[*International Conference on Machine Translation of Languages and Applied Language Analysis*,
National Physical Laboratory, Teddington, UK, 5-8 September 1961]

AUTOMATIC SENTENCE DIAGRAMMING *

by

WARREN PLATH

The Computation Laboratory of Harvard
University Cambridge, Massachusetts

1. INTRODUCTION

AS the reliability of techniques for automatic syntactic analysis has increased, the problem of how to represent the syntactic structure of analysed sentences has taken on considerable practical significance. Ideally, the output format of a syntactic analysis routine should be designed to realize two major objectives. The first is that it should contribute to the rapid and accurate evaluation of results, with a premium on the detection of residual errors and on the clear representation of the structures identified. The second objective relates not to the improvement of the analysis routine, but to the intended applications of the data which it provides. In this latter connection, it is highly desirable that any Information which will be required during further machine processing should be present in the output in a completely explicit, easily accessible form.

The experimental work described in this paper has led to the development of automatic techniques for producing structural diagrams of sentences. The diagramming program operates on Russian sentences which have been subjected to predictive syntactic analysis (Sherry, Bossert, 1960), reducing them to a predetermined canonical form.

Simple machine editing then produces sentence diagrams which lend themselves to rapid visual interpretation. Besides contributing to the evaluation of syntactic analyses, the diagrams show considerable promise as a point of departure for a variety of applications involving language data processing.

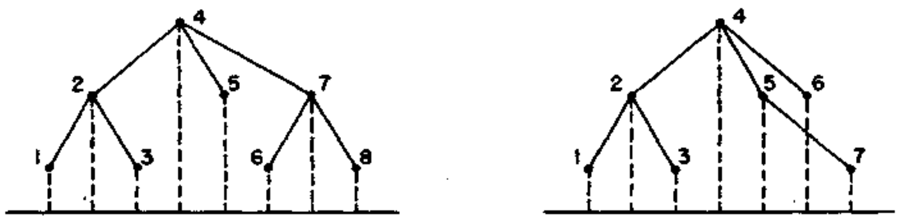
In order to explain the principles upon which the automatic diagramming system is based, it will prove convenient to consider first certain theoretical observations regarding two systems for the representation of sentence structure.

* This study has been supported in part by the National Science Foundation.

2. TREE STRUCTURE AND PARENTHETIC GROUPING

Le Cerf and Ihm (1960) have proved that for a certain class of formally defined languages, called ***projective languages***, it is always possible to construct structural tree diagrams which have the following properties: there is a one-to-one correspondence between the nodes of a given tree and the words of the corresponding sentence such that (a) nodes are ordered from left to right in the tree in sentence word order, (b) no two branches of the tree cross one another and (c) vertical projections from each node onto a horizontal line drawn below the tree do not intersect any of the branches of the tree. Another important criterion of projectivity as defined by Le Cerf and Ihm is that the projection of the ***domain*** of each word (the part of the tree which hangs down from the corresponding node, including the node itself) is a continuous line segment, corresponding to a continuous piece of linear text. (Hays [1960] uses the term "complete subtree" when referring to the domain of a word.)

Some of the properties of tree diagrams associated with projective and non-projective sentences can be illustrated with reference to ***Figure 1***. For the projective sentence whose diagram appears in (a), the domain of word 2 is the continuous string of words 1, 2, 3; the domain of word 4 is the entire sentence and the domain of word 5 is just word 5 itself. Using ***figure 1b***, the relationship between the projectivity of a sentence tree and the requirement that all word domains be continuous can be demonstrated by considering words 5, 6 and 7: Word 6 is not in the domain of word 5, so



TREE OF A PROJECTIVE SENTENCE (a) TREE OF A NON-PROJECTIVE SENTENCE (b)

(THE NUMBERED NODES CORRESPOND TO WORDS OF THE SENTENCE IN NORMAL TEXT ORDER. SOLID LINES CONNECTING NODES INDICATE GRAMMATICAL DEPENDENCY OF THE CORRESPONDING WORDS.)

