NATIONAL BUREAU OF STANDARDS REPORT

7882

THE APPLICATION OF AUTOMATA THEORY TO PROBLEMS IN INFORMATION RETRIEVAL (WITH SELECTED BIBLIOGRAPHY)

by

Russell A. Kirsch

A report from the Research Information Center and
Advisory Service on Information Processing
Data Processing Systems Division

To the

National Science Foundation

NBS

U.S. DEPARTMENT OF COMMERCE

NATIONAL BUREAU OF STANDARDS

THE NATIONAL BUREAU OF STANDARDS

The National Bureau of Standards is a principal focal point in the Federal Government for assuring maximum application of the physical and engineering sciences to the advancement of technology in industry and commerce. Its responsibilities include development and maintenance of the national standards of measurement, and the provisions of means for making measurements consistent with those standards; determination of physical constants and properties of materials; development of methods for testing materials, mechanisms, and structures, and making such tests as may be necessary, particularly for government agencies; cooperation in the establishment of standard practices for incorporation in codes and specifications; advisory service to government agencies on scientific and technical problems; invention and development of devices to serve special needs of the Government; assistance to industry, business, and consumers in the development and acceptance of commercial standards and simplified trade practice recommendations; administration of programs in cooperation with United States business groups and standards organizations for the development of international standards of practice; and maintenance of a clearinghouse for the collection and dissemination of scientific, technical, and engineering information. The scope of the Bureau's activities is suggested in the following listing of its four Institutes and their organizational units.

Institute for Basic Standards. Electricity. Metrology. Heat. Radiation Physics. Mechanics. Applied Mathematics. Atomic Physics. Physical Chemistry. Laboratory Astrophysics.* Radio Standards Laboratory: Radio Standards Physics; Radio Standards Engineering.** Office of Standard Reference Data.

Institute for Materials Research. Analytical Chemistry. Polymers. Metallurgy. Inorganic Materials. Reactor Radiations. Cryogenics.** Office of Standard Reference Materials.

Central Radio Propagation Laboratory.** Ionosphere Research and Propagation. Troposphere and Space Telecommunications. Radio Systems. Upper Atmosphere and Space Physics.

Institute for Applied Technology. Textiles and Apparel Technology Center. Building Research. Industrial Equipment. Information Technology. Performance Test Development. Instrumentation. Transport Systems. Office of Technical Services. Office of Weights and Measures. Office of Engineering Standards. Office of Industrial Services.

---

* NBS Group, Joint Institute for Laboratory Astrophysics at the University of Colorado.

** Located at Boulder, Colorado.

NATIONAL BUREAU OF STANDARDS REPORT

NBS PROJECT

12-12-12405
12101

NBS REPORT

7882

March 1, 1963

THE APPLICATION OF AUTOMATA THEORY TO PROBLEMS IN INFORMATION RETRIEVAL (WITH SELECTED BIBLIOGRAPHY)

by

Russell A. Kirsch

A report from the Research Information Center and
Advisory Service on Information Processing
Data Processing Systems Division

To the

National Science Foundation

NBS

U.S. DEPARTMENT OF COMMERCE

NATIONAL BUREAU OF STANDARDS

FOREWORD

This report has purposely not been cast in the conventional tutorial form of a state-of-the-art review. For this reason, it will serve only partially the needs of an interested newcomer desiring to familiarize himself with current activity in the field of automata theory. This limitation extends into the selection of the bibliography, for which the author carefully indicates the criteria used in making choices from the much larger bibliography on automata theory. In the interests of a more coherent presentation, the body of the report and the related bibliography adhere to topics for which the author can depict a plausible relationship to problems that arise out of information retrieval needs. Thus, this report serves as a state-of-the-art review for the serious information retrieval investigator having a prior contact with the formalizable aspects of his problem. In addition, it is directed toward the research worker in several of the scientific disciplines that the author considers applicable to some of the problems arising in information retrieval. It is hoped that the material presented here will elicit sufficient interest on their part to consider some of these topics as appropriate for their own research interests. It is also hoped that this advance indication of a dual objective will cause both groups to take this into consideration and find that their interests have been adequately served.

S. N. Alexander, Chief
Data Processing Systems Division

Contents

THE APPLICATION OF AUTOMATA THEORY TO PROBLEMS IN INFORMATION RETRIEVAL (WITH SELECTED BIBLIOGRAPHY)

by

Russell A. Kirsch

I. INTRODUCTION

The purpose of this survey is to suggest certain topics in automata theory that may be explored with a view toward achieving further understanding of what has been presystematically called "the information retrieval problem". The survey is presented in the form of selected problems that have direct connection with the information retrieval problem. For each of these problems there is presented a description of the appropriate topics in automata theory which furnish a proper formulation of appropriate questions and, in many cases, the answers to these questions. Along with those topics in automata theory, reference is made to relevant items in a bibliography of the published literature.

The problems in information retrieval which have been selected for study are not claimed to constitute a representative sample of problems or even a sample of the most important problems. The selection has, instead, been made according to the presumed utility of automata theory in treating these problems. To those readers who would prefer to proceed from practical versions of the information retrieval problem to more abstract versions, the procedure here may seem unwarranted. To those readers three justifications are offered. First, the automata theory literature exists and is growing rapidly. If possible, it may just as well be interpreted and applied as be allowed to develop with no connection to these practical applications. Secondly, many interesting and perhaps important questions about the behavior of information retrieval systems have been asked in a presystematic manner that does not admit of solution. In some of these cases, a formulation in terms of automata theory seems to lead to a solution which may be identified as an answer to the source question. Third, the conspicuous absence of any unifying theory of information retrieval creates a commonly felt desire to develop one. The automata theory literature may offer a starting point.

It should be clearly understood that in this survey, we are not formulating problems for solution and even less are we solving such problems. We are rather, simply suggesting areas in which others may carry on such tasks. In order to be formulated sufficiently precisely to warrant serious investigation, most of the problems suggested by the topics that are discussed here would require much more attention than can be afforded in this survey. Our contribution is one of selection of and motivation toward certain topics of investigation. It will remain for others to establish the significance of this contribution.

Some comments about what we include as automata theory are in order. At this point, we will simply state a prejudice in favor of those parts of the literature which view automata as language manipulating devices, which study the abstract information processing capabilities of automata, and which study the particular task of mechanizing inference in formal languages. Because of this parochial selection of topics in automata theory, the present survey should not be considered as a general survey of automata theory. A preliminary estimate has indicated that to have included all of what might reasonably be considered as automata theory would have increased the size of the bibliography by a decimal order of magnitude. We do claim, however, to have achieved reasonable comprehensiveness on the topics that have been chosen for inclusion.

Information retrieval (IR) means many things to many people. Quite uniformly, however, it is a characteristic of situations in which the IR problem occurs that we can find an occurrence of what DeSola Price calls a "petty illness of science - its superabundance of literature". 1/

The existence of the petty illness is a fact upon which most observers may agree. The point of departure arises when attempts are made to diagnose the illness and to prescribe cures. The cures may range from palliatives, many of which exist and are successful, to panaceas, none of which is commonly agreed to exist. The diagnosis also will differ with different observers. It is interesting to notice that the diagnosis usually anticipates the cure in the sense that those problems are identified by the diagnostician which he believes can be cured with the available devices of a conceptual or technological sort.

1/ D. J. DeSola Price, "Science Since Babylon," Yale Univ. Press, 1961, New Haven, Conn., p. 124.

Clearly, insofar as problems are formulated in terms of available intellectual devices, an expansion of the set of such devices offers the possibility of new formulations of the IR problem (hopefully more ambitious ones) and the possibility of solutions to these problems. Prior reference was made to the pre-systematic version of the IR problem to indicate that extant descriptions of the parts of the IR problem with which we are concerned in this survey are not usually presented in terms of formal mechanisms. There are a few notable exceptions, e.g., the use of Boolean algebra, the propositional calculus, and the calculus of classes for manipulating descriptors, or the use of graph theoretic techniques and their algebraic counterparts for explicating the notion of connection or relevance between documents or information items. To these few formal tools is added that of automata theory.

We suspect that the IR problems to which automata theory will find most powerful application are those that occur in some of the very large information collections, e.g., the U.S. Patent Office. In these applications, there is a need for handling documents over which very little control may be exercised during their generation. In particular, there is a need for manipulating the information content of the documents rather than the documents themselves, which have comparatively little function to perform. We find a need to handle natural language as an input source and a contrasting possibility of using artificial languages as communication mechanisms. We find a need for inferring from the contents of many documents the relevance of the contents of a particular document to an inquiry and we find a corresponding desirability for mechanizing this inference process. Finally, when we study such large IR problems, we are faced with the question whether any a priori assurance can be offered that exhaustive search is possible. We shall attempt to show in this survey that a reasonable possibility exists that investigations in automata theory will offer the opportunity to obtain precise counterparts of the questions that arise and to synthesize systems which satisfy the practical needs.

A final prefatory remark is in order. We cannot, in the space of this survey, attempt an exposition of automata theory as such. This may pose a problem to some readers, since we will refer to and use ideas developed in the literature. We therefore suggest either that the reader consult one of the reference surveys [59, 86, 87, 130, 135, 195, 202, 238, 241, 249, 263, 308] 1/, or

1/ Hereinafter these numbers identify specific bibliographic items in Section VI.

that he run the risk associated with interpreting our remarks in an informal manner even when we use notions very precisely defined in the literature, e.g., finite automaton, language, proof procedure, and decision procedure. Although it will be disconcerting to some authors to see their terminology quite frankly used out of context, it should be a source of satisfaction that the great care they generally use in selecting terminology results in a jargon that can communicate something of a motivational nature to the newcomer to the field. The serious worker will, of course, go on to a more careful study of the subject, with negligible damage having been done for the sake of motivation.

II. DEFINITION OF AUTOMATA THEORY - SCOPE NOTES

There are no commonly accepted definitions of automata theory and to adduce one from some standard source will serve no purpose here. The extensional definition of our subject is, of course, given by the bibliography which contains references through June 1962. A more intensional definition is given in the following paragraphs.

The literature of automata theory is intimately connected with the study of machines. In the very old literature, none of which is included here, the machines studied were intended to appear lifelike. The contemporary literature is, in a few cases, similarly motivated, but by a more circuitous route. What characterizes the modern literature of automata theory is the emphasis upon information processing. It is no accident that this literature has developed concurrently with the development of the stored program computer (and also with the development of telephone switching systems). As a consequence, a large part of the automata literature is directed toward the design of computer-like devices and of switching circuits. All such papers have been excluded here except certain ones that are included for other reasons. Specifically excluded therefore are papers on design of switching circuits (especially relay contact networks) and also references on switching theory which often includes work of an algebraic or combinatorial nature. This exclusion probably eliminates about a thousand papers.

One may move, in a series of small steps, from papers concerned with practical design problems in automata to papers that study the abstract properties of classes of automata. For example, one may study the behavioral properties of all automata, subject to the restriction that they possess a finite memory irrespective of the practical problems of size, error, difficulty of design, etc. [249] Our center of attention falls mostly in this more abstract area of automata theory. Thus on the subject of finite automata we have included literature concerned with problems in representing languages by finite automata, while we have excluded literature that studies relations between finite state languages and the memory requirements or structural constraint they impose on machines that generate or recognize them. We have attempted to include the literature on studies of the behavior of automata as viewed from exterior measurements.

Our particular emphasis has been upon papers which attempt to say something about the structural (syntactical) properties of the languages of inputs accepted by automata or outputs generated by them subject to the constraints imposed upon the automata themselves. Where these constraints are rigid, many properties can be derived that hold for the inputs and outputs (the languages) of the automata. However, it has been well established [70: see pg. 174, Theorem 6.5] that when the constraints imposed upon the automata are eased only slightly to allow the computing capability of Turing machines, it no longer becomes possible to say anything in a constructive manner about structural properties of the languages of these automata. For this reason, the subject of recursive function theory or computability theory which studies the behavioral properties of Turing machines or equivalent automata has very little to say about the structural properties of the languages manipulated by such automata. In spite of this, we have included certain references to recursive function theory, i.e., the general reviews of the subject, because of several important connections with automata theory which will be expanded upon later. Thus, although the subject of recursive function theory has intimate connections with automata theory, we have included it in this survey for other reasons that will be clear when we discuss the application to IR.

Another subject included in this survey also has a somewhat tenuous connection with automata theory, namely, mechanical theorem proving. The subject is usually considered a part of mathematical logic (it happens to have a direct connection with recursive function theory). The contemporary literature contains several papers which provide proof procedures for parts of logic by direct manipulation of the logical languages [225: see also 48, 67]. Here the use of such syntactic techniques relates directly to our main subject of interest, and consequently such papers are included. However, the majority of the literature on mechanical theorem proving techniques [best access is via 308] makes use of a technique based upon an expansion theorem due to Herbrand. The papers that use these expansion techniques rather than the syntactical ones are included in this survey because of the potential relevance to the problem of mechanizing inference in IR systems.

III. THE APPLICATION OF AUTOMATA THEORY TO PROBLEMS IN INFORMATION RETRIEVAL

A. IR Systems Which Must Process Natural Language

An important class of information retrieval problems occurs in systems where the information to be retrieved is in the form of natural language. (By natural language, we specifically include textual material, but we will subsequently extend the definition to include other classes of languages.) One can certainly find large and important IR systems in which the data to be manipulated are largely numerical or symbolic in one form or another and not natural language information in the sense of text. The fact is, however, that a large residue of important IR systems require that documents which have been prepared exclusively for consumption by people be processed by machines for purposes of information retrieval. The information content of these documents is partly in the form of natural language.

As for the processing of such natural language information by machines for purposes of information retrieval, there are at least three commonly held viewpoints. The first viewpoint holds that natural language is largely a capricious and unsystematic mechanism for conveying information, and that to this extent it is not possible to develop machine procedures for processing the natural language information in ways comparable to that in which people process such information. This viewpoint leads to a rejection of any possibility of mechanizing the processing of natural language. Occasionally, this viewpoint is defended not on the thesis that natural language is capricious, but rather that natural language is based on such an inordinately large number of regularities that to provide the descriptions of these regularities that underlie natural language is an insurmountable task, and regularities, not rules, [1/] are all that we can describe.

Although there is no reason to reject the thesis regarding the complexity of, say, English, there is also no corresponding reason to conclude that the complexity of the language is such as to transcend the data processing capabilities of conventional data processing systems.

1/ Paul Ziff, "Semantic Analysis", Cornell Univ. Press, Ithaca, N.Y., 1960, p. 34.

The second viewpoint on natural language processing is based on a somewhat equivocal stand on whether or not the complete description of a natural language is possible. This viewpoint, however, exploits a stochastic model of language to achieve some degree of success in auto-indexing and auto-abstracting. The primary tool used in such cases is the digital computer in its role as a numerical information processor.

The third viewpoint that we offer on the question of whether or not machines can process natural language text holds that in principle such a natural language as English is systematic and has regularities, and that formal systems are constructable whose rules explain these regularities. The possibility of creating such rules has an importance well beyond the intrinsic interest in explaining a natural phenomenon like English. The importance derives from the strong indication that complete formal description of a natural language like English will simultaneously provide at least a partial description of the process whereby natural language is interpreted and understood, and consequently will provide at least a partial explicatum of the information bearing function of natural language.

The strongest possible defense of this third viewpoint would, of course, be to exhibit the explanatory set of rules which describe the observed regularities in natural language and then to show the way in which these rules can be exploited to provide a systematic theory of the language. Since no one is known who can presently supply a complete set of rules, we will offer here an attempted explanation of why such rules are not forthcoming and the constructive suggestion of how automata theory may contribute to the formulation of such rules. Several examples of automata applications to English are described in the literature [53, 179, 140a, 317]. Our explanation for the unavailability of the rules that describe a language like English is based upon essentially a simple notion, that of the complexity of the language. It would not be a misconstruing of the whole tradition of modern descriptive linguistics to suggest that as natural language is now understood, it is essentially a complex process.

Complexity is understood here in the sense that the minimal description of the process is elaborate. This is not to say that the structural linguist cannot find sweeping generalizations and pervasive observations of the properties of natural language. It is, however, to suggest that a thorough comprehensive description of a language like English would necessarily involve a vast number of phenomena individually explained. We find, thus, that simple descriptions of language are available only to the normative grammarians, those who describe language as it should be, whereas the descriptive linguist finds himself confronted with an essentially complex process to describe.

The complexity of this process suggests a simple reason why the process has thus far not been described, namely the unavailability of suitable publication media for the results of such investigations.

Let us imagine that a descriptive linguist were possessed of all the necessary insight into the behavior of a spoken language to enable him to provide a complete structural characterization of the regularities in that language. This assumption may indeed be a valid one for certain linguists. We ask now in what form the results of the insight of such a linguist can be presented. Certainly they cannot be presented in the form of conventional publication for several reasons: First, there is no audience for the description of processes that are essentially complex; second, there is in conventional publication media no mechanism for achieving consistency and no insistence upon precision; and third, there is no requirement for results to be expressed unambiguously and unequivocally. It should be evident by now that the desirable characteristics which conventional publication media lack are for the first time available in a new form of publication as a medium for linguistic description. We are, of course, referring to the mechanism of the computer programming language. All indications are that through the medium of the computer program extremely complex processes may be described. Furthermore, these descriptions are operational descriptions in the sense that they may be used and interpreted.

