DOCUMENT RESUME

ED 149 727 IR 005 504

TITLE Abstracts of Research. July 1974-June 1975.
INSTITUTION Ohio State Univ., Columbus. Computer and Information Science Research Center.
PUB DATE 75
NOTE 92p.; For related documents, see ED 092 117, ED 097 025, ED 139 394

EDRS PRICE MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS *Abstracts; *Annual Reports; Artificial Intelligence; Cognitive Processes; Computational Linguistics; *Computer Programs; *Computer Science; Information Processing; Information Retrieval; *Information Science; Information Storage

ABSTRACT

Abstracts of research papers in computer and information science are given for 68 papers in the areas of information storage and retrieval; human information processing; information analysis; linguistic analysis; artificial intelligence; information processes in physical, biological, and social systems; mathematical techniques; systems programming; computer architecture and networks; joint programs; and computation theory. Abstracts are indexed by investigator and subject. Introductory material includes a description of the Ohio State University Computer and Information Science Research Center. (VT)

ABSTRACTS OF RESEARCH

JULY 1974 - JUNE 1975

The Computer and Information Science Research Center

The Ohio State University

Columbus, Ohio 43210

FOREWORD

The Computer and Information Science Research Center at The Ohio State University is a research organization which consists primarily of faculty, staff and graduate students of the Department of Computer and Information Science. Some of the research activities are performed in conjunction with other University departments as well as off-campus organizations.

This publication contains the abstracts of research which has been carried on during the 1974-75 academic year. This research has been supported in part by grants from governmental agencies as well as by The Ohio State University. Sponsorship with government agencies and with other units on the campus is identified at the end of an abstract.

A bibliography of the research reports published by the Center is included in this publication as Appendix F. Copies of some of these reports are still available on a complimentary basis from the Computer and Information Science Research Center, The Ohio State University, 2036 Neil Avenue Mall, Columbus, Ohio, 43210. Titles with PB or AD numbers may be obtained from The National Technical Information Center, The U.S. Department of Commerce, 5285 Port Royal Road, Springfield, Virginia, 22151, in paper copy, magnetic tape, or microfiche. There is a nominal charge for their service.

Marshall C. Yovits
Director, Computer and
Information Science Research Center

TABLE OF CONTENTS

I. THE OHIO STATE UNIVERSITY COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER

ORGANIZATIONAL STRUCTURE

The Computer and Information Science Research Center at The Ohio State University is an interdisciplinary activity involving the faculty, staff and graduate students from the Department of Computer and Information Science. Some of the research activities are performed in conjunction with other University departments as well as off-campus organizations. The Center also interacts closely with Battelle Memorial Institute, Chemical Abstracts Service, and the Ohio College Library Center, which are adjacent to the Ohio State Campus, as well as with a number of other organizations located in Columbus, Ohio which are engaged in computer and information science research activities, such as Bell Laboratories, Western Electric Corporation, and Rockwell International Corporation. Although the Research Center and the Department utilize many of the same personnel and have the same Director, they are two separate and distinct entities.

OBJECTIVES OF THE CENTER

The Computer and Information Science Research Center has the following objectives: (1) to develop a broad research program in computer and information science; (2) to develop, test, and evaluate practical applications of research in computer and information science; (3) to coordinate and integrate these functions with an academic program in computer and information science at The Ohio State University, as well as with other disciplines at the University. The Center is a focal point for a number of applied information processing activities on the campus.

SCOPE OF THE PROGRAM

The program in computer and information science at The Ohio State University has been defined broadly to encompass most of the analytical activities frequently considered to be part of this discipline. This approach has been chosen because it is felt that in order to generate the needed concepts, foundations, and generalized techniques, it is necessary to examine analytically a number of different areas of computer and information science. In this way a firm empirical and theoretical foundation may be established for generalized computer and information systems. Such a view commits the program not only to the theoretical study of computer and information systems, but also to the study of their realization and their impact on the user.

Those areas of study which are emphasized both in the academic program and in the research activities of the Center are as follows:

1. General theory of information
2. Information storage and retrieval
3. Theory of automata and theory of computation
4. Artificial intelligence
5. Pattern recognition
6. Computer programming, including systems programming.
7. Theory and processing of programming languages
8. Digital computer architecture and organization
9. Numerical analysis
10. Man-machine interaction and systems
11. Formal and computational linguistics
12. Management information and systems
13. Biological information processing
14. Social, economic, and psychological aspects of information production, processing, and use.

FACILITIES

The Computer and Information Science Research Center has a Digital Equipment Corporation DEC System-10 computer. It is a moderate size flexible time-sharing computer which is dedicated to research and education in the field of computer and information science. This computer provides valuable hands-on experience for the faculty and students of the Computer and Information Science Research Center and permits research activities involving non-standard and innovative applications of computers of both a hardware and software nature. Some of these research activities which are currently underway are:

1) Experimental and developmental research in time-sharing and multiprogramming systems.
2) Complex systems simulation research using graphical display devices.
3) On-line information retrieval systems studies.
4) Hardware modification and interface studies.
5) Software modification and development (e.g., PL/1 and COBOL compilers).
6) Man-machine interaction and psychophysical experiments.
7) Pattern recognition studies.
8) Computer simulation of language learning.
9) Speech analysis and synthesis.
10) Analysis and synthesis of human locomotion.
11) Computer data base systems.

The Research Center also has access to the University Computer Centers. They are: Instruction and Research Computer Center; Hospital Computer Center; and University Systems Computer Center.

Included in these centers are an IBM 370/165, two IBM 370/158, an IBM 1620 for plotting, and an IBM 7 for analog to digital conversion, as well as several

IBM 1130 machines, and a number of remote terminals. An IBM 370/168 will be available January 1, 1976.

Many specialized facilities and laboratories of the University are also available to the staff and students of the Research Center. Some of these are the Office of Computer Assisted Instruction, the Institute for Research in Vision, Telecommunications Center, Behavioral Science Laboratory, Listening Center, Communications and Control Systems Laboratory, Mershon Center for Education in National Security, and many others.

The Research Center also interacts with the Ohio College Library Center which is administratively independent of the University. The Center was formed by the Ohio College Association, and operates a common computerized library network connecting the Ohio colleges and universities (both private and state assisted) and many points outside the state of Ohio. These include sixteen regions with approximately 500 college, university, public and special libraries participating in the system. Two well established national information systems have units on the campus of The Ohio State University. The MEDLINE system is an automated on-line service to access medical journals of the previous three year period. The ERIC (Educational Research Information Center) system is an automated batch system to access research reports and journal literature in the field of education. These systems are available to the staff and students. Interaction has also been initiated with The Academy for Contemporary Problems.

The University has established a university-centered-information system. The information system, called the Mechanized Information Center (MIC) operates as a department of the University Libraries. MIC has developed a multi-disciplinary batch-mode information system from machine-readable data bases, primarily for the campus scientific community. The interface to the MIC system is decentralized as much as possible through the existing system of twenty-three libraries around the campus which serve sepcialized publics. MIC acquires data bases from commercial sources, as well as from professional societies and governmental agencies. Research activities in MIC are directed toward improving the services of MIC to its users through software refinement and development. This center works closely with the staff and students of the Department and the Center.

MIC provides two basic types of services: (1) computer-based current awareness to help people keep up-to-date with current publications, and (2) computer-based retrospective searches to bring people up-to-date with previously published information. There are three current awareness services: (1) multidisciplinary, mainly in physical and biological sciences and engineering, (2) social sciences, and (3) education. Two retrospective services, multi-disciplinary and education, are also available. For each of the five services, a specific data base is searched to select bibliographic citations that are pertinent to a person's interests. MIC now provides current awareness searches to more than 4,000 people and organizations and has performed more than 25,000 retrospective searches.

ACADEMIC PROGRAMS IN COMPUTER AND INFORMATION SCIENCE

The program at The Ohio State University emphasizes education, research and the professional practice and application of computer and information science. The academic program is offered within the Department of Computer and Information Science and encompasses undergraduate and graduate degrees through the Ph.D. Statistics showing the growth of the department are found in Table 1.

Organization of The Department of Computer and Information Science

The Department of Computer and Information Science is a separate academic unit located administratively in the College of Engineering, operating in part as an interdisciplinary program with the cooperation of many other departments and colleges throughout the University.

Objectives of The Department

The program at The Ohio State University emphasizes education, research and the professional practice and application of computer and information science. The educational program offers undergraduate and graduate degrees through the Ph.D. The research activities which are a central part of the program consist of a broad conceptual base supported by a number of contracts and grants as well as by the university. The broad core research program and these other research tasks interact to form an integrated framework.

Undergraduate Programs

Undergraduate degrees in computer and information science are available to students in the College of Engineering, the College of Mathematics and Physical Sciences of the College of the Arts and Sciences, and the College of Administrative Sciences. The particular program chosen depends upon the student's interests and career objectives.

The undergraduate program in the College of Engineering leads to the degree of Bachelor of Science in Computer and Information Science. This program is designed for the student who wants to specialize in computer and information science from within an engineering environment. Hence, the program provides the student with a core of computer and information science, mathematics, and engineering science Both depth and breadth in computer and information science, are assured by specific required course sequences in several areas of engineering and science yet, sufficient flexibility exists so that a student can elect a portion of his technical course work in order to develop his individual interests.

There are two undergraduate programs in the College of Mathematics and Physical Sciences. These programs lead either to the degree of Bachelor of Science or the degree of Bachelor of Arts with a major in computer and information science. The programs are cast in a liberal arts setting and are similar in content. The Bachelor of Science program provides a somewhat more technical and thorough education in computer and information science and mathematics while the Bachelor of Arts program is somewhat more flexible and provides an opportunity to relate computer and information science to some other discipline.

Table 1: Growth of Department of Computer and Information Science

	SEPT'67	SEPT'68	SEPT'69	SEPT'70	SEPT'71	SEPT'72	SEPT'73	SEPT'74	SEPT'75
A. Staff									
1. Full Time	5	11	14	15	18	18	18	20	21
2. Part Time	5	9	10	11	12	14	16	12	12
B. Graduate Students	32	89	114	151	165	187	209	198	209 (est)
C. Undergraduate Students	100	143	300	485	576	450	510	475	490 (est)
D. Course Enrollment (Autumn Quarter)	542	770	1059	1293	1447	1676	1728	1925	2050 (est)

	'67-'68	'68-'69	'69-'70	'70-'71	71'-72'	72'-73'	73'-74'	74'-75'	75'-76
Students Taught	1977	2892	3933	4703	5174	5600	6129	6876	7500 (est)
M.S. Degree Awarded	7	17	35	44	47	49	67	58	55 (est)
Ph.D. Degrees Awarded					4	8	4	7	18 (est)
Applications for Graduate Study	181	190	343	425	400	323	290	355	
Number of Graduate Students Supported	27	72	78	88	89	83	78	81	

The undergraduate program in the College of Administrative Science leads to the degree of Bachelor of Science in Business Administration with a major in computer and information science. This program is designed for the student that is business oriented and desires an education in computer and information science and a general education in the administrative sciences. The program's objective is not to make a computer specialist out of a student, but rather to enable him to recognize the opportunities to use the computer in his managerial activities, to know what to expect from it, and to know how to communicate effectively with computer specialists so that computerized projects will be properly handled from a technical as well as a managerial point of view.

Graduate Programs

The Department of Computer and Information Science offers graduate programs leading to both the Master's and Ph.D. degrees. The graduate program leading to the Master's Degree is available in seven options.

Option I for the student desiring a theoretical foundation in computer and information science.

Option II for the student specializing in information systems.

Option III for the student specializing in computer systems.

Option IV for the student specializing in numerical analysis.

Option V for the student specializing in operations research.

Option VI for the student specializing in biomedical information processing.

Option VII for the student specializing in administrative science.

Each of these options provides a background in several aspects of computer and information science, as well as additional mathematical sophistication appropriate to the student's interest. Each of the options may lead to the Doctoral program in computer and information science, and each may be taken with a thesis option or without a thesis option. (See Appendix A for a listing of courses by number and title.)

All courses of study at the Master's level require completion of a core program in computer and information science, together with the required courses specified for one of the options and additional courses as specified by the student's adviser. The core program includes courses on; Principles of Man-Machine Interaction, Numerical Analysis, Data Structures, Advanced Computer Programming, Digital Computer Organization, Mathematical Foundations of Computer and Information Science, Introduction to Linguistic Analysis, Modern Methods of Information Storage and Retrieval, and Advanced Seminar in Computer and Information Science.

The graduate program leading to the Doctoral Degree in Computer and Information Science is flexible in that it is tailored to the particular background and interests of the individual student. These interests may lie in any one of the research and instructional areas already listed as well as in many

other cognate areas. A cognate field is defined as a field supporting or closely related to the fourteen Departmental fields and is ordinarily specified by an integrated program of study in other departments of the University.

Course Offerings

Currently there are about 81 courses (each one quarter in length) offered by the Department, 22 of which are largely undergraduate with the remainder being upper level undergraduate and graduate courses. In addition to these courses there are over two hundred courses offered by a variety of departments of the University which are of interest to our graduate students who are encouraged to take these courses.

Faculty

The Department of Computer and Information Science has a full time faculty of twenty-one members at the assistant professor level and above. They have a wide range of backgrounds and experience. The above faculty is supplemented by staff who have joint appointments with other departments; by staff from other departments who teach courses primarily for Computer and Information Science students; and by adjunct staff people who are employed in off campus organizations who teach courses in the Department of Computer and Information Science (see Appendix B). There are currently a total of about 14 supplemental staff in this category.

INTERACTION WITHIN THE UNIVERSITY

Both the Research Center and the Department of Computer and Information Science interact with other departments and research programs within the University. This is essential because of the multi-disciplinary nature of the activities encompassed in this field. A number of the academic faculty have joint appointments in other departments. Staff members of the Department of Computer and Information Science have appointments in the following departments and organizations:

a. Accounting
b. Allied Medicine
c. Art
d. Biophysics
e. Electrical Engineering
f. Engineering Graphics
g. Instruction and Research Computer Center
h. Mathematics
i. Psychology
j. University Libraries
k. University Systems Computer Center

INTERACTION WITHIN THE COMPUTER AND INFORMATION SCIENCE COMMUNITY

Columbus, Ohio is one of the major centers for information science and for the transfer of information in the United States. A number of organizations are involved with the activities of computer and information science. This affords an opportunity for students and faculty to interact with appropriate personnel in these organizations. Some of these are:

a. Chemical Abstracts Service
b. Battelle Memorial Institute
c. Bell Laboratories
d. City National Bank
e. Columbus and Southern Ohio Electric Company
f. Western Electric Corporation
g. Rockwell International Corp.
h. Industrial Nucleonics
i. State of Ohio Department of Finance; Department of Highways
j. Columbus Board of Education.
k. Ohio College Library Center

There are a large number of scientists who come to Columbus in order to visit with the Department and Center and who usually present a seminar. (The seminars for the period of this report are listed in Appendix C.) The people cover virtually all phases of computer and information science.

In addition, our people interact at most of the major technical meetings in this country as participants giving papers, assisting on panels, as attendees, and as officials. Hardly a major technical meeting in the appropriate fields is held without a contribution from one or more of the personnel from the Ohio State Computer and Information Science Research Center. A list of these activities can be found in Appendix D.

Research efforts of the staff are disseminated to the professional community through several publication channels. A list of current publications of the Research Center staff is included as Appendix E. In addition, the Research Center issues a technical report series (see Appendix F).

DOCTOR OF PHILOSOPHY DEGREE

The Doctor of Philosophy degree was awarded to the following students during 1974-75. See Appendix G for a complete listing of Ph. D. dissertations.

Name	Dissertation
James L. Beug	Human Extrapolation of Strings Generated By Ordered Cyclic Finite State Grammars
Michael E. Doherty	A Heuristic For Minimum Set Covers Using Plausability Ordered Searches
Serge Fournier	The Architecture of a Grammar-Programmable High-Level Language
Oluwumi Longe	An Index of Smoothness for Computer Program Flowgraphs
Edwin J. McCauley	A Model for Data Secure Systems
Frederick E. Petry	Program Inference From Example Computations Represented by Memory Snapshot Traces.
Hui-Yang Su	Pagination of Programs for Virtual Memory Systems

RECENT TECHNICAL REPORTS

The Computer and Information Science Research Center began publishing a technical report series in 1968. A list of recent reports of the research center follows (see Appendix F for complete list.)

BIERMANN, A.W.: KRISHNASWAMY, R. Constructing programs from example computations. August 1974. [41]p. (OSU-CISRC-TR-74-5)

BUTTELMANN, H.W. Semantic directed translation of context free languages. September 1974. 36p. (OSU-CISRC-TR-74-6) (PB-242 854/AS)

HARTSON, H.R.: HSIAO, D.K. Languages for specifying protection requirements in data base systems (Part I). January 1975. 61p. (OSU-CISRC-TR-74-10). AD/A-006 280/2GI

HORGER, Jr.,W.A. Data base module verification - a certification method for data secure systems. June 1975. 125p. (OSU-CISRC-TR-75-3)

HSIAO, D.K.: BAUM R.I. Information secure systems. November 1974. 55p. (OSU-CISRC-TR-74-9)

HSIAO, D.K.: McCAULEY III, E.J. A model for data secure systems (Part II). October 1974. 40p. (OSU-CISRC-TR-74-7) (AD/A 506 276/0GA)

HSIAO, D.K.: KAFFEN, N.E. A model for data secure systems (Part III). November 1974. 32p. (OSU-CISRC-TR-74-8) (AD/A-004 676/3GI)

McCAULEY III, E.J. A model for data secure systems. March 1975. 119p. (OSU-CISRC-TR-75-2) (AD/A-011 359/7GI)

PETRY, F.E. Program inference from example computations represented by memory snapshot traces. February 1975. 143p. (OSU-CISRC-TR-75-1)

II. INFORMATION STORAGE AND RETRIEVAL

AUTOMATIC PROFILE ENHANCEMENT BASED ON ASSOCIATION MEASURES DERIVED FROM SDI SEARCH PROFILES

The use of association measures to improve information retrieval performance has been explored to a considerable extent in the past decade. Much of the work done to date has been based on small, static document collections where the documents in the collection have been used as a basis for deriving the association measures. However, this approach appears to be impractical for large or dynamic document collections such as those used in many SDI (selective dissemination of information) systems.

