## THERE DOES NOT EXIST AN ENUMERABLE FAMILY OF CONTEXT-FREE GRAMMARS THAT GENERATES THE CLASS OF SINGLE-VALUED LANGUAGES

K. Yu. Gorbunov

UDC 519.765

I. The concept of context-free grammars entails a consideration of a deducible word together with its deduction, where a deduction is understood as the syntactic analysis of a particlar word. It is natural to require uniqueness of the analysis, i.e., uniqueness of the deduction. This requirement means that the corresponding context-free grammar must be single-valued.

The class of single-valued context-free grammars possesses better algorithmic properties than does the class of all context-free grammars. For example, as has been shown by Semenov [1], given an arbitrary single-valued context-free grammar and an arbitrary regular contextfree grammar, it is possible to decide whether or not they are equivalent (i.e., whether they generate the same language). It is known that the problem of equivalence of a regular context-free grammar is not solvable for arbitrary context-free grammars.

Muchnik [2] raised the following question: Is it possible to enumerate a family of single-valued context-free grammars that generates all the single-valued languages? If the answer were in the positive, the family of single-valued languages would be just as well-defined as, say, the family of deterministic languages.

In the present article we will obtain a negative response to this question and, incidentally, construct a class of languages that possess certain interesting properties. That is, for the grammars specifying the languages of this class, the problem of equivalence with an arbitrary context-free grammar is solvable.

For purposes of illustration, recall that, as Hopcroft has proved [3], of all the regular languages only the bounded languages possess these properties, i.e., finite unions of concatenations of languages of the form  $\{u\}$ ,  $\{u\}^*$ , where u is some word. On the other hand, in this class the family of essentially non-single-valued languages is not co-enumerable.

II. We will consider finite transducers, both deterministic (abbreviated DT) and nondeterministic (abbreviated NDT), defined on regular sets. In other words, the set of terminal states will be distinguished in these transducers. Let us give formal definitions. A nondeterministic transducer will be understood to refer to the quintuple  $\langle \Sigma, Q, q_0, F, \delta \rangle$ , where  $\Sigma$  is a finite alphabet, Q a finite set of states,  $q_0$  an initial state,  $q_0 \in Q$ , F a set of terminal states,  $F \subseteq Q$ , and  $\delta$  a finite set of transformations, where  $\delta \subseteq (\Sigma^* \times Q \times \Sigma^* \times Q)$ . We will represent NDT in the form of a directed graph with labeled edges whose vertices are states. With every computation of a NDT we associate a path in the graph. By  $in(\gamma)$  we denote the word which the NDT reads along the path  $\gamma$ , and by  $out(\gamma)$  the word it writes on this path. The pair of words  $\langle u, v \rangle$  belongs to the <u>diagram</u> of a NDT if there exists a path  $\gamma$  starting from an initial state and ending in a terminal state such that  $u = in(\gamma)$  and  $v = out(\gamma)$ . A deterministic transducer is an NDT such that  $\delta \subseteq (\Sigma \times Q \times \Sigma^* \times Q)$  and such that for any  $q \in Q$  and any letter  $u \in \Sigma$  there exist unique  $w \in \Sigma^*$  and  $q_1 \in Q$  such that  $(u, q, w, q_1) \in \delta$ . It is known that the problem of equivalence of arbitrary DT is solvable (cf. [4, pp. 322-323], where it is shown that transducers are equivalent if their diagrams coincide).

On the other hand, the problem of equivalence for NDT is not solvable [4, pp. 322-326]. The diagram of a mapping specified by a transformation  $\mathfrak{A}$  will be denoted  $\Gamma_{\mathfrak{A}}$ . The following assertion will play a fundamental role in the discussion.

THEOREM 1. Suppose that  $\mathfrak{A}_1,\ \ldots,\ \mathfrak{A}_n$  are arbitrary DT and that  $\mathfrak{B}$  is an arbitrary NDT. Then

(a) the following equality is solvable:

 $\Gamma_{\mathfrak{V}}=\Gamma_{\mathfrak{U}_{1}}\cup\ldots\cup\Gamma_{\mathfrak{U}_{n}},$ 

Moscow State University. Translated from Matematicheskie Zametki, Vol. 50, No. 1, pp. 34-40, July, 1991. Original article submitted April 4, 1988; revision submitted November 10, 1989.