We wish to suggest, therefore, that the mechanism of the computer programming language provides an ideal publication medium for the research results of the descriptive linguist when he attempts to describe the regularities he observes in a natural language. Since this new tool has only comparatively recently become available to the descriptive linguist [319], we suggest that herein lies an explanation for the previous unavailability of any exhaustive structural description of common natural languages. We have digressed to offer one of the reasons for the present lack of structural description of a language like English. However, our main purpose is to provide a more constructive suggestion, namely, that automata theory can make a significant contribution to the description of natural languages.

1. Automata Theory as a Source of Mechanisms for the Description of English Text

Let us confine ourselves at first to the particular problem of describing the structure of English language text. We will inquire into the usefulness of automata theory in describing the regularities that are observed in printed natural language.

In an early paper [1o2], Kleene considered an abstract mathematical system consisting of a set of symbols and the sequences that could be formed from these symbols. The significant question asked and answered was whether a simple characterization could be obtained for the sequences of such symbols which could be generated or accepted by an automaton subject to the sole restriction that its memory be finite. Much of the automata literature has developed in the tradition established by Kleene's paper. More recently we find the interpretation in which the alphabet consists of words, and the sequences recognized or generated by automata are sequences of words and therefore properly called sentences.

A few seminal papers [14, 20, 22, 55, 60] quite explicitly introduce the language interpretation for the automata which they investigate, and much of the recent progress in the study of such automata has been motivated by the purpose of interpreting the results as results about languages. It is of considerable use to think in terms of such an interpretation and then to ask ourselves what the theoretical models furnish toward understanding the properties of the languages which are thereby defined.

From Kleene's paper, the development of the subject was largely motivated by mathematical considerations. Copi et al [68] furnished an interpretation of Kleene's results in terms of computers which provided a more readily understandable interpretation of Kleene's paper. The automata considered in these papers were the finite automata. Subsequently, the series of investigations initiated by Chomsky [51] and Bar-Hillel [22] eased the restriction on finiteness for the automata being used. A whole series of automata subsequently were created which ranged from the finite automata to the Turing machines in their computing capabilities [255].

A corresponding subject of considerable interest among mathematical linguists has been the problem of choice of the appropriate automaton to account for natural language data of the type that is observed and for extrapolations therefrom. Considerable attention has been devoted to the choice of the correct place in the hierarchy from the finite automata through the Turing machines at which common natural languages can be described [51]. Such investigations have led to purely linguistic questions of interest regarding foundational matters concerned with the appropriate corpus to describe in a grammar for a language.

Currently there exists controversy between the attempts to account for certain parts of natural language as Yngve does with finite automata [317], and the attempt on the part of others following in Chomsky's path to use models of greater power than the finite automata. Much of this controversy rests upon the appropriate choice of corpus to be described. There is no need, however, to concern ourselves with this controversy since our consideration here is the extent to which automata theory can contribute to a further understanding of these problems which are information retrieval problems qua linguistic problems.

Let us see then what benefits the descriptive linguist obtains in consulting the automata theory literature as a source of mechanisms for doing descriptive linguistic research. The first thing the linguist will find is a rich assortment of formal mechanisms for describing languages. These formal mechanisms will generally be abstractions of rather specialized computers, each having its own limited degree of computing capability. What is interesting and simultaneously important about these *formal mechanisms* is the extent to which they lack the general computing capability of an ordinary computer or programming language.

Previously, we suggested that the use of a programming language as a mechanism for presenting linguistic results offers certain advantages of consistency and a demand for precision, the ability to express complex processes, and a degree of unequivocalness hardly demanded by other mechanisms. We would expect that the linguist would accept the demands imposed upon him by the rigors of a programming language in order to achieve a degree of operational usefulness for the results of his investigation. Unfortunately, however, there is both a practical and a theoretical reason why the linguist who expresses his results in the form of statements in certain programming languages will encounter difficulties in achieving the degree of usefulness that he might hope for. The difficulties arise from the very power inherent in most programming languages.

Let us explore this difficulty. Suppose that a linguist obtains data from some corpus and attempts to provide a general description of the data that he has obtained and that he expects to obtain on further investigation. He thus expects some predictive value from his results in studying natural language. Suppose further that he states these results in the *form of processes* implemented in a general programming language. We have already argued that these natural language processes are inherently complex and in particular that they are probably sufficiently complex not to allow simple comprehension by observation alone. So the linguist quite obviously decides that, in order to evaluate the results he has obtained and to analyse the consequences of his predictions, he will forgo the possibility of manual checking and will rather make use of the same tool, namely the computer, as an aid in evaluating his own results. At this point, he encounters both the practical and theoretical difficulty mentioned above.

It is comparatively simple to write a process description for a general-purpose computer about which no other computer of even arbitrarily great computing capacity can answer certain precisely stated questions. For example, the linguist might decide to write a simple program that uses a finite vocabulary and that produces from this finite vocabulary a potentially unlimited number of sentences which he would like to identify as sentences in the language he is describing. He might then like to ask of some arbitrary sentence, whether it has been included in the set of sentences he has thus described. In general, this question can be answered neither by the linguist nor by any computer of arbitrarily great computing power [70, Theorem 2.3]. In individual isolated instances, the computer may indeed provide a tool which can answer such questions, but a general decision procedure of the type we mentioned does not and can not exist.

As a practical matter, it is also the case that writing computing programs to answer questions about other computer programs is an occupation with only limited rewards, particularly when the programs being investigated are complex. It thus appears that the ideal tool for presentation of descriptive linguistic results might be one whose ability to accept complex process descriptions is not greatly limited, but which nevertheless has the property that evaluation of the processes thereby described is still possible. Here is where the automata literature has many mechanisms to contribute, for example, finite state automata [20], simple phrase structure automata [55, 273], categorial grammars [22, 176, 208], context-dependent phrase structure grammars [55], linear bounded automata [214], dependency systems [102], and other mechanisms which have yet to be applied as linguistic models [140a, Chapter 6 re automata].

We have shown that it is important for the description of linguistic processes to be presented in terms of certain automaton mechanisms having less computing capability than general-purpose computers. It remains for the investigator to choose among competing models. Fortunately, his choice here can be a fairly well informed one. For most of the formal mechanisms which are presented in the literature, there is available a characterization of the structural properties of the languages which they can describe. Consequently, if the linguist knows the corpus that he intends to describe with a formal mechanism, he may be able to reject certain such mechanisms as of too limited computing capability to do the job and reject others as being unnecessarily powerful. We may consider an example of such a choice that might be made.

Let us give a recursive characterization of an infinite set of strings of English words which we will elect to consider as English sentences. First, we will consider the single word "Hello" to be an English sentence. Then we will allow, as an English sentence, any sequence that consists of "How do you say" followed by an English sentence, followed by "in German". All and only such sentences will be the sentences in the language we intend to account for. Thus, we are considering sentences like "Hello", or "How do you say "Hello" in German", or "How do you say "How do you say "Hello" in German" in German", etc. It has been shown [249] that no finite automaton can produce just this language. On the other hand, a simple phrase structure grammar or linear bounded automaton [214] can. Here then, is an example of an optimal choice of automaton model for describing a particular corpus.

Occasionally the choice may be of an appropriate corpus in light of a particular model, in just the opposite sequence from the one mentioned above. The investigator may decide that certain sentences (particularly those of arbitrary length) can be either included or excluded according to whichever decision makes the resulting theory simpler and according to whichever decision makes the resulting processing mechanisms simpler.

We have stated that, when one elects to describe results in the form of a general-purpose computer program, certain important processors that one might want to construct cannot in fact be built. However, when one restricts oneself to the use of more narrowly circumscribed automaton models, such as some of the grammars listed in a preceding paragraph, then many processors can be constructed although a few important ones are still impossible to construct. It is important to notice that the processors can be constructed essentially independent of the linguistic investigation which is concerned with constructing a theory for a particular corpus.

An example of a processor that can be constructed in all cases is a generator which produces objects of the language being described. Such a generator is of some use in the early stages of debugging a grammar insofar as it exhibits to the investigator certain of the consequences (usually randomly selected) of the statements that he has made in his linguistic theory [252, 321]. An even more important processor is a general-purpose parser or syntactic analyser [170].

Such a processor takes as input a grammar and a particular sentence, and determines of the sentence whether it is producible from the grammar and, if so, what its analysis with respect to the grammar is, that is to say, how the grammar could produce such a sentence. For most of the restricted automata considered, such a processor is capable of construction. There are other processors of use in debugging, for example, concordance makers which can generally be constructed for most of the automaton models.

From the practical standpoint, it is important to notice that these processors can all be constructed as an essentially independent investigation, usually of a mathematical or programming nature, disjoint from the essentially linguistic investigation. Furthermore, the processors can usually anticipate the linguistic results in that they can be very general in their capabilities and can be used continually during the course of a progressing linguistic investigation. The advantages in independent construction of these language processors should not be overlooked. Unfortunately, not all the processors that one might desire are available, even in principle, for some of the automata that one might elect to use. It has been shown, for example, that no processor can be constructed which will take two arbitrary phrase structure grammars and determine whether the languages producible from these grammars are equivalent [23]. This equivalence question is undecidable. One must use a more restricted automaton such as the finite automaton to have the possibility of such a processor.

There is one last advantage in the use of automaton models which does not occur in the literature. This is the possibility of designing mathematical procedures for a comparative evaluation of alternative grammars. It is usually the case that the investigator confronted with a particular corpus of data elects one description of these data over another because of its presumed greater simplicity or its greater correspondence with intuitive insight or with some operational procedure. It seems that when one is presenting results in formal automaton terms, a possibility opens up of reducing one's notion of simplicity to very precise terms and then of using mechanical procedures for comparing alternative descriptions submitted for the same data. The undecidability results, such as the one mentioned above, offer some caveats with respect to the generality of such evaluation procedures, but particularly for the more restricted automata the possibility of such evaluation procedures in at least a few important cases remains a good one for exploitation.

We have thus suggested five kinds of contributions that the automata theory literature has for the descriptive linguist attempting to account for the regularities in English text. First, we see that automata theory contributes certain precisely formulated descriptive mechanisms. Second, it provides theorems about these mechanisms from which known properties of the descriptions using these mechanisms may be derived. Third, the use of the automata theory models enables the investigator to design processors of very general capability such as parsers, generators, or concordance makers. Fourth, automata theoretic considerations lead to certain caveats regarding the impossibility of certain tasks and suggest to the linguist that, for certain kinds of questions, general processors can not exist and hence cannot be constructed. Finally, the automata theory literature enables the investigator to choose certain models and forces the rejection of certain others once he has fixed upon his corpus which is to be described.

2. Automata Theory for Describing Other Phenomena as Languages

In previous discussion, we have restricted ourselves to considering the particular form of natural language as it occurs in textual material. It will now be of interest to inquire whether the same automata theoretic tools that are useful in describing natural language text can also be used in describing other phenomena which may properly be considered as natural linguistic phenomena. Let us list some of the distinguishing characteristics that separate these phenomena which we call natural languages from other phenomena which occur in human communication situations. First, we find phenomena which, to unsophisticated observation, appear to be at best only probabilistically constrained but which, on further investigation, turn out to have some very pervasive rigid regularities. Secondly, we find a rather small alphabet of symbols juxtaposed (usually in a linear array) in complex manner. Third, we find that people can learn quite easily to behave according to the regularities of the language, although, paradoxically, they seem to be very poor [198] at describing the regularities that characterize their learned behavior. Fourth, and probably most important from the pragmatic standpoint, we find that these phenomena are used for the purpose of communicating meaning. The first two of these four properties are the ones which are exploited in applying automata theory to the description of such a language, and the last two are exploited in making non-mechanized, that is to say human, information retrieval systems work. If a parallelism between the natural language phenomena in English text and other phenomena which occur in information retrieval systems can be found we would hope by analogy to exploit the application of automata theory to these other phenomena in a similar manner to the way we use it in handling natural language text.

Four classes of phenomena which occur in information retrieval systems and which we suggest can also be handled profitably as natural language phenomena are (1) display information, such as mathematical equations that occur in text; (2) chemical nomenclature; (3) handwriting; and (4) geometrical configurations, such as occur in schematic diagrams. Let us first consider objects which usually occur embedded in English language text and which are generally referred to as mathematical display information. A typical example would be a mathematical equation.

Of the four characteristics of natural languages mentioned above, mathematical display information certainly has the first and the second. It also serves the communication purpose of property number four, but we might briefly ask whether it has the third of the properties mentioned, namely, that it is easily learned in an informal manner and not explicitly verbalized by users of this purported language. Some unpublished investigations by the author and others have showed that, notwithstanding the occasional explicit definitions of mathematical notations, such as appear primarily in books on logic and other formal texts [1], the typical instance of mathematical usage is an informal one based not on any explicit formal definition that has occurred but rather a definition that results from common usage. Thus, display information in mathematical text has the third of the four properties also.

Some may raise the objection that, when viewed as a language, display information fails to be one-dimensional. This is, of course, a valid objection. It suggests, therefore, that if automata theoretic techniques are to be used in providing formal models for the description of mathematical text as natural language, these models must be generalized to handle two-dimensional configurations, i.e., more generalized kind of concatenation than is usually studied in automata theory.

It is interesting to note that the need for handling such information as mathematical display information embedded in text is not generally recognized when consideration is given to the problem of input of natural language text. One typically thinks in terms of print-readers which can perform linear scan of documents and which can transcribe into data processing equipment the representation of the characters which have thus been linearly scanned. It does not seem possible, however, to characterize mathematical text in quite so simple a linear fashion. Consequently, one would suspect that linear scanning devices like character readers will require some degree of pre-editing to eliminate essentially nonlinear language information from the input before scanning techniques such as are commonly being used will work.

---

[1] F. Cajori, A History of Mathematical Notation.

It may even turn out that documents which contain information for retrieval purposes and which are nominally textual in nature will be discovered to have more information content than is denoted by the linear text strings in the documents and that such additional information will serve important purposes for information retrieval.

We are referring here to the kind of nominally textual information which is highly formatted and for which the formatting conveys additional information beyond that which is present in the occurrences of words in the text alone. Examples which come to mind are such things as figure captions and headings, page numbers, equations in mathematical texts, labels within figures, etc. The point we are trying to make, therefore, is that documents which are nominally textual in nature in fact contain occurrences of another kind of natural language which is two-dimensional and which serves an important communication purpose just as does the text itself. Insofar as automata theoretic tools can be applied to the study of the text language, so similarly it would appear these tools can be applied to the descriptions of this language inherent in document format.

A second kind of natural language which is occasionally found in certain kinds of documents and which has less practical importance is that of handwriting. There are some indications [83] that the four natural language properties are shared by handwriting and that a consequent application of automata theoretic tools can be made here. However, the practical consequences do not seem significant.

The third application is one of more practical consequences, namely, chemical nomenclature. Garfield [105] showed how techniques of structural linguistic description can be applied to the syntactic analysis of the names of chemical compounds. It certainly appears to be the case that one can produce a grammar for the language of chemical nomenclature. One practical consequence of so doing would be to solve once and for all the problem of ambiguity of nomenclature since it is quite possible to design machine procedures which will syntactically analyze a language with respect to its grammar and then identify as equivalent structures which are syntactically distinct.

The last example of a natural language phenomenon occurs in diagrams and schematics, and in general in pictorial objects which are symbolic rather than representational. The dividing line between the symbolic and the representational pictorial information sources is not an easy one to describe precisely. However, definitely included within the symbolic pictorial sources are electrical circuit diagrams, chemical structure diagrams, flow charts, printed pages, and certain mechanical drawings. The last example is perhaps a transitional one between the symbolic and representational. Representational sources, of course, include photographs and sketches.

In order to establish our thesis that such pictorial sources can profitably be viewed as natural languages and described with automata theoretic tools, let us consider one representative example, namely, that of electrical circuit diagrams. Since circuit diagrams are an ideal instance of a symbolic pictorial source, we would hope to find here the four properties that characterize the natural languages to which automata theoretic tools seem applicable. Indeed, we do find first of all a set of regularities which admit of variations but which only allow a choice among a set of quite definitely prescribed alternatives. Thus, the symbol for a capacitor or a resistor in a circuit diagram may be any one of a surprisingly large number of alternatives as may be the symbolic mechanism used to denote the fact that the capacitor and resistor are electrically connected. However, the formal mechanisms which may be used are rather rigidly prescribed, and any violation of them is immediately recognized by a user of this "language" as a violation of correct use.

Second, we note that the language may be described by a finite though rather large vocabulary of symbols with rules regulating the ways in which these symbols may be juxtaposed, although one seldom, if ever, sees an explicit statement of what the regularities are which the rules can explain. Third, we note that circuit diagrams are read and produced by people who usually learn them informally, and lastly, the communication function performed by circuit diagrams is, of course, apparent. It would be hoped that in applying the tools of language description provided by automata theory in this more generalized example of a language, one would thereby obtain techniques for using such devices as character sensing machines to handle as primary input to information retrieval systems this kind of information source which presently seems well beyond any possibility of input to an information retrieval system except by manual means.

