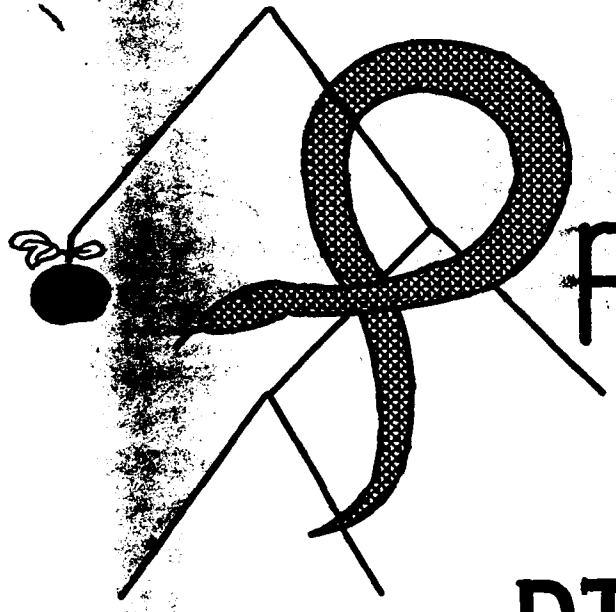
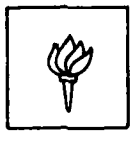
Research in Natural Language Processing
January 15, 1985 - March 31, 1990

Ralph Grishman

PROTEUS Project Memorandum #35
June 1990



PROTEUS PROJECT



Department of Computer Science
Courant Institute of Mathematical Sciences
New York University

(1)

Research in Natural Language Processing
January 15, 1985 - March 31, 1990

Ralph Grishman
PROTEUS Project Memorandum #35
June 1990

Final Technical Report for
Contract N00014-85-K-0163
with the Office of Naval Research

PROTEUS Project
Computer Science Department
Courant Institute of Mathematical Sciences
New York University

Research in Natural Language Processing
January 15, 1985 - March 31, 1990

Ralph Grishman, principal investigator

This report describes research done by the PROTEUS Project at New York University during the period January 15, 1985 to March 31, 1990. All of the activities described below were supported in part by the Strategic Computing Program of the Defense Advanced Research Projects Agency under Contract N00014-85-K-0163 from the Office of Naval Research. Accordingly, this memorandum serves as the technical portion of the final report on this contract. Some of the work reported here was also funded in part by other agencies; this is noted under the specific topics below.

This report is intended to provide only an overview and outline of the various research activities performed during this period. Technical details of the various activities are provided by the Proteus Project Memoranda cited herein. Many of these memoranda have also been published in journals and conference proceedings. A listing of these memoranda and their place of publication appears later in this report.

Research Goals

Our primary interest is in the development of systems which can automatically process natural language text concerning limited domains. At the outset of our research, ~~both~~ NYU ~~and~~ a number of ~~other~~ research groups had created systems which could analyze some moderately complex texts. However, these systems were just demonstrations operating on small bodies of text, and were too fragile to become prototypes of operational systems (the few exceptions either operated on extremely simple text or extracted only a few predefined types of facts from a text). ~~Our~~ central goal was to develop the techniques and tools which would allow us to create robust prototypes for operational systems. This central goal has led us to a wide variety of research in computational linguistics, from the development of parsers, grammars, and complete applications for text processing to basic research on issues of syntactic and semantic analysis. ~~We~~ summarize below these different research areas and provide pointers to the memoranda, publications, and dissertations produced as part of this research.

1. Parsing and Grammars

1.1. Core Syntactic Analyzer

(Developed by Jean Mark Gawron, Ngo Thanh Nhan, Michael Moore, and Ralph Grishman)

The PROTEUS Syntactic Analyzer was developed to provide an efficient, easy-to-use base for the various experiments in computational linguistics described below. The syntactic analyzer is based on the active chart parsing algorithm and uses a compositional technique based on lambda reduction for syntactic regularization. A PROTEUS Restriction Language compiler is provided which translates a declarative language suitable for stating grammatical constraints into LISP code to check these constraints during parsing.