An interesting alternative, which is being explored in this research, makes use of the intellectual effort expended on manual preparation of a fairly large, but relatively stable group of profiles used for an SDI system providing current awareness and retrospective search services on both large and dynamic document collections. Accordingly profile-profile associations and term-term associations derived from a profile-term matrix are being investigated to determine their usefulness for automatic profile enhancement of both existing profiles and new profiles entering the system.

A. E. Petrarca, G. J. Lazorick, B. J. Brinkman (Sponsor: National Science Foundation. GN 27458)

A CONCEPT SCATTERING FUNCTION FOR EVALUATING THE EFFECTIVENESS OF VOCABULARY CONTROL ALGORITHMS USED IN AUTOMATIC INDEXING

To improve vocabulary control in indexes derived automatically from natural language text it is desirable to have an objective and quantitative method of evaluating the effectiveness of different vocabulary control techniques. Automatic indexes produced without vocabulary control are generally beset by a certain amount of concept scattering resulting from the inflected forms of words or phrases representing the various concepts being indexed, and in extreme cases, some inflected forms pertaining to one concept are completely separated by inflected forms of words unrelated to that concept. This last phenomenon is illustrated by the alphamerically ordered set of keywords *plan, planar, planarity, planning, plans*, where the words denoting the concept of *planning* are separated by those denoting the concept of *planarity*.

The disorder associated with concept scattering represents a form of index entropy. Accordingly, a concept scattering function has been developed which is intuitively based on the difference in index entropy between any given index and its corresponding ideal index (having zero index entropy). Appropriate ideal indexes have been constructed and computer programs have been written to calculate values for this scattering function on indexes subjected to different automatic vocabulary control techniques. The results obtained thus far clearly

provide a quantitative and objective evaluation of the relative effectiveness of the particular vocabulary control algorithms studied. The concept scattering function is now being used for selection of a superior vocabulary control algorithm to be incorporated into an automatic indexing system.

A. E. Petrarca, W. S. Stalcup

A DISTANCE MEASURE FOR AUTOMATIC DOCUMENT CLASSIFICATION

This research has concentrated on the design of an efficient algorithm for automatic sequential document classification. It is assumed that a set of categories and a set of keywords describing these categories is given. Each keyword is viewed as a k-dimensional vector where k is the number of categories. This enables a document to be represented as a set of points in a k-dimensional space. A distance measure, called the Bayesian distance, is then defined on this space, and consists of a magnitude and a direction. It is shown that by studying the variation of this magnitude and direction as keywords are read from a document, noisy or inappropriate keywords can be isolated and clusters of similar keywords can be identified. By analyzing these clusters of keywords, primary and secondary classes for a document can be obtained. This algorithm has been implemented on the SPIN data base and encouraging results have been obtained.

L. J. White, G. Kar (Sponsor: National Science Foundation, Office of Science Information Service) GN 36340

GENERATION OF UNIQUE CONFIGURATIONAL REPRESENTATIONS OF CHEMICAL STEREOISOMERS FROM NON-UNIQUE UNAMBIGUOUS 3 DIMENSIONAL INPUT DESCRIPTIONS

During the past few decades, many techniques for computer-based storage and retrieval of chemical structural information have been developed, several of which are now being utilized for creation of large files of structural data. The two most prominent stores of such data are those being created and maintained by Chemical Abstracts Service (CAS) and the Institute for Scientific Information (ISI).

The basic objective of creating such stores is to enable researchers to interrogate them in various ways to obtain information which is difficult to retrieve by conventional printed indexes. Such interrogations may take the form of requests for (a) information on compounds having a common set of structural characteristics (e.g., for studies on structure-activity relationships); (b) information on compounds having a desired set of chemical, physical, or biological properties needed for some specific application; or (3) information on precursors needed for the synthesis of a desired chemical structure. The data bases also provide many interesting opportunities for pattern recognition studies.

There are several prototype systems available, or under development, to provide search capabilities such as the above, utilizing data bases such as those provided by CAS and ISI. These systems permit sophisticated substructure searches (i.e., searches for subgraphs of chemical structures) on most 2-dimensional aspects of chemical structure. However, stereospecific substructure searches

cannot be readily performed on most of the large data bases available because of the way in which the stereochemistry (i.e., the 3-dimensional structural relationships) is handled. The CAS and ISI data bases, for example, handle stereochemistry in a manner suitable primarily for making global distinctions between two molecular structures whose 2-dimensional representations are isomorphic.

Because of the importance of local as well as global stereochemistry in many areas of research (particularly in medicinal chemistry, biochemistry, etc.) a prototype system for handling these relationships has been developed. The system generates unique global and local descriptions for each structure exhibiting certain types of stereoisomerism in terms of a global operator combined with a set of unique configurational descriptors--one for each local site contributing to the overall stereochemistry of the molecule. The configurational descriptors are derived from non-unique, unambiguous input descriptions. The prototype system, which is an adaptation of that used by CAS, is currently being tested and extended to cover a wider range of geometrical configurations encountered in chemical structures.

A. E. Petrarca, D. G. Williams, K. J. Wells

IDENTIFICATION & MEASUREMENT OF COST & PRICING ELEMENTS OF COMPUTING RESOURCES

Dual economic measurement and allocation criteria for computing, "cost" and "price", are necessary due to the increasingly multitasking, shared resource characteristics of computing. "Cost" relates to resource utilized: "price" relates to the placement of value (utility) on a service rendered. Resource utilized is not necessarily congruous to service rendered; cost allocation and pricing may require separate, distinct measurement and allocation mechanisms. A methodology is being developed to identify cost and pricing centers and to synthesize measurement and allocation methods on the basis of that identification. In conjunction with this methodology, declarative and procedural capabilities will be added to the Information Processing System Simulator (IPSS) to provide the modeler with the capability to develop cost and service statistics in a manner similar to response time and queuing.

D. Isaacs, T. G. DeLutis (Sponsor: National Science Foundation. GN36622).

IMPROVEMENT OF AUTOMATIC VOCABULARY CONTROL IN INDEXES DERIVED FROM NATURAL LANGUAGE TEXT

One of the main purposes of vocabulary control in an indexing system is to reduce the uncertainty in locating information on a particular topic. This effect is generally achieved by restricting the number of synonyms or near synonyms used to describe a particular topic, and by providing cross references to the preferred synonym chosen to represent each topic.

Unfortunately most automatic indexing systems are totally lacking in vocabulary control so that information on a given topic may be scattered among many inflected forms of a word (or phrase) as well as under synonyms having totally different word roots. In extreme cases, some inflected forms pertaining to one concept are completely separated from each other by words unrelated to

that concept. This can be seen in the alphabetically ordered keyword sequence *rat, rate, rating, ratio, rats* where the words *rat* and *rats* are scattered from each other by unrelated keywords.

Automatic vocabulary control, like manual control, can be attained by replacement of the original natural language keyword by a preferred word looked up in a thesaurus or dictionary. However, there are many situations where thesauri are unavailable or too small for adequate vocabulary control, and other situations where the thesauri are too large to conveniently store, maintain, or use very efficiently. Many of these problems can be eliminated or minimized by stemming-recoding algorithms which not only recognize conceptually related keywords via removal of their inflectional endings, but create preferred index words by appending suitable suffixes to the word roots obtained from the stemming process. The use of such algorithms for improved automatic vocabulary control in an automatic indexing system is being explored in this research. The effectiveness of each algorithm chosen for study is being evaluated by a concept-scattering function (see separate abstract) which was designed to provide an objective and quantitative evaluation of techniques for providing automatic vocabulary control.

A. E. Petrarca, W. S. Stalcup

THE INFORMATION PROCESSING SYSTEM SIMULATOR -- IPSS

The Information Processing System Simulator (IPSS) will be the realization of a methodology developed for investigating the behavior of complex computer based information processing systems. IPSS is being designed to deal with the problems associated with Information Storage and Retrieval systems, and to serve three distinct users: the researcher, the information system analyst/designer, and the student. One important research area for which IPSS will be especially useful is the investigation of suitable data structures/data manipulation operations for generalized Data Base Management Systems (DBMS). The most important departure from current simulation packages is its emphasis on the interrelationship between the application, the system software and the I/O process. IPSS can be used to describe I/O activity for a wide variety of computer systems and applications. It is being designed in a modular fashion to assist continued development of both language and run control facilities, and to permit the simulation time system to be other than IBM S/370's.

T. G. DeLutis (Sponsor: National Science Foundation, Office of Science Information Services. GN 36622)

A METHODOLOGY FOR A MULTIPLE GOAL APPROACH TO COMPUTER SYSTEMS DESIGN

The computer systems design process is formalized for more effective analysis. The formalization follows a normal problem solving approach (goal determination, etc.) but requires quantification of goals and constraints at each level. Because the design of a computer system inherently has many levels of conflicting objectives, the multi-level, multi-goal structure of goal programming was selected for the methodology. This approach forces a designer to determine the value of his goals before testing his design, rather than

molding them to fit his results. The output of the goal programming approach yields more effective information to the decision-maker (designer) than other linear programming techniques. The goal programming appraoch will be extended to utilize data generated by the Information Processing System Simulator (IPSS) about a given computer system model, and then indicate which variables need modification in succeeding runs.

J. S. Chandler, T. G. DeLutis, (Sponsor: National Science Foundation. GN-36622).

OCCURRENCES OF TRAILING N-GRAMS AS A FUNCTION OF DATA BASE SIZE

To facilitate the development of generalized automatic procedures for the removal of suffixes from natural-language words, the occurrences of trailing n-grams in natural-language data bases are being studied. Digram and trigram occurrence counts have been obtained for 7 graduated-size samples of words taken from the natural-language titles of documents cited in a multidisciplinary data base. The 7 samples contain approximately 1000, 3333, 10000, ..., 10^6 title words. A plot of digram occurrences as a function of sample size clearly shows asympotic behavior, with likely occurrence of all theoretically possible digrams in a 10^7-word sample of natural-language titles. However, a similar plot of trigram occurrences suggests that a data base of approximately 10^9 title words would be required for the theoretical upper limit of trigram patterns to be reached, if at all. The data collected in this study is being used for evaluation and improvement of stemming-recoding algorithms to be used for automatic vocabulary control in indexes derived from natural language text (see separate abstract).

A. E. Petrarca, W. S. Stalcup (Sponsor: National Science Foundation. GN 27458)

PERFORMANCE EVALUATION OF DATA BASE MANAGEMENT SYSTEMS

A methodology for the performance evaluation of data base management systems is being developed. A data base management system (DBMS) is characterized in terms of: (1) the structure and contents of the data base, and (2) the capabilities provided by the DBMS for the creation, update, minipulation, and accessing of data in the data base. The methodology consists of declarative and procedural components for the characterization of a DBMS. There are declaratives for the logical description of a data base (in terms of the relational structure of the data and the desired access paths to the data). Declaratives and procedures are provided for mapping a logically-described data base to physical files. Special data management primitive operations are provided to characterize in a procedural manner the data manipulation facilities of a DBMS. The methodology is sufficiently general to model a wide variety of DBMS designs.

J. A. Aitken, T. G. DeLutis, (Sponsor: National Science Foundation. GN 36622).

III. HUMAN INFORMATION PROCESSING

HUMAN EXTRAPOLATION OF STRINGS GENERATED BY ORDERED CYCLIC FINITE STATE GRAMMARS

The problem of systematically investigating human behavior in relation to the complexity of sequential concepts can be decomposed into that of devising a suitable measure of human behavior and a measure of complexity for sequential concepts. In this study, human learning was measured in terms of errors made in extrapolating character strings. The strings used were ordered in terms of structural parameters of the grammars used to generate the strings. The structural parameters used were the size of the terminal alphabet and the level of embedding or chain length of the generating grammars. Two experiments were performed using human subjects in a computerized implementation of a string extrapolation paradigm called the character prediction task.

In the first experiment, it was found that increases in the size of the terminal alphabet and the level of embedding of the generating grammars produced statistically significant increases in the number of trials to last error and the number of errors made in extrapolating the strings. In the second experiment, it was found that increases in the chain length of the generating grammars produced statistically significant increases in the number of errors made in extrapolating the strings, corrected for those trials where there was insufficient information to correctly predict the strings. Finally, for all grammars with 12 characters in their terminal alphabets, a correlation of 0.695 was found between the corrected number of errors and the chain length of the generating grammars.

These results show strong agreement between difficulty of learning the strings and the complexity of the strings defined by the generating grammars. Cyclic finite state grammars of the kind used in this study may thus be considered as reasonable representations of tasks involving the attainment of sequential concepts where difficulty is a parameter.

J. L. Beug, R. L. Ernst

IV. INFORMATION ANALYSIS

AN EMPIRICAL ANALYSIS OF INFORMATION AND RELATIONS AMONG ELEMENTS OF A DECISION SYSTEM

The relationship among elements of a decision system is being analyzed by measuring the effect of different information displays on a decision task. In the decision task, a decision maker (DM) was placed in the role of director of an information retrieval system with the objective of maximizing the number of users of the system. The DM specified relative weights of outcomes, action-outcome pair probabilities, and course of action selection for the time frame of twelve consecutive decision periods. The information display was a medium for feedback of results of DM response after each decision period. Differentiation among information displays was made on the basis of level of information presented and combinations of these levels. Variation in information display and the associated performance is viewed as a transformation of data into different levels of information.

Three system functions transformed DM responses into three levels of information defined as follows: operational - information to aid the DM in operational or procedural control; management - information to aid the DM in managerial or policy control; and strategic - information to aid the DM in top level control. The system functions so defined imply that the information content of operational level information is contained in management level which in turn is contained in strategic level.

Three sets of values related to DM responses were examined: a) the actual pre-defined ordering of courses of action as specified by the system (i.e., state of nature), b) the DM-specified ordering of courses of action, and c) the DM course of action selection. System and DM ordering of acts were based on expected value computations which were determined from act-outcome pair probabilities and relative weights for the outcome set.

Several techniques were applied in the analysis of DM responses. DM performance measures were developed to determine the deviation from the optimal decision state as specified by the system. In addition, a mathematical model of concept learning was used to define and categorize DM responses for two methods of analysis. For the first method, this model evaluated DM responses over all twelve decision periods as steady-state transition probabilities. For the second, Bayesian techniques were used to analyze the rate of convergence to the optimal decision state from one decision period to the next. Thus, the impact of information display on DM responses was in terms of Markovian and Bayesian probabilities respectively. Both methods analyzed DM responses with respect to the prevailing state of nature as defined by the system and to the DM's perception of that state of nature.

The decision task and the information displays were incorporated into an

interactive simulation. A prototype model of this simulation was developed earlier in the form of a management game supported by the Mechanized Information Center of The Ohio State University and was reported elsewhere.

M. J. Lee, R. L. Ernst

A GENERAL THEORY OF INFORMATION FLOW AND ANALYSIS

A generalized framework for the development of a theory of information flow which permits the analysis and quantification of information has been suggested. In addition to its theoretical and conceptual interest, there are major and immediate implications for the development of information systems and networks as well as for the general understanding of information flow, retrieval, and transfer.

Numerous intuitive notions exist about the interrelationship between information and decision-making. At the pragmatic level, information has value to the extent that it is useful as a resource for purposeful activity. The primary "purposeful activity" in life is decision-making. Hence, information and decision-making are inextricably tied together. In our formulation, in fact, information is defined in just that way, as being data of value in decision-making.

As a consequence, a measure of the amount of information in a data set or message is defined in terms of a quantity called the decision state of a decision-maker. The decision state is a function of the determinism of the decision-maker, i.e., how easy it is for him to make a decision in a particular decision situation. We suggest a way of evaluating the decision state quantitatively.

We assume that the decision-maker can compute expected values for each alternative. The relative expected value of each alternative can then be computed and normalized so that the sum of all the relative expected values is 1. The *value of the decision state* is then defined as the summation of the expected values of all the possible courses of action weighted by the relative expected value of each course of action. A new measure for the *information contained in a particular decision state* is developed:

$$I = m \sum_{i=1}^{m} \{REV(a_i)\}^2 - 1,$$

where m is the number of courses of action available to a decision-maker and $REV(a_i)$ is the relative expected value of each course of action. The information is defined in terms of a two-choice deterministic situation which we call a "binary choice unit". This measure is universally applicable for all information that is concerned with the effectiveness of the data upon the recipient.

A measure of the *amount of information in a data set or message* can be arrived at by computing the difference in the amount of information in the decision state before and after receipt of the data. That is, the amount of information is arrived at by considering the impact this new data has on the

decision-maker's decision state. In symbolic terms, $I(D)$, the amount of information in data set D, is

$$I(D) = I_{t+1} - I_t,$$

where I_{t+1} and I_t are the amounts of information in the decision state after and before receipt of the data set.

The suggested information measure leads to a measure of the pragmatic information content of a data set for a *particular* decision-maker at a *particular* point in time. The data acquired, processed, stored, and disseminated by an information system may be used, however, as a resource by *various* decision-makers at *various* points in time. Hence, in the design and development of information systems, there exists a problem whose level of complexity is an order of magnitude above that of the primary problems addressed in this study -- the problem of quantifying the information contained in a data set in terms of its overall usefulness for a *range of decision-makers over a period of time.*

One possible approach to this problem of assigning a number to a data set to indicate its composite information content would be to start by determining the relationship between the effectiveness of a decision-maker and the information content of the data set. Since what is really desired is some indication of the information content of this data set for this decision-maker over a period of time, one may determine some index $I(D)$ of the average information contained in data set D over some period of time. Then, if one were to develop a measure of the effectiveness of each of the decision-makers for whom this data set serves as a resource, it would be possible to formulate an *information profile* for the data set. Such an information profile would indicate the average information content of a data set as a function of a decision-maker-effectiveness.

If such a profile could be determined for every data set to be stored in an information system, then some number derived from this profile could serve as an index of the composite value of this data set. This method would be of major importance for the development of a sound procedure for the design of more effective information systems.

M. C. Yovits, J. G. Abilock (Sponsor: National Science Foundation. GN 31628)

STRUCTURAL MODELS FOR DECISION SYSTEMS

This research is a continuation of previous efforts concerned with the development and application of generalized information systems model. Its major focus is on the structural properties and organization of elements of decision systems. To date, the research has been confined to the formalization of such systems.