B Artificial Languages in IR Systems

Thus far, we have been concerned primarily with the input problem to information retrieval systems We have been inquiring into the applications of particular techniques to characterizing information which has been produced independently of information retrieval purposes and which must be analysed by descriptive techniques. Such is the nature of the input problem in many retrieval systems where the information to be processed largely exists prior to and independent of the retrieval system but must, nevertheless, be manipulated within the system for retrieval purposes. There appears to be no way to avoid suitable analysis of the input information in such systems either by mechanical techniques, in which case we have suggested the application of automata theory, or by manual techniques, which of course represent the more common case in present retrieval systems.

It is necessary now that we draw a sharp distinction between the information to be manipulated in a retrieval system over which there is essentially no control because the information exists and has been produced for purposes independent of the system and the information which is produced at the system or within the system or for the purpose of being used by the system. Quite clearly, an example of such more manageable or controllable information occurs in search prescriptions that are addressed to an information retrieval system. Notice that although a search prescription may talk <u>about</u> natural language objects, there is no need for the search prescription itself necessarily to be <u>in</u> a natural language. The point is that interrogations made to a mechanized retrieval system can quite possibly be within some artificial language, although inevitably the artificial language must be able to manipulate natural language objects if it is to have any retrieval power at all. Our contention is, thus, that although the object language of a retrieval system must inevitably contain natural language within it, the syntax (or interrogation) language need not be a natural language and may in fact be a highly stylized artificial language.

We will grant that certain advantages accrue to making the interrogation language as well as the object language in a retrieval system a natural language. For example, as has been pointed out by others, processors of the object and syntax language can be shared in common. One would expect that in attempting to provide, within a search prescription, an operational definition of "relevance" one would have to make heavy use of the object language in formulating the search prescription. In such cases, a common language and common processors shared between the language <u>within</u> the system and the language <u>of</u> the system would have obvious benefits. However, the relatively primitive state of our understanding of how to interpret natural language commands suggests that during a certain interim period it may be necessary that the interrogation language within a retrieval system be at least to some extent stylized and artificial. It should be apparent by now that our main thesis in this section will be that artificial languages within IR systems can be so designed that they share some of the properties that have been discovered to hold for natural languages.

In areas of application other than IR, there are in existence today completely artificial languages which have been constructed to look like English. This appearance is achieved by sharing a common alphabet (or vocabulary) between the artificial language and English [265a]. It is well to recognize, however, that more is known about natural language than its vocabulary. One is led to suspect that if some of the syntactical properties which are known to hold for natural languages were built into artificial languages, the convenience of use and perhaps even the ease of implementation in, say, compiling would increase considerably.

Although the ALGOL language [217] was not designed for information retrieval purposes, a brief consideration of the automata theoretic questions regarding ALGOL may prove enlightening. The definition of ALGOL was presented as two essentially disjoint systems, the first being the syntactical specification of the language and the second being the semantics or intended interpretation of the syntactical system. Although such a practice is common in mathematical logic, it is rarely the case that in areas of practical application the syntax of a language is specified with such precision as is the syntax of ALGOL. We can see certain advantages which result from separating the definition of the syntax of the language from the interpretation or semantics. Irons [158] has been able to create a very economical compiler by exploiting the formal structure in which the ALGOL syntax is presented. This structure turns out to be partially that of a simple phrase structure language [118].

What is relevant about the ALGOL experience is that although the syntax is unique, the semantics is subject to much variation when the language is implemented upon different machines which use different arithmetics. It may appear surprising that the syntax of the language can be precisely defined while the semantics is left open or subject to different interpretations. That this can, in fact, be done follows from the wide-spread use of numerical processing techniques and the consequent existence of an informal language that arises to communicate these techniques. ALGOL is a formal language whose structure is motivated by corresponding structures in the informal language.

Now the possibility is available to us for constructing artificial languages for interrogating IR systems in very much the way a language like ALGOL is constructed. We cannot, of course, exploit so universal a language as that used for motivating ALGOL, but we can gather those things that are said by IR system customers in interrogating their systems and then construct artificial languages which say the same things. The implementation and interpretation of such languages will necessarily vary from one IR system to the next, but we would expect that the syntax of the formal language could remain constant. The language Query [46] is a simple example of a language defined in this fashion.

C. Inference in IR Systems

We have been concerning ourselves with methods whereby a description of the structural properties of natural language text may be obtained and made available to a machine system. We have also discussed certain processors that may be applied to such a linguistic description, the primary purpose of which is to help improve the description and make it more accurate and complete. Thus far, however, we have not concerned ourselves with the purpose of arriving at such a structural characterization of natural language text. In this section, we shall discuss one of the main applications of a linguistic theory such as we have been discussing, namely, the application to the process of formal inference. We shall consequently be discussing information retrieval systems that perform inference on formal objects that derive originally from natural language text.

We do not suggest that there exist any current information retrieval systems which can in any strong sense of the word be said to perform formal inference. Such systems would have to infer from assertions which appear in documents to consequences which need not appear explicitly. A possible future exception may be [166]. The nearest that we might come to that is in the case of information retrieval systems which perform some calculation before delivering an answer. It is well to contrast the kind of information retrieval systems which perform inference with those which might properly be called search and selection systems. The search and selection systems upon receipt of a search prescription find in storage some item which answers the prescription or perhaps some composite of several items which answer the search prescription. The essential limitation in such systems is on the size of the storage. For practical purposes, this storage limitation often is not significant and many ingenious mechanisms have been designed whereby access to such storage may readily be achieved.

In fundamental contrast to such systems, however, is the kind of information retrieval system in which we are primarily interested here. In such a system, a storage consisting of a finite and possibly even quite small set of formal objects is used in conjunction with a set of rules of inference to produce additional objects for storage. These additional objects may, in fact, be computed and stored but it would be more likely the case that the number of such objects being potentially infinite would preclude the possibility of storing all the derivable objects in the system. Consequently, in systems which are capable of performing inference, one would expect that the more common situation would be that items which answer a search prescription would be computed only after the search prescription had been furnished to the system.

There is a very precise but unfortunately artificial example that we may use to illustrate the contrast between the more conventional kind of retrieval system and the one capable of inference such as we are interested in here. Suppose the purpose of a retrieval system were to furnish customers with theorems in elementary plane geometry. The search prescription might consist of schematic representation of theorems and the answer to a search prescription might be a theorem of plane geometry which satisfies the schematic description. There are two kinds of retrieval systems which could operate in the way we just described. The first kind would consist of a large set of previously calculated theorems of plane geometry stored and so classified in storage that upon receipt of the search prescription, reasonably direct access could be had to an appropriate theorem which satisfies the search prescription. The number of such theorems stored would necessarily be finite but considerable ingenuity could be used to assure that all theorems reasonably called for by reasonable customers would, in fact, already have been anticipated and stored. The other kind of system which would behave in a similar fashion would be the one capable of performing inference. In such a system, a relatively small set of axioms for plane geometry would be stored along with formal rules of inference. Then upon receipt of a search prescription, the system would proceed to calculate theorems until a suitable theorem was obtained which satisfied the schema of the search prescription. Considerable ingenuity would be required to assure that useless theorems would be less often generated than useful ones. Heuristic search techniques might be employed but there would be no guarantee of success or of efficiency. However, the number of theorems potentially available in such a system would be literally infinite and the possibility of the system furnishing some quite remote theorem in answer to a search prescription would be considerably improved.

We have belabored this example to illustrate the difference between conventional information retrieval systems, even those which operate with natural language objects such as automatic indexing and automatic abstracting systems, and information retrieval systems which perform inference based upon natural language text. These latter systems must contain linguistic theories to govern the text that they manipulate. Furthermore they must be able to combine linguistically analyzed objects with each other to produce new ones which need not have been formerly within the system. We must mention here (but only mention) the important philosophical question regarding the extent to which human beings perform inference as an entirely linguistic process or as an extralinguistic process. Regardless of how that question is answered, however, it is clearly the case that a great deal of inference performed by human beings is essentially linguistic in nature. We wish to propose here that some part of this inference process may be carried out within an information retrieval system.

The traditional way of studying the process of inference has been to translate assertions from natural language into some artificial language and then within the artificial language to have rules of inference which enable the artificial language objects to be used in producing new ones. A subsequent translation back from the artificial language to the natural language enables the results of the inference to be stated in natural language. In the long history of attempts to formalize inference, we would expect that at least some of the results that would be needed to create a mechanized inference system have indeed been created. If we may be so presumptuous as to attempt a succinct summary of the state of the art of mechanizing inference in natural language, we might characterize it as follows: The problem of linguistic analysis of the objects to be used in the process of inference has only recently received the rigorous treatment required for mechanization. The problem of translating from the natural language, with or without its exhaustive linguistic analysis, into a formal language of logic has received much treatment but usually of a fragmentary nature. The study of the formal inference process in logical languages whether or not they derive from natural languages has received massive treatment and constitutes in fact a large part of the study of symbolic logic. Finally, the problem of translating from formal logic objects into natural language objects has again received at best only fragmentary treatment.

Of the many parts of the problem of performing natural language inference within a mechanized system, there is only one that we can treat within this survey and that is the problem of performing formal inference in logical languages with automata. Before so doing, however, we may mention just one remaining interesting possibility. We have assumed that there must be a translation from a natural language into a formal language before inference can be implemented. This is not necessarily the case. It may eventually become possible to state rules of inference which operate not on formal logical objects but rather on syntax language objects which represent the output of analysis programs operating on natural language. The possibility of performing inference directly on syntactically analyzed natural language objects is an important one because it is a more realistic model of the presumed inference process in people, but we cannot discuss this further here. Instead, we will assume that what inference is to be performed is performed upon formal logical objects which may have arisen in translation from natural language. The literature we are concerned with here is that of mechanical theorem proving.

Although there has been a history for many years in studying the process of inference, most of the study has been metatheoretic in the sense that theorems concerning inference were developed rather than applied to performing inferences. In the past few years, however, much interest has developed in the possibility of actually implementing theorem proving techniques using data processing machines. In a sense, this represents a new departure for the logician insofar as his previous concern with a study of the process of inference has now been slightly reoriented toward actually performing inferences even though the inferences are performed through the intermediary of a machine. The inference that has actually been implemented in machine procedures has been almost entirely applied to problems in mathematics or in logic itself. Of the investigations that have been reported in the literature, we can detect four essentially different types which have received differing amounts of investigation. The first of these four types of theorem proving may be considered to have been the first used on computers even though it was identified as such only recently by Lehmer [180]. The theorem proving technique here consists of using the data processing machine as a clerical aid to the mathematician engaged in an essentially mathematical investigation. The computer serves here as an extremely rapid calculator for testing conjectures by calculation and for suggesting conjectures to the mathematician who bears the burden of the investigation. Investigations in number theory have proved a fruitful area of application for this kind of theorem proving technique.

A second kind of technique has been applied to theorem proving in the propositional calculus. The emphasis here has been upon direct manipulation of the mathematical language, in some cases the language of the propositional calculus, and in obtaining proofs by explicitly applying rules of inference which generate new theorems from old theorems [306, 48, 67, 189, 259]. A much more radical attempt at theorem proving is represented by the third technique, usually identified as heuristic theorem proving. The first applications were to theorems in the propositional calculus [221, 225] and subsequent ones were in geometry [112]. The application to formal integration [281] can also be considered in this class, although the primary emphasis has not been on theorem proving.

The fourth kind of mechanical theorem proving which has had the greatest activity recently makes use of expansion procedures which provide formal interpretations of the syntactical objects which are the theorems to be proved. [73, 305, also 29, 71, 72, 74, 81, 82, 100, 113, 114, 168, 171, 242, 243, 245, 258, 260, 292, 306, 307, 308].

The techniques used here are semi-proof procedures in the sense that they are applicable to undecidable domains. They provide proofs for formulas which are theorems and lead to procedures which fail to terminate for formulas which are not theorems. Much of the recent emphasis in this last area has been on attempts to truncate the expansion procedures to make them more economical. It is interesting to note that these truncation procedures turn out to be, in effect, highly formalized versions of heuristics quite comparable to those used in the heuristic theorem proving techniques. The supposed dichotomy that exists, therefore, between the two classes of theorem proving techniques, heuristic and so-called algorithmic, is therefore a difference in name only and not one of substance [305].

We are interested here in the mechanical theorem proving techniques, not because we are interested in proving theorems in mathematics but because they represent the sole attempt thus far to carry out, on machines, highly elaborate processes of inference. It is important to know whether the same techniques that are developed for mechanical theorem proving may ultimately be of use in carrying out inference in formal languages that derive from natural language. [312, 313]

The present mechanical theorem proving techniques represent successful attempts at programming highly complex combinatorial manipulations of formal language objects. To this extent, we expect that the same techniques will be useful in carrying out inference in formal languages that derive from natural language. There is one consideration, however, that seems to limit the degree of application of present theorem proving techniques to those which we are ultimately most interested in. To understand this limiting factor, we must first notice that the mechanical theorem proving techniques presently being investigated are applied to what the logician calls "pure or uninterpreted formal systems". Formulas in these pure systems become theorems if under every interpretation of the formal object, the formulas are true (actually this holds only in those formal systems where completeness theorems can be proved). As an example, let us consider a formula from a first order functional calculus.

$$Ex\ (\ Fx \vee - Fx\ )$$

This formula turns out to be a theorem in the first order functional calculus for (among other reasons) the fact that no matter how one interprets the property F and no matter what individual is assigned to the variable x , the expression is satisfied. The language is pure insofar as the freedom of interpreting F and x is as wide as we mention.

Now, in the case of a formal language that derives ultimately from a natural language, we would expect quite the opposite to be the case. We would expect that the counterparts of the function letters and the object variables would be highly interpreted, very rigorously constrained objects subject only to some slight freedom in how values may be assigned to them. If this is the case, it would follow that those inferences which can be carried out in natural language as formalized in some formal language will represent very specialized versions of the inferences that can be carried out in the more pure languages that are commonly studied today.

It is not clear whether inference in highly interpreted languages, like the natural language counterparts, is actually an easier or a more complex task to program on a machine than is inference in pure languages. We can see reasons for both sides of the argument. The task would be simpler with constrained interpretations for the formal objects, because so many fewer proof possibilities have to be explored than in the unconstrained case. The process of performing inference in the formalized version of natural language would on the other hand be more complex because the task of calculating the specialized interpretations of the formal objects would impose a significant data processing burden on a machine. Whether the net result would be a gain or a loss for the natural language theorem proving over the logical language theorem proving is quite unclear at present. Our conclusion must thus be that the present mechanical theorem proving techniques may potentially prove useful in carrying out inference in languages that derive from natural language more closely than the logical languages currently studied but that the nature of the contribution and the degree of contribution is at present unclear.

D. Theoretical Feasibility and Complexity Questions

There is associated with certain information retrieval systems the requirement to provide certain assurances or guarantees of system performance. For example, in military command and control systems, there is a need to assure reliability. In U.S. Patent Office searching, there is the need to assure exhaustiveness of search. Investigations into the nature of such requirements, when prosecuted far enough, often lead into certain fundamental considerations which we would like to discuss briefly here.

Our thesis is that some of the fundamental questions asked about information retrieval systems are the counterparts of questions that arise in certain areas of automata theory, particularly in computability theory and that if one wishes seriously to investigate these questions, the appropriate investigation is within computability theory [70, 262, 263, 282, and also 42, 144, 161, 174, 181, 186, 212, 215, 237, 240, 241, 291, 297, 299]. The questions that we shall briefly be concerned with here all arise in systems that are infinite or potentially infinite, so we will briefly digress to discuss the appropriateness of finite or infinite IR system models.

The mathematical techniques that have thus far proved useful in studying IR systems, namely Boolean algebraic techniques and graph theory, have been useful exclusively in systems that are finite and indeed, often rather small. Thus, one conventionally studies the Boolean algebra of a finite set of index entries for a document collection or the Boolean algebra of subsets of a document set, or if one has a finite number of information items and some indexes associated with these items, one may study the association networks formed by the connections between the indexes and the items they index. The graph theoretic techniques that are applicable here are used particularly for finite systems. We are led to inquire then, whether any information retrieval systems exist which should properly be modeled with infinite models. It appears that at least two kinds of retrieval systems can appropriately be modeled in this way. The first kind of system is the system which grows in principle arbitrarily large with time. Thus, the literature through which Patent Office searches are to be made grows with time and even though estimates of exponential growth have been made, it is reasonable to simply approximate this growth by considering it to be unbounded. A more plausible case for the infinite IR system occurs when we allow the possibility of inference. It is immediately the case that systems which have even only a small number of formal objects and a few rules of formal inference can admit of literally an infinite number of inferences that can be performed. Thus, if we take an IR system and inquire not about the information tokens stored in the system but of the consequences or results of inferences that can be carried out based upon what is stored in the system, we immediately have an infinite amount of information which must be searched.