Figure 1

that the domain of word 5 is clearly discontinuous. As drawn, the projection of node 6 crosses the branch (5, 7) in violation of the rules for projective sentences. The only other way in which the branch (4, 6) can be drawn, without intersecting another branch or without changing the left-to-right ordering of nodes 5, 6 and 7, is to place it below the branch (4, 5, 7.). However, this latter configuration also violates the conditions for projectivity, this time because the projection of node 5 crosses the branch (4, 6). It is thus evident that a sentence with discontinuous word domains cannot be projective, since its tree diagram will fail to exhibit the required properties.

As Le Cerf and Ihm have noted, it is possible to represent the structure of sentences by marking off word domains in linear text with pairs of parentheses. Connectedness of all word domains insures that such parenthesization will involve no interleaving of parentheses (as occurs, for example, in the construction "(W[(XY)Z]", where "W-Y" and "X-Z" are discontinuous word domains). For projective sentences, the resulting configurations of parentheses will thus be "nested" in a regular manner, much in the way they occur in familiar algebraic and logical expressions. Hence projective languages - that is, languages with exclusively projective sentences - may also be described as regularly parenthesized languages.

On the basis of the sentences which they have examined so far, Le Cerf and Ihm have observed that a number of natural languages, including English, French, German and Russian, appear to be projective. A similar view has been adopted in some of the recent automatic translation research conducted at the Harvard Computation Laboratory, in that the system of predictive syntactic analysis used on Russian and English sentences has proceeded on the equivalent hypothesis that sentences in those two languages have a regularly nested syntactic structure. When, during predictive analysis, elements of an inner nest are found to be followed by an element belonging to a nest which began further to the left, the inner nest is assumed to have been completed, and any unfulfilled predictions associated with it are automatically removed from the prediction pool (Sherry, 1960; Bossert, Guiliano and Grant, 1960). This procedure reflects the expectation that only well-nested constructions will be found in normally acceptable sentences.

The relation between representations of sentence structure in terms of parenthetic grouping and representations in terms of tree diagrams has been investigated to a limited extent by Hays (1960.) Hays severely restricts his discussion by deciding, on the one hand, to ignore ***word-class*** distinctions and considerations of word order, and on the other hand, by identifying parenthetic grouping exclusively with a form of parenthesization corresponding to conventional immediate constituent analysis. Under the latter restriction, nesting occurs only in binary fashion, an expression in parentheses being partitioned, if at all, into precisely two components by the insertion of inner parentheses.

One of the conclusions reached in Hays' study is that certain features of syntax which can be represented in tree diagrams cannot be indicated by parenthetic grouping, whereas certain other syntactic properties expressible in the latter notation are lost in tree representations. The two parts of this conclusion are based respectively on the existence of examples of syntactic structures which have identical parenthetic representations, but distinct tree diagrams, and of syntactic structures with Identical tree diagrams, but different parenthetic representations. Upon careful inspection, however, it becomes apparent that the validity of Hays' examples is entirely dependent on his restrictive assumptions regarding word order,

TREE STRUCTURE

(a)

(((LITTLE)(JOHN))((ATE)(BREAKFAST)))   ((HE)((ATE)((HIS)(BREAKFAST))))

PARENTHETIC STRUCTURE
(IMMEDIATE CONSTITUENT ANALYSIS)

(b)

Figure 2

word classes and parenthetic grouping. Under those assumptions, for example, the two sentences "Little John ate breakfast" and "He ate his breakfast" have identical tree structures, but different parenthetic structures ***(Figure 2).*** As soon as word order considerations are introduced, however, the tree representations become distinct as well ***(Figure 3).***

To demonstrate that parenthetic grouping can reflect any distinctions of form among various tree structures, it is sufficient to describe a unique one-to-one mapping which carries any projective tree into a parenthetic expression. This mapping is accomplished by marking off the projection of the domain of each node in the tree with a pair of parentheses. Such a form of parenthesization is illustrated in ***Figure 3b.*** Comparison of ***Figure 2b*** with ***Figure 3b*** reveals that the grouping indicated in the latter differs considerably from that determined by traditional

