

AD-A035 165

BOLT BERANEK AND NEWMAN INC CAMBRIDGE MASS
SPEECH UNDERSTANDING SYSTEMS. VOLUME I. INTRODUCTION AND OVERVI--ETC(U)
DEC 76 W A WOODS, M BATES, B BROWN, B C BRUCE N00014-75-C-0533
BBN-3438-VOL-1

F/G 17/2

UNCLASSIFIED

NL

| OF |

AD
A035165



END

DATE
FILMED

3-77

BBN Report No. 3438

ADA035165

[Handwritten initials]

12

SPEECH UNDERSTANDING SYSTEMS

Final Report

November 1974 - October 1976

Volume I: Introduction and Overview

Sponsored by
Advanced Research Projects Agency
ARPA Order No. 2904

DDC
RECEIVED
FEB 8 1977
D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by ONR under Contract No. N00014-75-C-05330 ✓

The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER BBN 3438 - 3438 - Vol-1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) SPEECH UNDERSTANDING SYSTEMS, Volume I. Final Technical Progress Report 30 October 1974 to 29 October 1976 Introduction and Overview.	7. TYPE OF REPORT & PERIOD COVERED Final Prog. Report, 30 Oct 1974 to 29 Oct 1976	
5. AUTHOR(s) W. Woods, M. Bates, G. Brown, B. Bruce, C. Cook, J. Klovstad, J. Makhoul, B. Nash-Webber, R. Schwartz, J. Wolf, V. Zue.	8. CONTRACT OR GRANT NUMBER(s) N00014-75-C-0533	
9. PERFORMING ORGANIZATION NAME AND ADDRESS Bolt Beranek and Newman Inc. 50 Moulton Street Cambridge, Mass. 02138	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 5D30 (12) 90p.	
11. CONTROLLING OFFICE NAME AND ADDRESS ONR Department of the Navy Arlington, Virginia 22217	12. REPORT DATE (11) December 1976	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	13. NUMBER OF PAGES 82	
	15. SECURITY CLASS. (of this report) Unclassified	
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	
16. DISTRIBUTION STATEMENT (of this Report) Distribution of this document is unlimited. It may be released to the Clearinghouse Department of Commerce for sale to the general public.		
17. DISTRIBUTION STATEMENT (of this Report) (if different from Report) (10) Madeline Bates William A. Woods, Geoffrey Brown, Bertram C. Bruce, Craig C. Cook, John W. Klovstad		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Acoustic-phonetic experiment facility, acoustic-phonetic recognition, acoustic-phonetic rules, artificial intelligence, ATN grammars, audio- response generation, budget management, computational linguistics, control strategies, data base, dictionary expansion, discourse model, fact retrieval, formal command language, formant tracking grammars, HWIM, knowledge sources.		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) This five-volume final report is a review of the BBN speech understanding system, covering the last two years of the project from October 1974 through October 1976. The BBN speech understanding project is an effort to develop a continuous speech understanding system which uses syntactic, semantic, and pragmatic support from higher level linguistic knowledge sources to compensate for the inherent acoustic indeterminacies in continuous spoken utterances. These knowledge sources are integrated with attempts		

DD FORM 1473 EDITION OF 1 NOV 68 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

060100

19. Key words (cont'd.)

lexical retrieval, likelihood ratio, linear prediction, multi-component systems, natural language retrieval system, natural language understanding, parametric modeling, parsing, pattern recognition, phonetic labeling, phonetic segmentation, phonological rules, phonology, pragmatic grammar, pragmatics, probabilistic labeling, probabilistic lexical retrieval, prosodics, question-answering, recognition strategies, resource allocation, response generation, scoring philosophy, semantic interpretation, semantic networks, semantics, shortfall algorithm, shortfall density, signal processing, spectral matching, speech, speech analysis, speech generation, speech synthesis, speech recognition, speech understanding, SUR, syntax, synthesis-by-rule, system organization, task model, user model, word verification.

20. Abstract (cont'd.)

sophisticated signal processing and acoustic-phonetic analysis of the input signal, to produce a total system for understanding continuous speech. The system contains components for signal analysis, acoustic parameter extraction, acoustic-phonetic analysis of the signal, phonological expansion of the lexicon, lexical matching and retrieval, syntactic analysis and prediction, semantic analysis and prediction, pragmatic evaluation and prediction, and inferential fact retrieval and question answering, as well as synthesized text or spoken output. Those aspects of the system covered in each volume are:

- Volume I. Introduction and Overview
- Volume II. Acoustic Front End
- Volume III. Lexicon, Lexical Retrieval and Control
- Volume IV. Syntax and Semantics
- Volume V. The Travel Budget Manager's Assistant

ACCESSION No.		
NTIS	White Section	<input checked="" type="checkbox"/>
DTIC	Self Section	<input type="checkbox"/>
UNANNOUNCED		<input type="checkbox"/>
JUSTIFICATION		
BY		
DISTRIBUTION/AVAILABILITY CODES		
FORM	MAIL and/or SPECIAL	
A		

SPEECH UNDERSTANDING SYSTEMS

Final Report

November 1974 - October 1976

ARPA Order No. 2904

Contract No. N00014-75-C-0533

Program Code No. 5D30

Principal Investigator:

William A. Woods
(617) 491-1850 x361

Name of Contractor:
Bolt Beranek and Newman Inc.

Scientific Officer:
Marvin Denicoff

Effective Date of Contract:
30 October 1974

Title:
SPEECH UNDERSTANDING SYSTEMS

Contract Expiration Date:
29 October 1976

Editor:
Bonnie Nash-Webber
(617) 491-1850 x227

Amount of Contract: \$1,966,927

Sponsored by
Advanced Research Projects Agency
ARPA Order No. 2904

This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by ONR under Contract No. N00014-75-C-0533.

DDC
RECEIVED
FEB 8 1977
D

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

The BBN Speech Understanding System: Final Report

This is one of five volumes of the BBN Speech Understanding System Final Report. The Table of Contents for the entire set is given below.

Volume I. Introduction and Overview

- A. Introduction
- B. Design Philosophy of HWIM
- C. Overview of final system
- D. Design of final performance test and performance analysis overview
- E. Future
- F. References
- G. Appendices
 - 1. Sample set of sentence types
 - 2. Sample trace of an utterance being processed
 - 3. Publications
 - 4. Comprehensive Index to Technical Notes

Volume II. Acoustic Front End

- A. Acoustic Front End
- B. Acoustic-Phonetic Recognition
- C. A Speech Synthesis-by-Rule Program
- D. Verification
- E. References
- F. Appendices
 - 1. Dictionary Phonemes
 - 2. List of APR labels
 - 3. List of APR rules
 - 4. Parameters for Scoring

Volume III. Lexicon, Lexical Retrieval and Control

- A. Dictionary
- B. Phonological rules
- C. Dictionary Expansions
- D. Lexical Retrieval
- E. Control Strategy
- F. Performance
- G. References
- H. Appendices
 - 1. Annotated phonological rules
 - 2. Format and examples of dictionary files
 - 3. Result summaries for each token (TRAVELDICT and BIGDICT)
 - 4. Performance Results for Strategy Variations
 - 5. BIGDICT and TRAVELDICT listings
 - 6. Dictionary Expansion - A User's Guide

Volume IV. Syntax and Semantics

- A. Parsers
- B. Grammars
- C. Prosodics
- D. References
- E. Appendices
 - 1. Listing of MIDGRAM Grammar
 - 2. Sample Parse-Interpretations
 - 3. Parser trace

Volume V. TRIP

- A. Introduction
- B. The Travel Budget Manager's Assistant
- C. Flow of Control
- D. Linguistic Processing
- E. Execution
- F. Response Generation
- G. Conclusions
- H. References
- I. Appendices
 - 1. Data Base Structures
 - 2. Example Parses and Interpretations
 - 3. Methods
 - 4. Generation Frames
 - 5. A Generated Description of the Stored Trips

Acknowledgements

To the following people who contributed to the development of the BBN
Speech Understanding System:

William A. Woods, Principal Investigator
Madeleine Bates
Geoffrey Brown
Bertram C. Bruce
Laura Gould
Craig C. Cook
Gregory Harris
Dennis H. Klatt
John W. Klovstad
John I. Makhoul
Bonnie L. Nash-Webber
Richard M. Schwartz
Jared J. Wolf
Victor W. Zue

To our secretaries - Beverly Tobiason, Kathleen Starr and Angela Beckwith -
for the exceptional diligence, competence, and good humor they have shown
throughout the assembly of this report.

"Had we but world enough and time..."

Andrew Marvell, To His Coy Mistress

TABLE OF CONTENTS

	<u>page</u>
<u>Introduction and Overview</u>	
A. Introduction	1
B. Design Philosophy of HWIM	10
C. Overview of Final System	17
D. Final Performance Test of HWIM	27
E. Future	31
<u>References</u>	37
<u>Appendices</u>	
1. Sample Set of Sentence Types	38
2. Trace Listing Generated by HWIM	41
3. Publications	79
4. Comprehensive Index to Technical Notes	82

A. INTRODUCTION

1. Rationale for Speech Understanding Research

There are many reasons for wanting to use speech as a medium of communication between man and machine. A principal one is that it is the human being's most natural and effective output channel -- the word per minute rate of an ordinary speaker is approximately four times that of an average skilled typist and ten times that of an unskilled typist. In speech mode, a person's ability to convey ideas seems to be bounded only by his ability to conceive them, while in all other modes, he is bounded by his output channel. Other advantages are that the speaker already knows how to speak, thus making possible applications for people otherwise untrained, and speaking does not require the use of the speaker's hands or eyes, thus facilitating applications where the speaker's hands and eyes are busy. If telephone speech can be automatically understood by machines, then the existing telephone network can be used as an access path to computers via the ordinary telephone hand set. Recent advances in speech compression techniques and the potential for cheap mass produced telephone handsets that perform preliminary signal analysis for "packetized" speech make this possibility even more likely by providing higher quality speech information than the current telephone channels.

The number of potential application areas for speech understanding is enormous, with different applications depending on different characteristics of the speech communication channel. For example, pilots could use speech as a control channel while their hands and eyes were busy flying the plane. The same is true of tank commanders, foot soldiers, radar technicians, etc. Even for such tasks as taking inventory, the use of a cassette recorder, which frees up the hands and eyes of the taker and speeds up the inventory process has become a standard practice. However, someone eventually has to transcribe the tape. An automatic speech understander for this application would have many advantages, including the ability to cue the inventory taker for things that he had missed, to double check numbers that seem out of line, or to ask for repeats of information that is unintelligible. (In this situation, an offline transcriber has no such recourse.)

For a variety of management tasks, including command and control systems and intelligence applications, a significant problem is the entry of information into a data base. The capture of many kinds of data that one would like to have in a machine data base originates in a spoken or human perceptual form, and entry into a machine currently requires someone to type in the information. Speech understanding systems could eliminate this bottleneck and also the potential for errors, omissions, and other fumbles that arise from extra human agents as critical links in the chain.

2. Background and History

The report by Newell et al. [1971], which led to the creation of the ARPA Speech Understanding program, laid out excellent reasons for wanting continuous speech as a communication channel between men and machines, many of which have been mentioned above. Figure 1 gives a summary of the advantages of speech as a communication channel. Figure 2 gives a summary of the data rates for different communication channels. Recent work by Chapanis and Ochsman [Chapanis, 1975, Ochsman & Chapanis, 1974] makes a convincing case for the increased effectiveness of human communication in a variety of tasks when speech is permitted, compared to situations in which the interface requires typed or purely visual information. In their experiments, spoken communication permitted the solution of problem solving tasks in approximately half the time taken to solve the same problems with typed communication. (And interestingly, adding visual information produced very little additional improvement beyond that already provided by speech.)

For reasons such as the above, in the spring of 1970, Dr. Larry Roberts, then director of the Information Processing Technology (IPT) Office of ARPA, assembled a speech understanding study group to assess the feasibility of a program of focused research aimed toward the development of continuous speech understanding systems. The question at issue was whether the state of knowledge in the necessary fields was such that one could reasonably expect such a program to make progress. The conclusion of that group, expressed in its report [Newell et al., 1971], was that although the problem was difficult, the time was ripe for progress in this field, and a program was outlined for a reasonable first step aimed at producing demonstrable total systems in five years.

1. Most effortless encoding of all output channels.
(Especially if free language permitted.)
2. Higher data rate than other output channels.
3. Preferred channel for spontaneous output.
4. Does not tie up hands, eyes, feet or ears.
5. Can be used while in motion.
6. Can be used easily in parallel with other channels
or effectors.
(Possibly same as point 5, though possible
independent.)
7. Broadcast over short ranges (tens of feet).
8. Inexpensive and readily available terminal equipment.

Figure 1.

1. Reading out loud	~ 4	words/sec
2. Speaking (spontaneously)	~ 2.5	words/sec
3. Typing (record)	~ 2.5	words/sec
4. Typing (skilled)	~ 1	word/sec (~ 5 strokes/sec)
5. Handwriting	~ .4	word/sec
6. Hand printing	~ .4	word/sec
7. Telephone dialing (not touch-tone)	~ .3	word/sec (~ 1.5 digits/sec)
8. Mark sense cards	~ .1	word/sec (~ .5 digits/sec)

Figure 2. Speeds of various channels.

The Newell report made clear that the attainment of automatic speech understanding systems would require considerable research specifically focused on the task of constructing total speech understanding systems. As a first step, it proposed a specific set of goals and objectives for a five year program designed to constrain the problem to one of manageable proportions (for example, high quality speech input was assumed rather than, say, telephone speech; the vocabularies should be 1000 words, but the syntax and semantics were permitted to be limited and artificial in order to constrain the problem; multiple speakers were to be handled, but major dialect differences were not). A list of the target objectives from the Newell report, together with a synopsis of the performance of the BBN system on those dimensions is given in Figure 3.

The ARPA Speech Understanding Project has now completed its planned five years, during which time considerable progress has been made in the development of continuous speech understanding systems. In August and September of 1976, systems were demonstrated at System Development Corporation, at Carnegie-Mellon University, and at Bolt Beranek and Newman Inc., which were able to understand continuous speech of various levels of difficulty with varying degrees of success. These systems show the validity of the two major assumptions that underlay the initial speech understanding research program. The first is the basic feasibility of constructing continuous speech understanding systems by using a total system approach combining lexical, syntactic, semantic, and pragmatic constraints with the results of acoustical, phonetic, and phonological analysis. The second is the timeliness for focused research in this area -- that is, the necessary foundation had been laid by several decades of basic research in speech recognition, speech synthesis, and signal processing, and in computational linguistics and language processing, to make possible significant progress in their combination into total speech understanding systems. In this report we will present the results of BBN's part in this program, what we have learned, where we stand, and what remains to be done.

Goals (from 1971)

The system should:

- (1) accept continuous speech
- (2) from many
- (3) cooperative speakers of the general American dialect,
- (4) in a quiet room,
- (5) over a good quality microphone,
- (6) allowing slight tuning of the system per speaker,
- (7) but requiring only natural adaptation by the user,
- (8) permitting a slightly selected vocabulary of 1000 words
- (9) with a highly artificial syntax,
- (10) and a highly constrained task,
- (11) with a simple psychological model of the user,
- (12) providing graceful interaction,
- (13) tolerating less than 10% semantic error
- (14) in a few times real time,
- (15) [on a 100 MIPS machine]
- (16)
- (17)
- (18)
- (19) and be demonstrable in 1976 with a moderate chance of success.

HWIM (1976)

The system did:

- accept connected speech
- from three
- cooperative male general American speakers,
- in a somewhat quiet computer terminal room,
- over an ordinary close-talking microphone,
- with no tuning of the system for the individual speaker,
- and requiring no speaker adaptation,
- on 1097 words with no post-selections
- with a somewhat natural combined syntax and semantics (average branching >100),
- and a well-defined task,
- with no user model,
- and modest interaction capabilities,
- with 56% semantic error,
- in about 1350 times real time,
- on a .35 MIPS PDP-10,
- with 256K 36-bit words
- with a hierarchical system organization
- (cost per utterance not measured)
- and was operational 12 October 1976.

Figure 3.

3. Knowledge Components for Speech Understanding

When the ARPA speech understanding program first began in 1971, our state of knowledge about how an overall speech understanding system should be organized was extremely limited. The only existing example, the Vicens-Reddy system at Stanford University [Vicens, 1969] was so specialized that its understanding strategy relied on such system-specific facts as that most words began and ended with stop consonants and that all sentences contained the word "block". The techniques used were clearly not general enough to provide a model of how a speech understanding system should be organized. Looking instead to initial experience with human spectrogram readers [Klatt and Stevens, 1973] for our model, we became convinced that a potentially viable speech understanding system would involve some organization of the following conceptually distinct sources of knowledge:

a) Segmentation and Labeling

A process of detecting acoustic-phonetic events in the speech signal and characterizing the nature of the individual segments of the signal.

b) Lexical Retrieval

A process of retrieving candidate words from the lexicon that are acoustically similar to the labeled segments.

c) Word Matching

A process of determining some measure of the degree of similarity between a word hypothesis at a given point in the speech signal and the acoustic evidence there.

d) Syntax

The ability to determine if a given sequence of words is a possible subpart of a grammatical sentence and to predict possible continuations for such sentence fragments.

e) Semantics

The ability to determine if a given hypothesized sentence is meaningful or nonsensical (in addition to being grammatical).

f) Pragmatics

The ability to determine if a sentence is appropriate to the context in which it is uttered, given our knowledge of the particular speaker, the task he is trying to accomplish, and what has been said previously in the discourse.

In addition to these sources of knowledge, it was obvious that the spectrogram readers were making use of an additional ability that was considerably less overt. By some criteria, they were making decisions about which possible fragmentary hypotheses to rule out, which ones to pursue further by trying to find compatible interpretations of adjacent portions of the utterance, and when to return to a previously rejected hypothesis in light of new information. These decisions imply the existence of a control strategy for speech understanding, and from the beginning, our speech understanding system has been designed to facilitate the discovery and exploration of such strategies*.

4. Methodology

The method that we adopted to begin our research in speech understanding was one of "incremental simulation" [Woods and Makhoul, 1973] in which an overall system is "implemented" with some combination of computer programs and human simulators filling the roles of the different components. Initially, we implemented lexical retrieval and word matching components with computer programs, and then used one human to simulate a segmenting and labeling component, and another one, in another room, to simulate the syntax, semantics, pragmatics, and control components. Since this organization tended to break up the Gestalt recognition characteristic of a single person both forming word hypotheses and matching them against a spectrogram, it allowed us to explore the control strategy issue by making visible intermediate steps that would otherwise be subconscious. Our original experiences with this mode of exploration have now been widely supplemented by running a fully programmed system, with detailed tracings of intermediate results. These tracings, an example of which is given in

* Another ability, clearly significant for speech understanding, was not apparent in the behavior of the spectrogram readers. This is one's use of prosodic cues, such as the overall intonation and rhythm of the utterance. Where punctuation may give information in written text about clause boundaries, scoping of conjoined phrases, question/statement distinctions, etc., in spoken discourse, prosodic information gives similar information and much more -- it is considerably more versatile than written punctuation in resolving what would otherwise be syntactic ambiguity. However, speech prosodics is not well understood. Its absence from the human spectrogram reading was probably due to (a) no salient realization of fundamental frequency (or pitch) in the spectrogram and (b) little detailed understanding of prosodic cues and the significance of their presence or absence.

Appendix 2, are then subjected to detailed "failure analyses", as a result of which our understanding of speech understanding problems and control strategy issues has increased steadily during the project.

5. The Role of the Task

The difficulty of constructing a speech understanding system depends on an inherent level of difficulty of the task. Tasks differ in difficulty due to factors such as the amount of noise in the environment, the quality of the transmission channel through which the signal is acquired, the amount of individualized training of the speaker permitted, the number of different speakers and the differences of their dialects, the size of the vocabulary, and the inherent difficulty of the language that they are permitted to speak to the system. All of these factors except the last one are effectively controlled by the target goals of the ARPA speech understanding program. The difficulty of the language the speaker is permitted to use (i.e., the scope of the syntax that is permitted and the range of semantic concepts that can be referred to) was specifically left open, although permitted to be artificial and constraining. The selection of a task determines this language, including the basic vocabulary that will be recognized and the basic knowledge that will be used by the syntactic, semantic, and pragmatic components to judge the acceptability of hypothesized interpretations of the utterance.