Let us consider, then, that we have one or the other of these infinite information retrieval systems. What are some of the questions that one might ask about such systems, and what is the relevant theory that addresses itself to these questions? The first question arises out of a peculiar assymetry that exists between relevant and irrelevant items in an information collection. When one is engaged in a sequential search through an information collection, one often notices that certain information items are immediately recognized and accepted as being relevant (whatever that means) to a search prescription. The relevance can usually be exhibited. On the other hand, one seldom can assert of a particular item in an information collection that it is definitely irrelevant to a search prescription. As a matter of fact, a supposedly-irrelevant item may, later in a search, turn out to be quite relevant based upon information obtained from some subsequent item found in the search. This leads us to the following question. Suppose that in performing a sequential search through an information collection after some long period no item has been found which can be found relevant to the given search prescription. What conclusion should be drawn from the thus far executed search? One is tempted to suggest a rather far reaching analogy between this question and a question that is asked in computability theory. The counterpart question is as follows: In carrying out the systematic enumeration by an algorithmic process of the members of some infinite set, if a particular member being sought has not yet been encountered, what can be said about the ultimate possibility of encountering such a member? The answer to the counterpart question occurs in computability theory which exploits a theorem to the effect that for sets which are recursively enumerable but not recursive, there exists no decision procedure for membership [70]. The answer to the question is that nothing can be said about whether the desired member will be found or not found in such a search. One is lead by analogy, therefore, to suspect that in the information retrieval problem when performing a sequential search if a particular item has not yet been discovered to exist in the collection, only one of two possibilities remains, either that the item is not there or that it has simply not yet been shown to be in the collection even though it may, in fact, have already been encountered. One can think of practical examples of the situation. For instance, in a file arranged chronologically what is the relevance to a search prescription on communication devices of an information item published in the 1930's that relates to the preparation of purified germanium? Not until one encounters the solid state physical devices appearing later in such a file can one establish that the original item that occurred previously in the information collection was, in fact, crucially relevant. We wish to suggest, therefore, that an analogy and perhaps more exists between the properties of recursively-enumerable non-recursive sets in computability theory and information collections in which the non-existence of answers to search prescriptions can never be asserted with confidence.

The whole matter of relevance raises a second question where we suspect there may be a potential contribution of computability theory. It is generally recognized in information retrieval that no acceptable operational definition and certainly not one of any theoretical usefulness exists which explicates the notion of relevance. The notion of relevance is usually explained in terms of pragmatic matters or sometimes in terms of a semantic theory. We wish to suggest a purely syntactic notion of relevance.

Let us consider the artificial information retrieval system described in the previous section in which we have stored a set of theorems and in which the search prescription is in the form of a particular theorem which we may presume in this case does not exist within the system. There also is to be a set of rules of inference whereby proofs may be constructed. When the customer furnishes the search prescription to the system, he simultaneously supplies the system with a set of theorems which are common to him and the system or, informally, theorems which he knows to be true.

The problem of retrieval in this artificial system is to determine whether there exists a theorem stored in the system having the property that along with the theorems supplied by the customer, it may be used as a part of a proof of the theorem supplied as the search prescription. If such a theorem exists, we suggest that it is a relevant theorem to the proof of the theorem which is the search prescription. To be slightly more general, we suggest that a formal object stored within an information retrieval system is relevant to a search prescription if there exists an inference from certain commonly accepted objects common to the customer and the system to the given search prescription which makes use of the relevant formal object in the system. If there is no single such object but only a set of such objects, then that set is relevant and no single member of the set is relevant to the search prescription.

We do not pretend that this notion of relevance has a practical interpretation at the present time, largely because there does not presently exist either an information retrieval system in which inference can be performed nor even if such a system were to exist would we then have necessarily any ability to determine the information items common to the customer and the system which could be used to establish the relevance of a test item. Our suggestion here, then, is simply that if one wants to study this question of relevance in the abstract, one may study the question of whether or not items exhibited during a search belong to proofs of other objects which act as search prescriptions.

The last of the theoretical questions we intend to discuss is concerned with complexity measures. Before performing a search through a sequential file, and in particular in the case of an infinite file, it would be desirable if certain *a priori* measures of the complexity of the question could be obtained. For example, it might be useful if one could say that independent of the contents of the file, for two particular search prescriptions, one would necessarily involve more searching time than the other. There are extreme examples of course, where this can be done already. In a highly formalized IR system, one might encounter search prescriptions which turn out to be as formal statements tautologies. For such statements, no searching is necessary at all of course, and the complexity or cost associated with such a search would be minimal.

We ask, however, whether any *a priori* measures can be associated with search prescriptions in other than these trivial cases. For this question, there is a precise counterpart which occurs as a deep result in computability theory. The so-called Kleene-Mostowski hierarchy of arithmetical predicates is a hierarchy [70 chapter 9] which represents a partial ordering of certain formulas which may serve as a counterpart of search prescriptions in an information retrieval system. The use of the hierarchy theorem in computability theory is as follows: Certain questions put in the form of so-called recursive predicates can be answered by a minimal class of machines, namely the Turing machines. These questions are the counterparts of the questions which require no searching in IR systems. Next, by a very simple formal modification of the questions, actually by the addition of a single quantifier in front of a quantification schema, a new question is manufactured. If there is no way of eliminating this formal change, the question thereby manufactured becomes unanswerable by the first class of machines. If, by the intervention of some oracle, we provide this class of machines with the answer to one such question, a whole class of new questions now becomes answerable. However, the same modification on these questions can again be accomplished to produce a still further class of questions which are no longer answerable by the modified class of machines. An answer may again be externally furnished to a machine from such a class to provide a new class of answerable questions, and the process continues.

At the moment, we can only suggest an analogy between the IR system complexity question and the corresponding computability theory result. If however, we may anticipate the nature of such a correspondence that might be established, we would expect a series of search prescriptions to be classified in successively greater complexity classes according to the same formal properties that are used in the computability theory case. The particular form of property that is exploited is the number of changes from universal to existential quantifiers needed in order to specify the question. According to this analogy, the three search prescriptions listed below would be given successively greater complexity measures. Intuitively, it will be seen that the number of searches that must be performed through a file to find out whether the file has the property that satisfies the search prescriptions asked increases successively from the first to the third of the questions which follow.

Question 1.

Does the sentence "Starch is extracted from potatoes." occur in the file?

Formalization:

$$\exists x\ [x = \text{starch is extracted from potatoes}]$$

Question 2.

Does the sentence "Starch is extracted from t" occur in the file - when t denotes a word which never occurs in the file in a sentence of the form "t is a poison."?

Formalization:

$$\exists x\ \forall y\ [x = \text{starch is extracted from } t \text{ and } \\ -y = t \text{ is a poison}]$$

Question 3.

Does the sentence "Starch is u ed from t" occur in the file - where t denotes a word which never occurs in the file in a sentence of the form "t is a poison," and where the u in the past participle u ed is an allomorph of the u in some sentence which occurs in the file "u ion is a chemical process"?

Formalization:

$$\exists x \forall y\ \exists z\ [x = \text{starch is u ed from t and} \\ -y = \text{t is a poison and} \\ z = \text{u ion is a chemical process involving t}].$$

In summary, our thesis is that a purely formal analysis of search prescriptions addressed to a retrieval system may enable a partial ordering of these questions to be made with respect to some type of complexity measure which will be independent of the contents of the file to be searched.

We have thus seen that in at least three cases certain questions that are of a fundamental nature about the behavior of infinite IR systems correspond to questions that exist in computability theory. If one is sufficiently convinced that this correspondence is a strong one, a profitable investigation can be made into the computability theory results for the purpose of applying them to the theoretical problem in IR and hopefully thereby to obtain the answers to otherwise presystematically stated questions about IR systems.

IV. CONCLUSION

In conclusion, we would like to reiterate here that we have not treated the problems which we have discussed in a formal manner. It has not been our purpose to state questions which may be investigated as mathematical problems. That would be a more ambitious undertaking than we have attempted here. As a compromise between such an ambitious undertaking and attempting nothing in the direction of formalizing problems in information retrieval, we have offered here a discussion of certain problems for which there is practical motivation and we have attempted to show counterparts of these problems which have already received attention in the automata theory literature. Our medium has been analogy, our purpose motivation.

Quite evidently the major burden of accomplishing useful results is upon the reader who feels that a plausible case has been made for the productiveness of certain formal investigations. If it were possible, we would attempt to give an indication of the relative degree of refractoriness of the many questions raised in this report. With the present state of our knowledge, such indications can only serve as guises for prejudices. The matter of relative degrees of difficulty is perhaps a moot point, however, because there are most likely just two approaches that the serious investigator will take to the topics mentioned in this survey. He will either approach questions which are more interesting before those which are less so, or he will approach those questions whose importance to the information retrieval problem seem to him greater. "Which way is better? Neither way is better. Both ways are necessary. It is also necessary to make a choice between them." 1/

V. ACKNOWLEDGEMENTS

The author wishes to express his sincere appreciation to Mrs. Arlene Ruhl, Miss Armine Smith, and Miss Josephine Walkowicz who obtained the bibliographic items and to whom credit is due for the degree of accuracy that has been achieved on the citations. Errors are due to the impatience of the author. Thanks are also due to Robert D. Elbourn, B. Kirk Rankin, III, and Miss Walkowicz for helpful comments on the paper.

1/ T. S. Eliot: The Cocktail Party, Harcourt Brace & Co., New York, 1950, p. 141.

SELECTED BIBLIOGRAPHY

001. ACF Industries, Inc. "Investigations of Formal Systems of Reasoning Leading to Machine Processing." Alexandria, Va., undated, 13p., Final Report under Contract No. AF 19(604)-1582; Scientific Report No. AFCRC-TR-56-166. ASTIA doc. no. AD-110 103.

002. ACF Electronics. "Digitalization of the Propositional Calculus and the Monadic Functional Calculus." Alexandria, Va., December 1, 1955 - February 29, 1956, 33p., Contract No. AF 19(604)-1582; Scientific Report No. AFCRC-TN-56-176. ASTIA doc. no. AD-103 983.

003. ACF Industries, Inc. "Mechanization Techniques for Digitalized Logic." Alexandria, Va., June 1, 1956 to August 31, 1956, 35p., Contract No. AF 19(604)-1582; Scientific Report No. AFCRC-TN-56-593. ASTIA doc. no. AD-98 794.

004. Ackermann, William. "Solvable Cases of the Decision Problem" STUDIES IN LOGIC AND THE FOUNDATIONS OF MATHEMATICS, edited by L. E. J. Brouwer, E. W. Beth, and A. Heyting, Amsterdam, Netherlands, North-Holland Publishing Company, 1954, 114p.

005. Acton, Forman S. "Research on Digital Computer Language Problems." Princeton, N. J., Princeton University, December 1, 1960, 7p., Final Report on Contract Nonr 1858(22). ASTIA doc. no. AD-250 303.

006. Aizerman, M. A., L. A. Gusev, L. I. Rozonoer, I. M. Smirnova and A. A. Tal'. "Finite Automata I." AUTOMATION AND REMOTE CONTROL (The Soviet Journal AVTOMATIKA I TELEMEKHANIKA in English Translation), 21:2 (October 1960) 156-163. (Russian original dated February 1960, p. 224.)

____. "Finite Automata II." Ibid., 21:3 (November 1960) 248-254. (Russian original dated March 1960, p. 359.)

007. Amarel, Saul. "On the Automatic Formation of a Computer Program Which Represents a Theory." SELF-ORGANIZING SYSTEMS 1962, edited by Marshall C. Yovits, George T. Jacobi, and Gordon D. Goldstein, Washington, D. C., Spartan Books, 1962, p. 107-175.

008. Asey, Stan E. "On the Recursiveness of Sets Defined by Simple Phrase Structure Grammars." ZEITSCH. MAT. UND SPRACH, 2:3 (1928) 102-103.

009. Ashby, William R. AN INTRODUCTION TO CYBERNETICS. New York, John Wiley and Sons, 1957, 2nd edition, 295 p.

010. Asser Gunter. "Normal Post Algorithms." ZEITSCHRIFT FUR MATHEMATISCHE LOGIK UND GRUNDLAGEN DER MATHEMATIK, 5:3-4 (1959) 323-333.

011. ____________. "Turing Machines and Markov Algorithms." ZEITSCHRIFT FUR MATHEMATISCHE LOGIK UND GRUNDLAGEN DER MATHEMATIK, 5:3-4 (1959) 346-365.

012. Balestic, Ch.-J. "Analysis of Ideas Developed by Noam Chomsky in His Book: 'Syntactic Structures'." Gif-sur-Yvette (Seine et Oise), France, Centre d'Etudes Nucleaires de Saclay, February 1960, 67 p., CENS-DOC-AFD 2.

013. Bar-Hillel, Yehoshua. "On Syntactic Categories." JOURNAL OF SYMBOLIC LOGIC, 15:1 (March 1950) 1-16.

014. ____________. "A Quasi-Arithmetical Notation for Syntactic Description." LANGUAGE, 29:1 (January-March 1953) 47-58.

015. ____________. "Logical Syntax and Semantics." LANGUAGE, 30:2 (Part 1), (April-June 1954) 230-237.

016. ____________. "A Logician's Reaction to Recent Theorizing on Information Search Systems." AMERICAN DOCUMENTATION, 8:2 (April 1957) 103-113.

017. ____________. "The Mechanization of Literature Searching." MECHANISATION OF THOUGHT PROCESSES (Proceedings of a Symposium held at the National Physical Laboratory on 24th, 25th, 26th and 27th November 1958), London, Her Majesty's Stationery Office, 1959, Volume II, p. 789-800.

018. ____________. "Decision Procedures for Structure in Natural Languages." LOGIQUE ET ANALYSE 2:5 (January 1959) 19-29.

Also published as Appendix III of his "Report on the State of Machine Translation in the United States and Great Britain." Jerusalem, Hebrew University, February 15, 1959, 9 p., Technical Report No. 1, Contract Nonr 2578(00). N.R. 049-130. ASTIA doc. no. AD-215 139.

019. Bar-Hillel, Y. "A Demonstration of the Non-Feasibility of Fully Automatic High-Quality Machine Translation." Appendix IV of his "Report on the State of Machine Translation in the United States and Great Britain." Jerusalem, Hebrew University, February 15, 1959, 6p., Technical Report No. 1, Contract Nonr 2578(00). N.R. 049-130. ASTIA doc. no. AD-215 139.

020. Bar-Hillel, Yehoshua and E. Shamir. "Finite-State Languages: Formal Representations and Adequacy Problems." THE BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL; 8F:3 (February 1960) 155-166.

021. Bar-Hillel, Yehoshua. "Some Theoretical Aspects of the Mechanization of Literature Searching." Jerusalem, Hebrew University, April 1960, 74 p., Technical Report No. 3, Contract No. N62558-2214. ASTIA doc. no. AD-236 772.

022. Bar-Hillel, Yehoshua, C. Gaifman, and E. Shamir. "On Categorial and Phrase Structure Grammars." BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL, 9F:1 (June 1960) 1-16.

023. Bar-Hillel, Yehoshua, M. Perles, and E. Shamir. "On Formal Properties of Simple Phrase Structure Grammars." Jerusalem, Hebrew University, 8 July 1960, 58p., Technical Report No. 4, Contract No. N62558-2214.

024. Barnes, Robert F. "Language Problems Posed by Heavily Structured Data." COMMUNICATIONS OF THE ACM, 5:1 (January 1962) 28-34.

025. Belford, Geneva G. "Computer Logic Programs." Urbana, Ill., Digital Computer Laboratory, University of Illinois, Doctoral Thesis, June 1960, 69p., (Report No. 99). ASTIA doc. no. AD-236 974.

026. Belyakin, N.V. "Universality of a Calculating Machine with Potentially Infinite External Memory." Translation of PROBLEMY KIBERNETIKI (PROBLEMS OF CYBERNETICS), Moscow, State Publishing House of Physico-Mathematical Literature, 1961, Section II, p. 77-86; in "Foreign Developments in Machine Translation and Information Processing, No. 71," New York, N.Y., U.S. Joint Publications Research Service, January 19, 1962, 29 p., JPRS: 12050, distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

027. Berman, Richard, Joseph Sharp, and Lawrence Sturges. "Syntactical Charts of COBOL 61." COMMUNICATIONS OF THE ACM, 5:5 (May 1962) 260.

028. Beth, E. W. "Remarks on Natural Deduction." KONINKLIKE NEDERANDSE AKADEMIE von WETENSCHAPPEN. PROCEEDINGS. Amsterdam, North Holland Publishing Co., Volume 58, No. 3, Series A (Mathematical Sciences), 1955, p. 322-325.

029. __________. "On Machines Which Prove Theorems." SIMON STEVIN, 32 (1958).

030. __________. FORMAL METHODS, (an Introduction to Symbolic Logic and to the Study of Effective Operations in Arithmetic and Logic), New York, Gordon and Breach, 1962, 170 p.

031. Bohnert, H. G. "Comments" (on the "Pros and Cons of a Special IR Language"). COMMUNICATIONS OF THE ACM, 5:1 (January 1962) 10.

032. Buchi, J. Richard. "Weak Second-Order Arithmetic and Finite Automata." Ann Arbor, Michigan, College of Literature, Science and the Arts, University of Michigan, September 1959, 50 p., Technical Report (Report Nos. 2722, 2755, 2794-6-T), Contracts Nonr-1224(21), DA 20-018-ORD-16971 and DA 36-039-sc-78057. ASTIA doc. no. AD-225 753.

033. __________. "Regular Canonical Systems and Finite Automata." Ann Arbor, Michigan, College of Literature, Science and the Arts, University of Michigan, December 1959, 25 p., Report No. 03105; 2794-7-T, Contracts Nonr-1224(21), and DA 36-039-sc-78057. ASTIA doc. no. AD-233 478.