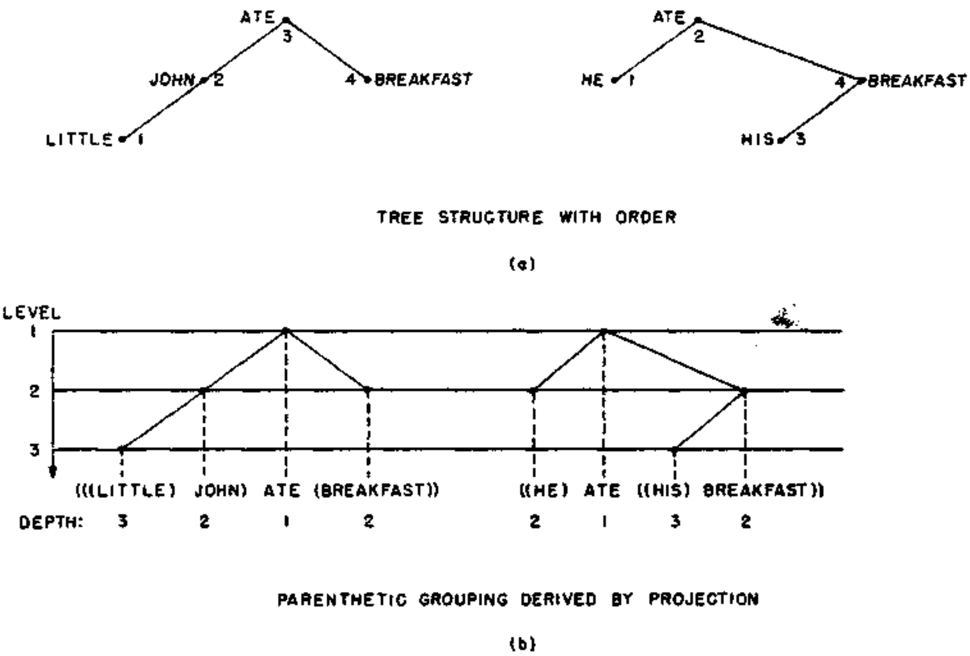


Figure 3

immediate constituent analysis, in that no attempt is made to force a division of each construction into two constituents. Rather, the parentheses introduced by projection indicate the hierarchy of nodal dependencies within the tree: the topmost node is enclosed within one pair of parentheses, the nodes immediately dependent on it are enclosed within two pairs of parentheses, and so forth. In every case, the depth of nesting of an item in parentheses is numerically equivalent to the level of the corresponding node in the tree diagram.

The new notation is considerably more general and flexible than the parenthetic notation treated by Hays, in that it readily mirrors the property of trees that any node may have any number of branches descending from it. Consider, for example, the sentence "John gave the book to him gladly", which would be extremely awkward to treat within the framework, of immediate constituent analysis. As soon as the tree diagram indicating the syntactic interdependencies of the words in the sentence has been constructed, however, the corresponding projective parenthetic representation of the sentence is readily obtained ***(Figure 4).*** If one removes the outermost parentheses of the expression in the example, there remains one item 'gave' without parentheses, together with four items - 'John', 'the book', 'to him', 'gladly' - within parentheses. This indicates that 'gave' serves in this sentence as a functor with four arguments, corresponding to the fact that the topmost node of the tree diagram has four nodes immediately dependent on it. As successive levels of parentheses are removed, the hierarchy which obtains among the constituents of each construction is always similarly unambiguous,

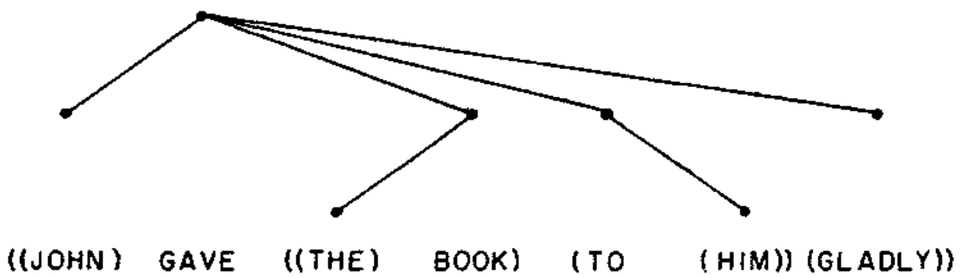


Figure 4

in that one item within each pair of parentheses will not be surrounded by further sets of parentheses, whereas all others will.