In the first speech understanding system developed at BBN, we worked in the context of an existing natural language question-answering system, LUNAR [Woods et al., 1972], which had a well-defined vocabulary, syntax, and semantics. This system answered English questions of the kind that would be asked by a lunar geologist ("selenologist", if you will) about the chemical analyses of the Apollo 11 moon rocks. Its vocabulary and syntax were extensive, but its range of semantic concepts was extremely narrow. The results of our work in this context have been published in previous reports and papers. The system we developed was called SPEECHLIS; it was characterized by a semantics-first understanding strategy that used a semantic network to identify the underlying semantic relationships possible among matches of content words found in the utterance and then used

syntactic information to predict and fill in small function words and check the overall syntactic consistency of the utterance. The SPEECHLIS system dealt with a vocabulary of 250 words (drawn from the 3500 word vocabulary of the LUNAR system).

After implementing SPEECHLIS and using it to develop insights into the speech understanding process and the requirements for the different knowledge components, we began to consider also the extension of the techniques to a different semantic domain, one that would be more comprehensible to non-geologists and would encompass a more diverse semantic range. We were also looking for a domain in which aspects of dialog structure would be more apparent, thus providing additional opportunities for higher-level linguistic constraints to help resolve the low-level acoustic phonetic ambiguities. The domain that was chosen is that of trips and travel budgets. The system, dubbed HWIM for "Hear What I Mean", is intended to be an assistant to a travel budget manager, handling information about planned and taken trips, travel budgets and their status, plane fares and perdiems, and other information important to planning. In addition to such static information, the system is also intended to support the consideration of alternative courses of actions and their consequences. ("If I cancel the trip to St. Louis and only send two people to San Diego, will I still be over budget?"). This task was chosen as a small and easily comprehensible version of a generalized management problem. A very brief sample dialogue with the system is shown below, with the travel budget manager's input preceded by ">". (For a detailed description of the processing involved in this dialogue, see Vol. V, Sec. B4.)

>When did Bill go to Mexico?

Do you mean BILLY DISKIN, BILL HUGGINS, BILL LEVISON,
BILL PATRICK, BILL PLUMMER, BILL RUSSELL, BILL MERRIAM,
OR BILL WOODS?

>Bill Woods.

BILL WOODS' TRIP from BOSTON MASSACHUSETTS to JUAREZ
MEXICO was from October 15th, 1975 to October 17th.

B. DESIGN PHILOSOPHY OF HWIM

1. Theories, Monitors, Notices, and Events -
A Computational Framework for Perception

The BBN speech understanding system has evolved within a general framework for viewing perceptual processes. Central to this framework is an entity called a theory. A theory represents a particular hypothesis about some or all of the sensory stimuli that are present. Perception is viewed as the process of forming a believable coherent theory that can account for all the stimuli. In our framework, this is arrived at by successive refinement and extension of partial theories until a best complete theory is found.

In general, the perception process requires the ability to recognize any member of a potentially infinite class of perceptible objects that are constructed out of elementary constituents according to known rules. That is, the object perceived is generally a compound object, constructed from members of a finite set of elementary constituents according to some kind of well-formedness rules. These elementary constituents, as well as the relationships among them that are tested in the well-formedness rules, must be directly perceptible. Thus, a perceptual system must incorporate some basic epistemological assumptions about the things that it can perceive and the rules governing their assembly. The well-formedness rules can be used to reject impossible interpretations of the input stimuli and may also be usable to predict other constituents that could be present if a given partial theory were correct.

This perception framework assumes mechanisms for using subsets of the input stimuli to form initial "seed" hypotheses for certain elementary constituents (stimulus-driven hypothesization) and mechanisms for deriving hypotheses for additional elementary constituents that could be compatible with a given partial theory (theory-driven, or predicted, hypothesization). It also assumes mechanisms for verifying hypotheses against the input stimuli and evaluating the well-formedness of a compound hypothesis to assign it some measure of quality and/or likelihood. A theory may therefore be thought of as a hypothesis that has been evaluated in this way and assigned a measure of confidence.

In the case of speech understanding, a theory can range from an elementary hypothesis that a particular word is present at a particular point in the input (a word match) to a complete hypothesis of a covering sequence with a syntactic and semantic interpretation. (In general, a theory is a set of compatible word hypotheses with possible gaps between them and with partial syntactic and semantic interpretations.) A partial theory may be able to generate predictions for appropriate words or classes of words either adjacent to the words already hypothesized or possibly elsewhere in the utterance.

Predictions are dealt with in our computational framework by two kinds of devices: monitors, which are dormant processing requests passively waiting for expected constituents, and proposals, which are elementary hypotheses that are to be evaluated against the input. Proposals result in actively seeking stimuli that would verify them, while monitors passively wait for such hypotheses to be formed. The functioning of monitors assumes that there is an organizing structure into which all derived partial hypotheses are placed as they are discovered and that the monitors can essentially set "traps" for the kinds of events that they are watching for. This is to be contrasted with polling or continuous parallel evaluation of "demons" or other active processes to watch for expected patterns in the input stream. Monitors perform no computation unless and until some other process makes an entry of the kind they are waiting for in some data structure.

In our first speech understanding system, SPEECHLIS, for example, a word match for "concentration" would set monitors on the concept nodes for SAMPLE and CHEMICAL ELEMENT in the semantic network. If a word such as "Helium" were found anywhere else in the utterance, a check in the semantic network starting with Helium would lead to the superset category CHEMICAL ELEMENT where it would wake up the monitor from "concentration", thus detecting the coincidence of a detected hypothesis and a predicted hypothesis [Nash-Webber, 1975].

When a monitor is triggered, an event is created calling for the evaluation of a new hypothesis that, if acceptable, will become a new

theory. We refer to this process as noticing an event. In general, a number of events are competing for service by the processor at any moment. In human perception, there may be full parallel processing of such events, but in a serial machine, these events must be queued up and given processing resources on the basis of some priority ordering. (Even in human perception, there is probably some sort of priority allocation of resources, since various kinds of interference can occur.) In our computational framework, events are maintained on a queue in order of priority, the top event being processed at each step. The processing of an event can result in new proposals being made, new monitors being set, and existing monitors being triggered to produce new events. Since so much hinges on the event chosen for processing, a major control strategy issue is that of assigning priorities to events in order to find the most likely interpretation of the input without exhaustively exploring the full range of possible hypothetical interpretations.

2. Scoring Philosophy

A major concern in our speech understanding research has been the development of a uniform scoring philosophy for combining the scores from different knowledge sources. The search for a rational score-combining function began in SPEECHLIS, where we explored the issue by allowing the Control component to form a total score for a theory as an arbitrary combination of scores from several different knowledge sources. Initially, each knowledge source constructed scores in its own way. For example, the Syntactic component gave arbitrary score adjustments, such as +10 or -15 to particular theories that it considered to be using unlikely constructions, while the Word Matching component used a combination of individual phoneme scores plus extra points for certain consistency tests and demerits for inconsistencies. It was clear that in order to avoid comparing apples with oranges, we needed a principle for deciding how many points of syntactic score were worth one point of lexical matching score, how many points of lexical matching score were worth one point of verification score, etc. One might assume that a good overall theory score would be some weighted sum of the scores assigned by the different knowledge sources, but the issue of concern would still be choosing the correct weights. (CMU adopted

this weighted sum strategy in their Hearsay-II system, but chose weights essentially by trial and error.)

In HWIM, we have adopted a uniform scoring philosophy in which the score of each different knowledge source is statistically calibrated to a uniform scoring dimension. Here they are appropriately weighted to make the total score of a theory the sum of the scores assigned by the individual knowledge sources. The scoring dimension chosen is that of log likelihood ratios, and the interpretation of a theory's score is an estimate of the log probability of its being correct. The computation of this score involves a straightforward application of Bayes' theorem, deriving the probability of a theory T_1 given evidence E_j by the formula:

$$\Pr(T_1|E_j) = \Pr(E_j|T_1) \cdot \Pr(T_1) / \Pr(E_j).$$

That is, the probability of a theory given the evidence is equal to the probability of the evidence given the theory times the a priori probability of the theory, divided by the a priori probability of the evidence (i.e., the probability of this particular evidence occurring independent of T_1).

When there are several types of evidence, each from a different knowledge source, or several components of evidence within a single knowledge source, this equation can be factored into components:

$$\Pr(T_1|E_{j_1} \& E_{j_2} \& \dots \& E_{j_n}) = [\Pr(E_{j_1}|T_1)/\Pr(E_{j_1})] \cdot [\Pr(E_{j_2}|T_1)/\Pr(E_{j_2})] \dots [\Pr(E_{j_n}|T_1)/\Pr(E_{j_n})] \cdot \Pr(T_1)$$

under the assumption that the different pieces of evidence are independent. (A corresponding equation can be derived that accounts for any significant dependencies.) In HWIM, this equation is used in combining evidence both from different knowledge sources (e.g., acoustic-phonetic, verification, and prosodic), and also within any given knowledge source (e.g., each relevant phonetic segment in the segment lattice contributing evidence for a particular word hypothesis).

We assume then that the score assigned to a theory by a given knowledge source K is an estimate of the ratio $\Pr(E_{jk}|T_1)/\Pr(E_{jk})$, where E_{jk} is the evidence that K consulted in assigning its score. A ratio of 1

corresponds to essentially no information. That is, the evidence is as likely to occur by chance in any theory as it is for the particular theory T_1 . A ratio less than 1 indicates that it is more likely to occur for some arbitrary theory than it is for this particular T_1 . For the sake of computational efficiency, each ratio is represented in HWIM as 100 times its log so that scores can be combined by small integer addition rather than multiplication of floating point numbers. Positive logs thus represent theories that in some sense "account for" the evidence, while negative logs correspond to theories that are to some extent contraindicated. It is quite possible, theoretically, for a correct theory to have a negative score, but it would not be regarded as likely as another theory with a higher score.

A potential problem with using an estimate of the theory's probability as the criterion for choosing the preferred interpretation of an utterance is that if the a priori probabilities of the different T_1 's are significantly different, system performance will be excessively colored by an attempt to guess the most likely thing to be said, possibly overriding the acoustic evidence. In particular, if there are two interpretations A and B with the acoustic evidence favoring A by a probability factor of 50%, but the a priori probability favoring B by more than this, then B would be chosen rather than A. This would of course maximize the overall success rate of such a system if its a priori probabilities were correct. But such a system would run a major risk of picking likely utterances with poor acoustic evidence over unlikely utterances even with perfect acoustic evidence. To avoid the above problems and obtain a system whose performance is based on its ability to hear and understand rather than its ability to guess what will be said, we impose an assumption that all utterances are equally likely to occur, thus eliminating the a priori term $Pr(T_i)$ from our scores.

3. Calibrating a Knowledge Source

Using the above scoring philosophy, the score returned by a given knowledge source should be the log of an estimate of the probability ratio $\Pr(E_{jk}|T_1)/\Pr(E_{jk})$ where E_{jk} is the evidence it examined. To estimate that ratio, we use a process called calibration. Suppose a given knowledge source can compute an arbitrary parameter $f(T_1)$, which we believe to be correlated with the likelihood that T_1 is a correct interpretation. (Assume for the moment that $f(T_1)$ has a finite set of possible values or a range that can be divided into a finite set of regions.) Then given a suitable representative data base, we can calibrate that particular parameter by constructing a histogram of its values for correct theories and another histogram of its values for all theories on which it could be evaluated at all. For any particular value of $f(T_1)$ then, we can approximate $P(f(T_1)|T_1)/P(f(T_1))$ as the ratio of the number of times correct theories had that value over the number of times arbitrary theories had it. This gives us an appropriately calibrated contribution of this particular parameter $f(T_1)$ to the likelihood of T_1 being the correct interpretation. In the case of continuous valued f 's, refining the mesh of the histogram and increasing the number of samples used for calibration can make this estimate arbitrarily precise. Alternatively, other techniques (such as those used for vector modification, (Vol. II, Sec. 3)) can be used to estimate the probability ratios without the noise introduced by the histogram technique, gaining more precise estimates with smaller number of samples at the cost of a more complex computation at run time.

If a given knowledge source has a set of independent parameters, such as the $f(T_1)$ discussed above, then a total score for a theory can be constructed as the log of the product of the individual likelihood ratios (sum of the individual log likelihood ratios). If two such parameters are not independent, then their joint distributions can be measured by experiment and used for determining the appropriate likelihood ratio for the particular combination of values observed. In this way, any given knowledge source and any given measurement can be calibrated so that its scores are compatible with the scores of all other knowledge sources.

To illustrate the process, consider the calibration of the Verification component. This component measures a parameter that is essentially a spectral difference between the acoustic waveform present and that of a synthesized waveform for a particular word. To transform this score into a log likelihood ratio that can be added to the scores of other components, we divide the range of possible spectral distances into a finite number of intervals and construct histograms of the possible scores both for word verification requests that are known to be correct hypotheses and for all word verification requests independent of whether they are correct or not. To obtain these statistics, we actually run the speech understanding system in a "good only" mode, which follows only correct partial theories, but makes many incorrect word verification requests in the process. Using these two histograms, we can assign an appropriate log likelihood ratio for any given spectral distance. This procedure is described more fully in Vol. II, Sec. E.

In a similar way, we have calibrated prosodic evidence checked by the grammar, using error statistics for Lea's prosodic boundary detector (Vol. IV, Sec. C), but we have not yet been able to run experiments to check the effectiveness of its use. More detailed discussion of the probabilistic calibration of individual knowledge sources will be found in the sections discussing those components.

C. OVERVIEW OF FINAL SYSTEM

This section will present an overview of the final version (12 October) of the HWIM speech understanding system. We start with a presentation of its structure, followed by brief descriptions of its components, which are the implementations of the sources of knowledge. We conclude with an outline of the final control strategy for using these components to understand an utterance. In all cases, the descriptions of the system offered in this section will be previews of the more detailed descriptions that will be found in the remaining four volumes of this report.

1. The Structure of HWIM

The structure of the HWIM speech understanding system is shown in Figure 1. In this figure, components of the system are shown as boxes, and data structures are shown as ovals. Arrows illustrate flow of control, except in the lower right and lower center, where they illustrate flow of data. (Not shown are the control linkages between the Speech Control and the APR, PSA, and RTIME components.)

The HWIM system is implemented on TENEX, a virtual memory time-sharing operating system for the PDP-10 [Bobrow et al., 1972]. HWIM is made up of a number of individual TENEX processes (called "forks") in a single multiple-process job structure. Each box shown in Figure 1 is implemented as a separate TENEX process (RTIME and PSA are actually separate processes).

The reasons for this multiprocess structure derive from:

1. Differences in the implementation languages of the various components (TRIP, Speech Control, Syntax, and the Dictionary Expander are written in INTERLISP, Lexical Retrieval and APR in BCPL, and Verifier, PSA, RTIME, and TALKER in combinations of BCPL, Fortran, and machine language.);
2. Storage demands (The program and data structure requirements of all the components far exceed the 262,144-word size of a single TENEX address space.);
3. Historical circumstance (This is not as frivolous as it might appear. Work on all the knowledge source implementations began, and has continued, as individual stand-alone programs, with which their creators interact in the course of development and testing. A small amount of additional control logic for system linkage allows these programs to function as parts of the complete speech understanding

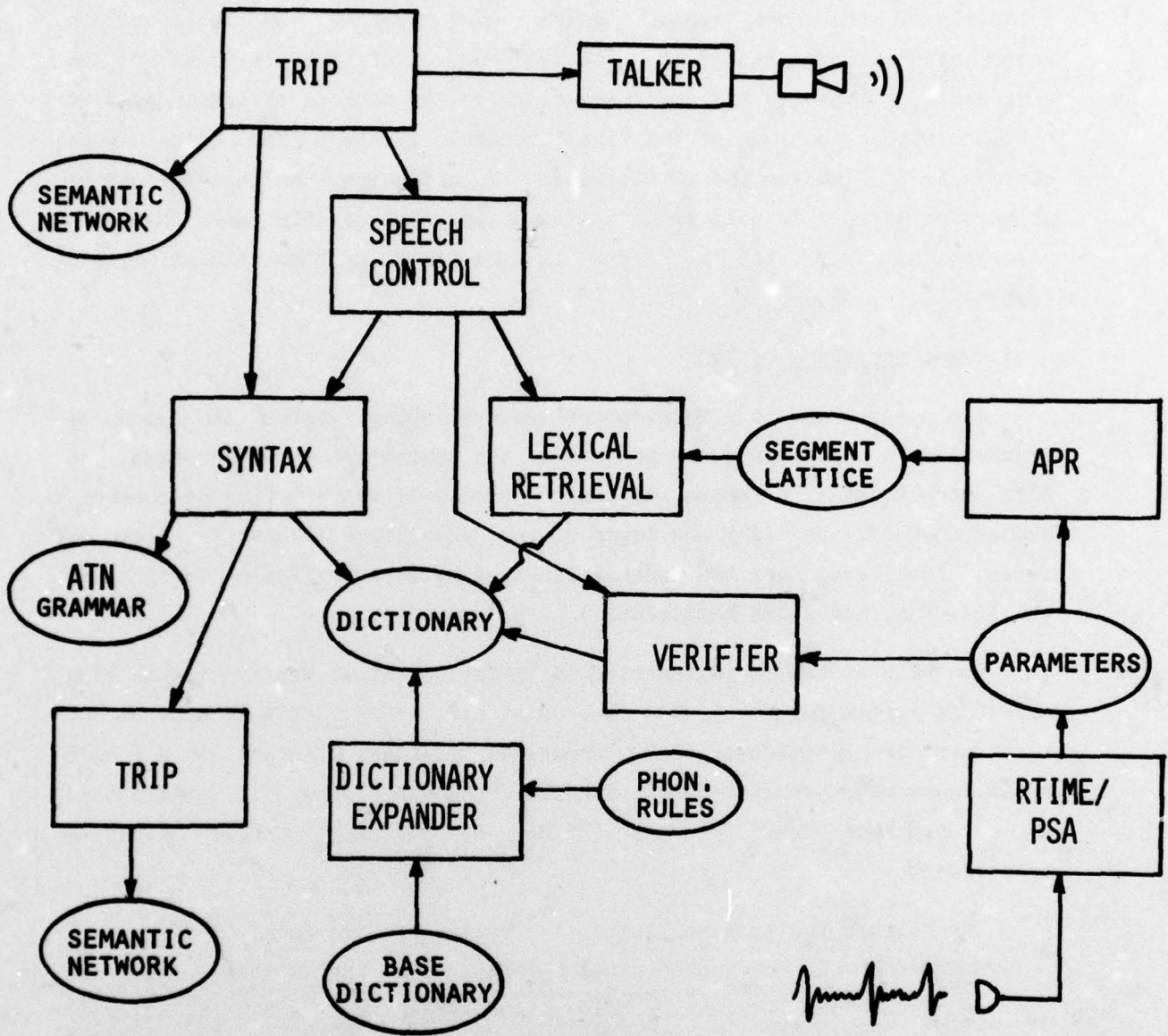


FIG. 1. HWIM SYSTEM ORGANIZATION.

system as well. In effect, the system is assembled from these components at run time.);

4. Convenience (As will be explained in Vols. 4 and 5, the TRIP component contains in its semantic network factual information useful to the parsing of an utterance. Rather than re-implement this information and its accessing routines in the Syntax component, we merely created a second instantiation of TRIP below the Syntax component, so that Syntax can communicate with it. The reason for a second instantiation rather than the same one is so that the Speech Control system can be run without the top TRIP fork.)

This convenience of implementation is not without its price in terms of run-time efficiency (see Vol. III, Sec. F).