The analyzer was originally coded in Franz LISP, ported to Symbolics Zetalisp and then to Common Lisp (now running on Symbolics LISP machines, on the SUN-3 and SUN-4 under Allegro Common Lisp, and on IBM PC/ATs under Golden Common Lisp). While the coding and porting of the system was substantially completed in the first two years of the project, at least one major enhancement has been made each year since then. The Restriction Language compiler has been entirely recoded and substantially speeded up; the syntactic regularization rules have been simplified; and a mechanism for computing paths in the context-free grammar (for attribute passing) has been added.

The program has been distributed to several sites: to the Naval Research Laboratory under a parser development contract with the laboratory, to Monmouth College (New Jersey) as part of our research in machine translation, to the University of Melbourne (Australia), and to Merck & Co. (New Jersey).

Development of the parser was also supported in part by the Naval Research Laboratory under Contract N00014-85-K-2028, and by the National Science Foundation under Grant DCR-85-01843.

Documentation: Proteus Project Memoranda #4, 5, and 9

1.2. English Grammar

(Developed by Ngo Thanh Nhan and Ralph Grishman)

We have been gradually building an English grammar for use with the PROTEUS Syntactic Analyzer. This grammar has been adapted to the needs of our various applications and the sublanguages they require; in turn these needs have led to a gradual broadening of the English grammar. It is an augmented context-free grammar which is based on Linguistic String Theory. The context-free component closely follows the grammar presented by Sager in *Natural Language Information Processing*. The regularization component converts the parse into a simpler structure based on clauses and noun phrases with labeled syntactic slots and features (roughly comparable to the f-structure of LFG).

1.3. Comparatives

(Dissertation research of Carol Friedman)

Comparative constructions ("more ... than ...") are generally regarded as one of the most difficult to handle in an English grammar because they occur in a variety of forms and syntactic environments. Most systems provide only a very limited coverage of comparatives. We therefore decided to implement a broad-coverage treatment of comparatives, both because of its importance as a component of the grammar and as a demonstration of the extensibility of our system. Extending prior work by Carol Friedman, we developed a treatment of comparatives based on their structural similarity to other constructs of English, such as coordinate conjunction. This treatment, along with the necessary semantic extensions, was incorporated into our question-

answering system (see section 5 below).

Publications: Proteus Project Memoranda #21 and 24

1.4. Parallel Parsing

(Research of Mahesh Chitrao)

Parsing remains one of the most time-consuming aspects of natural language processing. The rapid proliferation of small parallel processing systems appears to offer one possible route for substantially speeding up this task. The objective of our experiments was to determine whether, through relatively simple modifications to existing systems, we can obtain substantial speed-ups in parsing.

We modified the chart parser of PROTEUS to operate under ZLISP, a parallel LISP developed by Isaac Dimitrovsky as part of the NYU Ultracomputer Project. We ran a series of experiments, both under simulation and on the actual Ultracomputer, varying the complexity of the grammar, the length of the sentence, and the number of processors. Our results were generally encouraging; we obtained typical speed-ups of 5 to 7 with our largest grammar.

Publication: Proteus Project Memorandum #10

1.5. Stochastic Grammars

(Ongoing doctoral research of Mahesh Chitrao)

The grammars developed for some of our message processing applications (described below) allow for a variety of sentence fragments and other telegraphic constructs. This leads to a lot of ambiguity: many full sentences can also be analyzed as some combination of fragments. To reduce this ambiguity, we introduced scoring rules into the grammar (to prefer full-sentence analyses over fragments) and modified the parser to use these scores and perform a best-first search for the analysis with the highest score.

However, these scoring rules are ad hoc and difficult to assign. A more systematic approach which we have pursued is to derive these scores from the statistics of usage of the productions in our grammar when analyzing a "training" corpus. More frequently used productions are assigned higher scores and are therefore favored by the parser. Preliminary tests of this scheme have been encouraging, but extensive further experiments will be required.