034. __________. "On a Problem of Tarski." Ann Arbor, Michigan, College of Literature, Science and the Arts, University of Michigan, July 1960, 8 p., Technical Note, Contracts DA 36-039-sc-78057 and Nonr-1224(21).

035. __________. "Turing-Machines and the Entscheidungsproblem." Ann Arbor, Michigan, College of Literature, Science and the Arts, University of Michigan, December 1961, 25p., Report No. 03105-22-T, Contract Nonr-1224(21). ASTIA doc. no. AD-268 676.

036. Burks, Arthur W. "The Logic of Fixed and Growing Automata." PROCEEDINGS OF THE INTERNATIONAL SYMPOSIUM ON THE THEORY OF SWITCHING, 2-5 April 1957, PART I, Cambridge, Mass., Harvard University Press, 1959, p. 147-188. (The Annals of the Computation Laboratory of Harvard University, Vol. XXIX.)

037. ________________. "Computation, Behavior, and Structure in Fixed and Growing Automata." BEHAVIORAL SCIENCE, 6:1 (January 1961) 5-22.

038. ________________. "Programming and the Theory of Automata." PARALLEL COMPUTERS, AUTOMATA AND ADAPTIVE SYSTEMS, Ann Arbor, The University of Michigan Engineering Summer Conferences, Summer 1962, 29 p.

039. Burks, Arthur W. and Hao Wang. "The Logic of Automata." Part I, JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY: 4:2 (April 1957) 193-218; Part II, 4:3 (July 1957) 279-297.

040. Burks, Arthur W. and Jesse B. Wright. "Sequence Generators and Digital Computers." Ann Arbor, Michigan, College of Literature, Science and the Arts, University of Michigan, February 1961, 101 p., Report No. 03105-19-T, Contract No. Nonr-1224(21).

041. Burks, Arthur W., Don W. Warren and Jesse B. Wright. "An Analysis of a Logical Machine Using Parenthesis-Free Notation." MATHEMATICS TABLES AND OTHER AIDS TO COMPUTATION, 8:46 (April 1954) 53-57. (Now MATHEMATICS OF COMPUTATION.)

042. Cannonito, Frank B. and Gerald D. Fogel. "A Survey of the General Theory of Effectively Computable Functions." Bethpage, N.Y., Grumman Aircraft Engineering Corporation, February 1962, 115 p., Research Report RE-153. ASTIA doc. no. AD-273 300.

043. Carr, John W., III. "Programming and Coding." Chapter 2 in HANDBOOK OF AUTOMATION, COMPUTATION, AND CONTROL edited by E.M. Grabbe, S. Ramo, and D.E. Wooldridge, Vol. 2, COMPUTERS AND DATA PROCESSING, New York, John Wiley and Sons, Inc., 1959, p. 2-01 - 2-270.

044. Carnap, R. and Y. Bar-Hillel. "An Outline of a Theory of Semantic Information." Cambridge, Mass., Research Laboratory of Electronics, Massachusetts Institute of Technology, October 27, 1952, 49 p., Technical Report No. 247.

045. Carton, Maurice. "A Mechanization Procedure for Geometry." CYBERNETICA, 3:2 (1960) 83-116. (In French.)

046. Cheatham, T.E., Jr. and S. Warshall. "Translation of Retrieval Requests Couched in a 'Semiformal' English-Like Language." COMMUNICATIONS OF THE ACM, 5:1 (January 1962) 34-39.

047. Cheatham, T.E., Jr., G.O. Collins, Jr., and G.F. Leonard. "CL-I, An Environment for a Compiler." COMMUNICATIONS OF THE ACM, 4:1 (January 1961) 23-28.

048. Cherry, Colin and P.K.T. Vaswani. "A New Type of Computer for Problems in Propositional Logic, with Greatly Reduced Scanning Procedures." INFORMATION AND CONTROL 4:2-3 (September 1961) 155-168.

049. Chomsky, Noam. "Systems of Syntactic Analysis." JOURNAL OF SYMBOLIC LOGIC, 18:3 (September 1953) 242-256.

050. ________________. "Logical Syntax and Semantics: Their Linguistic Relevance." LANGUAGE, 31:1 (1955) 36-45.

051. ________________. "Three Models for the Description of Language." IRE TRANSACTIONS ON INFORMATION THEORY, IT-2:3 (September 1956) 113-124.

052. ________________. "Logical Structures in Languages." AMERICAN DOCUMENTATION, 8:4 (October 1957) 284-291.

053. ________________. SYNTACTIC STRUCTURES. The Hague, Netherlands, Mouton and Company, 1957, 116 p.

054. ________________. "Linguistics, Logic, Psychology, and Computers." COMPUTER PROGRAMMING AND ARTIFICIAL INTELLIGENCE (1958 Summer Session, University of Michigan), edited by John W. Carr III, Lecture I, p. 429-454.

055. ________________. "On Certain Formal Properties of Grammars." INFORMATION AND CONTROL, 2:2 (June 1959) 137-167.

056. ________________. "A Note on Phrase Structure Grammars." INFORMATION AND CONTROL, 2:4 (December 1959) 393-395.

057. ________________. "On the Notion 'Rule of Grammar'." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, p. 6-24.

058. Chomsky, Noam. "Context-free Grammars and Pushdown Storage." Cambridge, Mass., Research Laboratory of Electronics, Massachusetts Institute of Technology, Quarterly Progress Report No. 65, April 15, 1962, 187-194.

059. ______________. "Formal Properties of Grammars." To appear in Vol. 2 of HANDBOOK OF MATHEMATICAL PSYCHOLOGY edited by R.R. Bush, E.H. Galanter, and R.D. Luce, New York, John Wiley and Sons, Inc., 1962.

060. Chomsky, Noam and George A. Miller. "Finite State Languages." INFORMATION AND CONTROL, 1:2 (May 1958) 91-112.

061. ______________. "Introduction to the Formal Analysis of Natural Languages." To appear in Vol. 2 of HANDBOOK OF MATHEMATICAL PSYCHOLOGY edited by R.R. Bush, E.H. Galanter, and R.D. Luce, New York, John Wiley and Sons, Inc., 1962.

062. Church, Alonzo. "Special Cases of the Decision Problem." REV. PHIL. de LOUVAIN, 49:22, Third Series (May 1951) 203-221.

______________. Correction: Ibid., 50:26 (May 1952) 270-272.

063. ______________. INTRODUCTION TO MATHEMATICAL LOGIC. Princeton, N.J., Princeton University Press, 1956, 376 p. (Princeton Mathematical Series, 17).

064. ______________. "Application of Recursive Arithmetic to the Problem of Circuit Synthesis." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N.J., 2nd edition, 25 July 1960, p. 3-50.

065. ______________. "Application of Recursive Arithmetic in the Theory of Computers and Automata." ADVANCED THEORY OF THE LOGICAL DESIGN OF DIGITAL COMPUTERS, Ann Arbor, University of Michigan Summer Session, June-July 1958.

066. Collins, George E. "Tarski's Decision Method for Elementary Algebra." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N.J., 2nd edition, 25 July 1960, 64-70.

067. Copi, Irving M. and Robert W. Beard. "Programming an Idealized General-Purpose Computer to Decide Questions of Truth and Falsehood." Ann Arbor, Michigan, Willow Run Laboratories, University of Michigan, April 1959, 17 p., Report No. 2144-402-T, Contract DA 36-039-sc-52654, Project Michigan. ASTIA doc. no. AD-214 305.

068. Copi, Irving M., Calvin C. Elgot and Jesse B. Wright. "Realization of Events by Logical Nets." JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY, 5:2 (April 1958) 181-196.

069. Curry, Haskel B. "Some Logical Aspects of Grammatical Structure." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII) Providence, R.I., American Mathematical Society, 1961, 56-68.

070. Davis, Martin. COMPUTABILITY AND UNSOLVABILITY. New York, McGraw-Hill Book Co., 1958, 210 p.

071. ____________. "A Program for Presburger's Algorithm." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N.J., 2nd edition, 25 July 1960, 215-223.

072. ____________. "Eliminating the Irrelevant from Mechanical Proofs." Paper presented at the five hundred eighty-ninth meeting of the American Mathematical Society, Chicago, Illinois, April 12-14, 1962, 12 p. Work supported in part by AFOSR under contract AF 49(638)-995.

073. Davis, Martin and Hilary Putnam. "A Computing Procedure for Quantification Theory." JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY, 7:3 (July 1960) 201-215.

074. Davis, Martin, George Logemann and Donald Loveland. "A Machine Program for Theorem-Proving." New York, N.Y., Institute of Mathematical Sciences, New York University, June 8, 1961, 13 p., Report No. IMM-NYU-288, Contract AF 49(638)777; AFOSR-819. ASTIA doc. no. AD-269 378.

075. deLeeuw, K., E. F. Moore, C. E. Shannon and N. Shapiro. "Computability by Probablistic Machines." AUTOMATA STUDIES, edited by C. E. Shannon and J. McCarthy, Princeton, N. J., Princeton University Press, 1956, 183-212.

076. Detlovs, V. K. "Normal Algorithms and Recursive Functions." DOKLADY AKAD. NAUK USSR, 90:5 (1953) 723-725. (In Russian)

077. ____________. "Equivalence of Normal Algorithms and Recursive Functions." TRUDY MATEMATICHESKOGO INSTITUTA imeni V.A. STEKLOVA, 52(1958) 75-139. (Collected Works, edited by N.A. Shapin, of the V.A. Steklov Mathematical Institute "Problems of Constructive Directions in Mathematics 1," Moscow and Leningrad, Academy of Sciences, U.S.S.R.)

078. Dreben, Burton. "A Systematic Treatment of the Decision Problem." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N. J., 2nd edition, 25 July 1960, p. 363.

079. Dunham, B. "Symbolic Logic and Computing Machines - A Survey and Summary." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC; CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N. J., 2nd edition, 25 July 1960, 164-166.

080. ___________. "The Formalization of Scientific Languages. Part I. The Work of Woodger and Hull." IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 1:4 (October 1957) 341-348.

081. Dunham, Bradford and J.H. North. "Theorem Testing by Computer." Paper presented at the Polytechnic Institute of Brooklyn Symposium on the Mathematical Theory of Automata, New York, N.Y., April 24-26, 1962.

082. Dunham, Bradford, R. Fridshal and G.L. Sward. "A Non-Heuristic Program for Proving Elementary Logical Theorems." INFORMATION PROCESSING, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING, UNESCO, PARIS, 15-20 June 1959, edited by S. de Picciotto, Paris, UNESCO; Munchen, R. Oldenbourg; London, Butterworths, 1960. 282-285.

083. Eden, Murray. "On the Formalization of Handwriting." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECT (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, 83-88.

084. Eden, Murray and A. P. Paul. "The Generation of Handwriting." Cambridge, Mass., Research Laboratory of Electronics, Massachusetts Institute of Technology, Quarterly Progress Report No. 61, April 15, 1961, 163-166.

085. Eggan, L.C. "The Star-Height of Regular Expressions." Ann Arbor, Michigan, College of Literature, Science and the Arts, University of Michigan, March 1962, 10 p., Technical Note, Contracts Nonr 1224(21) and DA-SIG-36-039-61-G4, ORA Projects 03105 and 04422.

086. Elbourn, R.D., and W. H. Ware. "The Evolution of Concepts and Languages of Computing." PROCEEDINGS OF THE INSTITUTE OF RADIO ENGINEERS, 50:5 (May 1962) 1059-1066.

087. Elgot, Calvin C. "Lectures on Switching and Automata Theory." Ann Arbor, University of Michigan, 1958, 77 p. ASTIA doc. no. AD-211 893.

088. ________________. "Decision Problems of Finite Automata Design and Related Arithmetics." TRANSACTIONS OF THE AMERICAN MATHEMATICAL SOCIETY, 98:1 (January 1961) 21-51.

089. Elgot, C.C. and J.D. Rutledge. "Operations on Finite Automata, Extended Summary." SWITCHING CIRCUIT THEORY AND LOGICAL DESIGN (Proceedings of the Second Annual Symposium, Detroit, Mich., October 17-20, 1961, and Papers from the First Annual Symposium, Chicago, Ill., October 9-14, 1960), edited by Robert S. Ledley, New York, American Institute of Electrical Engineers, September 1961, 129-132.

090. Ershov, A.P. "On Operator Algorithms." DOKLADY AKAD. NAUK USSR, 122:6 (October 1958) 967-970. Translation in Automation Express 1:5 (January 1959) 20-22.

091. Fairthorne, R.A. "Some Clerical Operations and Languages." INFORMATION THEORY, THIRD LONDON SYMPOSIUM (Papers read at a Symposium on 'Information Theory' held at the Royal Institution, London, September 12th to 16th 1955) edited by Colin Cherry; New York, Academic Press, Inc.; London, Butterworths, 1956, 111-120.

092. Feigenbaum, E.A. "An Information Processing Theory of Verbal Learning." Santa Monica, Calif., The RAND Corporation, October 9, 1959, 161 p., p-1817.

093. ______________. "The Simulation of Verbal Learning Behavior." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers presented at the Joint IRE-AIEE-ACM Computer Conference, Los Angeles, Calif., May 9-11, 1961), Western Joint Computer Conference, 1961, p. 121-132.

094. Feigenbaum, E.A. and Herbert A. Simon. "Generalization of an Elementary Perceiving and Memorizing Machine." Paper prepared for presentation at the International Federation of Information Processing Congress 1962, Munich, August 27 to September 1, 1962, 14 p.

095. Fitch, Frederic B. "Representation of Sequential Circuits in Combinatory Logic." Abstract in JOURNAL OF SYMBOLIC LOGIC, 24:4 (December 1959) 313.

096. Floyd, Robert W. "A Note on Mathematical Induction on Phrase Structure Grammars." INFORMATION AND CONTROL, 4:4 (December 1961) 353-358.

097. ______________. "On Syntactic Analysis and Operator Precedence." Woburn, Mass., Computer Associates, Inc., Report No. CA-62-2, August 1962, 31 p., Contract AF 19(628)-419.

098. Freudenthal, Hans. "Lincos, Design of a Language for Cosmic Intercourse, Part I." STUDIES IN LOGIC AND THE FOUNDATIONS OF MATHEMATICS, edited by L. E. J. Brouwer, E. W. Beth and A. Heyting, Amsterdam, North-Holland Publishing Co., 1960, 224 p.

099. Friedman, Joyce. "A Decision Procedure for Computations of Finite Automata." NOTICES OF THE AMERICAN MATHEMATICAL SOCIETY, 9:1, Issue 59 (February 22, 1956) 27-28 Abstract (588-6) of Contributed Papers, the February Meeting in New York, February 22, 1962.

100. ______________. "Decision Procedures for the Functional Calculus." Alexandria, Va., ACF Industries, Inc., March 1, 1956 to May 31, 1956, 44 p., Contract AF 19(604)-1582; AFCRC-TN-56-387. ASTIA doc. no. AD-104 175.

______________. Supplementary Report, September 1956, 45 p.; AFCRC-TN-56-785. ASTIA doc. no. AD-110 123.

101. Friedman, Joyce. "Some Results in Church's Restricted Recursive Arithmetic." JOURNAL OF SYMBOLIC LOGIC 22:4 (December 1957) 337-342.

102. Gaifman, Haim. "Dependency Systems and Phrase Structure Systems." Santa Monica, Calif., The RAND Corporation, May 22, 1961, 64 p., p-2315.

103. Gardner, Martin. "Logic Machines." SCIENTIFIC AMERICAN, 186:3 (March 1952)68-73.

104. Garfield, Eugene. "Chemico-Linguistics: Computer Translation of Chemical Nomenclature." NATURE, 192:4798 (October 14, 1961) 192.

105. ______________. "An Algorithm for Translating Chemical Names to Molecular Formulas." JOURNAL OF CHEMICAL DOCUMENTATION, 2:3 (July 1962) 177-179.

106. Gelernter, H.L. "A Note on Syntactic Symmetry and the Manipulation of Formal Systems by Machine." INFORMATION AND CONTROL, 2:1 (April 1959) 80-89.

107. ______________. "Realization of a Geometry Theorem Proving Machine." INFORMATION PROCESSING, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING, UNESCO, PARIS, 15-20 June 1959, edited by S. de Picciotto, Paris, UNESCO; Muanchen, R. Oldenbourg; London, Butterworths, 1960, 273-82.

108. ______________. "Theorem Proving by Machine." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N.J., 2nd edition, 25 July 1960, 305-308.

109. ______________. "Machine Generated Problem Solving Graphs." Paper presented at the Polytechnic Institute of Brooklyn Symposium on the Mathematical Theory of Automata, New York, N.Y., April 24-26, 1962.

110. Gelernter, H.L. and N. Rochester. "Intelligent Behavior in Problem-Solving Machines." IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 2:4 (October 1958) 336-345.

111. Gelernter, H., J. R. Hansen and C. L. Gerberich. "A Fortran-Compiled List-Processing Language." JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY, 7:2 (April 1960) 87-101.

112. Gelernter, H., J.R. Hansen and D.W. Loveland. "Empirical Explorations of the Geometry Theorem Machine." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, San Francisco, Calif., May 3-5, 1960), Western Joint Computer Conference, 1960, p. 143-149.