2. The Components of HWIM

The TRIP component serves as the overall system controller, the data base manager and retriever, and the embodiment of HWIM's factual and discourse level knowledge sources. TRIP uses a semantic network to represent the travelers, destinations, meetings, trips, budgets, etc. that it must model. If it is presented with a typed request, it calls Syntax directly to parse and semantically interpret the request. If the request is spoken, then TRIP invokes the Speech Control component, passing to it discourse level expectations and receiving from it the semantic interpretation of the utterance. This interpretation, which is in a language akin to predicate calculus, is then manipulated by TRIP in its information retrieval role to carry out the user's request. The following are some examples of semantic interpretations:

Utterance: Give me a list of the untaken trips.
 Interpretation: (FOR: ALL A0001/ (FINDQ: DB/TRIP (TIME (AFTER NOW)))
 : T ; (OUTPUT: A0001))
 [English paraphrase: "For every member of the class of trips that start later than now, print it out."]

Utterance: How much is in the Speech Understanding budget?
 Interpretation: (FOR: THE A0046 / (FINDQ: DB/BUDGET
 (PROJECT (SPEECH UNDERSTANDING))) : T ;
 (OUTPUT: (GET: A0046 MONEY/REMAINING)))
 [English paraphrase: "Find the single entity that is the budget for the speech understanding project, and print out its amount of remaining money."]

Utterance: Who is going to IFIP?
 Interpretation: (FOR: THE A0014 / (FINDQ: DB/CONFERENCE (SPONSOR IFIP))
 : T ; (FOR ALL A0015 / (FINDQ: DB/TRIP (TO/ATTEND A0014))
 : T ; (OUTPUT: (GET: A0015 TRAVELER])
 [English paraphrase: "Find the single conference whose sponsor
 is IFIP. For each trip to attend that conference,
 print out its traveler."]

TRIP's response to the user can be either typed or spoken, as appropriate. For the latter, TALKER, a speech synthesis program, is invoked. (TRIP and TALKER are described in Vol. V; TALKER is further discussed in Vol. II, Sec. C.)

The Speech Control component directly and explicitly controls the remaining components of the system for the purpose of understanding utterances. We will defer description of the control strategy until Sec. C.3.

The Dictionary Expander, although not a part of the running speech understanding system, is employed at system loadup time to expand the dictionary of baseform pronunciations by means of phonological rules to encompass all predictable pronunciation variants. This also includes producing the data structure that enables the Lexical Retrieval component to apply across-word phonological rules during the matching process. (In fact, not one, but several versions of dictionary are produced, each specialized to the uses made of it.) The current dictionary starts with 1138 base words, which are expanded to 1363 words by means of regular inflections, of which 1097 words are currently known to the grammar. Some of these 1363 words have more than one pronunciation baseform (1789 in all), and these are expanded into 8642 pronunciations by the application of the phonological rules. (See Vol. III, Secs. A, B, and C.)

RTIME is the component used to acquire (digitize and store) the speech signal, and PSA is the signal processing component, which converts the speech signal into a parametric representation. This parametric form, used by the APR and Verification components, consists primarily of overall energy, energies in low-, mid-, and high-frequency bands, fundamental frequency, the first 3 formant frequencies, zero crossing rate, a spectral shape measure, and a 13-pole LPC spectral model (see Vol. II, Sec. A).

The Acoustic-Phonetic Recognizer (APR) operates on the parametric representation of the utterance with acoustic-phonetic rules to produce a bottom-up phonetic hypothesis called the segment lattice, illustrated in Figure 2 below.

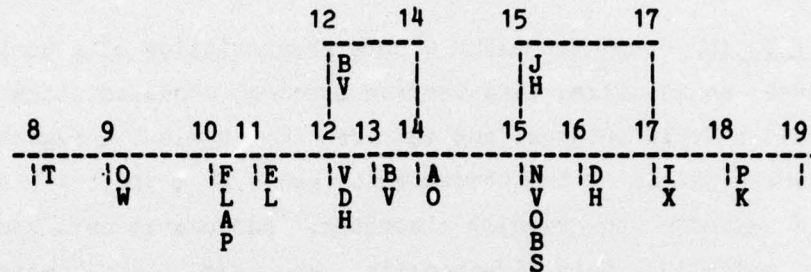


Fig. 2. Segment lattice fragment for the words "total budget". Time runs horizontally from left to right. The numbers denote segment boundaries, and the characters denote segment labels.

The segment lattice allows for the representation of alternate segmentation paths where the acoustic evidence is not sufficiently unambiguous to permit unique segmentation decisions at this level. Each segment label, although shown as a single mnemonic label for the convenience of the experimenter, is in fact a vector of likelihood ratios L_j , one for each phoneme, giving an estimate of the likelihood of each phoneme P_j in light of the acoustic evidence. That is,

$$L_j = \frac{\text{Pr}(\text{acoustic evidence} \mid \text{phoneme}_j \text{ is correct})}{\text{Pr}(\text{acoustic evidence})}$$

These likelihood ratios are derived from performance statistics of the APR operating on a data base of utterances by five male speakers (see Vol. II, Sec. B).

The Lexical Retrieval component, which embodies the lexical and phonological knowledge sources, is used to match dictionary word pronunciations against the segment lattice in two ways: (1) without regard to context, or (2) "anchored" off a previously found word match or set of word matches, in order that the effects of adjacent words be taken into account. In either case, it can seek either all words or only those in a specified set. It uses a tree-structured dictionary in order to do the

matching and across-word phonological rule application most efficiently. It combines the segment lattice scores using Bayes' Rule to estimate the probability of the word match given the observed signal (see Vol. III, Sec. D).

The Verification component takes a word pronunciation plus context (if any), generates an idealized time-varying spectral representation using a speech synthesis-by-rule program, and compares that against a region of the parametric representation of the utterance by means of a spectral distance metric and a dynamic programming algorithm. Because it uses generative rather than analytic acoustic-phonetic knowledge and matches a pronunciation against the parameters rather than the segment lattice, it provides a word matching capability largely independent of the Lexical Retriever, but at a rather higher computational cost (see Vol. II, Sec. D).

The Syntax component embodies syntactic, semantic, and pragmatic knowledge sources, and by invoking a second instantiation of the TRIP component, accesses the factual knowledge source as well. Its parser is built around a "pragmatic grammar", which accepts only those utterances that are grammatical in the usual sense, meaningful in the travel budget management domain, and appropriate to the pragmatic circumstance of a single speaker talking to a computer data management system and given the previous conversation. In this grammar, the usual NP, VP, and PP constituents are replaced by structures such as "meetings", "trips", and "budget items". This specialization considerably increases the predictive power of the grammar, since it does not permit phrases like "her next workshop", which, although formed from words in the HWIM vocabulary and acceptable in normal English grammar, is not appropriate to this task domain. The functions of the Parser are (1) to judge the grammaticality of a given word sequence, (2) to predict the possible extensions of a hypothesized word sequence, and (3) to build up a formal representation of the utterance, in this case, the semantic representation described earlier. It can parse utterances starting at any point in the utterance and working in both directions thereafter (see Vol. IV, Secs. A and B).

(A Prosodics component, not shown in Figure 1 because it was not included in the 12 October system, was also worked on during this contract. That work is described in Vol. IV, Sec. C).

3. LHSDVP: The Left-Hybrid, Shortfall Density, Verify-At-Pick Control Strategy

The basic notions of control strategy, and in particular, the notion of a central Control component that explicitly directs the process of hypothesis formation and extension, are derived from our early experiences with incremental simulation of speech understanding systems described in Sec. A. In HWIM, this Control component uses the other components much like subroutines, making all decisions concerning resource allocation.

The control strategy outlined below for using the knowledge source components is the one employed in the final performance testing of the HWIM system, but it is by no means the only one implemented within HWIM. Because we have viewed as a central task the exploration of different control strategies, the Control component itself contains about 25 flag variables that control as many strategy options. (Vol. III, Sec. E of this report describes many of these options, and Vol. III, Sec. F describes some experiments in strategy variations that were performed with HWIM.) We will first describe HWIM's basic control strategy and then the specific options employed in the final performance test.

The basic form of all of HWIM's present control strategies may be stated concisely as follows:

1. Having acquired, parameterized, and formed a segment lattice from the utterance, use Lexical Retrieval to scan the lattice for words that match well, independent of context. Place these seed matches, ordered by score, on an event queue.
2. Select the top-scoring event(s) from the queue and present it as a theory (or hypothesis) to the Parser.
 - a) If the theory is a complete sentence and spans the utterance, the Parser returns its semantic interpretation. This is accepted as the best model of the utterance (but in principle, the process could continue to find the second best interpretation, etc.).
 - b) If the theory is grammatically unacceptable, it is rejected and step 2 is repeated.
 - c) Otherwise, the Parser makes proposals for all words and semantic categories that are possible at each end of the theory.

3. Give the proposals to Lexical Retrieval for anchored scans from each end of the theory, taking into account any across-word phonological effects required by the words at each end. For each such word found, form a new word event, which is a one-word extension of the previous theory. Score each new event and place it in the proper position on the event queue.
4. Possibly rescore, remove, or add events on the queue, due to interactions between new and old events. (For instance, if the new word of an event is the same and is in the same position as a new word in another event, but is in the opposite direction, then form a new collision event, made up of the union of the word matches of the two "colliding" events. Also, if an event hits a possible end of the utterance, make an end event, which makes that hypothesis; an end event may include an ending score penalty if it's not at a definite end.)
5. Return to Step 2.

This cycle of syntactic proposals, scans for the words proposed in the context of the words already found, and formation of new events continues until an event is processed that spans the utterance, is acceptable to the grammar, and is consistent with the observed speech signal. Obviously, the number of such cycles (or new theories) required must be at least the number of words in the sentence plus 2 (for the end events), but such an optimum is rarely achieved. In some cases, the number of cycles required is on the order of 2 times the number of words; in others the number is much larger. (See the final performance results in Vol. III, Appendices 3 and 4.)

Among the variations possible on such a strategy are constraints on the directions that theories are allowed to grow, the use of the Verification component, and the method of scoring the events to determine their positions on the event queue.

If the control strategy operates in a general, "middle-out" fashion, in which seed word matches are sought anywhere in the utterance, then there may ensue quite a lot of resource competition among events in different regions of the utterance. This often leads to a large number of theory extensions (due to different competing theories) before a spanning theory is formed. The left-hybrid strategy is one way to cut down on this. Seed words are sought stemming only from possible left end boundaries of the utterance, and only those words that could occur on the left end of a sentence are sought. The resulting theories are constrained to grow leftwards until they become left-ended; only then are they allowed to grow

rightwards. An additional advantage is the additional syntactic constraint that comes from parsing sentence fragments that have a definite left end, as opposed to the more general case of sentence fragments anywhere in the middle of a sentence.

The Verification component yields a word match score that is largely independent of that given by Lexical Retrieval. However, Verification is not employed lightly, due to its high computational cost. Although one would like to verify each new word match found by Lexical Retrieval, in order to bring as much knowledge as possible to bear on the ordering of the event queue, such a strategy would involve excessive computation. The "Verify-at-Pick" strategy selects for verification (at the beginning of Step 2 above) only those events from the top of the event queue that are necessary to ensure that the top of the queue is properly ordered. That is, it goes down the queue, verifying new events and rescoreing them as necessary, only as far as necessary to be sure it hasn't missed one that could get promoted to the top of the queue by virtue of its verification score.

In a system such as HWIM, there need to exist two distinct scoring philosophies: (1) A measure of quality, which expresses how likely a partial hypothesis is, and (2) a measure of priority, which is used to determine the order in which partial hypotheses are to be processed. These two types of scores could be the same (and indeed, in earlier versions of HWIM, they were), but this does not necessarily lead to the best interpretation in the shortest time. In HWIM at this time, the quality-score is an estimate of the probability of the particular phoneme, word, or word sequence, given the observed speech signal. The objective is to find the most likely acceptable word sequence, given the observed signal. The priority-score used for the final performance test is a measure called shortfall density. Briefly, it consists of scoring word matches and word match sequences not by their lexical-matching (quality) scores, but by the extent to which their quality scores fall below a summed per-segment upper bound scoring function. This difference, called the shortfall, is then divided by the duration, to give a shortfall density. The items on the event queue are ordered by increasing shortfall density,

rather than by decreasing quality score. Scoring methods, including shortfall density, are discussed along with other control strategy issues in Vol. III, Sec. E.

In summary, the control strategy employed for HWIM's final performance testing was the basic strategy described at the beginning of this section, using:

1. the Left-Hybrid policy to constrain the formation of seed events and their extension,
2. the Verify-at-Pick method of employing the Verification component, and
3. the Shortfall Density method of computing priority scores for ordering the event queue.

Explaining the way that HWIM uses its various sources of knowledge to understand an utterance is a little bit like explaining how to high jump; the principles are few, but there are many practical details that are difficult to describe. Perhaps some "practical experience" in how HWIM understands utterances can be gained from Appendix 2 of this volume, which contains a trace listing of how HWIM handled the utterance, "Do we have a surplus?". The trace listing shows all of the steps described above, and it is extensively annotated.

D. FINAL PERFORMANCE TEST OF HWIM

The final task in BBN's speech understanding research project was to make a meaningful test of its performance on understanding sentence length utterances. For this purpose, we recorded a new set of 124 utterances spoken by the three male speakers from whom previous utterances had been solicited and processed these utterances with the latest version (12 October) of HWIM. This section summarizes these results.

The test set comprised two tokens each of 62 sentences. Thirty-one of these sentences had been used in test utterances processed by earlier versions of HWIM, and the others were entirely new sentence types. The sentences were chosen to fall within the limitations of a smaller grammar and vocabulary, so that we could use them to test the system on task domains of two sizes, the 409-word MIDGRAM/TRAVELDICTION and the 1097-word BIGGRAM/BIGDICTION.* One token of each of the 62 sentences was recorded by speaker WAW, and speakers JJW and RMS recorded one token of each of the 62 between them, 31 for each. The utterances range in length from 3 to 13 words; the average length is 6 words or 1.8 seconds. The test sentences are listed in Appendix 1, this volume.

These 124 new utterances were processed by both versions of HWIM using the LHSVDP strategy (left-hybrid strategy, using Shortfall Density scoring and the Verification component in verify-at-pick mode) (see Sec. C). The system performance is summarized in Table 1.

1. Scores

In Table 1, the term "Semantically correct" refers to utterances whose word list was not entirely correct, but whose semantic interpretation, as generated by the Parser, was identical to that of the correct sentence. This is regarded as "correctly understood" in the sense of Newell et al. (1973).

*The two figures correspond to the number of entries in TRAVELDICTION and BIGDICTION that are accessible to MIDGRAM and BIGGRAM, respectively.

	BIGDICT (1097 words)	MIDDICT (409 words)
Correctly understood	54 = 44%	65 = 52%
All words correct	51 = 41%	59 = 48%
Semantically correct	3 = 2%	6 = 4%
Incorrect	45 = 36%	40 = 32%
Close to correct	29 = 23%	25 = 20%
Less correct	13 = 10%	10 = 8%
Very wrong	3 = 2%	5 = 4%
No response	25 = 20%	19 = 15%
Gave up (150 theories)	24 = 19%	18 = 14%
System broke	1 = 1%	1 = 1%
Estimated average branching (words)	196	67
Speed (times real time)	1350	1050

Table 1. Summary of final performance results,
124 utterances by three speakers.

For example:

Utterance: How much is left?
Understood: What is left?
Semantic interpretation:
(FOR: ALL A0007 / (FINDQ: DB/BUDGET)
: T ; (OUTPUT: (GET: A0007 MONEY/REMAINING)))

(That is, "for each budget output the amount of money currently remaining there.")

Of the incorrectly understood utterances, most were quite close to the correct sentence. For example:

Utterance: Show me Bill's trip to Washington.
Understood: Show only Bell's trip to Washington.
(i.e., Alan Bell, rather than Bill Woods.)

Utterance: The registration fee is twenty dollars.
Understood: Their registration fee is twenty dollars.

Others were farther away from the correct sentence. For example:

Utterance: Give me a list of the untaken trips.
Understood: Give Bill's fifty untaken trips.

For some utterances, the system failed to make a response, as it was set to give up if it had processed as many as 150 theories without building a spanning theory. In all but a few such cases, at that point there was no hope that HWIM would correctly understand the utterance. In two other cases, the system was unable to proceed because of a bug in some component.

We feel that the 44% "correctly understood" performance is a low estimate of HWIM's capabilities. Development of the system continued right up to three weeks from the end of the contract, and although this final version of the system was tested on a variety of utterances, time prevented the more thorough testing that we would have preferred. Consequently, the 124 utterance final test runs uncovered many bugs, most of which did not bring the system to a halt, but definitely interfered with the process of finding the correct interpretation. Naturally, no changes could be made to the system during the test without invalidating the results already obtained and compromising the "new" status of the 124 utterance test set.

2. Estimated Average Branching

Wherever at the grammar's request the Lexical Retrieval component is called to scan the segment lattice for a set of proposed words and classes, the number of words possible at that point is noted. This number ranges from over 900, at the left end of the sentence, to 1, in cases where the context allows only a single word to follow. The "estimated average branching (words)" in Table 1 is the arithmetic mean of these branching factors.

3. Speed

The speed of the system is shown in Table 1 in terms of "number of times real time", i.e., the PDP-10 CPU time for the run divided by the duration of the sentence, for those utterances that were correctly or incorrectly understood. The time required to understand an utterance is divided among the various components of the system in the following rough proportions:

RTIME/PSA	5%
APR	0
Lexical Retrieval	40
Verification	15
Parser	20
Control	20

The results for the smaller, 409-word task domain show an increase in the correctly understood utterances and a corresponding decrease in the incorrect and "no response" cases, but the improvement is far from

dramatic. This suggests that the factor of 2.7 in vocabulary size (or the corresponding factor of 3 in average word branching) represents a much smaller change in the "difficulty" imposed by the task domain.

More details of the final performance testing may be found in Vol. III, Sec. F., and a tabulation of the results for each utterance may be found in Vol. III, Appendix 3.

E. FUTURE

1. Where We Are and How We Got Here

As mentioned in the preceding section, the performance of the current HWIM system is not a true indicator of its potential. This final report marks the expiration of the original five-year program; it does not in any sense represent closure or completeness in the system. During the last five months of the contract, performance increased at a rate of approximately 10% per month from a baseline of near zero as the individual components of the system began to come together, bugs and inconsistencies were ironed out, and new variations of control strategies and scoring algorithms were developed. There is no reason to expect that this trend of increasing performance has reached its apex. On the contrary, there is every reason to believe that considerable improvements remain to be made. For example, during the final weeks of the contract, the 1097 word vocabulary was run for the first time, containing approximately 600 words that had never been seen before by the system. The final test run was made on this vocabulary with no alteration of the individual word pronunciations from the form in which they were copied from a standard pronunciation reference [Kenyon and Knotts, 1944]. These words introduced never before seen acoustic-phonetic environments into the system, with the result that the APR component was functioning using an incomplete set of rules for the phenomena that it was being called on to label. The fact that the system could perform at all with its vocabulary more than doubled without any feedback is evidence for the robustness of the techniques being used. In fact, our performance decreased only slightly with the increased vocabulary.

There are many other places in the system where individual components are functioning far below their potential capabilities. The major thrust of our research program has been towards developing effective techniques for handling the speech understanding problem for multiple speakers without individual training of the system to a speaker. This pursuit has been a major evolutionary effort, and modification and improvement of individual components has been sufficiently continuous that none of the components has

been tuned to its full potential. For example, pronunciation likelihoods used in the dictionary base forms and in the phonological rules are intuitive human ones, not statistically measured estimates. Moreover, while a method of checking the compatibility between syntactic hypotheses and intonational information in the speech waveform has been implemented (Vol. IV, Sec. C), it has not yet been tested, much less tuned. A third source of untapped potential comes from the pronunciation features associated with individual word pronunciations. These are intended to be checked by the grammar to verify that the pronunciation of a word is consistent with a hypothesis about its use, but few places in the grammar have actually been annotated to make such checks, and the mechanism by which this information is made available to the parser has not yet had all the bugs ironed out of it. In general, the current performance can only be taken as an intermediate benchmark in a system that is still in a state of rapid development.

