The *Sixth European Meeting on Cybernetics and Systems Research* , organized by the Austrian Society for Cybernetic Studies, will be held April 13-16, 1982, at the University of Vienna, Austria. [See *AJCL 7:2*, pg. 123.] For a Preliminary Program contact the Chairman:

> Professor Robert Trappl
> Department of Medical Cybernetics
> University of Vienna
> Freyung 6/2
> A-1010 Vienna, AUSTRIA

A *Symposium on Artificial Intelligence*, sponsored by the Austrian Society for Artificial Intelligence and the Austrian Society for Cybernetic Studies, will be held at the Sixth European Meeting on Cybernetics and Systems Research, April 13-16, 1982, in Vienna. [See *AJCL 7:2*, pg. 123.] Extended abstracts must be submitted by December 1, 1981, to:

> Organizing Committee of the 6th EMCSR 82
> Austrian Society for Cybernetic Studies
> Schottengasse 3
> A-1010 Vienna, AUSTRIA

The *1982 National Computer Conference (NCC)*, sponsored by AFIPS, will be held June 7-10, 1982, in the Houston Astroarena. [See *AJCL 7:2*, pg. 122.] Papers must be submitted by October 31, 1981, to:

> Dr. Howard L. Morgan
> NCC'82 Technical Program Chairman
> Department of Decision Sciences
> The Wharton School/CC
> University of Pennsylvania
> Philadelphia, Pennsylvania 19104

The *Fifth European Conference on Electrotechnics* will take place in Copenhagen at The Technical University of Denmark, on June 14-18, 1982. [See *AJCL 7:2*, pg. 122.] Further details may be obtained from:

> DIEU, Danish Engineers' Post Graduate Inst.
> The Technical Univ. of Denmark, Bldg. 208
> DK-2800 Lyngby, DENMARK

The *Ninth International Conference on Computational Linguistics (COLING 82)* will be held July 5-10, 1982, in Prague, Czechoslovakia. It will be sponsored by the International Committee on Computational Linguistics, in association with the Linguistic Institute of L. Stûr, Slovak Academy of Science, Bratislava, and the Faculty of Mathematics and Physics, Charles University, Prague. [See *AJCL 7:2*, pg. 122.] Four copies of a 3-4 page, double-spaced summary must be submitted by December 1, 1981, to:

> COLING 82
> MFF UK, Linguistics
> Malostranské n. 25
> 118 00 Prague 1, CZECHOSLOVAKIA

Abstracts of Current Literature

KNET Extensions

Michael W. Freeman
ADO/FSSG
Burroughs Corporation
P.O. Box 517
Paoli, Pennsylvania 19301

Henry H. Leitner
Center for Research in Computing Technology
Aiken Computation Laboratory
Harvard University
Cambridge, Massachusetts 02138

This paper describes a number of extensions to Brachman's Structured Inheritance Network (SI-Net) formalism, with a view to creating an interactive system which will enable domain experts to develop and enhance knowledge networks (KNETs) representative of their particular application domains. The three principal extensions introduced are the QUA and VIZ links, and augmented event transition networks. The former permit us to partition the attribute space of conceptual entities in accordance with the defining roles played by these entities in certain kinds of socio-legal interactions. We use an event transition network to represent the value-class of so-called "dynamic attributes" in order to define all (and only all) the possible event sequences which the associated entity can participate in relative to the given attribute and still remain itself. The basic concept of simple succession of events in the transition network is then augmented by conditions on dependency relations which may exist across events among various roles (including bounds on the amount of time that can elapse between successive events). Since such event transition nets are themselves concepts within the KNET formalism, they can be placed along a specialization/generalization hierarchy. We show how these KNET extensions provide an extremely flexible basis for maintaining data base integrity and consistency, determining rollback dependencies during error recovery or historical perspectives, signalling possible irregularities in the attempted recording of new events which violate the legal event sequences defined by the net, as well as for generating derivative sets and mutually non-exclusive stative attribute values and associated interpretations for the wide range of ways this opens up for referencing such concepts in natural language.

Representing Knowledge in an Interactive Planner

Ann E. Robinson and David E. Wilkins

Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Proc. 1980 AAAI Conf., August 1980, 148-150.

This note discusses the representation for actions and plans being developed as part of the current planning research at SRI. Described is a method for uniformly representing actions that can take place both in the domain and during planning. The representation accommodates descriptions of abstract (hypothetical) objects.

A Computer Model of Child Language Learning

Mallory Selfridge

Electrical Engineering and Computer Science Dept.
The University of Connecticut
Storrs, Connecticut 06288

Proc. 1980 AAAI Conf., August 1980, 224-227.

A computer program modelling a child between the ages of 1 and 2 years is described. This program is based on observations of the knowledge this child had at age 1, the comprehension abilities he had at age 2, and the language experiences he had between these ages. The computer program described begins at the age 1 level, is given similar language experiences, and uses inference and learning rules to acquire comprehension at the age 2 level.

An Approach to Acquiring and Applying Knowledge

Norman Haas and Gary G. Hendrix

Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Proc. 1980 AAAI Conf., August 1980, 235-239.

The problem addressed in this paper is how to enable a computer system to acquire facts about new domains from tutors who are experts in their respective fields, but who have little or no training in computer science. The information to be acquired is that needed to support question-answering activities. The basic acquisition approach is "learning by being told." We have been especially interested in exploring the notion of simultaneously learning not only new concepts, but also the linguistic constructions used to express those concepts. As a research vehicle we have developed a system that is preprogrammed with deductive algorithms and a fixed set of syntactic/semantic rules covering a small subset of English. It has been endowed with sufficient seed concepts and seed vocabulary to support effective tutorial interaction. Furthermore,

the system is capable of learning new concepts and vocabulary, and can apply its acquired knowledge in a prescribed range of problem-solving situations.

Project EPISTLE: A System for the Automatic Analysis of Business Correspondence

Lance A. Miller

IBM Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, New York 10598

Proc. 1980 AAAI Conf., August 1980, 280-282.

The system described here is intended to provide the business executive with useful applications for the computer processing of correspondence in the office environment. Applications will include the synopsis and abstraction of incoming mail and a variety of critiques of newly-generated letters, all based upon the capability of understanding the natural language text at least to a level corresponding to customary business communication. Successive sections of the paper describe the background and prior work, the planned system output, and the implementation.

When Expectation Fails: Towards a Self-Correcting Inference System

Richard H. Granger, Jr.

Artificial Intelligence Project
Department of Information and Computer Science
University of California
Irvine, California 92717

Proc. 1980 AAAI Conf., August 1980, 301-305.

Contextual understanding depends on a reader's ability to correctly infer a context within which to interpret the events in a story. This "context-selection problem" has traditionally been expressed in terms of heuristics for making the correct *initial* selection of a story context. This paper presents a view of context selection as an ongoing process spread *throughout* the understanding process. This view requires that the understander be capable of recognizing and correcting erroneous initial context inferences. A computer program called ARTHUR is described, which selects the correct context for a story by dynamically re-evaluating its own initial inferences in light of subsequent information in a story.

Generating Relevant Explanations: Natural Language Responses to Questions about Database Structure

Kathleen R. McKeown

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Proc. 1980 AAAI Conf., August 1980, 306-309.

The research described here is aimed at unresolved problems in both natural language generation and natural language interfaces to database systems. How relevant information is selected and then organized for the generation of responses to questions about database structure is examined. Due to limited space, this paper reports on only one method of explanation, called "compare and contrast." In particular, it describes a specific constraint on relevancy and organization that can be used for this response type.

The Semantic Interpretation of Nominal Compounds

Timothy Wilking Finin

Computer and Information Science
Moore School of Electrical Engrg. D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Proc. 1980 AAAI Conf., Aug, 1980, 310-312.

This paper briefly introduces an approach to the problem of building semantic interpretations of *nominal compounds,* i.e. sequences of two or more nouns related through modification. Examples of the kinds of nominal compounds dealt with are: "engine repairs," "aircraft flight arrival," "aluminum water pump," and "noun noun modification."

Towards an AI Model of Argumentation

Lawrence Birnbaum, Margot Flowers, and Rod McGuire

Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

Proc. 1980 AAAI Conf., August 1980, 313-315.

This paper describes a process model of human argumentation, and provides examples of its operation as implemented in a computer program. Our main concerns include such issues as the rules and structures underlying argumentation, how these relate to conversational rules, how reasoning is used in arguments, and how arguing and reasoning interact.

Knowledge Representation for Syntactic/Semantic Processing

Robert J. Bobrow

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

Bonnie Lynn Webber

Computer and Information Science
Moore School of Electrical Engrg. D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Proc. 1980 AAAI Conf., August 1980, 316-323.

This paper describes the RUS framework for natural language processing, in which a parser incorporating a substantial ATN grammar for English interacts with a semantic interpreter to simultaneously parse and interpret input. The structure of that interaction is discussed, including the roles played by syntactic and semantic knowledge. Several implementations of the RUS framework are currently in use, sharing the same grammar, but differing in the form of their semantic component. One of these, the PSI-KLONE system, is based on a general object-centered knowledge representation system, called KL-ONE. The operation of PSI-KLONE is described, including its use of KL-ONE to support a general inference process called "incremental description refinement." The last section of the paper discusses several important criteria for knowledge representation systems to be used in syntactic and semantic processing.

Language and Memory: Generalization as a Part of Understanding

Michael Lebowitz

Computer Science Department
Columbia University
406 Mudd Building
New York, New York 10027

Proc. 1980 AAAI Conf., August 1980, 324-326.

This paper presents the Integrated Partial Parser (IPP), a computer model that combines text understanding and memory of events. An extended example of the program's ability to understand newspaper stories and make generalizations that are useful for memory organization is presented.

Failures in Natural Language Systems: Applications to Data Base Query Systems

Eric Mays

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Proc. 1980 AAAI Conf., August 1980, 327-330.

