One generalization of context-free grammars and its application to structural analysis of drawings

Ivan Aksak, Volodymir Kijko, 40, Prospect Akademika Glusckova
Viacheslav Matsello, Mychailo Schlesinger 252207, Kiev, Ukraine
Institute of Cybernetics, tel.: 266-25-69
Ukraine Academy of Sciences schles%image.kiev.ua@ts.kiev.ua

I.Aksak, V.Kijko, V.Matsello, M.Schlesinger. One generalization of context-free grammars and its application to structural analysis of drawings.

A problem of parsing of image elements set is considered. A new algorithm of hierarchical description of the set of image elements is proposed, and the result of its testing on real handdrawn pictures is also adduced. The general idea of the algorithm is, that while structural analysis of the image, it is necessary to take into consideration not only local interconnections between image's elements, but also more complex interconnections. The new algorithm is based on one generalization of context-free grammars, that has greater generating possibility than ordinary context-free grammars.

1. Description of applied problem.

The proposed algorithm is intended for making the descriptions of drawings and maps. Saying "description" we mean, that it is necessary to define that, for example, Fig.1 consists of: the first object, that consists of transistor graphical sign and inscription, inscription consists of transistor conventional name and number, transistor's conventional name consists of letter **V** and letter **T**, number consists of figure **1** and figure **2**; the second object, which consists of ....., etc. To produce such description, some set of unifications must be fulfilled: at first the letters **V** and **T** are united into the word **VT** (transistor's conventional name) and figures **1** and **2** are united into the number **13**. Then the word **VT** and number **13** are united into inscription, and at last the sign of transistor is united with the inscription as the final compound object, that has its own certain meaning in the image, and cannot be united with any object. An obvious approach when elements are united if: 1) they can be united in general (for example **V** with **T**, but not **V** with -▭-), 2) they are nearly located in the image, is not correct by more detailed consideration, and can solve the problem only in some special cases. The reason of that is because the whole image is unambiguous, but not (not always) its local parts. For example it is clear, that inscription **VT13** (Fig.1) relates to transistor drawn bold, because in other case it would be impossible to relate inscription **VT12** to any other transistor. It is easy for a human to parse such image, but to do it automatically the speculations of such type must be formalized.

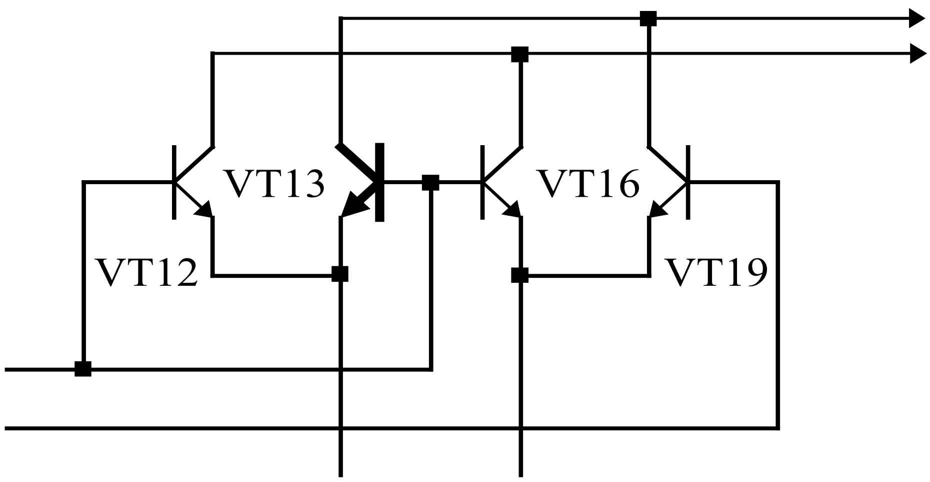

Fig.1

We consider the problem as a specific problem of syntactic analysis of objects, where the set of allowable objects is defined by certain type of attribute grammar. The object under parsing is the set of primary image elements, that are obtained using some procedure of image elements recognition, for example [1]. So the image in Fig.1 is supposed to be represented as a list, which consists of all transistor's signs, all figures and all letters, that exist in the picture.