The most significant conclusions that can be drawn from the current performance are that a system with a vocabulary of over 1000 words and a grammar with a branching ratio of 196, with no tuning of the system to individual speakers and no discourse information being used, can function at approximately the 50% sentence understanding level. Detailed analysis of our performance record (Vol. III, Appendix 3) will show that the speech confusions that are made are generally reasonable ones, many of which can be corrected by techniques already under development. Some of them could be legislated away by highly artificial grammars designed to eliminate speech confusions. HWIM's grammar makes no such attempts, and in fact permits such minimal difference pairs as "What is the registration fee" and "What is their registration fee" (where, it should be noted, one of the pronunciations of "their" is DH EH IO R, and an across-word gemination rule will merge the R with that of "registration"). It should also be pointed out that the system has been set up to use discourse context to resolve such pairs by using information such as whether there is a potential antecedent for the pronoun "their". However, the final demonstration test was performed on isolated sentences with the discourse model modified to return positive answers to any such questions. Thus, the expected

performance of the system in context, even with no improvements in the system would exceed that of the test conditions.

2. What We've Learned

In the course of our research, we believe that we have made significant advances in the technology required for continuous speech understanding. From an initial state of trial and error approximations and incremental simulations of human spectrogram readers, we have developed a number of theoretically interesting and demonstrably effective techniques. These include a general uniform scoring method for combining the scores of different knowledge sources (Sec. B), an efficient technique for incorporating generative phonological rules into a dictionary both within and across word boundaries (Vol. III, Sec. C), the use of pragmatic ATN grammars for combining syntactic, semantic, pragmatic, and prosodic information (Vol. IV, Sec. B), an efficient bi-directional parsing algorithm for ATN grammars that permits them to be used in middle-out parsing without sacrificing predictive capability (Vol. IV, Sec. A), a synthesis-by-rule Verification component that improves both performance and cost effectiveness by comparing word hypotheses with the input signal at the parametric level (Vol. II, Sec. D), and several overall control strategies, some of which can be guaranteed to find the best possible interpretation of an utterance without systematically enumerating all possibilities (Vol. III, Sec. E).

We have found that the vocabulary size problem does not seem to be as significant as one would expect. On the other hand, although we have made significant progress, we have not yet gained closure on the set of acoustic-phonetic and phonological effects that are necessary to handle new vocabulary items and new speakers without specific training. Consequently, we do not know what the eventual performance of the system would be if that were achieved. We have discovered a number of effective control strategies with interesting and useful theoretical capabilities, but there remain many strategy variations that we have not yet had the opportunity to try.

3. Continued Research Required

There are many obvious lines of research arising immediately from our current work that need to be pursued.

(1) The most pressing need is a continued extension of the acoustic-phonetic rules embedded in the APR component and a correction of bugs and inconsistencies uncovered in our test run (Vol. I, Sec. D; Vol. III, Sec. F). This should be followed by an additional testing phase to take a better reading on the potential for the techniques used.

(2) System tuning can account for a significant increase in the performance of such a system; an experiment is required to determine how much the system can be improved by thorough tuning.

(3) The data base on which the APR computes its statistics is currently much too small and needs to be extended. Similarly, the strategy variation experiments (described in Vol. III, Sec. F) were performed on test sets much too small to support any firm conclusions; these need to be repeated for larger test sets.

(4) All of the mechanisms have been implemented for testing the use of prosodic cues, but relevant experiments have not been tried.

(5) Experiments to test the effectiveness of the discourse model in improving system performance in context need to be set up and carried out.

(6) The search for improved control strategy options should be continued.

(7) More thorough testing of pronunciation features and pronunciation likelihoods needs to be done.

These are all examples of required research arising directly from our current work. In addition, there are related research problems that were not focused on in the current ARPA program that need to be pursued. These include dealing with the lower quality of telephone speech, uncooperative speakers, female speakers, speakers with significant dialect differences, dynamic acquisition of a model of an arbitrary speaker, noisy environments,

and fluent natural English syntax (although some of these problems have been considered to some extent within the current five-year program). These problems were recognized in the speech study group report (Newell et al., 1971) to be of sufficient difficulty that they would require additional research and would be prime candidates for future research, should the results of the first five year program appear promising.

Other problems not specifically envisioned in the speech study group report, but which now appear worthy of attention, are the synthesis of speech responses as part of a total speech understanding system and the integration of speech understanding and speech compression techniques as a possible technology for the telephone speech problem (and also for the very low bandwidth speech compression problems).

4. Future Speech Understanding Systems

The HWIM system is a large and cumbersome research vehicle, consuming significant computational resources on a fairly large and expensive computer. Future speech understanding systems that would arise from this work will operate on machines with even greater computational capabilities, but which are much smaller and cheaper than those available today. In Vol. III, Sec. F, a detailed projection is given for real-time implementation of such a system on hardware that is estimated to be available by 1980 at a cost of approximately \$63,500.

Future speech understanding systems will contain a large number of phonological rules dealing with different dialects and will be able to listen to a speaker for a few sentences and automatically adapt themselves to his speech. Precursors of such a capability are already present in HWIM. For example, HWIM computes an estimate of the speaker's vocal tract length from each utterance and uses this for certain speaker normalization transformations that are made in the APR component. Pronunciation features on lexical pronunciations can be used to indicate pronunciations produced by dialect dependent rules (the rules in HWIM can introduce such features), and the understanding of a sentence using such pronunciations can be used as evidence to increase the likelihood of subsequent pronunciations using the same dialect.

Future systems will contain sophisticated models of the task domain and the ongoing conversation so that pragmatic information can be used to resolve otherwise unreconcilable acoustic ambiguities. HWIM contains such capabilities in rudimentary form, but our ability to model conversations and codify pragmatic inference rules is still quite limited.

Future systems will permit the user to talk to a microphone placed somewhere near him while leaving him free to move around, unlike existing systems that require a fixed microphone placement with respect to the mouth. This will require modeling of room acoustics and reverberation effects not currently considered in HWIM.

Finally, future systems will make effective use of prosodic information both to resolve ambiguities and to eliminate much of the searching of false hypotheses. Our current level of understanding of prosodic rules is extremely limited.

In the very near future, much more limited systems will become available for use. These systems will have carefully chosen task-specific vocabularies and grammars that minimize the potential for phonetic ambiguity and will be tuned for an individual speaker by requiring him to utter samples of the vocabulary items. The computational demands of such systems, however, will not be as great as those of HWIM, and the machines necessary to perform the task will become available sooner and at less cost. To use such a system, a speaker will probably insert a cassette on which his pronunciations of the vocabulary items are stored and wear a headset that holds the microphone in a fixed position with respect to his mouth (perhaps with noise-canceling capabilities). Such systems will be useful for data entry, traffic monitoring, postal sorting, programming and debugging, information retrieval, computer-aided decision-making, and many other applications that cut across the spectrum of man-machine communication.

In conclusion, we believe that, eventually, speech understanding systems will become as common a piece of technology as the telephone and television are today.

References

- [1] Bobrow, D.G., J.D. Burchfiel, D.L. Murphy and R.S. Tomlinson (1972)
"TENEX, a Paged Time Sharing System for the PDP-10," CACM, Vol. 15, No. 3.
- [2] Chapanis, A. (1975)
"Interactive Human Communication," Scientific American, Vol. 232, No. 3, pp. 36-42.
- [3] Kenyon, J.S. and T.A. Knotts (1944)
A Pronouncing Dictionary of American English, G.C. Merriam Co., Springfield, MA.
- [4] Klatt, D.H. and K.N. Stevens (1973)
"On the Automatic Recognition of Continuous Speech: Implications from a Spectrogram-Reading Experiment," IEEE Trans. Audio Electroacoust. Vol. AU-21.
- [5] Nash-Webber, B.L. (1975)
"The Role of Semantics in Automatic Speech Understanding," in Representation and Understanding: Studies in Cognitive Science, D. Bobrow and A. Collins (eds.), Academic Press.
- [6] Newell, A., J. Barnett, J. Forgie, C. Green, D. Klatt, J.C.R. Licklider, J. Munson, R. Reddy, and W. Woods (1971)
Speech Understanding Systems: Final Report of a Study Group. Reprinted by North-Holland/American Elsevier, Amsterdam.
- [7] Ochsman, R.B. and A. Chapanis (1974)
"The Effects of 10 Communication Modes on the Behavior of Teams During Co-operative Problem-Solving," Int. J. Man-Machine Studies, 6, 579-619.
- [8] Vicens, P. (1969)
"Aspects of Speech Recognition by Computer," Ph. D. Thesis, Computer Science Dept., Stanford University.
- [9] Woods, W.A., R.M. Kaplan and B.L. Nash-Webber (1972)
"The Lunar Sciences Natural Language Information System: Final Report," BBN Report No. 2378, Bolt Beranek and Newman Inc., Cambridge, MA (available from NTIS as N72-28984).
- [10] Woods, W.A. and J. Makhoul (1973)
"Mechanical Inference Problems in Continuous Speech Understanding," Proc. Third Int'l. Joint Conference on Artificial Intelligence, pp. 200-207. (Reprinted in Artificial Intelligence, Vol. 5, No. 1, pp. 73-91, Spring, 1974).

Appendix 1 - Sample Set of Sentence Types

This appendix lists a sample of the sentence types acceptable to the HWIM speech understanding system. In fact, this list shows the 62 sentence types used in the final performance test described in Sec. D. These sentences are acceptable to both MIDGRAM and BIGGRAM.

- 102. What is the total budget figure?
- 106. List all trips to California this year.
- 110. What is the registration fee?
- 119. Who's going to IFIP?
- 125. When is the next ASA meeting?
- 126. How much have we already spent?
- 137. How much is left?
- 178. Show me a list of the remaining trips.
- 180. What is the registration fee for the next ACL conference?
- 183. Enter a trip for Bonnie Nash-Webber to Grenoble.
- 185. Please show me John Makhoul's three trips to Pittsburgh.
- 270. When did Craig go to Utah?
- 272. What is the plane fare to Ottawa?
- 275. Add a new budget item.
- 277. The registration fee is twenty dollars.
- 278. List the remaining untaken trips.
- 279. Which trips were canceled?
- 280. How many trips are there?
- 283. Cancel Lyn's trip to the ASA meeting.
- 285. Give me a list of the untaken trips.
- 292. Schedule a trip by train to New York.
- 312. What is the one way air fare from Boston to London?
- 315. Create a trip.
- 318. Show his trip.
- 320. Print her fare.
- 323. Show me her trips.
- 330. Was the trip expensive?

330. Do we have a surplus?
334. Why did he visit SDC?
335. Who went to IFIP?
339. Are we over the budget?
345. Show me Bill's trip to Washington.
346. When is the ACL conference?
347. Figure his expenses.
348. Display all taken trips.
349. I will fly to San Diego.
350. When is Chip's trip?
351. List all fares to Chicago.
352. How much did Geoff spend?
353. List all remaining trips.
354. Add a trip to France.
355. How many trips has Rich taken?
356. Who went to Los Angeles this year?
357. When is Jack going to San Francisco?
358. What is the round-trip fare to Chicago?
359. Who went to Santa Barbara in August?
360. Does the speech budget have a surplus?
361. Why did Jerry go to Univac?
362. How much did we spend in December?
365. Schedule a trip for Jack Klovstad to SDC.
366. How much money is in the current budget?
367. The robot budget has 7 K.
369. The trip number is 4 3 7 6.
370. What trips remain in the speech budget?
371. My trip to Mexico cost eight hundred dollars.
372. Please give me the actual cost of trip number 6 8 7 3.
373. What is the cost of my trip?
374. How much is in the speech understanding budget?
375. Where is the next ASA meeting?
376. Lyn's trip to France cost six hundred dollars.
377. Show me all trips to El Paso.
378. Enter a trip to Grenoble.

Appendix 2 - Trace Listing Generated by HWIM

This appendix contains a sample trace output of the kind normally generated by the system in the course of understanding an utterance. The utterance is named RMS333, and it is a token of the sentence "Do we have a surplus?" spoken by Richard Schwartz. The trace itself is shown on pages 42 to 72. The segment lattice for the utterance is shown on pages 73 to 77 in two forms: first, graphically, and then in a tabular form that shows the seven top scoring phonemes with their log likelihood scores. Page 78 is a summary of this run of the system.

12_LOAD (<SPSYS>FLAGS.LHSDVP) ←
FILE CREATED 14-Oct-76 22:24:05
FLAGSCOMS
(V: CONTROLFLAGS)
<SPSYS>FLAGS.LHSDVP;2

Load file of flags that determine the strategy, in this case, Left Hybrid, using Shortfall Density scoring, and Verify-at-Pick.

16_SYNTAX-DRIVEN (RMS333 Y) ←
<NAKHOUL>RMS333.INFO;1:

Call the speech understander on utterance RMS333, using segment lattice version Y.

Utterance: RMS333
Speaker: Rich Schwartz
Sentence: Do we have a surplus?
Record date: 1 Oct 76
Final test utterance. Recorded in the Imlac room, temporary doors, PDP11 off. Read. 2nd of 2 tokens. Fits MIDGRAM.
UTT: (DO WE HAVE A SURPLUS)

Name of "info" file, which contains the following information about the utterance:

RMS333 Y 17-Oct-76 14:02:08 Used 4 in 6 seconds. BBN-TENEXC
System, Group load avgs = 1.21 .814 Control-fcrk page faults = 371

Hatcher = <KLOVSTAD>CLASSLEX.SAV;13
Dictionary = <KLOVSTAD>DICTSTRUCT.BIGDICT-115;2
Verifier = <COOK>NEWHAPPER.SAV;47
Synfork = <SPSYS>SYNFORK.BIGDICT-115;12
Running on BBN-TENEXC

} Names and version numbers of the files that form the subsidiary processes of the HWIM system.

LEFT, RIGHT ENDS: 0/1/5 13/17/17 ←

Lexical retrieval returns the ending boundaries:

SHORTFALLFLAG = T
DENSITYFLAG = T
PRUNEFLAG = NIL
OTHERDIRFLAG = NIL
UNSHLPFLAG = T
GHOSTFLAG = T
CHOOSEDIRFLAG = NIL
BESTSPANFLAG = T
QUICKFLAG = NIL
GHOSTPSLSFLAG = T
ENDCHECKFLAG = T
CREDITFLAG = NIL
LIABILITYFLAG = NIL
PRONLIKEFLAG = T
PAUSECLAMP = 20
COLLISIONFLAG = NIL
CONTEXTFLAG = T
LEFTENDFLAG = NIL
KILLBADEVTSFLAG = T
GOODONLYFLAG = NIL

0 = leftmost nonpenalty left end
1 = rightmost " " "
5 = " possible " "
and similarly for the right end.

Values of the flags that control the strategy options.

PROSODYFLAG = NIL
VERIFYFLAG = VERIFY@PICK
VERIFYDEFAULT = 60
HYBRIDFLAG = 0
MAXTHRYN = 150
MINSTARTS = 0

} A few more flags

Initial scan -- left-to-right:
Rescan from 0 to 4

} Scan across entire utterance to form the MAXSEG profile

Doing hybrid initial scan
Initial scan from left end of utterance:
Left end branching factor = 947

Scan for seeds at left end only.
There are 947 words possible at the left end.

Creating seed event:
88 YOUR 2-3 -94 -16 (REDUCED) L

Creating seed event:
92 WORK 3-7 -14 0 -- L

Creating seed event:
96 WILL 3-7 -53 0 -- L
93 WILL 3-6 -24 -31 (REDUCED) L
74 WILL 1-4 -27 -16 (REDUCED) L

Here begins a listing of the "seed" words found in the initial scan, in reverse alphabetical order. Some seeds consist of a single word match; others contain several. (These are called fuzzy word matches.)

Creating seed event:
95 WHY 3-6 -24 0 -- L

Creating seed event:
69 WHOSE 1-6 5 0 -- L

Creating seed event:
70 WHOLE 1-4 1 0 -- L
32 WHOLE 1-5 -4 0 -- L

Creating seed event:
62 WHO 1-5 182 0 -- L
59 WHO 0-2 -184 0 -- L

Creating seed event:
94 WHERE 3-6 -24 0 -- L

Creating seed event:
98 WHEN 3-6 -77 0 -- L

Creating seed event:
99 WHAT 3-6 -78 -16 (REDUCED) L

Creating seed event:
91 WERE 3-6 -10 0 -- L

← A word match printout contains the following information:
95 the word match index
WHY the word
3-6 left and right boundary numbers
-24 lexical match score
0 pronunciation likelihood score
-- pronunciation features (none in this case)
L signifies a left-to-right match

; <RESULTS>SN7-12OCT/BIGDICT-115/LHS DVP.RMS333;3 Fri 3-Dec-76 2:14PM

Creating seed event:

90 WE 3-6 128 0 -- L
78 WE 2-6 -35 0 -- L

Creating seed event:

104 WAS 3-7 -129 -16 (REDUCED) L

Creating seed event:

72 THE 1-3 -7 -16 (REDUCED) L
82 THE 2-6 -78 -23 -- L

Creating seed event:

124 THAT 5-9 -95 -31 (REDUCED) L
110 THAT 4-9 -95 -31 (REDUCED) L

Creating seed event:

79 SHE 2-6 -38 -31 (REDUCED) L
102 SHE 3-6 -94 -31 (REDUCED) L

Creating seed event:

123 SEVEN 5-10 -94 -10 -- L
109 SEVEN 4-10 -94 -10 -- L

Creating seed event:

108 OUR 4-6 -91 -23 -- L
122 OUR 5-6 -91 -23 -- L

Creating seed event:

76 OUR 2-5 4 -23 -- L
86 OUR 2-3 -88 -31 (REDUCED) L

Creating seed event:

58 OUR 0-1 -183 -23 -- L

Creating seed event:

117 ON 4-7 -123 -16 (REDUCED) L
131 ON 5-7 -123 -16 (REDUCED) L

Creating seed event:

55 OHIO 0-5 -135 0 -- L

Creating seed event:

112 OH 4-6 -101 0 -- L
126 OH 5-6 -101 0 -- L

Creating seed event:

101 LYN 3-7 -92 0 -- L

Creating seed event:

97 LINA 3-7 -68 0 -- L

Creating seed event:

89 IT 2-4 -96 -16 (REDUCED) L

Creating seed event:

119 IRAQ 5-9 7 0 -- L

105 IRAQ 4-9 7 0 -- L

Creating seed event:

84 TN 2-4 -86 -16 (REDUCED) L

Creating seed event:

113 I 4-6 -101 0 -- L

127 I 5-6 -101 0 -- L

Creating seed event:

31 HDW 1-5 3 0 -- L

Creating seed event:

68 HIS 1-4 6 -16 (REDUCED) L

Creating seed event:

67 HER 1-3 15 -16 (REDUCED) L

71 HER 1-5 -1 -31 (REDUCED) L

Creating seed event:

81 HE 2-6 -67 0 -- L

103 HE 3-6 -123 0 -- L

Creating seed event:

63 HE 1-3 57 0 -- L

Creating seed event:

64 HAWAII 1-6 17 0 -- L

Creating seed event:

80 ERIE 2-6 -59 0 -- L

100 ERIE 3-6 -90 0 -- L

Creating seed event:

106 ERIC 4-7 -37 -16 -- L

120 ERIC 5-7 -37 -16 -- L

Creating seed event:

16 ELEVEN 5-10 84 -40 -- L

15 ELEVEN 4-10 84 -40 -- L

Creating seed event:

107 EACH 4-7 -58 0 -- L

121 EACH 5-7 -58 0 -- L

Creating seed event:

66 DO 1-5 17 0 -- L
65 DO 1-5 17 -31 (REDUCED) L
53 DO 0-5 -66 0 -- L
73 DO 1-3 -14 -31 (REDUCED) L
52 DO 0-5 -66 -31 (REDUCED) L
54 DO 0-3 -97 -31 (REDUCED) L

Creating seed event:

57 DELETE 0-7 -153 -31 -- L

Creating seed event:

60 DAYTON 0-5 -187 -10 -- L

Creating seed event:

63 BURTON 0-5 -187 -10 -- L

Creating seed event:

56 BILLY 0-6 -137 0 -- L

Creating seed event:

85 ARE 2-3 -88 -16 (REDUCED) L

Creating seed event:

132 ANOTHER 5-10 -132 0 -- L
119 ANOTHER 4-10 -132 0 -- L

Creating seed event:

116 AN 4-7 -123 -16 (REDUCED) L
130 AN 5-7 -123 -16 (REDUCED) L

Creating seed event:

83 AN 2-4 -78 -16 (REDUCED) L

Creating seed event:

111 AIR 4-6 -101 0 -- L
125 AIR 5-6 -101 0 -- L

Creating seed event:

115 AI@LAB 4-9 -116 -31 -- L
129 AI@LAB 5-9 -116 -31 -- L

Creating seed event:

114 A 4-6 -101 -16 (REDUCED) L
128 A 5-6 -101 -16 (REDUCED) L

Creating seed event:

75 A 2-3 14 -16 (REDUCED) L
77 A 2-5 -22 -31 (REDUCED) L

Creating seed event:

87 -PAUSE- 2-5 -92 0 -- L

Creating seed event:

51 -PAUSE- 0-1 20 0 -- L

MAXSEGS: This is the MAXSEG profile. The region from 0-1 has a MAX of 20, 1-2 has a MAX of 43, etc.

