AFOSR 69-3000 TR
DECEMBER 1969

AIR FORCE OFFICE of SCIENTIFIC RESEARCH
Directorate of Information Sciences

information sciences 1968

OFFICE of AEROSPACE RESEARCH ★ U.S. AIR FORCE
Arlington, Virginia 22209

AFOSR 69-3000TR

INFORMATION SCIENCES--1968

Being a Summary Report of the Contracts and Grants of the Directorate of Information Sciences active during all or a part of 1968.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

Office of Aerospace Research

Arlington, Virginia 22209

December 1969

TABLE OF CONTENTS

INTRODUCTION AND PROGRAM HIGH-LIGHTS

The AFOSR Information Sciences program exists to serve two related Air Force communities. One is concerned with improving the handling of scientific and technical information in Air Force libraries, information centers and management information systems. The other is concerned with meeting Air Force technical objectives in such applied areas as avionic communications, electromagnetic intelligence, ground-based surveillance, ground communications, information displays, information processing, reconnaissance and aerospace photography and optronics.

This report lists the progress made on 65 research contracts and grants that were active during 1968 in the four areas designated by the COSATI subject headings of documentation and information technology, language and linguistics, bionics, and information theory. Selected highlights of the program are listed in this section. Following this the reader will find brief statements of the progress made under all efforts monitored by this Directorate during the reporting period, together with citations to the 125 reports, journal articles, books and chapters in symposium proceedings they produced.

The names of the investigators cited in the following sections are followed by a number (e.g., 1.21) referring to the section of this publication listing their publications. Copies of cited publications may be obtained either through the Defense Documentation Center or the Clearinghouse for Federal Scientific and Technical Information.

DOCUMENTATION AND INFORMATION TECHNOLOGY

Robert A. Fairthorne, (1.19) State University of New York, Albany, stated the principles of Impotence that limit the behavior of information systems: Discursive Impotence, which states that systems must accept discourse as authors write it or readers make use of it or wish to make use of it: the Principle of Ignorance, which states that one cannot tell whether a statement is true just by reading it, and: the Principle of Delegation, which states that no one can substitute for someone else's reaction to discourse, but can only delegate for him.

A delegate must be told what to do explicitly, as well as being able to do what he is told.

Gerald Jahoda, (1.03) Florida State University, investigated the use of computer-prepared indexes to researchers' document collections. He found that despite initial interest in the keyword from title indexes, the researchers for whom these indexes were prepared preferred to go directly to the documents filed by subject.

William Goffman, (1.02) Case Western Reserve University, applied the mathematical theory of epidemics to the transmission of information among a community of scientific researchers... specifically the field of symbolic logic since 1847. He found that the growth of this field resembled a recurring epidemic process with peak points occurring at 25 year intervals, each peak higher than its immediate predecessor, and that each outburst of activity was made up of a sequence of overlapping epidemic processes each of which corresponds to a specific subject area--a phenomenon curiously akin to Brownian motion, in that random motion of individuals from field to field produces diffusion effects on the entire field. A theoretical paper modelled the interaction of a literature, its users and its creators as an epidemic process with an intermediate host. Within this model it was shown that for every literature there exists a point in time at which the literature becomes too large to support the natural flow of ideas within its subject area. A consequence of this theorem is that one can specify the exact point when an artificial device, e.g., an information retrieval system, may be needed in order to prevent the communication process from breaking down.

S. Watanabe, (1.06) University of Hawaii, made three major breakthroughs in investigating the methodological foundations of pattern recognition: he showed the formal relationship between pattern recognition and inductive inference; he produced a useful algorithm based on a philosophical investigation of object-predicate reciprocation, and; he made a major contribution to the empirical foundation of logic by his study of the continuously-valued characteristic function.

Donald Brick, (1.08) Information Research Associates, has been studying the application of Wiener's canonical expansions (where the structure of the expansion, or circuitry, is independent of input and

function) to finite-bandwidth, finite-time duration and/or finitely quantized (real) signals. He has applied these expansions to the study of sinusoidal signals in noise, where the processor acts as a discriminator of amplitude and, to some extent, frequency. This general procedure has applications to target classification, source identification and information extraction.

Laveen Kanal, (1.10) Lehigh University, has defined a new model derived from a class of infinite-state Markov chains for errors occurring in a real communication channel, and has shown that all preceding finite-state Markov models, and several others are but special cases of his model. He is now devising a general technique, based on his infinite state model, to evaluate the performance of error correcting coding techniques.

Herbert Freeman, (1.13) New York University, has been applying linguistic theory to pattern analysis. A pattern analyzer program has been written for which the grammar of a class of patterns can be specified in tabular form, and patterns then examined in a manner analogous to the parsing of an English-language sentence. The work has been applied to the analysis of bubble-chamber photographs, vein structures of tree leaves, and variously shaped alphanumeric characters.

He has also been investigating the problem of solving the "hidden-line" problem for solids bounded by quadric surfaces. Procedures have been developed that permit some of the key calculations to be performed in two rather than three dimensions, with significant reductions in computing time.

G. Krulee and Ben Mittman, (1.15) Northwestern University, have been working on the development of an effective remote-terminal based information retrieval system. INFOL, a language primarily directed toward structured data storage and retrieval, and TRIAL for bibliographic applications in searching and indexing, have been made operational for the CDC 6400. INFOL has been used for storage and selective retrieval of cardiology and dermatology cases. TRIAL has been used to produce bibliographies in geological literature and in computer sciences, among many other topics.

June Shoup, (1.17) Speech Communications Research Laboratory, has made three major advances in instrumentation for speech analysis: the construction of a moderately-priced TV system which permits a visual display such as a sound spectrogram to be displayed on a computer controlled CRT and to be traced with a light pen. This computer-aided measurement of sound spectrograms proved to be at least four times as fast as hand measurement. A special face mask permits simultaneous records to be obtained of air flow, of sub-glottal pressure, of the acoustic wave and of motion pictures of the true vocal folds. This mask allows the laboratory to be the first group to obtain data from the above-mentioned types of records simultaneously. A program implemented the fast Fourier transforms on a small computer, which was essential to expedite their experimental study of the acoustic properties of the fricatives and sibilants of English, and in the automatic tracking of acoustic parameters of the speech wave.

LANGUAGE AND LINGUISTICS

Philip Bagley, (2.18) University City Science Center, completed a two year study on ways of designing new computer programming languages independent of the hardware characteristics of current computers. He developed a new technique for expressing data values, data elements and data structures, but found that much more freedom of storage organization is needed than is provided by current machines.

Dick Bennett, (2.14) Signatron, developed BUILD (Bases for Uniform Language Definition) a base for the definition of a wide class of languages--e.g., by an advanced programmer using the language for himself for writing a complex program or would design a language for others in a particular applications area. If it works, BUILD promises to simplify and unify computer languages so that they may all be built on one universal base and may all be handled by one general translator.

Ran Banerji, (2.4) Case Western Reserve, working on game theory, has found that for any Koffman type positional game he can now construct efficiently the descriptions of forcing positions of depths about 10 and of various degrees of sublety (i.e., immediate threats, threats to threaten, threats of threats to threaten, etc.)

Von Glasersfeld, (2.07) Georgia Institute for Research, has made his automatic parsing procedure for English sentences fully operational on an IBM 360/65 computer. Its output has been developed into a graphic display of sentence structure which presents the parsings like tree diagrams. Sentences up to 16 words are now processed in 0.5--2.0 seconds.

David Rothenberg, (2.11) New York Research Group, has developed a pattern recognition model which maps a possibly infinite and continuous space, over which sensory stimuli range into a finite space of discrete points, the classifications of such stimuli. The model has been successfully applied to human auditory perceptions and to visual perception.

Saul Amarel, (2.13) RCA, has been working on a question answering system with a data base consisting of descriptions of geometrical patterns arranged in two dimensions and with a question language which is a fragment of English. The question language and the answer-generation techniques have been extended so that they can process questions involving arithmetic.

Seymour Ginsberg, (2.15) System Development Corporation, working on the theory of programming languages, has devised the theory of an abstract family of languages (AFL). The notion of an AFL turns out to be extremely important for language theory. It unifies the theory, allows algebraic proofs to replace machine arguments and points the way for new developments. In particular, isolated facts about closure properties and many diverse proofs are now part of a general algebraic theory with many special cases. Additionally, new facts on substitution, intersection and hierarchy relationships among well-known families of languages have been obtained.

Lew Clapp, (2.06) Computer Research Corporation, completed survey of 500 on-line computer systems and languages. On the basis of these data, a 2nd report on "Successful Techniques in Developing On-Line Systems" is being written. The data base, from 500 surveys, was studied and manipulated by an on-line system, written in SNOBOL!

BIONICS

Les Gerhardt and Johannes Goerner, (3.01) Bell Aerosystems, have

been applying fixed linear adaptive gain networks to a variety of practical problems in pattern recognition, mixture analysis, system modeling and adaptive control. They have developed an adaptive non-deterministic search procedure which is an improvement over presently known methods because of its compatibility between high speed and accuracy.

Gerhard Hollander, (3.04) Hollander Associates, has developed a system for classifying the various approaches to adaptive systems into six groups of increasing cost and effectiveness, based on the review of over 400 reports. This basic classification scheme can be expanded to all adaptive systems and can yield or be the basis for quantitative cost and effectiveness values, so that eventually the best system for a given application can be determined by scanning the characteristics of the systems and the quantitative description of the application.

Earl Gose, (3.05) University of Illinois, Chicago Circle, has been working on a variety of pattern recognition problems. A color-selective flying spot scanner was constructed to provide pictorial input to the computer. Programs were written to measure shape, color and textural properties, and to perform filtering and noise-reduction. Their utility was demonstrated in the classification of brain cells, blood cells and breast tumors, using electron micrographs, photomicrographs and X-rays.

Heinz von Foerster, (3.06) University of Illinois, and his graduate students have worked on a variety of problems. These include:

An improved list structure system of the ring type, designed for storing and manipulating relational data of arbitrary complexity within the computer core memory.

The first phase of a powerful program debugging tool which uses a semantic analysis to detect program faults at a higher level then those found by the usual syntax check.

Developed a mathematical model of neural behavior which seems to be both physiologically realistic and formally satisfactory.

Development of a Visual Image Processor, a special purpose electronic computer specifically designed to compute on two-dimensional visual images where the intensities of an image point and those of its near neighbors constitute the independent variables of a function to be chosen by the operator. This works on the detection and registration of topological and geometrical properties such as single or multiple-connectedness.

William Kilmer, (3.08) Michigan State University, developed a computer model of the reticular formation (the structure in vertebrate central nervous systems that commits an animal to one or another gross modes of behavior) which shows habituation and recovery, conditioning and extinction and generalization and discrimination. Other work is in progress on the application of modern pattern recognition methods to the analysis of electroencephalographs.

R. B. McGhee, (3.10) University of Southern California, discovered that of 5,040 theoretically possible quadreped walking gaits there is only one which maximizes the dynamic stability of low speed locomotion. These results were applied to the manually controlled four-legged truck built by General Electric Company under ARPA sponsorship, since serious doubts had been raised concerning the dynamic stability of the machine. The machine was successfully tested in June 1968 using the gait recommended by Dr. McGhee. The optimality of this gait was unknown until the AFOSR-sponsored investigation.

George Bekey, (3.10) University of Southern California, developed a novel sequential hybrid computer algorithm for the solution of integral equations of the Fredholm type. It seems possible that this approach may lead to a real time "on line" technique for solving such equations.

Gordon Pask, (3.11) Systems Research, Ltd., has developed a computer simulated learning model which is run against the model for an adaptive teaching machine. This is compared with the behavior of human subjects run against a real adaptive teaching machine with similar properties. In practice, it is found that students adopt inept learning strategies which they can't describe but hate to give up. There is significant improvement in teaching efficiency if the teaching machine (1) determines the student's current strategy (2) selects a logically good strategy for the skill (3) arrives at a compromise between

the student's strategy and the good strategy (4) imposes the compromise strategy as a tutorial prescription, and (5) gives the student reasons for doing so.

INFORMATION THEORY

Kurt Bing, (4.08) RPI, developed a linear system of natural deduction in which the usually cumbersome quantifier rules are replaced by simpler ones, constituting a more convenient linear formalism for logical deduction. With it, one can accomplish the purpose in tree type systems of combining several deductions into one, but without having to trace branches of deduction trees.

Frank Cannonito, (4.02) U of California, Irvine, has found properties of information retrieval analogous to solving the word problem in semi-groups. He has consulted for RADC on logic, recursive function theory, especially as it applies to information sciences and artificial intelligence.

Ray Smullyan, (4.03) City University of New York, summarized approximately eight years of research under AFOSR sponsorship in a definitive monograph on first order logic. As he writes: "Mathematical logic is noted for its formal rigor. But many mathematicians feel that the way mathematical logic has been classically presented, though rigorous and precise, is neither elegant nor economical... Most of the general mathematicians who are not specialists in logic are remarkably uninformed about the subject... If one follows classical treatments, it takes a fantastically long time wading through the most pedantic and trivial details before one gets to first base concerning the mathematically significant results of logic. No wonder the general mathematician gets discouraged... (as do) the applied mathematician and worker in the computer sciences, who find logic helpful but not easily available." Smullyan has now reorganized mathematical logic in a form which makes the main results far more easily available to and useable by the general mathematician and worker in the computer sciences. He has also shown how the techniques involved lead to new results in mathematical logic itself, which in many ways go beyond those obtained by classical methods.

Martin Davis, (4.05) New York University, has obtained new results in various branches of mathematical logic and computability theory.

Specific relations have been found between various hierarchies of notations for transfinite ordinals. These systems of notations for transfinite ordinal numbers play a key role in finitary consistency proofs for formal theories, which in turn can be expected to influence work in mechanical theorem proving.

Bernard Elspas, (4.10) Stanford Research Institute, conducted a month-long seminar to try to bridge the gap between switching theory and automata theory on the one hand, and their applications in computers. The relationship between the practical art of computers and its associated theory is unique in engineering since such factors as noise, impurity of materials, inaccuracies in fabrication and the vagaries of complex environments, which are integral to fields like electronics or control engineering may be effectively ignored by many cases of practical importance. Nevertheless there is a general feeling that the gap between theory and practice is growing. For example, switching theory has not been of significant use in the design of large logic networks, nor have the theories of automata and formal languages been of significant use in the design of computer systems nor in the design of executive programs. The seminar had the following goals: (1) to analyze the state of the current relationship between the theory and practice of digital systems: (2) to distinguish specific technical problems for which new theory is needed or where new techniques are needed for the application of existing theory. The final report, now in preparation should provide guidelines and more importantly, impetus for theoretical research workers in automata, languages, switching and combinatorics to pursue problems of considerable practical relevance to the art of digital systems design, and (2) to make available in one place a statement of recent results in these fields that are of actual or potential value to practical machine designers.

Bill Kautz and Jim Turner, (4.11) Stanford Research Institute, completed a comprehensive survey of all published Soviet literature on graph theory and its applications. The availability of this report should draw attention to this important collection of mathematical and technological results. Although Soviet activity in graph theory and its applications lags behind corresponding Western work in quantity and quality, the level of activity is increasing rapidly and there are many excellent Soviet contributions to the theory. The best Soviet work has been concerned with bounds on numerical indices associated with graphs,

properties of algebraic structures associated with graphs, and operations on graphs.