2. The formal statement of the problem.

2.1. Main definitions.

Let $\mathbf{N}$ be a finite alphabet of names, $\mathbf{N^P} \subset \mathbf{N}$ be a subset of primary names and $\mathbf{N^F} \subset \mathbf{N}$ be a subset of final names. Let $\mathbf{A}$ be a multidimensional linear space, in which the subset $\mathbf{A^F}$ is defined. We consider an object as a pair (n, a), where $n \in \mathbf{N}$, $a \in \mathbf{A}$. The pair (n, a) is called an identifier of the object. The element n of the identifier is a name, and element a is an attribute of the object. Objects, whose names $n \in \mathbf{N^P}$ are called primary, other objects we will call compound. Each compound object consists of some nonempty set of primary objects.

We will distinguish allowable and not allowable objects using the following formal construction, that is the certain generalization of context-free (CF) grammars by Homsky [2].

Let $\mathbf{P^3} \subset (\mathbf{N}\times\mathbf{A})\times(\mathbf{N}\times\mathbf{A})\times(\mathbf{N}\times\mathbf{A})$ and $\mathbf{P^2} \subset (\mathbf{N}\times\mathbf{A})\times(\mathbf{N}\times\mathbf{A})$ be two relations, that define an identifier to be referred to compound object, i.e. to the set of primary objects.

<u>Definition 1.</u> A set $\mathbf{O}$ of primary objects, i.e. compound object, can receive the identifier (n, a) in the system of relations $\mathbf{P^3}$ и $\mathbf{P^2}$, if one of the next two conditions is true:

1. The compound object $\mathbf{O}$ consists of the single primary object (n, a).
2. There exists such compound object $\mathbf{O'}$ with the identifier (n', a'), that $((n, a), (n', a')) \in \mathbf{P^2}$.
3. There exist such two compound objects $\mathbf{O'}$ и $\mathbf{O''}$ with the identifiers (n', a') и (n'', a''), that:

   a) $\mathbf{O'} \cap \mathbf{O''} = \varnothing$;

   b) $\mathbf{O'} \cup \mathbf{O''} = \mathbf{O}$;

   c) $((n, a), (n', a'),(n'', a'')) \in \mathbf{P^3}$;

We determine an allowability of compound object in the system of relations $\mathbf{P^3}$ и $\mathbf{P^2}$ via the following definition.

<u>Definition 2.</u> A compound object $\mathbf{O}$, i.e. a set of primary objects, is allowable in the system of relations $\mathbf{N}$, $\mathbf{N^P}$, $\mathbf{N^F}$, $\mathbf{A}$, $\mathbf{A^F}$, $\mathbf{P^2}$, $\mathbf{P^3}$ if the object $\mathbf{O}$ can receive an identifier $(n, a) \in (\mathbf{N^F}, \mathbf{A^F})$.

<u>Definition 3.</u> A set $\mathbf{O}$ of primary objects is an allowable image, if such subsets $\mathbf{O}_1, \mathbf{O}_2, \ldots, \mathbf{O}_n$ exist, that $\mathbf{O}_1 \cup \mathbf{O}_2 \cup \ldots \cup \mathbf{O}_n = \mathbf{O}$, $\mathbf{O}_i \cap \mathbf{O}_j = \varnothing$ $(i \neq j)$ and each $\mathbf{O}_i$ is an allowable compound object.

<u>Definition 4.</u> An image description is a set of derivation trees of the following type (Fig.2):

1) leaves ($\mathbf{L}_1$, $\mathbf{L}_2$, $\mathbf{L}_3$, $\mathbf{L}_4$ in Fig.2) of the tree correspond to primary objects. Each leave has one output. It means that corresponding primary object is contained in the object connected to the output;

2) root ($\mathbf{R}$ in Fig.2) of the tree corresponds to allowable compound object. Each root has one or two inputs. It means that corresponding allowable compound object consists of the objects connected to inputs;

3) each node ($\mathbf{N}_1$, $\mathbf{N}_2$, $\mathbf{N}_3$, $\mathbf{N}_4$ in Fig.2) corresponds to allowable compound object. Each node has one output and one or two inputs. It means that corresponding compound object is contained in another compound object, connected to the output, and consists itself of the objects connected to his inputs.