A decision system is a special case of an information system that possesses connected information acquisition and dissemination, decision-making, execution, and transformation functions as in the Yovits and Ernst Generalized Information System model. The major distinguishing characteristics of a decision system are that the decision-making situation is well defined such that the elements of the

decision task may be represented as elements of the information system. For a standard analysis of a decision task into acts, states, events, probabilities of states, values of act-state pairs, and consequences, correspondences between the decision elements and the Generalized Information System are: (a) information acquisition and dissemination--under environmental and prevailing decision-making policy (i.e., partitions of the set of consequences) inputs, an act state matrix whose values are pairs of state probabilities and values of acts conditional upon the occurrence of a state is generated, (b) decision-making--given the act-state matrix and its associated values, and a choice rule (relation), a subset of pairs within the matrix is defined and outputted as the course(s) of action, (c) execution--given the course(s) of action, a system consisting of a subset of the environment produces a subset of states (events), and (d) transformation--under the prevailing decision-making policy and environmental context, the event(s) are associated with a particular act generating an act-event pair. Partitions over possible act-event pair sets and the associated measures of the partitions define the decision state(s) for a particular decision system. The states so defined also specify the prevailing context of the environment and decision-making policy (consequence set partitions).

These correspondences have been cast as a system of reticulated linear stochastic sequential machines whose kernel is the act-state matrix. Preliminary investigation of the system has shown interesting tentative results bearing on information acquisition and utilization (learning), and uncertainty in decision-making when approached through mathematical models of learning and Bayesian probability analysis. Moreover, the research shows considerable promise for developing various design criteria for some kinds of command and control decision systems, and developing their simulation.

R. L. Ernst, M. J. Lee

V. LINGUISTIC ANALYSIS

THE AFFECTS OF PHYSICAL ENVIRONMENT ON THE INTERPRETATION OF LANGUAGE

The aim of this work is to explore the ways in which the spatial context of a speech act is involved in the assignment of an interpretation to the utterance and to allow for such involvement in natural language understanding systems. Linguistics have studied aspects of the interaction of physical environment and meaning under the title of space-deixis. Such work has centered on accounting for the affect the location of the speaker and addressee can have the interpretation of an utterance. However, the location of other objects can also have influence. This is seen when the orientation of one object affects the interpretation of spatial relations with respect to another, such as when a theatrical stage defines what is to the left of an actor without consideration of his orientation.

The study is also considering the ways in which objects take on spatial descriptions. This includes such conventions as assuming orientation based on position when utilized by humans and assuming orientation based on object movement.

The phenomena studied bear importantly on natural language man-machine interaction in such application as machine control and computerized consultants for physical tasks. One way to allow for them is to sufficiently restrict the linguistic forms such that the point of view and the spatial descriptions are well defined. One such set of English forms with appropriate rules of interpretation has been identified. A possibly preferable means of operation would be to allow the machine access to some of the nonlinguistic sources of information which humans utilize. Work to this end is now in progress in connection with other projects here at The Ohio State University.

N. K. Sondheimer

A FAMILY OF TRANSLATORS FOR PHRASE STRUCTURE PROGRAMMING LANGUAGES

Formal properties of translation between high-level programming languages are studied. The basis for this study is a formal model of programming languages called a "phrase-structure system" which has both a formal grammar and a formal semantics. The sets of sentences of phrase-structure systems are just the recursive sets, and the meanings of the sentences are the recursively enumerable sets. We define and study a family of "phrase-structure translators". Translation can either be "syntax-directed", or "semantic-directed" or a combination of both. Phrase-structure translations fall into the latter category. For each member of the family, we give sufficient conditions for the "correctness" of the translation. We show in certain cases that for a phrase-structure translation from language A to B, an "inverse" phrase-structure translation from B

to A may be effectively constructed. We prove a kind of "pumping" theorem for phrase-structure translations and use it to show that for certain members of the family and certain languages (eg., binary to ternary notation), no translators exist. Next we prove that for any two recursive phrase-structure languages, A and B, where there is an effective translation T from A to B, if there is a constant k such that for all w in A, k*length(w) < length(T(w)), then there is a phrase-structure translator which computes T. Finally, we investigate the relative time complexities of the members of the family of translators.

H. W. Buttelmann, A. Pyster (Sponsor: Air Force Office of Scientific Research 75-2811.)

FORMAL GRAMMARS WITHOUT SYNTACTIC VARIABLES

We examine the languages of pure grammars, first studied by Gabrelian (unpublished). Pure grammars are formal grammars with no nonterminal symbols (except the axiom, which may not appear on the right hand side of any production).

A hierarchy of grammars and languages, parallel to that of Chomsky's, is formed and investigated. The properties of the pure hierarchy are markedly different from those of the Chomsky hierarchy and many results on these differences are proven. In addition, we discuss some of the functions of nonterminals and why not having them in pure grammars affects the generated languages.

H. W. Buttelmann, A. Pyster, L. Reeker

A FORMAL THEORY OF THE SYNTAX, SEMANTICS AND TRANSLATION OF PHRASE-STRUCTURE LANGUAGES

A formal definition for a semantics for phrase structure grammars, called a phrase structure semantics, has been developed. It is a model of the following semantic philosophy: 1) it is phrases which have meaning, and 2) the meaning of a phrase is a function of its syntactic structure and of the meanings of its constituents. A pair (G,S) where G is a phrase structure grammar and S is a phrase structure semantics, is called a phrase structure language description. A language is not just a set of sentences, but a set of sentences with meanings assigned to them. The language of a psld is the set of all pairs (x,m) such that x is a sentence of the grammar and m is a non-empty meaning assigned to x by the language definition. The sets of sentences of the languages of psld's are just the recursively enumerable sets. It has been shown that for any psld with a type 0 or type 1 grammar, there exists a psld with a context free grammar that defines the same language. Translation is defined on the languages of psld's, and it has been shown that the translation function of the languages of arbitrary pairs of psld's is effectively computable.

H. W. Buttelmann (Sponsor: Air Force Office of Scientific Research 75-2811)

ON THE SYNTACTIC STRUCTURES OF UNRESTRICTED GRAMMARS. I: GENERATIVE GRAMMARS AND PHRASE STRUCTURE GRAMMARS

Formal definitions for the syntactic structures of unrestricted grammars are given. The traditional forms for grammar productions give rise to "generative

grammars" with "derivation structures" (where productions have the form $\alpha \rightarrow \beta$), and "phrase structure grammars" with "phrase structures" (where productions have the form $A \rightarrow \beta / \mu\_\nu$), two distinct notions of grammar and syntactic structure which become indistinguishable in the context free case, where the structures are trees. Parallel theories are developed for both kinds of grammar and structure. We formalize the notion of structural equivalence for derivations, extended to unrestricted grammars, and we prove that two derivations are structurally equivalent if and only if they have the same syntactic structure. Structural equivalence is an equivalence relation over the derivations of a grammar, and we give a simpler proof of a theorem by Griffiths that each equivalence class contains a rightmost derivation. We also give a proof for the uniqueness of the rightmost derivation, following a study of some of the properties of syntactic structures. Next, we investigate the relationship between derivation structures and phrase structures and show that the two concepts are non-isomorphic. There is a natural correspondence between generative productions and phrase structure productions, and, by extension, between the two kinds of grammars and between their derivations. But we show that the correspondence does not necessarily preserve structural equivalence, in either direction. However, if the correspondence from the productions of a phrase structure grammar to the productions of a generative grammar is a bijection, then structural equivalence on the generative derivations refines the image under the correspondence of structural equivalence on the phrase structure derivations.

H. W. Buttelmann

ON THE SYNTACTIC STRUCTURES OF UNRESTRICTED GRAMMARS. II: AUTOMATA

We define a generalization of the finite state acceptors for derivation structures and for phrase structures. Corresponding to the Chomsky hierarchy of grammars, there is a hierarchy of acceptors, and for both kinds of structures, the type 2 acceptors are tree automata. For $i = 0, 1, 2, 3$, the sets of structures recognized by the type i acceptors are just the sets of projections of the structures of the type i grammars, and the languages of the type i acceptors are just the type i languages. Finally, we prove that the set of syntactic structures of a recursively enumerable language is recursive.

H. W. Buttelmann

PRACTICAL ALGORITHMS FOR COMPUTER COLLATION OF LONG NATURAL TEXTS FOR SCHOLARLY USE

A collation of a text and its variant, say of a first and revised second edition of a novel, consists of an indication of the changes that were made from one text to the other. It is evidently desirable that the tedious task of preparing a collation be automated if possible; the scholar is more interested in using than in preparing collations. Unfortunately there are several problems which arise in attempting to automate the process. The data base in large, say twice 100,000 English words, and any inefficient algorithm could be expensive to use. The definitions of what constitutes a collation are somewhat subjective. Also the direct linear scan approach can easily get lost on texts that have received any but the simplest revisions.

We attack this problem by making precise definitions of global and local collation. We are developing an algorithm which realizes both definitions and uses newly developed techniques for fast pattern matching in attempting to produce useful collations at reasonable cost.

D. J. Moore

SOME PROPERTIES OF SYNTAX-DIRECTED TRANSLATIONS

We are investigating problems relating to the existence and computability of syntax-directed translations. The properties of the set of syntax trees generated by a finite generating set of trees, and maps on those sets of trees, are being studied, and we are investigating the following conjecture: Let L_a be a phrase-structure language of numbers written in radix a and L_b a phrase-structure language of numbers written in radix b. Then a finitely specified syntax-directed translation from L_a to L_b exists if $\log_a b$ is rational.

H. W. Buttelmann, F. J. Dickey (Sponsor: Air Force Office of Scientific Research 75-2811.)

SPS: A FORMALISM FOR SEMANTIC INTERPRETATION

This project is experimenting with a formalism, called SPS, for writing semantic processors for natural language understanding systems. SPS is intended for use in turning underlying syntactic structures in the form of constituent structure trees into underlying semantic structures in the form of nets composed of PLANNER-like assertions. The formalism is based on Woods-style "pattern → action" rules. The pattern element specifies tree fragments and various types of selectional restrictions. On the action side a variety of devices, including the use of registers, allow common reference to entities in the assertions produced. The registers used for reference can also be used to specify selectional restrictions across rules and for establishing default conditions for handling semantic ellipsis. Finally, SPS provides a control structure for the ordering of the application of the rules that interpret constituents and to control, in part, where the tree fragments are matched.

The power of SPS is seen in its unique ability to allow for the development of Case structures, especially the structures connected with the English prepositions that reference location, orientation and motion in space. These forms have always been troublesome for Case systems. Particularly difficult are the facts that 1) more than one of these prepositions can appear in a sentence in the same rule, 2) their appearance can correspond to the need for multiple predicator semantic structures, and 3) they exhibit complex distributional and semantic relations among themselves and with respect to other sentential elements. SPS can allow for each of these phenomena.

An interpretation system for SPS has been implemented in LISP 1.6 on the DEC System-10.

N. K. Sondheimer, R. Pardo, D. Perry

STRUCTURE OF COMPUTER AND NATURAL LANGUAGE ALGORITHMS

The ultimate aim of the research is to provide better techniques of storing and retrieving information in a computer data base; by developing a theory of the manner in which the information contained in such data bases is distributed. The contents of such data bases might be volumes of text written in some natural language, such as English, or algorithms written in some computer language, such as FORTRAN or PL/1. The basis of such a distribution theory for computer algorithms has already been proposed, but the area of applicability of the theory needs to be extended. Thus, the technique to be followed will include (1) extending the area of applicability of an existing distribution theory of computer algorithms, (2) developing a parallel theory applicable to natural language, and (3) implementing that theory via computer program and testing it against current natural language data samples.

S. H. Zweben

STUDIES IN DECIDABILITY, EXISTENCE, AND EFFECTIVENESS OF TRANSLATIONS ON PHRASE-STRUCTURE LANGUAGES

This research is directed towards an analysis of techniques of translation of formal languages. A formal model of semantics, from earlier work, is used to study various types of translation. Properties of the translators are studied, and conditions under which they translate are investigated. A set of conditions is provided in order to guarantee that the translators produce correct translations. It is proved that these conditions are a minimal set in order to provide this guarantee. Two types of translators are studied and compared. Finally, a study is made of a limited class of translators that can be algorithmically generated. An algorithm is developed to provide a translation mechanism which is a 'best possible' translator. It is then proved that no other translator of this type can translate any source language sentences that the translator produced by the algorithm cannot.

H. W. Buttelmann, R. Krishnaswamy (Sponsor: Air Force Office of Scientific Research 75-2811.)

STUDIES OF THE REPRESENTATION AND MANIPULATION OF LINGUISTIC INFORMATION

The computation and manipulation of information abstracted from linguistic forms for use in such tasks as Question-Answering or Robotics presents many interesting problems. This project is freely attacking whichever of these problems appear tractable. The representation of quantification, and spatial and temporal reference have been the problems most considered. The area of spatial reference has been the most successfully studied with results including a proposal for extensive revision of the Case deep structure representation. Work is also in progress on a theory of interpretation for semantic nets which would permit the nets to be considered as procedural information.

N. K. Sondheimer

SYNTAX-DIRECTED AND SEMANTIC-DIRECTED TRANSLATION OF PHRASE STRUCTURE LANGUAGES

The basis of this theory is a model, from earlier work, of language description called a phrase structure language description, which contains both syntax and semantic information. Certain translation algorithms for these languages are being studied, which promise to be more efficient than the general algorithm of earlier theory. These algorithms are either syntax-directed or semantic-directed, depending on whether they are controlled by syntax information exclusively, or by both syntax and semantic information. A particularly fast translation algorithm is a slight generalization of the syntax-directed translations, called a "phrase-structure translation," where the syntax control information can be specified in a simple, finite way. It has been shown that it is possible to use a semantic-directed translation scheme as a translator generator, to produce the finite specifications of a very fast phrase-structure translation, if such a translation exists.

Running computer programs of the translator and translator generator have been developed, and the programs are being used to develop and test definitions and translations of very simple languages.

H. W. Buttelmann (Sponsor: Air Force Office of Scientific Research 75-2811)

VI. ARTIFICIAL INTELLIGENCE

THE INTEGRATION OF GRAPHICAL AND SEMANTIC KNOWLEDGE

This work addresses the problem of integrating Graphical and Semantic Knowledge into a data-base for the support of Graphics and Natural Language Processing Systems. The goal of the project is to allow Graphical Information to be used in the interpretation of Linguistic forms, and Semantic information to be used in the interpretation and construction of Graphical Structures. Achieving this goal should facilitate man-machine interaction.

As part of this project, current programs and structures in the areas of Computational Semantics, Computer Graphics, and Computer Vision are being considered with the intention of isolating suitable structures and techniques or identifying the need for new ones. Suitable domains for use in development of a prototype system are also being considered.

D. C. Brown, B. Chandrasekaran, A. P. Lucido, N. K. Sondheimer

LOW-FREQUENCY RADAR AIRCRAFT DETECTION AND CLASSIFICATION

Previous research has demonstrated the feasibility of using multiple low-frequency radar returns for target classification. Simple object shapes have been successfully classified by linear techniques, but aircraft data poses greater difficulty, as in general such data are not linearly separable. Since two parameters specify the aspect angle, this data lies on a 2-dimensional surface in n-space, where n is the number of frequencies utilized. A bilinear approximation of this surface has provided interpolation capability, improved proximity information, and an intersection algorithm which determines whether the aircraft data is separable. Improved separability together with decreased susceptibility to additive noise was obtained by mixing horizontal and vertical polarization data at an optimum polarization angle. Studies have indicated the importance of a phenomenon known as bias, where in the identification of a test signal as one of two given aircraft, the probability of misclassification for one aircraft can be substantially larger for one aircraft than the other for high noise levels. Thus classification algorithms are currently being developed which jointly minimize the probability of misclassification and the bias effect. Another approach being investigated is multipoint classification, where the input is a sequence of independent radar measurements of the aircraft target to be identified. It has been shown that this approach can substantially reduce the effects of noise even for a modest sample size, and a number of classifiers are being investigated for this purpose.

L. J. White (Sponsor: Air Force Office of Science Research. Grant 69-1710)

PROGRAM INFERENCE FROM EXAMPLE COMPUTATIONS REPRESENTED BY MEMORY SNAPSHOT TRACES

Example computations are studied as an easy means of communicating a person's idea of an algorithm to the computer. A computational environment is developed in which the user can carry out a computation needing to express himself only by the changes in contents of the memory variables produced by the calculations. The sequence of these memory changes is captured by the system in a memory snapshot trace. Using the trace and the specification of the computational environment in which it was produced, two steps, decomputation and synthesis, are performed to infer the program from the given computation. Numerous experiments on the inference of various programs are described illustrating the use of the system.

F. E. Petry

A SYNTAX-DIRECTED METHOD OF EXTRACTING TOPOLOGICAL REGIONS FROM A SILHOUETTE FOR PATTERN RECOGNITION

The overall objective of the research is to develop a method for decomposing, hierarchically relating, and extracting topological regions from pictorial data. The objective is accomplished in two steps both of which necessarily depend on a binary-valued pictorial pattern--a silhouette.

In step 1, the picture decomposition, a silhouette, defined as a picture, can be decomposed into two types of topological regions: *surrounds* and *collages* (multiply-connected images). Each collage is then considered a picture unto itself, and it is, in turn, decomposed into surrounds and collages. The decomposition is directed by a top-down parser. The topological regions at any level of decomposition are related by the parser through a context-free grammar. Such a decomposition enables any picture to be expressed in terms of surrounds and collages where the collages are *solids* (simply-connected images).

In step 2, the solid decomposition, a solid is further decomposed into topological regions by separating it into disjoint subsets, which are formed by selecting points of separation that break the solid into meaningful subsolids. The problem is formulated in terms of a graph-theoretical problem of recursively selecting the weighted median and/or center nodes of a tree and its subtrees. A solid is first reduced to a tree by forming its skeleton or medial axis. This tree is then expressed as a weighted connection matrix from which the weighted median and/or center nodes of the tree can be calculated. The tree is then broken at the selected nodes forming subtrees (and submatrices) on which the decomposition procedure is repeated. As in step 1, the decomposition is directed by a top-down parser and the topological regions at any level of decomposition are related by the parser through a context-free grammar. The output of the parser is a string description which can be used to extract any indicated subset of the solid.