2. Domain Modeling for Language Analysis

(Dissertation research of Tomasz Ksiezyk, with contributions by Ngo Thanh Nhan, John Sterling, and Leo Joskowicz)

Most of our application tasks arose from the need to process various types of messages containing natural language narrative. Our original research interest lay in the creation of a language analysis system which posessed and could make effective use of a detailed domain model. We recognized that a full understanding of some of these narratives (including in particular the causal and temporal relations) would required such a detailed model. In addition, we felt that in the long term such a model, complemented by syntactic and general semantic knowledge, would be important for robust text analysis. We therefore focussed our research during the first several years (1985-88) on one type of message, CASREPs (equipment casualty reports), on developing detailed domain knowledge (a model of the equipment), and on using this knowledge for language understanding.

We selected one large piece of equipment, the starting air system for propulsion gas turbines, which is of substantial yet manageable complexity and has had frequent reported failures. We constructed a detailed model of this equipment, incorporating sufficient structural and functional information to permit a qualitative simulation of the system. Based on a corpus of 36

CASREPs involving this equipment, we developed a CASREP grammar (using the PROTEUS Syntactic Analyzer). In cooperation with UNISYS Defense Systems we developed a semantic analysis procedure. Finally, during 1987 and 1988 we developed a discourse analyzer capable of identifying implicit causal and temporal relationships among the events described in the message.

The domain model is used primarily by the noun phrase analyzer and the discourse analyzer. The noun phrase analyzer, which must cope with long, complex nominals, relies heavily on the structural and descriptive information in the equipment model. The recovery of causal relations in discourse analysis is based primarily on the simulation capabilities of the model. In addition, a dynamic graphical interface to the model provides direct visual feedback regarding the system's interpretation of a message.

This system has been demonstrated publically at several stages during its development: at the DARPA Strategic Computing -- Natural Language Processing meetings in Los Angeles, CA, May 1986 and in Philadelphia, PA, May 1987; and at the Assn. for Computational Linguistics Annual Meetings in New York, June 1986, in Palo Alto, CA, July 1987, and in Buffalo, NY, June 1988.

Publications: Proteus Project Memoranda #1, 2, 3, 6, 8, 11, 14, 17, and 31. (#31 provides the most recent brief overview; #11 provides a more detailed exposition of the system).

In addition, a 10-minute videotape entitled "Message Understanding through Equipment Simulation" was presented at IJCAI-89 and published by Morgan Kaufmann as part of a collection of tapes from that conference.

3. Robust Message Processing

(Research of John Sterling and Ralph Grishman)

Beginning in 1987, we began to look at some message processing tasks which did not require such a detailed or extensive domain model and thus be more suitable for the short-term development of prototypes. This work was stimulated by the MUCK (Message Understanding Conference) series, organized by the Naval Ocean Systems Center.

3.1. RAINFORMs

The first MUCK conference was held in June 1987. It was based on a corpus of Navy RAINFORM messages describing naval encounters and engagements. Participants were given a set of 10 messages one month prior to the conference and were asked to prepare a system to process these messages. The precise output was not specified, although most participants generated some representation of the "events" in the messages. We assembled our system from the PROTEUS parser, semantic code previously developed for CASREPs, and new code for reference resolution and some simple discourse analysis. At the conference, participants were asked to adapt their system to process one additional message and record any changes which were required to their system.

This effort provided a limited but significant demonstration of our ability to port the text processing software to a new and quite different domain. After the conference we spent some time extending, generalizing, and evaluating this system. In particular, one of the problems in processing these messages is the highly telegraphic input. In addition to the omission of top-level sentence constituents such as subject and tense (which was also observed in the CASREPs), there was frequent omission of function words such as prepositions, "to", and "as". We described a combination of extensions to the grammar, weighting rules in the grammar, and tight semantic constraints to handle such omissions.