(b) the following inclusion is solvable:

$$\Gamma_{\mathfrak{B}} \subseteq \Gamma_{\mathfrak{A}_1} \cup \ldots \cup \Gamma_{\mathfrak{A}_m};$$

(c) if the inclusion (1) holds, there exist regular languages  $R_1, R_2, \ldots, R_n$  such that  $\Gamma_{\mathfrak{B}} = \Gamma_{\mathfrak{A}_1 \upharpoonright R_1} \cup \ldots \cup \Gamma_{\mathfrak{A}_n \upharpoonright R_n}$ ; the languages  $R_1, \ldots, R_n$  may be effectively constructed; by  $\mathfrak{A} \upharpoonright R$  we denote the DT  $\mathfrak{A}_R$  whose diagram coincides with the set  $\{r, \mathfrak{A}(r) \mid r \equiv R\}$ ; it is obvious that such a DT exists;

(1)

(d) if  $\mathfrak{A}$  is a DT and  $\mathfrak{B}$  is a NDT, the inclusion problem  $\Gamma_{\mathfrak{A}} \subseteq \Gamma_{\mathfrak{B}}$  is not solvable.

<u>Proof.</u> Assertion (a) follows at once from (b) and (c). In fact, once the inclusion (b) has been verified, we may, by (c), expand  $\mathfrak{B}$  into a union of DT. Next we verify the inverse inclusion. To prove (d), note that Post's correspondence reduces to proving the inclusion  $\Gamma_{\mathfrak{A}_{1}} \subseteq \overline{\Gamma_{\mathfrak{A}_{2}}}$ , where  $\mathfrak{A}_{1}$  and  $\mathfrak{A}_{2}$  are DTs with a single state. It is known that  $\overline{\Gamma_{\mathfrak{A}}}$ , may be specified by means of an NDT. Let us describe this NDT. Suppose that  $\mathfrak{A}_{2}$  corresponds to a morphism g and that the letter a runs through the alphabet  $\Sigma$ . We enumerate all possible transitions, specifying the word just read to the left of the incremental state, and the word just written on the right:

$$\begin{array}{ll} q_1 \rightarrow aq_1g \ (a); & q_1 \rightarrow q_2a; & q_1 \rightarrow aq_3x; & q_1 \rightarrow aq_4y; \\ q_2 \rightarrow q_2a; & q_3 \rightarrow aq_3; & q_4 \rightarrow q_2; & q_4 \rightarrow aq_4. \end{array}$$

Here  $q_1$  is the initial state,  $q_2$  and  $q_3$  are terminal states, x runs through the set of words from  $\Sigma^*$  with length less than |g(a)|, and y is the set of words from  $\Sigma^*$  with length equal to |g(a)|, but not equal to g(a). Obviously, this NDT is the desired NDT. Step (d) is proved.

Let us prove (b). Suppose that  $\Sigma$  is an alphabet consisting of all the transducers we are considering. If  $w \in \Sigma^*$  and  $\mathfrak{A}$  is a DT,  $\mathfrak{A}(w)$  will denote the word that appears at the output of  $\mathfrak{A}$ , if the word w is fed to its input.

<u>Definition</u>. Suppose that  $\langle w_1, w_2 \rangle$  is a pair of words from  $\Sigma^*$  and let  $\mathfrak{A}$  be DT. An element of a free group with genetrices from  $\Sigma$  and equal to  $w_2^{-1}\mathfrak{A}(w_1)$  will be called the <u>defect</u> of  $\langle w_1, w_2 \rangle$  relative to  $\mathfrak{A}$ .

According to the basic property of the concept of defect,  $w_2 = \mathfrak{A}(w_1) \Leftrightarrow$  the defect of  $\langle w_1, w_2 \rangle$  relative to  $\mathfrak{A}$  is equal to 1. By the defect of a path  $\gamma$  we will understand the defect of the pair  $\langle in(\gamma), out(\gamma) \rangle$ . By a <u>deduction</u> in an NDT we will understand a path from an initial state to a terminal state. It may be assumed that, in all operational cycles, in an NDT  $\mathfrak{B}$  letters are added to one and only one of the words (either a word that has been read or one that has been written). An arbitrary path  $\gamma$  in an NDT will be said to be <u>admissible</u> relative to  $\mathfrak{A}$  if  $out(\gamma) = \mathfrak{A}(in(\gamma))$ .