113. Gilmore, Paul C. "A Program for the Production from Axioms of Proofs for Theorems Derivable Within the First Order Predicate Calculus." INFORMATION PROCESSING, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING, UNESCO, PARIS, 15-20 JUNE 1959, edited by S. de Picciotto, Paris, UNESCO; Munchen R. Oldenbourg; London, Butterworths, 1960, p. 265-273.

114. ________________. "A Proof Method for Quantification Theory: Its Justification and Realization." IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 4:1 (January 1960) 28-35.

115. ________________. "Book Surveys: Quantificational Logic." INFORMATION AND CONTROL, 3:2 (June 1960) 197-203.

116. Ginsburg, Seymour. "Sets of Tapes Accepted by Different Types of Automata." JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY, 8:1 (January 1961) 81-86.

117. ________________. "Examples of Abstract Machines." Santa Monica, Calif. Systems Development Corporation, June 13, 1961, 15 p., Technical Memorandum 517/000/01.

118. Ginsburg, S. and H. G. Rice. "Two Families of Languages Related to ALGOL." Santa Monica, Calif., System Development Corporation, July 1961, 42 p., TM-578.

119. Ginsburg, Seymour and Gene F. Rose. "A Comparison of the Work Done by Generalized Sequential Machines and Turing Machines." Santa Monica, Calif., Systems Development Corporation, March 1961, 16 p., TM-604.

120. ______________________________. "Operations which Preserve Definability in Languages." Santa Monica, Calif., System Development Corporation, October 1961, 43 p., SP-511.

121. Ginsburg, S. and E. H. Spanier. "Distinguishability of a Semi-Group by a Machine." PROCEEDINGS OF THE AMERICAN MATHEMATICAL SOCIETY, 12:4 (August 1961) 661-668.

122. Giuliano, Vincent E. "The Logic of Automatic Formula Synthesis." PROCEEDINGS OF THE NATIONAL SYMPOSIUM ON MACHINE TRANSLATION (Held at the University of California, Los Angeles, February 2-5, 1960), Englewood Cliffs, N.J., Prentice-Hall, Inc., 1961, p. 462-471.

123. __________________. (Project Director). "Studies for the Design of an English Command and Control Language System." Cambridge, Mass., Arthur D. Little, Inc., June 1962, 1 v., ESD-TR-62-45, Report No. CACL-1, Contract AF 19(628)-256.

124. Giveon, J. "Boolean Matrices and Their Application to Finite Automata." Jerusalem, Hebrew University, September 3, 1960, 13 p., Technical Report No. 5, Contract No. N62558-2214.

125. __________. "Lattice Matrices and Finite Automata." Jerusalem, Hebrew University, February 1962, 44 p., Technical Report No. 8, Contract N62558-2214. ASTIA doc. no. AD-275 096.

126. Glebskii, Yu. V. "Realizable Sequences in Finite Automata." Translation of PROBLEMY KIBERNETIKI (PROBLEMS OF CYBERNETICS), Moscow, State Publishing House of Physico-Mathematical Literature, 1961 Section VIII, p. 279-282; in "Foreign Developments in Machine Translation and Information Processing, No. 69," New York, N.Y., U.S. Joint Publications Research Service, January 19, 1962, 10 p., JPRS:12048; distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

127. Glennie, A.E. "On the Syntax Machine and the Construction of a Universal Compiler." Pittsburgh, Pa., Computation Center, Carnegie Institute of Technology, July 10, 1960, 64 p. Technical Report No. 2, Contract Nonr-76018, Project NR 049-141. ASTIA doc. no. AD-240 512.

128. Gluesing, Eugene C. "Symbolic Language Translation." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, San Francisco, Calif., March 3-5, 1959), New York, N.Y., The Institute of Radio Engineers, 1959, p. 288-291.

129. Glushkov, V.M. "Abstract Automata and the Decomposition of Free Semigroups." DOKLADY AKAD. NAUK S.S.S.R., 136 (1961) 765-767. Translation in SOVIET MATHEMATICS, 2:1 (January 1961) 121-123.

130. ______________. "The Abstract Theory of Automata." USPEKHI MATEMAT. NAUK, 16:4-62, No. 5 (September-October 1961) 3-62. (In Russian.) Author's Introduction (in English) in AUTOMATION EXPRESS, 4:5 (1962) 13-14.

131. Goodstein, R.L. "On the Formalisation of Indirect Discourse." JOURNAL OF SYMBOLIC LOGIC, 23:4 (December 1958) 417-419.

132. Gorn, Saul. "On the Logical Design of Formal Mixed Languages." Section 5 of "Common Programming Language Task. Part I," by S. Gorn, J. Crozier and others. Philadelphia, Pa., Moore School of Electrical Engineering (in cooperation with Institute for Cooperative Research), University of Pennsylvania, July 31, 1959, 164 p., Report No. AD59UR1, Vol. 4, Task D, Contract DA36-039-sc-75047. ASTIA doc. no. AD-236 997.

133. __________. "The Treatment of Ambiguity and Paradox in Mechanical Languages." Philadelphia, Pa., Office of Computer Research and Education, University of Pennsylvania, April 1961, 33 p., Contract AF 49(638)951; AFOSR-603-61. ASTIA doc. no. AD-259 782.

134. ___________. "Some Basic Terminology Connected with Mechanical Languages and Their Processors, (A Tentative Base Terminology Presented to ASA 3. as a Proposal for Subsequent Inclusion in a Glossary)." COMMUNICATIONS OF THE ACM, 4:8 (August 1961) 336-339.

135. ___________. "Specification Languages for Mechanical Languages and Their Processors - A Baker's Dozen (A Set of Examples Presented to ASA X3.4 Sub-Committee)." COMMUNICATIONS OF THE ACM, 4:12 (December 1961) 532-542.

136. Gorn, Saul and E. Parker. "Common Programming Language Task, Part 1." Philadelphia, Pa., Moore School of Electrical Engineering (in cooperation with Institute for Cooperative Research), University of Pennsylvania, 1 July 1959 to 30 June 1960, 1 v., Report No. AD60UR1; Final Report on Task D, Contract No. DA-36-039-sc-75047. ASTIA doc. no. AD-248 110.

137. Gray, Harry J., Jr., Oliver Aberth, Carl Beckman, Garry T. O'Halloran, Noah S. Prywes, Jerome D. Sable, and Ralph Flood. "Information Retrieval and the Design of More Intelligent Machines." Philadelphia, Pa., Moore School of Electrical Engineering (in cooperation with Institute for Cooperative Research), University of Pennsylvania, 1 May 1958 to 30 June 1959, 200 p.; Final Report No. AD59UR1, July 31, 1959, Contract No. DA-36-039-sc-75047. ASTIA doc. no. AD-235 999.

138. Green, Bert F., Jr. "Computer Languages for Symbol Manipulation." IRE TRANSACTIONS ON HUMAN FACTORS IN ELECTRONICS, HFE-2:1 (March 1961) 3-8.

139. Green, Bert F., Alice K. Wolf, Carol Chomsky, and Kenneth Laughery. "Baseball: An Automatic Question-Answerer." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, Los Angeles, Calif., May 9-11, 1961), Western Joint Computer Conference, 1961, p. 219-224.

140. Guillaume, Marcel. "Semantic Tableaux of the Restricted Predicate Calculus." Paris, Seminaire Bourbaki, 1957-1958. Review 4902 in MATHEMATICAL REVIEWS, 21:7 (July-August 1960) 898.

140a. Harris, Zelligs. String Analysis of Sentence Structure. The Hague, Mouton and Co., 1962, 70 p.

141. Hays, David G. "Grouping and Dependency Theories." Santa Monica, Calif., The RAND Corporation, September 8, 1960, 13 p. RM-2646, Contract AF 49(638)700, Project RAND. ASTIA doc. no. AD-250 237.

142. ______________. "On the Value of a Dependency Connection." Santa Monica, Calif., The RAND Corporation, January 20, 1961, 20 p., RM-2712-AFOSR, Contract AF 49(638)737, Project 9769; AFOSR-150. ASTIA doc. no. AD-257 286.

143. ______________. "Automatic Language-Data Processing." Chapter 17 in COMPUTER APPLICATIONS IN THE BEHAVIORAL SCIENCES, edited by H. Borko, Englewood, N.J., Prentice-Hall, Ind., 1962, p. 394-423.

144. Hermes, Hans. COUNTABILITY, DECIDABILITY, COMPUTABILITY. Berlin, Springer-Verlag, 1961, 246 p.

145. Herzberger, Hans G. "The Joints of English." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, p. 99-103.

146. Hilbert, David and Wilhelm Ackermann. PRINCIPLES OF MATHEMATICAL LOGIC. New York, Chelsea Publishing Co., 1950, 172 p. (Translated from the German by Lewis M. Hammond, George G. Leckie, and F. Steinhardt.)

147. Hillman, Donald J. "The Boolean Algebra Model." Report No. 2 of "Study of Theories and Models of Information Storage and Retrieval" Bethlehem, Pa., Lehigh University, August 3, 1962, 57p. National Science Foundation Grant No. G24070.

148. ______________. "A Positive Model for Systems of Special Classification." Report No. 3 of "Study of Theories and Models of Information Storage and Retrieval," Bethlehem, Pa., Lehigh University, August 29, 1962, 20 p. National Science Foundation Grant No. G24070.

149. Hiz, Henry. "Steps Toward Grammatical Recognition." INFORMATION RETRIEVAL AND MACHINE TRANSLATION edited by Allen Kent, Vol. III (in 2 parts) of ADVANCES IN DOCUMENTATION AND LIBRARY SCIENCE, Jesse H. Shera, General Editor. New York, Interscience Publishers, Inc., 1960, p. 811-822, Part 2.

150. __________. "Congrammaticality, Batteries of Transformations and Grammatical Categories." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, p. 43-50.

151. Hodes, Louis. "Some Results from a Pattern Recognition Program Using LISP." Cambridge, Mass., Research Laboratory for Electronics and Massachusetts Institute of Technology Computation Center, Artificial Intelligence Project Memo 18, 3 p., undated.

152. ____________. "Pattern Recognition with LISP." Cambridge, Massachusetts, Research Laboratory of Electronics, Massachusetts Institute of Technology, Quarterly Progress Report No. 60, January 15, 1961, p. 166-167.

153. Holland, John H. "Cycles in Logical Nets." Ann Arbor, Michigan, College of Literature, Science, and the Arts, University of Michigan, August 1959, 61 p., Report No. 2722,2794-4-T, Contract DA-36-039-sc-78057, NSF Grant No. G-4790. ASTIA doc. no. AD-229 237.

154. Hormann, Aiko M. "Programs for Machine Learning." Santa Monica, Calif., System Development Corporation, May 29, 1962, 75 p., Technical Memorandum 669/000/01.

155. Huffman, David A. "The Design and Use of Hazard-Free Switching Networks." JOURNAL OF THE ASSOCIATION OF COMPUTING MACHINERY, 4:1 (January 1957) 47-62.

156. Ihm, P. and Y. Lecerf. "A Theory of G-Orders " Translation of section of GRISA Report No. 2, EURATOM, Group for Research on Automatic Scientific Information, May 1960; in "Foreign Developments in Machine Translation and Information Processing, No. 34, France." New York, N.Y., U.S. Joint Publications Research Service, October 11, 1961, p. 8-12, JPRS: 10367; distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

157. Ingerman, Peter Z. "Towards a Theory of Recursive Processors." PREPRINTS OF SUMMARIES OF PAPERS PRESENTED AT THE 16th NATIONAL MEETING OF THE ASSOCIATION FOR COMPUTING MACHINERY AT LOS ANGELES, CALIFORNIA, September 5-8, 1961, New York, ACM Headquarters, p. 5B-2, 1-4.

158. Irons, Edgar T. "A Syntax Directed Compiler for ALGOL 60." COMMUNICATIONS OF THE ACM, 4:1 (January 1961) 51-55.

159. Kaphengst, Heinz. "An Abstract Programmed Computer." ZEITSCHRIFT FUR MATHEMATISCHE LOGIK UND GRUNDLAGEN DER MATHEMATIK, 5:3-4 (1959) 366-379.

160. Kazakov, V.D. and O.P. Kuznetsov. "List of Foreign Works Published in 1958 on the Theory of Relay Actions and Finite Automata." AUTOMATION AND REMOTE CONTROL (The Soviet Journal AVTOMATIKA I TELEMEKHANIKA in English Translation), 21:9 (April 1961) 941-944. (Russian original dated September 1960, p. 1332.)

161. Kleene, Stephen C. INTRODUCTION TO METAMATHEMATICS. Princeton, N.J., D. Van Nostrand Co., Inc., 1952, 550 p.

162. Kleene, Stephen C. "Representation of Events in Nerve Nets and Finite Automata." AUTOMATA STUDIES, edited by C. E. Shannon and J. McCarthy, Princeton, N.J., Princeton University Press, 1956, p. 3-41.

163. Klein, Sheldon and Robert F. Simmons. "Automated Analysis and Coding of English Grammar for Information Processing Systems." Santa Monica, California, System Development Corporation, August 28, 1961, 17 p., SP-490.

164. ______________________________. "A Computational Approach to Grammatical Coding of English Words." Santa Monica, Calif., System Development Corporation, 22 February 1962, 25 p., SP-701.

165. Klingbiel, Paul H. "Language Oriented Retrieval Systems." Arlington, Va., Armed Services Technical Information Agency, February 1962, 105 p., AD-271 600.

166. Kochen, Manfred. "Adaptive Mechanisms in Digital 'Concept' processing." 1962 JOINT AUTOMATIC CONTROL CONFERENCE (Preprints of Papers Presented, New York, N.Y., June 27-29), New York, American Institute of Electrical Engineers, 1962, Session 17, p. 1-9.

167. Kreisel, Georg. "Hilbert's Programme." DIALECTICA, 12:3/4 (1958) 346-372.

168. Kripke, Saul A. "A Completeness Theorem in Modal Logic." JOURNAL OF SYMBOLIC LOGIC, 24:1 (March 1959) 1-14.

169. Kulagina, O.S., "One Method of Defining Grammatical Concepts on the Basis of the Theory of Sets." Translation of PROBLEMY KIBERNETIKI (PROBLEMS IN CYBERNETICS) Moscow, State Publishing House of Physico-Mathematical Literature, 1958, No. 1, p. 203-214; in "Soviet Developments in Machine Translation, The Automatization of Translation of One Language to Another," New York, N.Y., U.S. Joint Publications Research Service, 10 April 1959, p. 1-18, JPRS:646-D; distributed by U.S. Department of Commerce, Office of Technical Services.

_______________. PROBLEMS IN CYBERNETICS I, A.A. Lyapunov, Chief Editor; Richard Goodman, Translation Editor; London, Pergamon Press Ltd., 1960, p. 228-242.

170. Kuno, Susumu and Anthony G. Oettinger. "Multiple-Path Syntactic Analyzer." PREPRINT OF THE PROCEEDINGS OF THE IFIP CONGRESS 62, International Federation for Information Processing, Munich, August 27 to September 1, 1962, Amsterdam, North Holland Publishing Co., 1962, p. 128-133.

171. Kuroda, Sigekatu. "An Investigation on the Logical Structure of Mathematics XIII, A Method of Programming Proofs in Mathematics for Electronic Computing Machines." NAGOYA MATHEMATICAL JOURNAL 16 (February 1960) 195-203.

172. Kuznetsov, A.V. "On the Possibility of Utilizing Information-Logical Languages for a Machine Search of Theorems and Deduction of Effects." Translation of TEZISY SOVESHCHANIYA PO MATEMATICHESKOY LINGVISTIKE (ABSTRACTS OF THE CONFERENCE ON MATHEMATICAL LINGUISTICS), Leningrad, 1959 in "Soviet Developments in Information Processing and Machine Translation, Abstracts of the Conference on Mathematical Linguistics, Leningrad, 15-21 April 1959, p. 87-89, JPRS:893-D; distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

173. Kuznetsov, A.V., E.V. Paducheva, and L.M. Yermolayeva. "On Information Language for Geometry and Algorithm for Translation from Russian into Information Language." Translation of LINGVISTICHESKIYE ISSLEDOVANIYA PO MASHINNOMU PEREVODU (LINGUISTIC STUDIES ON MACHINE TRANSLATION), Moscow, All-Union Institute of Scientific and Technical Information Publishing House, 1961, Issue No. 2; in "Foreign Developments in Machine Translation and Information Processing, No. 80," New York, N.Y., U.S. Joint Publications Research Service, 27 March 1962, p. 52-98, JPRS: 13173; distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

174. Lacombe, Daniel. "The Theory of Recursive Functions and Its Applications." BULL. SOC. MATH. FRANCE 88(1960) 393-468. Review A60, MATHEMATICAL REVIEWS, 23:1A (January 1962) 8.

175. Lambek, Joachim. "The Mathematics of Sentence Structure." AMERICAN MATHEMATICS MONTHLY, 65:3 (March 1958) 154-170.

176. ________________. "On the Calculus of Syntactic Types." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, p. 166-178.

