

Other Conferences

The *Fifth Conference on Automated Deduction*, organized by INRIA, will be held in Les Arcs, France, July 8-11, 1980. For further information contact:

INRIA
Service des Relations Extérieures
Domaine de Voluceau - Rocquencourt
B.P. 105
78150 Le Chesnay, FRANCE

A workshop on *Logic Programming* and closely related topics will be held July 14-16 in Budapest, sponsored by the von Neumann computer science society. [See *AJCL* 6,1, pg. 53.] For further information contact:

Sten-Ake Tärnlund
Department of Computer Science
University of Stockholm
106 91 Stockholm, SWEDEN

The Institute of Informatics of Warsaw University is organizing the *First International Workshop on Natural Communication with Computers*, to take place September 9-12, 1980, in Warsaw, Poland. [See *AJCL* 6,1, pg. 52.] Further information can be obtained from the Workshop Secretariat:

Miss Ludmila Rózańska
Institute of Informatics
Warsaw University
PKiN pok. 850
00-950 Warsaw, POLAND

IFIP Congress '80, sponsored by IFIP, will be held in Kyoto, Japan, and Melbourne, Australia, October 6-17, 1980. For further information contact:

IFIP Congress '80
G.P.O. Box 880G
Melbourne, Victoria
AUSTRALIA 3001

An *International Congress on Applied Systems Research and Cybernetics*, sponsored by The School of Computer Science, University of Windsor, and four societies, will be held in Acapulco, Mexico, December 12-15, 1980. The main theme of the Congress is "The quality of life and how to improve it." The Congress will include a session on Computers and the Humanities. For further information contact:

Dr. George E. Lasker
Congress President
School of Computer Science
University of Windsor
Windsor, Ontario N9B 3P4 CANADA

The 1981 *Office Automation Conference*, sponsored by AFIPS, will be held in Houston, Texas, March 23-25, 1981. For further information contact:

Ms. Carol Sturgeon
AFIPS - OAC
Suite 800
1815 North Lynn Street
Arlington, Virginia 22209

The *Fifth International Conference on Computers and the Humanities*, sponsored by the Association for Computers and the Humanities, will be held in Ann Arbor, Michigan, in May 1981. For further information contact:

Professor Joseph Raben
Computers and the Humanities
Queens College, CUNY
Flushing, New York 11367

Abstracts of Current* Literature
Making Preferences More Active

Yorick Wilks

Department of Language and Linguistics
University of Essex
Wivenhoe Park
Colchester CO4 3SQ ENGLAND

Artificial Intelligence 11 (1978), 197-223.

The paper discusses the incorporation of richer semantic structures into the Preference Semantics system: they are called *pseudo-texts* and capture something of the information expressed in one type of frame proposed by Minsky (q.v.). However, they are in a format, and subject to rules of inference, consistent with earlier accounts of this system of language analysis and understanding. Their use is discussed in connection with the phenomenon of *extended* use: sentences where the semantic preferences are broken. It is argued that such situations are the norm and not the exception in normal language use, and that a language understanding system must give some *general* treatment of them. A notion of *sense projection* is proposed, leading on to an alteration of semantic formulas (word sense representations) in the face of unexpected context by drawing information from the pseudo texts. A possible implementation is described, based on a new semantic parser for the Preference Semantics system, which would cope with extended use by the methods suggested and answer questions *about the process of analysis itself*. It is argued that this would be a good context in which to place a lan-

* Editor's note: The abstracts in this issue are not as "current" as I would like, but, because we did not publish the *Journal* in 1979, there is still much material to get caught up on.

guage understander (rather than that of question-answering about a limited area of the real world, as is normal) and, moreover, that the sense projection mechanisms suggested would provide a test-bed on which the usefulness of frames for language understanding could be realistically assessed.

On the Use of Framed Knowledge in Language Comprehension

Eugene Charniak

Department of Computer Science
Brown University
Providence, Rhode Island 02912

Artificial Intelligence 11 (1978), 225-265.

Notions like "frames," "scripts," etc. are now being used in programs to understand connected discourse. We will describe a program in this vein which understands simple stories about painting. (Jack was painting a chair. He dipped a brush into some paint. Q:Why?) In particular, problems of matching, read time inference, and undoing false conclusions will be stressed. The program makes heavy use of real world knowledge, and there is an extensive discussion of various issues in knowledge representation and how they affect frame representations: modularity, the need for problem solving, worldly vs control knowledge, and cleanliness. The paper concludes with an extensive discussion of the program's shortcomings.

An Artificial Intelligence Approach to Language Instruction

Ralph M. Weischedel

Department of Computer and Information Sciences
University of Delaware
Newark, Delaware 19711

Wilfried M. Voge

Department of German
University of California
Irvine, California 92717

Mark James

School of Social Science
University of California
Irvine, California 92717

Artificial Intelligence 10 (1978), 225-240.