Software has been written to simulate the array processor and to form cell-group variant euler skeletons.

J. Amoss, K. J. Breeding, B. Chandrasekaran (Sponsor: Air Force Office of Scientific Research. Grant 72-2351)

VII: INFORMATION PROCESSES IN PHYSICAL, BIOLOGICAL AND SOCIAL SYSTEMS

DISCRETE COMPUTER SIMULATION MODELS FOR ANTHROPOLOGICAL STUDY OF SETTLEMENT AND SUBSISTENCE SYSTEMS

We attempt to apply the techniques of computer-modelling using discrete simulation to the problem of understanding patterns of settlement and subsistence of human groups. We concentrate on prehistoric populations which are small and isolated, and for which the archeological data are abundant.

A model is being developed for a particular aboriginal tribe that lived in the Glenwood locality of southern Iowa from 900 to 1300 A.D. During this period a transition occurred from dispersed to nucleated communities, and there is ample archeological evidence that the manner of subsistence changed concurrently. These transitions accompanied a continent-wide climactic transition.

The simulation parameters include climatic, ecological and kinship patterns. Hypotheses are being formed concerning the effects of these parameters on the settlement and subsistence patterns. They are to be tested on a SIMSCRIPT program for the model.

The main reason for the research is to study the usefulness of this type of modelling in anthropology, and to develop a methodology for constructing and validating such models. Thus careful attention is being paid to the problems of proper utilization of artifacts in the inference step, and the place of simulation models in this process.

D. J. Moore

VIII. MATHEMATICAL TECHNIQUES

AN ALGORITHM FOR SUB-OPTIMAL SET COVERS WITH CONFIDENCE MEASURES OF OPTIMALITY

An effective heuristic has been developed for the minimum set cover problem: min cx such that $Ax \geq 1$, $x = 0$ or 1, A binary. Two algorithms have been developed: one for the general costs problem (arbitrary c vector), and one for the cardinality problem (all c_i equal). Both algorithms are implemented in computer programs in FORTRAN and assembler. Test results for the cardinality problem exhibit very good performance on test problems from the literature. An analysis has been completed which can be applied to give a measure of confidence of obtaining a global optimum solution.

L. J. White, M. E. Doherty

BINARY STRINGS AND TOPOLOGY

As little time was available for this research during the past year, only one significant result was obtained. The procedure for generating strings (binary decimals) over (0,1) corresponding in the limit to points in a space admitting both metric and measure, was applied to the following case. Consider a rectangle in the Euclidean plan whose diagonal divides it into 30°-60°-90° triangles, and drop perpendiculars on that diagonal from the other two vertices. Identify hypothenuses of equal length for each of the two pairs of congruent triangles, whereby the original rectangle is split into a larger and a smaller rectangle, both similar to it. The points in the larger rectangle have their strings start with 1, the others with 0,. Continuing this process ad infinitum associates a unique point with every string. At every step dimensionality is preserved but connectivity increases. In the limit we have an infinitely connected "supertoroid", a 2-D manifold whose geometry and topology are mind-boggling. The discovery of such spaces and very preliminary studies of their properties comprise the year's activity in this area.

J. Rothstein

COMPARATIVE STUDIES TOWARDS THE PERFORMANCE EVALUATION OF SOFTWARE FOR SOLVING SYSTEMS OF NONLINEAR EQUATIONS

The success of algorithms for solving systems of nonlinear equations is very dependent on the nature of the particular system being solved. This research was directed towards parametrizing nonlinear systems of equations so that one could more effectively select which algorithm is best suited for a given system. The parameters investigated included the size of the system, the goodness of the initial approximation, the nonlinearity of the system, and the number of potential nonconvergent regions inherent in the system. Comparative studies were conducted

for six different algorithms. These studies showed that the above parameters did influence the efficiency and likelihood of success of each of the algorithms. Based on these comparisons, a user-oriented guideline for the choice of algorithm relative to these parameters was suggested, and a model for comparing algorithms for solving systems of nonlinear equations was developed.

In addition to parametrization, two additional related topics were persued. First, properties for a measure of the degree of nonlinearity inherent in a system of nonlinear equations were established and one such measure was developed and tested. Secondly, a special type of nonconvergent region was discovered and described; then, appropriate steps for identifying and dealing with such regions were established.

D. L. Kalmey, D. S. Kerr, L. J. White

MINIMUM-REDUNDANCY VARIABLE LENGTH CODING

This research is aimed at obtaining efficient solution techniques for solving prefix encoding problems involving cost associated code alphabets. An efficient algorithm has been developed for finding a most economical encoding for equiprobable messages using arbitrary size alphabets. For the case of non-equiprobable messages and a binary alphabet, two techniques have been developed--a selective enumeration technique that outperforms the previously known techniques, and an algorithmic technique that is better suited for large problems. Based on a preliminary study of the latter technique, it is conjectured to obtain optimal solutions under general conditions. Development of a proof of this conjecture is currently under study.

L. J. White, V. Santhanam

SELF-SOLVING QUASIGROUPS

Properties of groupoids satisfying the condition that ab=c implies a=bc have been investigated. They are necessarily quasigroups. The squares of all elements are both two-sided identities for the elements and idempotents. They exist for all n (n an integer, the order of the quasigroup), and for all n, are at least as numerous as groups of the same order. With any group $(G, \cdot)$ we can introduce the associated self-solving quasigroup $(G,*)$ defined by $a * b = c$ if and only if $a \cdot b = c^{-1}$. There are self-solving quasigroups which are not associated with any group for all $n \geq 3$. Only for $n = 2^m$, can a self-solving quasigroup be group, and it must be a direct product of cyclic groups of order 2. It is believed that these quasigroups can be useful in cryptography and data security.

J. Rothstein

SOME L^P-NORM MAXMIN PROBLEMS: ALGORITHMS AND APPLICATIONS

This research examines a few closely related maxmin nonlinear programming problems that have applications in such diverse areas as location theory, infor-

mation theory and pattern recognition. One investigation involves the location of a point in a given convex polyhedron which maximizes the minimum Euclidean distance from a given set of m points in the polyhedron. This problem is also equivalent to finding the minimal radius of m hyperspheres centered at the given points such that each point in the polyhedron is "covered" by one or more of the hyperspheres. The existence of a finite candidate set solution for the problem is demonstrated. The computational growth of an algorithm which generates this finite candidate set solution is shown to be m^{n+1} where m and n are the cardinality of the set of given points and the dimensionality of the space respectively. However, efficient algorithms are suggested for the problem in the plane, and in three dimensions when the convex polyhedron corresponds to the convex hull of the given points. Several heuristics were employed in achieving this efficiency. The potential relevance of the above problem to pattern classification is shown. In particular, algorithms are suggested to generate nearest-neighbor decision surfaces for an arbitrary number of disjoint pattern sets in two and three dimensions as an application of the above maxmin problem.

L. J. White, B. Dasarathy

IX. SYSTEMS PROGRAMMING

THE ANIMA GRAPHICS PROGRAMMING LANGUAGE

In conjunction with Professor C. Csuri, Department of Art and Director of the Computer Graphics Research Group, the graphics programming language ANIMA is being developed. ANIMA users will be able to:

1. Treat picture descriptions in a fashion analogous to the treatment given other structured data types, such as numeric array;

2. program in a device independent manner, in that a file and a graphic input device may be used interchangably to assist in, say, picture definition;

3. have a compiler available for the language. Also, the intent is to base the language implementation on a minicomputer.

With the inclusion of a data type specifically for the storage of pictures descriptions, and verbs in the language that are patterned for effective manipulation of these data types, ANIMA becomes a true graphics programming language.

A. P. Lucido (Sponsor: National Science Foundation.)

COMPILATION OF SEQUENTIAL PROGRAMS FOR PARALLEL EXECUTION

This research derives techniques for the practically optimal compilation of a subclass of the class of acyclic programs for execution on an idealized parallel computer. The computer, partially described by Hellerman in 1966, has a single central processing unit controlling several arithmetic units. Each arithmetic unit can perform any of the computer's arithmetic instructions in one time quantum. Completion of a design optimal for a given work load would require information on the number of arithmetic units required for minimal time execution of programs representative of that work load, information derivable from optimal compilation of programs. Even in the absence of a "Hellerman computer" compilation of sequential programs for it could give insight into the degree of parallelism inherent in the programs, setting apart array parallelism.

Hellerman offered two techniques for the compilation of single algebraic expressions: earliest stage assignment and latest stage assignment; preferring the latter to minimize simultaneously the number of time quanta and the number of arithmetic units required for execution of the expressions. In this research, an algorithm has been developed, called "balanced assignment", intermediate between Hellerman's techniques. Balanced assignment is proved to be optimum for single expressions and for maximal sequences of branch free code, "basic

blocks". Prosser's 1959 definition of dominance is employed to generalize this algorithm to some branching programs. The notion of the "compound block" is further defined as the set of program statements dominated by some program statement. In loop free programs these compound blocks are proved to be nested like programming language blocks if and only if the program has a flowchart that is a series-parallel network with two-way branches and junctions. This property is named "acyclic smooth". The property is independent of the programming language used, but it is possible to specify acyclic smooth programming languages; i.e., languages that can express only acyclic smooth programs. The balanced assignment rule is generalized to acyclic smooth programs, but the optimality of the generalization is proved only within a subclass of the class of all permissible arrangements of the program, a class called the class of linear level assignments. Generalization of balanced assignment to looping programs appears possible once a suitable definition of cyclic smoothness has been formulated, a topic left to future research.

C. R. Foulk, O. C. Juelich

THE COMPUTER EXPERIMENTATION FACILITY

A computer experimentation facility is being assembled that will support research into areas such as computer graphics, programming languages and their translators, operating systems, and computer architecture. Currently available equipment includes an Applications Group, Inc., model AG-60 plasuro panel based graphics terminal, a touch panel for the AG-60, a sonic tablet, Hewlitt-Packard 2115A minicomputer, data communications equipment to allow long distance 2400 based rate communications between the AG-60/H-P 2115A and the DECIO of the CIS Department, a plotter, and more. The main idea is to provide an environment that will foster and aid research into some of the applied areas of computer science.

A. P. Lucido, C. Kearns (Sponsor: French Fellowship of the Engineering Graphics Department.)

COMPUTER SYSTEM FOR SECURE DATA PROCESSING

The work investigates four logical levels of computer protection structure: system protection, data communication and processing, protection of stored data and programs and user identification and authorization procedures. The complete secure computer system is designed which is general enough so that it enables all operations that present day computer systems offer. The protection mechanism is realized in hardware for better security and it is independent of any internal program, machine or data structure. The main problems in such a system (execution of programs, data set manipulations, hardware modifications and security procedures) are all given in detail. The system is evaluated in terms of hardware and operational overhead.

In addition, cryptographic techniques for computer use are investigated, classified and compared. The design of the generalized homophonic cipher system, its secrecy and efficiency are given too. Finally, necessary hardware

modifications are all designed and their operations in the system are all described.

M. T. Liu, S. Muftic, J. Rothstein

CONCURRENCY IN REAL-TIME INFORMATION SYSTEMS

A method of synchronizing processes of a real-time information system is developed. Two requirements of the system are that if a query Q is submitted at time T_i, all queries submitted before T_i are not affected by Q and all queries submitted after T_i are affected by Q. Many current systems satisfy these requirements by sophisticated scheduling algorithms and data base lockout mechanisms; our synchronization method guarantees these requirements without either of these mechanisms. We have also identified the system design requirements that will allow the maximum amount of multiprocessing. Assuming there is only one process that can change the value of a variable, it is desirable for variables to have common reference sets. It also has been shown that if, for each variable, the only process that references a variable directly follows the process that changes that variable, concurrent processing of all processes can occur.

H. S. Koch

CONTEXT PROTECTION IN DATA BASE SYSTEMS

The goal of this research is the development of a new type of protection mechanism known as the context protection mechanism for interactive data secure systems.

In all currently known protection mechanisms access control decisions are made a priori on the basis of access rules related to users, data and data base operations. In this research an attempt will be made to include protection mechanisms in which access control decisions depend upon additional factors such as past access history and the content of data units.

Plans for implementing the new type of mechanisms in an experimental data base system will be considered.

D. Cohen, D. K. Hsiao, M. T. Liu (Sponsor: Office of Naval Research, N00014-75-C-0573)

COST-EFFECTIVE ACCESS CONTROL IN DATA BASE SYSTEMS

In every data base system that has been proposed, an unavoidable consequence has been the introduction of system performance 'overhead' attributed to the logical access control mechanisms provided by the data base system. This overhead possesses the characteristics that the cost of providing adequate data security for each user of the system is directly proportional to the effectiveness and sophistication of

that security. A data secure system that can provide the user with his required level of security and still keep the cost of information processing and transfer at an acceptable level is termed "cost-effective". The objective of this research is to conduct an examination of the requirements for providing cost-effective data security in data base systems.

J. M. Hennings, D. K. Hsiao (Sponsor: Office of Naval Research. N00014-75-C-0573)

DATA BASE ACCESS CONTROL IN THE PRESENCE OF CONTEXT DEPENDENT PROTECTION REQUIREMENTS

Although the capability of the access control mechanisms to regulate field, record, and file security has been recognized as indispensible in advance data base systems, there is the need of more subtle protection which we shall call context dependent protection.

Context dependent protection of data enables us to change the accessibility of a data unit (field, record, or file) dynamically when some other data units have been accessed. A context dependent protection requirement specifies how and under what condition such a change is to be made for a particular data unit.

In this research, two basic types of context dependent protection requirements will be identified and studied. The first type of requirement, called an access decreasing requirement, will reduce the accessibility of a data unit when the condition specified by it is satisfied. The other type of requirement, called an access increasing requirement, will make a data unit more accessible when the condition specified is satisfied. Our research shows that both types of requirements can be studied by means of certain built-in relationships among the data units involved. Therefore, the first goal of this research is to investigate these relationships and their enforcement. Since the introduction of context dependent protection will definitely have some effects on the access control problem as a whole, the second goal is to study theoretically the overall impact of context dependent protection on protecting data in data base systems. A graph-theoretic approach will be used since directed graphs are good candidates for representing simple relations.

C. J. Nee, D. K. Hsiao, D. S. Kerr

DATA GENERATION FOR COMPUTER GRAPHICS

In this project, which is performed in conjunction with the Computer Graphics Research Group on campus (Prof. Charles Csuri, Director), the problem of simplifying what is currently a chore and a drudgery, namely, the generation of three-dimensional data specifically for use in Computer Animation, is being investigated. Some of the problem constraints are: the use of a minicomputer (PDP 11/45) and the requirements of user interaction and real-time.

A scenario we are working towards can be described as follows: The user sketches in free hand (using, say, the sonic pen in its 2-D mode) two or three orthographic projections of the object he has in mind. The system constructs the 3-D version of the object, which will necessarily be only approximately the object desired. This, however, involves not only the interpolation of the object in three space from two-dimensional information, but also the preprocessing required in cleaning up the inaccuracies and inconsistencies associated with hand-drawn pictures. The user might also, for classes of polyhedral objects, use the system's projected capability to obtain plane edges from free-form lines drawn in orthographic projections. Notice that these capabilities involve formation of hypotheses by the machine and drawing upon stored knowledge about different kinds of objects. Once the approximate 3-D version is created, then the user might sculpt the finished product by calling upon a variety of 3-D warping and sculpting routines.

B. Chandrasekaran, R. Parent (Sponsor: National Science Foundation. Grant DCR-74-00768 A01)

A DATA SECURE COMPUTER ARCHITECTURE

A new computer architecture to support a very large, highly secure attribute-oriented data base management system (DBMS) has been developed. A secure DBMS must be designed from the outset with security in mind. The system's security mechanisms must be an integral part of its architecture. The complex software currently necessary for a secure DBMS is a primary cause of unacceptable performance, and reliability in many systems. The architecture developed here overcomes this by providing unconventional hardware that inherently simplifies the algorithms of a secure DBMS. The proposed architecture contains four major components: the directory memory, the intersector, the mass storage and a pair of interface processors. The directory memory maintains attributes of the data base, the intersector executes set manipulation operations, the mass storage contains the data base, and the interface processors control the other components. The directory memory, intersector and mass storage are implemented with segment associative memories. This new kind of memory does not ascribe to the extremes of easy-to-use "location-addressed" memories. Segment associative memories of many different sizes and speeds--including ones of sufficient capacity to contain a very large data base--are well within the capabilities of near-term technology. By using segment associative memories it is possible to build an inherently more reliable and better performing DBMS that can support a wide range of contemporary data management activities.

R. Baum, D. K. Hsiao (Sponsor: Office of Naval Research. N00014-75-C-0573.)

DESIGN AND IMPLEMENTATION OF A DATA SECURE SYSTEM

The goal of this research is to design and implement an experimental data base management system with multi-level access control capabilities. The design work has been completed. The implementation work is being done on the DEC-10 computer. Languages for the implementation include both MACRO-10 and PREST4, which is a form of structured FORTRAN that was created for this effort. The three levels of logical access control in the system, the files, records,

and record fields, are each partitioned into those which the user can access and those which the user cannot access. The system is based on some known research results such as the Generalized File Structure, the Parallel Access Algorithm, and the Security Atom concept. User queries to the system are in the form of Boolean combinations of keywords. The system also supports full update capabilities. When finished, it will be used as a test bed for other concepts in data security.

N. Kaffen, L. Breene, T. Rodeheffer, D. Schmaltz, D. K. Hsiao (Sponsor: Office of Naval Research, N00014-75-C-0573)

DESIGN OF A DATA BASE MANAGEMENT SYSTEM FOR CONCURRENT PROCESSING

We have described how a DBMS can be designed to be structured to allow concurrent processing of that system as it is processing various queries. That is, we want to resolve queries by multiprocessing the components of the DBMS. It is assumed that the DBMS will be operating in a real-time environment, which puts further constraints on the system. Normally, in a concurrent processing environment, variables are shared between processes and the critical sections are monitored to solve the mutual exclusion problem. However, this solution is not sufficient for a DBMS that operates in a real-time environment. It has been shown why this solution is not sufficient. A model has been developed that will be able to satisfy real-time constraints.