A significant class of failures in interactions with data base query systems is attributable to misconceptions or incomplete knowledge regarding the domain of discourse on the part of the user. This paper describes several types of user failures, namely, intensional failures of presumptions. These failures are distinguished from extensional failures of presumptions since they are dependent on the structure rather than the contents of the data base. A knowledge representation has been developed for the recognition of intensional failures that are due to the assumption of nonexistent relationships between entities. Several other intensional failures which depend on more sophisticated knowledge representations are also discussed. Ap-

propriate forms of corrective behavior are outlined which would enable the user to formulate queries directed to the solution of his/her particular task and compatible with the knowledge organization.

Organizing Memory and Keeping It Organized

Janet L. Kolodner
Information and Computer Science Department
Georgia Institute of Technology
Atlanta, Georgia 30332

Maintaining good memory organization is important in large memory systems. This paper presents a scheme for automatically reorganizing event information in memory. The processes are implemented in a computer program called CYRUS.

Narrative Text Summarization

Wendy Lehnert
Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

In order to summarize a story it is necessary to access a high level analysis that highlights the story's central concepts. A technique of memory representation based on affect units appears to provide the necessary foundation for such an analysis. Affect units are conceptual structures that overlap with each other when a narrative is cohesive. When overlapping intersections are interpreted as arcs in a graph of affect units, the resulting graph encodes the plot of the story. Structural features of the graph then reveal which concepts are central to the story. Affect unit analysis is currently being investigated as a processing strategy for narrative summarization.

Meta-Planning

Robert Wilensky
Electronics Research Laboratory
Computer Science Division
Department of EECS
University of California, Berkeley
Berkeley, California 94720

This paper is concerned with the problems of planning and understanding. These problems are related because a natural language understander must apply knowledge about people's goals and plans in order to make the inferences necessary to explain the behavior of a character in a story. Thus while a story understander is not a planner, it must embody a theory of planning knowledge. I have developed such a theory in the construction of PAM (Plan Applier Mechanism), a story understanding program. This paper is concerned not with the understanding mechanism itself, but that part of its planning knowledge which is independent of whether that knowledge is used to explain someone's behavior or to generate a plan for one's own use.

Issues in the Development of Natural Language Front-Ends

J. Hendler, T.P. Kehler, P.R. Michaelis,
B. Phillips, K.M. Ross, and H.R. Tennant
Texas Instruments
P.O. Box 225936
MS 371
Dallas, Texas 75023

This paper will discuss some issues we believe to be important to the design of a natural language front-end. These are divided into three categories: conceptual coverage, linguistic coverage, and implementation issues. The section on conceptual coverage discusses the use of a domain expert, which understands what the user is saying even though the system to which the front-end is interfaced might not be able to properly do what the user wants. The section on linguistic coverage discusses attempts to allow a natural language interface to handle natural, interactive human communication. Two solutions are explored: first, the design of a robust natural language understanding system composed of many experts that know about some aspect of the organization of language is considered; second, because the design of a robust system is a large task, the intermediate goal of limiting the vocabulary and constructions that can be used while retaining all the user-oriented benefits of natural language is considered. The implementation issues considered are the design of a system in which the grammar and the domain of discourse can be easily extended and which can be used for more than one domain without extensive rewriting.

Text-Critiquing with the EPISTLE System: An Author's Aid to Better Syntax

Lance A. Miller, George E. Heidorn, and Karen Jensen
IBM Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, New York 10598

The experimental EPISTLE system is ultimately intended to provide office workers with intelligent applications for the processing of natural language text, particularly business correspondence. A variety of possible critiques of textual material are identified in this paper, but the discussion focuses on the system's capability to detect several classes of grammatical errors, such as disagreement in number between the subject and the verb. The system's error-

detection performance relies critically on its parsing component which determines the syntactic structure of each sentence and the grammatical functions fulfilled by various phrases. Details of the system's operations are provided, and some of the future critiquing objectives are outlined.

Shifting to a Higher Gear in a Natural Language System

Bozena Henisz Thompson and Frederick B. Thompson
California Institute of Technology
Pasadena, California 91125

We have completed the development of the REL System, a system for communicating with the computer in natural language concerning a relational database. We have been using that system in a series of experiments on how people actually do communicate in solving an intellectual task. These experiments, together with our general experience with REL and related work elsewhere, have led us to the specification and development of a new system, the POL (Problem Oriented Language) System. POL is an evolutionary extension of REL, preserving what has worked, and extending and adding new capabilities to meet observed needs. These improvements include more responsive diagnostics, handling of sentence fragments, inter-knowledge-base communications, and new facilities for building and extending the knowledge bases of users. This paper introduces POL.

A Practical Comparison of Parsing Strategies

Jonathan Slocum
Linguistics Research Center
P.O. Box 7247
University of Texas
Austin, Texas 78712

Although the literature dealing with formal and natural languages abounds with theoretical arguments of worst-case performance by various parsing strategies, there is little discussion of comparative performance based on actual practice in understanding natural language. Yet important practical considerations do arise when writing programs to understand one aspect or another of natural language utterances. Where, for example, a theorist will characterize a parsing strategy according to its space and/or time requirements in attempting to analyze the worst possible input according to an arbitrary grammar strictly limited in expressive power, the researcher studying natural language processing can be justified in concerning himself more with issues of practical performance in parsing sentences encountered in language as humans actually use it, using a grammar expressed in a form convenient to the human linguist who is writing it.

This paper has two purposes. One is to report an evaluation of the performance of several parsing strategies in a real-world setting, pointing out practical problems in making the attempt, indicating which of the strategies is superior to the others in which situations, and most of all determining the reasons why the best strategy outclasses its competition in order to stimulate and direct the design of improvements. The other, more important purpose is to assist in establishing such evaluation as a meaningful and valuable enterprise that contributes to the evolution of natural language processing from an art form into an empirical science. That is, our concern for parsing efficiency transcends the issue of mere practicality.

In this paper we detail our experimental setting and approach, present the results, discuss the implications of those results, and conclude with some remarks on what has been learned.

Computational Complexity and Lexical Functional Grammar

Robert C. Berwick
AI Laboratory
Massachusetts Institute of Technology
545 Technology Square
Cambridge, Massachusetts 02139

Recently, a new theory of grammar has been advanced with the explicitly stated aim of meeting the dual demands of learnability and parsability — the Lexical Functional Grammars (LFGs) of Bresnan. The theory of Lexical Functional Grammars is claimed to have all the descriptive merits of transformational grammar, but none of its computational unruliness. In LFG, there are no transformations (as classically described); the work formerly ascribed to transformations such as "passive" is shouldered by information stored in lexical entries associated with lexical items. The elimination of transformational power naturally gives rise to the hope that a lexically-based system would be computationally simpler than a transformational one.

The main result of this paper is to show that certain Lexical Functional Grammars can generate languages whose recognition time *is* very likely computationally intractable, at least according to our current understanding of what is or is not rapidly solvable. Briefly, the demonstration proceeds by showing how a problem that is widely conjectured to be computationally difficult — namely, whether there exists an assignment of 1's and 0's (or "T"s and "F"s) to the literals of a Boolean formula in conjunctive normal form that makes the formula evaluate to "1" (or "true") — can be re-expressed as the problem of recognizing whether a particular string is or is not a mem-

ber of the language generated by a certain lexical functional grammar.

This paper also discusses the relevance of this technical result for more down-to-earth computational linguistics.

Corepresentational Grammar and Parsing English Comparatives

Karen Ryan
Linguistics Department
142 Klaeber Court
University of Minnesota
Minneapolis, Minnesota 55455

Marcus (1980) notes that the syntax of English comparative constructions is highly complex, and claims that both syntactic and semantic information must be available for them to be parsed. This paper argues that comparatives can be structurally analyzed on the basis of syntactic information alone via a strictly surface-based grammar. Such a grammar is given in Ryan (1981), based on the corepresentational model of Kac (1978). While the grammar does not define a parsing algorithm per se, it nonetheless expresses regularities of surface organization and its relationship to semantic interpretation that an adequate parser would be expected to incorporate. The central problem in parsing comparatives involves identifying the arguments of comparative predicates, and the relations borne by these arguments to such predicates. A corepresentational grammar is explicitly designed to assign predicate-argument structure to sentences on the basis of their surface syntactic organization. This paper discusses four problem areas in the description of comparatives and outlines the sections of the grammar of Ryan (1981) that apply to them.

Performance Comparison of Component Algorithms for the Phonemicization of Orthography

Jared Bernstein
Telesensory Speech Systems
3408 Hillview Avenue
Palo Alto, California 94304

Larry Nessly
University of North Carolina
Chapel Hill, North Carolina 27514

A system for converting English text into synthetic speech can be divided into two processes that operate in series: a text-to-phoneme converter, and a phonemic-input speech synthesizer. The conversion of orthographic text into a phonemic form may itself comprise several processes in series, such as formatting text to expand abbreviations and non-alphabetic expressions, parsing and word class determination, segmental phonemicization of words, word and clause level stress assignment, word internal and word boundary allophonic adjustments, and duration and fundamental frequency settings for phonological units.

Comparing the accuracy of different algorithms for text-to-phoneme conversion is often difficult because authors measure and report system performance in incommensurable ways. Furthermore, comparison of the output speech from two complete systems may not always provide a good test of the performance of the corresponding component algorithms in the two systems, because radical performance differences in other components can obscure small differences in the components of interest. The only reported direct comparison of two complete text-to-speech systems (MITALK and TSI's TTS-X) was conducted by Bernstein and Pisoni. This paper reports one study that compared two algorithms for automatic segmental phonemicization of words, and a second study that compared two algorithms for automatic assignment of lexical stress.

PHONY: A Heuristic Phonological Analyzer

Lee A. Becker
Department of Computer Science
Indiana University
101 Lindley Hall
Bloomington, Indiana 47401

PHONY is a program to do phonological analysis. Within the generative model of grammar the function of the phonological component is to assign a phonetic representation to an utterance by modifying the "underlying representations" (URs) of its constituent morphemes. URs are abstract entities which contain the idiosyncratic information about pronunciations of morphemes.