This paper describes an implemented, prototype system for a sophisticated, intelligent tutor for instruction in a foreign language. The system is an application of artificial intelligence research in natural language, but it implements several ideas that depart from standard approaches to natural language understanding. For instance, the semantic analyzer diagnoses several kinds of comprehension problems and semantic errors that a student might make. Some fine distinctions in meaning are represented to detect misuse of words. Not only is a model of good syntax included in the tutor, but also a model of incorrect forms, rich

enough to pinpoint specific syntactic mistakes. Finding the intended interpretation is complicated by the likelihood of student errors. Therefore, perfect syntactic form is not necessary for semantic analysis of the student's input. The problems discussed and solutions presented are closely related to the more general problem of how to respond to a natural language input that surpasses the computer's model of language or of context.

A Critical Perspective on KRL

Wendy Lehnert

Department of Computer Science
Yale University
New Haven, Connecticut 06520

Yorick Wilks

Department of Language and Linguistics
University of Essex
Wivenhoe Park
Colchester CO4 3SQ ENGLAND

Cognitive Science 3 (1979), 1-28.

Bobrow and Winograd have presented to the AI community two descriptions of KRL (Bobrow & Winograd, 1977, Bobrow, Winograd et al., 1977) which explicate both a high level AI programming language and a theory of knowledge representation. In actual practice, the line between these roles is necessarily vague. As is the case with all programming languages, commitments made to specific data formats or control structures profoundly affect design decisions made by the user. In KRL, there are additional commitments to knowledge representation in the programming language as well. While these commitments are neutrally presented as convenient features of high level language, their impact on the user would be far less neutral. To a user who has not previously investigated problems of knowledge representation first-hand, KRL either suggests a particular approach or imposes that same approach. In either case, the user is liable to be unconscious of the continual trade-off between low level design options and high level programming convenience.

KRL Another Perspective

Daniel G. Bobrow

Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, California 94304

Terry Winograd

Computer Science Department
Stanford University
Stanford, California 94305

Cognitive Science 3 (1979), 29-42.

Wendy Lehnert and Yorick Wilks (pp. 1-28 of this issue of *Cognitive Science*) have written a lengthy paper raising a number of issues concerning KRL. [See

abstract above.] Much of their paper is an excellent explanation of some of the features and problems of KRL, and will serve to clarify things which we have explained poorly or not at all in previous papers. Other parts of what they say we find more contentious, and much of this response will be an argument against views of theirs which we feel are confused or wrong. The decision to focus on the disputes does not imply a general rejection of the paper. It was clearly intended in the spirit of constructive criticism and makes a number of valid and important points. We feel it is useful to write a response, not as a defense, but as a further step in a dialog through which we will all come to a better understanding of language and cognition.

Narrative Models of Action and Interaction

Robert De Beaugrande

University of Florida

Benjamin N. Colby

University of California
Irvine, California 92717

Cognitive Science 3 (1979), 43-66.

This paper explores some issues which a humanlike story system ought to encompass, but which are usually not in the main focus of narrative models since Propp. We argue that knowledge about actions and interactions can account not only for how stories are constructed, but also for why some stories are more *interesting* and *enduring* than others. We analyze a traditional English folktale in these terms, and show how classes of surface expressions make recoverable the underlying structures and dependencies that schema-based understanders comprise.

Coherence and Coreference

Jerry R. Hobbs

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Cognitive Science 3 (1979), 67-90.

Coherence in conversations and in texts can be partially characterized by a set of coherence relations, motivated ultimately by the speaker's or writer's need to be understood. In this paper, formal definitions are given for several coherence relations, based on the operations of an inference system; that is, the relations between successive portions of a discourse are characterized in terms of the inferences that can be drawn from each. In analyzing a discourse, it is frequently the case that we would recognize it as coherent, in that it would satisfy the formal definition of some coherence relation, if only we could assume certain noun phrases to be coreferential. In such cases, we will simply assume the identity of the entities referred to, in what might be called a "petty conversational

implicature," thereby solving the coherence and coreference problems simultaneously. Three examples of different kinds of reference problems are presented. In each, it is shown how the coherence of the discourse can be recognized, and how the reference problems are solved, almost as a by-product, by means of these petty conversational implicatures.

Handling Complex Queries in a Distributed Data Base

Robert C. Moore

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

Technical Note 170, Oct. 1979.

As part of the continuing development of the LADDER system, we have substantially expanded the capabilities of the data base access component that serves as the interface between the natural-language front end of LADDER and the data base management systems on which the data is actually stored. SODA, the new data base access component, goes beyond its predecessor IDA, in that it accepts a wider range of queries and accesses multiple DBMSs. This paper is concerned with the first of these areas, and discusses how the expressive power of the query language was increased, how these changes affected query processing in a distributed data base, as well as what are some limitations of and planned extensions to the current system.

Computational Models of Beliefs and the Semantics of Belief-Sentences

Robert C. Moore and Gary G. Hendrix

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

Technical Note 187, June 1979.

This paper considers a number of problems in the semantics of belief sentences from the perspective of computational models of the psychology of belief. We present a semantic interpretation for belief sentences and show how this interpretation overcomes some of the difficulties of alternative approaches, especially those based on possible-world semantics. Finally, we argue that these difficulties arise from a mistaken attempt to identify the truth conditions of a sentence with what a competent speaker knows about the meaning of a sentence.

Focusing and Description in Natural Language Dialogues

Barbara J. Grosz

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

Technical Note 185, April 1979.

When two people talk, they focus their attention on only a small portion of what each of them knows or believes. Both what is said and how it is interpreted depend on a shared understanding of this narrowing of attention to a small highlighted portion of what is known.