H. S. Koch

DETERMINING THE STABILITY OF DISCRETE SIMULATION MODELS

When modellers write discrete simulation programs using SIMSCRIPT or GPSS they usually do not evaluate the stability of the model directly. This is a very difficult analytic task and the best they can do is to obtain an empirical estimate of the stability. They execute the simulation program many times, varying slightly the parameters of the model and changing the seed of the pseudo-random number generator, determining how much the output is affected. This can be a very expensive process, and it is not as informative as many would like.

We are attempting to develop a way to analytically determine the stability properties of such models. Consider a differential equation (d.e.) and a numerical integration scheme for obtaining the approximate solution of its initial value problem. There are substantial theories dealing with the stability properties of the d.e. (Lyapunov's for example) and the numerical stability, roundoff and turncation problems of the integration method. We use these theories as follows: let a simulation modelling program be written in a normal form which is so designed that one may conveniently view it as a numerical integration scheme for some d.e., and from that normal form program extract the d.e. in question. Analyze the stability behavior of the d.e. and the integration scheme and apply the results to the simulation model.

To facilitate the study we define a simple simulation language, SSL, which has the two important properties that any discrete simulation model can

be expressed in the language, and that any program in the language can be easily transformed into the normal form.

D. J. Moore, P. Miller

A FAMILY OF PROTECTION LANGUAGES FOR SPECIFYING DATA BASE PROTECTION REQUIREMENTS

In recent years, the issue of security and privacy in large integrated data bases has attracted much attention. Along with the installation of access control systems to protect stored information, there is the need to develop a protection language for the users to communicate to the system their security requirements. Due to increasing demand for more effective communication and sophisticated security requirements, the scope of protection languages has evolved from simple commands that result in a limited range of features to a language with sophisticated constructs, providing highly discriminating security control.

This research develops a set of protection language component parts (PLCP's) that embodies access control features at many different levels of sophistication. The PLCP approach "compartmentalizes" the language constructs, allowing concise language descriptions with emphasis on salient protection features. We show that the various features in the PLCP's can be selected as the basis of a family of protection languages for use in different system environments. An interactive protection language for use in an attribute-based data base is implemented. Experience with the system reveals extensions for further study, especially in the area of enforcement dynamics.

D. H. Wong, D. K. Hsiao

LANGUAGES FOR SPECIFYING PROTECTION REQUIREMENTS IN DATA BASE SYSTEMS--A SEMANTIC MODEL

This work develops a model to provide a semantic base for constructs of protection languages at many levels of sophistication, accommodating a wide range of protection policies.

The basic sets of the model are presented and subsets of states are defined by Boolean expressions (conditions) of arithmetic relations on the resource values. History keeping is used to allow a sensitivity to the context of access history so that access decisions can be based on the occurrence or non-occurrence of previous operations. Provision is made for the specification of auxiliary program invocations to achieve additional protection measures before, during, and after the access decision making-process. The concept of extended resources allows a user to invoke procedures having greater access rights than does the user himself.

A family of protection languages are proposed as the key to expression of policies and rules of access by authorizers. Translation of protection languages

consists of lexical (syntactic) translation into entities of the model and semantic translation by the authorization and enforcement processes of the model. Protection language constructs fall into four categories: definitional, operational, advanced protection services, and semantic parameters. The definitional category is for specifying definitions of users, user groups, resource units, and conditions. The operational category says who can do what operations, on what resources, and under what conditions. The advanced protection services contain protection constructs such as history keeping, auxiliary program invocations, and extended resources. Semantic parameter specification provides the means to define protection policies at the highest levels of sophistication.

The work is now completed and has resulted in a Ph.D. dissertation.

H. Rex Hartson, David K. Hsiao (Sponsor: Office of Naval Research. N 00014-75-C-0573)

A MODEL FOR A SECURE DATA BASE AND A FRAMEWORK FOR THE APPLICATION OF CRYPTOGRAPHIC TRANSFORMATIONS IN A DATA BASE SYSTEM

In our research we introduce a top-down approach to data security. Starting with the general subject area we show why the security problems in data base systems should be discussed separately from those in operating systems. We have defined some basic concepts such as: protection specification, protection mechanism and protection binding. Using these concepts we analyzed the current models of data base system (mainly the CODASYL model) and pointed out some serious disadvantages related to security, in particular the problem of centralization. We then developed a levels structured model which incorporates in it many of the security models known today and we have shown that with this model decentralization can be achieved if desired. We developed some notation for the levels model and using this notation we were able to deal with the question of the application of cryptographic transformations in a data base. More research has still to be done on the problem of how the cryptographic transformations are going to be used. More research has to be done also on the security engineering problem, the evaluation of different protections mechanisms and in particular the evaluation of cryptographic transformations.

E. Gudes, H. S. Koch, J. Rothstein, F. A. Stahl

STRUCTURED DATA BASE GENERATION.

As part of data base management system functions, the generation of a variety of new data bases is a large effort. In order to minimize the amount of effort for each data base generation, we propose to structure the process into four basic steps.

The first step requires a special-purpose program that converts raw data into records in the form of attribute-value pairs. During the conversion from raw data to attribute-value pairs, a standard statistical package is used to gather information about the raw data for the data base administrator. Since the conversion must be tailored to the raw data format and is used only once

for each data base generation, a one-time program is needed. Subsequent steps can be standardized, for they depend only on well defined input formats. Once the data base administrator knows the required record organization and file structure of the new data base, he can proceed with the second step by building record templates and specifying the physical sizes of the attribute dictionary segment, the atom directory segment, and the storage cell. This step is done interactively in order to aid the data base administrator to review his work. Step three involves converting records from groups of attribute-value pairs to the system's physical organization. This process involves the use of record templates to validate attributes and values and to verify the hierachical organization of the records. The final step is to insert the records (in system's physical organization) into a cell of data base storage. This process includes updating attribute dictionaries and atom directories to reflect the presence of new records.

R. Knablein, D. K. Hsiao (Sponsor: Office of Naval Research. N00014-75-C-0573)

STRUCTURED FORTRAN

PREST4 is a program on the DECSystem-10 that takes an input file containing source code written in the PREST4 language, and generates an output file in standard DEC FORTRAN. As an option, it also generates a listing. Basically, the PREST4 language is an extension of FORTRAN with changes made to facilitate the writing of structured programs. Thus the name PREST4 - PREprocessor for STructured FOR(4)TRAN, is coined.

PREST4 recognizes four types of structures: an IF-THEN-ELSE structure for specifying conditional execution; a DO-END structure for grouping a series of statements into a unit and for specifying a repetition criteria; a READ structure for reading data, and for handling the end of file by specifying an assignment or a subroutine call; and a simple structure which consists of a single statement.

Besides program statements, the user may supply preprocessor control statements in the PREST4 source code. These control statements cause PREST4 to perform some specific action (e.g., eject a page in the listing) and do not cause themselves to be generated in the output file. Every control statement has a percent sign (%) in Column 1.

PREST4 can be told to TRACE a program, in which case it places a subroutine call to the tracing package (passing the line number as an argument) into the output file before examining each statement of the program. The tracing package records at run time the last 500 calls and can list the line numbers upon request. Thus in the event of an error the user can determine which statements have been recently executed.

T. Rodeheffer, D. K. Hsiao (Sponsor: Office of Naval Research. N00014-75-C-0573)

X. COMPUTER ARCHITECTURE AND NETWORKS

THE ARCHITECTURE OF A GRAMMAR-PROGRAMMABLE HIGH-LEVEL LANGUAGE MACHINE

This research is concerned with system design of a grammar-programmable high-level language machine. The term grammar-program is applied to the context of High-Level Language Machine and defines an intermediate level between the basic hardware/firmware functions of a computer system and its software language translators. Through grammar-programs, the syntax and semantics of various programming languages can be specified to the Grammar-Programmable Machine which uses these specifications to process directly the users' high-level language programs.

Taking advantage of its intermediate position between software compilation and hardware interpretation of high-level languages, the grammar-programmable machine emphasizes the strong points of interpreters to overcome and even enhance the best features of compilers, and vice-versa. For this reason, it enables the realization of the potential which neither a compiler nor an interpreter can achieve individually.

M. T. Liu, S. Fournier

THE DESIGN OF THE DISTRIBUTED LOOP COMPUTER NETWORK

This research is concerned with the design of the Distributed Loop Computer Network (DLCN) using a new transmission mechanism proposed in previous work. This new transmission mechanism is much more efficient and sophisticated than those in current use and has the following advantages: 1) concurrent transmission of variable-length messages is allowed, 2) nearly immediate loop access is guaranteed, regardless of load, and 3) automatic regulation of message traffic is provided, all accomplished in a completely distributed network.

The design requirements, properties and hardware implementation, using standard logic circuits, of the loop interface are investigated. Following that, several novel solutions to message types and formats, message acknowledgement, error detection and retransmission, lost message detection and removal, priority message insertion and lockout prevention are added to the interface design. In addition, software system requirements for the Distributed Loop Operating System (DLOS) are studied, including network protocol, interprocess communication, resource allocation and management, and error detection and recovery.

The main goal of DLCN is to develop an integrated hardware/software/communication system with distributed control that provides efficient, felexible, reliable and fail-soft service at low cost in the face of constantly changing

user demands.

M. T. Liu, C. C. Reames

A NOVEL MODEL FOR A MIXED VOICE/DATA TRANSMISSION SYSTEM FOR COMPUTER COMMUNICATION

This research is concerned with the development of a novel model for a mixed model for a mixed voice/data transmission system for computer communication which operates by embedding digital data signals into a set of voice channels during the silent periods of voice conversation. Several previous works are briefly reviewed. Mini/micro computer technology is used in the control of the digital signal embedding procedures. Design considerations such as interfacing to the existing telephone channels, voice detection, voice/data recognition and separation, routing strategies, modems, message protocols, data security, etc. are discussed in sufficient detail. It is expected that the proposed model will not only lower communication costs for computer networking but will also improve the efficiency of channel resource utilization.

M. T. Liu, J. T. Wang

OPTIMAL PROCESS ALLOCATION IN DISTRIBUTED HETEROGENEOUS COMPUTER NETWORKS

This research is concerned with optimal process allocation in distributed heterogeneous computer networks so as to minimize hardware/software redundancy. A mathematical model is proposed for analysis using integer and dynamic programming which allows different service rates for each process. A simulation model is to be implemented to verify the results. It is expected that the goal of minimizing redundancy can not only increase overall system efficiency, but can also self-tune the system according to the constantly changing request pattern.

M. T. Liu, T. T. Cheng

A PREDICTIVE RESPONSE TIME MONITOR FOR COMPUTER NETWORKS

A predictive response time module to assist users of a heterogeneous network of computers was designed. The network user is able to query this dynamically updated software module to obtain current information relating to the busyness of time sharing systems on the network and the amount of time required on these systems to run various classes of computing applications.

The implementation of the module is proposed as an extension of a Network Access Machine (NAM) developed by the Computer Networking Section at the National Bureau of Standards. The NAM is a minicomputer which assists the network user in obtaining network services. The response time module will accept from users the designation of time sharing systems on which they wish to do work and will return either a narrative description of the general response time characteristics

of those systems or comparative lists of response times for running short and long FORTRAN, COBOL, and BASIC jobs on those systems.

The feasibility of the design and implementation plan was verified by response time measurement experiments conducted on a DEC System-10. A measure of system busyness called "Percent CPU Idle Time" was found to be a satisfactorily accurate parameter on which to base response time predictions for the FORTRAN, COBOL and BASIC jobs mentioned above.

S. Mamrak (Sponsor: The National Bureau of Standards)

XI: JOINT PROGRAMS

THE STATUS OF WOMEN AND MINORITIES IN ACADEMIC COMPUTER SCIENCE

A survey of women and minority students and faculty in computer science during the years 1971 to 1975 was conducted. Analysis of the data indicated that effective affirmative action programs for recruitment into graduate degree programs are needed to enlarge the number of women and minorities qualified for later employment in computer science. Also, possible discrimination in employment of women and minority graduate students was revealed.

S. Mamrak, R. G. Mentanelli (Sponsor: The University of Illinois at Urbana-Champaign)

XII. COMPUTATION THEORY

THE COMPUTATIONAL COMPLEXITY OF ORACULAR ENUMERATION PROCEDURES

We study enumeration procedures in which an oracle is used which supplies a stream of information at no cost. The model developed is suitable for a study of the abstract complexity properties of the situation.

The questions asked concern the existence of oracular sets which are very helpful, and output sets which can be helped by no recursively enumerable, oracle. We investigate the influence of the order of the stream of oracular data on its usefullness.

We attempt to study questions analogous to those answered by Nancy Lynch in the case of recursive oracles used by oracular Turing Machines.

Using techniques developed by Paul Young, Albert Meyer and Pat Fischer we have shown that the effectiveness of an enumeration oracle can depend very heavily upon the order in which the oracular information is presented. We conjecture that there are interesting situations in which the oracular information is so redundantly coded that it is useful no matter what the order of presentation might be.

We discuss the applicability of this work to theorem proving.

D. J. Moore

THE COMPUTATIONAL COMPLEXITY OF RECURSIVELY ENUMERABLE TOTAL ORDERS

In the study of abstract complexity theory many disturbing pathological characteristic of recursive functions have been discovered. Here we investigate the hypothesis that sets which are graphs of total orders, and thus highly structured, may possess less severely pathological properties. We study in particular the computational complexity of the problem of recognizing membership in the graphs of recursively enumerable (r.e) total orders.

We show that, independently of the order type and underlying field, there are (1) r.e. total orders for which it is arbitrarily difficult to recognize infinitely many elements, and (2) r.e. total orders for which any membership recognition procedure can be sped up by an arbitrary amount on infinitely many arguments. Also, we show that whether or not there are r.e. total orders possessing these properties on all but finitely many arguments is strongly dependent on order type. Total orders with a least or greatest element possess the infinitely often, but not the almost everywhere complexity properties; however those total orders with no least or greatest element possess the almost

everywhere complexity properties. The proofs that the infinitely often complexity properties hold for order type ω make use of finite injury priority arguments.

D. J. Moore

PARALLEL RECOGNITION OF FORMAL LANGUAGES

We have generalized cellular automata and iterated switching networks to what we call bus automata (BA). These can switch programmably variable communication paths between automata (cells). This permits speed-up in computation by working on different parts of a problem in parallel at different locations.

Immediate formal languages (IML) are defined as those accepted by BA in a finite number of state-change intervals independent of length of input string. IML includes all regular languages and important classes of CF and CS languages also. Closure properties of IML languages were investigated.

Several relationships were established to previously investigated language systems including the "Chomsky Hierarchy", various parallel acceptance language families previously studied, and the polynomial time complexity families PTIME and NPTIME. A hierarchy of languages amenable to parallel acceptance emerges.

J. Rothstein, M. Moshell

THE STRUCTURE OF THE SET OF FEASIBLY COMPUTABLE FUNCTIONS

Functions whose values can be calculated in a time which is a polynomial function of the length of the input are considered feasibly computable. We investigate the structure of the set of feasibly computable functions in several ways.

We define a polynomial hierarchy analogous to the arithmetic hierarchy of recursion theory. This notion has been linked to the concept of oracular polynomial computations. We use both characterizations in looking for natural problems which lie at the various levels of the hierarchy.

We define a polynomial T-predicate, and analogies of Rice's Theorems, recursion theorems, creative and simple sets. Any successful developments of this nature can help in attempts to solve the problem of whether P=NP.

D. J. Moore, W. Leggett

TOWARD AN ARITHMETIC FOR CELLULAR AUTOMATA AND PARALLEL COMPUTATION

The base 2 number system suggested by groupoid patterns (which was reported on last year) has been further investigated and found to have properties at once fascinating, frustrating, and encouraging. It is frustrating in the sense that the simplest arithmetic operations can so far only be done in a complicated

manner, fascinating in the sense that many complicated combinatorial relationships can be dealt with simply, and encouraging because the operations examined to date can be done simultaneously, in large numbers, and without mutual interference, by cellular automata.

An earlier generalization of base 2 to an infinitude of integer representations using 0 and 1 also grew out of the groupoid research. They correspond, in the limit of large numbers, to multiplication by a fixed number between 1 and 2 when the number of digits is increased by one. It appears likely that all of these systems will have counterpart "parallel" systems in the same sense that ordinary base 2 has its counterpart in the system of the previous paragraph.

J. Rothstein

TURING UNIVERSALITY AND PARALLELISM OF GROUPOID STRING ALGORITHMS

The chief theorem to emerge during the past year describes a general speed-up of Turing machine calculations by an I-D cellular automaton. It builds on the groupoid "parent string" parallel computation earlier reported. We simulate a given Turing machine, augmented by the parallelism capability of the I-D CA. Its cells control switching paths running parallel to the CA; this device is called a bus automaton (BA). The theorem is that a bus automaton can perform exactly the same computation as a Turing machine in a number of state-change intervals exceeding the number of tape reversals of the original Turing machine by unity. Furthermore, if the distance between reversals is bounded for the class of computations that Turing machine does, then the entire computation can be done in one interval.