LEMMA 1.1. Let  $\gamma_1$  and  $\gamma_2$  be paths in some NDT into some state q, and let  $\delta_1$  and  $\delta_2$  be paths in this NDT emanating from q. Suppose that the DT  $\mathfrak{A}$  reaches the words in( $\gamma_1$ ) and in ( $\gamma_2$ ) while in the same state q'. Then if three of the four paths  $\gamma_1\delta_1$ ,  $\gamma_1\delta_2$ ,  $\gamma_2\delta_1$ ,  $\gamma_2\delta_2$  are admissible relative to  $\mathfrak{A}$ , so is the fourth path.

<u>Proof.</u> We let  $\beta_1$  (correspondingly,  $\beta_2$ ) denote the result of the operation  $\mathfrak{A}$ , beginning with state q', on the input in  $(\delta_1)$  [correspondingly, in  $(\delta_2)$ ]. Then the lemma may be restated in the following way. If three of the following equalities

 $\begin{array}{l} (\operatorname{out} (\delta_1))^{-1} (\operatorname{out} (\gamma_1))^{-1} \mathfrak{A} (\operatorname{in} (\gamma_1)) \ \beta_1 = 1, \\ (\operatorname{out} (\delta_2))^{-1} (\operatorname{out} (\gamma_1))^{-1} \mathfrak{A} (\operatorname{in} (\gamma_1)) \ \beta_2 = 1, \\ (\operatorname{out} (\delta_1))^{-1} (\operatorname{out} (\gamma_2))^{-1} \mathfrak{A} (\operatorname{in} (\gamma_2) \ \beta_1 = 1, \\ (\operatorname{out} (\delta_2))^{-1} (\operatorname{out} (\gamma_2))^{-1} \mathfrak{A} (\operatorname{in} (\gamma_2)) \ \beta_2 = 1 \end{array}$ 

are true, then so is the fourth. In this formulation the lemma is self-evident. If, for example, the first three equalities hold, we find from the first and third equality that  $(out(\gamma_2))^{-1}\mathfrak{A}(in(\gamma_2)) = (out(\gamma_1))^{-1}\mathfrak{A}(in(\gamma_1))$ , and this result, together with the second equality, yields the fourth equality.

<u>Ramsey's Graph Theorem.</u> For any pair of natural numbers m and n, there exists a number k such that any complete graph (i.e., a graph possessing all possible edges) with at least k vertices whose edges are colored in n colors possesses a complete subgraph with m vertices whose edges are colored in a single color.

Obviously, there exists a sorting algorithm that computes k, given m and n. We will use this theorem only for m equal to 3. We denote the computable function that yields k given n in this case by R(n).

To prove (b), it suffices to show that, given  $\mathfrak{A}_1, \ldots, \mathfrak{A}_n, \mathfrak{B}$ , it is possible to find a constant  $C_1$  such that if, for all deductions  $\gamma$  in  $\mathfrak{B}$  of length not greater than  $C_1$ , the pair  $\langle \operatorname{in}(\gamma), \operatorname{out}(\gamma) \rangle$  belongs to  $\Gamma_{\mathfrak{A}_1} \cup \cdots \cup \Gamma_{\mathfrak{A}_n}$ , then all of  $\Gamma_{\mathfrak{B}}$  is embedded in  $\Gamma_{\mathfrak{A}_1} \cup \cdots \cup \Gamma_{\mathfrak{A}_n}$ .

LEMMA 1.2. Given any finite set of DT A and any NDT  $\mathfrak{B}$ , it is possible to find a constant C with the following property. If there is a path  $\gamma_0$  in  $\mathfrak{B}$  into some state q such that for all paths  $\gamma$  of length at most C out of state q into a terminal state, the path  $\gamma_0\gamma$  is admissible with respect to at least one DT from A, then for any path  $\gamma$  in  $\mathfrak{B}$  issuing from q and leading to a terminal state, the path  $\gamma_0\gamma$  is admissible with respect to at least one DT from A, then for each terminal state one DT belonging to A. Replacing A by any one of its subsets will not increase the value of C.