Focusing is an active process. As a dialogue progresses, the participants continually shift their focus and thus form an evolving context against which utterances are produced and understood. A speaker provides a hearer with clues of what to look at and how to look at it - what to focus on, how to focus on it, and how wide or narrow the focusing should be. As a result, one of the effects of understanding an utterance is that the listener becomes focused on certain entities (both objects and relationships) from a particular perspective.

Focusing clues may be linguistic or they may come from knowledge about the relationships between entities in the domain. Linguistic clues may be either explicit, deriving directly from certain words, or implicit, deriving from sentential structure and from rhetorical relationships between sentences.

This paper examines the relationships between focusing and definite descriptions in dialogue and its implications for natural language processing systems. It describes focusing mechanisms based on domain-structure clues which have been included in a computer system and, from this perspective, indicates future research problems entailed in modeling the focusing process more generally.

Utterance and Objectives: Issues in Natural Language Communication

Barbara J. Grosz

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

Technical Note 188, June 1979.

Communication in natural language requires a combination of language-specific and general common-sense reasoning capabilities, the ability to represent and reason about the beliefs, goals, and plans of multiple agents, and the recognition that utterances are multifaceted. This paper evaluates the capabilities of natural language processing systems against these requirements and identifies crucial areas for future re-

search in language processing, common-sense reasoning, and their coordination.

A Framework for a Portable Natural Language Interface to Large Data Bases

Kurt G. Konolige

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

Technical Note 197, Oct. 1979.

A framework is proposed for developing a portable natural language interface to large data bases. A discussion of problems arising from portability leads to the identification of a key concept of the framework: a conceptual schema for representing a user's model of the domain as distinct from the data base schema. The notions of conceptual completeness and linguistic coverage are shown to be natural consequences of this framework. An implementation of the framework, called D-LADDER, is presented, and some preliminary performance results reported.

Theoretical Foundations of Linguistics and Automatic Text Processing

Jane J. Robinson

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

Technical Note 199, Oct. 1979.

Texts are viewed as purposeful transactions whose interpretation requires inferences based on extralinguistic as well as on linguistic information. Text processors are viewed as systems that model both a theory of text and a theory of information processing. The interdisciplinary research required to design such systems has a common center, conceptually, in the development of new kinds of lexical information, since words are not only linguistic objects, they are also psychological objects that evoke experiences from which meanings can be inferred. Recent developments in linguistic theory seem likely to promote more fruitful cooperation and integration of linguistic research with research on text processing.

DIAGRAM: A Grammar for Dialogues

Jane J. Robinson

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

Technical Note 205, Feb. 1980.

This paper presents an explanatory overview of a large and complex grammar, DIAGRAM, that is used in a computer system for interpreting English dialogue.

DIAGRAM analyzes all of the basic kinds of phrases and sentences and many quite complex ones as well. It is not tied to a particular domain of application, and it can be extended to analyze additional constructions, using the formalism in which it is currently written. For every expression it analyzes, DIAGRAM provides an annotated description of the structural relations holding among its constituents. The annotations provide important information for other parts of the system that interpret the expression in the context of a dialogue.

DIAGRAM is an augmented phrase structure grammar. Its rule procedures allow phrases to inherit attributes from their constituents and to acquire attributes from the larger phrases in which they themselves are constituents. Consequently, when these attributes are used to set context-sensitive constraints on the acceptance of an analysis, the contextual constraints can be imposed by conditions on dominance as well as conditions on constituency. Rule procedures can also assign scores to an analysis, rating some applications of a rule as probable or as unlikely. Less likely analyses can be ignored by the procedures that interpret the utterance.

In assigning categories and writing the rule statements and procedures for DIAGRAM, decisions were guided by consideration of the functions that phrases serve in communication as well as by considerations of efficiency in relating syntactic analyses to propositional content. The major decisions are explained and illustrated with examples of the rules and the analyses they provide. Some contrasts with transformational grammars are pointed out and problems that motivate a plan to use redundancy rules in the future are discussed. (Redundancy rules are meta-rules that derive new constituent-structure rules from a set of base rules, thereby achieving generality of syntactic statement without having to perform transformations on syntactic analyses.) Other extensions of both grammar and formalism are projected in the concluding section. Appendices provide details and samples of the lexicon, the rule statements, and the procedures, as well as analyses for several sentences that differ in type and structure.

Interpreting Natural-Language Utterances in Dialogs About Tasks

Ann E. Robinson, Douglas E. Appelt, Barbara J. Grosz, Gary G. Hendrix, and Jane J. Robinson
 Artificial Intelligence Center
 SRI International
 333 Ravenswood Avenue
 Menlo Park, California 94025

Technical Note 210, March 1980.

This paper describes the results of a three-year research effort investigating the knowledge and proc-

esses needed for participation in natural-language dialogs about ongoing mechanical-assembly tasks. Major concerns were the ability to interpret and respond to utterances within the dynamic environment effected by progress in the task, as well as by the concomitant shifting dialog context.

The research strategy followed was to determine the kinds of knowledge needed, to define formalisms for encoding them and procedures for reasoning with them, to implement those formalisms and procedures in a computer system called TDUS, and then to test them by exercising the system.