The input to PHONY is pronunciations of words and phrases upon which a preliminary morphological analysis has been completed. They have been divided into morphemes, and different instances of the same morpheme have been associated. These are represented as strings of phonetic symbols including morpheme- and word-boundaries. Indices are used to associate various instances of the same morpheme. The output of PHONY is a set of phonological rules or regularities in the data, as well as a set of underlying representations for the morphemes. The phonological rules generate the various pronunciations of the morphemes from their underlying representations.

Two Discourse Generators

William C. Mann
USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, California 90291

The task of discourse generation is to produce multisentential text in natural language which (when heard or read) produces effects (informing, motivation, etc.) and impressions (conciseness, correctness, ease of reading, etc.) which are appropriate to a need or goal held by the creator of the text. In comparing two AI discourse generators here we can do no more than suggest opportunities and attractive options for future exploration. We describe them only in terms of a few of the techniques which they employ, partly because these techniques seem more valuable than the system designs in which they happen to have been used. The systems which we describe are PROTEUS, by Anthony Davey at Edinburgh and KDS by Mann and Moore at ISI, both of which are severely limited and idiosyncratic in scope and technique. First we identify particular techniques in each system which contribute strongly to the quality of the resulting text. Then we compare the two systems discussing their common failings and the possibilities for creating a system having the best of both.

A Grammar and a Lexicon for a Text-Production System

Christian Matthiessen
USC/Information Sciences Institute
4676 Admiralty Way
Marina del Rey, California 90291

In a text-production system high and special demands are placed on the grammar and the lexicon. This paper views these components in such a system. First, the subcomponents dealing with semantic information and with syntactic information are presented separately. The problems of relating these two types of information are then identified. Finally, strategies designed to meet the problems are proposed and discussed. One of the issues that is illustrated is what happens when a systemic linguistic approach is combined with a KL-ONE like knowledge representation — a novel and hitherto unexplored combination.

Language Production: The Source of the Dictionary

David D. McDonald
Computer and Information Science
University of Massachusetts
Amherst, Massachusetts 01002

Ultimately in any natural language production system the largest amount of human effort will go into the construction of the *dictionary:* the data base that associates objects and relations in the program's domain with the words and phrases that could be used to describe them. This paper describes a technique for basing the dictionary directly on the semantic abstraction network used for the domain knowledge itself, taking advantage of the inheritance and specialization mechanisms of a network formalism such as KL-ONE. The technique creates considerable economies of scale, and makes possible the automatic description of individual objects according to their position in the semantic net. Furthermore, because the process of deciding what properties to use in an object's description is now given over to a common procedure, we can write general-purpose rules to, for example, avoid redundancy or grammatically awkward constructions.

Analogies in Spontaneous Discourse

Rachel Reichman
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

This paper presents an analysis of analogies based on observations of natural conversations. People's spontaneous use of analogies provides insight into their implicit evaluation procedures for analogies. The treatment here, therefore, reveals aspects of analogical processing that are somewhat more difficult to see in an experimental context. The work involves explicit treatment of the discourse context in which analogy occurs. A major focus here is the formalization of the effects of analogy on discourse development. There is much rule-like behavior in this process, both in underlying thematic development of the discourse and in the surface linguistic forms used in this development. Both these forms of regular behavior are discussed in terms of a hierarchical structuring of a discourse into distinct, but related and linked, context spaces.

Investigation of Processing Strategies for the Structural Analysis of Arguments

Robin Cohen
Department of Computer Science
University of Toronto
Toronto, Ontario CANADA M5S 1A7

This paper outlines research on processing strategies being developed for a language understanding system, designed to interpret the structure of arguments. For the system, arguments are viewed as trees, with claims as fathers to their evidence. Then understanding becomes a problem of developing a representative argument tree, by locating each proposition of the argu-

ment at its appropriate place. The processing strategies we develop for the hearer are based on expectations that the speaker will use particular coherent transmission strategies and are designed to be fairly efficient (work in linear time). We also comment on the use by the speaker of linguistic clues to indicate structure, illustrating how the hearer can interpret the clues to limit his processing search and thus improve the complexity of the understanding process.

What's Necessary to Hide?: Modelling Action Verbs

James F. Allen

Department of Computer Science
University of Rochester
Mathematical Sciences Building
Rochester, New York 14627

Proc. 19th Annual ACL Conf., June 1981, 77-83.

This paper considers what types of knowledge one must possess in order to reason about actions. Rather than concentrating on how actions are performed, as is done in the problem-solving literature, it examines the set of conditions under which an action can be said to have occurred. In other words, if one is told that action A occurred, what can be inferred about the state of the world? In particular, if the representation can define such conditions, it must have good models of time, belief, and intention. This paper discusses these issues and suggests a formalism in which general actions and events can be defined. Throughout, the action of hiding a book from someone is used as a motivating example.

A Rule-based Conversation Participant

Robert E. Frederking

Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania 15213

Proc. 19th Annual ACL Conf., June 1981, 83-87.

The problem of modelling human understanding and generation of a coherent dialog is investigated by simulating a conversation participant. The rule-based system currently under development attempts to capture the intuitive concept of "topic" using data structures consisting of declarative representations of the subjects under discussion linked to the utterances and rules that generated them. Scripts, goal trees, and a semantic network are brought to bear by general, domain-independent conversational rules to understand and generate coherent topic transitions and specific output utterances.

Search and Inference Strategies in Pronoun Resolution: An Experimental Study

Kate Ehrlich

Department of Psychology
University of Massachusetts
Amherst, Massachusetts 01003

Proc. 19th Annual ACL Conf., June 1981, 89-93.

The process of assigning a referent to a pronoun can be viewed as utilizing two kinds of strategies. One strategy is concerned with selecting the best referent from among the candidates available. The other strategy is concerned with searching through memory for the candidates. These two types of strategies, which will be referred to mnemonically as inference and search strategies, have different kinds of characteristics. A search strategy dictates the order in which candidates are evaluated, but has no machinery for carrying out the evaluation. The inference strategy helps to set up the representation of the information in the text against which candidates can be evaluated, but has no way of finding the candidates. In this paper, the way these strategies might interact is explored and the results of two studies that bear on the issues are reported on.

Presupposition and Implicature in Model-Theoretic Pragmatics

Douglas B. Moran

Department of Computer Science
Oregon State University
Corvallis, Oregon 97331

Proc. 19th Annual ACL Conf., June 1981, 107-108.

Model-theoretic pragmatics is an attempt to provide a formal description of the pragmatics of natural language as effects arising from using model-theoretic semantics in a dynamic environment. The pragmatic phenomena considered here have been variously labeled *presupposition* and *conventional implicature*.

The models used in traditional model-theoretic semantics provide a complete and static representation of knowledge about the world. However, this is not the environment in which language is used. Language is used in a dynamic environment — the participants have incomplete knowledge of the world and the understanding of a sentence can add to the knowledge of the listener. A formalism which allows models to contain incomplete knowledge and to which knowledge can be added has been developed.

Some Computational Aspects of Situation Semantics

Jon Barwise

Department of Philosophy
Stanford University
Stanford, California 94305

Proc. 19th Annual ACL Conf., June 1981, 109-111.

Can a realist model theory of natural language be computationally plausible? Or, to put it another way, is the view of linguistic meaning as a relation between expressions of natural language and things (objects, properties, etc.) in the world, as opposed to a relation between expressions and procedures in the head, consistent with a computational approach to understanding natural language? The model theorist must either claim that the answer is yes, or be willing to admit that humans transcend the computationally feasible in their use of language.

Until recently the only model theory of natural language that was at all well developed was Montague Grammar. Unfortunately, it was based on the primitive notion of "possible world" and so was not a realist theory, unless you are prepared to grant that all possible worlds are real. Montague Grammar is also computationally intractable, for reasons discussed in the paper.

John Perry and I have developed a somewhat different approach to the model theory of natural language, a theory we call "Situation Semantics". Since one of my own motivations in the early days of this project was to use the insights of generalized recursion theory to find a computationally plausible alternative to Montague Grammar, it seems fitting to give a progress report here.

A Situation Semantics Approach to the Analysis of Speech Acts

David Andreoff Evans

Stanford University
Stanford, California 94305

During the past two decades, much work in linguistics has focused on sentences as minimal units of communication, and the project of rigorously characterizing the structure of sentences in natural language has met with some success. Not surprisingly, however, sentence grammars have contributed little to the analysis of discourse. Human discourse consists not just of words in sequences directed by a speaker to an addressee, used to *represent situations* and to *reveal intentions*. Only when the addressee has apprehended both these aspects of the message communicated can the message be interpreted. The analysis of discourse that emerges from Austin's work, grounded in a theory of action, takes this view as central, and the concept of the speech act follows naturally. This paper describes a situation semantics approach to the analysis of speech acts.

Problems in Logical Form

Robert C. Moore

Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

This paper surveys what we at SRI view as some of the key problems encountered in defining a system of representation for the logical forms of English sentences, and suggests possible approaches to their solution. We first look at some general issues related to the notion of logical form, and then discuss a number of problems associated with the way information involving certain key concepts is expressed in English. Although our main concern here is with theoretical issues rather than with system performance, this paper is not merely speculative. The DIALOGIC system currently under development in the SRI Artificial Intelligence Center parses English sentences and translates them into logical forms embodying many of the ideas presented here.

A Case for Rule-Driven Semantic Processing

Martha Palmer

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

The primary task of semantic processing is to provide an appropriate mapping between the syntactic constituents of a parsed sentence and the arguments of the semantic predicates implied by the verb. This is known as the Alignment Problem. Section One of this paper gives an overview of a generally accepted approach to semantic processing that goes through several levels of representation to achieve this mapping. Although somewhat inflexible and cumbersome, the different levels succeed in preserving the context sensitive information provided by verb semantics. Section Two presents the author's rule-driven approach which is more uniform and flexible yet still accommodates context sensitive constraints. This approach is based on general underlying principles for syntactic methods of introducing semantic arguments and has interesting implications for linguistic theories about case. These implications are discussed in Section Three. A system that implements this approach has been designed for and tested on pulley problem statements gathered from several physics text books.