J. Rothstein

APPENDIX A

COMPUTER AND INFORMATION SCIENCE COURSE LISTING

BY NUMBER AND TITLE

100 Computers in Society

201 Elementary Digital Computer Programming

211 Computer Data Processing I

212 Computer Data Processing II

221 Programming and Algorithms I

222 Programming and Algorithms II

294 Group Studies

294/223 Introduction to Computer Systems

311 Introduction to File Design and Analysis

411 Design of On-Line Systems

422 Topics in Computing for Engineers

494 Group Studies (Discontinued 7/75)

505 Fundamental Concepts of Computer and Information Science

509 Survey of Computer & Information Science for High School Teachers

541 Survey of Numerical Methods

542 Introduction to Computing in the Humanities

543 Intermediate Digital Computer Programming

548 Digital Computer Programming for High School Teachers

549 Numerical Analysis for High School Teachers

550 Introduction to Information Storage and Retrieval

555 Survey of Programming Languages

594 Group Studies (Discontinued 7/75)

610 Principles of Man-Machine Interaction

640 Numerical Analysis

641 Computer Systems Programming I

642 Numerical Linear Algebra

643 Linear Optimization Techniques in Information Processing

644 Advanced Computer Programming

652 Modeling of Information Systems

675 Digital Computer Organization

676 Minicomputer Evaluation and Selection

677 Computer Networks

680 Data Structures

693 Individual Studies

694 Group Studies

705 Mathematical Foundation of Computer and Information Science

706 Information Theory in Behavioral Science

712 Man-Machine Interface

720 Introduction to Linguistic Analysis

726 Theory of Finite Automata

727 Turing Machines and Computability

728 Topics in Theory of Computing

730 Basic Concepts in Artificial Intelligence

735 Statistical Methods in Pattern Recognition

740 Computer Systems Programming II

741 Comparative Operating Systems

745 Numerical Solution of Ordinary Differential Equations

746 Advanced Numerical Analysis

750 Modern Methods of Information Storage & Retrieval

751 Fundamentals of Document-Handling Information Systems

752 Techniques for Simulation of Information Systems

753 Theory of Indexing

754 Language Processing for Information Storage and Retrieval

755 Programming Languages

756 Compiler Design & Implementation

760 Selected Topics in the Mathematics of Information Handling

765 Theory of Management Information Systems

775 Advanced Computer Organization

780 File Structures

781 Aspects of Computer Graphics Systems

788 Intermediate Studies in Computer & Information Science

788.01 Theory of Information

788.02 Information Storage & Retrieval

788.03 Theory of Automata

788.04 Artificial Intelligence

788.05 Pattern Recognition

788.06 Computer Systems Programming
788.06A OS-MUT

788.07 Programming Languages

788.08 Computer Organization

788.09 Numerical Analysis

788.10 Man-Machine Interaction

788.11 Fornal Languages

788.12 Management Information Systems

788.13 Biological Information Processing

788.14 Socio-Psychological Aspects of Information Processing

793 Individual Studies

794 Group Studies (Discontinued 7/75)

797 Interdepartmental Seminar

805 Information Theory in Physical Science

806 Cellular Automata & Models of Complex Systems

812 Computer & Information Science Research Methods

820 Computational Linguistics

835 Special Topics in Pattern Recognition

840 Operating System Implementation

845 Numerical Solution of Partial Differential Equations

850 Theory of Information Retrieval I

851 Theory of Information Retrieval II

852 Design and Analysis of Information Systems Simulations

855 Formal Theory of Programming Languages

865 Seminar on Socio-Psychological Aspects of the Information Sciences

880 Advanced Theory of Computability

888 Advanced Studies in Computer & Information Science

888.01 Theory of Information

888.02 Information Storage & Retrieval

888.03 Theory of Automata

888.04 Artificial Intelligence

888.05 Pattern Recognition

888.06 Computer Systems Programming

888.07 Programming Languages

888.08 Computer Organization

888.09 Numerical Analysis

888.10 Man-Machine Interaction

888.11 Formal Languages

888.12 Management Information Systems

888.13 Biological Information Processing

888.14 Socio-Psychological Aspects of Information Processing

889 Advanced Seminar in Computer & Information Science

899 Interdepartmental Seminar

999 Research

APPENDIX B

COMPUTER AND INFORMATION SCIENCE FACULTY

Marshall C. Yovits, Ph.D., (Yale University).
Professor and Chairman of Department of Computer and Information Science and Professor of Electrical Engineering. Director, C.I.S. Research Center. Information systems, theory of the flow of information, self-organizing systems.

Ranko Bojanic, Ph.D., (Mathematical Institute of the Serbian Academy of Science).
Professor of Computer and Information Science and Professor of Mathematics. Mathematical analysis, theory of approximation.

Charles A. Csuri, M.A., (The Ohio State University)
Professor of Computer and Information Science and Professor of Art. Advancement of computer graphics technology in software and hardware (language algorithms, data generation or inputs), use of computer technology in telecommunications.

Richard I. Hang, M.S., (The Ohio State University).
Professor of Computer and Information Science and Professor of Engineering Graphics. Computer graphics, engineering application of computers.

Clyde H. Kearns, M.S., (The Ohio State University).
Professor of Computer and Information Science and Professor of Engineering Graphics. Computer graphics, engineering application of computers.

Robert B. McGhee, Ph.D., (University of Southern California).
Professor of Computer and Information Science and Professor of Electrical Engineering. Control theory, switching theory, logical design.

Harold B. Pepinsky, Ph.D., (University of Minnesota).
Professor of Computer and Information Science and Professor of Psychology. Clinical and socio-cultural psychology.

Roy F. Reeves, Ph.D., (Iowa State University).
Professor of Computer and Information Science and Professor of Mathematics. Director, Instruction and Research Computer Center. Numerical analysis and programming.

Jerome Rothstein, A.M., (Columbia University).
Professor of Computer and Information Science and Professor of Biophysics. Informational problems in science, methodology, biocybernetics.

Charles Saltzer, Ph.D., (Brown University).
Professor of Computer and Information Science and Professor of Mathematics. Coding theory, numerical analysis, automata theory.

Hugh Atkinson, M.A., (University of Chicago).
Associate Professor of Library Administration and Director of Libraries. Library on-line automation, data processing for libraries, technical service processing models.

Kenneth Breeding, Ph.D., (University of Illinois).
Associate Professor of Computer and Information Science and Associate Professor of Electrical Engineering. Computer organization and switching theory.

H. William Buttelmann, Ph.D., (University of North Carolina).
Associate Professor of Computer and Information Science (Summer 1975).
Automata theory, computer architecture and programming languages.

Balakrishnan Chandrasekaran, Ph.D., (University of Pennsylvania).
Associate Professor of Computer and Information Science. Pattern recognition and artificial intelligence, learning automata theory, finite memory decision theory and game theory.

Ronald L. Ernst, Ph.D., (University of Wisconsin).
Associate Professor of Computer and Information Science and Associate Professor of Psychology. Human performance theory and engineering, complex information processing and systems evaluation.

Clinton R. Foulk, Ph.D., (University of Illinois).
Associate Professor of Computer and Information Science. Programming languages, systems programming, programming heuristics.

David K. Hsiao, Ph.D., (University of Pennsylvania).
Associate Professor of Computer and Information Science. Systems programming, information storage and retrieval systems, file systems, data base management systems, access control and privacy protection of data, data definition language and processor, system architectures.

Douglas S. Kerr, Ph.D., (Purdue University).
Associate Professor of Computer and Information Science. Numerical analysis and programming.

Ming-Tsan Liu, Ph.D., (University of Pennsylvania).
Associate Professor of Computer and Information Science. Computer organization, switching and automata theory, mathematical programming, computer architecture, pseudo-Boolean programming, threshold logic.

Anthony E. Petrarca, Ph.D., (University of New Hampshire).
Associate Professor of Computer and Information Science. Automatic indexing, chemical structural information processing, automated search systems, other aspects of information storage and retrieval.

James B. Randels, Ph.D., (The Ohio State University).
Associate Professor of Computer and Information Science and Assistant Director, Learning Resources Computer Center. Computer operating systems and utilities, telecommunications applications, subroutine libraries, programming languages.

James E. Rush, Ph.D., (University of Missouri).
Adjunct Associate Professor of Computer and Information Science. Indexing theory, automated language processing, organization of information, and parallel processing.

Cellanna I. Taylor, B.S.L.S., (Graduate School of Library Science, Case-Western Reserve University).
Senior Research Associate and Associate Professor of Library Administration. Information dissemination and utilization systems, information centers, library systems and management.

Lee J. White, Ph.D., (University of Michigan).
Associate Professor of Computer and Information Science and Associate Professor of Electrical Engineering. Mathematical programming, data structures, organization of information.

Ronald L. Wigington, Ph.D., (University of Kansas).
Adjunct Associate Professor of Computer and Information Science and Director of R. & D., Chemical Abstracts Service. Computer system design.

Thomas G. DeLutis, Ph.D., (Purdue University).
Assistant Professor of Computer and Information Science. Design and evaluation of information systems, systems programming.

Donald L. Kalmey, Ph.D., (The Ohio State University). Appointed Summer 1975.
Assistant Professor of Computer and Information Science. Numerical analysis, computer architecture and organization, programming.

Harvey S. Koch, Ph.D., (Pennsylvania State University).
Assistant Professor of Computer and Information Science. Data definition language, data base management, programming languages and compiler design.

Frederick S. Koehl, Ph.D., (The Ohio State University).
Adjunct Assistant Professor of Computer and Information Science and Math Analyst, Instruction and Research Computer Center. Sorting techniques, topological groups, compiler design.

Anthony P. Lucido, Ph.D., (Iowa State University).
Assistant Professor of Computer and Information Science. Computer architecture, compiler design, interactive computer graphics.

Sandra Mamrak, Ph.D., (University of Illinois). Appointed Autumn 1975.
Assistant Professor of Computer and Information Science. Performance evaluation, computer networks, systems programming.

Robert F. Mathis, Ph.D., (The Ohio State University).
Assistant Professor of Computer and Information Science and Assistant Dean and Secretary of the Graduate School. Programming languages, numerical analysis.

Daniel J. Moore, Ph.D., (University of Kansas).
Assistant Professor of Computer and Information Science. Complexity theory, recursion theory, artificial intelligence.

Lawrence L. Rose, Ph.D., (Pennsylvania State University). Appointed Autumn 1975.
Assistant Professor of Computer and Information Science. Information retrieval, simulation, data structures, programming languages.

Norman K. Sondheimer, Ph.D., (University of Wisconsin).
Assistant Professor of Computer and Information Science. Natural language processing, artificial intelligence, information storage and retrieval.

Frederick A. Stahl, Ph.D., (University of Illinois).
Assistant Professor of Computer and Information Science. Computational security, cryptography, information retrieval, computers in the humanities, and in the law, artificial intelligence.

Stuart H. Zweben., (Purdue University).
Assistant Professor of Computer and Information Science. Programming languages, compilers, data structures, operating systems.

Ernest Staveley, B.S., (U.S. Naval Postgraduate School).
Administrative Assistant and Assistant Director, C.I.S. Research Center.

APPENDIX C

COMPUTER AND INFORMATION SCIENCE SEMINAR SERIES

October 10, 1974 "Real-time Computer Animation," Charles A. Csuri, Professor of Art and Director, Computer Graphics Research Group, The Ohio State University.

October 17, 1974 "Banyan Trees and Live Research," Jerome Rothstein, Professor of Computer and Information Science, The Ohio State University.

October 24, 1974 "A Heuristic Algorithm for the Minimum Set Cover Problem Using Plausibility Ordered Search Neighborhoods," Michael E. Doherty, Ph.D. Candidate, Computer and Information Science, The Ohio State University.

October 31, 1974 "An Overview of Debugging Tools and Techniques," Robert F. Mathis, Assistant Professor of Computer and Information Science, The Ohio State University.

November 7, 1974 "Microprocessors and Microcomputers," Fred A. Hatfield, President, Computer Data Systems.

November 14, 1974 "What is Industrial Computer Science--One Man's View," Phillip S. Dauber, IBM Thomas J. Watson Research Center.

November 21, 1974 "A Model for Data Secure Systems," Edwin J. McCauley, Ph.D. Candidate, Computer and Information Science, The Ohio State University.

December 5, 1974 "Interpretive Structural Modeling," John Warfield, Battelle, Columbus, Ohio.

January 7, 1975 "Comparative Response Times of Time-Sharing Systems on the ARPA Network," Sandra Mamrak, University of Illinois.

January 9, 1975 "Where Did All the Computer Power Go?" Herbert R. J. Grosch, Computerworld.

January 16, 1975 "Emulation of Computer Networks by Microprogrammable Microcomputers," David Cohen, Ph.D. Candidate, Computer and Information Science, The Ohio State University.

January 23, 1975 "Protection Structures in Languages and Systems," Anita Jones, Assistant Professor, Computer Science, Carnegie-Mellon University.

January 28, 1975 "The Proper Role of 'Retention' in Programming Languages," Howard Ross Hale, Jr., University of Delaware, Newark, Delaware.

January 30, 1975 "Some Thoughts on the Care and Feeding of Research Sponsors," McIver W. Woody, Associate Director for Development, Research Foundation, The Ohio State University.

February 6, 1975 "Performance Measures of Derived Search Keys," Edward T. O'Neill, Assistant Dean, School of Information and Library Studies, State University of New York at Buffalo.

February 13, 1975 "Careers in Computing," Robert Porter, Jr., Corporate College Recruiting, IBM Corporation, and David Knox, Programming Manager, Office Products Equipment, IBM Corporation.

February 20, 1975 "Automatic Programming: Inference of Program from Memory Snapshots of Sample Calculations," Fred Petry, Ph.D. Candidate, Computer and Information Science, The Ohio State University.

February 27, 1975 "Organization of Stored-Program Controlled Switching Systems--An Overview," Santanu Das, Head, Maintenance and Diagnostics Research Group; Paul H. Henson Research Center, North Electric Company.

March 6, 1975 "Adaptive Systems Modeling," Daniel Howland, Professor, Management Science, The Ohio State University.

March 10, 1975 "Automatic Keywords Selection for Reference Retrieval Systems," Chang-Shu Hubert Yang, Ph.D. Candidate in Computer Science, Cornell University.

March 12, 1975 "SERAPS: A Self-Repairing Automatic Programming System," Michael J. Kessler, Instructor of Computer Science, State University of New York at Buffalo.

March 13, 1975 "Computer Scheduling Strategies and Their Effects," Walter Doherty, Manager of Computer Usage Measurement and Evaluation, IBM Corporation.

March 31, 1975 "Automatic File Management Heuristics," Lawrence L. Rose, State University of New York at the Binghamton University Center.

April 9, 1975 "The Design of Programs for Asynchronous Multiprocessors," Philip H. Mason, Carnegie-Mellon University.

April 10, 1975 "Teacher Control in Computer Assisted Instruction," Peter Calingaert, Professor of Computer Science, University of North Carolina.

April 14, 1975 "An Interactive Analysis System for Execution-Time Errors," Alan M. Davis, Ph.D. Department of Computer Science, University of Illinois at Urbana-Champaign.

April 16, 1975 "Graphically-Enhanced Data Base Management System Design," Wayne D. Dominick, Vogelback Computing Center and Department of Computer Science Northwestern University.

April 17, 1975 "Parallel Recognition of Formal Languages by Cellular Automata," J. Michael Moshell, Ph.D. Candidate, Computer and Information Science, The Ohio State University.

April 24, 1975 "Real-Time Process Control--The Programmer's Cloud Nine," James P. Shaffer, Senior Analyst Programmer, Industrial Nucleonics.

May 1, 1975 "Compilation of Sequential Programs for Parallel Execution," Otto C. Juelich, Ph.D. Candidate, Computer and Information Science, The Ohio State University.

May 8, 1975 "High Capacity Optical Storage Systems," Carl Verber, Research Leader, Solid State and Optical Sciences Section, Battelle Memorial Institute.

May 15, 1975 "Data Security and Privacy," Robert H. Courtney, Jr., Manager, Data Security and Privacy, IBM Corporation.

May 22, 1975 "The Ohio College Library Center Network," Larry L. Learn, Director, Computer Facilities Division, The Ohio College Library Center.

May 29, 1975 "Extended Syntax-Directed Translation of Programming Languages," Arthur B. Pyster, Ph.D. Candidate, Computer and Information Science, The Ohio State University.

APPENDIX D

RELATED ACTIVITIES OF THE STAFF OF

COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER

J. G. Abilock presented an invited paper entitled "A Semiotic Framework for Information Science Leading to the Development of a Quantitative Measure of Information" (Co-author: M. C. Yovits) at the 37th American Society for Information Science Annual Meeting, Atlanta, October 15. The paper appears in Information Utilities; Proceedings of the 37th ASIS Annual Meeting, Vol. 11.

J. A. Aitken presented a paper entitled "Performance Evaluation of Data Base Management Systems (Co-author: T. G. DeLutis) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

J. A. Aitken presented a paper entitled "An On-Line Interactive Data Base Management System" (Co-authors: J. S. Chandler and T. G. DeLutis) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

B. J. Brinkman presented a paper entitled "Use of Association Measures Based on SDI Search Profiles" (Co-authors: G. J. Lazorick and Anthony E. Petrarca) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

H. W. Buttelmann presented a seminar on "Applied Research in Automated Language Processing" at Chemical Abstracts Service on October 24, 1974.

J. S. Chandler presented a paper entitled "An On-Line Interactive Data Base Management System" (Co-authors: J. A. Aitken and T. G. DeLutis) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

J. S. Chandler presented a paper entitled "A Methodology for a Multiple Goal Approach to Computer Systems Design" (Co-author: T. G. DeLutis) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

B. Chandrasekaran presented a paper entitled "Distance Functions for Independent Measurements and Finite Sample Size" at the II International Joint Conference on Pattern Recognition, Lyngby, Denmark, August 14, 1974. Co-author and co-presenter was A. K. Jain, Department of Mathematics, Wayne State University. The paper appears in the proceedings of the conference.

E. I. Cohen presented a paper entitled "Goals and Directions for Research in Program Verification" at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

B. Dasarathy presented a paper entitled "Some Maxmin Location and Pattern Classifier Problems: Theory and Algorithms" (Co-author: L. J. White) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

T. G. DeLutis was the speaker for a session on "Simulation Techniques" at the Association for Computing Machinery East Central Regional Conference, Troy, Michigan on November 1-2, 1974.

T. G. DeLutis presented papers entitled "Simulation of Data Base Files: Definition Independence" and "Simulation Employing Multiple Computers" (Co-authors: J. A. Aitken and J. S. Chandler) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

C. R. Fouk was the chairman of a panel discussion on "Structured Programming" at the October 9, 1974, meeting of the Central Ohio Chapter of the Association for Computing Machinery in Columbus.

M. S. Friedman presented a paper entitled "Design and Implementation of a Method for Controlling Postmortem Diagnostic Output" at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

E. Gudes presented a paper entitled "Cryptography and Data Base Security" at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

D. K. Hsiao presented a paper entitled "Minicomputers as Communication Subnets on Computer Network" at the International Meeting on Mini-Computers and Data Communication, Liege, Belgium, on January 1, 1975. Co-authors are Roy Reeves and Thomas Wyrick.

D. K. Hsiao presented an invited talk entitled "Data Base Models--Attribute -Based and Relational" at the Computer Science Colloquium held at Harvard University on January 13, 1975.