Publication: Proteus Project Memorandum #22

3.2. OPREPs

The second MUCK Conference, held in the spring of 1989, involved a corpus of OPREP (OPerational REPort) messages quite similar in content to the earlier RAINFORMs. However, the task was somewhat more ambitious, reflecting the growing experience of the participants. 125 messages were distributed 3 months prior to the conference to aid in developing the systems; an additional 5 messages were used for on-site evaluation of the final systems. The system's task was to create data base entries from the messages; each entry corresponded to one "event" in a message, and recorded various features of the event, such as agent, patient, time, and location. Expected system output was specified, along with a scoring function for rating each system's output.

Our system performed relatively well on this task (because of disagreements about ground rules and about evaluation, precise comparisons were not published). One reason for our good performance was our emphasis on mechanisms for error correction and constraint relaxation. For example, our system included a spelling corrector and a grammar capable of handling a variety of fragments and run-on sentences. Most important was our ability to deal with gaps in the system's semantic knowledge. In our prior systems, a sentence would not be accepted unless every subject-verb-object and head-modifier combination was recognized as an instance of a valid semantic pattern. However, collecting all possible valid semantic patterns was unrealistic, particularly given the time constraints and the complexity of the domain. So we instead used a system of preference semantics, which preferred parses containing valid semantic patterns but did not impose any absolute constraints. The benefits obtained by using preference semantics were dramatic.

As part of our cooperation with other DARPA contractors, a copy of this system was provided to BBN Systems and Technologies Corp.

Publications: Proteus Project Memoranda #26, 28, and 32

4. Question Answering and User Feedback

4.1. Question-Answering System

A small question-answering system (natural language interface for data base retrieval), operating on a small data base of student transcripts, was previously built for research and teaching purposes. This system was substantially modified to operate with the PROTEUS syntactic analyzer. It has been used in our graduate course in computational linguistics and for research on user feedback (see next section) and on comparatives (see section 1.3).

4.2. Feedback for Semantic Overshoot

(Research of Ping Peng)

One of the most difficult tasks in developing a natural language interface involves collecting a complete set of semantic relations and the forms in which they may be expressed. For a large domain and complex application the task may be impossible -- the domain may not be closed or cleanly delineated, and completeness may be beyond reach. We are therefore faced with the problem of handling semantic overshoot: user input which exceeds the semantic model incorporated in the system. This aspect of our research has focussed on providing helpful user feedback in cases of semantic overshoot.

Our approach in cases of semantic overshoot has been to identify the closest variants of the user's input which would be acceptable to the system (i.e., would be within the system's semantic model), and to provide these variants as suggestions to the user. We have conducted a series of experiments using a small question-answering system operating on a domain of student transcripts and course prerequisites. Because of the limitations of the system in terms of vocabulary

and syntax, we have found that semantic overshoot is not the primary reason for the rejection of user input; however, in cases of semantic overshoot our system was able, in the majority of cases, to provide appropriate feedback so that the user could reformulate his query.

Research supported in part by the National Science Foundation under Grant No. DCR-8501843.

Publication: Proteus Project Memorandum #7

5. Machine Translation

(Research of Michiko Kosaka [Monmouth College], Virginia Teller [Hunter College and the Graduate Center, The City University of New York], Ralph Grishman, and Ping Peng)

At its heart, machine translation is a specialized form of text processing. Admittedly, the task of transferring information from one language to another (at some level of representation) is not shared by other text applications, and the need for generation is shared by only a few. However, the greatest hurdle to high-quality translation lies in succesful source language disambiguation, both syntactic and lexical. It is here that we may hope that our experience with text analysis may be of benefit to the problems of machine translation.

This work has been done in collaboration with Monmouth College (New Jersey) and Hunter College (part of the City University of New York). Because one of our collaborators is a native speaker of Japanese, and because of the strong interest in the translation of Japanese technical material, all of our work has been in the translation of Japanese to English.

This research is supported in part by the National Science Foundation under Grant IRI-8902304.