Yehoshafat Give'on, (4.13) Stanford University, has found several methods developed in standard automata theory useful in other branches of mathematical systems theory. The most outstanding of these methods are in respect to minimization of systems, simulation of systems, and the reduction of simulation of systems to some well-known algebraic theory (e.g., semigroup theory for simulation of sequential machines.) He has been using category theory as a framework to study these methods. His work in algebra automata (an algebraic model for a certain kind of parallel information processing) and on linear systems (which are the core of control theory) has been directed towards finding the particular properties of the application of automaton theoretic methods.

STAFF PUBLICATIONS

In addition to the monitoring of contracts and grants, the Directorate staff produced the following publications during 1968.

Rowena W. Swanson, Information, An Exploitable Commodity. Invitational paper for the Fourth Annual Meeting of the Information Processing Assn. of Israel. April 1968. IPA Monograph No. 5, AFOSR 68-0652, AD 677197.

Rowena W. Swanson, "Information Services for Small-Scale Industry." American Documentation, Vol. 19, No. 4 (Oct 1968) 412-413.

Harold Wooster, "Our Nervous Science Publishers," The Torch, (July 1968) 14, 28-32; AFOSR 68-1880, AD 674991.

Harold Wooster, "Towards A Uniform Federal Report Numbering System and a Cuddly Microfiche Reader--Two Modest Proposals," Paper presented at the Third Annual Northeastern Documentation Center/Industry Users Conference, Waltham, Massachusetts, 17 April 1968; AFOSR 68-0772, AD 669204.

Harold Wooster and Eliot Sohmer, Information Sciences - 1967, (Scientific Report, January 1968; AFOSR 68-0006).

Harold Wooster, "Machina Versatilis--A Modern Fable" Library Journal, 94, 725 (15 Feb) 1969.

Harold Wooster, "The Future of Scientific Data" OAR Research November 1968.

Harold Wooster, "Airlie House Conferences", in Encyclopedia of Library and Information Science, Vol I. Dekker, NY 1968.

Section I

DOCUMENTATION AND INFORMATION TECHNOLOGY

The problem really begins when someone, we will assume altruistically motivated, creates something called data. Much of this data is in tabular, graphic, or symbolic form. At this instant a highly predictable, and absolutely irreversible, process takes place. A convention of long standing dictates that the results of this labor be written up for presentation at meetings and publications in numerous journals. As a result, additional data, often referred to as verbiage, is generated to connect the original data together. Behold - a document has been created.

Sometime after the creation has taken place, most authors decry the time lag, the document is abstracted by a "competent" abstractor who is well versed in the subject matter. This is also the time when key words are identified and isolated for future use in locating the document. This process is assured since the key words are very significant in the eyes of the abstractor. Should the author disagree, then it is obvious that the author never fully appreciated the content of his paper and indeed may not have been competent to write it in the first place.

At this point the paper is stored on magnetic tape or photographic film. It is connected to the outside world at the lowest level by the original document, at the next level by the abstract, and at the highest level by the key words. We have now arrived at the far end of the communication spectrum, and, once again, we are dealing with data - not documents. A great deal of effort is constantly being expended to strengthen the weak links of this data chain.

Some of the brightest lights appear to be new concepts in data storage allocation, both area and level, and highly efficient retrieval algorithms. Concept coordination, permutated indexes, structure tables and systems languages, and automatic indexing and abstracting techniques have all contributed toward better and more reliable documentation and information technology. Even with these initial successes, the problem still remains far from solved for the large scale systems, and successful implementation on small, low power, light weight systems is barely visible on the horizon.

1.1 INFORMATION PROCESSING PERSONNEL SURVEY AND DATA BASE ANALYSIS

H.G. Asmus and Isaac D. Nehama
American Federation of Information Processing Societies (AFIPS), New York, New York
Current Contract Number* F44620-67-C-0092

Summary of Recent Research Results

A final report resulted which presented tabular and graphic summaries of the data collected during a 1968 survey conducted by AFIPS of personnel engaged in the information processing field. The data were acquired by questionnaire. The questionnaire was distributed by seven professional members of AFIPS to their memberships. The survey reported on data from 29,826 returns, and pertain to the following five categories of information: personal data, education, employment, professional activities, salary and income. Among the respondents, 19% are in the 25-29 year age range, 22% in the 30-34 year age range, and 20% in the 35-39 year age range; 87% are male. Educational achievements were as follows: Ph.D.'s, 7.5%; M.S.'s, 24%; and, B.S.'s, 37%. Engineering was the principal discipline, mathematics was second, and the physical sciences were third. The primary occupational specialty was programming, while the chief areas of application were scientific and engineering, and business and administration. An exhaustive analysis of regional and national salary variations was included.

Reports and Publications

M.R. Davis and I.D. Nehama, "Information Processing, Personnel Survey, 1968," Final Scientific Report. AFOSR 69-1434TR AD 688937

*This research effort is supported by the Advanced Research Projects Agency (ARPA) and monitored, both technically and administratively, by the Directorate of Information Sciences, AFOSR.

1.2 APPLICATIONS OF A MODEL FOR COMMUNICATION INFORMATION TRANSMISSION PROCESSES

William Goffman
Case Western Reserve University, Cleveland, Ohio
Contract and Grant Numbers AF49(638)-357, AF-AFOSR-35-62, AF-AFOSR-403-63, AF-AFOSR-403-64, AF-AFOSR-403-65, AF-AFOSR-403-66

Summary of Recent Research Results

The mathematical theory of epidemics is being applied to the information transmission phenomena evidenced by the evolution of concepts in the field of symbolic logic as they have become manifest through the literature since the work of Boole and DeMorgan. Peak points appear to have occurred since 1847 at 25 year intervals, each succeeding outburst of activity being greater in intensity than its immediate predecessor. Each outburst is composed of a sequence of overlapping epidemic processes, each of which represents a specific subject area in symbolic logic as defined by Church's subject index. Although the field seems to have developed according to a specific pattern, i.e., one area and then another drew the attention of researchers, individual researchers have not tended to move from one area to another in a predictable manner. A secondary effort has involved the completion of an experiment with searching strategies. Theory has been related to information retrieval systems.

Reports and Publications

W. Goffman, "Mathematical Approach to the Spread of Scientific Ideas--The History of Mast Cell Research" Nature 212, 449-452 1966 AFOSR 69-0030TR, AD 686885

1.3 ANALYSIS OF PERSONAL INDEX STRUCTURES AND USES

Gerald Jahoda
Florida State University, Tallahassee, Florida
Grant Numbers AF-AFOSR-895-65, AF-AFOSR-895-67

Summary of Recent Research Results

Searches were conducted on three indexes to a collection of 3200

documents in the field of chemistry. The indexes were (a) keyword from title without added entries, (b) keyword from title with added entries, and (c) a single access point per document alphabetic subject index. Search results are characterized in terms of recall, precision, search time, and three other single figure measures. A measure of index performance based on recall, precision, and search time was developed. No statistically significant difference was found between search results with multiple access points per document keyword from title index and the single access point per document alphabetic subject index. Statistically significant differences in search results were found between (a) and (b). The effect of the size of the delete word list and of elements of vocabulary control in keyword form title indexes were also studied.

Reports and Publications

G. Jahoda and M. L. Stursa "Tests of Indexes--A Comparison of Keyword from Title Indexes With and Without Added Keywords and a Single Access Point Per Document Alphabetic Subject Index" AFOSR-69-0445TR AD 683750

Gerald Jahoda, "Planning Improved Library Service for Scientists in Universities." College and Research Libraries 28, 343-346, 1967. AFOSR-68-1238 AD 670645

1.4 STRATEGIES FOR MANIPULATING UNIVERSAL DECIMAL CLASSIFICATION RELATIONSHIPS FOR COMPUTER RETRIEVAL

T. W. Caless
George Washington University, Washington, D.C.
Contract Number F44620-68-C-0035

Summary of Recent Research Results

Special subject analysis matrices are being developed for an earth sciences document collection of 1000 documents that will serve as the data base for testing the effectiveness of the Universal Decimal Classification for computer retrieval. Relator schema, to be used as a UDC colon substitute, have been shown to be a convenient device for making fine distinctions between classes of knowledge, and a

simplified notational technique has been developed for ease of application. UDC rules of formation, procedures for preparing UDC schedule revisions, and computer search strategies are being developed.

Reports and Publications

T. W. Caless, "Searchability," Presented to the Reclassification Conference sponsored by the School of Library Science and Information Services, The University of Maryland on 4-6 April 1968. Proceedings to be published in December 1968. 5 pp.

T. W. Caless, "Proposed Revision to Perreault's Relator-Schema," submitted to the International Federation for Documentation in June 1968, issued in 068-18 to the FID/CCC dated 22 August 1968, p. 8-9.

T. W. Caless, "Perreault's Relators Applied to Earth Sciences Collection Classified with UDC," to be published in 1969 by VINITI in Papers Contributed to International Information Congress Scheduled for 16-18 September 1968 (suspended). 9 pp.

T. W. Caless, "Subject Analysis Matrices for Classification with UDC," presented to the UDC Seminar held in Copenhagen on 2-6 September 1968, to be published in the Proceedings in 1969, 5 pp.

T. W. Caless, "Plan for Implementing an Information System for Marine Sciences Literature," presented to the International Marine Information Symposium sponsored by the Marine Technology Society on 31 October - 1 November 1968. Proceedings to be published in 1969. p. 49-58. AFOSR 68-2578 AD 678837.

T. W. Caless, "Subject Analysis Matrices for Document Classification," submitted to the Classification Society, North American Branch for presentation at the Annual Meeting in connection with the American Association for the Advancement of Science to be held in Dallas, Texas on 28 December 1968.

1.5 ANALYSIS OF INFORMATION SCIENCE TERMINOLOGY

Charles W. Shilling and Patricia Fuellhart
George Washington University, Washington, D.C.
Current Grant Number AF-AFOSR-1325-67

Summary of Recent Research Results

This study is based on an analysis of the terms and concepts contained in a collection of lexical resources in Information Sciences--thesauri, vocabularies, term lists and classification schemes. These resources have been analysed with the purpose of tabulating the occurrence of terms. This study, therefore, has been aimed at a quantitative review of lexical aids to information science language with the objective of determining its content and the frequency with which concepts have been recorded. This analysis, accomplished from a matrix using the IBM S 360/40 has enabled the researchers to quantify the terminology and to prepare a series of tables illustrating the frequency of both generic and specific terms appearing in a collection of U.S. and European lexical aids to the language of publication and discourse as it has been recorded in lexical form. An inventory and annotation has been prepared to identify and describe the lexical resources available for review. The publication dates of the source materials range from 1960 to 1967.

Reports and Publications

P.O. Fuellhart and D.C. Weeks, "Compilation and Analysis of Lexical Resources in Information Sciences." June 1968. AFOSR 68-1375 AD 671148.

1.6 SELECTION OF VARIABLES IN CLUSTERING AND PATTERN RECOGNITION

Michael S. Watanabe
University of Hawaii, Honolulu, Hawaii
Grant Numbers AF-AFOSR-1187-67, AF-AFOSR-1466-68

Summary of Recent Research Results

Pattern recognition methodology was extended in a paper entitled,

"Mathematical Explication of Classification of Objects" by M. S. Watanabe. It clearly showed the relation of pattern recognition to inductive inference. A useful algorithm, based on an investigation of object-predicate reciprocation, was produced which significantly reduced the required dimensionality of feature space. This permitted the analysis of problems which had been hardware limited. A major contribution to the empirical foundation of logic was made through the study of the continuously-valued characteristic function. The SELFIC and CLAFIC feature extraction programs were extended and now converge to a well-balanced decision function. This was brought about by generating a separating hyperplane which was spaced between two, or more, clusters based on proportionality and constraints on its normal. It should be stressed that this process is adaptive, in that the data may be presented sequentially as the algorithm converges to a solution. Other investigators are now beginning to employ some of these methods. The use of this algorithm is now being extended to geography, meteorology, medical diagnosis, and psychology.

Reports and Publications

T. Kaminuma, T. Takekaw and S. Watanabe, "Reduction of Clustering Problem in Pattern Recognition." January 1968. AFOSR 69-0751TR AD 684625.

S. Watanabe, "Object-Predicate Reciprocity and its Application to Pattern Recognition." December 1968. AFOSR 69-0763TR, AD 684594.

1.7 CONFERENCE ON METHODOLOGIES OF PATTERN RECOGNITION

Michael S. Watanabe
University of Hawaii, Honoluly, Hawaii
Grant Number AF-AFOSR-1379-68

Summary of Recent Research Results

A volume containing the abstracts of papers presented at the Conference held at the University of Hawaii in January 1968 has been published. The papers are concerned with theoretical foundations of pattern recognition methodologies, and consider, inter alia, the

selection of variables, definitions of class features, limitations on domains of applicability of algorithms, assessments of decision-making procedures, and measures of goodness of variables. A proceedings volume is in press.

Reports and Publications

M. S. Watanabe, "International Conference on Methodologies of Pattern Recognition, Abstracts." January 1968. AFOSR 68-515 AD 666679.

1.8 APPLICATION OF WIENER CANONICAL EXPANSIONS TO PATTERN RECOGNITION

Donald B. Brick
Information Research Associates, Inc., Lexington, Mass.
Contract Number AF49(638)-1631

Summary of Recent Research Results

Work with the feature processors and preprocessors has shown that: (1) all processes must be non-deterministic; (2) certain continuity requirements must be met; (3) all processes derived via integrable operations on white noise may be handled; (4) the behavior of orthogonal-increment processes is well defined; and, (5) that symmetrically-distributed processes simplify the operation. It has been shown that 18 Laguerre and Hermite terms are still sufficient for the discrimination of sinusoidal signals in white noise. The primary parameteric discriminant is amplitude, however frequency is valuable in some instances. The design of analog Laguerre filters was undertaken and completed so they could be readily constructed if needed in future investigations. The theoretical efforts have shown how this work can be extended into multiple-random process inputs and outputs, inputs composed of noisy digital waveforms and inputs composed of damped sinusoids in noise. It was further shown that: (1) the procedure extends in a straightforward manner to the multi-dimensional case; (2) noisy digital waveforms lead to digital realizations in terms of fundamental digital building blocks; and, (3) damped sinusoids in noise can be handled through several different approaches. As a direct result of this investigation, it is now possible to develop

tapped delay lines and majority-logic units which have valuable and unique applications in digital pattern recognizers, encoders, decoders, digital logic systems, function generators, and communication and computing systems.

Reports and Publications

Donald B. Brick, "On the Applicability of Wiener's Canonical Expansion for the Likelihood Function of a Continuous Martingale" (Scientific Technical Report Number 8, 20 March 1968, AFOSR-69-0950 AD 650193).

1.9 OPTIMIZATION AND EVALUATION OF MEMORY ORGANIZATIONS FOR LARGE SCALE FILING SYSTEMS

Kenneth B. Krohn
Krohn-Rhodes Research Institute, Inc., Washington, D.C.
Contract Numbers F44620-67-C-0112, F44620-69-C-0020

Summary of Recent Research Results

Work has continued on the study of combinatorial semigroups, that is, by studying classes of combinatorial semigroups closed under wreath products and division, or in terms of automata classes of combinatorial machines all of which can be realized by series-parallel or cascade combinations of certain basic semigroups.