In the examples in ***Figures 3*** and ***4,*** both of the representational forms used have indicated only word order and a generalized form of syntactic dependency. Word order is represented in both instances by maintaining the usual left-to-right sequence of words. Syntactic dependency is indicated in the tree diagrams by lines connecting the appropriate pairs of nodes, with the dependent node always placed one level lower than the one upon which it depends. In the parenthetic expressions, a word which depends directly on another word will be included in all pairs of parentheses which enclose the latter, plus one additional pair. Thus the dependent word has a depth which is greater by one than that of the word upon which it depends.

The two notations just discussed are inadequate, in their present form, for representing the syntactic structure of sentences in natural languages. This is due to the fact that they do not distinguish among the different types of syntactic linkages between words. Moreover, inclusion of information regarding the part of speech of each word, while it would greatly increase the representational power of the notations in question, would still fail to allow a complete characterization of the syntactic structure of all sentences. That this is the case can be seen quite clearly from a comparison of the two Russian sentences Красный стол имеет ноги and Красные ноги имеет стол which have identical sequences of parts of speech as well as identical generalized dependency structures (word 1 depends on word 2, which depends on word 3, and so forth). Yet the noun-verb-noun string in the first sentence is syntactically a subject verb-object sequence, where the first noun agrees with the verb in number and the second does not; whereas the noun-verb-noun string in the second sentence is on object-verb-subject sequence, with the agreement relationship correspondingly reversed.

One simple method of augmenting tree notation so that the different types of syntactic linkages can be distinguished is to assign the names of the appropriate syntactic roles (such as 'subject', 'indirect object', 'preposition complement', and so forth) to the nodes in a diagram. This can be accomplished either by direct labelling on the diagram, or by placing the labels in a line above or below the tree at those points where the vertical projections of the corresponding nodes intersect the line. This latter approach is particularly convenient when the nodes of the diagram are already labelled with the words to which they correspond. Similar techniques can be employed in labelling parenthetic expressions, but their essentially one-dimensional nature is destroyed if both words and syntactic roles are indicated.

In concluding these observations regarding the use of tree diagrams and parenthetic notation for the representation of syntactic structure, it would be well to point out that although one can construct a parenthetic expression containing the same information as a given tree, there are considerable differences in the relative accessibility of the information in the two different forms. Tree diagrams lend themselves to rapid visual interpretation, whereas interpretation of the more compressed notation of parentheses, in the case of complex structures, often requires a preliminary process of decipherment. For purposes of further machine processing, however, the advantage appears to lie with the parenthetic notation.

3. CONSTRUCTION OF THE DIAGRAMS

In the previous section, a method was introduced for transforming tree diagrams of sentence structure into equivalent parenthetic representations by a process of projection. The automatic sentence diagramming procedure about to be described performs what is essentially the inverse transformation, in that it carries parenthetic expressions into equivalent tree structures. An important key to this latter process is the fact, already noted earlier, that "depth" or "level" remains invariant in passing from one notation to the other.

Construction of the diagrams proceeds in three distinct phases, each of which is associated with a separate machine pass. During the first pass, each Russian word is provided with coding indicating all constructions within which it is included. As will be shown shortly, the set of codes assigned to the words of an entire sentence marks the boundaries of constructions in much the same way as would pairs of parentheses. A depth indication derived from the configuration of the "parentheses" is also assigned to each word in the text during the first pass. In the second pass, the words in each sentence are sorted into a standard order on the basis of the syntactical relationships indicated by the "parentheses". The "canonical" sentences thus produced are finally converted into tree diagrams by means of a simple editing pass which translates the numerical depth indications into the physical level of nodes within trees.

In order to explain how the codes assigned during the first pass function, on the one hand, as sets of parentheses, and on the other hand as keys for sorting the words of each sentence, it will be necessary to refer to a few details regarding both the format of the output and the internal structure of the codes. The primary printed output of the predictive syntactic analyzer ***(Figure 5)*** has a basically columnar format, with separate columns for the Russian words and for their English correspondents, grammatical coding and attributed syntactic roles. A text can be read by scanning the

FIRST ENGLISH EQUIVALENT	RUSSIAN WORD (TRANSLITERATED)	GRAMMATICAL CODING	LINKAGE AND SYNTACTIC ROLE
ON	NA	R	INF PREP
ANODIC	ANODNOJ	AD00000	594 R COMP
LOAD	NAGRUZKE	NDI1F000	595 R COMPM
TUBE	LAMPY	NDI2F000	596 N COMP
⋮	⋮	⋮	⋮

Figure 5

(OUTPUT OF THE PREDICTIVE SYNTACTIC ANALYZER) column of Russian words vertically, from the top downward. During the first pass of the diagramming routine, this format is modified only to the extent of inserting a code word of fixed length for each Russian word. Since the code words are always placed in the same horizontal position, they form a distinct column in the output of the first pass.