<u>Proof.</u> Suppose that  $A = \{\mathfrak{D}_1, \ldots, \mathfrak{D}_l\}$ . We denote the set of paths from q into a terminal state by M. The length of the path  $\gamma$  is denoted by  $|\gamma|$ . Suppose that K is the set of states in  $\mathfrak{B}$ , and let m be the number of pairs  $\langle i, \text{ state } \mathfrak{D}_i \rangle$ , where  $i \in \{1, 2, \ldots, l\}$ . We let  $C = K \cdot R(\mathfrak{m})$ . Let us assume that the assertion claimed by Lemma 1.2 is false. Then for some  $\gamma \in M$  such that  $|\gamma| > C$ , the pair  $\langle in(\gamma_0 \gamma), out(\gamma_0 \gamma) \rangle$  belongs to  $\Gamma_{\mathfrak{T}_1} \cup \ldots \cup \Gamma_{\mathfrak{T}_l}$ , but for all  $\gamma' \in M$  such that  $|\gamma'| < |\gamma|$ , does not. Since  $|\gamma| > K \cdot R(\mathfrak{m})$ , there exists a state q of the transducer  $\mathfrak{B}$ , through which  $\gamma$  passes at least  $R(\mathfrak{m})$  times. Let us determine this state. With each  $1 \leq i \leq R(\mathfrak{m})$  we associate  $\gamma_i$ , the segment of  $\gamma$  until the i-th pass through q and  $\delta_i$  the segment of  $\gamma$  after the i-th pass through q. Let us consider the path  $\gamma_i \delta_j$  belongs to  $\Gamma_{\mathfrak{T}_1} \cup \ldots \cup \Gamma_{\mathfrak{T}_l}$ . We associate with every pair i  $\langle j | \text{the pair } (\operatorname{in}(\gamma_0 \gamma_i \delta_j) \rangle$  belongs to  $\Gamma_{\mathfrak{T}_1} \cup \ldots \cup \Gamma_{\mathfrak{T}_l}$ . We associate with every pair i  $\langle j | \text{the path } (\gamma_i \gamma_i \delta_j) \rangle$  belongs to  $\Gamma_{\mathfrak{T}_1} \cup \ldots \cup \Gamma_{\mathfrak{T}_l}$ . We associate with every pair i  $\langle j | \text{the path } (\gamma_0 \gamma_i \delta_j) \rangle$  belongs to  $\Gamma_{\mathfrak{T}_1} \cup \ldots \cup \Gamma_{\mathfrak{T}_l}$ . We associate with every pair i  $\langle j | \text{the pair } (\operatorname{in}(\gamma_0 \gamma_i \delta_j) \rangle$  and the states of this DT in  $\operatorname{in}(\gamma_0 \gamma_i \delta_j)$ . By Ramsey's theorem there exist numbers  $n_1 < n_2 < n_3$  such that the same pair (i, q) corresponds to the three pairs  $\langle n_1, n_2 \rangle, \langle n_1, n_3 \rangle, \langle n_2, n_3 \rangle$ . Then Lemma 1.1 is applicable to the path  $\gamma_0 \gamma_{n_1} \delta_{n_2}$ ,  $\delta_{n_3}$ , since by construction the three paths  $\gamma_0 \gamma_{n_1} \delta_{n_2}$ ,  $\gamma_0 \gamma_{n_1} \delta_{n_3}$ ,  $\gamma_0 \gamma_{n_2} \delta_{n_3}$  are admissible with respect to  $\mathfrak{D}_i$ . Lemma 1.2 is proved. To prove (b), it suffices to apply Lemma 1.2, where  $\gamma_0$  is the empty path, q an initial state, and  $A = \{\mathfrak{A}_1, \ldots, \mathfrak{A}_n\}$ .

Let us prove (c). The length of a defect will be assumed to equal the length of its noncontractable representation.

LEMMA 1.3. There exists a constant  $C_2$  that is computable in  $\mathfrak{B}, \mathfrak{A}_1, \ldots, \mathfrak{A}_n$  with the following property:

If the inclusion (1) holds, then for any deduction  $\gamma$  in  $\mathfrak{B}$ , a number i may be found such that  $\langle in(\gamma), out(\gamma) \rangle \subset \Gamma_{\mathfrak{A}_i}$ , and such that for any initial segment  $\gamma_1$  of  $\gamma$ , the length of the defect of  $\gamma_1$  relative to  $\mathfrak{A}_i$  is not greater than  $C_2$ .