1.10 TOPICS IN PATTERN RECOGNITION AND COMMUNICATION SCIENCE

Laveen N. Kanal
Lehigh University, Bethlehem, Pennsylvania
Grant Number AF-AFOSR-1390-68

Summary of Recent Research Results

A new model was defined which was derived from a class of infinite-state Markov chains and applied to errors occurring in a real communications channel. It was shown that all preceeding finite-state Markov models and several other models, including the often proposed Pareto model, are but special cases of this new model. It was further

shown how to synthesize a model from only the Error-Free-Run and Error-Cluster-Run distributions. The fundamental significance of these distributions was demonstrated as well as their relationship to other models. This work is now being extended to produce a general technique for the evaluation of the performance of error correcting ccding techniques using the new model. An analytic technique was also developed to evaluate the performance of the non-linear receiver structure which minimizes the output errors when the intersymbol interference, which occurs when transmitting information through real channels at or near Nyquist rate, is known. Prior to this effort, only a computer simulation was available to approximate a solution. It is now possible to calculate the performance of some simple optimum receivers. New light was also shed on some questions of dimensionality and sample size which are important in the statistical design of pattern classification systems. This has shown that hitherto published results are but extremes in a range of possibilities.

Reports and Publications

Laveen N. Kanal, (Editor), Pattern Recognition, (Book), Thompson Book Co., Washington, D.C., 1968

T.J. Harley, L.N. Kanal, and N.C. Randall, "System Considerations for Automatic Imagery Screening," pp. 15-31 in Pictorial Pattern Recognition, edited by G.C. Cheng et all, Thompson Book Co., Washington, D.C. 1968

L. Kanal and B. Chandrasekaran, "Recognition, 'Machine Recognition' and Statistical Approaches," to appear in the Proceedings of International Conference on Methodologies of Pattern Recognition held at Honolulu, Hawaii, January 1968, edited by S. Watanabe; to be published by Academic Press 1969

L. Kanal and B. Chandrasekaran, "On Dimensionality and Sample Size in Statistical Pattern Classification," Proceedings 1968 National Electronics Conference, p. 2-7, NEC, Inc., Oak Brook, Illinois, 1968

L. Kanal, "Automatic Imagery Screening," invited lectures given to the Nato Advanced Study Institute on Automatic Classification and Interpretation of Images, Pisa-Tirrenia, Italy, August 1968

B. D. Fritchman, "A Binary Channel Characterization Using Partitioned Markov Chains," IEEE Trans. on Inf. Theory, pp. 221-227; April 1967

K. Abend, T. J. Harley, B. D. Fritchman and C. Gumacos, "On Optimum Receivers for Channels Having Memory," IEEE Trans. on Inf. Theory, November 1968

1.11 AIR FORCE SCIENTIFIC RESEARCH BIBLIOGRAPHY

Tom Goodwin
Library of Congress, Washington, D. C.
ISSA Numbers ISSA-00005-58, ISSA-00002-59, ISSA-00004-60, ISSA-00002-62, ISSA-00002-63, ISSA-00013-64, ISSA-00001-65, ISSA-00007-66, ISSA-00013-66, ISSA-0010-67

Reports and Publications

T. C. Goodwin, V. G. Hooker, D. C. Yates, P. A. Patrick, A. H. Hatch and J. F. Lindsay: Air Force Scientific Research Bibliography 1962. AFOSR 700-VI. 1968.

1.12 INTERNATIONAL FEDERATION FOR INFORMATION PROCESSING CONGRESS 68

L. W. Cohen
National Research Council, Washington, D. C.
Grant Number AF-AFOSR-1511-68

Summary of Recent Research Results

The triennial Congress of the International Federation for Information Processing (IFIP) was held in Edinburg in August 1968. About 3700 delegates attended, 585 from the United States, with some 47 countries represented. Of the 246 technical papers delivered, 112 were from the U.S. Of 281 applications submitted to a review committee for travel support, 74 were accepted, of which 18 were furnished through AFOSR sponsorship and 42 by NSF. The technical program consisted of sessions on the following topics: compilers, information retrieval, mathematical linguistics, design automation, mathematical analysis,

management of applications programming, data communications, approximation techniques, education in information processing, systems planning, formal languages and definitions, linear algebra, analog and hybrid systems, ALGOL 68, numerical analysis, component technology, applications in physical science, management aids, partial differential equations, design of software, real-time search experiments, computer system organization, artificial intelligence, learning and teaching methods, operating system implementation, analysis of computer systems, discrete mathematics, applications in marketing and production, social science applications, mathematical programming, application languages, file structure, theory of computation, memory systems, file management and data banks, automata theory, picture processing, medical applications, computer-aided instruction, etc.

1.13 DIGITAL COMPUTER PROCESSING OF GRAPHICAL DATA

Herbert Freeman
New York Universkty, New York, New York
Grant Numbers AF-AFOSR-152-62, AF-AFOSR-24-63, AF-AFOSR-24-64, AF-AFOSR-24-65, AF-AFOSR-24-66, AF-AFOSR-1367-68

Summary of Recent Research Results

Four distinct problems in computer graphics are being investigated. The first consists of applying linguistic theory to pattern analysis. A pattern analyzer program has been written for which the grammar of a class of patterns can be specified in tabular form and patterns then examined in a manner analogous to the parsing of an English-language sentence. The results provide a basis for classifying patterns, as well as detecting syntactical errors in pattern structures. The work has been applied to the analysis of bubble-chamber photographs, vein structures of tree leaves, and various shaped alphanumeric characters.

The second investigation is concerned with optimal allocation of two-dimensional shapes. This is the problem encountered in sheet-metal layout, land-development, parking-lot layout, etc. Procedures for handling such problems under a wide variety of practical constraints has been developed.

The third problem is that of reconstructing the description of a three-dimensional, planar-faced object from two or more of its perspective projections. Results to date consist of techniques for correlating faces and vertices appearing in different projections, and computer procedures for manipulating the graphical data.

The fourth investigation has the objective of efficiently solving the "hidden-line" problem for solids bounded by quadric surfaces. Procedures have been developed that permit some of the key calculations to be performed in two rather than three dimensions, with significant reductions in computing time. It has been found that knowledge of invisibility can be transferred from one line to an adjacent one in the same manner in which this has been done successfully for planar-faced objects.

Reports and Publications

P. T. Brady, "A Stochastic Model of Message Interchange in a Channel with Transmission Delay." IEEE Transactions on Communication Technology 15, 405-412. 1967. AFOSR-67-2130, AD 658225.

H. Freeman and P. P. Loutrel, "An Algorithm for the Solution of the Two-Dimensional 'Hidden-Line' Problem." IEEE Transactions on Electronic Computers 16, 784-790. 1967. AFOSR-68-0871, AD 674077.

J. Feder, "Languages, Automata and Classes of Chain-Encoded Patterns" August 1967. AFOSR 68-0794. AD 668801.

S. P. Morse, "A Mathematical Model for the Analysis of Contour-Line Data." Journal of the Association of Computing Machinery 15, 205-220, 1968. AFOSR-68-1391. AD 671068.

S. P. Morse, "Computer Storage of Contour Map Data," Proceedings of the National Conference, ACM. 1968. AFOSR 68-2577, AD 678543.

1.14 A STUDY OF DIGITAL ENCODING SYSTEMS

J.B. O'Neal, Jr.
North Carolina State University, Raleigh, North Carolina
Contract Number F44620-69-C-0033*

Summary of Recent Research Results

Shannon's rate distortion bound has been modified to derive a bound on quantizing noise, and Kalman filtering techniques were used for predictive quantizing. Improved error correcting codes were developed for data transmission. Adaptive feedback encoders were examined which detected overload and corrected the condition by adaptive methods such as high information delta modulation and delayed differential pulse code modulation. Using a 31 level quantizer, it was found that adaptive quantization was of little value except for two-level systems and that quantizer expansion proved to be of little benefit. Kalman prediction offered no gain over simple linear predictors. A large vocabulary of words was digitized, and the speech was submitted to an analog computer which produced a tape for suitable input to an IBM 360/65 computer. A set of analog filters was tested for use in simulating a wide variety of systems. Hardware was obtained for tests involving the transmission of video signals over limited bandwidths, and a delta modulator was developed for the audio range. A computer simulation model was encoded which received 50 words, each spoken 10 times, and extracted recognition parameters. An upper bound was developed for the probability of correct classification of speech in pattern recognition, and mathematical models of the peripheral auditory system were created to extract recognition features. The human factors group compiled a word list representing a cross-section of speech sounds in the English language, and they produced a methodology for evaluation of a speech pattern recognition system. The possibility of forming a signal processing laboratory, with emphasis on computer graphics, was examined.

*Supported under Project THEMIS

1.15 MAN-MACHINE INTERACTION IN INFORMATION RETRIEVAL SYSTEMS

G.K. Krulee
Northwestern University, Evanston, Illinois
Grant Number AF-AFOSR-1598-68

Summary of Recent Research Results

Two batch-processing retrieval systems, INFOL and TRIAL, that were previously developed, were made operational for the CDC 6400 as tools for research on remote terminal-based information retrieval systems. INFOL is primarily directed toward structural data storage and retrieval. TRIAL is intended for bibliographic applications in searching and indexing. A preliminary design was developed for a storage and retrieval language and data structure to incorporate features now present in INFOL and TRIAL and additional programming capability for a terminal-based environment. The system is called RIMS (Remote Information Management System). RIMS is being implemented in two phases, a batch-processing phase to experiment with file establishment, data structures, retrieval and report-generation characteristics, and user-language interfaces, and a remote-terminal phase to experiment with browsing capabilities, interactive instruction in system usage, and user interaction with the system. Language specifications are being defined in a modified Backus normal form metalanguage. TRIAL was successfully used in a selective dissemination of information system. It is especially adaptable to large files of bibliographic data from which selective listings or various forms of indexes are needed.

Reports and Publications

L. Borman and D. Dillman, "Trial: An Information Retrieval System for Creating, Maintaining, Indexing, and Retrieving from Files of Textual Information--Users Manual." December 1968. AFOSR 69-0743TR AD 684626

B. Mittman, J. Vallee and R. Chalice, "INFOL for the CDC 6400: Information Storage and Retrieval System." November 1968. AFOSR-69-0755TR AD 684627.

1.16 INFORMATION SYSTEM DESIGN

Morris Rubinoff
University of Pennsylvania, Philadelphia, Pennsylvania
Contract and Grant Numbers AF-AFOSR-190-63,
AF-49(638)-1252, AF-49(638)-1421

Summary of Recent Research Results

Work has centered about three major tasks: (1) rewriting the software in a more universal set of higher symbolic languages; (2) statistical processing of narrative text to derive indexing vocabularies, topics and concept classes within broader subject areas, and real-time determination of searcher interest profile; and, (3) extension of the man-machine conversational medium to an almost unrestricted form of English with a dictionary, grammar, parser, and pragmatic analyser in the computer. The previously computer dependent programs have been rewritten in relatively independent FORTRAN IV with COBOL being employed where necessary in the input-output operations. The mass store is now an IBM 1301 disk storage unit, and the IBM 7040 computer is being replaced by an RCA Spectra 70/46. Extensive data manipulation of narrative text has established a firm statistical base with respect to indexing vocabularies, word clusters that represent topics or concept classes within a subject area, and computer evaluation of user interest profiles in a real-time conversational mode. The original data base consisted of 1500 documents from the Repository of the Association for Computing Machinery, and it has now been extended with magnetic tape entries from the Defense Documentation Center and the Avionics projects from the Moore School. Word clusters are now constructed from both high frequency occurrence of single words and co-occurrence of terms. Real English has far outstripped its predecessor, Easy English. The grammar has been successfully programmed and debugged, and the parser is currently being encoded.

Reports and Publications

H. Cautin, F. Rapp, T.C. Lowe, and M. Rubinoff, "An Experimental On-Line Information Retrieval System," (Scientific Report, May 1967; AFOSR-68-1600, AD 671962).

H. Cautin and F. Rapp, "Description of Easy English," (Scientific Report, Moore School Report No. 67-22, April 1967; AFOSR 68-0545, AD 660569).

M. Fogel, "On-Line Typewriter Access to Classification Tables On Drum Storage," (Scientific Report, April 1967; AFOSR 68-1604, AD 671978).

M. Rubinoff, W. Franks, and D.C. Stone, "Description of an Experiment Investigating Term Relationships as Interpreted by Humans," (Scientific Report, June 1967; AFOSR 68-1599, AD 671906).

J.M. Smith, "An Oral Experiment on Retrieval Dialogue," (Scientific Report, June 1967; AFOSR 68-1836, AD 674058).

J.M. Smith, "A Written Experiment of Retrieval Dialogue," (Scientific Report, August 1967; AFOSR 68-1839, AD 673900).

D.C. Stone, "Word Statistics in the Generation of Semantic Tools for Information Systems," (Moore School Report Number 68-23, Dec. 1967; AFOSR 68-0237, AD 664915).

1.17 SPEECH COMMUNICATION AND AUTOMATIC SPEECH RECOGNITION

June E. Shoup
Speech Communications Research Lab., Inc., Santa Barbara, California

Contract and Grant Numbers AF-49(638)-00492, AF-AFOSR-302-62, AF-AFOSR-595-64, AF-AFOSR-595-65, AF-AFOSR-595-66, AF-AFOSR-1215-67, AF-AFOSR-1252-67

Summary of Recent Research Results

Three major advances in instrumentation for speech analysis were made during this period: (1) the construction of a moderately-priced TV system with light pen which produced a cathode ray tube display of digitized data (i.e a sound spectrogram); (2) the development of a special face mask to obtain simultaneous records of air flow,

subglottal pressure, the acoustic speech wave, and motion pictures of the true vocal folds; and, (3) the development of a program for frequency analysis which is based on the fast Fourier transform (FFT). The FFT has been found very useful in exploring the fricatives and sibilants of English as well as in tracking acoustic parameters of the speech wave. Data is still being collected on French and Surinam phonology. Dr. Huttar completed his study on the relation between acoustic prosodic parameters and the emotional characteristics of speech. It is thought that continued effort in the prosodies is essential if it is to be applied to the problem of automatic speech recognition. The principle efforts continue to be in physiological, acoustic, and perceptual studies with special emphasis in phonology and grammar of the spoken word.

Reports and Publications

R. A. Houde, "A Study of Tongue Body Motion During Selected Speech Sounds." August 1968. AFOSR 69-0318TR, AD 682966.

G. E. Peterson, "The Speech Communication Process." Manual of Phonetics 7, 155-172. 1968. AFOSR 69-0346TR, AD 682869.

G. E. Peterson, "The Relation of Descriptive Phonetics to Research in Speech Communication." Glossa 2, 1-10. 1968. AFOSR 68-2052, AD 677178.