A Taxonomy for English Nouns and Verbs

Robert A. Amsler

Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

The definition texts of a machine-readable pocket dictionary were analyzed to determine the disambiguated word sense of the kernel terms of each word sense being defined. The resultant sets of word pairs of defined and defining words were then computationally connected into two taxonomic semi-lattices ("tangled hierarchies") representing some 24,000 noun nodes and 11,000 verb nodes. The study of the nature of the "topmost" nodes in these hierarchies and the structure of the trees reveal information about the nature of the dictionary's organization of the language, the concept of semantic primitives and other aspects of lexical semantics. The data proves that the dictionary offers a fundamentally consistent description of word meaning and may provide the basis for future research and applications in computational linguistic systems.

Interpreting Natural Language Database Updates

S. Jerrold Kaplan and Jim Davidson

Department of Computer Science
Stanford University
Stanford, California 94305

Although the problem of *querying* a database in natural language has been studied extensively, there has been relatively little work on processing database *updates* expressed in natural language. To interpret update requests, several linguistic issues must be addressed that do not typically pose difficulties when dealing exclusively with queries. The primary difficulty with interpreting natural language updates is that there may be several ways in which a particular update can be performed in the underlying database. Many of these options, while literally correct and semantically meaningful, may correspond to bizarre interpretations of the request. While human speakers would intuitively reject these unusual readings, a computer program may be unable to distinguish them from more appropriate ones. If carried out, they often have undesirable side effects on the database. Our approach to this problem is to generate a limited set of "candidate" updates, rank them according to a set of domain-independent heuristics that reflect general properties of "reasonable" updates, and either perform the update or present the highest ranked options to the user for selection. This paper briefly examines some of the linguistic problems encountered, and describes an implemented system that performs simple natural language database updates.

Dynamic Strategy Selection in Flexible Parsing

Jaime G. Carbonell and Philip J. Hayes

Department of Computer Science
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213

Robust natural language interpretation requires strong semantic domain models, "fail-soft" recovery heuristics, and very flexible control structures. Although single-strategy parsers have met with a measure of success, a multi-strategy approach is shown to provide a much higher degree of flexibility, redundancy, and ability to bring task-specific domain knowledge (in addition to general linguistic knowledge) to bear on both grammatical and ungrammatical input. A parsing algorithm is presented that integrates several different parsing strategies, with case-frame instantiation dominating. Each of these parsing strategies exploits different types of knowledge; and their combination provides a strong framework in which to process conjunctions, fragmentary input, and ungrammatical structures, as well as less exotic, grammatically correct input. Several specific heuristics for handling ungrammatical input are presented within this multi-strategy framework.

A Construction-Specific Approach to Focused Interaction in Flexible Parsing

Philip J. Hayes

Department of Computer Science
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213

A flexible parser can deal with input that deviates from its grammar, in addition to input that conforms to it. Ideally, such a parser will correct the deviant input; sometimes, it will be unable to correct it at all; at other times, correction will be possible, but only to within a range of ambiguous possibilities. This paper is concerned with such ambiguous situations, and with making it as easy as possible for the ambiguity to be resolved through consultation with the user. We show the importance of asking the user for clarification in as focused a way as possible. Focused interaction of this kind is facilitated by a construction-specific approach to flexible parsing, with specialized parsing techniques for each type of construction, and specialized ambiguity representations for each type of ambiguity that a particular construction can give rise to. A construction-specific approach also aids in task-specific language development by allowing a language definition

that is natural in terms of the task domain to be inter-
preted directly without compilation into a uniform
grammar formalism, thus greatly speeding the testing
of changes to the language definition.

Controlled Transformational Sentence Generation

Madeleine Bates

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts, 02138

Robert Ingria

Department of Linguistics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Proc. 19th Annual ACL Conf., June 1981, 153-158.

This paper describes a transformational sentence
generator that was built primarily to focus on syntactic
form and syntactic relationships. Our main goal was
to produce a tutorial system for the English language;
the intended users of the system are people with lan-
guage delaying handicaps such as deafness, and people
learning English as a foreign language. For these pop-
ulations, extensive exposure to standard English con-
structions (negatives, questions, relativization, etc.)
and their interactions is necessary. The purpose of the
generator was to serve as a powerful resource for tuto-
rial programs that need examples of particular con-
structions and/or related sentences to embed in exer-
cises or examples for the student. The focus of the
generator is thus not so much on what to express as on
how to express it in acceptable English. The generator
is composed of three major parts: a base component
that produces base trees, a transformer that applies
transformational rules to the trees to derive a surface
tree, and a set of mechanisms to control the operation
of the first two components. We discuss each of these
components separately.

Transportable Natural-Language Interfaces to Databases

Gary G. Hendrix and William H. Lewis

Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Proc. 19th Annual ACL Conf., June 1981, 159-162.

Over the last few years a number of application
systems have been constructed that allow users to
access databases by posing questions in natural lan-
guages, such as English. When used in the restricted
domains for which they have been especially designed,
these systems have achieved reasonably high levels of
performance. Such systems as LADDER, PLANES,
ROBOT, and REL require the encoding of knowledge
about the domain of application in such constructs as

database schemata, lexicons, pragmatic grammars, and
the like. The creation of these data structures typical-
ly requires considerable effort on the part of a com-
puter professional who has had special training in
computational linguistics and the use of databases.
Thus, the utility of these systems is severely limited by
the high cost involved in developing an interface to
any particular database.

This paper describes initial work on a methodology
for creating natural-language processing capabilities
for new domains without the need for intervention by
specially trained experts. Our approach is to acquire
logical schemata and lexical information through sim-
ple interactive dialogues with someone who is familiar
with the form and content of the database, but unfa-
miliar with the technology of natural-language inter-
faces. To test our approach in an actual computer
environment, we have developed a prototype system
called TED (Transportable English Datamanager). As
a result of our experience with TED, the NL group at
SRI is now undertaking the development of a much
more ambitious system based on the same philosophy.

Chart Parsing and Rule Schemata in PSG

Henry Thompson

Department of Artificial Intelligence
University of Edinburgh
Hope Park Square
Meadow Lane, Edinburgh EH8 9NW

Proc. 19th Annual ACL Conf., June 1981, 167-172.

MCHART is a flexible, modular chart parsing
framework I have been developing (in LISP) at Edin-
burgh, whose initial design characteristics were largely
determined by pedagogical needs. PSG is a grammati-
cal theory developed by Gerald Gazdar at Sussex, in
collaboration with others in both the U.S. and Britain,
most notably Ivan Sag, Geoff Pullum, and Ewan Klein.
It is a notationally rich context free phrase structure
grammar, incorporating meta-rules and rule schemata
to capture generalizations.

In this paper I describe how I have used MCHART
in beginning to construct a parser for grammars ex-
pressed in PSG, and how aspects of the chart parsing
approach in general and MCHART in particular have
made it easy to accommodate two significant aspects
of PSG: rule schemata involving variables over cate-
gories, and compound category symbols ("slash" cate-
gories). To do this I briefly introduce the basic ideas
of chart parsing, describe the salient aspects of
MCHART, give an overview of PSG, and finally pres-
ent the interesting aspects of the parser I am building
for PSG using MCHART.

Generating Descriptions and Explanations: Applications to Questions about Database Structure

Kathleen R. McKeown

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-9, 1980.

The research being proposed is within the area of natural language generation. The generation process may be regarded as consisting of two major phases of computation; the first will determine what is to be said and how it is to be said, and the second will translate that message from an internal representation to English. Emphasis will be placed on the problems involved in the first component of generation.

The application for the generation of natural language is within a natural language interface to a database system. To date, the kinds of answers which can be generated by such systems are restricted to lists or tables of objects in the database. I am proposing generating responses to questions about the structure of the database. The kinds of responses that will be generated include descriptions of classes of objects in the database, differences between the classes, relations that hold between classes, information available in the database, and definitions of classes of objects, among others.

A Technique for Managing the Lexicon in a Natural Language Interface to a Changing Data Base

S. Jerrold Kaplan, Eric Mays, and Aravind K. Joshi

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-10, August 1979.

A difficulty in designing a Natural Language (NL) interface to a Data Base (DB) management system is insuring that DB updates do not obsolete the NL components. Of particular concern is the *lexicon* (the list of words that the system can process), mainly because NL queries can contain terms that appear as values in the DB, and hence are subject to change as the contents of the DB changes. For example, to process the question "Does John Jones still work for the company?", it is necessary to identify the string "John Jones" as a (potential) value in the "EMPLOYEE-NAME" field (assuming a suitable DB). If such names must appear in the lexicon for the system to process the query, then the lexicon will go out of date as the company's personnel shifts. Using the DB itself as an extension of the lexicon is equally problematic: the system will be unable to parse and provide a nega-

tive answer to such a question if the name does not appear at all in the DB.

The approach suggested here is to infer a plausible field from the context of the query and semantic information about the domain that is implicitly encoded in the structure of the DB. This technique has been implemented in CO-OP, a NL DB query system that provides cooperative responses and operates with a typical CODASYL DB system. CO-OP treats the problem of selecting a plausible field as a special case of resolving *semantic ambiguities*. Examples drawn from the implementation are presented.

Centered Logic: The Role of Entity Centered Sentence Representation in Natural Language Inferencing.

Aravind K. Joshi and Steve Kuhn

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-11, August 1979.

We will briefly describe the role of entity centered structure (ECS) of sentences in natural language inferencing. The basic structure in discourse generally singles out an entity, to be called center, among all those which are the arguments of the main predicate. ECS makes n-ary predicates look like monadic ones by temporarily masking their structure, thereby affecting the relative ease with which certain inferences are made and information is retrieved. This short paper deals with a preliminary formulation of a system designed to capture these ideas and contains several examples of how some natural language inferences can be represented in the system. Formal properties of the system are under investigation.