D. K. Hsiao presented an invited talk entitled "Information Secure Systems" at the Computer Science Colloquium, Rutgers University on April 3, 1975.

D. K. Hsiao has accepted invitations from the Programme Committee of IFIP Congress 74 to serve as Chairman of the session on Applications of Interactive Computing and as a panelist in the session on Privacy and Computers. The Congress will be held on August 5-10, 1975, in Stockholm, Sweden.

D. K. Hsiao has been appointed a member of the Publications Committee of the Institute of Electrical and Electronics Engineers (IEEE) Computer Society and as a member of the Editorial Board of the Computer Magazine, a monthly publication of the same Society.

D. K. Hsiao accepted an invitation to serve on the Data Management Panel of the Committee on Computer Science and Engineering Research Study chaired by Professor Bruce Arden, Princeton University. The committee is preparing a definitive volume on computer science and engineering research of which data management is one of the ten areas of research specialization.

D. Isaacs presented a paper entitled "Identification and Measurement of Cost and Pricing Elements of Computing Resources" (Co-author T. C. DeLutis) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

G. Kar presented a paper entitled "Grammatical Inference Problem" (Co-author: L. J. White) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

D. S. Kerr was an invited participant in a Symposium/Workshop on "Planning for Action - The Privacy Mandate" co-sponsored by the Institute for Computer Sciences and Technology of the National Bureau of Standards and by the MITRE Corporation, held in Washington, D.C., April 2-4, 1975.

D. S. Kerr was elected Vice-Chairman of the Special Interest Group on Computer Science Education of the Association for Computing Machinery.

H. S. Koch presented a paper entitled "A Missing Component of Current Data Base Management Systems - Data Base Reorganization" at the 37th American Society for Information Science Annual Meeting, Atlanta, October 17, 1974. The paper appears in Information Utilities: Proceedings of the 37th ASIS Annual Meeting, Vol. 11.

H. S. Koch served as Secretary of the Subschema Task Group of the CODASYL DDLC April 74-April, 1975.

R. Krishnaswamy and Alan W. Biermann, Computer Science, Duke University, co-authored a paper entitled "A System for Program Synthesis from Examples" which Dr. Biermann presented at the Institute of Electrical and Electronics Engineering Systems, Man, and Cybernetics 1974 International Conference, Dallas, Texas, on October 2-4, 1974.

M. T. Liu presented a paper entitled "Emulation of Computer Networks by Micro-programmable Microcomputers" at the 7th Annual Workshop on Microprogramming held in Palo Alto, California on September 30-October 2, 1974. Co-author was David Cohen. The paper is published in the Proceedings of the Workshop, pp. 159-167.

M. T. Liu co-authored a paper "Variable-Length Message Transmission for Distributed Loop Computer Networks" presented by C. C. Reames at the 2nd Annual Symposium on Computer Architecture, Houston, Texas, January, 1975.

M. T. Liu presented "Minicomputers as Communications Subnets in Computer Networks" (Co-authors: D. K. Hsiao, R. F. Reeves, and T. P. Wyrick) at the International Meeting on Mini-Computers and Data Communication, Liege, Belgium, January, 1975.

A. P. Lucido was the Session Chairman for graphics languages at the First Annual Conference on Computer Graphics and Interactive Techniques.

A. P. Lucido was the Technical Program Chairman, and Session Organizer for the Second Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH'75)

A. P. Lucido is Editor of the Proceedings of SIGGRAPH'75. In addition, he is Editor of SIGGRAPHITI, the newsletter of SIGGRAPH, the Special Interest Group on Computer Graphics, and Interactive Techniques, of the Association for Computing Machinery.

D. A. Marik presented a paper entitled "Grammatical Inference Problem" (Co-author: L. J. White) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

P. F. Mathis presented an invited paper entitled "An Overview of Debugging Tools and Techniques" at the Association for Computing Machinery East Central Regional Conference, Troy, Michigan, November 1, 1974.

R. F. Mathis presented a paper entitled "Pre-Execution, Batch, Interactive, and Post-Mortem Debugging" at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

D. J. Moore presented a paper entitled "An Algorithm to Collate Long Natural Language Texts for Scholarly Use (Co-author: F. R. Horowitz) at the Computer Science Conference, Washington, D.C., February 20, 1975.

D. J. Moore presented a paper entitled "Simulation of Settlement/Subsistence Systems: A SIMSCRIPT Model for the Glenwood Locality" (Co-author: L. Zimmermann) at the Thirty-First Plains Anthropology Conference, November, 1974.

M. Moshell presented a paper entitled "Parallel Recognition of Formal Languages by Cellular Automata" (co-author: J. Rothstein) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

J. Rothstein was invited to speak at a Symposium on Parallel Processing in Artificial Intelligence at the Courant Institute of Mathematical Sciences (Department of Computer Science), New York, January 17, 1975. The symposium was co-sponsored by the Office of Naval Research. Professor Rothstein spoke on "Universality and Hierarchiality in Parallel Processing and Their Importance in Pattern Recognition and Artificial Intelligence."

J. Rothstein was elected secretary of the local Institute of Electrical & Electronics Engineers (IEEE) chapter of the Professional Group on Computers.

J. Rothstein and Carl Weiman, New York University, co-authored a paper entitled "Fast Algorithms for Generating, Translating, and Rotating Straight Line Paths on Grids for Computer Graphics" which Dr. Weiman presented at the Computer Science Conference, Washington, D.C., February 18, 1975.

J. Rothstein presented papers entitled "Groupoid Strings, Turing Universality, and Parallel Processing in Cellular Automata" and "Parallel Recognition of Formal Languages by Cellular Automata" (Co-author: M. Moshell) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

V. Santhanam presented a paper entitled "An Algorithm for Efficient Generation of Optimal Prefix Codes" (Co-author: L. J. White) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

J. Smith presented a paper entitled "Sequential Document Classification Techniques: Experiments and Experimental Results" (Co-author: L. J. White) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

S. N. Srihari presented a paper entitled "Statistical Classifier Design for Aircraft Identification" (Co-author: L. J. White) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

F. A. Stahl accepted an invitation from UNESCO, the United Nations Educational, Scientific and Cultural Organization, to advise and assist in the development of graduate studies and research in Computer Science at Universidad Simon Bolivar, Caracas, Venezuela, under the National System of Engineering Education for Industry Project, from September to December 1974.

W. S. Stalcup presented a paper entitled "Automatic Vocabulary Control in Printed Indexes: Evaluation Through Use of a Quantitative Measure of Concept Scattering" (Co-author: A. E. Petrarca) at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

L. J. White presented an invited paper entitled "An Efficient Algorithm for Maximum k - Matching in Weighted Graphs" at the 12th Annual Allerton Conference on Circuit and System Theory sponsored by the Department of Electrical Engineering and The Coordinated Science Laboratory, University of Illinois, Urbana, Illinois, October 2-4, 1974. The paper appears in the Proceedings of the Conference.

M. C. Yovits presented an invited paper entitled "A Semiotic Framework for Information Science Leading to the Development of a Quantitative Measure of Information" (Co-author: J. G. Abilock) at the 37th American Society for Information Science Annual Meeting, Atlanta, October 15, 1974. The paper appears in Information Utilities; Proceedings of the 37th ASIS Annual Meeting, Vol. 11. Dr. Yovits also chaired a session at the ASIS Meeting.

M. C. Yovits was Chairman of a session on Education at the Computer Science Conference in Washington, D.C., February 18-20, 1975.

M. C. Yovits presented a paper entitled "Graduate Education in Computer Science and its Relationship to Industry" at the National Computer Conference, Anaheim, California, May 20, 1975.

M. C. Yovits was a panelist on "Information Transfer Today and Tomorrow" at the American Society for Engineering Education National Conference, at Rensselaer Polytechnic Institute, June 17-20, 1975.

M. C. Yovits is co-editor of Advances in Computers, Volume 13, Academic Press, 1975 with M. Rubinoff.

M. C. Yovits is a member of the American Committee for the Second World Conference in Education (WCCE '75), sponsored by the International Federation for Information Processing, September 1-5, 1975, Marseille, France.

M. C. Yovits is a member of the Accreditation Committee of the Association for Computing Machinery beginning March 1975.

M. C. Yovits is a member of the Curriculum Committee on Computer Science of the Association for Computing Machinery, beginning May 1975.

S. H. Zweben presented an invited paper entitled "A Recent Approach to the Study of Algorithms" at the Association for Computing Machinery Annual Conference, San Diego, November 11, 1974. The paper appears in the Proceedings of the ACM Annual Conference, San Diego, 1974, p. 747-748. Dr. Zweben was also a member of a panel on "Software Physics" at the Conference.

S. H. Zweben will be a reviewer for the World Conference on Computers in Education, Marseille, France, to be held in September, 1975.

S. H. Zweben will be a reviewer at the ACM 75 National Conference, Minneapolis, Minn., to be held in October, 1975.

S. H. Zweben was awarded a $4,000 University Research Grant from the Graudate School for his research on "Structure of Computer and Natural Language Algorithms."

APPENDIX E

PUBLICATIONS OF THE COMPUTER AND INFORMATION SCIENCE RESEARCH CENTER STAFF[1]

ABILOCK, J. G.: YOVITS, M. C. A semiotic framework for information science leading to the development of a quantitative measure of information. In: Information Utilities; Proceedings of the 37th ASIS Annual Meeting, Vol. 11, Atlanta, October 15, 1974.

ATKINSON, H. C. Extension of new services and the role of technology. Library Trends, October, 1974.

ATKINSON, H. C. Priorities for new librarian of congress. Library Journal, July, 1974.

ATKINSON, H. C. The Ohio State University mechanized information center. In: Proceedings of EDUCOM 1973 Fall Conference.

BAMMEL, S. E.; ROTHSTEIN, J. The number of 9 x 9 Latin squares. Discrete Mathematics, Vol. 11, 1975, pp. 93-95.

BUTTELMANN, H. W. Semantic directed translation of context free languages. American Journal of Computational Linguistics, microfiche no. 7, November, 1974.

CHANDRASEKARAN, B.; and YOVITS, M. C. Artificial intelligence. In: Encyclopedia of Computer Science and Technology, Vol. 2. Marcel Dekker, Inc., New York, 1975.

CHANDRASEKARAN, B. Artificial intelligence - The past decade. In: Advances in Computers, Vol. 13, 1975, pp. 170-225.

CHANDRASEKARAN, B.; JAIN, A. A. Distance functions for independent measurements and finite sample size. In: Proceedings of the II International Joint Conference on Pattern Recognition, Lyngby, Denmark, August 14, 1974.

CHANDRASEKARAN, B.; LAM, C. C. A finite memory deterministic algorithm for the symmetric hypothesis testing problem. In: Transactions on Information Theory, published by the Institute of Electrical and Electronics Engineers, Inc., January, 1975.

HSIAO, D. K.; BAUM, R. A Semantic model for protection mechanisms in data base systems. In: The Proceedings of the 8th Hawaii International Conference on System Science, Honolulu, January 6, 1975.

KOCH, H. S. A missing component of current data base management systems - data base reorganization. In: Information Utilities; Proceedings of the 37th ASIS Annual Meeting, Vol. 11, Atlanta, October 15, 1974.

---

[1] See Appendix F for publications issued as part of the Computer and Information Science Research Center technical report series.

LIU, M. T.; COHEN, D. Emulation of computer networks by microprogrammable micro-computers. In: Proceedings of the ACM 7th Annual Workshop on Microprogramming, Palo Alto, California, September-October, 1974. pp. 159-167.

LIU, M. T.; REAMES, C. C. Variable-length message transmission for distributed loop computer networks. In: The Conference Proceedings of the Second Annual Symposium on Computer Architecture, Huston, Texas, January, 1975, pp. 7-12.

MOORE, D. J.; TUGGLE, D. Computer solution of verbal analogy problems. Computer Studies in the Humanities and Verbal Behavior, Vol. VI, No. 3, October, 1973.

PETRY, F. E.; BAUM, R.I.; BIERMANN, A. Speeding up the synthesis of programs from traces. IEEE Transactions on Computers, Vol. C-24, No. 2, February, 1975, pp. 122-136.

STAHL, F. Cryptography in the age of automation. In: The Proceedings of the IEEE International Symposium on Information Theory, October, 1974.

STAHL, F. Cryptography and computation. In: Encyclopedia of computer science. Edited by A. Ralston. Auerback, 1974.

WHITE, L. J.; KSIENSKU, A. A. Aircraft identification using a vilinear surface representation of radar data. Pattern Recognition, Vol. 6, 1974, pp. 35-45.

WHITE, L. J.; GILLENSON, M. L. An efficient algorithm for minimum k-covers in weighted graphs. Mathematical Programming, Vol. 8, 1975, pp. 20-42.

WHITE, L. J.; KSIENSKI, A. A. and REPJAR, A. G. Object Identification from multi-frequency radar returns. The Radio and Electronic Engineer, Vol. 45, No. 4, April, 1975, pp. 161-167.

WHITTEMORE, B. J.; YOVITS, M. C. A generalized concept for the analysis of information. Information Science: Search for Identity; Proceedings of the 1972 NATO Advanced Study Institute in Information Science, A. Debons, ed., Marcel Dekker, Inc., New York, 1974.

WHITTEMORE, B. J.; YOVITS, M. C. The Quantification and analysis of information used in decision processes. Information Sciences, Vol. 7, April, 1974, pp. 171-184.

YOVITS, M. C.; RUBINOFF, M., EDS. Advances in Computers, Vol. 13. Academic Press, 1975.

YOVITS, M. C. Graduate education in computer science and its relationship to industry. In: AFIPS Conference Proceedings National Computer Conference, Vol. 44, Anaheim, California, 1975.

YOVITS, M. C. A Theoretical framework for the development of information science. In: Information Science, Its Scope, Objects of Research and Problems, edited by A. I. Mikhailov. Publication of International Federation of Documentation Committee on Research on the Theoretical Basis of Information, Moscow, 1975.

PAPERS ACCEPTED FOR PUBLICATION

BUTTELMANN, H. W. On the syntactic structures of unrestricted grammars. I. Generative grammars and phrase structure grammars. Information and Control, Vol. 28, July, 1975. 52 pp.

BUTTELMANN, H. W. On the syntactic structures of unrestricted grammars. II. Automata. Information and Control, Vol. 28, July, 1975. 21 pp.

CHANDRASEKARAN, B.; JAIN, A. K. Independence, measurement complexity and classification performance. Transactions on Systems, Man & Cybernetics.

LIU, M. T.; REAMES, C. C. The design of the distributed loop computer network. In: Proceedings of the 1975 International Computer Symposiu,, Taipei, Taiwan, Republic of China, August, 1975.

LIU, M. T.; WANG, J. T. A novel model for a mixed voice/data transmission system for computer communication. In: Proceedings of the 1975 International Computer Symposium, Taipei, Taiwan, Republic of China, August, 1975.

MUFTIC, S. Social aspects of computer networks. Management Information, Journal of IFIP. November, 1975.

ROTHSTEIN, J.; WEIMAN, C. Parallel and sequential specification of a context sensitive language for straight lines on grids. Computer Graphics and Image Processing, 1st issue. 1976.

PAPERS SUBMITTED FOR PUBLICATION

KERR, D. S.; KALMEY, D. L. The bachelors and masters computer science graduate. Communications of the ACM.

KERR, D. S. Recent results on the attribute based data model--a tutorial. International Conference on Very Large Data Bases, Boston, Mass., September, 1975

LIU, M. T.; REAMES, C. C. Simulation of the distributed loop computer network. The 3rd Annual Symposium on Computer Architecture, Clearwater, Florida, January, 1976.

LIU, M. T.; FOURNIER, S. System design of a grammer-programmable high-level language machine. The 3rd Annual Symposium on Computer Architecture, Clearwater, Florida, January, 1976.

MOORE, D. The computational complexity of recursively enumerable total orders. Theoretical Computer Science.

MUFTIC, S. Social aspects of computer newtorks. Management Informatics.

MUFTIC, S. Secrecy systems for numerical data. Journal of the Association for Computing Machinery.