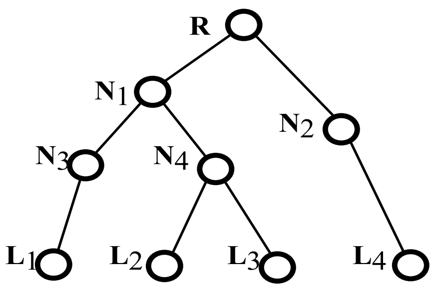

Fig.2

The described construction generalizes the conception of CF languages, because it allows to define any CF language. Really, let some CF grammar be defined by the terminal alphabet $\mathbf{B^T}$, nonterminal alphabet $\mathbf{B^N}$, $\mathbf{B} = \mathbf{B^N} \cup \mathbf{B^T}$, the axiom $\sigma \in \mathbf{B^N}$, the set $\mathbf{S^3}\subset\mathbf{B}\times\mathbf{B}\times\mathbf{B}$ of substitutions of type $b\to b'b''$ and the set $\mathbf{S^2}\subset\mathbf{B}\times\mathbf{B}$ of substitutions of type $b\to b'$. If one considers primary object as a pair (n, i), where $n \in \mathbf{B^T}$ and i is a positive integer number, then string is a specific compound object. In this case element of a string (i.e. symbol) is a primary object, n is a name, and i is a number of primary object in a string. The set of allowable compound objects coincides with the language of CF grammar if:

1. $\mathbf{B}$ is taken as $\mathbf{N}$, $\mathbf{B^T}$ is taken as $\mathbf{N^P}$, $\{\sigma\}$ is taken as $\mathbf{N^F}$;
2. a set of pairs of integers is taken as $\mathbf{A}$;
3. a set of pairs of type (l, n), $n \geq 1$ is taken as $\mathbf{A^F}$.

4. a set $\mathbf{P}^3$ is formed using the following rule:

$((n, a, b), (n', a', b'), (n'', a'', b'')) \in \mathbf{P}^3$, if $n \to n'n''$ belongs to the set of substitutions $\mathbf{S}^3$ and $a = a'$, $b = b''$, $a'' = b' + 1$.

5. a set P^2 is formed as follows:

$((n, a, b), (n', a', b')) \in \mathbf{P}^2$, if $n \to n'$ belongs to the set of substitutions $\mathbf{S}^2$ and $a = a'$, $b = b'$. .

The proposed construction has greater generating possibility than ordinary CF grammars. Thus, using this construction it is possible to define the set of sequences like $1^n0^n1^n$, that is not a CF language.

To solve the applied problem described above, we use the proposed construction in the following way. The set of attributes **A** is the set of 4-tuples of numbers, i.e. each objects represented by 5-tuple (n, l, u, r, d). The first element n is a name, and numbers l, u, r, d are coordinates of the left, up, right and down sides of the minimal rectangle, which envelopes the corresponding object in the image.

The relation **P3** is defined as the set of rules of type:

$n_i = n_i'n_i''\ \Psi_i$,

where i is a number of the rule, n_i, n_i', n_i'' are the names, Ψ_i is a relation defining the set of allowable 8-tuples of numbers $\{l', u', r', d', l'', u'', r'', d''\}$ for which this rule can be used. The relation Ψi has a form of a system of linear inequalities, that relates the elements of this 8-tuple. Attributes of the object in the left side of substitution are determined as the following: $l = \min(l', l'')$, $u = \min(u', u'')$, $r = \max(r', r'')$, $d = \max(d', d'')$.

The relation **P2** is defined as the set of rules of type:

$n_j = n_j\,'\Psi_j$,

where j is a number of rule, n_j, $n_j\,'$ are the names, Ψ_j is a relation defining the set of allowable 4-tuples $\{l', u', r', d'\}$ for which this rule can be used. The relation Ψj also has a form of a system of linear inequalities, which relate the elements of this 4-tuple. Attributes of the object in the left side of substitution are the same as ones in the right side, i.e. $l = l'$, $u = u'$, $r = r'$, $d = d'$.

2.2. Statement of the problem.

Let image be represented as the set of primary objects and the system of sets **N**, $\mathbf{N}^\mathbf{P}$, $\mathbf{N}^\mathbf{F}$, **A**, $\mathbf{A}^\mathbf{F}$, $\mathbf{P}^3$, $\mathbf{P}^2$ is defined. The task is to make a conclusion about allowability of the image, and if it occurs allowable, then make a description of the image.

This problem is difficult.

We propose a plausible algorithm, that finds an image description if this description is unique. The algorithm generates a description of the input image automatically, if this image is allowable and has only one possible description. The algorithm calls for human operator, when during the analysis it finds that the input image is not allowable, or that the input image can’t be unambiguously described (has more than one possible description).