The internal structure of the code words is indicated in schematic form in ***Figure* 6**. The portion of the code labelled "Structural Co-ordinate" designates all constructions, from the clause level down to the word level, which include the corresponding Russian word. The clause number, which appears in the left-most field of the code, is assigned on the basis of a

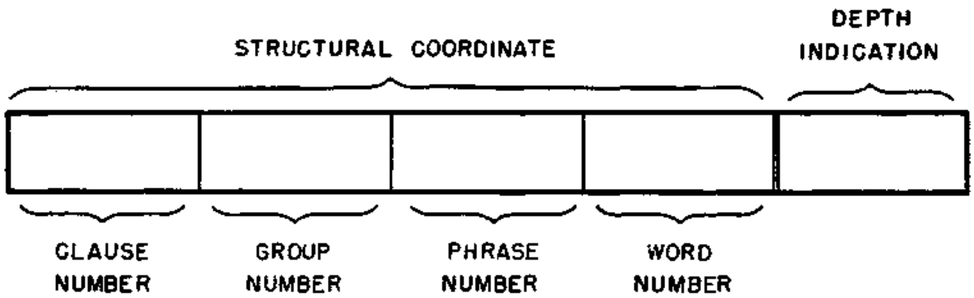


Figure 6

(CODE WORD FORMAT)

consecutive numbering of the clauses within each sentence. The group number indicates which of the major clause constituents - such as subject, predicate head", object and indirect object - contain the word. Such major clause constituents, which in general consist of one or more phrases, will be referred to as "groups" in the remainder of this paper, for want of a standard generic term. The two low-order fields of the structural co-ordinate indicate the position of the phrase within the group, and of the word within the phrase, respectively.

The relation of the structural co-ordinates to a parenthetic representation of sentence structure becomes evident when one considers the general appearance of a column of codes for an entire sentence ***(figure 7).*** For the simple one-clause sentence shown, the string of identical digits in the clause number column may be viewed as a bracket surrounding the entire clause. Within this outer bracket, there are shorter strings of identical digits in the group number column which mark the limits of the various groups. As is shown in the figure, the beginning of each bracket is inter-

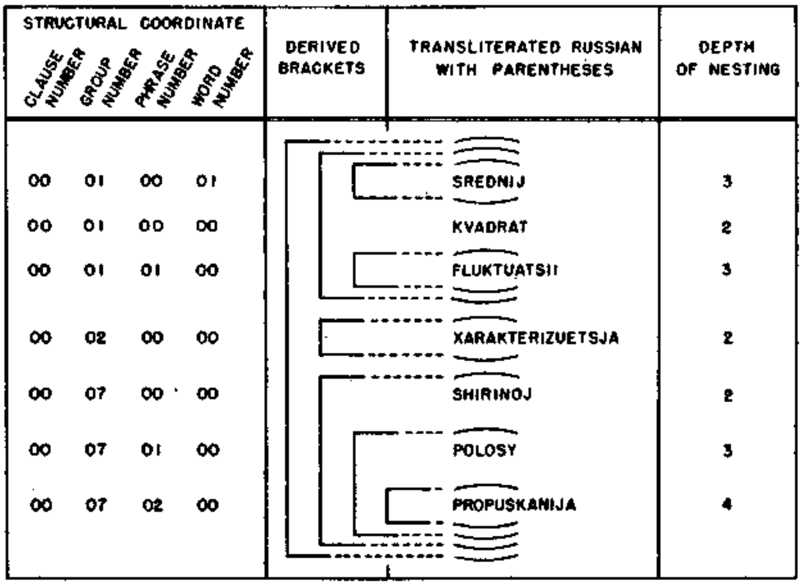


Figure 7

preted as a left parenthesis and the end of each bracket as a right parenthesis. Interpretation of the phrase number and word number coding in terms of brackets is somewhat less straightforward: Each successive phrase in a group is assumed to be nested within all previous phrases of the group, in order to reflect syntactic dependency in the manner described in the previous section. This approach is illustrated by the treatment of the three-phrase string "SHIRINOJ POLOSY PROPUSKANIJA" in the example in ***figure 7***.