Principal accomplishments include: development of a framework for encoding knowledge about linguistic processes; encoding of a grammar for recognizing many of the syntactic structures of English; development of the concept of "focusing," which clarifies a major role of context; development of a formalism for representing knowledge about processes, and procedures for reasoning about them; development of an overall framework for describing how different types of knowledge interact in the communication process; development of a computer system that not only demonstrates the feasibility of the various formalisms and procedures, but also provides a research tool for testing new hypotheses about the communication process.

A Plan-Based Approach to Speech Act Recognition

James F. Allen

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

Technical Report No. 131/79, Feb. 1979.

This thesis concerns how people use language to communicate. While the literal meaning and syntactic structure of a sentence tell us what the speaker said, often they do not capture what was actually meant. For instance, when we hear "Can you pass the salt?" at the dinner table, why do we take this to be a request to pass the salt rather than a query about our salt-passing ability? The answer, in part, is that understanding what was meant requires recognizing what the speaker intended, i.e. what he hopes to achieve. Thus, we consider "Can you pass the salt?" to be a request because we recognize that the speaker said it intending to get us to pass the salt.

In addition, people often communicate using only short phrases or even single words. For instance, were you to walk into my office, I could request that you close the door by simply saying "the door".

This thesis examines the language comprehension process using a model of goal-directed behaviour. Utterances are considered to be instances of speech acts, which are actions that can be reasoned about in a

planning system. A large part of a hearer's understanding of an utterance involves recognizing the plan (of the speaker) that produced it. This plan may contain *obstacles*, goals that the speaker cannot achieve without assistance. The responses are formed on the basis of these obstacles.

Speech acts are defined in terms of the plan that the hearer believes the speaker intended him to recognize. Speech act identification is then accomplished by the plan recognition process. This method accounts for many sentences that are not intended with their literal meaning (the indirect speech acts). In particular, this thesis concentrates on the indirect forms of requests and informs.

There are other linguistic phenomena that can be explained by this model and that are considered in the thesis. Answers that provide more information than explicitly asked for often arise from the method of basing responses on the obstacles in the speaker's plan. Some sentence fragments can identify the speaker's plan and obstacles. Thus, these can be answered appropriately without ever having to construct a syntactically complete sentence from the fragment. This work also extends naturally to explain the generation and participation in subdialogues whose purpose is to clarify or correct previous utterances.

Finally, it examines possible extensions to this model. In particular, it considers the generation (the planning) of indirect forms of speech acts and discusses how this work relates to the other classes of speech acts.

Dialogue Games

W.C. Mann

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

Research Report RR-79-77, Nov. 1979.

Natural dialogue does not proceed haphazardly; it has an easily recognized "episodic" structure and coherence that conform to a well developed set of conventions. This report represents these conventions formally in terms related to speech act theory and to a theory of action. The major formal unit, the Dialogue Game, specifies aspects of the communication of both participants in a dialogue. We define the formal notion of Dialogue Games and describe some of the important Games of English. Dialogue Games are conventions of interactive goal pursuit. Using them, each participant pursues his own goal in a way that sometimes serves the goals of the other. The idea of Dialogue Games can thus be seen as part of a broader theoretical perspective characterizing virtually all communication as goal pursuit activity. We also define and exemplify the property of Motivational Coherence

of dialogues. Motivational coherence can be used as an interpretive principal in explaining language comprehension. Actual dialogue games have a kind of casual connectedness that is not a consequence of their formal properties. This is explained in terms of a theory of action, which is also seen to explain a similar attribute of speech acts.

Computer as Author - Prospects and Results

William C. Mann and James A. Moore

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

Research Report RR-79-82, Jan. 1980.

For a computer program to be able to compose text is interesting both intellectually and practically. Artificial Intelligence research has only recently begun to address the task of creating coherent texts containing more than one sentence.

Our recent research has produced a new paradigm for organizing and expressing information in text. This paradigm, called Fragment-and-Compose, has been used in a pilot project to create texts from semantic nets. The method involves dividing the given body of information into many small propositional units, and then combining these units into smooth coherent text. So far the largest example written by Fragment-and-Compose has been two paragraphs of instruction about what a computer operator should do in case of indications of a fire.

This report describes the text generation problem and anticipates a specific way to disseminate and use technical developments. It presents the research that led to creation of Fragment-and-Compose, including the largest example of computer-produced text. It also discusses the immediate problems and difficulties of elaborating Fragment-and-Compose into a general and powerful method.

ATNs and the Semantic/Pragmatic Control of the Analysis and Generation of Natural Language

W. Wahlster

Universität Hamburg
Germanisches Seminar
Von-Melle-Park 6
D-2000 Hamburg 13 WEST GERMANY

HAM-RPM Report No. 11, March 1979, (In German).

This paper discusses adaptations of the ATN formalism for semantic and pragmatic processing in natural language understanding and generation. First, the various uses of ATNs are classified and a set of definitions of terms like 'semantically guided parsing', 'semantic grammar' and 'pragmatic grammar' is proposed. The use of ATNs for semantic and pragmatic

processing is investigated by discussing the generation of definite noun phrases. It is shown that when we try to describe the generation of noun phrases as a cognitive process whose results enable the listener to identify an object intended by the speaker - as it is done in the dialogue system HAM-RPM - we are confronted with several limitations of the ATN formalism. Finally, ten critical theses are formulated which summarize the strengths and limitations of ATNs within the framework of semantic and pragmatic processing.