1 20
2 43
3 75
4 21
5 43
6 73
7 33
8 63
9 32
10 23
11 73
12 30
13 36
14 30
15 36
16 51
17 63

The initial event queue is shown below. All events at this point are seed events, each a single word.

QUAL = quality score (lexical match)
DUR = duration in frames (centiseconds)
SHORT = shortfall score
SD = shortfall density
NLX = "non-lexical" score = prosodic & verification & ending scores

* signifies that the event has not yet been verified and that a default score (of 60 points) will be added to the NLX score shown as an estimated upper bound

SCORE = composite score that determines queue position
REGION = left and right boudaries of the event.

Sorting EVENTQUEUE

RMS333 Y 17-Oct-76 14:11:36 Used 327 in 574 seconds. BBN-TENEXC
System, Group load avgs = 2.34 1.37 Control-fcrk page faults = 3151

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	20	3	0	0	0*	20	0-1	-PAUSE- !
2	128	14	9	.643	0*	3.64	3-6	WE ! ← correct seed
3	182	17	0	0	0*	3.53	1-5	WHO (! indicates a word
4	57	11	61	5.55	0*	-.0909	1-3	HE sequence that is in the
5	-2	7	77	11	0*	-2.43	2-3	A ! utterance, according to
6	44	37	180	4.86	0*	-3.24	5-10	ELEVEN the info file. Usually
7	7	31	194	6.26	0*	-4.32	5-9	IRAQ it's correct but there's
8	-1	11	119	10.8	0*	-5.36	1-3	HER no check that it's in
9	1	13	138	10.6	0*	-6	1-4	WHOLE the right
10	17	17	165	9.71	0*	-6.18	1-5	DO ! ← correct seed place.)
11	-10	14	147	10.5	0*	-6.21	3-6	WERE
12	-14	19	184	9.68	0*	-6.53	3-7	WORK
13	-10	13	149	11.5	0*	-6.85	1-4	HIS
14	3	17	179	10.5	0*	-7	1-5	NOW
15	17	25	238	9.52	0*	-7.12	1-6	HAWAII
16	-24	14	161	11.5	0*	-7.21	3-6	WHERE
17	-24	14	161	11.5	0*	-7.21	3-6	WHY
18	-104	37	328	8.86	0*	-7.24	5-10	SEVEN
19	-23	11	141	12.8	0*	-7.36	1-3	THE
20	-19	13	158	12.2	0*	-7.54	2-5	OUR

21	5	25	250	10	0*	-7.6	1-6	WHOSE
22	-132	37	356	9.62	0*	-8	5-10	ANOTHER
23	-53	17	202	11.9	0*	-8.35	4-7	ERIC
24	-53	19	223	11.7	0*	-8.58	3-7	WILL
25	-126	31	327	10.5	0*	-8.61	5-9	THAT
26	-58	17	207	12.2	0*	-8.65	4-7	EACH
27	-68	19	238	12.5	0*	-9.37	3-7	LIMA
28	-147	35	391	11.2	0*	-9.46	4-9	AIOLAB
29	-59	21	271	12.9	0*	-10	2-6	ERIE
30	-67	21	279	13.3	0*	-10.4	2-6	HE

plus 24 additional sorted and 0 unsorted events.

Verscore = N.A. / (14 17)
for 51 -PAUSE- 0-1 20 0 -- L
in -PAUSE- !
New EVT.SCORE = 0.0

Enough of the top events on the event queue (3 in this case) are given to verification to ensure that the top of the queue is properly ordered.

Verscore = (715 36 28 44) / (28 42)
for 90 WE 3-6 128 0 -- L
in WE !
New EVT.SCORE = 1.928571

Verscore = (765 60 12 30) / (17 34)
for 62 WHO 1-5 182 0 -- L
in WHO
New EVT.SCORE = 3.529412

The format of the Verification results is:

- 765 = spectral distance score (not used)
- 60 = log likelihood score
- 12,30 = left, right times found by the Verifier
- 17-34 = left, right times found by Lexical Retrieval (from the segment lattice)

About to do event 1: WHO

Looking for
62 WHO 1-5 182 0 -- L
from right to left

The seed WHO is the first event considered by Syntax.

Doing syntax evaluation of theory.

CREATING THEORY#1
 SCORES = (182 10.70588 0 0.0 (0 60 0))
 62 WHO 1-5 182 0 -- L
 133 WHO 1-5 182 0 -- R
 134 WHO 1-3 15 0 -- R
 135 WHO 1-4 -12 0 -- R
 59 WHO 0-2 -184 0 -- L

The format of a theory's scores is:
182, 10.7 = quality, quality density
0, 0.0 = shortfall, shortfall density
(0 60 0) = prosodic, verification, ending scores

DOING PROPOSALS:

On the left:
Words: -PAUSE-

These are the proposals made by Syntax on both ends of the event WHO

Branching factor is 1

DOING PROPOSALS:

On the right:
Words: ATTEND-ED PLEW IS TRAVEL-ED VISIT-ED WENT WILL

Branching factor is 7

Since the left boundary of this event is at a non-penalty left end, we ask Syntax to consider making it a left end.

Trying left end event.

Doing left end event for THEORY#1 creating THEORY#2
[WHO
SCORES = (182 10.70588 0 0.0 (0 60 0.0))

The [signifies that
the theory includes a
left end hypothesis.

Reusing right proposals. ← Syntax makes the same proposals on the right

Noticing word event:

- 140 WILL 2-7 -51 0 -- L (62 . -79)
- 139 WILL 2-6 -22 -31 (REDUCED) L (62 . -79)
- 138 WILL 2-5 -7 -26 (REDUCED) L (62 . -79)
- 142 WILL 2-6 -112 -31 (REDUCED) L (62 . -79)
- 149 WILL 1-5 -90 -26 (REDUCED) L (59 . -411)
- 148 WILL 1-4 -87 -16 (REDUCED) L (59 . -411)
- 147 WILL 1-4 -146 -16 (REDUCED) L (62 . -79)
- 143 WILL 2-3 -116 -26 (REDUCED) L (62 . -79)
- 150 WILL 1-2 -144 -26 (REDUCED) L (59 . -411)

These "event notices"
show the words found by
Lexical Retrieval to the
right of [WHO.

following:

- 62 WHO 1-5 182 0 -- L
- 133 WHO 1-5 182 0 -- R
- 134 WHO 1-3 15 0 -- R
- ** et alia **

forming:

[WHO WILL

Noticing word event:

- 146 WENT 2-9 -134 -31 -- L (62 . -79)
- 144 WENT 2-7 -116 0 -- L (62 . -79)
- 141 WENT 2-6 -75 -31 -- L (62 . -79)

← The dotted pairs shown on these
word matches show the word
adjacency constraints found by
Lexical Retrieval. (62 . -79)
means that it is adjacent only
to word match number 62, and
a score of -79 must be added
to account for the word-ending
effects.

following:

- 62 WHO 1-5 182 0 -- L
- 133 WHO 1-5 182 0 -- R
- 134 WHO 1-3 15 0 -- R
- ** et alia **

forming:

[WHO WENT

Noticing word event:

- 145 IS 2-7 -127 0 (CONTRACT STRIPRIC) L (62 . -79)

following:

- 62 WHO 1-5 182 0 -- L
- 133 WHO 1-5 182 0 -- R
- 134 WHO 1-3 15 0 -- R
- ** et alia **

forming:

[WHO IS

RMS333 Y 17-Oct-76 14:13:52 Used 379 in 710 seconds. BBW-TENEXC
System, Group load avgs = 1.56 1.27 Control-fcrk page faults = 4623

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	128	14	9	.643	36	1.93	3-6	WE !
2	70	17	112	6.59	60*	.471	1-5>	[WHO WILL
3	20	3	0	0	0	0	0-1	-PAUSE- !
4	57	11	61	5.55	0*	-.0909	1-3	HE
5	-2	7	77	11	0*	-2.43	2-3	A !
6	44	37	180	4.86	0*	-3.24	5-10	ELEVEN
7	7	31	194	6.26	0*	-4.32	5-9	IRAQ
8	-1	11	119	10.8	0*	-5.36	1-3	HER
9	1	13	138	10.6	0*	-6	1-4	WHOLE
10	17	17	165	9.71	0*	-6.18	1-5	DO !
11	-10	14	147	10.5	0*	-6.21	3-6	WERE
12	-24	30	312	10.4	60*	-6.4	1-7>	[WHO IS
13	-14	19	184	9.68	0*	-6.53	3-7	WORK
14	-62	48	445	9.27	60*	-6.77	1-9>	[WHO WENT
15	-10	13	149	11.5	0*	-6.85	1-4	HIS
16	3	17	179	10.5	0*	-7	1-5	HOW
17	17	25	238	9.52	0*	-7.12	1-6	HAWAII
18	-24	14	161	11.5	0*	-7.21	3-6	WHERE
19	-24	14	161	11.5	0*	-7.21	3-6	WHY
20	-104	37	328	8.86	0*	-7.24	5-10	SEVEN
21	-23	11	141	12.8	0*	-7.36	1-3	THE
22	-19	13	158	12.2	0*	-7.54	2-5	OUR
23	5	25	250	10	0*	-7.6	1-6	WHOSE
24	-132	37	356	9.62	0*	-8	5-10	ANOTHER
25	-53	17	202	11.9	0*	-8.35	4-7	ERIC
26	-53	19	223	11.7	0*	-8.58	3-7	WILL
27	-126	31	327	10.5	0*	-8.61	5-9	THAT
28	-58	17	207	12.2	0*	-8.65	4-7	EACH
29	-68	19	238	12.5	0*	-9.37	3-7	LIMA
30	-147	35	391	11.2	0*	-9.46	4-9	AI@LAB

plus 26 additional sorted and 0 unsorted events.

About to do event 1: WE !

Looking for

90 WE 3-6 128 0 -- L ←

Doing syntax evaluation of theory.

Now the event queue shows the 3 new events properly ordered with respect to the previous events.

Angles at end of the region specification indicate direction of the event, -- i.e. the end at which a new word is being added.

The seed event WE was verified above, so no call to Verification here.

There weren't any right-to-left matches for WE, so we scan for some before continuing. A seed not at an end must have word matches in both directions, or else it can't grow in both directions.

CREATING THEORY#3

SCORES = (128 9.142857 9 .6428571 (0 36 0))

90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R
78 WE 2-6 -35 0 -- L
153 WE 2-5 -20 0 -- R

Syntax makes proposals on the left and right, but the left-hybrid strategy permits only left-going events to be noticed until a left end hypothesis is included.

DOING PROPOSALS:

On the left:

Words: -PAUSE- ARE CAN DID DO DOES HAVE WERE WILL WOULD

Branching factor is 10

DOING PROPOSALS:

On the right:

Words: -PAUSE- AFFORD ALLOW ALREADY ARE ARRANGE-ED ATTEND
ATTEND-ED ATTEND-ING BE BUDGET-ED CANCEL-EE DEPART FLEW
FLY FLY-ING GO GO-ING HAVE IN LEAVE LEFT MAKE OVER OVER-BUDGET-ED
PLAN PLAN-ED RECENTLY REMAIN-ING SCHEDULE SCHEDULE-ED SPEND
SPLIT START TAKE TAKEN TRAVEL TRAVEL-ED TRAVEL-ING UNDER
UNDER-BUDGET-ED VISIT VISIT-ED VISIT-ING WANT WENT WERE
WILL

Branching factor is 48

Noticing word event:

154 DO 1-4 63 -31 (REDUCED) R (152 . -77)
164 DO 1-5 17 0 -- R (151 . -79)
171 DO 1-5 -82 0 -- R (153 . -148)
165 DO 1-6 -122 -31 (REDUCED) R (151 . -79)
158 DO 1-4 -73 -31 (REDUCED) R (152 . -77)

preceding:

90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R

** et alia **

forming:

DO WE !

Noticing word event:

155 ARE 2-4 -11 -16 (REDUCED) R (152 . -77)
162 ARE 3-4 -102 -16 (REDUCED) R (152 . -77)

preceding:

90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R

** et alia **

forming:

ARE WE

Noticing word event:

170 -PAUSE- 0-2 -52 0 -- R (153 . -148)
161 -PAUSE- 2-4 -101 0 -- R (152 . -77)

preceding:

90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R
** et alia **

forming:

-PAUSE- WE

Already noticed: WE !

Noticing word event:

159 HAVE 1-4 -74 -16 (REDUCED) R (152 . -77)
163 HAVE 1-4 -105 0 -- R (152 . -77)

preceding:

90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R
** et alia **

forming:

HAVE WE

Noticing word event:

167 WERE 2-6 -159 0 -- R (151 . -79)
157 WERE 2-4 -69 0 -- R (152 . -77)
169 WERE 1-5 -182 0 -- R (151 . -79)

preceding:

90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R
** et alia **

forming:

WERE WE

Noticing word event:

156 WILL 2-4 -39 -26 (REDUCED) R (152 . -77)
166 WILL 2-6 -146 -26 (REDUCED) R (151 . -79)
168 WILL 1-5 -179 -26 (REDUCED) R (151 . -79)

preceding:

90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R
** et alia **

forming:

WILL WE

Noticing word event:

160 CAN 2-4 -96 -26 (REDUCED) R (152 . -77)

preceding:

90 WE 3-6 128 0 -- L
 152 WE 3-6 128 0 -- R
 151 WE 2-6 130 0 -- R
 ** et alia **

forming:

CAN WE

Noticing left end event

for: [WE !

Endscore = -.8110236

WE is at a possible left end boundary, but there's an ending penalty associated with the hypothesis that we end there. The end score is divided by the utterance duration, and a new ending event is formed and put back on the event queue.

Sorting EVENTQUEUE

Found better MAXSEG -- recomputing shortfalls

MAXSEGS:

1 20
 2 43
 3 75
 4 21
 5 43
 6 73
 7 33 1
 8 63 19
 9 32 9
 10 23
 11 73
 12 30
 13 36
 14 30
 15 36
 16 51
 17 63

In the course of matching words on the right of the previous event, new word matches were found that cause the system to increase MAXSEGS, our estimate of the upper bound, for the segments ending 7,8,9. This then requires recomputing shortfall scores for events covering this region and possibly reordering the event queue.
 (1, 19, 9' are the increases to be added to 33, 63, 32 respectively.)

 RMS333 Y 17-Oct-76 14:16:13 Used 460 in 851 seconds. BBN-TENEXC
 System, Group load avgs = 1.5 1.24 Control-fork page faults = 5288

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	100	21	132	5.74	36*	1.72	<2-6	ARE WE
2	159	25	116	4.3	36*	1.54	<1-6	DO WE !
3	128	14	9	.643	-67	1.12	3-6	[WE !
4	62	21	170	7.39	36*	.898	<2-6	WILL WE
5	70	17	112	6.59	60*	.471	1-5>	[WHO WILL
6	20	3	0	0	0	0	0-1	-PAUSE- !
7	57	11	61	5.55	0*	-.0909	1-3	HE
8	5	21	227	9.87	36*	-.342	<2-6	CAN WE

9	37	25	238	8.81	36*	-.9	<1-6	HAVE WE
10	-27	21	380	10.9	36*	-1.96	<2-6	WERE WE
11	-2	7	77	11	0*	-2.43	2-3	A !
12	45	37	208	5.65	0*	-4.03	5-10	ELEVEN
13	7	31	223	7.19	0*	-5.26	5-9	IRAQ
14	-1	11	119	10.8	0*	-5.36	1-3	HER
15	1	13	138	10.6	0*	-6	1-4	WHOLE
16	17	17	165	9.71	0*	-6.18	1-5	DO !
17	-10	14	147	10.5	0*	-6.21	3-6	WERE
18	-24	30	313	10.4	60*	-6.43	1-7>	[WHO IS
19	-14	19	185	9.74	0*	-6.58	3-7	WORK
20	-10	13	149	11.5	0*	-6.85	1-4	HIS
21	3	17	179	10.5	0*	-7	1-5	HOW
22	17	25	238	9.52	0*	-7.12	1-6	HAWAII
23	-24	14	161	11.5	0*	-7.21	3-6	WHERE
24	-24	14	161	11.5	0*	-7.21	3-6	WHY
25	-23	11	141	12.8	0*	-7.36	1-3	THE
26	-62	48	474	9.87	60*	-7.37	1-9>	[WHO WENT
27	-19	13	158	12.2	0*	-7.54	2-5	OUR
28	5	25	250	10	0*	-7.6	1-6	WHOSE
29	-104	37	357	9.65	0*	-8.03	5-10	SEVEN
30	-53	17	203	11.9	0*	-8.41	4-7	ERIC

plus 32 additional sorted and 0 unsorted events.

Verscore = N.A. / (21 30) ← A verification score of "N.A." means
 for 155 ARE 2-4 -11 -16 (REDUCED) R (152 . -77) "not applicable," because
 in ARE WE the event was too short
 New EVT.SCORE = -1.13354 to verify.

Verscore = (823 67 15 31) / (17 30) ← DO gets a very good verification
 for 154 DO 1-4 63 -31 (REDUCED) R (152 . -77) score of 67, so the
 in DO WE ! event DO WE gets to the
 New EVT.SCORE = 1.82 top of the queue and
 gets processed.

About to do event 1: DO WE !

CREATING THEORY#4 FROM THEORY#3

SCORES = (159 5.888889 116 4.296296 (0 103 0))

154 DO 1-4 63 -31 (REDUCED) R (152 . -77)

164 DO 1-5 17 0 -- R (151 . -79)

171 DO 1-5 -82 0 -- R (153 . -148)

** et alia **

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

DOING PROPOSALS:

On the left:

Words: -PAUSE- MONEY MUCH WHAT

Branching factor is 4

Again,

Syntax makes proposals to the left and right, but we only consider the ones to the left.

DOING PROPOSALS:

On the right:

Words: HAVE

Branching factor is 1

Noticing word event:

183 WHAT 0-1 -147 -31 (REDUCED) R (158 . 0) (165 . 0) (171 . 0) (164 . ** 0)

preceding:

154 DO 1-4 63 -31 (REDUCED) R (152 . -77)

164 DO 1-5 17 0 -- R (151 . -79)

171 DO 1-5 -82 0 -- R (153 . -148)

** et alia **

forming:

WHAT DO WE

DO WE is at a nonpenalty left end, so we ask Syntax to consider ending it now instead of creating an ending event.

Trying left end event.

Doing left end event for THEORY#4 creating THEORY#5

[DO WE !

SCORES = (159 5.888889 116 4.296296 (0 103 0.0))

Reusing right proposals.

Noticing word event:

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0) Lexical Retrieval.

following:

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

forming:

[DO WE HAVE !