Paraphrasing Using Given and New Information in a Question-Answer System

Kathleen R. McKeown

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-13, 1980.

The design and implementation of a paraphrase component for a natural language question-answer system (CO-OP) is presented. A major point made is the role of given and new information in formulating a paraphrase that differs in a meaningful way from the user's question. A description is also given of the transformational grammar used by the paraphraser to generate questions.

Natural Language Interaction with Dynamic Knowledge Bases: Monitoring as Response

E. Mays, S. Lanka, A.K. Joshi, and B.L. Webber
Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-46, 1980.

In this communication, we discuss an interesting aspect of natural language interaction with dynamically changing knowledge bases - the ability to monitor for relevant future changes in that knowledge. We also indicate the status of our current work in this area and the overall goals of our research on question-answering and monitoring dynamic knowledge bases.

Control of Inference: Role of Some Aspects of Discourse Structure-Centering

Aravind K. Joshi and Scott Weinstein
Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-47, 1980.

The purpose of this communication is to examine one particular aspect of discourse structure, namely, a discourse construct called *center* of a sentence (utterance) *in discourse* and its relation to the larger issue of control of inference. We have described very briefly the notion of center(s) of a sentence in discourse and discussed how the centering phenomenon might be incorporated in a formal model of inference and its relation to the intrinsic complexity of certain inferences.

Varieties of Cooperative Responses in Question-Answer Systems

Aravind K. Joshi
Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-48, 1980.

In this paper, certain types of cooperative responses desirable in question-answer systems in a data base environment have been briefly reviewed. In particular, cooperative responses have been considered dealing with the situation when the user's view and the system's view of the structure and/or content of the data base are disparate. Responses explaining the structure of the data base as well as responses which are implicit requests for monitoring states of dynamic data bases have also been discussed. Finally, a general formulation of a particular type of cooperative behavior has been briefly described.

Phrase Structure Trees Bear More Fruit Than You Would Have Thought

Aravind K. Joshi
Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-49, 1980.

Several results concerning phrase structure trees have been presented. These results show that phrase structure trees when viewed in certain ways have much more descriptive power than one would have thought. A brief account of local constraints on structural descriptions and an intuitive proof have been presented. The local constraints approach has been compared to some aspects of Gazdar's framework and that of Peters and Karttunen. Some results on skeletons (phrase structure trees without labels) have been presented also. It can be shown that phrase structure trees even when deprived of the labels retain in a certain sense all the structural information. This result has implications for grammatical inference procedures.

Parasession on Topics in Interactive Discourse: Influence of the Problem Text

Aravind K. Joshi
Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-50, 1980.

This paper was presented in the parasession on topics in interactive discourse, which was a special session organized in conjunction with the 18th Annual Meeting of the Association for Computational Linguistics (June, 1980, Philadelphia). The paper consists of comments in response to several issues raised by Barbara Grosz, the panel chairperson. All the comments pertain to the primary issue of how the purpose of the interaction of the problem context affects what is said and how it is interpreted. Wherever possible, the issues are discussed more in the context of "information seeking" interaction and the data base domain.

Mutual Beliefs in Question-Answer Systems

Aravind K. Joshi
Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

Technical Report MS-CIS-80-51, 1980.

In this paper, we will briefly discuss the role of mutual beliefs with respect to some specific aspects of man-machine interactions. Cooperative or helpful behavior can be defined in various ways; however in this paper we will deliberately limit ourselves to some

very particular aspects of cooperation and relate them to the more technical definitions of mutual beliefs. We will be particularly concerned with the case where cooperativeness involves both giving a truthful and informative response and "squaring away" the relevant mutual beliefs. In this context, we will suggest that the "squaring away" contribution of the interaction can be explained by a suitable modification of one of the maxims of cooperative conversation.

We also discuss a related topic dealing with excess (or surplus) information and how and when the system should assimilate this information, i.e., implicitly update itself without being explicitly told to do so. These considerations require that mutual beliefs should not be regarded just as a set of propositions but rather with some structure over them where the structuring of information itself is a mutual belief.

Correcting Misconceptions About Data Base Structure

Eric Mays

Computer and Information Science
Moore School of Electrical Engineering D2
University of Pennsylvania
Philadelphia, Pennsylvania 19104

This paper presents a method for computation of intensional failures of presumptions in queries to a natural language interface to a data base system. These failures are distinguished from extensional failures since they are dependent on the structure rather than the content of the data base. A knowledge representation has been investigated that can be used to recognize intensional failures. When intensional failures are detected, a form of corrective behavior is proposed to inform the user about possibly relevant data base structure that is related to the failure.

BORIS -- An Experiment in In-Depth Understanding of Narratives

Wendy Lehnert, Michael G. Dyer, Peter N. Johnson, C.J. Yang, and Steve Harley

Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

BORIS is a story understanding and question answering system which involves the specification and interaction of many sources of knowledge. Unlike skimmers, which simply extract the "gist" of a story in a top-down manner and ignore everything else, BORIS attempts to understand everything that it reads to as great a depth as possible. This report focuses on how the BORIS program handles a complex story involving a divorce.

Conceptual Information Retrieval

Roger Schank, Janet Kolodner, and Gerald DeJong

Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

If we want to build intelligent information retrieval systems, we will have to give them the capabilities of understanding natural language, automatically organizing and reorganizing their memories, and using intelligent heuristics for searching their memories. These systems will have to analyze and understand both new text and natural language queries. In answering questions, they will have to direct memory search to reasonable places. This requires good organization of both the conceptual content of texts and knowledge necessary for understanding those texts and accessing memory.

The CYRUS and FRUMP systems (Kolodner (1978), Schank and Kolodner (1979), DeJong (1979)) comprise an information retrieval system called CyFr. Together, they have the analysis and retrieval capabilities mentioned above. Frump analyzes news stories from the UPI wire for their conceptual content, and produces summaries of those stories. It sends summaries of stories about important people to CYRUS. CYRUS automatically adds those stories to its memory, and can then retrieve that information to answer questions posed to it in natural language.

This paper describes the problems involved in building such an intelligent system. It proposes solutions to some of those problems based on recent research in artificial intelligence and natural language processing, and describes the CyFr system, which implements those solutions. The solutions we propose and implement are based on a model of human understanding and memory retrieval.

Finding Objects With Given Spatial Properties

Drew McDermott

Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

An important class of queries for a spatial-retrieval system is the retrieval of an object given some desired properties, such as that it be located in a given region, or point in a given direction. Handling this kind of query requires using a discrimination tree, which breaks the space of possibilities down repeatedly, ultimately into manageable buckets. In a spatial system, a useful kind of retrieval is "spiral search," in which the system searches the tree by starting from the "best" cell and gradually enlarging its attack. A good way of

implementing such a module is to create a domain-independent system, data-driven by specialized user functions. In this scheme, the user functions can decide how to index an object, how to search a node of the tree, and how to reorganize a node when it is not discriminated properly. Special problems arise when this program is applied to the domain of simple shaped objects; a shaped object can fall into more than one bucket, and a shaped object can have different lengths in different directions. The user must tell the system what frames of reference to use for computing object coordinates.

A Temporal Logic for Reasoning About Processes and Plans

Drew McDermott

Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

Much previous work in artificial intelligence has neglected representing time in all its complexity. In particular, it has neglected continuous change and the indeterminacy of the future. To rectify this, I have developed a first-order temporal logic, in which it is possible to name and prove things about facts, events, plans, and world histories. In particular, the logic provides analyses of causality, continuous change in quantities, the persistence of facts (the frame problem), and the relationship between tasks and actions. It may be possible to implement a temporal-inference machine based on this logic, which keeps track of several "maps" of a time line, one per possible history.

What's the Point?

Roger C. Schank, Gregg C. Collins, Ernest Davis, Peter N. Johnson, Steve Lytinen, and Brian J. Reiser

Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

Understanding dialogue usually requires determining the intent, or point, of the utterance. Finding the point serves to constrain further processing. We present a categorization of points, and we propose algorithms and heuristics for deriving the point of a given utterance.

MAGPIE: A Goal-Based Model of Conversation

Peter N. Johnson and Scott P. Robertson

Department of Computer Science
Box 2158 Yale Station
Yale University
New Haven, Connecticut 06520

The importance of intention in conversation has been considered by many researchers in artificial intelligence and psychology. However, most models of conversation have been limited to pursuing the transfer of knowledge between the system and a user. We propose that conversational goals can address communication at a number of other levels, such as the conversants' emotions, their relationship, and their attitudes. MAGPIE (Multiple Active Goal Processor in Interactive Exchanges) is a computer model of a conversant that acquires and pursues conversational goals at a number of levels, including the goal of seeking dominance in its relationship with the other conversant. At the heart of the program is a set of tracking procedures, each of which monitors a specific level of communication flow in a conversation. These procedures are coupled with a conversational goal planner which generates responses that simultaneously pursue a number of goals. Currently, MAGPIE is able to model a wife during a short marital dispute with her husband. Normative data from human subjects is presented which supports the conversational goals proposed in our analysis.

Knowledge Organization and Distribution for Medical Diagnosis

Fernando Gomez and B. Chandrasekaran

Department of Computer and Information Science
The Ohio State University
Columbus, Ohio 43210

A diagnostician, when he arrives at a diagnosis or diagnoses, has invoked some concepts. They can be diseases, causes of them, or other notions that are relevant to the diagnosis. These concepts form a hierarchical structure similar to a botanical or zoological classification. The diagnostician's knowledge is distributed through this hierarchy. The concepts in the hierarchy provide the criteria to organize under them small pieces of knowledge represented in the form of production rules. Thus concepts may be viewed as clusters of production rules. They extend the capabilities of production rules to more complex problem solving situations. The rules under each concept are further organized into three subgroups: exclusionary, confirmatory, and recommendation rules. During the problem solving process, the concepts are taken as *specialists*. They interact and communicate among themselves by means of a blackboard.