1.18 DATA COLLECTION REDUCTION ANALYSIS AND DISSEMINATION

DeWitt O. Myatt and B. K. Farris
Science Communication Inc., Washington, D. C.
Current Contract Number* F44620-67-C-0022

Summary of Recent Research Results

The study established how the various types and forms of data were

*This research was conducted under the auspices of the National Systems Task Group of the Committee on Scientific and Technical Information of the Federal Council for Science and Technology. Funds were provided by the Advanced Research Projects Agency; it is being monitored administratively by the Directorate of Information Sciences of AFOSR.

acquired, stored, retrieved, packaged, and disseminated for various specific types of users. Further, it ascertained why these packaging methods were adopted, what changes in storage, retrieval, packaging, and dissemination would be most useful in the near future. The final report was set forth in three volumes. Volume I presented a plan for the study and implementation of national scientific and technical data systems concepts. Volume II presented scenarios of data activities in ten selected fields of science and technology covering characteristics of data, data flows, formal data efforts, and representative data-related problems. Volume III was a preliminary census of 226 formal data efforts which were representative of those currently operating in the U.S., including data service centers, data-document depositories, data program development and coordination and data handling service organizations.

Reports and Publications

D.O. Myatt, and B.K. Farris, "Study of Scientific and Technical Data Activities in the United States," Volume I (AFOSR-68-1403, Pt 1; AD670606), Volume II Part A&B (AFOSR-68-1403, Pt. 2; AD 670607) Volume II Part C (AFOSR-68-1403, Pt. 3; AD 607608).

1.19 UNIFICATION OF THEORY AND EMPIRICISM IN INFORMATION RETRIEVAL

Robert A. Fairthorne
State University of New York at Albany, Albany, New York

Summary of Recent Research Results

The strategy has been to isolate and identify essential and controllable activities within the observable field of information. In parallel, and outside the field, these activities identify and isolate the agents that superficially are controllable by those within it, but in fact are not. This double approach to the problem expurgates the IR field of irrelevant concepts and impossible goals. This leaves a minimal structure that can be meaningfully approached and modified as required to attain reasonable ends. Earlier researches showed that IR activities were compounded from twenty basic activities. Each of these irreducible activities used three of the generalized elements Message, Source, Destination, Code, Channel, and Designation. The subject of IR has

been found to involve fifteen triads which do not contain a combination of Destination and Message. Identification of these triads with particular library, communication, and managerial activities has brought out the fundamental nature of these triads. Distinct cases are produced by varying which of the elements was taken as given and whether they were interpreted in their extensional or intensional senses, and this has brought out essential unities in many diverse managerial and informational activities. It has been found that Designation is meaningful only in association with two of the other five elements and, in Information Retrieval, only in association with Destination and one of the other four, allows analysis of the distinct types of "aboutness" of value for IR. The Principles of Discursive Impotence, Ignorance, and Delegation have been invoked to clearly show what areas of future research are reasonable and which elements will be measurable.

Reports and Publications

R.A. Fairthorne, Critique of Soergel, D. "Remarks on Information Languages" in International Symposium on Relational Factors in Classification, June 1966, University of Maryland. Information Storage and Retrieval. 3, 4, December 1967, pp 293-294.

R.A. Fairthorne, Essay-review of Farradane, et al., "Report on Research in Information Retrieval by Relational Indexing....." J. Documentation, 24, 2 June 1968, pp 127-131.

R.A. Fairthorne, "The Limits of Information Retrieval," J. of Library History, Philosophy, and Comparative Librarianship. 3, October 1968, pp 363-374.

R.A. Fairthorne, "Functional Analysis of Information Retrieval," Final Scientific Report, School of Library Science, State University of New York at Albany (August 1968; AFOSR-68-2198, AD 677289).

R.A. Fairthorne, "The Scope and Aims of the Information Sciences and Technologies," Invited paper for the Committee FID/RI, 34th Meeting of F.I.D., Moscow, USSR (to be published). AFOSR 68-2484, AD 677858.

1.20 COMPENDIUM OF THE DISTRIBUTIONS OF MATHEMATICAL STATISTICS AND APPLICATIONS

Samuel Kotz
Temple University, Philadelphia, Pennsylvania
Grant Number AF-AFOSR-1411-68

Summary of Recent Research Results

A comprehensive compilation of statistical distributions is being made that provides a reference source for research in statistics and statistical applications. A compendium for discrete distributions was completed; that for continuous distributions is being completed. Approximately 40 major distributions and over 400 individual distributions are being analyzed. The description for each assesses the properties of estimated parameters and provides examples from major areas of application. The volume of information for continuous distributions necessitates separate reporting of univariate and multivariate distributions. The compilation has required the resolution of classification problems and is leading to the consideration of a system for storage and retrieval and a procedure for updating the corpus.

Reports and Publications

N. L. Johnson and S. Kotz, "Tables of Distributions of Quadratic Forms in Central Normal Variables I," (Inst. of Statistics Mimeo Series No. 557, December 1967; AFOSR-68-0849, AD 668479.

N. L. Johnson and S. Kotz, "Tables of Distributions of Quadratic Forms in Central Normal Variables II," (Scientific Report, Inst. of Statistics Mimeo Series No. 543, Dec 1967; AFOSR-68-0793, AD 668480.

(A shortened version of these Tables with an extensive introduction is being published in Sankhya, Ser. B, early in 1969)

S. Kotz, Invited paper on Twenty-five years of progress in Information Theory, (35 pp.), prepared for the University of California Centennial Volume (jointly with Professor J. Wolfowitz) Aug-Sep 1968. AFOSR 69-0037TR, AD 681447

S. Kotz and Srinivasan, Report entitled "Distribution of Product and Quotient of Bessel Function Variates," (18 pp.), jointly with Dr. Srinivasan) Temple University Research Series, 1968, AFOSR 68-2182, AD 677288. (Submitted for publication in the Annals of the Institute of Statistical Mathematics).

1.21 AUDITORY SIGNAL DETECTION, CORRELATION AND TRANSMISSION

James P. Egan
University of Washington, Seattle, Washington
Grant Number AF-AFOSR-1631-69
(Continuation of research at the University of Indiana under Grant AF-AFOSR 67-0548A)

Summary of Recent Research Results

There are two functional relations that are particularly informative in the determination of the properties of a detection system. One of these relations is called the receiver-operating characteristic, or ROC. The ROC shows the relation between the observer's proportion of the "correct acceptances" and his proportion of "incorrect acceptances" with the stimulus conditions (signal and noise) held at constant average values. The second relation is called the psychometric function. In the context of signal detection and biaural unmaksing, such a function shows the relation between detection performance (such as per cent correct) and the energy of the signal.

It is known from signal detection theory that both of the functions described above (ROCs and psychometric functions) take on different forms depending upon a number of factors, such as the statistical properties of the signal and of the noise, the filter bandwidth, the method of processing, etc. Therefore, ROCs and psychometric functions were determined for a variety of conditions of masking some of which show large amounts of binaural unmasking.

Receiver-operating characteristics were determined by D. S. Emmerich (Ph. D. dissertation) with three different interaural conditions. Three signal levels were employed for each of these interaural conditions, and data were obtained for four listeners. The smooth curves that were

fitted to the data were based upon a particular model for the ROC. This model assumes that the ROC is generated by varying a cut off along a decision axis upon which are plotted two normal density functions.

Reports and Publications

David S. Emmerich, "ROCs Obtained With Two Signal Intensities Presented in Random Order, and a Comparison Between Yes-No and Rating ROCs" Perception and Pscychophysics, 3 35-40, 1968. AFOSR 68-2678, AD 679223.

William A. Lindner, "Recognition Performance as a Function of Detection Criterion in a Simultaneous Detection-Recognition Task." Journal of the Acoustical Society of America 44, 204-211, 1968. AFOSR 68-2676, AD 680056.

Dennis McFadden, "Masking-Level Differences With and Without Interaural Disparities in Masker Intensity." Journal of the Acoustical Society of America 44, 212-223. AFOSR 68-2677, AD 679221.

Section 2

LANGUAGE AND LINGUISTICS

Communication is carried out between two entities through the use of language. The channel employed by a given language need not be two-way, but to be meaningful it is necessary that the intelligence be detectable by the recipient. The form of the intelligence is usually a mutually understood code consisting of one or more elements which may be impressed either serially or in parallel, with respect to time, upon the receiver. Such a concept is relevant to machine as well as man.

Information sciences is involved with basic research leading to a more efficient transfer of information in such significant diads as man-man, man-machine, and machine-machine. Efforts to define an abstract family of languages promise to identify fundamental principles leading to a better understanding of both internal and external machine communications. Studies in improved translators and compilers suggest that a user may soon tailor a general purpose, high-level symbolic language to his specific needs rather than be limited to the instruction set and macros supplied by a vendor.

The increase in information and general scientific data is following the earlier predictions, and the need for more automatic techniques to process this data is not being met. Research is continuing in semantics to provide a formal structure and powerful algorithms to assist in better abstracting and indexing. This area of interest may already have reached the point where the shortcomings of the indexing and abstracting algorithms are offset by the obvious advantages of a consistent set of rules for both information storage and retrieval.

2.1 INFORMATION PROCESSING MODELS AND COMPUTER AIDS FOR HUMAN PERFORMANCE

John A. Swets
Bolt, Beranek and Newman
Current Contract Number* F44620-67-C-0033

Summary of Recent Research Results

Analyses have been conducted on the processes by which rules of a second language are learned to enable the design of teaching procedures and to determine the feasibility of developing an advanced computer-based system for computer-aided language teaching. The research has been extended to include investigations of man-computer system dynamics, the learning of programming languages, and mechanisms of human processing of semantic information. Measures of dynamic system performance and quality are being developed that will enable prediction of user demands on man-machine systems, system response capabilities, and system acceptability to users. A programming language teaching experiment is being conducted to elicit information on formal reasoning and problem solving procedures. A computer model of the reading comprehension process and of the organization and use of semantic memory is being used to study cognitive processing and information retrieval.

2.2 FORMAL ANALYSIS OF NATURAL CONVERSATION IN THAILAND

Michael Moerman
University of California, Los Angeles, California
Current Grant Number* AF-AFOSR 68-1428

Summary of Recent Research Results

Tape recorded natural conversations in the Lue dialect of Thai are analyzed by procedures that do not depend on knowledge of a real

*This research effort is supported by the Advanced Research Projects Agency (ARPA) and monitored, both technically and administratively, by the Directorate of Information Sciences of AFOSR.

world external to the conversations analyzed. Results show an orderliness of conversation that makes apparent the influence of the conversational situation features to which members are shown to orient themselves. Two classes of such features are conversational sequencing (tieing rules) and conversational activities (taking sides). Lue utterances often lack pronouns or other lexical indicators of actors. In both Lue and English, and presumably other languages, linguistic resources are insufficient for the conversational task of aligning actions and actors. To accomplish this task, members additionally use shared and sactioned knowledge of the social world. Some of this knowledge appears capable of being analyzed as category-bound activities and as context-bound typifying ascription. Both of these concepts make use of the notion of categorization labels. The utility of these concepts for the analysis of conversation and interaction by both Lue and Americans is illustrated in the analysis. The social sciences in general, and ethnography in particular, were also studied. Results suggest that rules of correctness are insufficient to account for the member knowledge that produces social action. Rules of correctness are sometimes unnecessary. Members sometimes miscategorize, presumably deliberately, to accomplish insult, praise, etc. The use, social consequences, member knowledge, participant analysis, and other components of the meaning of a word are all heavily influenced by the setting and, particularly, by the sequence of tied utterances.

Reports and Publications

M. Moerman, "Analysis of Lue Conversation: Providing Accounts, Finding Breaches, and Taking Sides," (Final Scientific Report, August, 1968; AFOSR-68-1942, AD 676009).

M. Moerman, "Being Lue: Uses and Abuses of Ethnic Identification," (Center for Southeast Asia Studies Reprint No. 275, Sept 1968) AFOSR 68-2183, AD 677145.

2.3 CRITERIA FOR EXCLUSIONS AND CORRELATIONS FOR AUTOMATIC LANGUAGE ANALYSIS

Jehane B. Burns
University of Cambridge, Cambridge, England
Current Grant Number AF-EOAR-29-67

Summary of Recent Research Results

An automatic procedure was developed that associates formulas exhibiting content structure with natural language English sentences. Data input was derived from an existing automatic syntax analyzer. The formulas that have been produced are intended as an autonomous semantic component of the analyzer. The formulas behave as a semantic control and reduce multiple outputs from the analyzer. The formulas are also potentially of relevance to problems involving the automatic handling of data from natural language text.

Reports and Publications

Jehane Burns, "A Semantic Component in Automatic Sentence Analysis," December 1968. AFOSR 69-0727TR, AD 684589.

2.4 PROBLEM SOLVING AND PATTERN RECOGNITION LANGUAGES

Ranan B. Banerji
Case Western Reserve University, Cleveland, Ohio
Grant Numbers AF-AFOSR-0125-63, AF-AFOSR-0125-64, AF-AFOSR-0125-65, AF-AFOSR-0125-67

Summary of Recent Research Results

A mathematical theory is being formulated that clearly indicates bases for the success of game-playing computer programs. Methods are being developed for characterizing the properties of description languages that are capable of evaluating a language's effectiveness for a given problem. The methods enable the analysis of well-known problem solving procedures in terms of their effectiveness and efficiency. Several new procedures have been developed for solving a new class of competitive games. Parts of a pattern recognition language that is capable of self-modification have been implemented on a computer. A tree structure has been formulated that is useful in storing the descriptions of concepts, not only in terms of the names of the properties and values but also in terms of the descriptions of properties and values. Game decomposition research has led to the identification of necessary and sufficient conditions for the decomposability of unlabelled graphs. An efficient search strategy has been found for proving theorems by the refutation method.

Reports and Publications

R. B. Banerji, "Some Results in a Theory of Problem Solving," (Scientific Report Number SRC 126-A-67-59, November 1967; AFOSR 68-0346, AD 666043).

R. B. Banerji, "A Language for Pattern Recognition." Pattern Recognition, 6, 63-74, 1968. AFOSR 68-2003, AD 676800.

R. B. Banerji, "Some Results in a Theory of Problem Solving (II)." November 1967. AFOSR 68-2083, AD 676293.

R. H. Sherman, "A Model of Concept Learning," June 1968. AFOSR 68-2084, AD 676294.

2.5 A FORMAL ANALYSIS OF NATURAL AND ORAL VERBAL COMMUNICATION

Allen Barton and E. A. Schegloff
Columbia University, New York, New York
Current Contract Number* F44620-68-C-0040

Summary of Recent Research Results

Sequencing rules are being developed that describe the patterns of communication and information transmission used by members of society in normal and emergency situations. A distribution rule has been formulated that is being used to discover the resources it provides for keeping track of the developing course of a conversation. It has been observed, for example, that the structure of summons-answer sequences is more constraining than the structure of question-answer sequences. A property directly related to the non-terminality of summons-answer sequences is their non-repeatability. The non-terminality property also arises with misunderstandings of the use of names. The property of conditional relevance is being applied for a

*This research effort is supported by the Advanced Research Projects Agency (ARPA) and monitored, both technically and administratively, by the Directorate of Information Sciences of AFOSR.

rigorous description of items as sequenced pairs, and for a rigorous account of the absence of an item. Patterns are being developed that show features of orderliness, the coordination within conversations, and characteristics of constraints and the attitudinal availability of participants.