Sorting EVENTQUEUE

RMS333 Y 17-Oct-76 14:18:18 Used 546 in 976 seconds. BBN-TENEXC
System, Group load avgs = 3.99 1.68 Control-fcrk page faults = 5940

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	128	14	9	.643	-67	1.12	3-6	[WE !
2	314	48	115	2.3	103*	1.1	1-9>	[DO WE HAVE !
3	62	21	170	7.39	36*	.898	<2-6	WILL WE
4	70	17	112	6.59	60*	.471	1-5>	[WHO WILL
5	20	3	0	0	0	0	0-1	-PAUSE- !
6	57	11	61	5.55	0*	-.0909	1-3	HE
7	5	21	227	9.87	36*	-.342	<2-6	CAN WE
8	-39	28	372	10.9	103*	-.687	<0-6	WHAT DC WE
9	37	25	230	8.81	36*	-.9	<1-6	HAVE WE
10	100	21	132	5.74	36	-1.13	<2-6	ARE WE
11	-27	21	300	10.9	36*	-1.96	<2-6	WERE WE
12	-2	7	77	11	0*	-2.43	2-3	A !
13	45	37	200	5.65	0*	-4.03	5-10	ELEVEN
14	7	31	223	7.19	0*	-5.26	5-9	IRAQ
15	-1	11	119	10.8	0*	-5.36	1-3	HER
16	1	13	138	10.6	0*	-6	1-4	WHOLE
17	17	17	165	9.71	0*	-6.18	1-5	DO !
18	-10	14	147	10.5	0*	-6.21	3-6	WERE
19	-24	30	313	10.4	60*	-6.43	1-7>	[WHO IS
20	-14	19	185	9.74	0*	-6.58	3-7	WORK
21	-10	13	149	11.5	0*	-6.85	1-4	HIS
22	3	17	179	10.5	0*	-7	1-5	HOW
23	17	25	238	9.52	0*	-7.12	1-6	HAWAII
24	-24	14	161	11.5	0*	-7.21	3-6	WHERE
25	-24	14	161	11.5	0*	-7.21	3-6	WHY
26	-23	11	141	12.8	0*	-7.36	1-3	THE
27	-62	48	474	9.87	60*	-7.37	1-9>	[WHO WENT
28	-19	13	158	12.2	0*	-7.54	2-5	OUR
29	5	25	250	10	0*	-7.6	1-6	WHOSE
30	-104	37	357	9.65	0*	-8.03	5-10	SEVEN

The top event is the ending event [WE made 3 pages before this. Its NLX score includes both a positive Verification score and a negative ending score.

plus 33 additional sorted and 0 unsorted events.

About to do event 1: [WE !

 Doing left end event for THEORY#3 creating THEORY#6
 [WE !
 SCORES = (128 9.142857 9 .6428571 (0 36 -.8110236))

DOING PROPOSALS:

On the right:

Words: ARE ATTEND-ED FLEW TRAVEL-ED VISIT-ED WANT WENT WERE
WILL

Branching factor is 9

Noticing word event:

174 WENT 6-9 33 -31 -- L (78 . 0) (90 . 0)

188 WENT 6-9 -163 0 -- L (78 . 0) (90 . 0)

following:

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

forming:

[WE WENT

Noticing word event:

175 ARE 4-6 -33 -16 (REDUCED) L (78 . 0) (90 . 0)

186 ARE 6-8 -150 0 -- L (78 . 0) (90 . 0)

187 ARE 6-7 -150 -16 (REDUCED) L (78 . 0) (90 . 0)

following:

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

forming:

[WE ARE

Noticing word event:

184 WERE 6-8 -122 0 -- L (78 . 0) (90 . 0)

following:

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

forming:

[WE WERE

Noticing word event:

185 WILL 6-8 -136 -31 (REDUCED) L (78 . 0) (90 . 0)

following:

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

forming:

[WE WILL

Sorting EVENTQUEUE

 RMS333 Y 17-Oct-76 14:20:05 Used 600 in 1083 seconds. BBN-TENEXC
 System, Group load avgs = 4.26 1.93 Control-fcrk page faults = 6570

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	314	48	115	2.3	103*	1.1	1-9>	[DO WE HAVE !
2	62	21	170	7.39	36*	.898	<2-6	WILL WE
3	70	17	112	6.59	60*	.471	1-5>	[WHO WILL
4	20	3	0	0	0	0	0-1	-PAUSE- !
5	57	11	61	5.55	0*	-.0909	1-3	HE
6	5	21	227	9.87	36*	-.342	<2-6	CAN WE
7	80	14	174	6.69	-67*	-.646	3-6>	[WE ARE
8	-39	28	372	10.9	103*	-.687	<0-6	WHAT DC WE
9	37	25	238	8.81	36*	-.9	<1-6	HAVE WE
10	100	21	132	5.74	36	-1.13	<2-6	ARE WE
11	-27	21	380	10.9	36*	-1.96	<2-6	WERE WE
12	-2	7	77	11	0*	-2.43	2-3	A !
13	130	37	154	4.43	-67*	-2.65	3-9>	[WE WENT
14	45	37	208	5.65	0*	-4.03	5-10	ELEVEN
15	7	31	223	7.19	0*	-5.26	5-9	IRAQ
16	-1	11	119	10.8	0*	-5.36	1-3	HER
17	6	31	247	7.97	-67*	-5.68	3-8>	[WE WERE
18	1	13	138	10.6	0*	-6	1-4	WHOLE
19	17	17	165	9.71	0*	-6.18	1-5	DO !
20	-10	14	147	10.5	0*	-6.21	3-6	WERE
21	-24	30	313	10.4	60*	-6.43	1-7>	[WHO IS
22	-14	19	195	9.74	0*	-6.58	3-7	WORK
23	-10	13	149	11.5	0*	-6.85	1-4	HIS
24	3	17	179	10.5	0*	-7	1-5	HOW
25	17	25	238	9.52	0*	-7.12	1-6	HAWAII
26	-39	31	292	9.42	-67*	-7.13	3-8>	[WE WILL
27	-24	14	161	11.5	0*	-7.21	3-6	WHERE
28	-24	14	161	11.5	0*	-7.21	3-6	WHY
29	-23	11	141	12.8	0*	-7.36	1-3	THE
30	-62	48	474	9.87	60*	-7.37	1-9>	[WHO WENT

plus 36 additional sorted and 0 unsorted events.

Verscore = (758 60 47 64) / (42 65)

for 172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)
 in [DO WE HAVE !

New EVT.SCORE = 1.095833

HAVE gets a good Verification
 score of 60.

About to do event 1: [DO WE HAVE !

CREATING THEORY#7 FROM THEORY#5

SCORES = (314 6.28 115 2.3 (0 163 0.0))

154 DO 1-4 63 -31 (REDUCED) R (152 . -77)

164 DO 1-5 17 0 -- R (151 . -79)

171 DO 1-5 -82 0 -- R (153 . -148)

** et alia **

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

DOING PROPOSALS:

On the right:

Syntax makes these proposals on the right.

Words: A LESS MORE

Classes: DIGITS NUM-ADJ TEENS TENS

Branching factor is 36

Noticing word event:

189 A 8-10 96 -31 (REDUCED) L (172 . -93)

following:

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

forming:

DO WE HAVE A !

Noticing word event:

190 MORE 8-10 -14 0 -- L (172 . -93)

198 MORE 8-13 -157 0 -- L (172 . -93)

following:

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

forming:

DO WE HAVE MORE

Noticing word event:

192 OVER 8-13 -47 0 -- L (172 . -93)

196 OVER 8-12 -137 0 -- L (172 . -93)

following:

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

forming:

DO WE HAVE OVER

Noticing word event:

191 UNDER 8-13 -43 -31 -- L (172 . -93)

195 UNDER 8-12 -133 -31 -- L (172 . -93)

197 UNDER 8-11 -157 -31 -- L (172 . -93)

following:

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

forming:

DO WE HAVE UNDER

; <RESULTS>SH7-12OCT/BIGDICT-!15/LHSDVP.RMS333;3 Fri 3-Dec-76 2:14PM

Noticing word event:

193 LESS 8-11 -74 0 -- L (172 . -93)

following:

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

forming:

[DO WE HAVE LESS

Noticing word event:

194 EIGHT 8-10 -78 0 -- L (172 . -93)

following:

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

forming:

[DO WE HAVE EIGHT

Sorting EVENTQUEUE

Found better MAXSEG -- recomputing shortfalls

MAXSEGS:

1 20

2 43

3 75

4 21

5 43

6 73

7 34

8 82

9 41

10 23 9

Another MAXSEG adjustment.

11 73

12 30

13 36

14 30

15 36

16 51

17 63

RMS333 Y 17-Oct-76 14:21:14 Used 638 in 1152 seconds. BBN-TENEXC
System, Group load avgs = 2.47 1.66 Control-fcrk page faults = 7104

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLY	SCORE	REGION
1	285	54	175	3.12	163*	1	1-10> [DO WE HAVE A !
2	62	21	170	7.39	36*	.898	<2-6 WILL WE
3	70	17	112	6.59	60*	.471	1-5> [WHO WILL
4	20	3	0	0	0	0	0-1 -PAUSE- !
5	57	11	61	5.55	0*	-.0909	1-3 HE
6	5	21	227	9.87	36*	-.342	<2-6 CAN WE
7	206	54	254	4.54	163*	-.406	1-10> [DO WE HAVE MORE
8	80	14	174	6.69	-67*	-.646	3-6> [WE ARE

9	-39	28	372	10.9	103*	-.687	<0-6	WHAT DO WE
10	37	25	238	8.81	36*	-.9	<1-6	HAVE WE
11	100	21	132	5.74	36	-1.13	<2-6	ARE WE
12	142	54	318	5.68	163*	-1.55	1-10>	[DO WE HAVE EIGHT
13	-27	21	380	10.9	36*	-1.96	<2-6	WERE WE
14	145	66	387	5.69	163*	-2.31	1-11>	[DO WE HAVE LESS
15	-2	7	77	11	0*	-2.43	2-3	A !
16	172	77	426	5.39	163*	-2.5	1-13>	[DO WE HAVE OVER
17	130	37	164	4.43	-67*	-2.65	3-9>	[WE WENT
18	145	77	453	5.73	163*	-2.84	1-13>	[DO WE HAVE UNDER
19	44	37	218	5.89	0*	-4.27	5-10	ELEVEN
20	7	31	223	7.19	0*	-5.26	5-9	IRAQ
21	-1	11	119	10.8	0*	-5.36	1-3	HER
22	6	31	247	7.97	-67*	-5.68	3-8>	[WE WERE
23	1	13	138	10.6	0*	-6	1-4	WHOLE
24	17	17	165	9.71	0*	-6.18	1-5	DO !
25	-10	14	147	10.5	0*	-6.21	3-6	WERE
26	-24	30	313	10.4	60*	-6.43	1-7>	[WHO IS
27	-14	19	185	9.74	0*	-6.58	3-7	WORK
28	-10	13	149	11.5	0*	-6.85	1-4	HIS
29	3	17	179	10.5	0*	-7	1-5	HOW
30	17	25	238	9.52	0*	-7.12	1-6	HAWAII

plus 41 additional sorted and 0 unsorted events.

Verscore = (837 59 63 75) / (59 71)

for 189 A 8-10 96 -31 (REDUCED) L (172 . -93)
in [DO WE HAVE A !

A gets a good verification score of 59.

New EVT.SCORE = .9861111

About to do event 1: [DO WE HAVE A !

CREATING THEORY#8 FROM THEORY#7

SCORES = (285 5.089286 175 3.125 (0 222 0.0))

154 DO 1-4 63 -31 (REDUCED) R (152 . -77)

164 DO 1-5 17 0 -- R (151 . -79)

171 DO 1-5 -82 0 -- R (153 . -148)

** et alia **

90 WE 3-6 128 0 -- L

152 WE 3-6 128 0 -- R

151 WE 2-6 130 0 -- R

** et alia **

172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)

189 A 8-10 96 -31 (REDUCED) L (172 . -93)

DOING PROPOSALS:

On the right:

Words: DEFICIT HUNDRED SURPLUS THOUSAND

Branching factor is 4

Noticing word event:
 22 SURPLUS 10-17 64 0 -- L (189 . 0)
 following:
 189 A 8-10 96 -31 (REDUCED) L (172 . -93)
 forming:
 [DO WE HAVE A SURPLUS !

SURPLUS is matched with a fair score of 64, but that's appreciably below the MAXSEGS in the region of 10-17, so it gets a lot of shortfall. Consequently, this event falls down to #11 on the event queue, as shown below.

Noticing word event:
 199 THOUSAND 10-15 -195 -10 -- L (189 . 0)
 following:
 189 A 8-10 96 -31 (REDUCED) L (172 . -93)
 forming:
 [DO WE HAVE A THOUSAND

Sorting EVENTQUEUE

 RMS333 Y 17-Oct-76 14:22:04 Used 657 in 1202 seconds. BBN-TENEXC
 System, Group load avgs = 1.95 1.54 Control-fcrk page faults = 7587

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	62	21	170	7.39	36*	.898	<2-6	WILL WE
2	70	17	112	6.59	60*	.471	1-5>	[WHO WILL
3	20	3	0	0	0	0	0-1	-PAUSE- !
4	57	11	61	5.55	0*	-.0909	1-3	HE
5	5	21	227	9.87	36*	-.342	<2-6	CAN WE
6	206	54	254	4.54	163*	-.406	1-10>	[DO WE HAVE MORE
7	80	14	174	6.69	-67*	-.646	3-6>	[WE ARE
8	-39	28	372	10.9	103*	-.687	<0-6	WHAT DC WE
9	37	25	238	8.81	36*	-.9	<1-6	HAVE WE
10	100	21	132	5.74	36	-1.13	<2-6	ARE WE
11	345	124	430	3.41	222*	-1.14	1-17>	[DO WE HAVE A SURPLUS !
12	142	54	318	5.68	163*	-1.55	1-10>	[DO WE HAVE EIGHT
13	-27	21	380	10.9	36*	-1.96	<2-6	WERE WE
14	145	66	387	5.69	163*	-2.31	1-11>	[DO WE HAVE LESS
15	-2	7	77	11	0*	-2.43	2-3	A !
16	172	77	426	5.39	163*	-2.5	1-13>	[DO WE HAVE OVER
17	130	37	164	4.43	-67*	-2.65	3-9>	[WE WENT
18	145	77	453	5.73	163*	-2.84	1-13>	[DO WE HAVE UNDER
19	78	88	585	6.5	222*	-3.3	1-15>	[DO WE HAVE A THOUSAND
20	44	37	218	5.89	0*	-4.27	5-10	ELEVEN
21	7	31	223	7.19	0*	-5.26	5-9	IRAQ
22	-1	11	119	10.8	0*	-5.36	1-3	HER
23	6	31	247	7.97	-67*	-5.68	3-8>	[WE WERE
24	1	13	138	10.6	0*	-6	1-4	WHOLE
25	17	17	165	9.71	0*	-6.18	1-5	DO !
26	-10	14	147	10.5	0*	-6.21	3-6	WERE
27	-24	30	313	10.4	60*	-6.43	1-7>	[WHO IS
28	-14	19	185	9.74	0*	-6.58	3-7	WORK
29	-10	13	149	11.5	0*	-6.85	1-4	HIS
30	3	17	179	10.5	0*	-7	1-5	HOW

plus 42 additional sorted and 0 unsorted events.

Verscore = N.A. / (21 30)
for 156 WILL 2-4 -39 -26 (REDUCED) R (152 . -77)
in WILL WE
New EVT.SCORE = -1.959627

Two more events are given to verification. -PAUSE- can't be verified, and WILL gets a bad score of -73.

Verscore = (240 -73 26 47) / (21 47)
for 140 WILL 2-7 -51 0 -- L (62 . -79)
in WHO WILL
New EVT.SCORE = -7.352941

About to do event 1: -PAUSE- !

Looking for
51 -PAUSE- 0-1 20 0 -- L
from right to left

Doing syntax evaluation of theory.

CREATING THEORY#9

SCORES = (20 6.666667 0 0.0 (0 0 0))

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

DOING PROPOSALS:

On the left:

Words: -PAUSE- AGAIN ALSO AMOUNT ANSWER ANYBCDY ANYONE ANYWHERE
ARE ARRANGE-ED AVERAGE AWAY BE BEGIN BUCKS BUDGET BUDGETTED
CANCEL-ED CENT CENT-S COAST CONFERENCE COST COST-PAST COST-S
COSTS DEFICIT DOLLAR DOLLAR-S END EVERYBODY EVERYONE EXPENSE
EXPENSE-S EXPENSIVE FARE FARE-S FEE FIGURE FIGURE-S FLEW
FLY FLY-ING GO GO-ING GROUP HAVE HUNDRED INEXPENSIVE IS
IT ITEM ITEM-S K LAST LEFT LONG MAKE ME MEETING NORMAL NOW
NUMBER OH OVER-BUDGET-ED OVERHEAD PEOPLE PERDIEM PERSON
PLAN PLAN-ED PRESENTLY PREVIOUSLY PRICE PROJECT-N RECENT
RECENTLY REGISTRATION REGULARLY REMAIN-ING SCHEDULE SCHEDULE-ED
SEVENTY SHORT SIXTY SOMEBODY SOMEONE SOMETIME SOON SPEND
SPENT START SURPLUS SYMPOSIUM TAKE TAKEN THAT THEN THEN
THERE THOUSAND TICKET TICKET-S TODAY TOMORROW TOO TOOK TRAVEL
TRAVEL-ED TRAVEL-ING TRIP TRIP-S TYPICAL UNANTICIPATED
UNBUDGETTED UNDER-BUDGET-ED UNEXPECTED WAS WE WENT WERE WORKSHOP
YEAR YESTERDAY

Classes: CITY CONVEYANCE COUNTRY DIGITS DURATION FIRSTNAME
LASTNAME MONTH NPR ORD PRO-OBJ SEASON SPONSOR STATE TEENS
TENS WEEKDAY

Branching factor is 850

Many, many words are proposed on both sides of the -PAUSE-
Hybrid strategy doesn't look for them on the right yet, and of course none are found left of boundary 0.

DOING PROPOSALS:

On the right:

Words: -PAUSE- A ABORT ACTUAL ADD ADD-ED AFTER AIR ALL AN AN AND ANOTHER ANY ANYBODY ANYONE ARE ARRANGE ASSUME AUTO AVERAGE BE BEFORE BETWEEN BUDGET BUDGETTED BUS BY CALCULATE CAN CANCEL CAR CHARGE COMPUTE CONTINUE COST COST-PAST COST-S COSTS CREATE DELETE DID DISPLAY DISREGARD IVIDE-ED DO DOES DURING EACH ENTER ENTIRE ESTIMATE-V ESTIMATE-V-ED EVERYBODY EVERYONE EXACT EXPENSE EXPENSE-S EXPENSIVE FARE FARE-S FIGURE FINAL FIND FOR FORGET FROM GET GIVE GO HAS HAVE HE HER HIS HOW I IN INEXPENSIVE INITIAL IS IT ITEM ITEM-S ITEMIZE LAST LESS LIST LONG MAKE MINUS MORE MULTIPLY-ED MY NEXT NO NONE NORMAL OK ON ONE ONE-WAY OUR OVERHEAD PEOPLE PERDIEM PLAN PLANE PLEASE PLUS PRECISE PREVIOUS PRICE PRINT QUIT RECENT REGISTRATION REMAIN ROUND-TRIP SCHEDULE SHE SHORT SHOW SINCE SOME SOMEBODY SOMEONE STOP SUBTRACT-ED SUPPOSE TELL THANKYOU THAT THE THESE THEY THIS THOSE TICKET TICKET-S TILL TIMES TO TODAY TOMORROW TOTAL TRAIN TRIP TRIP-S TYPICAL UNANTICIPATED UNBUDGETTED UNEXPECTED UNTIL WAS WE WERE WHAT WHEN WHERE WHICH WHO WHOLE WHY WILL YES YESTERDAY ZERO

Classes: CITY COUNTRY DIGITS FIRSTNAME INTEGER LASTNAME NPR NUM-ADJ POSS PROJECT SPONSOR STATE TEENS TENS TRIP-ADJ

Branching factor is 972

Because -PAUSE- is at a non-penalty left end boundary, it is given to Syntax for consideration as a left end.