The Anatomy of the Natural Language Dialogue System HAM-RPM

W. v. Hahn, W. Hoepfner, A. Jameson, W. Wahlster

Universität Hamburg
Germanisches Seminar
Von-Melle-Park 6
D-2000 Hamburg 13 WEST GERMANY

HAM-RPM Report No. 12, May 1979.

HAM-RPM is a dialogue system which converses with a human partner in colloquial German about limited, but interchangeable, scenes. The objective of this report is to give a detailed, complete and self-contained description of the system in its present state of implementation. After a discussion of the goals and methodological principles which guide our research and a short introduction to the implementation language, an overview of the system's architecture, of its knowledge base and of the domains of discourse is given. Then each processing phase from the analysis of natural-language input to the generation of a natural-language utterance is described in detail. The examples used during these descriptions are supplemented by transcripts of complete dialogue sessions. Finally HAM-RPM's programming environment is described.

Cooperative Dialogue Behaviour in the Natural Language System HAM-RPM.

W. Hoepfner, A. Jameson

Universität Hamburg
Germanisches Seminar
Von-Melle-Park 6
D-2000 Hamburg 13 WEST GERMANY

HAM-RPM Report No. 13, April 1979, (In German).

This paper describes certain cooperative features of the natural language dialogue system HAM-RPM. Two of the system's components employed in analysing the natural partner's input are discussed with particular attention to cooperative dialogue behaviour. The realization of such behaviour in response to unknown forms encountered during lexical analysis presents several sequentialization problems. The second component - NP-resolution - provides for clarification dialogues in natural language and verbalizations of time-consuming processing steps ('Thinking aloud').

Implementing Fuzziness in Dialogue Systems

Wolfgang Wahlster

Universität Hamburg
Germanisches Seminar
Von-Melle-Park 6
D-2000 Hamburg 13 WEST GERMANY

HAM-RPM Report No. 14, Nov. 1979.

Techniques for dealing with fuzziness are described. First, the components of a natural language dialogue system in which fuzziness can figure importantly are enumerated and the communicative and cognitive functions of vagueness are surveyed. Second, some techniques employed within the dialogue system HAM-RPM for handling fuzziness during the analysis and generation phases are illustrated. Finally, a new model of fuzzy reasoning is presented which is based on a many-sorted fuzzy logic and includes a corroboration procedure for multiple derivations.

Natural Language Processing and Information Retrieval The Problem of Complexity

Jürgen Krause

IBM Scientific Center
Tiergartenstrasse 15
6900 Heidelberg WEST GERMANY

Technical Report 79.04.002, April 1979.

It is argued that research in powerful grammar structures opens one possible way to solve the problem of growing complexity in natural language processing for endusers. In automatic indexing there are first indications that analysis can be reduced to "powerful grammar structures", which are seen as function of statistical frequency, simplicity of algorithmic solution and high information value. In the area of query languages the idea of using a subset of language proves feasible. This was shown by evaluating USL, a natural language based question answering system.

Preliminary Results of a User Study with the 'User Specialty Languages' System and Consequences for the Architecture of Natural Language Interfaces

Jürgen Krause

IBM Scientific Center
Tiergartenstrasse 15
6900 Heidelberg WEST GERMANY

Technical Report 79.04.003, May 1979.

The question is asked whether the concept of using subsets of natural languages as query languages for data bases turns out to be feasible in actual applications using the question answering system 'USER SPECIALTY LANGUAGES' (USL). Methods of evaluating a natural language based information system are discussed. The results (error and language structure evaluation) give first indications of how to form the general architecture of application systems which use a subset of German as query language.

Natural Language Systems: A User's View

John M. Morris

4 Proctor Ave.
Clinton, New York 13323*SISTM Quarterly II, 3 (Spring 1979), 16-20.*

Labelling a data base management system as "natural language" is a little like labelling a package "natural food": the label doesn't tell us much about the contents. For the purposes of this paper, I want to adopt an absolutely minimal definition of "natural language," and then go on to what seems to me to be the more important question for the user: How well does it work? How much help will it give me in solving my problems?

I want to suggest several criteria for comparison, evaluation, and selection of natural language query systems for use with structured data bases. The criteria should also be useful in comparing natural language systems with artificial languages in similar applications. More importantly, I want to suggest that there are "natural language" systems which are working, and which are capable of serving useful purposes, and that it is perhaps best to look at them from the point of view of the practical user, rather than from that of someone trying to simulate the full range of human cognition.