APPENDIX F

TECHNICAL REPORT SERIES

1968

YOVITS, M. C.; ERNST, R. L. Generalized information systems: Some consequences for information transfer. October, 1968. 47p. (OSU-CISRC-TR-68-1) (PB-180 929)

FILLMORE, C. J.; LEHISTE, I. Working papers in linguistics no. 2. November, 1968. 128p. (OSU-CISRC-TR-68-3) (PB-182 596)

FRIED, J. B.; LANDRY, B. C.; LISTON, JR., D. M.; PRICE B. P.; VAN BUSKIRK, R. C.; WASCHSBERGER, D. M. Index simulation feasibility and automatic document classification. October, 1968. 21p. (OSU-CISRC-TR-68-4) (PB-182 597)

ROTHSTEIN, J. Thermodynamics & information: Before, in and beyond quantum mechanics. December, 1968. 21p. (OSU-CISRC-TR-68-5) (PB-183 738)

FINLEY, JR., M. R. The development of a basic language for artificial intelligence. January, 1969. 24p. (OSU-CISRC-TR-68-6) (PB-182 305)

1969

COLOMBO, D. S.; RUSH, J. E. Use of word fragments in computer-based retrieval systems. February, 1969. 7+[9]p. (OSU-CISRC-TR-69-1) (PB-184 104)

WHITE, L. J. Minimum covers of fixed cardinality in weighted graphs. March, 1969. 14p. (OSU-CISRC-TR-69-2) (PB-183 737)

JACKSON, D. M. The construction of retrieval environments and pseudo-classifications based on external relevance. April, 1969. 74p. (OSU-CISRC-TR-69-3) (PB-184 462)

ELLIOT, D. E.; HUANG, S.; LANGENDOEN, D. T.; LEE, P. G.; LEHISTE, I. Working papers in linguistics no. 3. June, 1969. 181p. (OSU-CISRC-TR-69-4) (PB-185 855) (ED-060 689)

BRIGGS, G. E. Reaction time and uncertainty in human information processing. March, 1969. 36p. (OSU-CISRC-TR-69-5) (PB-184 135)

WHITE, L. J.; RUSH, J. E. Linear lists for spiro graphs. June, 1969. 69p. (OSU-CISRC-TR-69-6) (PB-194 402)

PETRARCA, A. E.; LAY, W. M. The double KWIC coordinate index. A new approach for preparation of high-quality printed indexes by automatic indexing techniques. April, 1969. 12 +[17]p. (OSU-CISRC-TR-69-7)*

*Journal of Chemical Documentation, 9, 256 (1969)

YOVITS, M. C. Information science: Toward the development of a true scientific discipline. June, 1969. 27p. (OSU-CISRC-TR-69-8) (PB-187 983)

PETRARCA, A. E.; LAY, W. M. The double KWIC coordinate index. II. Use of an automatically generated authority list to eliminate scattering caused by some singular and plural main index terms. August, 1969. 13p. (OSU-CISRC-TR-69-9)*

MCCULLOUGH, J. L. The acquisition of information across cultures: I. Persuasive role play, counterargument and attitude change. August, 1969. 18p. (OSU-CISRC-TR-69-10) (PB-197 568)

ERNST, R. L.; YOVITS, M. C. Information science as an aid to decision making. September, 1969. 22p. (OSU-CISRC-TR-69-13) (PB-189 666) (ED-054 782)

LANDRY, B. C. An indexing and re-indexing simulation model. June, 1969. 50p. (OSU-CISRC-TR-69-14) (PB-198 115)

SALVADOR, R. Automatic abstracting and indexing. June, 1969. 93p. (OSU-CISRC-TR-69-15)

STEVENS, D. W. A computer program for the reduction of flow tables. June, 1969. 97p. (OSU-CISRC-TR-69-16) (PB-189 679)

COLOMBO, D. S. Automatic retrieval systems and associated retrieval languages. 1969. 69p. (OSU-CISRC-TR-69-17) (PB-198 116)

SCHLESSINGER, J. D.; WHITE, L. J. Optimum prefix encoding. August, 1969. 85p. (OSU-CISRC-TR-69-18) (PB-198 117)

DAY, R.; WHITE, L. J. Hebbian neural simulation: Computer program documentation. (OSU-CISRC-TR-69-19) (PB-204 003)

REEKER, L. H. Extended finite state representation of infinite machines. September, 1969. 36p. (OSU-CISRC-TR-69-20) (PB-187 949)

WILLIAMS, N. T.; ERNST, R. L. A computer simulation of human short-term memory. 1969. 62p. (OSU-CISRC-TR-69-22) (PB-197 874)

BEZDEK, R. R. The acquisition of information across cultures: II. Social science research in a different culture. III. Cross-cohort activity and attitude change, by J. L. McCullough. January, 1970. 36p. (OSU-CISRC-TR-69-23) (PB-197 876)

1970

UNKLESBAY, M. K. A one step version of Younger's algorithm for bounded context grammars. 1970. 41p. (OSU-CISRC-TR-70-1) (PB-197 603)

*Proceedings of the American Society for Information Science, Vol. 6, 1969, 277-282.

LI, Y. Information structure and optimal policy. September, 1970. 18p. (OSU-CISRC-TR-70-2) (PB-197 605)

DILLON, S. R. Some procedures for finding substitution property partitions, substitution property covers, and cover pairs for finite state sequential machines. 1970. 79p. (OSU-CISRC-TR-70-3) (PB-197 643)

MATHIS, B. A.; WHITE, L. J.; JACKSON, D. M. Stability analysis of term similarities for information classification theory. July, 1970. 79p. (OSU-CISRC-TR-70-4) (PB-195 376)

MCGHEE, R. B.; DILLON, S. R. A paull-unger procedure for substitution property partitions. April, 1970. 16p. (OSU-CISRC-TR-70-5) (PB-192 120)

DAY, R. G.; WHITE, L. J. Study of a random search method for function minimization. March, 1970. 77p. (OSU-CISRC-TR-70-6) (PB-194 404)

PHARES, R.; WHITE, L. J. Identification of circuits in chemical structures. June, 1970. 73p. (OSU-CISRC-TR-70-7) (PB-194 396)

HARALSON, K. M.; WHITE, L. J. Optimal prefix codes for ensembles of N equiprobable messages using a binary alphabet. May, 1970. 105p. (OSU-CISRC-TR-70-8) (PB-197 642)

PETRARCA, A. E.; LAITINEN, S. V.; LAY, W. M. Use of the double KWIC coordinate indexing technique for chemical line notations. 1970. 14+[17]p. (OSU-CISRC-TR-70-9) (PB-198 269)

LAY, W. M.; PETRARCA, A. E. Modified double KWIC coordinate index. Refinements in main term and subordinate term selection. 1970. 11+[11]p. (OSU-CISRC-TR-70-10) (PB-197 567) (ED-054 815)

LYONS, J. J. The speed-accuracy trade-off in processing different classes of material. 1970. 38p. (OSU-CISRC-TR-70-11) (PB-198 114)

FILLMORE, C. J.; LEHISTE, I.; MELTZER, D.; TATHAM, M. A.; THOMPSON, S. A. Working papers in linguistics no. 6. September, 1970. 132p. (OSU-CISRC-TR-70-12) (PB-194 829)

ROTHSTEIN, J. Information generalization of entropy in physics. February, 1970. 22p. (OSU-CISRC-TR-70-24) (PB-192 128)

JACKSON, D. M. Basis for an improvability measure for retrieval performance February, 1970. 31p. (OSU-CISRC-TR-70-25) (PB-197 812)

DRACHMAN, B.; EDWARDS, M. L.; FILLMORE, C. J.; LEE, G.; LEE, P.; LEHISTE, I.; ZWICKY, A. M. Working papers in linguistics no. 4. May, 1970. 164p. (OSU-CISRC-TR-70-26) (PB-192 163)

1971

ROTHSTEIN, J. Patterns and algorithms. January, 1971. 8p. (OSU-CISRC-TR-71-1) (PB-197 604)

GROSU, A.; LEE, G. Working papers in linguistics no. 7. February, 1971. [243]p. (OSU-CISRC-TR-71-2) (PB-198 278) (ED-060 688)

CHANDRASEKARAN, B.; KANAL, L. On linguistic, statistical, and mixed models for pattern recognition. March, 1971. 33,A5p. (OSU-CISRC-TR-71-3) (PB-198 279)

WHITTEMORE, B. An example of the application of generalized information systems concepts to the quantification of information in a decision system: The examination of quantified information flow in an industrial control problem. May, 1971. 51p. (OSU-CISRC-TR-71-4) (PB-202 621)

OSTROM, T. M.; STEELE, C. M.; SMILANSKY, J. Information and attitudes: The effects of information context and perceived discrepancy on attitudes. May, 1971. 23p. (OSU-CISRC-TR-71-5) (PB-202 622)

LI, Y. Equipment replacement models: A generalization and extension. May, 1971. 17p. (OSU-CISRC-TR-71-6) (PB-200 548)

ELLIOTT, D.; GEIS, M.; GROSU, A.; NOBEL, B.; ZWICKY, ANN; ZWICKY, ARNOLD. Working papers in linguistics no. 8. June, 1971. 197p. (OSU-CISRC-TR-71-7) (PB-202 724)

BOND, Z. S.; GREGORSKI, R.; KEREK, A.; LEHISTE, I.; SHOCKEY, L.; WENDELL, M. V. Working papers in linguistics no. 9. July, 1971. 232p. (OSU-CISRC-TR-71-8) (PB-204 002) (ED-060 686)

JOHNSEN III, A. M. Performance in memory scan task under conditions of fixed versus varied memory sets. November, 1971. 49p. (OSU-CISRC-TR-71-9) (PB-210 432)

MEHAFFEY III, L. The spectral sensitivity of the turtle Pseudemys scripta elegans. November, 1971. 89p. (OSU-CISRC-TR-71-10) (PB-210 433)

WILLIAMS, J. D. Stimulus encoding and human information processing. November, 1971. 65p. (OSU-CISRC-TR-71-11)

SANDERS, S. A. A modification of a method of generating random numbers using a combination of two congruential generators. December, 1971. 76p. (OSU-CISRC-TR-71-12)

LANDRY, B. C. A theory of indexing: Indexing theory as a model for information storage and retrieval. December, 1971. 271p. (OSU-CISRC-TR-71-13) (PB-205 829) (ED-057 843)

OSTROM, T. M.; EDWARDS, J. D.; ROSENBLOOD, L. K. Integration of discrepant information in interpersonal attitudes. December, 1971. 43p. (OSU-CISRC-TR-71-[illegible]) (PB-210 500)

1972

WHITE, L. J.; GILLENSON, M. L. Optimum center location. January, 1972. 69,A6,B32p. (OSU-CISRC-TR-72-1)

MELTZER, D. Speech synthesis by haar functions with comparison to a terminal analog device. January, 1972. 135p. (OSU-CISRC-TR-72-2)

CHANDRASEKARAN, B.; JAIN, A. K. Quantization of independent measurements and recognition performance. March, 1972. 14p. (OSU-CISRC-TR-72-3) (AD-747 706)

CAMERON, J. S. Automatic document pseudoclassification and retrieval word frequency techniques. March, 1972. 165p. (OSU-CISRC-TR-72-4) (PB-210 435)

OSTROM, T. M.; SLOAN, L. R.; MG CULLOUGH, J. L. Information and attitudes: The effects of repetition and amount of information. April, 1972. 38p. (OSU-CISRC-TR-72-5) (PB-209 802) (ED-061 968)

LEHISTE, I.; MELTZER, D.; SHOCKEY, L.; GREGORSKI, R. Working papers in linguistics no. 12. June, 1972. 88p. (OSU-CISRC-TR-72-6) (PB-210 781) (ED-069 162)

EKONG, V. J. U. Rate of convergence of hermite interpolation based on the roots of certain Jocabi polynomials. June, 1972. 54p. (OSU-CISRC-TR-72-7) (PB-211 237)

WEIMAN, C. F. R.; ROTHSTEIN, J. Pattern recognition by retina-like devices. July, 1972. 154p. (OSU-CISRC-TR-72-8) (PB-214 665/2) (ED-065 144)

CHANDRASEKARAN, B.; REEKER, L. H. Artificial intelligence - a case for agnosticism. August, 1972. 28p. (OSU-CISRC-TR-72-9) (ED-069 296)

ROTHSTEIN, J. Loschmidts's and Zermelo's paradoxes do not exist. October, 1972. 9p. (OSU-CISRC-TR-72-10) (PB-213 712/8) (PB-214 666/0)

YOVITS, M. C.; CHANDRASEKARAN, B. Artificial intelligence. December, 1972. 39p. (OSU-CISRC-TR-72-11) (AD-760 782)

HSIAO, D. K.; DENICOFF, M.; BERGART, J. G. An annotated and cross-referenced bibliography on computer security and access control in computer systems. November, 1972. 57p. (OSU-CISRC-TR-72-12) (AD-755 225) (ED-078 879)

FOULK, C. R.; JUELICH, O. C. Smooth programs and languages. November, 1972. 16p. (OSU-CISRC-TR-72-13) (PB-218 805/0) (ED-071 683)

YOVITS, M. C.; WHITTEMORE, B. A generalized conceptual development for the analysis and flow of information. December, 1972. 31p. (OSU-CISRC-TR-72-14) (PB-215 530/7)

MATHIS, B. A. Techniques for the evaluation and improvement of computer-produced abstracts. December, 1972. 262p. (OSU-CISRC-TR-72-15) (PB-214 675/1) (ED-071 686)

1973

LAY, W. M. The double KWIC coordinate indexing technique: theory, design, and implementation. February, 1973. 250p. (OSU-CISRC-TR-73-1) (PB-228 181/4GI) (ED-072 826)

YOUNG, C. E. Development of language analysis procedures with application to automatic indexing. April, 1973. 298p. (OSU-CISRC-TR-73-2) (PB-227 088/2GI) (ED-078 843)

STRONG, S. J. An algorithm for generating structural surrogates of English text. April, 1973. 140p. (OSU-CISRC-TR-73-3) (PB-227 071/8GI) (ED-078 844)

HSIAO, D. K. Logical access control mechanisms in computer systems. July, 1973. 24p. (OSU-CISRC-TR-73-4) (AD-765 727/3) (ED-078 880)

WANG, P. T. R. Bandwidth minimization, reducibility decomposition, and triangularization of sparse matrices. September, 1973. 162p. (OSU-CISRC-TR-73-5) (PB-227 396/9GI)

BIERMANN, A. W.; BAUM, R.; KRISHNASWAMY, R.; PETRY, R. E. Automatic program synthesis reports. October, 1973. 62p. (OSU-CISRC-TR-73-6) (AD-771 332/4GI) (ED-086 167)

WYRICK, T.; HSIAO, D. K. System configuration study on the interconnection of the IBM 370/165, DEC System-10 and MICRO 1600/21 computer systems. October, 1973. 26p. (OSU-CISRC-TR-73-7) (AD-770 611/2GI)

1974

HSIAO, D. K.; KERR, D. S.; MC CAULEY III, E. J. A model for data secure systems (Part I). February, 1974. 39p. (OSU-CISRC-TR-73-8) (AD-778 657/7GI) (ED-090 983)

HSIAO, D. K.; KERR, D. S.; NEE, C. J. Context protection and consistent control in data base systems (Part I). February, 1974. 28p. (OSU-CISRC-TR-73-9) (AD-778 658/5GI) (ED-090 982)

BAUM, R. I.; HSIAO, D. K. A data secure computer architecture. July, 1974. 22p. (OSU-CISRC-TR-73-10) (AD-787 580/0GI)

RUSH, J. E.; PEPINSKY, H. B.; LANDRY, B. C.; MEARA, N. M.; STRONG, S. M.; VALLEY, J. A.; YOUNG, C. E. A computer assisted language analysis system February, 1974. 71p. (OSU-CISRC-TR-74-1) (ED-090 948)

REAMES, C. C.; LIU, M. T. Variable-length message transmission for distributed loop computer networks (Part I). June, 1974. 9p. (OSU-CISRC-TR-74-2) (PB-242 913/2GI) (ED-092 152)

KIRSCH, B. M. An improved error diagnositcs system for IBM System/360-70 assembler program dumps. June, 1974. 60p. (OSU-CISRC-TR-74-3) (PB-235 060/1GI) (ED-092 153)

COHEN, D.; LIU, M. T. Emulation of computer networks by microprogrammable microcomputers. June, 1974. 17p. (OSU-CISRC-TR-74-4) (PB-242 912/4GI)

BIERMANN, A. W.; KRISHNASWAMY, R. Constructing programs from example computations. August, 1974. 41p. (OSU-CISRC-TR-74-5)

BUTTELMANN, H. W. Semantic directed translation of context free languages. September, 1974. 36p. (OSU-CISRC-TR-74-6) (PB-242 854/AS)

HSIAO, D. K.; KAFFEN, N. E. A model for data secure systems (Part II). October, 1974. 40p. (OSU-CISRC-TR-74-7) (AD/A-006 276/0GI)

HSIAO, D. K.; KAFFEN, N. E. A model for data secure systems (Part III). November, 1974. 32p. (OSU-CISRC-TR-74-8) (AD/A-004 676/3GI)

HSIAO, D. K.; BAUM, R. I. Information secure systems. November, 1974. 55p. (OSU-CISRC-TR-74-9)

1975

HARTSON, H. R.; HSIAO, D. K. Languages for specifying protection requirements in data base systems (Part I). January, 1975. 61p. (OSU-CISRC-TR-74-10) (AD-006 280/2GI)

PETRY, F. E. Program inference from example computations represented by memory snapshot traces. February, 1975. 143p. (OSU-CISRC-TR-75-1)

MC CAULEY III, E. J. A model for data secure systems. March, 1975. 119p. (OSU-CISRC-TR-75-2) (AD/A-011 359/7GI)

HORGER, JR., W. A. Data base module verification--a certification method for data secure systems. June, 1975. 125p. (OSU-CISRC-TR-75-3)

RODEHEFFER, T.; KAFFEN, N. PREST4--A highly structured FORTRAN language for systems programming. July, 1975. 36p. (OSU-CISRC-TR-75-4)

APPENDIX G

DOCTOR OF PHILOSOPHY DEGREE

1971

LANDRY, B. C[illegible]S A Theory of Indexing: Indexing Theory as a Model for Information Storage and Retrieval.

1969

CAMERON, JAMES S. Automatic Document Pseudoclassification and Retrieval by Word Frequency Techniques

EKONG, VICTOR J. Rate of Convergence of Hermite Interpolation Based on the Roots of Certain Jacobi Polynomials

GORDON, ROBERT The Organization and Control of a Slave Memory Hierarchy

HARRIS, DAVID R. GOLDA: A Graphical On-Line System for Data Analysis

MATHIS, BETTY ANN Techniques for the Evaluation and Improvement of Computer-Produced Abstracts

WEIMAN, CARL F. R. Pattern Recognition by Retina-Like Devices

WHITTEMORE, BRUCE J. A Generalized Decision Model for the Analysis of Information

1973

CHAN, PAUL SUI-YUEN An Investigation of Symetric Radix for Computer Arithmetic

DEFANTI, THOMAS A. The Graphics Symbiosis System - an Interactive Mini-Computer Animation Graphics Language Designed for Habitability and Extensibility

GELPERIN, DAVID H. Clause Deletion in Resolution Theorem Proving

HEPLER, STEPHEN PHILIP Use of Probabilistic Automata as Models of Human Performance

LAY, W. MICHAEL The Double-KWIC Coordinate Indexing Technique: Theory, Design, and Implementation

WANG, PAUL TIING RENN Bandwidth Minimization, Reducibility Decomposition, and Triangulation of Sparse Matrices

YOUNG, CAROL E. Development of Language Analysis Procedures with Application to Automatic Indexing

1974

BEUG, JAMES L. Human Extrapolation of Strings Generated by Ordered Cyclic Finite State Grammars

DOHERTY, MICHAEL E. A Heuristic For Minimum Set Covers Using Plausability Ordered Searches

GILLENSON, MARK L. The Interactive Generation of Facial Images on a CRT Using a Heuristic Strategy

LONGE, OLUWUMI An Index of Smoothness for Computer Program Flowgraphs

SU, HUI-YANG Pagination of Programs for Virtual Memory Systems

INVESTIGATOR INDEX

SUBJECT INDEX