Trying left end event.

Doing left end event for THEORY#9 creating THEORY#10

[-PAUSE- !

SCORES = (20 6.666667 0 0.0 (0 0 0.0))

DOING PROPOSALS:

On the right:

Words: A ABORT ACTUAL ADD AFTER AIR ALL AN AN AND ANOTHER ANY ANYBODY ANYONE ARE ARRANGE ASSUME AUTO BEPCRE BETWEEN BUDGET BUS BY CALCULATE CAN CANCEL CAR CHARGE COMEUTE CONTINUE COST COST-S CREATE DELETE DID DISPLAY DISREGARD DO DOES DURING EACH ENTER ENTIRE ESTIMATE-V ESTIMATE-V-ED EVERYBODY EVERYONE EXACT EXPENSE EXPENSE-S FARE FARE-S FIGURE FINAL FIND FOR FORGET FROM GET GIVE GO HE HER HIS HOW I IN INITIAL IS IT ITEM ITEM-S ITEMIZE LAST LESS LIST MAKE MORE MY NEXT NO NONE NORMAL OK ON ONE ONE-WAY OUR OVERHEAD PEOPLE PERDIEM PLAN PLANE PLEASE PRECISE PREVIOUS PRICE PRINT QUIT REGISTRATION ROUND-TRIP SCHEDULE SHE SHOW SINCE SOME SOMEBODY SOMEONE STOP SUPPOSE TELL THANKYOU THAT THE THESE THEY THIS THOSE TICKET TICKET-S TILL TODAY TOMORROW TOTAL TRAIN TRIP TRIP-S TYPICAL UNTIL WAS WE WERE WHAT WHEN WHERE WHICH WHO WHOLE WHY WILL YES YESTERDAY ZERO

Classes: CITY COUNTRY DIGITS FIRSTNAME INTEGER LASTNAME NPR NUM-ADJ POSS PROJECT SPONSOR STATE TEENS TENS TRIP-ADJ

Branching factor is 946 ← Note that the previous right proposals permitted 972 words. The left ending constrains them slightly.

Noticing word event:

62 WHO 1-5 182 0 -- L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- WHO

Already noticed: [WHO

Noticing word event:

63 HE 1-3 57 0 -- L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- HE

Noticing word event:

202 OUR 1-5 58 -23 -- L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- OUR

Noticing word event:

64 HAWAII 1-6 17 0 -- L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- HAWAII

Noticing word event:

66 DO 1-5 17 0 -- L (51 . 0)

65 DO 1-5 17 -31 (REDUCED) L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- DO !

; <RESULTS>SH7-12OCT/BIGDICT-115/LHSDVP.RMS333;3 Fri 3-Dec-76 2:14PM

Noticing word event:

204 A 1-5 14 0 -- L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- A

Noticing word event:

69 WHOSE 1-6 5 0 -- L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- WHOSE

Noticing word event:

67 HER 1-3 15 -16 (REDUCED) L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- HER

Noticing word event:

68 HIS 1-4 6 -16 (REDUCED) L (51 . 0)

following:

51 -PAUSE- 0-1 20 0 -- L

136 -PAUSE- 0-1 20 0 -- R (158 . 0) (165 . 0) (171 . 0) (164 . 0) (135
** . 0) (134 . 0) (133 . 0)

201 -PAUSE- 0-2 -193 0 -- R

forming:

[-PAUSE- HIS

Sorting EVENTQUEUE

RMS333 Y 17-Oct-76 14:29:03 Used 907 in 1621 seconds. BBN-TENEXC
System, Group load avgs = 1.44 1.09 Control-fcrk page faults = 9061

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	77	14	61	4.36	0*	-.0714	0-3>	[-PAUSE- HE
2	57	11	61	5.55	0*	-.0909	1-3	HE
3	5	21	227	9.87	36*	-.342	<2-6	CAN WE
4	206	54	254	4.54	163*	-.406	1-10>	[DO WE HAVE MORE
5	80	14	174	6.69	-67*	-.646	3-6>	[WE ARE
6	-39	28	372	10.9	103*	-.687	<0-6	WHAT DC WE
7	37	25	238	8.81	36*	-.9	<1-6	HAVE WE
8	100	21	132	5.74	36	-1.13	<2-6	ARE WE
9	345	124	430	3.41	222*	-1.14	1-17>	[DO WE HAVE A SURPLUS !
10	142	54	318	5.68	163*	-1.55	1-10>	[DO WE HAVE EIGHT
11	62	21	170	7.39	36	-1.96	<2-6	WILL WE
12	-27	21	380	10.9	36*	-1.96	<2-6	WERE WE
13	145	66	387	5.69	163*	-2.31	1-11>	[DO WE HAVE LESS
14	-2	7	77	11	0*	-2.43	2-3	A !
15	172	77	426	5.39	163*	-2.5	1-13>	[DO WE HAVE OVER
16	130	37	164	4.43	-67*	-2.65	3-9>	[WE WENT
17	145	77	453	5.73	163*	-2.84	1-13>	[DO WE HAVE UNDER
18	78	88	585	6.5	222*	-3.3	1-15>	[DO WE HAVE A THOUSAND
19	19	14	119	8.5	0*	-4.21	0-3>	[-PAUSE- HER
20	44	37	218	5.89	0*	-4.27	5-10	ELEVEN
21	55	20	147	7.35	0*	-4.35	0-5>	[-PAUSE- OUR
22	37	20	165	8.25	0*	-5.25	0-5>	[-PAUSE- DO !
23	7	31	223	7.19	0*	-5.26	5-9	IRAQ
24	-1	11	119	10.8	0*	-5.36	1-3	HER
25	34	20	168	8.4	0*	-5.4	0-5>	[-PAUSE- A
26	10	16	149	9.31	0*	-5.56	0-4>	[-PAUSE- HIS
27	6	31	247	7.97	-67*	-5.68	3-8>	[WE WERE
28	1	13	138	10.6	0*	-6	1-4	WHOLE
29	17	17	165	9.71	0*	-6.18	1-5	DO !
30	-10	14	147	10.5	0*	-6.21	3-6	WERE

plus 49 additional sorted and 0 unsorted events.

Verscore = (323 -66 12 24) / (17 28)
 for 63 HE 1-3 57 0 -- L (51 . 0)
 in [-PAUSE- HE
 New EVT.SCORE = -9.071429

The top 7 events all get poor scores from Verification, which knocks them down the event queue.

Verscore = (323 -56 12 24) / (17 28)
 for 63 HE 1-3 57 0 -- L (51 . 0)
 in HE
 New EVT.SCORE = -11.54545

Verscore = N.A. / (21 30)
 for 160 CAN 2-4 -96 -26 (REDUCED) R (152 . -77)
 in CAN WE
 New EVT.SCORE = -3.198758

Verscore = (-4 -116 58 76) / (59 71)
 for 190 MORE 8-10 -14 0 -- L (172 . -93)
 in [DO WE HAVE MORE
 New EVT.SCORE = -3.665344

: <RESULTS>SH7-12OCT/BIGDICT-115/LHSDVP.RHS333;3 Fri 3-Dec-76 2:14PM

Verscore = (-149 -116 25 37) / (30 42)
for 175 ARE 4-6 -33 -16 (REDUCED) L (78 . 0) (90 . 0)
in [WE ARE
New EVT.SCORE = -13.21762

Verscore = N.A. / (14 17)
for 183 WHAT 0-1 -147 -31 (REDUCED) R (158 . 0) (165 . 0) (171 . 0) (1
**64 . 0)
in WHAT DO WE
New EVT.SCORE = -2.8302

Verscore = (163 -104 13 30) / (17 30)
for 159 HAVE 1-4 -74 -16 (REDUCED) R (152 . -77)
in HAVE WE
New EVT.SCORE = -7.46

About to do event 1: ARE WE

CREATING THEORY#11 FROM THEORY#3
SCORES = (100 4.347826 132 5.73913 (0 36 0))
155 ARE 2-4 -11 -16 (REDUCED) R (152 . -77)
162 ARE 3-4 -102 -16 (REDUCED) R (152 . -77)
90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R
** et alia **

This event had been verified earlier. (note absence of * in the above queue printout;

DOING PROPOSALS:
On the left:
Words: -PAUSE- WHEN WHERE

Again, proposals are made on both sides, but only left-going ones are considered now.

Branching factor is 3

DOING PROPOSALS:
On the right:
Words: ATTEND-ING FLY-ING GO-ING OVER OVER-BUDGET-ED SCHEDULE-ED
TRAVEL-ING UNDER UNDER-BUDGET-ED VISIT-ING

Branching factor is 10

Noticing word event:
205 -PAUSE- 0-2 -124 0 -- R (155 . 0)
206 -PAUSE- 2-3 -170 0 -- R (162 . 0)
200 -PAUSE- 1-2 -117 0 -- R (155 . 0)

preceding:
155 ARE 2-4 -11 -16 (REDUCED) R (152 . -77)
162 ARE 3-4 -102 -16 (REDUCED) R (152 . -77)

forming:
-PAUSE- ARE WE

The two events -PAUSE- ARE WE and ARE WE are considered identical if they're at the left end.

Already noticed: ARE WE

Noticing left end event
for: [ARE WE
Endscore = -.2362205

Sorting EVENTQUEUE

RMS333 Y 17-Oct-76 14:30:21 Used 947 in 1699 seconds. BBN-TENEXC
System, Group load avgs = 1.39 1.11 Control-fcrk page faults = 9700

Event queue now is:

#	QUAL	DUR	SHORT	SD	NLX	SCORE	REGION	
1	345	124	430	3.41	222*	-1.14	1-17>	[DO WE HAVE A SURPLUS !
2	100	21	132	5.74	6	-1.37	2-6	[ARE WE
3	142	54	318	5.68	163*	-1.55	1-10>	[DO WE HAVE EIGHT
4	62	21	170	7.39	36	-1.96	<2-6	WILL WE
5	-27	21	380	10.9	36*	-1.96	<2-6	WERE WE
6	145	66	387	5.69	163*	-2.31	1-11>	[DO WE HAVE LESS
7	-2	7	77	11	0*	-2.43	2-3	A !
8	172	77	426	5.39	163*	-2.5	1-13>	[DO WE HAVE OVER
9	130	37	164	4.43	-67*	-2.65	3-9>	[WE WENT
10	-39	28	372	10.9	103	-2.83	<0-6	WHAT DC WE
11	145	77	453	5.73	163*	-2.84	1-13>	[DO WE HAVE UNDER
12	5	21	227	9.87	36	-3.2	<2-6	CAN WE
13	78	88	585	6.5	222*	-3.3	1-15>	[DO WE HAVE A THOUSAND
14	206	54	254	4.54	47	-3.67	1-10>	[DO WE HAVE MORE
15	19	14	119	8.5	0*	-4.21	0-3>	[-PAUSE- HER
16	44	37	218	5.89	0*	-4.27	5-10	ELEVEN
17	55	20	147	7.35	0*	-4.35	0-5>	[-PAUSE- OUR
18	37	20	165	8.25	0*	-5.25	0-5>	[-PAUSE- DO !
19	7	31	223	7.19	0*	-5.26	5-9	IRAQ
20	-1	11	119	10.8	0*	-5.36	1-3	HER
21	34	20	168	8.4	0*	-5.4	0-5>	[-PAUSE- A
22	10	16	149	9.31	0*	-5.56	0-4>	[-PAUSE- HIS
23	6	31	247	7.97	-67*	-5.68	3-8>	[WE WERE
24	1	13	138	10.6	0*	-6	1-4	WHOLE
25	17	17	165	9.71	0*	-6.18	1-5	DO !
26	-10	14	147	10.5	0*	-6.21	3-6	WERE
27	37	28	238	8.5	0*	-6.36	0-6>	[-PAUSE- HAWAII
28	-24	30	313	10.4	60*	-6.43	1-7>	[WHO IS
29	-14	19	185	9.74	0*	-6.58	3-7	WORK
30	25	28	250	8.93	0*	-6.79	0-6>	[-PAUSE- WHOSE

plus 49 additional sorted and 0 unsorted events.

Verscore = (750 43 66 136) / (71 141)
for 22 SURPLUS 10-17 64 0 -- L (189 . 0)
in [DO WE HAVE A SURPLUS !
New EVT.SCORE = -1.275602

The correct event has worked
its way to the top of the
event queue again. It gets
a good Verification score
of 43.

About to do event 1: [DO WE HAVE A SURPLUS !

```

*****
CREATING THEORY#12 FROM THEORY#8
SCORES = (345 2.738095 430 3.412698 (0 265 0.0))
154 DO 1-4 63 -31 (REDUCED) R (152 . -77)
164 DO 1-5 17 0 -- R (151 . -79)
171 DO 1-5 -82 0 -- R (153 . -148)
  ** et alia **
90 WE 3-6 128 0 -- L
152 WE 3-6 128 0 -- R
151 WE 2-6 130 0 -- R
  ** et alia **
172 HAVE 6-9 189 -31 (REDUCED) L (78 . 0) (90 . 0)
189 A 8-10 96 -31 (REDUCED) L (172 . -93)
22 SURPLUS 10-17 64 0 -- L (189 . 0)

```

DOING PROPOSALS: ← Syntax proposes only -PAUSE- on the right
 On the right:
 Words: -PAUSE-

Branching factor is 1
 Trying right end event. ← Because the right boundary of the event is a definite right end, we ask Syntax to consider that. It does, sees a spanning event, and returns the complete parse.

```

*****
Doing right end event for THEORY#12 creating THEORY#13
[DO WE HAVE A SURPLUS] !

```

Found complete parse:

```

(POR: SOME A0010 / (FINDQ: DB/GROUP (MEMBERS SPEAKER))
  : T ; (POR: SOME A0009 / (FINDQ: DB/CONTRACT (GROUP A0010))
        : T ;
        (POR: THE A0011 / (FINDQ: DB/BUDGET
                          (BUDGET/OF A0009))
        : T ;
        (TEST: (GREATERP (GET: A0011
                          MONEY/REMAINING)
                  0))))))

```

```

S Q
SUBJ NP PRO YOU
      FEATS NU SG
AUX TNS PRESENT
      VOICE ACTIVE
VP V HAVE
   OBJ NP N SURPLUS
      FEATS NU SG

```

Semantic interpretation, which can be roughly translated as:

"Find a group of which the speaker is a member; find a contract that belongs to that group; find the budget of that contract; test if the money remaining in that budget is greater than zero, and repeat until a non-zero budget is found or there are no more groups and contracts for the speaker.

"Traditional" parse tree of the sentence. This is a vestigial output of the parser. It no longer has a function in the system.

SCORES = (345 2.738095 430 3.412698 (0 265 0.0))

These are the scores of the spanning theory.

Spanning theory found.

 RMS333 Y 17-Oct-76 14:31:17 Used 971 in 1755 seconds. BBN-TENEXC
 System, Group load avgs = 1.38 1.14 Control-fcrk page faults = 10145

FUNCTIONS	TIME	# CALLS	PER CALL	%	
LEPTENDSCAN	63.650	1	63.650	7	} L
INITIALSCAN	36.649	1	36.649	4	
JACKPROPOSE-IN-INITIALSCAN	39.203	1	39.203	4	
JACKPROPOSE-IN-SYNDRIV	0.000				
JACKPROPOSE-IN-DOPROPOSALS	112.280	15	7.485	13	} C
RESCORE-EVTS	4.100	3	1.367	0	
REVERSEDIR?	7.642	3	2.547	1	} L
VERIFY	76.526	16	4.783	9	
DOEVT	0.000	0	0.000	0	} C
SEDEVTHANDLER	184.769	3	61.590	21	
WORDEVTHANDLER	16.966	5	3.393	2	} S
COLLISIONEVTHANDLER	0.000	0	0.000	0	
ENDEVTHANDLER	66.737	5	13.347	8	} C
NOTICELEFT	3.485	4	0.871	0	
NOTICERIGHT	1.930	10	0.193	0	
NOTICE!	22.791	90	0.253	3	
SHORTORDERP	3.228	993	0.003	0	} C
SORTQUEUE	2.415	9	0.268	0	
SYNDRIV	93.542	1	93.542	11	
PRINTEVENTQ	36.121	9	4.013	4	
ENDSCAN	113.665	60	1.894	13	} L
TOTAL	885.699	1229	0.721		

This is a breakdown of the time spent in the principal functions of the control fork, which includes the time spent in subsidiary processes. For this utterance, time was spent as:

Lexical Retrieval	42%
Verification	9%
Syntax	31%
Control	18%

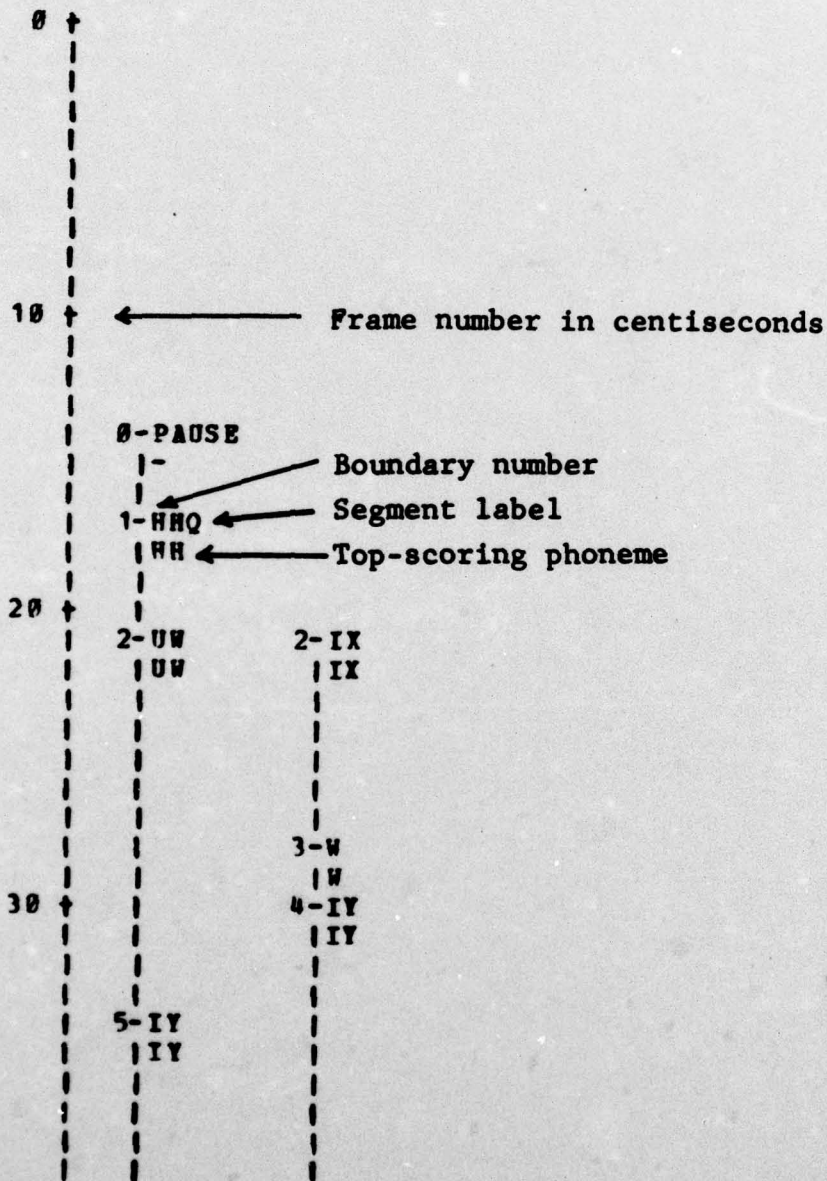
THEORY# 1	WHO	This is a recapitulation of the theories formed during the understanding of this utterance.
THEORY#2(1)	[WHO	
THEORY#3	WE !	
THEORY#4(3)	DO WE ?	
THEORY#5(4)	[DO WE !	
THEORY#6(3)	[WE !	
THEORY#7(5)	[DO WE HAVE !	
THEORY#8(7)	[DO WE HAVE A !	
THEORY#9	-PAUSE- !	
THEORY#10(9)	[-PAUSE- !	
THEORY#11(3)	ARE WE	
THEORY#12(8)	[DO WE HAVE A SURPLUS !	
THEORY#13(12)	[DO WE HAVE A SURPLUS] !	