Toward a Theory of Distributed Word Expert Natural Language Parsing

Chuck Rieger and Steve Small

Department of Computer Science
University of Maryland
College Park, Maryland 20742

An approach to natural language meaning-based parsing in which the unit of linguistic knowledge is the word rather than the rewrite rule is described. In the word expert parser, knowledge about language is distributed across a population of procedural experts, each representing a word of the language, and each an expert at diagnosing that word's intended usage in context. The parser is structured around a coroutine control environment in which the generator-like word experts ask questions and exchange information in coming to collective agreement on sentence meaning. The word expert theory is advanced as a better cognitive model of human language expertise than the traditional rule-based approach. The technical discussion is organized around examples taken from the prototype LISP system which implements parts of the theory.

Integrating Knowledge Sources for Computer "Understanding" Tasks

Richard E. Cullingford

Department of Electrical Engineering and Computer Science
University of Connecticut
Storrs, Connecticut 06268

Recent research in artificial intelligence has identified a number of knowledge sources which appear to be needed for effective automatic "understanding" of connected natural language speech and text. These include word- and phrase-level semantics, models for actors and objects, and inference techniques for using script- or goal-oriented knowledge structures. A unified model of the understanding process can be defined using the distributed-computing viewpoint, provided some way can be found to integrate and control a collection of "experts," each one associated with a certain kind of knowledge source. A technique is described, called hierarchical task management, for constructing computer language-processing systems comprising an arbitrary number of distinct, potentially distributed, processes. The technique is based upon the repeated activation and expansion of data structures called tasks, which define important components of the understanding process in terms of a controlled interaction among the experts. The tasks are maintained on several "agendas," and are manipulated by a uniform monitor called the task manager. The process of task management is illustrated in a multiprocess story understander called a distributable script applier mechanism (DSAM), which reads and summarizes newspaper stories about plane crashes.

A Re-Evaluation of Story Grammars

Alan M. Frisch and Donald Perlis

Department of Computer Science
Mathematical Sciences Building
The University of Rochester
Rochester, New York 14627

Black and Wilensky (1979) have made serious methodological errors in analyzing story grammars, and in the process they have committed additional errors in applying formal language theory. Our arguments involve clarifying certain aspects of knowledge representation crucial to a proper treatment of story understanding.

Particular criticisms focus on the following shortcomings of their presentation: (1) an erroneous statement from formal language theory, (2) misapplication of formal language theory to story grammars, (3) unsubstantiated and doubtful analogies with English grammar, (4) various non sequiturs concerning the generation of non-stories, (5) a false claim based on the artificial distinction between syntax and semantics, and (6) misinterpretation of the role of story grammars in story understanding. We conclude by suggesting appropriate criteria for the evaluation of story grammars.

A Model for Planning in Complex Situations

Robert Wilensky

Electronics Research Laboratory
Computer Science Division
Department of EECS
University of California, Berkeley
Berkeley, California 94720

A model of planning applicable to complex, commonplace activities is being developed. This model differs from previous approaches in that it is based on the following assumptions: (1) a planning agent must be able to infer its own goals in addition to being able to generate plans for these goals; (2) everyday planning is primarily concerned with reasoning about the interactions between plans and goals; (3) *metaplanning* (formulating knowledge about how to plan abstract plans and goals, and having the planner use this knowledge to solve its own planning problems) is used as a driving principle; (4) *projection* (simulating hypothetical futures based on current plans and world knowledge) is used to infer goals and debug plans; and (5) planning knowledge should be equally available for understanding as well as for planning.

Coupled together, the mechanisms for implementing these features give rise to a system of considerable

power. For example, the ability to infer one's own goals is needed in an autonomous planning agent since it must deal with unexpected situations. However, together with meta-planning knowledge and the ability to project hypothetical futures, this feature enables the planner to detect and reason about complicated goal interactions, anticipate problems with proposed plans, and make choices in the face of competing alternatives.

This model has been developed in detail for the detection and resolution of goal conflicts. In doing so, we postulate several goal conflict resolution strategies, or meta-plans, called RE-PLAN, CHANGE-CIRCUMSTANCE, and SIMULATE-AND-SELECT. The structure and application of these meta-plans is explored in the context of both decision making and understanding the actions of other planners.

It's for Your Own Good:
A Note on Inaccurate Reference

C. Raymond Perrault
Department of Computer Science
University of Toronto
Toronto, Ontario CANADA M5S 1A7

Philip R. Cohen
Department of Computer Science
Oregon State University
Corvallis, Oregon 97331

In his book *Speech Acts,* Searle suggests that his analysis of illocutionary acts can be extended to account for the "propositional act" of reference. He wants to give necessary conditions for a speaker successfully referring to some entity by using a certain referring expression in an utterance to a certain hearer. We point out here some inadequacies in Searle's conditions, in particular how they fail to account for cases of successful reference through expressions which speaker (and hearer) may believe are not true of their intended referent. More specific conditions based on the notion of mutual belief are proposed.

Stories within Stories

Bertram Bruce
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

What appears to be a single story is often a complex set of stories within stories, each with its distinct author and reader. Examples of such stories within stories are presented. Results of analyses of basal readers and trade books in terms of embedded stories are also discussed. These suggest that a greater variety of stories could and should be made available to children.

Higher-Level Features in Children's Stories:
Rhetorical Structure and Conflict

Cindy Steinberg and Bertram Bruce
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

Traditional surveys of children's literature have examined features such as text structure and topic, but have failed to take into account rhetorical elements such as author-reader distance, commentary, point of view, and insight into characters' minds. Similarly, they have glossed over aspects of character-to-character interaction such as responses to interpersonal conflict. These "higher-level features" of stories may be what makes stories interesting to read. They are also principal contributors to story complexity, and hence, to difficulty for beginning readers. With regard to both interestingness and complexity, it is important to come to a better understanding of these features.

To concretize our discussion, we first present two examples showing the importance of higher-level text features. Second, we sketch a theory of higher-level story features. Then, we briefly describe how we are applying our analysis to a selection of children's stories. Finally, we discuss some implications of this work.

Strategies for Controlling Hypothesis Formation in Reading

Bertram Bruce and Andee Rubin
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

Reading is a process of forming and evaluating hypotheses to account for the data in a text. Because of its complexity, the task of reading requires strategies for controlling the proliferation of hypotheses. Four of these strategies, (a) jumping to conclusions, (b) maintaining inertia, (c) relying on background knowledge, and (d) working backwards from the goal, are generally effective, but they occasionally create reading problems, rather than alleviating them. Examples from protocols of readers reading a reading test passage are presented. These examples show both the effective use of the strategies and some problems that may arise from their use.

A New Point of View on Children's Stories

Bertram Bruce

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02138

Recent work on text analysis at the Center for the Study of Reading and elsewhere has produced surprising results regarding the texts that children read in school. These results support the hypothesis that part of the difficulty children encounter in making the transition from beginning to skilled readings lies in an abrupt shift in text characteristics between lower and upper elementary school. Moreover, a comparison between school texts and popular trade books shows that the school texts may provide inadequate preparation for the texts that skilled readers need to master. Thus, characteristics of the texts that children are expected to read may hinder rather than help in the attainment of educational goals.

Processing Models for Children's Story Comprehension

Robert de Beaugrande and Genevieve W. Miller

English Department
University of Florida
Gainesville, Florida 32611

It is argued that in abandoning trace-abstraction models of story comprehension in favor of schema-based ones, we have left some issues unresolved in regard to the treatment of bottom-up input. Using a real children's story employed in empiric tests, we review evidence that schemas undergo specification and modification via bottom-up input as the story is progressively read: a phenomenon describable as *procedural attachment*. There seems to be some trace-abstraction during comprehension but in a peripheral role and with unreliable or inconsistent results across a test population. We conclude that accuracy of recall is a less crucial question than the identification of strategies active during recall of both abstractive and constructive nature.

The Pragmatics of Discourse Planning

Robert-Alain de Beaugrande

English Department
University of Florida
Gainesville, Florida 32611

It is argued that the basic notions of natural language pragmatics cannot be the same as those of syntax and semantics as developed so far. Instead, pragmatics must be an empirically oriented theory of action and interaction. The role of sentences and predications is secondary.   .

The most promising approaches for such a pragmatics are: (1) conceptual dependency theory, in which language is a form of actions specified by goal-directed plans (e.g. Schank); (2) plan theory, in which the analysis of tasks and resources leads to the specification of a planned sequence of steps (e.g. Sacerdoti); (3) problem-solving theory, in which points or states in a problem space have to be connected by a successful pathway (e.g. Newell and Simon); and (4) procedural theory of discourse, in which language elements and systems are investigated with respect to how people utilize them in communication and processing.

The paper offers the framework of a natural language pragmatics along these lines and applies the resulting theory to a study of a scene from a stage play by Sidney Howard. It is shown that the actions and discourse actions of the scene are indeed generated by the characters' plans and goals.

Linguistic Theory and Metatheory for a Science of Texts

Robert De Beaugrande

English Department
University of Florida
Gainesville, Florida 32611

This article explores the typical reactions which occur when an established science confronts a new object of inquiry, as we find when linguistic theory encounters the text. The usual discussions are not productive as long as the old "paradigm" is still accepted as the framework for achievement. The issues are therefore re-examined in terms of the metatheory of science (e.g. Sneed, Stegmüller, Lakatos, Feyerabend, Hempel), and some general solutions are expounded for the problems of validating theories on the basis of empirical content. A paradigmatic example is then presented in order to show a possible role for logical linguistics in future theories: a computer grammar that parses text sentences into a progressive network and back again via theorem-proving, with further capacities for applying schemas, answering questions, and generating summaries. This example serves as an application of general design values and criteria for preferring and comparing alternative theories.

Theoretical Foundations of the Automatic Production and Processing of Technical Reports

Robert de Beaugrande

English Department
University of Florida
Gainesville, Florida 32611

The following treatise surveys the issues and approaches for designing a computer system capable of reading, understanding, and writing technical reports. Recent progress in computer science and artificial intelligence research is used to specify the nature of the modules in the system. The processing of a sample text is observed during the phases of reading and writing a report on the origin of sunspots. The author advances some proposals for correlating syntax and semantics of English from a procedural standpoint. The discussion is illustrated with structural diagrams.