2.6 SURVEY OF ON-LINE COMPUTER LANGUAGES AND SYSTEMS

L.C. Clapp
Computer Research Corporation, Newton, Massachusetts

Summary of Recent Research Results

A survey of 500 on-line computer systems and languages was conducted, and a series of programs were developed to access and manipulate the resulting data base. The programs were used to produce tables and charts for reporting purposes. A study is being completed on an analysis of techniques being employed in the development of on-line systems. A comparative study is being made of languages for handling problems involving large data bases for information storage and retrieval. Procedures for testing the languages are being devised and applied to the languages that will enable their analysis and comparison. The programs were written in SNOBOL and are the nucleus of a specialized data management system that can be easily modified, searched, added to or deleted from through a text-editing language in a time-shared system.

2.7 AUTOMATIC ENGLISH SENTENCE ANALYSIS

Ernst Von Glasersfeld
Georgia Institute for Research, Inc., Athens, Georgia
Grant Number AF-AFOSR-1319-67

Summary of Recent Research Results

The Multistore procedure (automatic parsing program for English sentences) has become fully operational on an IBM 360/65 computer. Its output has been developed into a graphic display of sentence

structure which presents the parsings in a way equivalent to tree diagrams. The system now handles sentences up to a length of 16 words and its capacity to deal with punctuation marks has been increased: it can now handle question marks and certain types of comma. Several kinds of syntactic restraint (i.e. rules which prevent output of word combinations which, although syntactically and semantically acceptable as such, are not relevant in the interpretation of a given sentence) are now operative. Owing to the complete implementation of the significant address method in the machine program, processing times have been greatly reduced; sentences up to 16 words long are now processed in 0.5 to 2.0 seconds. The results of previous preposition analysis have been partially implemented, and a detailed study of infinitive-governing adjectives and clause-governing verbs has been made. Work is currently underway to efficiently employ the Multistore parser.

Reports and Publications

E. V. Glasersfeld and P. P. Pisani, "The Multistore System MP-2," November 1968. AFOSR 69-0467TR, AD 684165.

Brian Dutton, "An Introduction to the Theory and Practice of Correlational Grammar," October 1968. AFOSR 69-0509TR, AD 684166.

2.8 CONTRIBUTIONS TO AN UNDERSTANDING OF THE LANGUAGE COMMUNICATION PROCESS

Y. Bar-Hillel
Hebrew University, Jerusalem, Israel
Contract Number F61052-68-C-0036

Summary of Recent Research Results

Examination of the process of understanding as it occurs in communication through natural languages is showing that the essentially pragmatic character of natural languages produces partial interpretation. Two studies are being made into this aspect of language. The first is examining the validity of the distinction raised by Chomsky between

the theory of competence and the theory of performance. The second is concerned with the problem of subjunctive and counterfactual conditionals, explicitly based on strictly epistemic concepts. Papers have been written on dictionaries and meaning rules and on the structure of language. Papers on applications of epistemic logic and a taxonomy of conditions are in the process of being completed.

2.9 A FORMAL THEORY AND ALGORITHMS FOR INTELLIGENT ACTIVITIES

E. Mark Gold
Institute for Formal Studies, Los Angeles, California
Previous Grant Numbers AF-AFOSR 856-65, AF-AFOSR 856-66; Current Contract Number F44620-67-C-0018

Summary of Recent Research Results

Three approaches were investigated toward the development of an artificial sensory learner: enumerative learners, grammar learners, and parametric network learners. Grammar learning appears the most promising approach. Parametric network learning appears the most descriptive for biological sensory learning, but it appears necessary to express this learning in terms of rules of grammar learning to understand the operations involved. A negative result was obtained in an attempt to apply the enumerative learning method to the construction of a parametric network learner. A data reduction technique has been hypothesized that uses arbitrary functions of experimental results as state variables; an experimental result is a real-valued function of the state of the system. It is assumed that the system can be described by a finite number of state variables. The technique was applied to homogeneous and inhomogeneous ensembles of deterministic and semi-probabilistic systems. The technique was explored for applicability to learning to understand (not speak) a foreign language.

Reports and Publications

E. M. Gold, "Language Identification in the Limit," Information and Control, Volume 10, Number 5, pp. 447-474 (May 1967; AFOSR 67-2815, AD 662826).

Yorick Wilkes, "Computable Semantic Derivations and Their Applications to Natural Language Texts." January 1968. AFOSR 68-0796, AD 668556.

2.10 CROSS-CULTURAL COMMUNICATION THROUGH NONVERBAL BEHAVIOR

Paul Ekman
Langley Porter Neuropsychiatric Institute, San Francisco, Calif.
Current Grant Number* AF-AFOSR-1229-67

Summary of Recent Research Results

Nonverbal behavior data have been acquired from subjects in Brazil, New Guinea, and Japan, as well as in the United States, and are being comparatively analyzed with respect to emotions, traits, attitudes, personalities, etc. A study of the gestures in the different cultures is being made, that is, testing hypotheses about the basis for pan-cultural gestures, i.e., movements that convey the same specific message across literate and preliterate cultures. A glossary of pan-cultural gestures is being compiled. Nonverbal reactions to stress in isolation and interaction are being analyzed from videotapes of Japanese subjects; tapes for comparative analyses of American subjects will be made. Videotapes, measures of simultaneous heart rate and galvanic skin response, and verbal reports have been acquired for a study of the information value of nonverbal behavior in comparison with other channels of communication. This will be a two-culture study of Japanese and Americans.

2.11 MODELS FOR ADAPTIVE SYSTEMS AND HUMAN COMMUNICATION SYSTEMS

David Rothenberg
New York Research Group, Inc., New York, New York
Contract and Grant Numbers AF-AFOSR-0881-65, AF-49(638)-1738, AF-AFOSR-1596-68

*This research effort is supported by the Advanced Research Projects Agency (ARPA) and monitored, both technically and administratively, by the Directorate of Information Sciences of AFOSR.

Summary of Recent Research Results

A pattern recognition model has been developed that maps a possibly infinite and continuous space (over which sensory stimuli range) into a finite space of discrete points, the classification of such stimuli. The input to the model is a partial preordering of points in the former space together with criteria that determine the number of classifications required. The model chose a finite subset of the stimulus space that defines a function, and a range about each element of the subset such that the union of all such ranges is maximal and the classification of pairs of points contained in such union is unambiguous. Restrictions on the choice of such subset derives from the limitations in information-carrying capacity of the resulting classification system. Successful applications to human auditory perception have been made, and visual applications are in progress. A model for feedback-controlled data reduction has also been developed in which input data sequentially presented through a finite number of input channels is represented by a hierarchic structure of property-derived classes. Feedback is applied to a synthesis of the input data generated by the model, which also has the capacity for interrogation of the trainer. Methods for task-oriented property generation are also being developed.

Reports and Publications

J. Korein, A. L. Bender, D. Rothenberg and L. J. Tick, "Computer Processing of Medical Data by Variable-Field-Length Format." Journal American Medical Association 196, 957-963. 1966. AFOSR-69-0131TR, AD 682113.

2.12 LANGUAGES FOR AUTOMATIC PROGRAMMING TOOLS

Alfonso Caracciolo di Forino
University of Pisa, Pisa, Italy
Current Contract Number F61052-67-C-0097

Summary of Recent Research Results

Basic concepts are being assessed to provide a formal definition for programming languages in general and for direct machine tool languages in particular. An attempt is being made to establish a

unique methodology for the various programming languages that range from high-level algorithmic languages like Fortran, Algol, and Simula to basic computer languages and machine tool languages. A rough model for machine tool languages has been developed that describes a language by a recursive function that yields, for each input tape, initial state, and initial workpiece configuration, the final state and the final workpiece configuration. The model has suggested a different approach from that of APT, and has been used in a detailed study based on extended Markov algorithms that coincides with ideas developed for the application of symbol manipulation languages. A method has been proposed for developing a formal definition of APT-like languages that incorporates an explicit definition of the rules of interpretation of the language. The method proposes defining a sequence of structural transformation rules to be applied to an abstract string resulting from the concatenation of two parts, the first representing the program to be interpreted and the second the present machine state that includes all relevant information at any given stage of the interpretational process and the present workpiece configuration. The algorithm is complex, and there is no unique way of defining it.

Reports and Publications

A. Caracciolo and A. Camera, "On a Formal Definition of Direct Machine Tool Languages." July 1968. AFOSR 68-2770, AD 679568.

2.13 PROBLEMS IN MACHINE PROBLEM SOLVING

Saul Amarel
Radio Corporation of America, Princeton, New Jersey
Contract Numbers AF 49(638)-1184, F44620-68-C-0012

Summary of Recent Research Results

Detailed studies of relationships between choice of problem representation and the relative success of problem solving by machine suggest that certain appropriate shifts in problem representation can be mechanized via the utilization of some types of information about the problem in the formulation of improved reduction procedures

for its solution. Work on a conceptual framework for problem solving has shown the theoretical and practical significance of considering several alternative modes of describing a solution in the problem solving process. Points of contact have been established between concepts in problem solving and problems in program semantics, computer linguistics, and advanced information systems. The problem of proving the correctness of programs was explored from an algebraic point of view; ideas were tried successfully in the proof of correctness of a compiler for expressions. Research on the formulation of a representation system for modeling parsing procedures provides a conceptual unification of existing processes for syntactic analysis and suggests new types of parsers. An extension of this research to deterministic parsers has clarified certain essential aspects of their organization. A formal system for describing computer language translations was specified with promising preliminary results. Recent studies of question-answering systems suggest ways of extending the domain of discourse of a system in an evolutionary way.

Reports and Publications

S. Amarel, "An Approach to Heuristic Problem Solving and Theorem Proving in the Propositional Calculus," in Systems and Computer Science (University of Toronto Press, Canada, 1967; AFOSR 68-0689, AD 667505) pp 125-220.

S. Amarel, "On Representations and Modeling in Problem Solving and on Future Directions for Intelligent Systems" (Scientific Report Number 2, June 1968; AFOSR 68-1580, AD 671992).

S. Amarel, "On Representations of Problems of Reasoning About Actions," in Machine Intelligence 3, D. Michie, ed. (Edinburgh University Press, Scotland, 1968; AFOSR 68-1327, AD 670558) pp 131-171.

S. Amarel, "Problem-Solving Procedures for Efficient Syntactic Analysis." May 1968. AFOSR 68-1305, AD 670575.

M. Kochen, "On the Representation of Limited Information By Means of Pictures, Trees, and English-Like Sentences," (Scientific Report, May 1968; AFOSR 68-1343, AD 670555).

M. Kochen, "Automatic Question-Answering of English-Like Questions About Simple Diagrams" (Scientific Report, May 1968; AFOSR 68-1303, AD 670545).

M. Kochen, "Automatic Answering of English-Like Questions About Arithmetic," Nov 1968. AFOSR 69-0272TR, AD 682339.

P. J. Landin and R. M. Burstall, "Programs and Their Proofs, An Algebraic Approach." Nov 1968. AFOSR 69-1235TR, AD 687498.

D. A. Walters, "A Representation System for Parsing Procedures" (Scientific Report, May 1968; AFOSR 68-1301, AD 670577).

D. A. Walters and A. J. Korenjak. Sept 1968. AFOSR 69-0086TR, AD 681131.

R. O. Winder, "An Evaluation of Heuristics for Threshold-Function Test-Synthesis" (Scientific Report, May 1968; AFOSR 68-1297, AD 670556).

2.14 SELF-EXTENDING COMPUTER LANGUAGES AND THEIR TRANSLATION

Richard K. Bennett
Signatron, Inc., Lexington, Massachusetts
Contract Number F44620-68-C-0007

Summary of Research Results

This effort has resulted in the development of BUILD (Base for Uniform Language Definition) which is an initial meta-language providing for artificial language definitions. The artificial languages so created are highly efficient in that they are tied directly to the mechanisms of the translator process. This has been accomplished by isolating and utilizing the fundamental elements of computer languages in general together with the fundamental mechanisms of assemblers and compilers.

This has resulted in a versatile, powerful base from which specific languages and translators can be developed. Thus, advanced programmers, employing the techniques given in the BUILD I Manual, could produce languages specific to such diverse interests as mathematicians, engineers, operations analysts, and library scientists. It should be clearly pointed out that only the user dictates differences between language levels. BUILD recognizes no difference between a meta-language and an object language except at the direction of the user.

Reports and Publications

R.K. Bennett, "BUILD ... A Base For Uniform Language Definition," Final Technical Report, (1968; AFOSR 69-0072TR, AD 681169).

2.15 THEORY OF PROGRAMMING (ALGORITHMIC) LANGUAGES

Seymour Ginsburg
System Development Corporation, Santa Monica, California
Grant Number AF-AFOSR-1203-67

Summary of Recent Research Results

Effort continued to examine an abstract family of languages in an attempt to identify and isolate principles germane to the information sciences. It has now been shown that the abstract family of languages is extremely important in that it unifies the theory, allows algebraic proofs to replace machine arguments, and provides the basis for new development. Specifically, isolated facts about closure properties and many diverse proofs now are part of an algebraic theory with many special cases. This study has uncovered numerous new facts on substitution, intersection, and hierarchy relationships among well-known families of languages. A grammar for such a language has been formulated, and new rules, as well as new ways to use old rules, have been found. This has culminated in a mathematical model of transformational grammars. Different devices were examined, and the related languages were considered. In particular, the elimination of end-markers for two-way devices was studied in depth.

2.16 STATUS OF COMPUTATIONAL LINGUISTICS AND RELATION TO MACHINE TRANSLATION

Winifred P. Lehmann
University of Texas, Austin, Texas
Grant Number AF-AFOSR-1320-67

Summary of Recent Research Results

Resulting monograph will be published as a book under the title Machine Translation Today. The mechanical translation systems that have been developed and put into use are described in these categories: (1) Systems based primarily on syntactic descriptions, (2) Systems based on semantic as well as syntactic descriptions, and (3) Long-range research programs. The underlying aim is to present a statement which will provide for non-specialists a general view of the undertakings, accomplishments and future work in mechanical translation. This project has also permitted the establishment of a comprehensive bibliography in machine-retrievable form.

2.17 COMPARATIVE ANALYSIS OF LANGUAGES FOR MACHINE PROCESSING

Ernesto Zierer
Universidad Nacional de Trujillo, Trujillo, Peru
Grant Number AF-AFOSR-1470-68

Summary of Recent Research Results

English, German, Spanish, and Japanese are being analyzed to establish equivalences among the particles in each language. The semantic analysis of 110 occurrences of constructions with 'de' in written and spoken Japanese shows 18 different equivalents in English, 19 in German, and 15 in Spanish, indicating the wider range in the Japanese. Matrix and binary tree displays are being developed for the correlations. An algorithm for generating the particle occurrences is being formulated. For the particle 'ni,' 115 occurrences were found in Japanese that yielded 20 equivalents in English, 18 in German, and 19 in Spanish. The particle was found to express 31 different relations. Semantic analysis of 95 occurrences of the

Japanese 'e' showed 8 equivalents in English, 6 in German, and 3 in Spanish. The particle 'wo' is presently being studied. These results have implications both in understanding communication in these languages and in their machine manipulation.