The Question Answering System PHLIQA1W.J.H.J. Bronnenberg, H.C. Bunt, S.P.J. Landsbergen,
R.J.H. Scha, W.J. Schoenmakers, E.P.C. van UtterenPhilips Research Laboratories
Eindhoven, THE NETHERLANDS*Report M.S. 10.933, 1979.*

PHLIQA1 is an experimental computer program which answers English questions about a limited subject domain. The design of the system is based on the distinction between different, successively "deeper" levels of semantic analysis. At each of these levels, a different formal language is used to express the meaning of a question. The languages have the same syntactic and semantic structure, but they differ in the constants they contain. The most important representation languages are: (1) the English-oriented Formal Language, with constants that correspond to the "content words" of English, (2) the World Model Language, with constants that represent the primitive concepts that characterize the subject domain of the system, and (3) the Data Base Language, with constants that represent the "record types", "attributes" and "sets" that constitute the data base. The system consists of a series of modules which translate a question from one level to the next lower level, until it finally results in an answer. The paper describes the overall structure of the system, the formal languages used for semantic representation, and the translation

steps between the different levels. It also shows how the multilevel semantics set-up makes it possible to treat ambiguous and unanswerable questions in a careful way.

An Anatomy of Graceful Interaction in Spoken and Written Man-Machine Communication

Phil Hayes and Raj Reddy

Department of Computer Science
Carnegie-Mellon University
Schenley Park
Pittsburgh, Pennsylvania 15213*Technical Report CMU-CS-79-144, Sept. 1979.*

There have recently been a number of attempts to provide natural and flexible interfaces to computer systems through the medium of natural language. While such interfaces typically perform well in response to straightforward requests and questions within their domain of discourse, they often fail to interact gracefully with their users in less predictable circumstances. Most current systems cannot, for instance: respond reasonably to input not conforming to a rigid grammar; ask for and understand clarification if their user's input is unclear; offer clarification of their own output if the user asks for it; or interact to resolve any ambiguities that may arise when the user attempts to describe things to the system.

We believe that graceful interaction in these and the many other contingencies that can arise in human conversation is essential if interfaces are ever to appear cooperative and helpful, and hence be suitable for the casual or naive user, and more habitable for the experienced user. In this paper, we attempt to circumscribe graceful interaction as a field for study, and identify the problems involved in achieving it.

To this end we decompose graceful interaction into a number of relatively independent skills: skills involved in parsing elliptical, fragmented, and otherwise ungrammatical input; in ensuring robust communication; in explaining abilities and limitations, actions and the motives behind them; in keeping track of the focus of attention of a dialogue; in identifying things from descriptions, even if ambiguous or unsatisfiable; and in describing things in terms appropriate for the context. We claim these skills are necessary for any type of graceful interaction and sufficient for graceful interaction in a certain large class of application domains. None of these components is individually much beyond the current state of the art, and we outline the architecture of a system that integrates them all. Thus, we propose graceful interaction as an idea of great practical utility whose time has come and which is ripe for implementation. We are currently implementing a gracefully interacting system along the lines presented; the system will initially deal with typed input, but is eventually intended to accept speech.

A Semantic Network of Production Rules in a System for Describing Computer Structures

Michael D. Rychener

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

Technical Report CMU-CS-79-130, June 1979.

A novel implementation of the basic mechanisms of a semantic network is presented. This constitutes a merging, in terms of the underlying language architecture, of a powerful problem-solving mechanism, production-rule systems, with a proven representation formalism. Details are presented on the most basic aspects of the network, namely on representing nodes and on mechanisms for their access. Commands for definition, modification, and search-based displays of network information are discussed. The relations of the network are divided into six groups: taxonomic, structural, functional, descriptive, means-ends, and physical. The "further specification" relation is put forward as an improvement over concepts such as "ISA", superset, instantiation, individuation, and the type-token distinction. The importance of uniformly representing methods and network as rules and the importance of distinguishing temporary from permanent states are discussed. Since the system is rule-based, it includes a simple but powerful augmentation capability, embodied in a language for expressing methods. Though evidence is not provided here for advanced semantic net capabilities, there is sketched a production system position on a number of relevant issues for those capabilities. The domain of application is the symbolic description and manipulation of computer structures at the PMS (processor-memory-switch) level. The system will ultimately be used for computer-aided design activities.

Computer Models of Human Personality Traits

Jaime G. Carbonell

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

Technical Report CMU-CS-79-154, Nov. 1979.

A goal-based analysis of human personality traits is presented with the objective of developing a comprehensive simulation model. It is shown that understanding trait attributions is an integral part of story understanding and therefore much of natural language processing. The model of personality traits is derived from the goal trees of the POLITICS system, the notion of social prototypes, and planning/counterplanning strategies.

An Information Retrieval System Based on a Computer Model of Legal Knowledge

Carole D. Hafner

Computer Science Department
General Motors Research Labs.
Warren, Michigan 48090

University of Michigan Ph.D. thesis, 1978.

A document retrieval system that has knowledge of the subject matter of its data base is used to explore the relationship between memory structures and information retrieval processes. Six semantic relations are used to define a semantic network model of legal knowledge in the area of Negotiable Instruments Law. These are: set/member relations, superclass/subclass relations, constituent relations, property relations, role relations, and event-condition relations. A formal language is developed that allows the user to create complex relational expressions describing situations of interest to lawyers. A data base of 400 cases and statutes is defined using this language, called the Situation Description Language (SDL), to describe the situations that are discussed in each document.

The retrieval system accepts queries in the Situation Description Language, and uses its knowledge of legal relationships to determine when a SDL expression that has been entered as part of a document descriptor satisfies another SDL expression that has been entered as a query. Some logical and semantic relationships between sets of SDL expressions are explored in the context of different types of queries, and the realization of these relationships is demonstrated. The system is implemented in LISP on the Michigan Terminal System.