~L

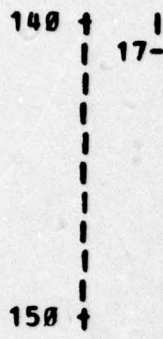
Program to plot .SL files
Extension: SLY Output to TTY? Yes
Look at actual probabilities? No
Also output dictionary mapping? Yes
Also output hand labels? No
Plot only part of sentence? No

Segment lattice for
utterance RMS333. Time
runs down the page.

Utterance: RMS333
; <MAKHOUL>RMS333.SLY



40	↑			
		6-IVSON	6-	
		RH		
		7-PH		
		AE		
50	↑			
		8-NVOBS		
		V		
		9-UH		
		UH		
1				
0	70	↑	10-S	
			Z	
H				
S				
F				
r				
a				
m				
e				
s				
	80	↑	11-ER	11-AX
			ER	AX
			12-R	
			R	



More for this SL file? Yes
 Only change region? No
 Look at actual probabilities? Yes
 Compute Labeling Performance statistics? No
 Use a scoring Offset to simulate matcher score? Yes
 Offset: 70 Only look at the top few? Yes
 How many? 7 Also output dictionary mapping? No
 Also output hand labels? No
 Plot only part of sentence? No

This is a tabulation, for each segment, of the top 7 phonemes and their log-likelihood scores.

segment boundary numbers
 segment boundary times

0-1[PAUSE]14-17 115	Stress: 10						
20 -	-83 HH	-130 G	-147 OY	-147 ENX	-147 URB	-147 EA	
1-2[HHQ]17-21 204	Stress: 36						
103 HH	54 Q	-21 DH	-58 OY	-58 ENX	-58 URB	-58 EA	
2-5[UW]21-34 229	Stress: 42						
79 UW	62 EH	42 EL	30 UR	16 EN	-22 AH	-26 AO	
2-3[IX]21-28 143	Stress: 39						
48 IX	14 AX	-16 INX	-25 EN	-40 IH	-46 IY	-65 LIH	
3-4[W]28-30 173	Stress: 34						
77 W	-3 L	-16 Q	-89 OY	-89 ENX	-89 URB	-89 EA	
4-6[IY]30-42 141	Stress: 45						
51 IY	16 INX	7 UY	-67 LIH	-87 ER	-101 AX	-101 OW	
5-6[IY]34-42 141	Stress: 45						
51 IY	16 INX	7 UY	-67 LIH	-87 ER	-101 AX	-101 OW	
6-7[IVSON]42-47 213	Stress: 35						
27 HH	26 TX	24 Q	4 DH	-4 G	-22 W	-27 YH	
7-8[EH]47-59 141	Stress: 49						
69 AE	42 EH	34 RH	-22 LIH	-55 IX	-56 AY	-65 AX	
8-9[NVOBS]59-65 160	Stress: 41						
93 V	73 DH	26 DX	22 D	10 TS	4 B	0 Q	

; <RESULTS>SH7-120CT/BIGDICT-!15/LHSDVP.RMS333;3 Fri 3-Dec-76 2:14PM

9-10[UH]65-71 174 Stress: 45
72 UH 47 LIH 39 UW 23 EL 3 AH -11 OW -11 AXR

10-11[S]71-83 166 Stress: 38
97 Z 63 TH 47 S -37 F -60 SH -82 TCH -96 OY

11-13[ER]83-94 175 Stress: 45
73 ER 65 AXR -20 EH -24 UH -28 UW -31 RIH -52 AA

11-12[AX]83-88 167 Stress: 45
50 AX 9 IX 8 RIH -18 OW -25 AXR -61 ER -95 OY

12-13[R]88-94 144 Stress: 45
63 R -70 V -81 DX -87 AXR -118 OY -118 ENX -118 URB

13-14[URPLOS]94-99 248 Stress: 37
98 URG 77 URK 64 URT 28 K 3 G -14 OY -14 ENX

13-15[JHT]94-105 223 Stress: 37
70 JH 67 TV 14 T 12 P -39 OY -39 ENX -39 URB

14-15[JHT]99-105 223 Stress: 28
70 JH 64 ST 14 T -33 TV -39 OY -39 ENX -39 URB

15-16[AH]105-121 151 Stress: 45
53 AH 49 EH -16 AE -17 AY -20 ER -24 OW -29 AA

16-17[SZ]121-141 141 Stress: 32
78 S 33 TH 0 Z -32 CH -121 OY -121 ENX -121 URB

More for this SL file? No

Utterance: STOP Halt? Yes

-L

; <RESULTS>SH7-12OCT/BIGDICT-115/LHSDVP.RMS333;3 Fri 3-Dec-76 2: 14PM

HWIM-RUN SUMMARY 17-Oct-76 14:02:02 Utterance: RMS333 Y

Control: (<SPSYS>SH7-12OCT.BIGDICT-115;1 . <LISP>LISP.SAV;83)
Matcher: <KLOVSTAD>CLASSLEX.SAV;13
Dictionary: <KLOVSTAD>DICTSTRUCT.BIGDICT-115;2
Verifier: <COOK>NEWMAPPER.SAV;47
Synfork: <SPSYS>SYNFORK.BIGDICT-115;12

SHORTFALLFLAG = T	DENSITYFLAG = T	PRUNEFLAG = NIL
OTHERDIRFLAG = NIL	UNSHLPFLAG = T	GHOSTFLAG = T
CHOOSDIRFLAG = NIL	BESTSPANFLAG = T	QUICKFLAG = NIL
GHOSTPSLSFLAG = T	ENDCHECKFLAG = T	CREDITFLAG = NIL
LIABILITYFLAG = NIL	PRONLIKEFLAG = T	PAUSECLAMP = 20
COLLISIONFLAG = NIL	CONTEXTFLAG = T	LEPTENDFLAG = NIL
KILLBADEVTSFLAG = T	GOODONLYFLAG = NIL	PROSODYFLAG = NIL
VERIFYFLAG = VERIFY@PICK		
	VERIFYDEFAULT = 60	HYBRIDFLAG = 0
MAXTHRYN = 150	MINSTARTS = 0	

Sentence: Do we have a surplus?

Result: CORRECT spanning theory found. Score = 345
DO WE HAVE A SURPLUS !

Run on BBN-TENEXC	N	Word	Seen?
Used 971 in 1755 seconds.	Seeds: 2	WE	yes
13 theories, 74 sec/theory.	10	DO	no
10145 control page faults.			
Arithmetic mean branching = 241			
Geometric mean branching = 19.4			
Utterance duration = 1.27 sec.			
Speed = 783 times real time.			

This page is a summary of the run, showing the files that were used in the system, the state of the control flags, the sentence and its outcome, and some runtime statistics.

Appendix 3 - Publications

(November 1974 - October 1976)

- [1] Bates, M. (1975)
"Syntactic Analysis in a Speech Understanding System", Report No. 3116, Bolt Beranek and Newman Inc., Cambridge MA.
- [2] Bates, M. (1975)
"Syntactic Processing in the BBN Speech Understanding System", presented at the 13th Annual Meeting of the Association for Computational Linguistics, Boston MA.
- [3] Bates, M. (1975)
"The Use of Syntax in a Speech Understanding System", IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-23, No.1, pp.112-117.
- [4] Bruce, B.C. (1975)
"Belief Systems and Language Understanding", Report No. 2973, Bolt Beranek and Newman Inc., Cambridge MA.
- [5] Bruce, B.C. (1975)
"Case Systems for Natural Language", Artificial Intelligence 6, pp.327-360. Also as Report No. 3010, Bolt Beranek and Newman Inc., Cambridge MA.
- [6] Bruce, B.C. (1975)
"Discourse Models and Language Comprehension", American J. of Computational Linguistics, Microfiche 35, pp.19-35, 1975(4).
- [7] Bruce, B.C. (1976)
"Discourse Influence on Language Generation", presented at the 1976 International Conference on Computational Linguistics (COLING-76), Ottawa Canada.
- [8] Bruce, B.C. (1975)
"Pragmatics in Speech Understanding", Proc. Fourth International Joint Conference on Artificial Intelligence, Tbilisi USSR.
- [9] Bruce, B. (1975)
"Generation as a Social Action", in B. Nash-Webber and R. Schank (eds.), Theoretical Issues in Natural Language Processing, 10-13 June 1975, Cambridge MA, pp.64-67.
- [10] Burton, R. and W.A. Woods (1976)
"A Compiling System for Augmented Transition Networks," Proc. 1976 International Conference on Computational Linguistics (COLING-76), Ottawa Canada.
- [11] Cook, C.C. and D.H. Klatt (1976)
"Word Verification in a Speech Understanding System", J. Acoust. Soc. Am. 60 S11 (Abstract). Presented at the 92nd Meeting of the Acoustical Society of America, 15-19 November 1976, San Diego CA.
- [12] Cook, C.C. and R.M. Schwartz (forthcoming)
"Advanced Acoustic Techniques in Automatic Speech Understanding". (To be presented at the 1977 IEEE International Conf. on Acoustics, Speech, and Signal Processing, 9-11 May 1977, Hartford CT.)
- [13] Cook, C.C. (1976)
"Word Verification in a Speech Understanding System", Proc. IEEE International Conf. on Acoustics, Speech, and Signal Processing, 12-14 April 1976, Philadelphia PA, pp.553-556.
- [14] Klovstad, J.W. (1976)
"Probabilistic Lexical Retrieval Component with Embedded Phonological Word Boundary Rules", Proc. of the International Conference on Acoustics, Speech, and Signal Processing, April 12-14, Philadelphia PA (Abstract).

- [15] Makhoul, J.I. and C.C. Cook (1974)
 "Optimal Number of Poles in a Linear Prediction Model", J. Acoust. Soc. Am. 56, S14 (Abstract). Presented at the 88th Meeting of the Acoustical Society of America, 4-8 November 1974, St. Louis MO.
- [16] Makhoul, J.I. (1975)
 "Linear Prediction: A Tutorial Review", Proc. IEEE, Vol. 63, No. 4, pp. 561-580.
- [17] Makhoul, J.I. (1975)
 "Spectral Linear Prediction: Properties and Applications", IEEE Trans. Acoustics, Speech and Signal Processing, Vol. ASSP-23, No.3, pp.283-296.
- [18] Makhoul, J.I. (1975)
 "Linear Prediction in Automatic Speech Recognition", in Speech Recognition: Invited Papers presented at the IEEE Symposium, D.R. Reddy (ed.), New York: Academic Press, pp.183-220.
- [19] Makhoul, J.I. (1976)
 "New Lattice Methods for Linear Prediction", Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, Philadelphia PA.
- [20] Makhoul, J.I. (1976)
 "Methods for Nonlinear Spectral Distortion of Speech Signals", Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, April 12-14, Philadelphia PA.
- [21] Nash-Webber, B.L. and B.C. Bruce (1976)
 "Evolving Uses of Knowledge in a Speech Understanding System", Proc. 1976 International Conference on Computational Linguistics (COLING-76), Ottawa Canada.
- [22] Nash-Webber, B.L. (1975)
 "The Role of Semantics in Automatic Speech Understanding", in D. Bobrow and A. Collins (eds.), Representation and Understanding: Studies in Cognitive Science, New York: Academic Press.
- [23] Nash-Webber, B.L. (1975)
 "Semantic Support for a Speech Understanding System", IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-23, No.1, pp.124-128.
- [24] Nash-Webber, B.L. (1976)
 "Experiences with Semantic Interpretation", presented at the 1976 National Computer Conference, New York City, 7-10 June 1976.
- [25] Nash-Webber, B.L. (1976)
 "Semantic Interpretation Revisited", Report No. 3335, Bolt Beranek and Newman Inc. (Presented at the 1976 International Conference on Computational Linguistics (COLING-76), Ottawa Canada.)
- [26] Rovner, P.D., B.L. Nash-Webber, and W.A. Woods, (1975)
 "Control Concepts in a Speech Understanding System", IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-23, No.1, pp.136-139.
- [27] Schwartz, R.M. and V. Zue (1976)
 "Acoustic-Phonetic Recognition in BBN SPEECHLIS", Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, April 12-14, Philadelphia, pp. 21-24.
- [28] Schwartz, R.M., J. Klovstad, V. Zue, J. Makhoul, and J. Wolf (1976)
 "Phonetic and Lexical Processing in the HWIM Speech Understanding System", J. Acoust. Soc. Am. 60 S11 (Abstract). Presented at 92nd Meeting of the Acoustical Society of America, 15-19 November 1976, San Diego CA.
- [29] Schwartz, R.M. (1976)
 "Acoustic-Phonetic Experiment Facility for the Study of Continuous Speech", Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, April 12-14, Philadelphia, pp.1-4. (Also J. Acoust. Soc. Am. 58

S105 (Abstract). Presented at the 90th Meeting of the Acoustical Society of America, 3-7 November 1975, San Francisco CA.)

[30] Wolf, J.J. and W.A. Woods (1976)
"The HWIM Speech Understanding System - Overview and Performance", J. Acoust. Soc. Am. 60 S11 (Abstract). Presented at 92nd Meeting of the Acoustical Society of America, 15-19 November 1976, San Diego CA.

[31] Wolf, J.J. (1976)
"Speech Recognition and Understanding", in K.-S. Fu (ed.), Digital Pattern Recognition, New York: Springer-Verlag.

[32] Wolf, J.J. (1976)
"Knowledge, Hypotheses, and Control in the HWIM Speech Understanding System", Conference Record, 1976 Joint Conference on Pattern Recognition and Artificial Intelligence, (IEEE order number 76CH1169-2C), pp.113-125.

[33] Woods, W.A. and V.W. Zue (1976)
"Dictionary Expansion via Phonological Rules for a Speech Understanding System", Proc. of the International Conference on Acoustics, Speech, and Signal Processing, April 12-14, Philadelphia PA, pp.561-564.

[34] Woods, W.A., M. Bates, G. Brown, B.C. Bruce, J.W. Klovstad, B.L. Nash-Webber (1976)
"Uses of Higher Level Knowledge in a Speech Understanding System: A Progress Report", Proc. of the International Conference on Acoustics, Speech, and Signal Processing, April 12-14, Philadelphia PA, pp.438-441.

[35] Woods, W.A., M. Bates, G. Brown, and J.J. Wolf (1976)
"Linguistic Processing and Control Strategy in the HWIM Speech Understanding System", J. Acoust. Soc. Am. 60 S11 (Abstract). Presented at the 92nd Meeting of the Acoustical Society of America, 15-19 November 1976, San Diego CA.

[36] Woods, W.A. (1975)
"Some Methodological Issues in Natural Language Understanding", in B. Nash-Webber and R. Schank (eds.), Theoretical Issues in Natural Language Processing, 10-13 June 1975, Cambridge MA, pp.134-139.

[37] Woods, W.A. (1975)
"Syntax, Semantics and Speech", in D.R. Reddy (ed.), Speech Recognition: Invited Papers presented at the IEEE Symposium, New York: Academic Press.

[38] Woods, W.A. (1975)
"What's in a Link: Foundations for Semantic Networks", in D. Bobrow and A. Collins (eds.), Representation and Understanding: Studies in Cognitive Science, New York: Academic Press.

[39] Woods, W.A. (1975)
"Motivation and Overview of SPEECHLIS: An Experimental Prototype for Speech Understanding Research", IEEE Trans. Acoustics, Speech, and Signal Processing, Vol. ASSP-23, No.1, pp.2-10.

Appendix 4 - Comprehensive Index to Technical NotesQTPR 1. 1 November 1974 to 1 February 1975 (BBN Report No. 3018)

1. Acoustic Segmentation and Labeling (R.M. Schwartz)
2. Lexical Retrieval (J.W. Klovstad)
3. The Syntactic Component (M. Bates)
4. SEMNET - The Network Utility Package (B.L. Nash-Webber)

QTPR 2. 1 February 1975 to 1 May 1975 (BBN Report No. 3080)

1. Acoustic-Phonetic Research (R.M. Schwartz)
2. Speaker Normalization (J.I. Makhoul)
3. PCOMPILER - A Language for Stating Phonological and Phonetic Rules (D.H. Klatt, C.C. Cook, and W.A. Woods)
4. Prosodics (M. Bates and J.J. Wolf)
5. Retrieving Information from the Net (B.C. Bruce)

QTPR 3. 1 May 1975 to 1 August 1975 (BBN Report No. 3115)

1. Time Expressions (M. Bates and B.C. Bruce)
2. Procedural Semantics in the Travel System (B.C. Bruce and G. Harris)
3. Control Primitives (B.L. Nash-Webber)

QTPR 4. 1 August 1975 to 29 October 1975 (BBN Report No. 3188)

1. Parametric Modeling Of Probability Distributions (J.I. Makhoul and R.M. Schwartz)
2. Digital Spectrograms (C.C. Cook)
3. A Phonological Rule System for Dictionary Expansion (W.A. Woods)
4. Audio Response Generation (C.C. Cook, B.C. Bruce and L. Gould)

QTPR 5. 30 October 1975 to 31 January 1976 (BBN Report No. 3240)

1. New Lattice Methods for Linear Prediction (J.I. Makhoul)
2. Methods for Nonlinear Spectral Distortion of Speech Signals (J.I. Makhoul)
3. Acoustic-Phonetic Recognition in BBN SPEECHLIS (R.M. Schwartz and V.W. Zue)
4. Acoustic-Phonetic Experiment Facility for the Study of Continuous Speech (R.M. Schwartz)
5. Word Verification in a Speech Understanding System (C.C. Cook)
6. Uses of Higher Level Knowledge in a Speech Understanding System: A Progress Report (W. Woods, M. Bates, G. Brown, B. Bruce, J. Klovstad, B. Nash-Webber)

QTPR 6. 1 February 1976 to 30 April 1976 (BBN Report No. 3303)

1. Evolving Uses of Knowledge in a Speech Understanding System (B.L. Nash-Webber and B.C. Bruce)
2. Probabilistic Lexical Retrieval with Embedded Phonological Word Boundary Rules (J.W. Klovstad)
3. Shortfall Scoring Strategies for Speech Understanding Control (W.A. Woods)

Official Distribution List

Contract N00014-75-C-0533

Defense Documentation Center
Cameron Station
Alexandria, Virginia 22314

Office of Naval Research
Information Systems Program
Code 437
Arlington, Virginia 22217

Office of Naval Research
Code 1021P
Arlington, Virginia 22217

Office of Naval Research
Branch Office, Boston
495 Summer Street
Boston, Massachusetts 02210

Office of Naval Research
Branch Office, Chicago
536 South Clark Street
Chicago, Illinois 60605

Office of Naval Research
Branch Office, Pasadena
1030 East Green Street
Pasadena, California 91106

New York Area Office
715 Broadway - 5th Floor
New York, New York 10003

Naval Research Laboratory
Technical Information Division
Code 2627
Washington, D.C. 20375

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D.C. 20380

Office of Naval Research
Code 455
Arlington, Virginia 22217

Office of Naval Research
Code 458
Arlington, Virginia 22217

Naval Electronics Lab. Center
Advanced Software Technology Division
Code 5200
San Diego, California 92152

Mr. E. H. Gleissner
Naval Ship Research and
Development Center
Computation and Mathematics Dept.
Bethesda, Maryland 20084

Captain Grace M. Hopper
NAICOM/MIS Planning Branch (OP-916D)
Office of Chief of Naval Operations
Washington, D.C. 20350

Mr. Kin B. Thompson
Technical Director
Information Systems Division (OP-91T)
Office of Chief of Naval Operations
Washington, D.C. 20350

Advanced Research Projects Agency
Information Processing Techniques
1400 Wilson Boulevard
Arlington, Virginia 22209

Commanding Officer
Naval Air Development Center
Warminster, Pennsylvania 18974

Professor Omar Wing
Dept. of Electrical Engineering
Columbia University
New York, New York 10027

Assistant Chief for Technology
Office of Naval Research
Code 200
Arlington, Virginia 22217