2.18 GENERALIZATION AND EXTENSION OF PROGRAMMING LANGUAGE CONCEPTS

Philip R. Bagley
University City Science Center, Philadelphia, Pennsylvania
Current Contract Number F44620-67-C-0021

Summary of Recent Research Results

An analysis of concepts employed in the construction of computer programming languages was made for insight into ways of designing new languages having increased flexibility and complexity. Such concepts as variables, procedure call mechanisms, and program sequence controls were examined. A technique of expressing data values, data elements, and data structures was developed that provides for the construction of arbitrarily complex data elements and for arbitrarily chosen relationships between data elements. All expressions in a program that cause the language processor to take some action (including declarations) are viewed as transformations (procedures). A basic set of these transformations is proposed. The two main classes of transformations are transformations of data that create, test, modify, and destroy data elements, and transformations of sequence control, including control of iteration and of conditional execution. Functions are definable in two ways, in terms of other transformations and by enumeration in the form of function tables. The analysis suggests that more freedom of storage organization is needed in machine design that may be provided, in part, by large-scale associative memories.

Reports and Publications

P.R. Bagley, "Extension of Programming Language Concepts." November 1968. AFOSR 69-0023TR, AD 680815.

Section 3

BIONICS

This section is called "bionics" because bionics is a COSATI subject heading, not because we necessarily believe that it is politically wise nor scientifically accurate. Much of the work included in this section was reported under the rubric "Adaptive and Self-Organizing Systems" in previous volumes of this report.

The term "bionics" was originally coined to represent an area of scientific exploration between the formal disciplines of biology and electronics. The term "cybernetics", defined by Wiener as "the science of control and communication, in the animal and the machine," also claims sovereignty over this domain. Regardless of the label, such elements as coordination, regulation, transmission, feedback, and control are fundamental to any serious work in the information sciences.

The literature abounds with examples of accomplishment brought about by study in one field of science which significantly advanced the frontiers of another; infinitesimal calculus and astronomy, the virus and the protein molecule, the chromosomes and heredity, etc. The information sciences are continually searching for more fundamental and general algorithms by which progressively more and more complex systems may be studied. Such systems are no longer man or machine, rather they are man-machine.

Even now the first interplanetary spaceship is in the design phase. The research in bionics has already paved the way by providing a firm basis from which answers to many basic questions can be formulated. What portion of the duties aboard the spaceship should be assigned to the crew and what to the computer? What is the most efficient way for humans to talk to the computer, and what is the best way for the computer to respond?

3.1 OPTIMIZATION OF LEARNING NETWORKS FOR PATTERN CLASSIFICATION

L. A. Gerhardt
Bell Aerosystems Company
Contract Numbers AF-49(638)-1449, AF-49(638)-1627

Summary of Recent Research Results

The application of cascaded linear adaptive nets to the analysis of mixtures of prime patterns was shown to yield solutions for all linear mixtures. These solutions were demonstrated to be stable in the presence of general noise and therefore preferable to any solution by single layer nets. In addition, the training sets necessary to obtain a solution were shown to require only an analyzed set of independent mixtures and not pure samples of each constituent pattern. It was shown that linear nets with parameteric error-correcting training yield the best estimate of a random output in terms of a set of random inputs, and, by non-parameteric techniques, the best estimate of the mean value and variance of the output in terms of a set of independent inputs. A general adaptive non-deterministic search procedure has been developed that is applicable to the abstract network structures. It was shown to be an improvement over other presently known methods because of its compatibility between high speed and accuracy that is not found in other methods. Results of computer experiments are presented supporting the mathematical conclusions. The research is continuing with particular emphasis on the classification of patterns from classes with unknown statistics including the distribution shape as well as its parameters.

Reports and Publications

Johannes G. Goerner, Lester A. Gerhardt, K. W. Drake, J. A. Cadzow, and F. D. Powell, "The Influence of Network Structure on the Performance of Learning Systems" (BAC Scientific Report Number 9500-920089, October 1967; AFOSR-68-1350, AD 670845).

F. D. Powell, "On Adaptive Nets" (presented at Seminar, University of Texas, Austin, April 1967; AFOSR-68-1347, AD 656764).

3.2 INTEGRATION OF THEORY AND EXPERIMENT INTO A UNIFIED CONCEPT OF VISUAL PERCEPTION

Heinz Von Foerster and Humberto S. Maturana
University of Georgia, Athens, Georgia
Contract Number AF-49(638)-1680

Summary of Recent Research Results

Recent anatomical and neurophysiological findings of structure and function of the afferent fibers in vertebrates on various levels of the optic tract are being integrated with theoretical results obtained in the study of computing networks toward the development of a unified concept of visual perception. A theorem has been formulated that specifies the conditions in terms of luminosity vectors under which chromaticity changes. An experiment has been designed that demonstrates the theorem. A mathematical model for the network physiology that underlies the phenomenon is being developed. The present state of the theory of computing neural nets has been summarized. Special attention is being given in model development to the structural and functional properties of networks that compute invariants in the distribution of their stimuli.

3.3 COMPUTER SIMULATION OF NATURAL PATTERN GENERATION PROCESSES

Dan Cohen
Hebrew University, Jerusalem, Israel
Grant Number AF-EOAR-45-67

Summary of Recent Research Results

Toward generating a realistic simulation of the growth of a branching system, various modes of interactions between the branches have been tried. Study is being concentrated on two types of interactions. The first is an influence that spreads along the branches themselves and regulates the spacing between branching points and the dominance relations between branches of different orders. The second is an interaction in the space of the branching system. Any one branch

may shade another branch, or it may consume resources necessary for growth. A simulation of mutual shading is being tried in which growth and branching are influenced by the amount and direction of light at any one point. An approach is being attempted to account for the development of branched vascular systems in a tissue or an organ (or a drainage system) based on the hypothesis that the transport of a substance produced within the system eventually tends to be restricted to a system of branched connected channels because of the self-reinforcing positive feedback effect of the flow of the substance in the channels. A third approach is attempting to account for the size and shape of two dimensional organs in terms of local interactions at the boundary where most of the growth occurs.

3.4 CRITERIA FOR ADAPTIVE SYSTEMS

Gerhard L. Hollander
Hollander Associates, Fullerton, California
Contract Number AF-49(638)-1664

Summary of Recent Research Results

Over 400 reports of diverse approaches to the development of adaptive systems, particularly in areas of control systems, pattern recognition, and heuristics, have been classified into six groups of increasing cost and effectiveness. The classification scheme is multidimensional. The top hierarchic level concerns the goal of the adaptive system that determines the degrees of freedom for self-adjustment and the a priori assumptions about the system and the environment that must be satisfied for the "optimum." The three system goals could be invariant behavior, pre-specified dynamic optimum, and search for system optimum. Thus, if performance does not increase from the simpler to the more "intelligent" systems, the added cost should call the systems into question. These categories could (theoretically) be subdivided based on the adjustment in the basic loop, i.e., no adjustment, open-loop adjustment, and closed-loop adjustment, though the no-adjustment system is not adaptive in the true sense. The classification shows that many adaptive systems are of low order.

3.5 EVOLUTIONARY AND MULTILEVEL INFORMATION PROCESSING NETWORKS

Earl E. Gose
University of Illinois, Chicago, Illinois
Grant Numbers AF-AFOSR-0751-65, AF-AFOSR-0978-65, AF-AFOSR-0978-66, AF-AFOSR-1341-68

Summary of Recent Research Results

A technique for the adaptation of multilevel logical networks of two-state elements was developed. The effects of connectivity, element type, problem type, and the adaptation parameters on the rate of adaptation were determined by computer simulation. A new technique for choosing pattern recognition properties was based on the probability distributions of single properties and their correlations; another sequentially selects additional properties which separate the most often confused classes in many-class categorization problems. A color-selective flying spot scanner was constructed to provide pictorial input to the computer. Programs were written to measure shape, color, and textural properties and to perform filtering and noise-reduction. Their utility was demonstrated in the classification of brain cells, blood cells, and breast tumors, using electron micrographs, microphotographs, and x-rays. The use of weighted sums of random functions in information processing networks was investigated. A theory for the expected error versus the number of functions was derived and verified experimentally. A generalized theory for finding optimum replacement schedules for evolving populations was derived. The replacement cost, value function, and quality distribution of the newly-selected elements may be different in each generation, and the total number of replacements may be constrained.

3.6 ANALYSIS AND SYNTHESIS OF COGNITIVE PROCESSES AND SYSTEMS

Heinz Von Foerster
University of Illinois, Urbana, Illinois
Grant Numbers AF-AFOSR-0007-63, AF-AFOSR-0007-64, AF-AFOSR-0007-66, AF-AFOSR-0007-67

Summary of Recent Research Results

The central issue of this research is the establishment of the epistemological and theoretical framework in which the structural and functional organization of cognitive systems becomes sufficiently evident to become expressible in explicit formalisms that can be implemented in computer hardware and software. A new interpretation of neural mechanisms was advanced that stresses the interpretative aspects of perception and introduces the concept of "experience spaces," for example, depth, color, inter-sensory references, self-reference, and inference. New theorems of information transfer in complex regulatory systems have been developed. Several novel computer program structures called "cylinders" were devised for representing relational structures of arbitrary depth as encountered in the symbol system of natural languages. A program named Lejardin for graphic manipulation at the crt enables complex arrays of graphic elements to be formed and rearranged through no more than variation of numeric parameters in an otherwise fixed composite structure. Progress continues on other projects including measurement of the time dependent parameters of the speech waveform, the application of tensor calculus and differential geometry to neural modeling, and the identification and synthesis of linear automata from their state transition tables and output tables.

Reports and Publications

W. R. Ashby, "Information Processing in Everyday Human Activity." Bioscience 18, 190-192, 1968. AFOSR 68-1521, AD 672395.

H. Bielowski, "An Objective Function for Computer Systems Scheduling" October 1968. AFOSR 68-2583, AD 678164.

R. C. Conant, "Channel Capacity of Moore Automata." Information and Control 12, 453-465. 1968. AFOSR 68-2801, AD 679957.

R. C. Conant, "Information Transfer in Complex Systems with Applications to Regulation" AFOSR 68-0727, AD 667812.

A. Inselberg, H. Von Foerster and P. Weston, "Memory and Inductive Inference," Proceedings of Symposium on Cybernetics Problems in Bionics, pp 31-68. AFOSR 68-2800, AD 677956.

M.R. Gardner, "Critical Degenerateness in Large Linear Systems." June 1968. AFOSR 68-1525, AD 671963.

K. Kokjer, "Cutaneous Channel Capacity." September 1966. AFOSR 69-0034TR, AD 680816.

L. Lofgren, "Emergent Properties of Systems." Proceedings of Hawaii International Conference on System Sciences, January 1968. pp 330-332. AFOSR 68-1696, AD 674693.

Lars Lofgren, "An Axiomatic Explanation of Complete Self-Reproduction," Bulletin of Mathematical Biophysics 30, 415-425, 1968. AFOSR 68-2579, AD 678415.

J.K. Russell, "A Visual Image Processor," IEEE Transactions on Computers 17, 635-639, 1968. AFOSR 68-2802, AD 679750.

J. Schill, "Automatic Determination of Invariance in Machine Coded Speech." December 1968. AFOSR 69-0454TR, AD 683747.

H. Von Foerster, "Numbers of Man," April 1966. AFOSR 69-0035TR, AD 680821.

H. Von Foerster, "Biological Principles of Information Storage and Retrieval," in Electronic Handling of Information Testing and Evaluation (Academic Press, 1967). AFOSR 68-0774, AD 668284.

M.G. Wilkins, "A New Homeostat." June, 1968. AFOSR 69-0316TR, AD 683048.

3.7 INTERNATIONAL WORKSHOP ON PATTERN RECOGNITION

R. Emberson and Laveen Kanal
Institute of Electrical and Electronic Engineers, New York, NY
Contract Number F44620-68-C-0063

Summary of Recent Research Results

A workshop on pattern recognition was held in August 1968 in Delft,

Holland. A substantive summary of the papers and discussion has been issued. The summary classifies the presentations into four categories: character recognition, pattern classification and feature selection theory, picture processing, and other applications. The other applications concern neural networks and crop classification from multispectral remote sensor data. Although work abounds on theoretically optimum classification, and, in particular, stochastic approximation methods, problems of experimental design, sample size, feature dimensionality, system complexity, and design of cascaded decision nets remain essentially open even in the statistical approach to pattern classification.

Reports and Publications

B. Chandrasekaran, L. Kanal and G. Nagy, "Report on the 1968 IEEE Workshop on Pattern Recognition" AFOSR 69-1322TR, AD 638378.

3.8 THEORETICAL MODELS FOR THE RETICULAR FORMATION

William L. Kilmer
Michigan State University, East Lansing, Michigan
Grant Numbers AF-AFOSR-1023-66, AF-AFOSR-1023-67

Summary of Recent Research Results

Two reticular formation models have been built, S-RETIC and STC-RETIC. Despite its 8000-parameter complexity, S-RETIC properly switches from mode to mode as a function of the signals coming in over its 60 binary input lines. It thus serves as a crude model of how, in a fraction of a second, the million or more neurons in a reticular formation reach a workable consensus as to the proper mode of total commitment of an animal. STC-RETIC enlarges the behavioral capabilities of the model to include development, generalization, discrimination, habituation, classical conditioning, extinction, avoidance conditioning, and trial-and-error learning modes of behavior. STC-RETIC is a 500-instruction computer program. Pattern recognition algorithms for translating EEG signals into decisions about the mode of the EEG (resting, hyperventilating, sleeping) are also being studied. A method for determining the sleep

stage of a subject by computer has been developed that employs frequency histograms and Weibull distributions. A modification of STC-RETIC is being developed for segmenting speech signals into a sequence of basic sound units. Several theorems have been derived for designing computer control circuits as a cascade of simpler circuits and for resynchronizing such circuits after noise or power failure has put them in the wrong state.

Reports and Publications

P.S. Ulinski and William L. Kilmer, "On Relating the Behavior of a System to the Behavior of its Constituents A Theory of Cellular Differentiation," (Interim Scientific Report Number 1, January 1968; AFOSR 68-0217, AD 666469).

W.L. Kilmer, W.S. McCulloch and J. Slum, "An Embodiment of Some Vertebrate Command and Control Principles." Currents in Modern Biology 2, 81-97, 1968. AFOSR 68-2603, AD 678386.

W.L. Kilmer, W.S. McCulloch, J. Slum, E. Craighill and D. Peterson, "On a Cybernetic Theory of the Reticular Formation." Cybernetic Problems in Bionics, pp 431-479. Gordon & Breach, New York 1968. AFOSR 68-2642, AD 678387.

B.G. Reynolds and W.L. Kilmer, "Structure of Events and Automata." August 1968. AFOSR 68-2187, AD 677291.

W.F. Cutlip and W.L. Kilmer, "Synchronization and Decomposition of Finite Automata." August 1968. AFOSR 68-2319, AD 677759.

3.9 CONFERENCE ON LEARNING, REMEMBERING, AND FORGETTING

Frank Fremont-Smith

Current Contract Number* F49604-67-C-0154

*This research effort is supported by the Advanced Research Projects Agency (ARPA) and monitored, both technically and administratively, by the Directorate of Information Sciences of AFOSR.