As has been noted earlier, depth of nesting of an expression in parenthetic notation is here taken as the number of pairs of parentheses which include the expression. One very simple way of determining the depth of each word in a parenthesized sentence is to scan the sentence from beginning to end, keeping a running count of the number of left parentheses encountered minus the number of right parentheses. As can be seen by reference to ***figure 7***, such a count gives the depth, of each word, since it records at each point in the sentence the number of sets of parentheses which have been opened, but which have yet to be closed. This general technique, carried out in terms of changes in the various columns of the structural co-ordinates, allows the concurrent generation of depth indications and structural co-ordinates during the first pass of the diagramming routine.

Once the depth of each word in a sentence is known, it is a relatively easy task to devise an automatic method for placing the corresponding nodes at the proper levels of a two-dimensional output array. Since text order is reflected in the vertical dimension, it is convenient to represent depth horizontally, from left to right. Once the nodes have been assigned to the appropriate positions, however, it is not possible to complete the tree diagram by connecting nodes with lines in conventional fashion, since the output printer cannot simulate lines with an arbitrary orientation. Thus some other notational device must be adopted for representing syntactic dependency. The scheme employed here involves a preliminary reordering of the sentence such that each structure in the sorted version is followed directly by the substructures which depend on it. Such an order of sentence components corresponds to the Polish parenthesis-free notation (Łukasiewicz, 1957), which has been used for expressions in symbolic logic and other formal languages.

The principal changes brought about by the prescribed reordering are relatively minor ones, consisting chiefly in the placement of all adjectival and participial modifiers after the nouns which they modify, regardless of their initial location. Such a rearrangement is provided for during generation of the structural co-ordinates (which serve as sorting keys in the second pass) by the simple expedient of assigning a word number code of zero to the noun in each noun phrase ***(Figure 7).*** This ensures that the noun will be sorted to the top of its phrase during the second pass.

In addition to their provision for the necessary reordering of words within phrases, the structural co-ordinates are designed to bring about an optional reordering of groups within clauses. Since the order chosen - subject, predicate head, direct object, indirect object - is a common one for main clauses in English, the sorting procedure performs some of the word order transformations which may be necessary in the process of Russian-English structural transfer (Foust and Walkling, 1961). Once again, the code assignment scheme which provides for the desired reordering is a very simple one: whatever the original order of the sentence may be, the subject and its dependent structures are assigned the group number '01', verbal predicate heads are assigned the group number '02', and so forth.

It should be noted here that reordering of groups will give a legitimate parenthesis-free ordering of a clause only if all groups are assumed to occur on a single syntactic level and to depend on the clause indicator. This assumption, which contrasts with the common practice of treating subject, object and indirect object and indirect object as dependent on the verb, has been adopted in the experimental part of this study for two reasons: First, the major clause constituents are treated in a similar manner during the original syntactic analysis; second, the resulting output is more readable than it would be under the more common interpretation, since the verb need not be placed at the beginning of each clause, but can appear in normal position following the subject.

The assignment of fixed numbers to the different groups in a sentence results in a group number column which can readily be scanned, either visually or automatically, to determine the basic structure of each clause in terms of its major constituents. Subsequent sorting into fixed order has the further effect of identifying clauses with the same group components, but different orders (for example, subject-verb-object and object-verb-subject clauses), as having similar syntactic structure. The net result is that if one uses the sorted output as the starting point for the application of translation algorithms, the number of different structures which must be dealt with is considerably reduced. While such word order changes may obscure shades of stylistic emphasis which it might be possible to