3. Brief description of the algorithm.

At first a *binding up* procedure is executed over the list of image objects. Initially this list consists of only primary objects, but while operating the new compound ones are added to the list. The matter of the procedure is the following. All objects from the list are looked over. If, according to relation $\mathbf{R}^3$, it is possible to create a compound object from the pair of objects and these objects do not embrace common objects, then the new compound object is created and added to the list. The objects from this pair will be considered as initial objects for new compound object, and from the other side, new object will be considered as a derivative for each object of the pair. If according to relation $\mathbf{R}^2$, it is possible to create a compound object from some existing object, then the new compound object is created and added to a list. By analogy, the object from which the new one was created will be considered as an initial object for the new compound object, and from the other side, new object will be considered as a derivative. Procedure is finished when no new object can be created.

Then the procedure *of obliterating* is applied to the list formed by the binding up procedure. All objects of the list except primary and allowable compound objects are looked over. The objects which have no derivative ones are obliterated. Naturally, when some object was obliterated, its initial objects no more have it as a derivative. Obliterating procedure is finished when no more objects are possible to obliterate.

After that all primary objects of the list are inspected. If primary object has no derivative one, it means that the image is not allowable. In this case the message about the necessity of correction is sent to an operator. After correction the algorithm starts from the very beginning.

Then the following procedure of *deleting* [3] is executed.

The primary objects contained in the single allowable compound object are looked for. If such primary object was found, then we consider corresponding allowable compound object to be correctly recognized. That allows to delete all other allowable compound objects containing the primary objects, which are also contained in correctly recognized allowable compound object. As a result some more primary objects, contained in a single allowable compound object may appear. There are two possible consequences of image nonallowability in this step of the algorithm. The first is that primary objects appear which are contained in no one allowable compound object. In this case the message about the necessity of correction is sent to an operator. After correction the algorithm starts from the very beginning. The second case: it is impossible to delete any allowable compound object, but some primary objects contained in more than one allowable compound objects remain. It means that more than one ways of image representation as a set of nonintersectioned allowable compound objects exist. In this case the message about the necessity of choosing a single representation among all is sent to the operator.

If neither the firs nor the second case occurred it means that the image is allowable in the grammar, and the list of objects is the description of the image, where the root of each tree is allowable compound object and the leaves are primary objects.

Concluding remarks.

We have tested the algorithm on real handdrawn pictures. One of them is shown on Fig.3. This picture was scanned, preprocessed, and then the recognition procedure [1] was executed over the picture. Finally the procedure of parsing was applied to the results of recognition. The parsing procedure ran fully automatically (there were no calls to operator), and took 660 msec on Pentium 90 system. The result was checked and was founded to be correct.

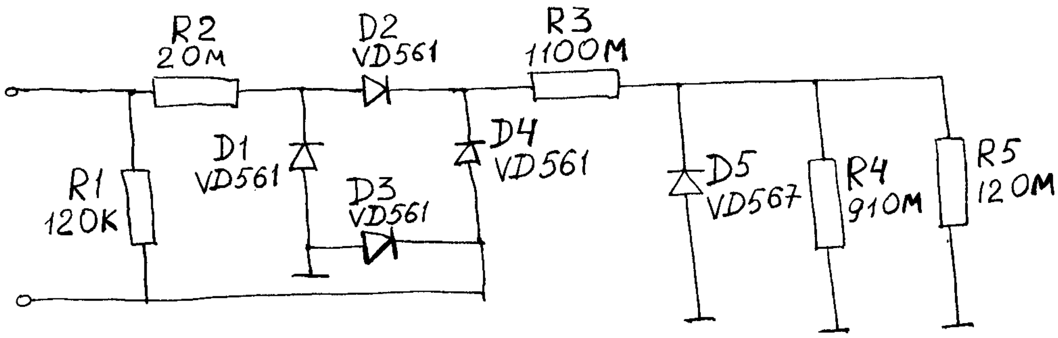

Fig.3

References

1. V.M.Kijko, Recognition of Objects in Images of Paper Based Line Drawings, // Proc. Of Third Intern. Conf. On Document Analysis and Recognition. - Montreal, 1995. - V.11, pp. 970 - 973.
2. Alfred V. Aho, Jeffrey D. Ullman. The Theory of Parsing, Translation and Compiling. - Prentice-Hall, Inc., Englewood Cliffs, N.J., 1972 - V.1.
3. M.I. Schlesinger, Formale Zugaenge fuer die Strukturanalyse von Grafishen Bildern, // Musterkennung 1992, 14.DAGM-Symposium. - Dresden, 14.-16. September 1992. - pp.10-19.