Summary of Recent Research Results

The fourth annual interdisciplinary Conference on Learning, Remembering, and Forgetting (sponsored jointly by AFOSR and the New York Academy of Sciences Interdisciplinary Communications Program) was held October 9-12, 1966 in Pacific Palisades, California. The proceedings of this conference were grouped under the following topical headings: Action contingent development of vision in neonatal animals; Some ways in which experience affects learning; and, The effects of hormones in infancy on central nervous system organization. The thrust of the conference was the elicitation of information on how an organism's capacity was shaped by its experience. The proceedings included discussions on such topics as the relationship between cybernetics and perceptual learning, analysis of behavioral phenotype and genotype, and interactions during sensitive periods. A bibliography was appended.

Reports and Publications

D. P. Kimble (editor), "Experience and Capacity," 4, Proceedings of the Fourth International Interdisciplinary Conference on Learning, Remembering, and Forgetting, Pacific Palisades, California, 9-12 Oct 66 (Final Scientific Report, AFOSR 68-2208, AD 681246).

3.10 ASYNCHRONOUS AUTOMATA AND DISCRETE-TIME SYSTEMS

George A. Bekey and R. B. McGhee
University of Southern California, Los Angeles, California
Grant Numbers AF-AFOSR 1018-66, AF-AFOSR 1018-67

Summary of Recent Research Results

Research in identification of discrete-time systems has concentrated on the problem of determining the unknown sampling frequency, and other parameters, of discrete systems with continuous inputs and outputs. The research has been successfully applied to the synthesis of a sampled data model of a human pilot whose inputs and outputs are a continuous function of time, both in the presence of observational

noise and parameter noise. This is probably the most powerful tool yet utilized for the construction of mathematical models of human operator performance. The mathematical aspects of the theory (stochastic approximation) which relate to the convergence of the identification process have been developed and published. The research in human operator information processing has also resulted in the development of an adaptive finite-state model of the human controller. This promises to be a starting point for significant new investigations in this area.

Reports and Publications

George A. Bekey and E.S. Angel, "Asynchronous Finite State Models of Manual Control Systems," in Proceedings of the Annual Conference on Manual Control, Cambridge, Mass., 1966 (AFOSR 68-0455, AD 666340).

A. A. Frank, "Automatic Control Systems for Legged Locomotion Machines," (U.S.C.E.E. Scientific Report No. 273, May 1968; AFOSR 68-1340, AD 670601).

R. B. McGhee, "Finite State Control of Quadruped Locomotion" in Simulation, Vol. 9, No. 3, pp 135-140 (Sept 1967; AFOSR 68-1320, AD 673577).

R. B. McGhee and A. A. Frank, "Optimum Quadruped Creeping Gaits," in Scientific Report, (July 1968; AFOSR 68-1846, AD 675256).

R. B. McGhee, "Some Finite Aspects of Legged Locomotion," in Mathematical Biosciences, Vol 2, pp 67-84 (1968; AFOSR 68-1764, AD 673912).

W. S. Meisel, "A Numerical Integration Formula Useful in Fourier Analysis," in Communications of the ACM, Volume 11, No. 1 pp 51 (January 1968; AFOSR 68-0653, AD 667321).

W. S. Meisel, "Hazards in Combinational and Sequential Switching," (U.S.C. Electrical Engineering Department Report Number 174, May 1966; AFOSR 68-0370, AD 665864).

L. R. Nardizzi and G. A. Bekey, "Synthesizing Optimal Controls for Modulated Discrete-Time Systems." International J Control 8, 571-590, 1968. AFOSR 69-1595TR, AD 690294.

L. R. Nardizzi and G. A. Bekey, "Optimization of Discrete-Time Systems with Combined Modulation." IEE Transactions on Automatic Control 13, 286-289. 1968. AFOSR 68-2694, AD 679262.

3.11 CYBERNETIC INVESTIGATION OF LEARNING AND PERCEPTION

Gordon Pask
System Research Ltd., Richmond, England
Contract Numbers AF-61(052)640, F61052-67-C-0010

Summary of Recent Research Results

Experiments have continued on developing methods of determining (measuring) the different strategies used by subjects in proceeding through a learning task. A method has been incorporated into a learning model to score for subjective estimates of correct responses, rather than for actual responses, that is applicable to many learning situations in which more than one set of response alternatives can be appropriate. In a teaching system based on the learning model, an executive controls learning and problem-solving activities. The subject, not the executive, through his responses, partitions the task into learnable segments. Two experiments were designed to explore this approach. One develops a code learning skill and the other employs a Telcomp time-sharing terminal. The experiments are intended to measure the efficiency of the strategy selected by the subject in comparison with a known standard, and the degree to which the subject adheres to his selected strategy. The experimental data are being used to acquire a better understanding of the requirements for artifically intelligent cognitive systems. Though the organization of an attention-directing system appears to be the converse of that required for an evolutionary system, this distinction is actually a matter of degree because all subsystems must maintain their activity continuously to enable the executive to sense and evaluate their properties.

3.12 MACHINE LEARNING FOR GENERAL PROBLEM SOLVING

David L. Johnson
University of Washington, Seattle, Washington
Contract and Grant Numbers AF-49(638)-1070, AF-AFOSR-366-62, AF-AFOSR-468-64, AF-AFOSR-468-65, AF-AFOSR-939-65, AF-AFOSR-939-67, AF-AFOSR-939-67B

Summary of Recent Research Results

Several learning system models are being developed that have, as a common aspect, the use of hierarchies of heuristics that govern system algorithms. Three areas of application are being examined: game learning as demonstrated in chess and bridge, learning for network designing, and pattern recognition and classification. Game learning does not aim for "best" performance but for self-improving procedures in actual situations. Results include the development of tree-pruning heuristics, the formation and use of generalizations for decision making in previously unencountered situations, the generation of new decision parameters, and the use of inferential reasoning to modify procedures. Design synthesis research has provided additions to the design of threshold logic networks and knowledge concerning the incorporation of interactive methods of system development and heuristics into a learning model. Threshold theoretic models have been applied to pattern recognition. Starting with a curve-fitting model, heuristics were developed for more general patterns with unspecified distributions and clusters of data points in the pattern space. Programs were formulated that are being used in nuclear engineering research.

Reports and Publications

K. H. O'Keefe, "Modularity in Design: The Application of Shift Registers in Secondary State Assignment." January 1968. AFOSR 68-1300, AD 670550.

Section 4

INFORMATION THEORY

The term, "information theory" has a precise scientific meaning which applies to few, if any, of the efforts herein reported. It is also the only COSATI subject heading in the information sciences which carries any connotation of theory. Hence its use to title this section.

Any machine, which in the broad sense includes man, must make constant use of information. This would naturally include sensing, channeling, integrating, storing, outputting, and feeding back the processed information. Many of the hard, and soft, disciplines have had long standing efforts to attempt to formulate broad, basic structures in information theory. Combinatorial algebras and new advances in logic and mathematical notation may eventually lead to information storage and retrieval techniques which will permit both a significant reduction in hardware storage media and vastly improved cycle speed.

Improvements in information theory already point to better use of research funds, in that recently developed theories and algorithms not only point out the areas wherein the probability of return is highest, but such studies have also pointed out boundary conditions outside of which there is no hope of success. Some of the more productive centers are recursive function theory, combinatorial algebras, group theoretic concepts, self-referential systems, and recursively enumerable sets to name but a few.

Vendors are already sufficiently motivated to support research directed toward improved computer capacity, faster cycle times, and more powerful instruction sets. It is becoming more apparent that research to provide the most efficient use of the hardware will have to be slanted along user interests.

4.1 ALGEBRAIC METHODS IN AUTOMATA THEORY

Marcel P. Schutzenberger and Maurice P. Nivat
Association Pour L'Etude Et La Recherche En Traitement
De L'Information, Paris, France
Contract Number AF-61(052)-00945

Summary of Recent Research Results

Research has continued in three topic areas, algebraic properties of monoids, properties of rational and algebraic subsets of monoids, and the computing power of finite automata. Rational subsets of any commutative monoid have been characterized. Further results have been obtained on the commutative closure of context-free languages. A general scheme is being developed for the syntactic analysis of context-free languages. Many computations on graphs have been found possible through the use of a finite automaton that processes the incidence matrix of the graph or through the use of a number of isomorphic automata attached to each vertex operating simultaneously.

Reports and Publications

M. P. Schutzenberger "On Products of Finite Dimensional Stochastic Matrices" Proceedings of American Mathematical Society 18, 850-853, 1967. AFOSR 68-0369, AD 665618.

M. P. Schutzenberger, "On a Question Concerning Certain Free Monoids" Journal of Combinatorial Theory 1, 437-442, 1966. AFOSR 68-0143, AD-663965.

M. P. Schutzenberger "On Mathematical Research for Artificial Intelligence" Proceedings Conference on Intelligence and Intelligent Systems, U of Georgia, 16 Jan 1967. AFOSR 68-1524, AD 671959.

4.2 COMBINATORIAL ALGEBRA WITH APPLICATIONS TO INFORMATION SCIENCE

F. B. Cannonito
University of California, Irvine, California
Grant Number AF-AFOSR-1321-67

Summary of Recent Research Results

Tietze transformations have been applied to a class of groups which showed them to be "one-sixth" in the Greendlinger sense. The conjugacy problem was then shown to be recursively solvable, and this was extended to a hierarchy of increasingly complex groups with solvable conjugacy problems. Recent findings suggest that separated groups may exist, and that an explicit construction may be possible to produce such a group. Work is underway to develop such a construction. The properties of information retrieval were found to be analogous to solving the word problem in semigroups, and the recent increase in ability to handle more complex semigroups suggests the technique is sufficiently powerful to cope with the known structure of groups.

Reports and Publications

F. Cannonito and M. Finkelstein, "On Primitive Recursive Permutations and Their Inverses." Accepted by The Journal of Symbolic Logic, Dec 1968. AFOSR 69-0762TR AD 684607.

Final status report, Jan 1969. AFOSR 69-0748TR.

4.3 SYMBOLIC LOGIC AND RECURSIVE FUNCTION THEORY

R. M. Smullyan
City University of New York, New York
(formerly of Yeshiva University, New York)
Grant Number AF-AFOSR-68-1375

Summary of Recent Research Results

A monograph was produced which concerned itself with updating and streamlining first-order logic through elementary quantification theory and presenting new methods on "cut-free" methods. One of the highlights was a tableau technique which proved to be a distinct improvement upon known tableau schemes by E. W. Beth and by K. J. J. Hintikka. This technique was used to prove that: (1) any formula provable by the tableau method is a tautology; (2) every tautology is provable by the

tableau method; (3) if X is true in all Boolean valuations which satisfy S, then X is deducible from S. These three major points respectively demonstrate the qualities of consistency, completeness, and compactness.

Reports and Publications

R. M. Smullyan, First-Order Logic, Springer-Verlag, New York, 1968, 158 pp.

R. M. Smullyan, "The Continuum Hypothesis," to appear in COSRIMS.

R. M. Smullyan, "Uniform Gentzen Systems," to appear in Journal of Symbolic Logic.

R. M. Smullyan, "Analytic Cut," to appear in Journal of Symbolic Logic.

R. M. Smullyan, "Abstract Quantification Theory," to appear in Symposium in Proof Theory and Intuitionism, Buffalo 1968, North Holland Publishing Company.

4.4 RELATIONS BETWEEN LOGICAL STRUCTURE, TIME, AND NATURAL NUMBERS

Gotthard Gunther
University of Illinois, Urbana, Illinois
Grant Numbers AF-AFOSR-480-64, AF-AFOSR-1391-68

Summary of Recent Research Results

A closer proximity between transclassic logic and mathematics has been made apparent through the development of a mapping of the Peano sequence of natural numbers onto kenogrammatic structure. The discovery that composition (addition) of numbers can be interpreted either in an iterative or an accretive sense enabled the formulation of a composition theory of proto numbers that is the starting point for the kenogrammatic structure. Research on a composition theory for

deutero structure showed that information theory, per se, is logically satisfied by deutero structure. This establishes that transclassic logic displays structural properties not covered by information theory, at least in its present form. No trito structure can be identified by only the frequency of occurrence of a certain symbol; identification also requires precise information about the location of every symbol in the context. With the discovery of an affinity between the first numbers of the binary system and the eight morphograms of two-valued logic, a number sequence was generated that potentially includes all possible tritograms of any length or structure. In the sense that numbers may be regarded as representations of structures, the step from one number to its successor signifies an increase in structure such that the interval between one number and the next becomes increasingly greater with respect to complexity. The number sequence is shown to represent an arithmetic place-value system of binary systems that, in turn, can produce ternary, quaternary, and generally n-nary systems.

4.5 RESEARCH ON TOPICS IN MATHEMATICAL LOGIC RELATED TO MECHANICAL MATHEMATICS

Martin D. Davis
New York University, New York, New York
Grant Numbers AF-AFOSR-995-66, AF-AFOSR-995-67, AF-AFOSR-1482-68

Summary of Recent Research Results

New results have been obtained in various branches of mathematical logic and computability theory. Specific relations have been found between various hierarchies of notations for transfinite ordinals used in algorithm finding proof-theoretic investigations. New reductions have been found of the problem of existentially defining all possible algorithmic processes in terms of addition and multiplication, in particular to the properties of a simple fourth degree Diophantine problem. The word problem for semigroups has been solved in various special cases involving a single relation. Systems of unary

second order logic and other extensions of predicate logic have been analyzed. Most scientific computer programs are compute bound. The results of this work will significantly reduce the amount of time needed to reach identical solutions obtained under conventional encoding methods and techniques.

Reports and Publications

Martin Davis, "One Equation to Rule Them All." Invited address before New York Academy of Science, November 1967. In press: Proceedings of the New York Academy of Sciences. AFOSR 69--363TR AD 683206.

Hilbert Levitz, "A Simplification of Takeuti's Ordinal Diagrams of Finite Order." Zeitschrift fur Mathematische Logik und Grundlagenforschungen (In Press).

Hilbert Levitz, "On the Relationship Between the Finsler and Tarski-Donner Hierarchies of Arithmetical Operations." Commentari Mathematicae Helveticae (In Press).

Hilbert Levitz, "On the Relationship Between Takeuti's Ordinal Diagrams O(n) and Schutte's System of Ordinal Notations $\sum(n)$, Studies in Logic. North Holland Press. (In Press).

Mitsuru Yasuhara, "On Minimal Ordinal Types of Theories," Koninklijke Nederlandse Akademie van Wetenschappen, vol 71 (1968), pp 87-94.

Mitsuru Yasuhara, "Incompleteness of L_p Languages," Fundamenta Mathematicae (In Press).

Mitsuru Yasuhara, "An Extension of the W-Completeness," submitted to the Journal of Symbolic Logic.