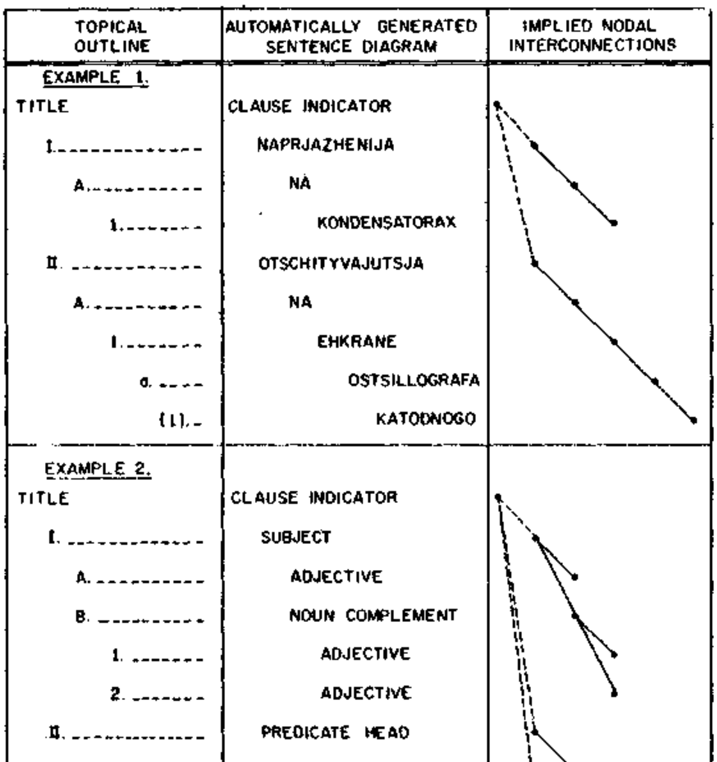


Figure 8

reflect in a highly polished English translation, the simplification thus achieved (without changing the basis syntactic relationships in any way) is an appealing one at the present stage of development of automatic translation techniques.

Due to the standardization of the order of syntactic units during the sorting procedure, the sentences comprising the output of the second pass may conveniently be thought of as having been reduced to a canonical form. Construction of tree diagrams from such canonical sentences is a very straightforward operation. Rather than using dots for nodes in the diagrams, either the words of each Russian sentence or the names of their syntactic roles are used directly, giving the effect of labelled nodes ***(Figure 8).*** If the depth of a given word in the canonical output is n, the third pass places it on magnetic tape in such a position as to appear in the nth column of the output print. Thus the arrangement of words in a two-dimensional diagram is accomplished in two separate stages: vertical rearrangement is completed during the second pass, whereas horizontal arrangement takes place during the third pass.

As can be observed from the examples in ***Figure 8,*** the sentence diagrams produced by the system are somewhat unusual, both in that they branch exclusively downward and to the right, and in that there are no lines between nodes. The total effect, however, can be seen to be a very familiar one - that of the standard topical outline, minus its letters and numerals. The nodal interconnections, which are unambiguously determined by the parenthesis-free ordering and the level of the nodes, are included explicitly at the right for purposes of comparison. Although most people familiar with conventional outline format should have no trouble filling in the connections by inspection, one can describe the process formally as follows: each word in column n is to be thought of as connected to all following (lower) words in column n+l. up to the first word which is in column n or a column to the left of column n.

4. APPLICATIONS OF AUTOMATIC DIAGRAMMING TECHNIQUES

As has previously been remarked, one great advantage of tree diagrams lies in the relative degree of ease with which they can be interpreted visually. In two dimensions, one can see at a glance relationships which are obscured in a one-dimensional format, even with brackets or parentheses included. The representation of syntactic level in terms of discrete physical distances in one dimension of the diagrams can be particularly advantageous when one is interested in comparing the structure of various sentences. If one wishes, in making such a comparison, to ignore the fine structure of the sentences below a certain level, one can readily accomplish this by considering only the portion of the diagrams to the left of the appropriate column in the output print.

Beyond their general usefulness as devices for representing sentence structure, the tree diagrams described here have also shown promise as aids in the detection of errors made during the original syntactic analysis. Whenever a syntactic link has been tentatively established by the predictive analysis program as holding between two words in a sentence, it is usually reflected in the tree diagram by placing the dependent word in a position immediately to the right of and below the other, however far apart the words may have been in the original sentence. If the assumed link is not a correct one, juxta-position of this sort generally causes the error to stand out, often through semantic incompatibility of the words in question.