Checking for Spelling and Typographical Errors in Computer-Based Text

Thomas N. Turba

Sperry Univac
Language Systems
Roseville Development Center
Roseville, Minnesota 55113

Sigplan Notices 16, 6 (June 1981), 51-60.

This paper addresses the problems and techniques of checking for spelling and typographical errors in computer-based text. To some extent, the paper is a combination of a report of work done by the author and a survey of other work which, although not all used by the author, is of equal value and interest. Some of the material presented is related to other aspects of text processing such as data compaction and the efficient searching of very large dictionaries.

Computer Aids for Writers

Lorinda Cherry

Bell Telephone Laboratories
600 Mountain Avenue
Murray Hill, New Jersey 07974

Sigplan Notices 16, 6 (June 1981), 61-67.

For many people, writing is painful and editing one's own prose is difficult, tedious, and error-prone. It is often hard to see which parts of a document are difficult to read or how to transform a wordy sentence into a more concise one. It is even harder to discover that one overuses a particular linguistic construct. The system of programs described here helps writers to evaluate documents and to produce better written and more readable prose. The system consists of programs to measure surface features of text that are important to good writing style as well as programs to do some of the tedious jobs of a copy editor. Some of the surface features measured are readability, sentence and word length, sentence type, word usage, and sentence openers. The copy editing programs find spelling errors, wordy phrases, bad diction, some punctuation errors, double words, and split infinitives.

Data-Base and Query Systems: New and Simple Ways to Gain Multiple Views of the Patterns in Text

Linda D. Misek-Falkoff

IBM Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, New York 10598

Research Report RC 8769, March 1981, 52 pages.

The goal of this paper is to suggest a spontaneous, "problem solving" approach to textual analysis, in a user-friendly milieu made possible by modern data-base facilities. The study of language is ubiquitous, with interactive systems increasingly supporting studies which previously might have been run in batch mode: the language scholar today is less reliant on mediating technical support. He can enter and query his data directly. He can, but need not, be a "programmer," and can personally control the degree of complexity entailed by his studies, according to his own felt needs and goals. This freedom of inquiry should greatly aid interpretation and inference, since the scholar can more easily alter his methods and critical hypotheses. Interesting connections between language studies at large and the practice of data-base administration may be gleaned, and the growing concord between humanism and technology further enhanced.

A Knowledge Engineering Approach to Natural Language Understanding

Jeannette G. Neal

Department of Computer Science
State University of New York at Buffalo
4226 Ridge Lea Road
Amherst, New York 14226

Technical Report 179, June 1981.

This paper presents the results of preliminary study of a knowledge engineering approach to natural language understanding. The KE system used is the SNePS semantic network processing system (Shapiro, 1979a). As a part of the study, a SNePS front-end system, called the NL-system, was implemented to enable the NLU expert to enter linguistic knowledge into the network in natural language and experiment with this base of knowledge.

This paper discusses the capabilities of the SNePS-based NL-system and illustrates these capabilities by example. The NL-system enables the NLU user-expert to: (a) input his rules about NLU and his lexicon into the semantic network knowledge base in natural language; (b) trace the inference processes, which utilize his rules, in natural language; and (c) employ an experimental rule-based generator to express knowledge built into the network via his NLU rules.

An example is presented to illustrate: the representation of both surface strings and semantic knowledge in the network; the input and representation of rules; how rules build conceptual structures from surface strings; natural language tracing of an NLU process; the use of knowledge for disambiguating an earlier, partially understood sentence; and rule-based generation.

Combining Path-Based and Node-Based Inference In SNEPS

Rohini K. Srihari

Department of Computer Science
State University of New York at Buffalo
4226 Ridge Lea Road
Amherst, New York 14226

Technical Report 183, June 1981, 52 pages.

This paper describes a recent enhancement made to SNePS to include path-based inference. The previous version of SNePS used only node-based inference and was thus limited in its capabilities. Combining path-based inference and node-based inference in SNePS has greatly increased the scope of the system. The paper first motivates the need for such a change and then proceeds to outline the theory as well as the implementation details. Finally, a series of examples is provided which supports the claim that the system is indeed more powerful and efficient now.

Formal Semantics for Time in Databases

James Clifford and David S. Warren

Department of Computer Science
State University of New York at Stony Brook
Stony Brook, New York 11794

Technical Report 81/025, June 1981.

The concept of an *historical database* is introduced as a tool for modelling the dynamic nature of some part of the real world. Just as first-order logic has been shown to be a useful formalism for understanding the underlying semantics of the relational database model, intensional logic is presented as an analogous formalism for understanding the temporal semantics involved in an historical database. The various components of the relational model, as extended to include historical relations, are discussed in terms of the model theory for the logic IL formulated by Richard Montague. The concepts of *intensional* and *extensional* data constraints and queries are introduced and contrasted. Finally, the potential application of these ideas to the problem of natural language database querying is discussed.

More Notes on an Intensional Logic for English: Reversing Extensional Forms

Keith Brian Gallagher

Department of Computer and Communication Sciences
221 Angell Hall
The University of Michigan
Ann Arbor, Michigan 48109

Computer Studies in Formal Ling. N-25, May 1981.

In a series of papers, Joyce Friedman and David Warren gave sufficient conditions for obtaining the extensional forms of words in Richard Montague's *The Proper Treatment of Quantification in Ordinary English* (PTQ). Given these extensionalized forms, one desires to return to the logically equivalent lambda normal form, as a prelude to obtaining the sentence that yielded this form. A procedure for doing this with a proof of its correctness is given.

Sentence Analysis Programs Based on Montague Grammar

Bipin Indurkhya

Philips International Institute of Technological Studies
Eindhoven NETHERLANDS

Masters Thesis, 1981.

The paper investigates computational aspects of the work in theoretical linguistics which is being done in the framework of "Montague Grammar." Grammars in this framework give precise descriptions of the relation between the surface forms of English sentences, their syntactic structures, and their meanings as represented by logical formulas. On the basis of such grammars effective analysis programs for English sentences may be constructed.

Focalizers, the Scoping Problem, and Semantic Interpretation Rules in Logic Grammars

Michael C. McCord

Computer Science Department
University of Kentucky
Lexington, Kentucky 40506

Technical Report 81-81, August 1981.

(To appear in Proceedings of the International Workshop on Logic Programming for Intelligent Systems, August 1981, Logicon, Woodland Hills, Calif.)

This paper deals with a system for semantic interpretation of natural language within the framework of logic programming. Of special interest are a class of grammatical items called *focalizers,* and the problem of determining their scopes in logical form. Focalizers include quantificational determiners, certain adverbs, and abstract items relating to discourse structure.

An Efficient Easily Adaptable System For Interpreting Natural Language Queries

David H.D. Warren

Department of Artificial Intelligence
University of Edinburgh
Forrest Hill
Edinburgh EH1 2QL SCOTLAND

Fernando C.N. Pereira

CAAD Studies, Department of Architecture
University of Edinburgh
Forrest Hill
Edinburgh EH1 2QL SCOTLAND

DAI Research Paper 155, February 1981, 18 pages.

This paper gives an overall account of a prototype natural language question answering system, call Chat-80. Chat-80 has been designed to be both efficient and easily adaptable to a variety of applications. The system is implemented entirely in PROLOG, a programming language based on logic. With the aid of a logic-based grammar formalism called extraposition grammar, Chat-80 translates English questions into the PROLOG subset of logic. The resulting logical expression is then transformed by a planning algorithm into efficient PROLOG, cf. "query optimization" in a relational database. Finally the PROLOG form is executed to yield the answer. On a domain of world geography, most questions within the English subset are answered in well under one second, including relatively complex queries.

Some Problems In Early Noun Phrase Interpretation

C.S. Mellish

Department of Artificial Intelligence
University of Edinburgh
Forrest Hill
Edinburgh EH1 2QL SCOTLAND

DAI Research Paper No. 147, 1980, 5 pages.

How does a piece of text provide the information necessary for generating a symbolic "meaning" and how can a computer program be organized to pick up that information? The work described here aims to investigate some of the constraints on the timing of semantic interpretation. In particular, we are interested in seeing to what extent the meaning can be built up in an incremental way as the analysis proceeds from left to right. We look at some problems of noun phrase interpretation in such a scheme and indicate some representational ideas that help to overcome them. This paper is a brief summary of a forthcoming Ph.D thesis (Mellish 80).

Semantic Long Term Memory and the Understanding of Language

M. Wettler

ISSCO
Universite de Geneve
17 Rue de Candolle
CH-1205 Geneve SWITZERLAND

Working Paper 37, 1978 (in German).

This is a study of the formal representation of conceptual knowledge in relation to the understanding and production of utterances. After the initial historical introduction, the final sections deal with the structure of schemata; inference and consistency of schemata; topic shift in dialogues; and the generation of descriptions of surface structure.

Six Lectures from Recursive Function Theory to Artificial Intelligence

G. Trautteur

ISSCO
Universite de Geneve
17 Rue de Candolle
CH-1205 Geneve SWITZERLAND

Working Paper 38, 1979.

This paper is based on a series of lectures given at ISSCO by Trautteur 1977/78. There are 3 sections. The first introduces the theory of effective procedures, leading up to a statement of Church's Thesis. The second investigates the relationship between language and metalanguage, bringing in the notions of a decision problem and incompleteness. The last section discusses various attempts to formalize the idea of complexity, and this is followed by the presentation of inductive inference, language identification, and dialogue systems from the point of view of recursive function theory. Finally, a discussion of the role of analog machinery in the theory of effective procedures is used to define a position for AI on various philosophical problems.

Some Consideration in Mapping Natural Language Queries on to Database Systems

L. Mazlack

ISSCO
Universite de Geneve
17 Rue de Candolle
CH-1205 Geneve SWITZERLAND

Working Paper 39, 1979.