5.1. Sublanguage Analysis

Most translation systems consist of an analysis phase, a transfer phrase, and a generation phase. By analyzing the source language to a relatively "deep" syntactic level, and generating the target language from this level, we can account for most of the systematic syntactic differences between the languages. However, we still need to describe the lexical correspondence between the languages (the word correspondence is rarely 1-to-1) and the idiosyncratic patterns of particular words (for example, the prepositions governed by verbs).

These correspondences are typically captured by lexical transfer rules, which test the immediate syntactic environment of a verb (e.g., the subject and object of a verb). These rules have generally been built up in a rather ad hoc manner, creating new (semantic) features for the rules as necessary. We have explored instead the possibility of developing such rules more systematically through sublanguage analysis: examining the co-occurrence patterns of sample texts in a domain in order to identify the sublanguage classes and patterns. We analyzed a small segment of a programming language manual, in both the original English and a faithful Japanese translation. We identified the sublanguage patterns in the two languages and the correspondence between the patterns in the two languages. We found a very close correspondence, particularly for the domain-specific words, thus suggesting that sublanguage analysis may indeed be a useful approach to building transfer patterns.

Publications: Proteus Project Memoranda #15 and 16

5.2. Japanese Grammar Construction

We are gradually constructing a broad-coverage Japanese grammar for our machine translation efforts. This work has proceeded in several stages as we have addressed successively more complex texts. We began by constructing a grammar for a Japanese version of our question-answering system. More recently, we have extended the grammar to cover a dialog on conference registration. As a next stage, the grammar will be extended to handle portions of a programming language manual. The grammar has been constructed in the augmented context-free framework provided by the Proteus parser.

5.3. Japanese Question-Answering System

(Developed by Ping Peng)

As an initial effort at testing our Japanese-language capabilities, we constructed a Japanese version of the question-answering system described above. We then extended this to produce English translations of questions posed in Japanese.

5.4. Reversible Grammar

(Research of Tomek Strzalkowski and ongoing doctoral research of Ping Peng)

One potential advantage of non-procedural grammars, such as unification grammars, is that they can operate bidirectionally, to either analyze or generate sentences. This has frequently been touted as an advantage of PROLOG definite-clause grammars. In practice, however, grammars which are efficient for parsing turn out to be very inefficient (if they function at all) for generation. To convert a parsing definite-clause grammar to an efficient generation grammar has typically required a restructuring of the grammar. We have developed algorithms to do this automatically by analyzing the definition-use patterns between the literals of the PROLOG clauses and then rearranging the literals in each clause. This algorithm has been tested on a moderate-sized English grammar (a unification version of the PROTEUS English grammar), and has been shown to produce a very efficient generator.

Publications: Proteus Project Memoranda #25, 29, 30, and 33

6. Formal semantics and pragmatics of natural language

6.1. Inter-sentential dependencies and non-singular concepts

(Research of Tomek Strzalkowski)

The goal of this research is to develop a computational model of processing and understanding of natural language, called the Stratified Model. In this design various stages (strata) in natural language processing have been identified, and related to corresponding world models that provide a denotational base for linguistic objects. In the research to date, two problems, in particular, received a great deal of attention: computing of inter-sentential dependencies in discourse, and representing the meaning of non-singular concepts in natural language. The first of these problems concerns translating sentences occurring within a larger discourse into a formal representation in logic. This work on these issues resulted in development of a collection of formal rules of translation into logic.

The research on meaning representation for non-singular concepts in natural language discourse investigates the problem of reference to things and entities at varying levels of aggregation: at generic level (birds fly), group level (the eggs are rotten), individual level (Tweety is a bird), and even subindividual level (Bush [here and now] is a president). The levels are related by coordinates, special functions that decompose higher level objects into their instances at a level below. Since an object can be decomposed with various coordinates, there will usually be several such instance levels. Among the coordinates the various space-time functions are probably most frequent, but other, more abstract (specimen, genera) are also found.