The canonical sentence output described in the previous section may well find its most important application in furnishing an objective basis for the classification of sentences according to structural type. In accordance with the observations made in the previous section regarding potential uses of the group number codes, one might wish, as an initial step in classification, to bring together all sentences with identical structure down to and including the group level. Not only would a catalogue of the sentence structure types found be linguistically interesting, but a study of possible translations on a category-by-category basis would tend to speed the discovery of generally valid translation algorithms based on syntactic criteria. With some minor format modifications, the current output would lend itself quite readily to such a classification process, which could be accomplished either manually or with the aid of a machine program for summarizing the contents of the clause and group columns for each sentence.

In addition to the important contribution which automatic sentence classification techniques would make toward the description of entire language systems and toward the development of translation algorithms, they might also have considerable value when applied in statistical studies of literary style. Traditional statistics of style, which have been based almost exclusively on non-syntactical criteria such as vocabulary distribution and sentence length, could be supplemented with data on the types of sentence structure preferred by a given author. Moreover, a number of potentially useful statistics relating to lower-level constructions might easily be obtained from the structural co-ordinates. Counts of the number of words per phrase and of the number of phrases per group might be included among such statistics.

Another possible application of the outputs of the sentence diagramming program is their employment as an aid in language data processing for purposes of information retrieval, particularly in systems for automatic literature abstracting of the sort proposed by Luhn (1958). The feature of the tree diagrams which is pertinent here is that the main components of a clause, including subject, verb and object, always correspond to the "main topics" in an outline, and are therefore located at the upper levels of the tree. When the words on these upper levels are considered apart from the lower-level structures which modify them, they often summarize the content of the sentence in a sort of "newspaper headline" or "telegraphic" style. Each level added fills in the details a little more completely, but at the expense of including more and more of the sentence.

Although sentences occur in which the key term or phrase lies buried deep down in the structure, preliminary observations indicate that there are many others in which the semantic hierarchy closely parallels that of the syntactic structure. This suggests that more sensitive vocabulary statistics for purposes of automatic abstracting may be obtainable by considering only words occurring in positions above a predetermined cut-off level in the sentence structure. Alternatively, one might count occurrences of words on each level, and then multiply by a fixed weighting factor in each instance before taking the overall totals. In either case, the necessary classification of word occurrences in a text according to syntactic level could be readily accomplished automatically using the depth indications provided for each word in the first pass of the diagramming routine.

In conclusion, it can be stated that the research described in this paper has not only demonstrated the feasibility of diagramming sentences by machine, but has also suggested the use of the diagrams and related machine outputs in a number of different areas. It should be noted, however, that the effectiveness of the method of diagramming presented here ultimately depends on the accuracy of the underlying syntactic analysis. Any final evaluation of the usefulness of this general approach in the field of application suggested here will therefore have to await further improvements in the techniques of automatic syntactic analysis.

REFERENCES

BOSSERT, W. (1960), "The Implementation of Predictive Analysis," *Mathematical Linguistics and Automatic Translation,* Report No. NSF-4, Harvard Computation Laboratory, Section VIII.

BOSSERT, W., Guiliano, V.E., and Grant, S.E. (1960), "Automatic Syntactic Analysis of English," *Mathematical Linguistics and Automatic translation,* Report No. NSF-4, Harvard Computation Laboratory, Section VII.

FOUST, W.D., and Walkling, J.R. (196), "A Preliminary Structural Transfer System", *Mathematical Linguistics and Automatic Translation,* Report No. NSF-6, Harvard Computation Laboratory,

HATS, D.G. (1960), "Grouping and Dependency Theories", Rand Corporation, P-1910, Santa Monica, California.

LE CERF, Y., and Ihm, P. (1960), "Elements pour une Grammaire Generale des Langues Projectives," *Rapport GRISA,* No. l, pp. 11-29.

LUHN, H.P. (1958), "The Automatic Creation of Literature Abstracts," *IBM Journal of Research and Development,* Vol. 2, pp. 159-165.

ŁUKASIEWICZ, J. (1957), *Aristotle's Syllogistic* (Oxford).

SHERRY, M.E. (1960), *Syntactic Analysis in Automatic Translation,* Doctoral Thesis, Harvard University, Cambridge, Massachusetts.