177. LeCerf, Y. "Program of Conflict - Model of Conflict." Translation of GRISA (Group for Research on Automatic Scientific Information) Report No. 4, EURATOM, August 1960, 26 p.; in "Foreign Developments in Machine Translation and Information Processing, No. 34, France." New York, N.Y., U.S. Joint Publications Research Service, October 11, 1961, p. 13-37, JPRS: 10367; distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

178. Lee, C.Y. "Automata and Finite Automata." BELL SYSTEM TECHNICAL JOURNAL, 39:5 (September 1960) 1267-1295.

179. Lees, R.B. "Automata and the Generation of Sentences" ANTHROPOLOGICAL LINGUISTICS, 1:4 (1959) 1-4.

179a. Lees, Robert B. "The Grammar of English Nominalizations," INTERNATIONAL JOURNAL OF AMERICAN LINGUISTICS, 26:3 Part II (July 1960) 205 p.

180. Lehmer, D.H., E. Lehmer, W.H. Mills, and J. L. Selfridge. "Machine Proof of a Theorem on Cubic Residues." MATHEMATICS OF COMPUTATION, 16:80 (October 1962) 407-415.

181. Levien, R.E. "Studies in the Theory of Computational Algorithms - 1. Formalization, Computability, Representation, and Analysis Problems." Santa Monica, Calif., The RAND Corporation, January 1962, lv., RM-3007-PR.

182. Lindsay, Robert K. "Inferential Memory as the Basis of Machines which Understand Natural Language." Chapter in forthcoming book COMPUTERS AND THOUGHT, Vol. I, ARTIFICIAL INTELLIGENCE, edited by E. Feigenbaum and J. Feldman.

183. Lukasiewicz, Jan. ARISTOTLE'S SYLLOGISTIC FROM THE STANDPOINT OF MODERN FORMAL LOGIC. Oxford, Clarendon Press, 1957, 2nd edition, 222 p.

184. Lyapunov, A.A. "Certain General Problems in Cybernetics." PROBLEMS IN CYBERNETICS I, A.A. Lyapunov, Chief Editor; Richard Goodman, Translation Editor; London, Pergoman Press Ltd., 1960, p. 1-21.

185. Lywbpenko, G.G. "Methods for Determining the Identical Truth or Falsity of the Calculus of Assertion Formulae of Bivalent Logic." DOPOVIDI AKAD. NAUK UKRAIN. RSR (1958) 1153-1156, Abstract 647 in MATHEMATICAL REVIEWS 21:2 (February 1960) 119.

186. Markov, A.A. THEORY OF ALGORITHMS. Works of the Mathematical Institute imeni V. A. Steklov, Vol. 42, Moscow, Academy of Sciences of the USSR, 1954, 444 p. Translated from the Russian; published for the National Science Foundation and the Department of Commerce by the Israel Program for Scientific Translations, 1961, OTS 60-51085.

187. Maron, M.E. "Handling of Non-Numerical Information." Chapter 11 in HANDBOOK OF AUTOMATION, COMPUTATION, AND CONTROL, edited by E.M. Grabbe, S. Ramo and D.E. Wooldridge,Vol. 2, COMPUTERS AND DATA PROCESSING, New York, John Wiley and Sons, Inc., 1959, p. 11-01 - 11-20.

188. Martin, R.M. "A Formalization of Inductive Logic." JOURNAL OF SYMBOLIC LOGIC, 23:3 (September 1958) 251-256.

189. McCarthy, John. "The Wang Algorithm for the Propositional Calculus Programmed in LISP." Cambridge, Mass., Research Laboratory for Electronics and Massachusetts Institute of Technology Computation Center, undated, 12 p., Memo 14, Artificial Intelligence Project.

190. ________________. "Programs with Common Sense." MECHANISATION OF THOUGHT PROCESSES (Proceedings of a Symposium held at the National Physical Laboratory on 24th, 25th, 26th and 27th November 1958), London, Her Majesty's Stationery Office, 1959, Volume I, p. 75-84.

191. ________________. "A Proposal for Solving the 'Baseball Problem' in LISP." Lexington, Mass., Lincoln Laboratory, Massachusetts Institute of Technology, 8 December 1959, 8 p., Group Report 54-11.

192. ________________. "Recursive Functions of Symbolic Expressions and Their Computation by Machine, Part I." COMMUNICATIONS OF THE ACM, 3:4 (April 1960) 184-195.

193. ________________. "A Basis for a Mathematical Theory of Computation, Preliminary Report." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, Los Angeles, Calif., May 9-11, 1961), Western Joint Computer Conference, 1961, p. 225-238.

194. McNaughton, Robert. "Symbolic Logic and Automata." Philadelphia, Pa., Moore School of Electrical Engineering, University of Pennsylvania, July 1960, 31 p., WADD Technical Note 60-244, Contract AF33 (616) 5886. ASTIA doc. no. AD-249 737.

195. ____________________. "The Theory of Automata, A Survey." ADVANCES IN COMPUTERS, Vol. 2, edited by F. L. Alt, New York, Academic Press, 1961.

196. McNaughton, Robert and H. Yamada. "Regular Expressions and State Graphs for Automata." IRE TRANSACTIONS ON ELECTRONIC COMPUTERS, EC-9:1 (March 1960) 39-47.

197. Medvedev, I. T. "On a Class of Events Representable in a Finite Automaton." Lexington, Mass., Lincoln Laboratory, Massachusetts Institute of Technology, June 30, 1958, 20 p. Group Report 34-73. Translation by Jacques J. Schorr-Kon of an article which appeared as one of two supplements to the Russian translation of AUTOMATA STUDIES edited by C. Shannon and J. McCarthy and published in early 1956. The Russian version, edited by A. A. Liapanov, appeared the same year and was published by the Publishing Agency for Foreign Literature, Moscow.

198. Miller, George A. and Noam Chomsky. "Pattern Conception." Paper presented at the University of Michigan Conference on Pattern Recognition, Ann Arbor, October 1957, 12 p.; AFCRC-TN-57-57.

199. ____________________. "Finitary Models of Language Users." To appear in Vol. 2 of HANDBOOK OF MATHEMATICAL PSYCHOLOGY edited by R. R. Bush, E. H. Galanter, and R. D. Luce, New York, John Wiley and Sons, Inc., 1962.

200. Minsky, Marvin. "Some Methods of Artificial Intelligence and Heuristic Programming." MECHANISATION OF THOUGHT PROCESSES (Proceedings of a Symposium held at the National Physical Laboratory on 24th, 25th, 26th and 27th November 1958), London, Her Majesty's Stationery Office, 1959, Volume I, p. 3-27.

201. ____________________. "Recursive Unsolvability of Post's Problem of 'Tag' and Other Topics in the Theory of Turing Machines." Cambridge, Mass., Lincoln Laboratory, Massachusetts Institute of Technology, June 9, 1960, Report No. 54G-0023.

202. Minsky, Marvin. "Steps Toward Artificial Intelligence." PROCEEDINGS OF THE IRE, 49:1 (January 1961) 8-30.

203. ________________. "A Selected, Descriptor-Indexed Bibliography to the Literature on Artificial Intelligence." IRE TRANSACTIONS ON HUMAN FACTORS IN ELECTRONICS, HFE-2:1 (March 1961) 39-55.

204. ________________. "Descriptive Languages and Problem Solving," PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, Los Angeles, Calif., May 9-11, 1961), Western Joint Computer Conference, 1961, p. 215-218.

205. ________________. "Physical Machines and Their Abstract Counterparts." Draft of part of Chapter II of a forthcoming book on artificial intelligence, 116 p., undated.

206. Mitchell, R. P. Sunnyvale, Calif., Lockheed Aircraft Corporation, Contract AF 30 (602)-1889:

______________. "Word Correlation Study II," 8 p., undated.

______________. "Word Correlation Study II - Supplement," 4 p., undated.

______________. "Word Correlation Study III," 10 p., undated.

______________. "Word Correlation Study." Final Report, 19 p., September 1959, LMSD-311696.

207. Mitchell, R. P. "On Properties of a Class of Categorial Grammars." Sunnyvale, Calif., Lockheed Aircraft Corporation, 55 p., undated.

208. ______________. "A Note on Categorial Grammars." Sunnyvale, Calif., Lockheed Aircraft Corporation, August 1961, 31 p., Report No. 6-90-61-21-A. ASTIA doc. no. AD-264 997.

209. Moore, E.F. "Gedanken-Experiments on Sequential Machines." AUTOMATA STUDIES, edited by C.E. Shannon and J. McCarthy, Princeton, N.J., Princeton University Press, 1956, p. 129-153.

210. More, Trenchard, Jr. "A Unified Propositional Calculus. Part of Deductive Logic for Automata." JOURNAL OF SYMBOLIC LOGIC, 23:1 (March 1958) 103. (Abstract of paper presented at the twenty-second annual meeting of the Association for Symbolic Logic held on December 27, 1957, at Harvard University, Cambridge, Massachusetts in conjunction with the fifty-fourth annual meeting of the Eastern Division of the American Philosophical Association.)

211. ____________________. "Procedures of Proof in Implicationally Ordered Formal Systems." Cambridge, Mass., Massachusetts Institute of Technology, High Speed Computer System Research, Quarterly Progress Report No. 9, January 31, 1960, p. 21-43.

212. Mullin, A.A. "Some Recursive Function Theory and Certain of Its Related Topics." Urbana, University of Illinois, privately published, October 1959, 81 p.

213. Myhill, J. "Finite Automata and the Representation of Events." Section V in "Fundamental Concepts in the Theory of Systems." Chicago, Illinois, Systems Research, U. of Chicago, November 1957, p. 112-137, WADC Technical Report 57-624, Contract No. AF 33(616)-2797, Project 7060 ASTIA doc. no. AD-155 741.

214. __________. "Linear Bounded Automata." Philadelphia, Pa., Moore School of Electrical Engineering, University of Pennsylvania, June 1960, 28 p., Report No. 60-22, Contract AF 33(616) 5886, Project 7062; WADD TN 60-165. ASTIA doc. no. AD-246 595.

215. __________. "Notes for a Series of Lectures on Recursive Functions." PARALLEL COMPUTERS, AUTOMATA AND ADAPTIVE SYSTEMS, Ann Arbor, The University of Michigan Engineering Summar Conferences, Summer 1962, 12 p.

216. Nagornyi, N.M. "Certain Generalization of the Notions of Normal Algorithm." TRUDY MATEMATICHESKOGO INSTITUTA imeni V.A. STEKLOVA, 52 (1958) 7-65. (Collected Works, edited by N.A. Shapin, of the V.A. Steklov Mathematical Institute "Problems of Constructive Directions in Mathematics 1," Moscow and Leningrad, Academy of Sciences, U.S.S.R.)

217. Naur, Peter (Editor). "Report on the Algorithmic Language ALGOL 60." COMMUNICATIONS OF THE ACM, 3:5 (May 1960) 299-314.

218. Nerode, A. "Continuous Mechanisms." Section III in "Fundamental Concepts in the Theory of Systems." Chicago, Illinois, Systems Research, University of Chicago, November 1957, p. 64-90, WADC Technical Report 57-624, Contract No. AF 33(616)-2797, Project 7060," ASTIA doc. no. AD-155 741.

219. __________. "Linear Automaton Transformations." PROCEEDINGS OF THE AMERICAN MATHEMATICAL SOCIETY, 9:4 (August 1958) 541-544.

220. Netherwood, Douglas B. "Logical Machine Design: A Selected Bibliography." IRE TRANSACTIONS ON ELECTRONIC COMPUTERS, EC-7:2 (June 1958) 155-178; corrections in EC-7:3 (September 1958) 250.

__________. "Logical Machine Design II. A Selected Bibliography." _Ibid._, EC-8:3 (September 1959) 367-380.

221. Newell, Allen and J. C. Shaw. "Programming the Logic Theory Machine." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, Los Angeles, Calif., February 26-28, 1957), New York, N.Y., The Institute of Radio Engineers, 1957, p. 230-240.

222. Newell, Allen and Herbert A. Simon. "The Logic Theory Machine, A Complex Information Processing System." IRE TRANSACTIONS ON INFORMATION THEORY, IT-2:3 (September 1956) 61-79.

223. Newell, Allen and F. M. Tonge. "An Introduction to Information Processing Language V." COMMUNICATIONS OF THE ACM, 3:4 (April 1960) 205-211.

224. Newell, Allen, J.C. Shaw, and Herbert A. Simon. "Empirical Explorations of the Logic Theory Machine: A Case Study in Heuristics." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, Los Angeles, Calif., February 26-28, 1957), New York, N.Y., The Institute of Radio Engineers, 1957, p. 218-230.

225. __________. "A Variety of Intelligent Learning in a General Problem Solver." SELF-ORGANIZING SYSTEMS (Proceedings of an Interdisciplinary Conference, May 1959) edited by Marshall C. Yovits and Scott Cameron, New York, Pergamon Press, 1960, p. 153-189.

226. Newell, Allen, J.C. Shaw and H. A. Simon. "Report on a General Problem-Solving Programme." INFORMATION PROCESSING, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING, UNESCO, PARIS, 15-20 June 1959, edited by S. de Picciotto, Paris, UNESCO; München, R. Oldenbourg; London, Butterworths, 1960, p. 256-264.

227. Nugent, William R. and Alexander Vegh. "Automatic Word Coding Techniques for Computer Language Processing." Lexington, Massachusetts, Itek Laboratories, February 1962, 104 p., Appendix, Report No. IL 9018-1 (Final), Contract AF 30(602)-2377; RADC-TDR-62-13, Vol. I. ASTIA doc.no. AD-272 401.

228. Oettinger, A.G. "Automatic Syntactic Analysis and the Push-down Store." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, p. 104-129.

229. Oettinger, A.G. and M.E. Sherry. "Current Research on Automatic Translation at Harvard University and Predictive Syntactic Analysis." PROCEEDINGS OF THE NATIONAL SYMPOSIUM ON MACHINE TRANSLATION (Held at the University of California, Los Angeles, February 2-5, 1960, edited by H. P. Edmundson. Englewood Cliffs, N.J. Prentice-Hall, 1961, p. 173-182.

230. Parikh, R.J. "Language-Generating Devices." Cambridge, Massachusetts, Research Laboratory of Electronics, Massachusetts Institute of Technology, Quarterly Progress Report No. 60, January 15, 1961, p. 199-212.

231. Patterson, G.W. "Logical Syntax and Transformation Rules." PROCEEDINGS OF A SECOND SYMPOSIUM ON LARGE-SCALE DIGITAL CALCULATING MACHINERY, (The Annals of the Computation Laboratory of Harvard University, Volume XXVI), Cambridge, Mass., Harvard University Press, 1951, p. 125-133.

232. _______________. "What is a Code?" COMMUNICATIONS OF THE ACM, 3:5 (May 1960) 315-318.

233. Paz, Azaria. "Homomorphisms Between Finite Automata." BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL, 10F:3 (February 1962) 92-100.

234. Perlis, A.J. and K. Samelson. "Preliminary Report - International Algebraic Language." COMMUNICATIONS OF THE ASSOCIATION FOR COMPUTING MACHINERY, 1:12 (December 1958) 8-22.

235. Perlis, A. J. and Charles Thornton. "Symbol Manipulation by Threaded Lists." COMMUNICATIONS OF THE ACM, 3:4 (April 1960) 195-204.

236. Perlis, A. J., J. W. Smith and H. R. Van Zoeren. "Internal Translator (IT), A Compiler for the 650." Pittsburgh, Pa., Computation Center, Carnegie Institute of Technology, 1 v., undated, 650 Library Program File Number 2.1.001.

237. Peter, Rozsa. "Rekursive Funktionen." Budapest, Akademiai Kiado, 1951.

238. Plath, Warren. "Mathematical Linguistics." Chapter in TRENDS IN EUROPEAN AND AMERICAN LINGUISTICS, edited by C. Mohrmann, A. Sommerfelt and J. Whatmough, Utrecht, Spectrum Publishers, 1961.

239. Porte, Jean. "A Simplification of the Theory of Turing." PROCEEDINGS, FIRST INTERNATIONAL CONGRESS ON CYBERNETICS, NAMUR, June 26-29, 1956. Paris, Gauthier-Villars, 1958, p. 251-280. (In French)

240. __________. "Post Systems, Markov Algorithms." CYBERNETICA, 1:2 (1958) 114-149. (In French)

241. Post, Emil L. "Recursively Enumerable Sets of Integers and their Decision Problems." BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY, 50:5 (May 1944) 284-316.

242. Prawitz, Dag. "An Improved Proof Procedure." THEORIA (A SWEDISH JOURNAL OF PHILOSOPHY AND PSYCHOLOGY), 26 (1960) 102-139.

243. Prawitz, Dag, Hakan Prawitz, and Neri Voghera. "A Mechanical Proof Procedure and its Realization in an Electronic Computer." JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY, 7:2, (April 1960) 102-128.

244. Putnam, Hilary. "Some Issues in the Theory of Grammar." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, p. 25-42.

245. Quine, W. "A Proof Procedure for Quantification Theory." JOURNAL OF SYMBOLIC LOGIC, 20:2 (June 1955) 141-149.

246. Rabin, M.O. "Two-Way Finite Automata." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957, printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N.J., 2nd edition, 25 July 1960. p. 366-369.

247. ___________. "Degree of Difficulty of Computing a Function and a Partial Ordering of Recursive Sets." Jerusalem, Hebrew University, April 25, 1960, 18 p., Technical Report No. 2, Contract No. 62558-2214. ASTIA doc. no. AD-250 784.