Further work was done on the problems of adequate representation of various types of generic sentences, and capturing the differences in their truth conditions. The notions of transparently generic sentences (birds fly) and opaquely generic sentences (physicists win Nobel Prizes) have been introduced. It has been proposed that a new system of quantifiers needs to be defined for first order logic in order to capture the meaning of these sentences.

Publications: Proteus Project Memoranda #19, 20, and 27

6.2. Intra-sentential temporal analysis

(Research of Sashidhar Reddi)

Most of our research has involved *narrative* text. For narratives, determining the temporal relations between events is an important part of the overall text analysis process. The analysis of temporal structure was a major part of our work on CASREPs (see section 2 above). In addition, we have studied the problem of interpreting temporal subordinating conjunctions ("after", "until", ...). We have developed a classification of predications into different types of events, processes, and states, and shown how this classification can be useful in interpreting subordinating conjunctions.

Publication: Proteus Project Memorandum #18

Proteus Project Memoranda

Cumulative Listing to June 1, 1990

1 **PROTEUS and PUNDIT: Research in Text Understanding**
R. Grishman and L. Hirschman
April 1986
Published in Computational Linguistics 12 (2), 141-145, 1986.

2 **Model-based Analysis of Messages about Equipment**
R. Grishman, T. Ksiezyk, and N. T. Nhan
April 1986
Also issued as Computer Science Dept. Technical Report #236

3 **An Equipment Model and its Role in the Interpretation of Nominal Compounds**
T. Ksiezyk and R. Grishman
April 1986
Also issued as Computer Science Dept. Technical Report #237
Superseded by PPM #6

4 **PROTEUS Parser Reference Manual**
R. Grishman
July 1986
Rev. A - June 1989
Rev. B - August 1989
Rev. C - May 1990

5 **Syntactic Regularization in Proteus**
J. M. Gawron
September 1986

6 **An Equipment Model and its Role in the Interpretation of Noun Phrases**
Tomasz Ksiezyk, Ralph Grishman, and John Sterling
January 1987
Rev. A - April 1987
Published in Proceedings IJCAI-87

7 **Responding to Semantically Ill-formed Input**
Ralph Grishman and Ping Peng
January 1987
Rev. A - September 1987
Rev. B - November 1987
Published in Proc. Second Conf. on Applied Natural Language Processing, Austin, TX, Feb. 1988

8 **Finding Causal and Temporal Relations in Equipment Failure Messages**
Leo Joskowicz, Ralph Grishman, and Tomasz Ksiezyk
February 1987
Superseded by PPM #13

9 **An Introduction to the PROTEUS Parser**
Ralph Grishman
March 1987
Rev. A - August 1989

10 **Evaluation of a Parallel Chart Parser**
Ralph Grishman and Mahesh Chitrao
September 1987
Rev. A - November 1987
Published in Proc. Second Conf. on Applied Natural Language Processing, Austin, TX, Feb. 1988

11 **Equipment Simulation for Language Understanding**
Tomasz Ksiezyk and Ralph Grishman
September 1987
Rev. A - June 1988
Published in International Journal of Expert Systems 2 (1) 33-78, 1989.

12 **Research in Natural Language Processing: Jan. 15, 1985 - Sept. 15, 1987**
Ralph Grishman
November 1987

13 **Deep Domain Models for Discourse Analysis**
Leo Joskowicz, Ralph Grishman, and Tomasz Ksiezyk
November 1987
Rev. A - March 1989
Published in The Annual AI Systems in Government Conference, Washington, DC, March 27-31, 1989

14 **Domain Modeling for Language Analysis**
Ralph Grishman
February 1988
Presented at 1988 Duisburg Symposium on Linguistic Approaches to Artificial Intelligence

15 **A Comparative Study of Japanese and English Sublanguage Patterns**
Virginia Teller [Hunter College & CUNY Graduate Center], Michiko Kosaka [Monmouth College], and Ralph Grishman
June 1988
Published in Proc. 2nd Int'l Conf. on Theoretical and Methodological Issues in Machine Translation of Natural Languages, Pittsburgh, PA, June 12-14, 1988.