The aim of this paper is to outline the design of an NL interface for existing database systems. Several examples are discussed where the same database query is expressed by different NL queries. A two-stage design is proposed which incorporates the notion of an intermediate database-independent query language which is then mapped into queries for the particular database under consideration.

Case in Linguistics and Cognitive Science

M. Rosner and H. Somers

ISSCO
Universite de Geneve
17 Rue de Candolle
CH-1205 Geneve SWITZERLAND

Working Paper 40, 1980.

The meaning of "case" as applied to sentences, verbs, events, and event types is quite different. Examples of the use of case in each of these four categories are exemplified and contrasted in detail. An attempt is made to show that comparisons of systems that use case should be avoided when in reality they only have trivial surface characteristics in common. It is concluded that there is no one Case Grammar: in reality the use of case is always relativized to a specific application.

ISSCO*PTOSYS: Brief Description and User Manual

H. Somers

ISSCO
Universite de Geneve
17 Rue de Candolle
CH-1205 Geneve SWITZERLAND

Working Paper 41, 1980.

This paper is a description of PTOSYS (ptotic system), a computer program which builds case frame representations of input English sentences. A central feature of the system is the incremental assembly of a dictionary structure which is built up through an interactive discourse with the user. The system invokes this process whenever it discovers a word it does not recognize, using its linguistic knowledge to guess the most pertinent questions.

The Use of Verb Features in Arriving at a "Meaning Representation"

H. Somers

ISSCO
Universite de Geneve
17 Rue de Candolle
CH-1205 Geneve SWITZERLAND

Working Paper 42, 1981.

A critical examination of the linguistic theory underlying the notion that in trying to map English sentences on to a case representation of deep structure, it is possible to use semantic features attached to the main verb to infer the correct type of case frame.

A Preliminary Study on Linguistic Implications of Resource Control in Natural Language Understanding

J.S. Bien

ISSCO
Universite de Geneve
17 Rue de Candolle
CH-1205 Geneve SWITZERLAND

Working Paper 44, 1980.

The paper presents some hypotheses concerning the organization of language processing by human and computer, which allow us to view a variety of apparently unrelated linguistic phenomena in terms of the sophisticated interactions between a few basic components. In particular, the fact that English articles are rendered in Slavonic languages mainly by word order and vice versa, which has up till now remained completely mysterious, is seen assuming different ways of controlling the depth of nominal phrase processing. Although not yet sustained satisfactorily, the hypothesis offers a general, intuitively appealing approach to natural language research.

Erfahrungen Mit Zwei Natürlich-Sprachlichen Abfragesystemen

Wilfrid Kettler and Arno Schmidt

Technical University of Berlin
Berlin, WEST GERMANY

Magdalena Zoeppritz

IBM Scientific Center
Tiergartenstrasse 15
D-6900 Heidelberg, WEST GERMANY

Technical Report 81.01.001, January 1981, 18 pages.

This paper reports on experiences and results of a joint study between the Technical University of Berlin and the Heidelberg Scientific Center. The purpose of the study was to implement the application "Rooms and Office Space Allocation" of the University Central Administration with the User Specialty Languages System (USL) developed at the Heidelberg Scientific Center and with the Berlin Semantic Oriented Translation System (BEAST) developed at the Technical University of Berlin to provide an on-line natural language facility for the administration, to evaluate the USL-System (not the BEAST-System because the DBMS to which it interfaces was not complete) in an actual work environment, and to study the effects of the different approaches to natural language system design taken in the two systems. The paper outlines the application and describes the BEAST and USL systems. Then follows an account of the experience gained from implementing the application with each of the two systems and results and observations during the implementation and exploration phases with the USL-System.

Proceedings of the Workshop on Data Abstraction, Databases and Conceptual Modelling

Michael L. Brodie
Computer Science Department
University of Maryland
College Park, Maryland 20742

Stephen N. Zilles
IBM Research Laboratory
5600 Cottle Road
San Jose, California 95193

A workshop on data abstraction, databases and conceptual modelling was held June 23-26, 1980, in Pingree Park in the Colorado Rockies. The meeting was intended as a forum in which artificial intelligence, database and programming language researchers could exchange ideas on conceptual modelling of systems of complex data. This proceedings consists of tutorials, edited transcripts of the workshop sessions and position papers prepared by the participants. Tutorial papers on artificial intelligence, databases and programming languages present current research problems, results, goals and terminology. Summaries of the tutorial presentations including questions and answers are also included as they contain additional material and illustrate some of the confusion which arises in discussions among the three research communities. The subject of the workshop is treated in seven topical sessions and a concluding summary session. The transcriptions of these sessions follow the general order of the presentations made at the workshop. The position papers are based on papers submitted for selecting workshop participants, which have been revised to reflect the participants' experience at the workshop.

Computational Strategies for Analyzing the Organization and Use of Information

Donald E. Walker
Artificial Intelligence Center
SRI International
333 Ravenswood Ave.
Menlo Park, California 94025

This chapter describes new developments in computer-based procedures that can improve our understanding of how people organize and use information. Relevant recent research in information science, computational linguistics, and artificial intelligence is reviewed. A program of research is presented that is producing systems that make it possible to study the organization and use of information and, at the same time, provide more effective support for people engaged in those activities. Finally, several current projects that are part of this longer-term program are discussed.

The Command Language Grammar: A Representation for the User Interface of Interactive Computer Systems

Thomas P. Moran
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, California 94304

This article introduces and discusses a specific grammatical structure — the Command Language Grammar (CLG) — as a representational framework for describing the user interface aspects of interactive computer systems. CLG partitions a system into a Conceptual Component (tasks and abstract concepts), a Communication Component (command language), and a Physical Component (display, keyboard, etc.). The components are further stratified into distinct Levels — a Task Level, a Semantic Level, a Syntactic Level, and an Interaction Level — each Level being a complete description of the system at its level of abstraction. Each Level's description contains procedures for accomplishing the tasks addressed by the system in terms of the actions available at that Level. That is, the system is described by progressive refinement. An extensive example, a small message-processing system, is described at all Levels in the CLG notation.

CLG is discussed from three points of view: The Linguistic View sees CLG as elaborating the structure of the system's user interface and of the communication between the user and the system. The principal goal of CLG in this view is to lay out the space of command language systems. The Psychological View sees CLG as describing the user's mental model of the system. The main concern in this view is with the psychological validity of the CLG description. The Design View sees CLG as a series of representations for specifying the design of a system. CLG proposes a topdown design process in which the conceptual model of the system is first specified and then a command language is created to communicate with it.

A Experiment in English-Spanish Automated Translation of Medical Language Data

Isabel Garcia-Hidalgo and George S. Dunham
Laboratory of Statistical and Mathematical Methodology
Building 12A, Room 3041
National Institutes of Health
Bethesda, Maryland 20205

An English-to-Spanish translation procedure and its associated dictionaries were developed and implemented for the 1,426 terms of the morphology section of

the International Classification of Diseases for Oncology. Morphological substitutions and respelling rules permit translation of most of the ICD-O vocabulary composed of Latin and Greek terms, which are cognate in the source and target languages, without the construction of a large word lexicon.

A fairly simple classification of words, which could be implemented by recognition of terminal morphemes, and which classifies them both syntactically and semantically, served as an adequate basis for translation, and sheds some light on the linguistic structure of this type of complex noun phrase seen universally in medical writings and communications. A set of 17 reduction-transformation rules based on the word classification provide syntactic control of the translation process.

A Method for Rapidly Applying Context Sensitive Phonological Rules

Robert L. Mercer and Paul S. Cohen
IBM Thomas J. Watson Research Center
P.O. Box 218
Yorktown Heights, New York 10598

IBM Research Report RC 8889, June 1981, 25 pages.

The application of phonological rules to phonemic strings to create phonetic graphs is a time-consuming process. Since many such graphs must be constructed during the decoding phase of automatic speech recognition, it is valuable to be able to rapidly construct phonetic graphs for strings of words from phonetic graphs for the individual words in the string. However, because many phonological rules operate across word boundaries or require interword context, it is not possible to determine a unique phonetic graph for a word independent of the context in which it occurs. This report describes a method for determining the phonetic graph for a word in isolation together with auxiliary information to allow phonetic graphs for different words to be rapidly interconnected to form a phonetic graph for a string of words.

A Model for Automated Phonemicization

M. Boot, E. Maat, and J. Renkers
Instituut Voor Toegepaste Taalkunde
en Computerlinguistiek
Wilhelminapark 11
3581 NC Utrecht

Research Report ISBN 90-6244-512-8, 1980, 91 pages.

In this report a computer program is designed that should be able to perform a transcription of written text into the phonematic format according to the principles of phonematic transcription. This report focuses on the design of the program and answers questions concerning the relation between the technical part (implementation) and the linguistic considerations behind this computer program. It is argued that in the past a computer program performing this linguistic task was principally designed from the implementation point of view. This has led to a computer program which has a strong ad hoc kind of problem solving part of it. Therefore, this computer program turns out to be not adaptable to new situations and unforeseen mistakes. In this paper it is argued that the implementation of linguistic problem solving algorithms itself is a secondary problem and it is also argued that the definition of the task as well as the kind of system being created should be founded in intelligent linguistic considerations. "Definitions" and "system" are summarized with the notion "design" in this report. Thus, in the report, it is argued that the design of computer programs for linguistic operations should be principally based on intelligent linguistic considerations. This is the fundamental difference between other computer programs proposed in the past for the phonemicization of Dutch. It is argued that such a design is more prosperous as far as the results of the phonemicization program are concerned. The principals of this type of design are applied to the concrete problem of transcription of written Dutch in a phonematic format. First results of the program FONGRAF are documented and discussed in this report.

FONGRAF: A Computer Program for Automated Phonemicization

J. Renkers and M. Boot
Instituut Voor Toegepaste Taalkunde
en Computerlinguistiek
Wilhelminapark 11
3581 NC Utrecht

Research Report ISBN 90-6244-513-6, 1980, 59 pages.

This report provides a listing of the PL/I program for automated phonemicization described in the abstract above.