248. Rabin, M. and D. Scott. "Remarks on Finite Automata." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N.J., 2nd edition, 25 July 1960, p. 106-112.

249. ______________________. "Finite Automata and Their Decision Problems." IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 3:2 (April 1959) 114-125.

250. Rado, T. "On Non-Computable Functions." BELL SYSTEM TECHNICAL JOURNAL, 41:3 (May 1962) 877-884.

251. Randolph, John W. "The Theory of Finite Automate." Princeton, N.J., Digital Systems Laboratory, Princeton University, May 1961, 50 p., Technical Report No. 6. (To be submitted to the Department of Mathematics, Princeton University, in partial fulfillment of the requirements for the degree of Bachelor of Arts.)

252. Rankin, B. Kirk, III. "A Programmable Grammar for a Fragment of English for Use in an Information Retrieval System." Washington, D.C., National Bureau of Standards, June 30, 1961, 16 p. (Unpublished Report).

253. Resnik, Michael D. "The Formalization of the Foundations of Geometry Within the System of Symbolic Logic R." Bedford, Mass., Air Force Cambridge Research Laboratories, September 1961, 115 p., Report No. AFCRL-704, Project 5632. ASTIA doc. no. AD-264 864.

254. Revzin, I. I. "On Some Notions of the So-Called Set-Theory Conception of Language." Translation of VOPROSY YAZYKOZNANIYA (PROBLEMS OF LINGUISTICS), No. 6, November-December 1960, p. 88-94; in "Foreign Developments in Machine Translation and Information Processing, No. 25," New York, N.Y., U.S. Joint Publications Research Service, 17 March 1961, p. 1-10, JPRS:6938; distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

255. Ritchie, Robert W. "Classes of Recursive Functions of Predictable Complexity." Princeton, N.J., Princeton University Dissertation, 1960, 85 p., appendices.

________________. DISSERTATION ABSTRACTS XXII:7 (January 1962) 2407.

256. ________________. "The Functions Computable by Linear Bounded Automata." AMERICAN MATHEMATICAL SOCIETY NOTICES, 8:2, Issue 53 (April 1961) 145. Abstracts (578-20) of Contributed Papers, the April Meeting in New York, New York, April 5-8, 1961.

257. Robinson, Abraham. "Relative Model-Completeness and the Elimination of Quantifiers." SUMMARIES OF TALKS PRESENTED AT THE SUMMER INSTITUTE OF SYMBOLIC LOGIC, CORNELL UNIVERSITY 1957; printed and distributed by Communications Research Division, Institute of Defense Analysis, Princeton, N.J., 2nd edition, 25 July 1960, p. 155-159.

258. ________________. "Proving a Theorem (As Done by Man, Logician, or Machine)." Ibid., p. 350-352.

259. ________________. "On the Mechanization of the Theory of Equations." BULLETIN OF THE RESEARCH COUNCIL OF ISRAEL, 9F:2 (November 1960) 47-70.

260. Robinson, John A. "Gamma I; A General Theorem-Proving Program for the IBM 704." Chicago University, Illinois, November 1961, 46 p., AEC Research and Development Report ANL-6447, Contract W-31-109-eng-38.

261. Rochester, N., S.H. Goldberg and D. J. Edwards. "Machine Manipulation of Algebraic Expressions." Cambridge, Mass., Research Laboratory of Electronics, Massachusetts Institute of Technology, Quarterly Progress Report No. 55, October 15, 1959, p. 132-134.

262. Rogers, Hartley, Jr. "Theory of Recursive Functions and Effective Computability." Cambridge, Mass., Department of Mathematics, Massachusetts Institute of Technology, 1957, 155 p., Volume I.

263. ________________. "The Present Theory of Turing Machine Computability." JOURNAL OF THE SOCIETY FOR INDUSTRIAL AND APPLIED MATHEMATICS, 7:1 (March 1959) 114-130.

264. Rose, Gene F. "Output Completeness in Sequential Machines." Santa Monica, Calif., System Development Corporation, February 6, 1961, 7 p., Technical Memorandum TM-584.

265. Sable, J. D. "Use of Semantic Structure in Information Systems." IRE TRANSACTIONS ON ENGINEERING WRITING AND SPEECH, EWS-4:3 (December 1961) 83-90.

265a. Sammet, Jean E., "Basic Elements of COBOL 61." COMMUNICATIONS OF THE ACM, 5:5 (May 1962) 237-253.

266. Scheinberg, S. "Note on the Boolean Properties of Phrase Structure Grammars." INFORMATION AND CONTROL, 3:4 (December 1960) 372-375.

267. Schutzenberger, M. P. "On an Application of Semi-Groups Methods to Some Problems in Coding." IRE TRANSACTIONS ON INFORMATION THEORY, IT-2:3 (September 1956) 47-60.

268. ________________. "On the Definition of a Certain Class of Automata." Chapel Hill, N.C., Institute of Statistics, University of North Carolina, January 1961, 34 p., Mimeograph Series No. 273; AFOSR 87.

269. ________________. "On a Special Class of Recurrent Events." Chapel Hill, N.C., Institute of Statistics, University of North Carolina, June 1961, 15 p., Mimeograph Series No. 291; AFOSR 891. ASTIA doc. no. AD-260 711.

270. ________________. "A Remark on Finite Transducers." INFORMATION AND CONTROL, 4:2-3 (September 1961) 185-196.

271. ________________. "On the Definition of a Family of Automata." Ibid. 245-270.

272. ________________. "Some Remarks on Chomsky's Context-Free Languages." Cambridge, Mass., Research Laboratory of Electronics, Massachusetts Institute of Technology, Quarterly Progress Report No. 63, October 15, 1961, p. 155-170.

273. Schützenberger, M. P. and N. Chomsky. "The Algebraic Theory of Context-Free Languages." Mimeograph, undated, 63p., footnotes, references.

274. Semarne, H. M. "Symbolic Logic in Language Engineering." PROCEEDINGS OF THE WESTERN JOINT COMPUTER CONFERENCE (Papers Presented at the Joint IRE-AIEE-ACM Computer Conference, San Francisco, Calif., May 3-5, 1960), Western Joint Computer Conference, 1960, p. 61-71.

275. Shamir, E. "On Sequential Languages." Jerusalem, Hebrew University, November 1961, 18 p., Technical Report No. 7, Contract No. N62558-2214.

276. Shannon, Claude E. "Von Neumann's Contributions to Automata Theory." BULLETIN OF THE AMERICAN MATHEMATICAL SOCIETY, 64:3 Part 2 (May 1958) 123-129.

277. Shepherdson, J.C. "The Reduction of Two-Way Automata to One-Way Automata." IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 3:2 (April 1959) 198-200.

278. Sherry, M.E. and A. G. Oettinger. "A New Model of Natural Language for Predictive Syntactic Analysis." INFORMATION THEORY (Papers read at a Symposium on Information Theory held at the Royal Institution, London, August 29th to September 2nd, 1960), London, Butterworth and Co., Ltd., 1961, p. 458-468.

279. Simmons, P. L. and R. F. Simmons. "The Simulation of Cognitive Processes: An Annotated Bibliography." IRE TRANSACTIONS ON ELECTRONIC COMPUTERS, EC-10:3 (September 1961) 462-483.

280. Simmons, R. F., S. Klein and K. McConlogue. "Towards the Synthesis of Human Language Behavior." Santa Monica, Calif., System Development Corporation, 27 September 1961, 15 p., SP-466.

281. Slagle, James R. "A Heuristic Program that Solves Symbolic Integration Problems in Freshman Calculus, Symbolic Automatic Integrator (SAINT)." Cambridge, Massachusetts, Lincoln Laboratory, Massachusetts Institute of Technology, May 10, 1961, 110 p., Report No. 5G-0001.

282. Smullyan, Raymond M. THEORY OF FORMAL SYSTEMS. Princeton, N.J., Princeton University Press, 1961, 142 p.

283. Solomonoff, R.J. "An Inductive Inference Machine." 1957 IRE NATIONAL CONVENTION RECORD, Vol. 5, Part 2 - Circuit Theory; Information Theory, p. 56-62.

284. ______________. "A New Method for Discovering the Grammars of Phrase Structure Languages." Cambridge, Mass., Zator Co., April 1959, 13 p., Report No. ZTB-124, Contract AF 49(638)376; AFOSR TN-59-110; ASTIA doc. no. AD-210 390.

______________. INFORMATION PROCESSING, PROCEEDINGS OF THE INTERNATIONAL CONFERENCE ON INFORMATION PROCESSING, UNESCO, PARIS, 15-20 June 1959, edited by S. de Picciotto, Paris, UNESCO; Munchen, R. Oldenbourg; London, Butterworths, 1960, p. 282-290.

285. ______________. "The Mechanization of Linguistic Learning." Cambridge, Mass., Zator Co., April 1959, 16 p., Report No. ZTB-125, Contract AF 49(638)376; AFOSR TN-59-246. ASTIA doc. no. AD-212 226.

286. ______________. "On Machines to Learn to Translate Languages and Retrieve Information." Cambridge, Mass., Zator Co., October 1959, 17 p., Report No. ZTB-134, Contract AF 49(638)376; AFOSR TN-59-646; ASTIA doc. no. AD-228 203.

287. ______________. "Progress Report, Research in Inductive Inference for the Period 1 April 1959 to 30 November 1960." Cambridge, Mass., Zator Co., January 1961, 18 p., Report No. ZTB 139, Contract AF 49(638)376; AFOSR 160. ASTIA doc. no. AD-252 446.

288. ______________. "A General Theory of Inductive Inference. Preliminary Report." Cambridge, Mass., Zator Co., November 1960, 26 p., Report No. ZTB-138, (Revision of Report No. V-131), Contract AF 49(638)376; AFOSR TN 60-1459. ASTIA doc. no. AD-252 445.

289. ______________. "A Coding Method for Inductive Inference." Cambridge, Mass., Zator Co., April 1961, 11 p., Report No. ZTB-140, Contract AF 49(638)376; AFOSR-510. ASTIA doc no. AD-256 886.

290. Solomonoff, R. J. "A Formal Theory of Inductive Inference." Cambridge, Mass., Zator Co., May 1962, 76 p., Report No. V-143.

291. Suryani, Janos. "Reduction Theory of the Decision Problem in Predicate Calculus and Other Subjects." Budapest, Verlag der Ungarischen Akademie der Wissenschaften, 1959, 215 p.

292. Tarjan, R. "On the Instrumentation of Logical Problems." ACTA TECH. ACAD. SCI. HUNGAR. 27(1959) 371-383. Abstract 3123 in MATHEMATICAL REVIEWS, 22:4B (April 1961) 539.

293. Tennenbaum, S. "Digital Processes." Section 1 of "Fundamental Concepts in the Theory of Systems." Chicago, Illinois, Systems Research, University of Chicago, November 1957, p. 1-32, WADC Technical Report 57-624, Contract No. AF 33(616)-2797, Project 7060; ASTIA doc. no. AD-155 741.

294. Trakhtenbrot, B.A. ALGORITHMS AND MACHINE SOLUTION OF PROBLEMS. Moscow, Gosudarstv. Izdat. Fiz.-Mat. Lit., 1960, 119 p. (In Russian.) Review 10906 in MATHEMATICAL REVIEWS, 22:11A (November 1961) 1855.

295. Travis, Larry E. "Analytic Information Retrieval." In NATURAL LANGUAGE AND THE COMPUTER, edited by P. Gavin, to be published by McGraw-Hill, New York, N.Y.

296. Tsetlin, M. L. "On the Behavior of Finite Automata in Random Media." AUTOMATION AND REMOTE CONTROL (The Soviet Journal AVTOMATIKA I TELEMEKHANIKA in English Translation), 22:10 (March 1962) 1210-1219. (Russian original dated October 1961, p. 1345.)

297. Turing, A. M. "On Computable Numbers with an Application to the Entscheidungs-problem." PROCEEDINGS OF THE LONDON MATHEMATICAL SOCIETY, Series 2, 42:3 (November 30, 1936) 230-265.

____________. "A Correction." Ibid., 43:7 (December 30, 1937) 544-546.

298. Uspenskii, V.A. "The Problem of Constructing a Machine Language for an Information Machine." Translation of PROBLEMY KIBERNETIKI (PROBLEMS OF CYBERNETICS), No. 2, 1959, p. 39-50; in "Soviet Developments in Information Processing and Machine Translation," New York, N.Y., U.S. Joint Publications Research Service, 9 February 1960, p. 1-17, JPRS:1145-D; distributed by U.S. Department of Commerce, Office of Technical Services, Washington, D.C.

299. Uspenskii, V.A. LECTURES ON COMPUTABLE FUNCTIONS. Moscow, Gosudarstvennoe Isdatel'stvo Fixiko-Matematicheskoi Literatury, 1960, 492 p. (In Russian).

300. Von Neumann, John. "The General and Logical Theory of Automata." CEREBRAL MECHANISMS IN BEHAVIOR; THE HIXON SYMPOSIUM, edited by L. A. Jeffress, New York, John Wiley and Sons, 1951, p. 1-31.

301. ________________. "Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components." AUTOMATA STUDIES, edited by C. E. Shannon and J. McCarthy, Princeton, N.J., Princeton University Press, 1956, p. 43-98.

302. ________________. THE COMPUTER AND THE BRAIN. New Haven, Conn., Yale University Press, 1958, 83 p.

303. Wang, Hao. "A Variant to Turing's Theory of Computing Machines." JOURNAL OF THE ASSOCIATION FOR COMPUTING MACHINERY, 4:1 (January 1957) 63-92.

304. __________. "The Axiomatization of Arithmetic." JOURNAL OF SYMBOLIC LOGIC, 22:2 (June 1957) 145-158.

305. __________. "Toward Mechanical Mathematics." IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 4:1 (January 1960) 2-22.

306. __________. "Proving Theorems by Pattern Recognition - I." COMMUNICATIONS OF THE ACM, 3:4 (April 1960) 220-234.

307. __________. "Computers as an Aid to Theorem-Proving." Paper, Symposium on the Interactions Between Mathematical Research and High Speed Computing, Five Hundred Ninetieth Meeting, American Mathematical Society, Atlantic City, April 16-19, 1962.

308. __________. "Proving Theorems by Pattern Recognition - II." BELL SYSTEM TECHNICAL JOURNAL, 40:1 (January 1961) 1-41.

309. Watanabe, Satosi. "Information-Theoretical Aspects of Inductive and Deductive Inference." IBM JOURNAL OF RESEARCH AND DEVELOPMENT, 4:2 (April 1960) 208-231.

310. Wegstein, J.H. and W. W. Youden. "A String Language for Symbol Manipulation Based on ALGOL 60." COMMUNICATIONS OF THE ACM, 5:1 (January 1962) 54-61.

311. Wells, Mark B. "MADCAP: A Scientific Compiler for a Displayed Formula Textbook Language." COMMUNICATIONS OF THE ACM, 4:1 (January 1961) 31-36.

312. Williams, Thyllis. "Translating from Ordinary Discourse into Formal Logic, A Preliminary Systems Study." Alexandria, Va., ACF Industries, Inc., September-November 1956, 110 p., Scientific Report, Contract AF 19(604) 1582; AFCRC-TN-56-770. ASTIA doc. no. AD-98 813.

313. Williams, T.M., R. F. Barnes and J. W. Kuipers. "Discussion of Major Features of a Restricted Logistic Grammar for Topic Representation." Lexington, Massachusetts, Itek Laboratories, February 1962, 23 p., Report No. IL 5206-26.

314. Wolpe, Harold. "Algorithm for Analyzing Logical Statements to Produce a Truth Function Table" COMMUNICATIONS OF THE ASSOCIATION FOR COMPUTING MACHINERY, 1:3 (March 1958) 4-13.

315. Yamada, Hisao. "Counting by a Class of Growing Automata." DISSERTATION ABSTRACTS, XXI:4 (October 1960) 833.

316. Yanov, Yu I. "The Logical Schemes of Algorithms" PROBLEMS IN CYBERNETICS I, A.A. Lyapunov, Chief Editor; Richard Goodman, Translation Editor; London, Pergamon Press, Ltd., 1960, p. 82-140.

317. Yngve, Victor H. "A Model and an Hypothesis for Language Structure." PROCEEDINGS OF THE AMERICAN PHILOSOPHICAL SOCIETY, 104:5 (October 1960) 444-466.

_______________. Cambridge, Mass., Research Laboratory of Electronics, Massachusetts Institute of Technology, August 15, 1960, Technical Report 369.

318. _______________. "The Depth Hypothesis." STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS (Proceedings of Symposia in Applied Mathematics, Volume XII), Providence, R.I., American Mathematical Society, 1961, p. 130-138.

319. _______________. "COMIT as an IR Language." COMMUNICATIONS OF THE ACM, 5:1 (January 1962) 19-28.

320. Yngve, Victor H. "Computer Programs for Translation." SCIENTIFIC AMERICAN, 206:6 (June 1962) 68-76.

321. _______________. "Random Generation of English Sentences." PROCEEDINGS OF THE 1961 INTERNATIONAL CONGRESS ON MACHINE TRANSLATION OF LANGUAGES AND APPLIED LANGUAGE ANALYSIS. Teddington, England, National Physical Laboratory, 1962 (in press).