Representation of Knowledge in a Legal Information Retrieval System

Carole D. Hafner

Computer Science Department
General Motors Research Labs.
Warren, Michigan 48090

Technical Report, 1979.

The Legal Research System is a knowledge-based computer system for legal information retrieval. The system uses a semantic model of legal knowledge to understand and interpret user queries, extending them to include terms that are implied, but not mentioned by the user. This paper analyzes some of the reasons why subject area knowledge is needed in an information retrieval system, describes how legal knowledge is encoded in the Legal Research System, and gives examples of the kind of inference the system can perform.

Retrieving Information from an Episodic Memory or Why Computers' Memories Should be More Like People's

Roger C. Schank and Janet Kolodner

Department of Computer Science
Yale University
New Haven, Connecticut 06520

Research Report 159, Jan. 1979.

The problem of how to organize particular experiences in a long-term memory has been largely neglected in Natural Language Understanding research. This paper addresses that problem. It shows why good memory organization is necessary for doing intelligent tasks such as story understanding and conversation. A computer memory organization modeled after human memory is proposed, as well as strategies for accessing the computer memory (again based on human information retrieval). This proposal can also be viewed as a new approach to the organization of intelligent databases. The CYRUS system, a computer memory model which implements this theory of memory organization, is described.

Problems in Conceptual Analysis of Natural Language

Lawrence Birnbaum and Mallory Selfridge

Department of Computer Science
Yale University
New Haven, Connecticut 06520

Research Report 168, Oct. 1979.

This paper reports on some recent developments in natural language analysis. We address such issues as the role of syntax in a semantics-oriented analyzer, achieving a flexible balance of top-down and bottom-up processing, and the role of short term memory. Our results have led to improved algorithms capable of analyzing the kinds of multi-clause inputs found in most text.

Definitional Mechanisms for Conceptual Graphs

John F. Sowa

IBM Systems Research Institute
205 East 42nd Street
New York, New York 10017

Graph Grammars and Their Application to Computer Science and Biology, edited by V. Claus, H. Ehrig, & G. Rozenberg, Lecture Notes in Computer Science, No. 73, Springer Verlag, Berlin, 1979, 426-439.

Conceptual graphs formalize the semantic networks used to represent meaning in artificial intelligence and computational linguistics. This paper presents mechanisms for defining new types of concepts, conceptual relations, and composite entities having other entities as parts. Type expansion allows graphs containing high-level concepts to be expanded into graphs con-

taining only low-level primitives; alternatively, type contraction allows graphs containing many low-level primitives to be contracted into graphs with a smaller number, but higher level, of concepts and relations. Inferences can then be performed on any level that is appropriate: the high-level concepts in the contracted form or the low-level primitives in the expanded form.

Notes on an Intensional Logic for English III: Extensional Forms

Joyce Friedman and David S. Warren

Computer and Communication Sciences
2076 Frieze Building
105 South State Street
University of Michigan
Ann Arbor, Michigan 48104

Computer Studies in Formal Ling. N-13, May 1979.

Montague (1973) translates English into an intensional logic which is an extension of the typed lambda-calculus. In this paper we consider the ways in which Montague's Meaning Postulates can be used to justify the introduction of extensional forms of translations. We define a function XTOU, which replaces quantification over individual concepts by quantification over entities, and a function FIXEXT, which then introduces extensional constants in place of the original intensional forms of lexical items. These obtain the forms of translations displayed in examples in *PTQ*. We prove that the two functions preserve logical equivalence in all models satisfying the Meaning Postulates.

Using Slots and Modifiers in Logic Grammars for Natural Language

Michael C. McCord

Department of Computer Science
915 Patterson Office Tower
University of Kentucky
Lexington, Kentucky 40506

Technical Report No. 69-80, April 1980.

In this paper, ideas are presented for the expression of natural language grammars in clausal logic, following the work of Colmerauer, Kowalski, Pereira, and Warren. A uniform format for syntactic structures is proposed, in which every syntactic item consists of a central *predication*, a cluster of *modifiers*, a list of *features*, and a *determiner*. The modifiers of a syntactic item are again syntactic items (of the same format), and a modifier's determiner shows its function in the semantic structure. In the rules for syntax, the notions of *slots* and *slot-filling* play a central role, following previous work by the author. The Appendix contains an example grammar and samples of parses and semantic interpretations into logical form. The system is implemented in Prolog.

An Experimental Study of Natural Language Programming

**Alan W. Biermann, Bruce W. Ballard,
and Anne M. Holler**

Department of Computer Science
Duke University
Durham, North Carolina 27706

Technical Report CS-1979-9, July 1979.

An experiment is described which gives data related to the usefulness and efficiency of English as a programming language. The experiment was performed with the NLC system described herein and used twenty-three paid volunteers from a first course in programming. The subjects were asked to solve two problems, one on the experimental system and one using the PL/C language studied in their course. The subjects typed a total of 1581 English sentences, 81 percent of which were processed correctly. The remaining 19 percent were rejected because of questionable user syntax or system inadequacies which are discussed.