16 **A Sublanguage Approach to Japanese-English Machine Translation**
Michiko Kosaka [Monmouth College], Virginia Teller [Hunter College & CUNY Graduate Center], and Ralph Grishman
August 1988
Published in New Directions in Machine Translation, Proc. of the Conference, Budapest, Aug. 18-19, 1988.

17 **Simulation-Based Understanding of Texts about Equipment**
Tomasz Ksiezyk *(doctoral dissertation)*
September 1988

18 **Extraction of Intra-Sentential Temporal Information**
Sashidhar Reddi
December 1988

19 **A Multi-level Model for Representing Generic and Habitual Sentences**
Tomek Strzalkowski
December 1988

20 **Meaning Representation for Generic Sentences**
Tomek Strzalkowski
December 1988

21 **A Computational Treatment of the Comparative**
Carol Friedman *(doctoral dissertation)*
January 1989

22 **Analyzing Telegraphic Messages**
Ralph Grishman and John Sterling
February 1989
Published in Proc. of the [DARPA] Speech and Natural Language Workshop, Philadelphia, PA (Feb. 21-23, 1989), Morgan Kaufman Publishers.

23 **A Very Brief Introduction to Computational Linguistics**
Ralph Grishman
February 1989

Published in Proc. of the [DARPA] Speech and Natural Language Workshop, Philadelphia, PA (Feb. 21-23, 1989), Morgan Kaufman Publishers.

24 **A General Computational Treatment of the Comparative**
Carol Friedman
April 1989
Published in Proc. 27th Annual Meeting Assn. Computational Linguistics, Vancouver, B.C. (June 26-29, 1989).

25 **Automated Inversion of a Unification Parser into a Unification Generator**
Tomek Strzalkowski
September 1989

26 **Preference Semantics for Message Understanding**
Ralph Grishman and John Sterling
October 1989
Published in Proc. [DARPA] Speech and Natural Language Workshop, Harwich Port, MA, Oct. 15-18, 1989, Morgan Kaufman Publishers.

27 **Extra-sentential Dependencies, Meaning Representation, and Generics**
Tomek Strzalkowski
November 1989
Published in Proc. Fourth Portugese Conf. on Artificial Inteligence, Lecture Notes in Artificial Intelligence #390, pp. 210-221, Springer.

28 **Towards Robust Natural Language Analysis**
Ralph Grishman and John Sterling
February 1990
Prepared for the AAAI Symposium on Text-Based Intelligent Systems, Stanford Univ., March 27-29, 1990.

29 **An Implementation of a Reversible Grammar**
Ping Peng and Tomek Strzalkowski
May 1990
To appear in Proc. 8th Canadian Conf. on Artificial Intelligence, Ottawa, May 1990.

30 **Automated Inversion of Logic Grammars for Generation**
Tomek Strzalkowski and Ping Peng
May 1990
To appear in Proc. 28th Annual Meeting Assn. Computational Linguistics, Pittsburgh, 6-9 June, 1990.

31 **Causal and Temporal Text Analysis: The Role of the Domain Model**
Ralph Grishman and Tomasz Ksiezyk
May 1990
To appear in Proc. 13th Int'l Conf. Computational Linguistics (COLING 90), Helsinki, August 20-25, 1990.

32 **Information Extraction and Semantic Constraints**
Ralph Grishman and John Sterling
May 1990
To appear in Proc. 13th Int'l Conf. Computational Linguistics (COLING 90), Helsinki, August 20-25, 1990.

33 **How to invert a natural language parser into an efficient generator: an algorithm for logic grammars**
Tomek Strzalkowski
May 1990
To appear in Proc. 13th Int'l Conf. Computational Linguistics (COLING 90), Helsinki, August 20-25, 1990.

Send requests for memoranda to

Prof. Ralph Grishman
Computer Science Dept.
New York University
715 Broadway, Room 703
New York, NY 10003

Internet: grishman@nyu.edu