4.6 THEOREM PROVING BY COMPUTER

Bruce Kallick
Northwestern University, Evanston, Illinois
Contract and Grant Numbers AF-49(638)1349, AF-49(638)-1458, AF-AFOSR-1125-66

Summary of Recent Research Results

A decision procedure was developed for predicate calculus formulas in the $E_mA_2E_n$ prefix class. The procedure is a hybrid based on Friedman's decision procedure and on Robinson's resolution method. It resembles the latter in that resolvent clauses are formed from the clauses of a formula expressed in conjunctive normal form and from other resolvent clauses until two complementary one-literal clauses resolve into the empty clause. Unlike the general resolution method, the procedure does not generate literals with nested Skolem functions. There is, therefore, a bound on literal length and the procedure terminates even if the empty clause is not generated. Results suggest implications for theorem-proving procedures based on the resolution method in the direction of curtailing literal length at the expense of increasing the required number of resolvents if in doing so the overall computation effort can be reduced, e.g., by avoiding character manipulation. The procedure may be modified to decide the two matrix-defined classes of Skolem normal-form formulas whose solvability was established by Friedman's method. The procedure was programmed in Algol for the restricted class of formulas having only dyadic predicates.

Reports and Publications

Bruce Kallick "Proof Procedures and Decision Procedures Based on the Resolution Method." Aug 1968. AFOSR 68-1823 AD 674004.

Bruce Kallick "A Decision Procedure Based on the Resolution Method" Paper presented at IFIP Congress 1968. AFOSR 68-1584, AD 671994.

4.7 LOGIC AND RATIONAL PROCESSING IN SCIENTIFIC DISCOVERY

Michael Scriven
PEC, Inc., Athens, Georgia
Contract Number AF-49(638)-01672

Summary of Recent Research Results

Some of the final thinking of Norwood Russell Hanson (deceased), the

original principal investigator on this research, has been summarized. Tasks of philosophy of science were being reconceptualized and extended to include an explicit analysis of the concepts peculiar to the domain of discovery, both the concept of discovery itself and concepts of creativity, innovation, and development. The summary attempts to demonstrate the value of bringing elements from the field of discovery into the analysis of the basic concepts of completed science and the enrichment resulting from unification of subjective and objective elements in science. The contention is made that there is a set of teachable procedures that can be described as the logic of science that can increase the efficiency of scientific practice and accelerate the teaching of scientific methods. The reconceptualizations are particularly relevant to research designs and strategies and to the field of computer simulation of machines having capabilities resembling those of the human intellect. Notions of probability, cause, law, and information were reviewed with suggestions of ways of introducing new methods of quantification and cross-connection for them.

Reports and Publications

M. Scriven, "Logic of Discovery." 1968. AFOSR 68-2796, AD 679605.

4.8 SYSTEMS OF NATURAL DEDUCTION WITHOUT ESSENTIAL RESTRICTIONS ON VARIABLES

Kurt Bing
Rensselaer Polytechnic Institute, Troy, New York
Current Contract Number F44620-67-C-0014

Summary of Recent Research Results

A linear system of natural deduction was constructed that incorporates two simplified quantifier rules. No variables are flagged. The system is designed to mirror the process of intuitive linear reasoning. A step in this process is conceived as producing a new conclusion from a set of assumption formulas which may be the same as a set of assumption

formulas previously given or may be changed in the step, rather than as passing from previous conclusions to a new conclusion through the same set of assumption formulas. Thus the same purpose is accomplished as in tree-type systems by combining several given deductions into a new one. A method of representing a deduction in the system that may be amenable to computer application was developed for the condition that all sets of assumption formulas are finite. The linear system is complete as to logical consequence and admits the positive, minimal, and intuitionistic subsystems. A subformula property for suitable deductions (and hence the separation property for such deductions) was shown to hold for a minimal subsystem and is an improvement on a result of Leblanc.

Reports and Publications

Kurt Bing, "Systems of Natural Deduction Without Essential Restrictions on Variables." November 1968. Final Report. AFOSR 68-2765 AD 679600.

4.9 LOGIC FOR SELF-ADAPTING AUTOMATA AND UNIVERSAL ELEMENTS

John F. Randolph
University of Rochester, Rochester, New York
Grant Numbers AF-AFOSR-464-64, AF-AFOSR 464-65, AF-AFOSR 481-66

Summary of Recent Research Results

A logical analysis was made of a general theory of the structure of human psychodynamics that showed, from Korzybskian point of view, a post-relational structure in human behavior amenable to analysis as a logical system. The doctrine was axiomatized in set theory terminology and notation. The formulation facilitates semantic analysis of such concepts as identification, allness, and linear thinking and enables natural language discourse to be more formally evaluated.

4.10 RELATIONSHIPS BETWEEN THEORY AND PRACTICE FOR DIGITAL SYSTEMS

B. Elspas
Stanford Research Institute, Menlo Park, California
Current Contract Number F44620-69-C-0012

Summary of Recent Research Results

A summer study was conducted to examine gaps between theory and practice in the design and use of digital systems. Important specific problems have been defined in areas of switching theory, languages and formal grammars, and design automation algorithms. An extensive annotated list of problem formulations, graded according to importance and difficulty, is being prepared as a guide to future research and applications. Switching theory, for example, has not been of significant use in the design of large logic networks, in the distribution of logical functions in computer systems, or in the physical partitioning and placement of devices under combined constraints of speech, number of terminals, package complexity, etc. Theories of automata and formal languages have not been applied in the design of computer systems or in the design of executive programs. Alternatively, the popularity of man-computer approaches may indicate inadequacies of theory for the practical design of large systems. Such practice is common in technology, though instances such as the use of algebraic models for formulating error-correcting codes suggest the likelihood that stronger links can be forged between theory and practice for the improvement of digital systems.

4.11 STUDY AND ASSESSMENT OF PROGRESS IN GRAPH THEORY

J. B. Turner and William Kautz
Stanford Research Institute, Menlo Park, California
Contract Numbers F44620-68-C-0024, F44620-67-C-0026

Summary of Recent Research Results

A comprehensive survey was made of all published Soviet literature on graph theory and its applications that totaled over 230 technical articles and books appearing up to June 1968. A report summarizes

the results and compares the Soviet state of the art with that in the West. Although Soviet activity in graph theory lags behind that in the West in quality and quantity, the level of activity and quality are increasing rapidly. The best Soviet work is concerned with bounds on numerical indices associated with graphs, properties of algebraic structures associated with graphs, and operations on graphs. Very little work has been reported on connectivity properties of graphs, matroid theory, the exact enumeration of graphs having prescribed properties, isomorphism testing, graph coloring, and the use of graphs for modeling in the social sciences. A world-wide indexed bibliography on graph theory has been compiled and will be published.

Reports and Publications

James Turner and William H. Kautz, "A Survey of Progress in Graph Theory in the Soviet Union," Final Report, Project 6885, Stanford Research Institute (Menlo Park, California, November 1968). AFOSR 69-0745TR AD 684156.

James Turner, "A Key-Word Indexed Bibliography on Graph Theory," Appendix to "Proof Techniques in Graph Theory," Proceedings of the Ann Arbor Conference on Graph Theory, 1968, Academic Press (New York, 1969). AFOSR 69-2812TR.

William H. Kautz and Karl N. Levitt, "A Survey of Progress in Coding Theory in the Soviet Union," Final Report, Project 6313, Stanford Research Institute (Menlo Park, California, December 1967). (This report was issued as a result of a similar survey made during the previous year, on contract F44620-67-C-0026, and is to be published as a special issue of the IEEE Transactions on Information Theory in early 1969). AFOSR 68-0590 AD 666917.

4.12 RESEARCH IN AUTOMATA THEORY AND ITS MODELING APPLICATIONS

Michael A. Arbib
Stanford University, Stanford Electronics Labs, Stanford, Calif.
Current Grant Number AF-AFOSR 1198-67

Summary of Recent Research Results

Contributions were made to the mathematical theory of automata and to applications of automata theory to embryology and psychology. Algebraic techniques, particularly those of category theory, were applied to the study of simple models of parallel computation. The treatment of algebraic automata was extended to automata employing algebras over arbitrary theories with special attention to the free theories that correspond, in ordinary automata theory, to the free monoids. The relationship was established between the algebraic automata theory approach to loop-free computation and the Krohn-Rhodes semigroup theoretic approach to sequential machines. A method was developed for analyzing multiplication in groups and semigroups, and the computation of various finite functions. A basic lemma was proved for obtaining new and improved bounds on the computation time required by a (d, r)-circuit to multiply elements of groups or semigroups. Models developed previously for analyzing self-reproducing automata were modified to increase their relevance to biology. Theoretical questions about the embedding of automata in arrays are being examined that should be relevant to the organizational problems of cellular arrangement in embryology. Models have been proposed for the child development process that suggest correspondences between such disorders as infantile autism and schizophrenias and differences in information processing mechanisms.

Reports and Publications

M. Arbib and Y. Giveon, "Algebra Automata I: Parallel Programming as a Prolegomena (sic) to the Categorical Approach." Jan 1968. AFOSR 68-0453, AD 666677.

Y. Giveon and M. Arbib, "Algebra Automata II: The Categorical Framework for Dynamic Analysis." Jan 1968. AFOSR 68-0454, AD 666457.

4.13 APPLICATIONS OF ALGEBRAIC AUTOMATA THEORY TO SYSTEM THEORY

Y. Giveon
Stanford University, Stanford, California
Grant Number AF-AFOSR-1636-69

Summary of Recent Research Results

Several methods that were developed in standard automata theory have been found useful in other branches of mathematical systems theory. The most outstanding of these are with respect to minimization of systems, simulation of systems, and the reduction of simulations to well-known algebraic theory (for example, semigroup theory for the simulation of sequential machines). These methods are being studied within the framework of category theory. Research in algebraic automata (an algebraic model for a certain kind of parallel information processing) and in linear systems (that are the core of control theory) is directed toward finding the particular properties of the application of automaton theoretic methods. Category theoretic tools are being developed for studying the essence of these methods in their full generality. Major results have been obtained on defining the notion of a category being concrete over another category, and on establishing axioms for concrete categories that enable the introduction of general system theory as a development of category theory. Results are also being obtained on establishing the notion of a wreath product for algebraic theories and characterizing the irreducible algebraic automata as those with irreducible algebraic theories.

4.14 FUNDAMENTAL AND UNIFICATIONAL TOPICS IN AUTOMATA THEORY

Dana S. Scott
Stanford University, Stanford, California
Contract Number F44620-68-C-0030

Summary of Recent Research Results

Necessary conditions have been developed for languages to be context-free, deterministic context-free, stack generable, stack decidable, and non-erasing stack generable. The result for context-free languages is a generalization of the necessary condition of Bar-Hillel, but it provides an alternative approach to inherent ambiguity in context-free languages. The infiniteness problem for stack generable languages was shown to be solvable.

Stack decidable and non-erasing stack generable languages were shown as not closed under reversal. Two unsolvability results were obtained. The question of whether a stack generable language is inherently ambiguous as a stack language is undecidable. The question of whether the reversal of a non-erasing stack generable language is also non-erasing stack generable is undecidable. A mathematical theory was formulated for the kinds of transformations required in the transformation grammars of Chomsky. The notion of a transducer operating on tree structures was defined that generalizes the machines that usually operate on strings. Several closure and decidability properties were proved for the T languages determined by the transducer and initial context-free grammars. T languages are conjectured always to be context sensitive. The theorems demonstrate that progress in structural analysis of languages is obtainable from work on the level of trees rather than strings.

Reports and Publications

W. C. Rounds, "Trees, Transducers and Transformations." Aug 1968. AFOSR 69-0089TR, AD 681174.

W. F. Ogden, "Intercalation Theorems for Pushdown Store and Stack Languages." AFOSR 69-0321TR, AD 682981.

Section 5

INDEX TO INSTITUTIONS

Section 6

INDEX TO PRINCIPAL INVESTIGATORS

Section 7

INDEX TO ARPA SUPPORTED PROJECTS MONITORED BY THE DIRECTORATE OF INFORMATION SCIENCES

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Air Force Office of Scientific Research
Office of Aerospace Research, USAF
Arlington, Virginia 22209

2a. REPORT SECURITY CLASSIFICATION
UNCL

2b. GROUP

3. REPORT TITLE
INFORMATION SCIENCES--1968

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Scientific Special

5. AUTHOR(S) (First name, middle initial, last name)
Harold Wooster
Russell B. Ives

6. REPORT DATE
December 1969

7a. TOTAL NO. OF PAGES
85

7b. NO. OF REFS
125

8a. CONTRACT OR GRANT NO. In-House

b. PROJECT NO. 9769

c. 6144501F

d. 681304

9a. ORIGINATOR'S REPORT NUMBER(S)
AFOSR 69-3000TR

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT
1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES
TECH OTHER

12. SPONSORING MILITARY ACTIVITY
Air Force Office of Scientific Research(SRI)
Directorate of Information Sciences
Arlington, Virginia 22209

13. ABSTRACT

This is the fifth in a series of annual reports of the basic research program monitored by the Directorate of Information Sciences, Air Force Office of Scientific Research. It reports the progress made on 65 contracts and grants active during all or part of calendar year 1968; in the four areas designated by the COSATI subject headings: documentation and information technology, language and linguistics, bionics, and information theory. Selected highlights of the program are given in the introductory chapter. This is followed by brief statements of progress and publications for each work unit. Indexes to institutions, principal investigators and ARPA supported efforts monitored by the Directorate staff are included.

DD FORM 1 NOV 1473

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A ROLE	LINK A WT	LINK B ROLE	LINK B WT	LINK C ROLE	LINK C WT
Information sciences						
Programming (computers)						
Bibliographies						
Digital computers						
Computers						
Scientific research						
Computer logic						
Classification						
Documentation						
Computer storage devices						
Information retrieval						
Language						
Speech						
Hearing						
Nervous system						
Mathematical logic						
Pattern recognition						
Speech transmission						
Machine translation						
Set theory						
Cybernetics						
Social communication						
Nerve cells						
Psychoacoustics						
Bionics						
Artificial intelligence						
Perception (psychology)						
Auditory perception						
Autonomic nervous system						
Speech recognition						
Man-machine systems						
Libraries						
Automata						
Information theory						
Linguistics						
Semantics						
Syntax						
Computational linguistics						
Context free grammars						
Subject indexing						

GPO 887-123

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)*
Air Force Office of Scientific Research
Office of Aerospace Research, USAF
Arlington, Virginia 22209

2a. REPORT SECURITY CLASSIFICATION
UNCL

2b. GROUP

3. REPORT TITLE

INFORMATION SCIENCES--1968

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Scientific Special

5. AUTHOR(S) *(First name, middle initial, last name)*
Harold Wooster
Russell B. Ives

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
December 1969	85	125

8a. CONTRACT OR GRANT NO	In-House	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO	9769	
c.	6144501F	9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
d.	681304	AFOSR 69-3000TR

10. DISTRIBUTION STATEMENT

1. This document has been approved for public release and sale; its distribution is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
TECH OTHER	Air Force Office of Scientific Research(SRI) Directorate of Information Sciences Arlington, Virginia 22209

13. ABSTRACT

This is the fifth in a series of annual reports of the basic research program monitored by the Directorate of Information Sciences, Air Force Office of Scientific Research. It reports the progress made on 65 contracts and grants active during all or part of calendar year 1968; in the four areas designated by the COSATI subject headings: documentation and information technology, language and linguistics, bionics, and information theory. Selected highlights of the program are given in the introductory chapter. This is followed by brief statements of progress and publications for each work unit. Indexes to institutions, principal investigators and ARPA supported efforts monitored by the Directorate staff are included.

DD FORM 1 NOV 68 1473

UNCLASSIFIED

Security Classification