None of the standard concerns about natural language programming related to vagueness, ambiguity, verbosity, or correctness was a significant problem, although minor difficulties did arise occasionally. The time required to solve the problems using English was much less than for PL/C. However, the natural language dialogues do not presently produce programs as general as their PL/C counterparts.

Using Pragmatic Knowledge for Natural Language Understanding: One Realization on Cooking Recipes

M.O. Cordier

Laboratoire de Recherche en Informatique
Bâtiment 490
Université de Paris - Sud
F-91405 Orsay, FRANCE

Research Report No. 48, 1979.

A program understanding a natural language requires a large pragmatic knowledge on the domain (in addition to semantic and syntactic ones). How to transit it? How to use it? That is the central problem of our work. The program which we have written is oriented to robot control. It analyses a text written in a natural language, infers all the necessary information, and outputs the elementary actions to be executed by the robot. The context used for the evaluation is that of cooking recipes. This system is designed to be highly adaptable to a new context. Therefore the large amount of necessary knowledge is entirely provided as data, via an appropriate easily usable language.

Natural Language: Some Aspects of it

D. Kayser

Laboratoire de Recherche en Informatique
Bâtiment 490
Université de Paris - Sud
F-91405 Orsay, FRANCE

Research Report No. 54, 1979.

The attempts to model the process of natural language comprehension with a computer have shed some light on the very nature of natural language itself. We discuss here some issues concerning language processing and show that any computational choice entails a commitment about what natural language really is. We conclude that what makes a language to be natural is its ability to be interpreted at variable depths.

Some Speculations on Language

Bernard Meltzer

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

DAI Research Paper No. 123, 1979.

Artificial Intelligence, as well as neurophysiological studies of recent decades, suggests a unitary view of language, in which natural language constitutes only part of the total language of the organism; the latter is a single but complex structure containing also the symbol-systems and their transformations responsible for mental processes. Following Sloman, the dogma that communication is the main function of language is opposed, and implications of the unitary thesis in respect to a number of linguistic and psychological questions are briefly discussed.

Pragmatic Problems of Man/Computer Dialogues

Jouko Seppänen

Helsinki University of Technology
Computing Centre
SF-02150 Espoo 15, FINLAND

Research Report 12, May 1979.

Some pragmatic problems of man/computer tactile display dialogues are discussed in the light of experience from interactive computer systems used at the Helsinki University of Technology. Various dialogue support features, techniques and concepts are described and assessed from human engineering and system design point of view.

Description Semantique et Dynamique du Discours

Thomas R. Hofmann

Department of English
Shimane University
Matsue City, Shimane, JAPAN 690

Universite de Paris IV Dissertation, 1978.

The aims of this study are first to provide epistemological foundations for a scientific study of descriptive or propositional semantics, independent of particular theories of syntax, and then to develop certain aspects of semantic theory to show that for a semantic theory to be adequate, it must integrate the propositional meanings of successive sentences together into larger units. Such theories are called integrative, and the last major aim is to present one such integrative semantic theory, the C-net theory developed by the author and J-P. Paillet, in a semi-formal fashion, and to develop some of its implications about the lexicon. This work shows that integrative semantics is both necessary and possible, and it adds some detail to the structure of an integrative semantic theory.

Clues to Vocabularic Structure: Comparative Suffixal Productivity in the Five Books of Rabelais

Heath Tuttle

North Carolina Educational Computing Service
Research Triangle Park, North Carolina 27709

Lingua e Stile, Anno XIV, n. 1, March 1979.

Quite typically in word-formation studies, and more particularly in studies of suffixal derivation, productivity, i.e. the extent to which a suffix is or is not used to form new derivatives, is described in qualitative rather than quantitative terms. By using a computer, one can group words together by suffix facily, and then date the derivatives using historical-etymological dictionaries. So for a given time span all new derivative productions can be obtained and counted, subject, of course, to their datability in such dictionaries.

The some 12,000 words in the Marty-Laveaux Glossaire were punched onto computer cards. A computer program was written to reverse alphabetize the words. The reverse dictionary thus produced contained the words spelled in their normal order, but alphabetized from the right instead of from the left. The result was that the words were grouped together by suffix rather than by prefix (as in a normal dictionary). This grouping allowed all the words in Rabelais' vocabulary with any particular suffix to be examined. Given a list of the suffixes active in the sixteenth century, the final, major task was to isolate those words with these suffixes which were actually derived in the sixteenth century, the period investigated; they had to have been produced by the joining of a suffix to a root in the sixteenth century, and not be merely a sixteenth-century spelling of an already existing French, Latin, or Greek word. To determine whether or not a word was of sixteenth century genesis, standard dictionaries were consulted. Graphs were then generated for the some seventy suffixes considered, showing the number of sixteenth-century derivatives in each of the five books.

An ASL Dictionary in APL

David V. Moffat

Department of Computer Science
North Carolina State University
P.O. Box 5972
Raleigh, North Carolina 27650

SIGLASH Newsletter 13, 1 (March 1980), 2-20.

This paper summarizes a project to develop a computerized dictionary of American Sign Language. The topics discussed are: project scope and development, the resulting system of programs, applications in sign language studies, and problems for future development. Appendices describe details of the notation used for sign descriptions, show an example dictionary, and tell how some of the programs are used.