure;
- Linguistic and semantic aspects of subject indexes, information transfer theory;
- Typographical layout, design, readability, relevance predictability;
- Field and laboratory evaluation and testing of indexes, methodology, performance criteria and measures;
- Special indexes: citation, learned journal, back-of-book, concordances, directories.
- Indexes in microform.

There will be a number of invited papers from speakers of several countries, and also opportunity for participants to contribute short papers.

Direct inquiries and/or papers to M. Keen in Aberystwyth.

(from *ASIS SIG/CR Newsletter*, (1974) August)

The Classification Society, England

This year's Annual General Meeting – Scientific Meeting of The Classification Society took place at the Royal Society of Medicine on Oct. 1, 74 in London. The general topic: "Mathematical models in classification". Three papers were presented: 1) G. J. S. Ross, Department of Statistics, Rothamsted Experimental Station, Harpenden, Herts.: The relevance of statistical models to numerical classification. 2) M. O. Hill, Institute of Terrestrial Biology, Penrhos Road, Bangor, Caernarvon: The ecological niche and its relevance to the classification of plant communities. 3) Mrs. P. F. M. Smart, Department of Biology, Southampton University: Uses and abuses of different classificatory models in ecology.

(Addendum to Intern. Classificat. 1 (1974) No. 1, p. 45: here the address of the European Secretary of The Classification Society should have been added. It is Dr. A. J. Willmott, University of York, Department of Computation, Heslington, York, Y 01 5 DD, England)

UDC Editions in 23 Languages

The tables of the Universal Decimal Classification are now available – in some cases not yet for all classes – (or being prepared) as full, medium or abridged editions in 23 languages:

	F	M	A
Czech			
Dutch	x	–	(x)
English	–	–	x
Finnish	x	(x)	x
French	–	–	x
German	–	x	x
Hebrew	x	x	x
Hungarian	–	–	x
Italian	x	–	x
Japanese	x	x	x
Korean	x	(x)	x
Macedonian	–	–	x
Polish	–	(x)	x
Portuguese	x	x	x
Romanian	x	(x)	x
Russian	(x)	–	–
Serbo-Croat	x	x	–
Slovak	x	x	x
Slovene	–	(x)	–
Spanish	–	x	x
Swedish	x	(x)	x
(with Danish and Norwegian indexes)	–	–	x

For full details ask for the free 1974 Catalogue of FID Publications, available from the FID Secretariat.

(from *FID News Bulletin* 23 (1974) No. 7, p. 89)

Intern. Classificat. 1 (1974) No. 2 News – Nachrichten