<u>Proof.</u> We take  $C_2 = C \cdot S$ , where C is the constant of Lemma 1.2 for the case  $A = \{\mathfrak{A}_1, \ldots, \}$  $\mathfrak{A}_n\},$  and S is the greatest number such that the length of a word written in the DT  $\mathfrak{A}_1,\ldots,\mathfrak{A}_n$ may increase as a single letter is being read. Let us determine an arbitrary deduction  $\gamma$  in  $\mathfrak{F}$ . Suppose that  $A = {\mathfrak{A}_1, \ldots, \mathfrak{A}_n}$ , and let B be the set of all those DT in A whose diagrams belong to  $\langle in(\gamma), out(\gamma) \rangle$ . Let us assume that our assertion is false in the case of  $\gamma$ . With every transducer  $\mathfrak{A}_j igodot B$  we associate some initial segment of  $\gamma$  whose defect relative to  $\mathfrak{A}_j$ is greater than  $C_2$ . Let us order all DTs belonging to B in increasing length of the initial segments corresponding to them:  $\mathfrak{A}_{i_1}, \mathfrak{A}_{i_2}, \ldots, \mathfrak{A}_{i_{|R|}}$  (if several DTs correspond to a single initial segment of  $\gamma$ , we arrange them in arbitrary order). We denote by  $\gamma_k$  the initial segment of  $\gamma$  corresponding to the DT  $\mathfrak{A}_{i_k}$ . By induction on k, we prove that any deduction in  $\mathfrak{B}$  containing its own initial segment  $\boldsymbol{\gamma}_k$  is admissible with respect to any DT belonging to the set  $A \searrow \{\mathfrak{A}_{i_1},\ldots,\mathfrak{A}_{i_k}\}$ . Suppose that the assertion has been proved in the case k' < k. Let us consider all deductions in  ${\mathfrak B}$  with initial segment  $\gamma_k$  and continuation of length not greater than C2. Since the length of the defect  $\gamma_k$  relative to  $\mathfrak{A}_{i_k}$  is greater than C2, none of these deductions is admissible with respect to  $\mathfrak{A}_{i_k}$ , since the length of a defect may decrease by at most one step in moving along the path. By the induction hypothesis, these deductions are admissible with respect to any DT belonging to  $A \setminus \{\mathfrak{A}_{i_1}, \ldots, \mathfrak{A}_{i_{k-1}}\}$  [if k = 1 this follows from (1)]. Consequently they are admissible also with respect to any DT that belongs to  $A \setminus \{\mathfrak{A}_{i_1}, \ldots, \mathfrak{A}_{i_k}\}$ . By Lemma 1.2, all deductions with initial segment  $\gamma_k$  are admissible with respect to some DT in  $A \setminus \{\mathfrak{A}_{i_1}, \ldots, \mathfrak{A}_{i_k}\}$ . For k = |B|, we obtain a contradiction. Lemma 1.3 is proved.

It is easily seen that for any C and any i, the set of words that are read on those deductions in  $\mathfrak{B}$  in which the length of the defect of any initial segment relative to  $\mathfrak{A}_i$  is at most C is regular, and that this set may be effectively constructed from C. In fact, an automaton that recognizes this set possesses as its states the triple <state  $\mathfrak{B}$ , state  $\mathfrak{A}_i$ , defect>, where the length of the defect does not exceed C. Since a defect and a pair of increments to inscribed words determine a new defect in unique fashion, the transitions of the automaton are determined in obvious fashion.

From Lemma 1.3 it follows that the desired regular sets  $R_i$ , i = 1,...,n, constitute the set of words that are read on those deductions in  $\mathfrak{B}$ , in which the length of the defect of any initial segment relative to  $\mathfrak{A}_i$  is not greater than  $C_2$ . Step (c), and along with it Theorem 1, are proved.

Let us now consider context-free grammars. With every transducer  $\mathfrak{B}$  we associate a language  $L_{\mathfrak{B}} = \{w \mapskip u^r | \langle w, u \rangle \in \Gamma_{\mathfrak{B}}\}$ , here  $u^r$  denotes the word which is the mirror image of u. Recall that a c-linear context-free grammar is understood to refer to a linear context-free grammar in whose alphabet a special terminal symbol # (called a <u>marker</u>) is provided and such that the right side of any rule either does not contain a marker but contains a nonterminal symbol, or is in fact a marker. It is easily seen that for any  $\mathfrak{D}$  the language  $L_{\mathfrak{B}}$  is generated by some c-linear context-free grammar, and conversely any c-linear context-free grammar generates a language of this form. Moreover, if  $\mathfrak{A}_1, \ldots, \mathfrak{A}_n$  are arbitrary transducers, a c-linear context-free grammar that generates the language  $L_{\mathfrak{B}}$ ,  $\bigcup \ \ldots \ \sqcup \ \mathfrak{L}_{\mathfrak{A}_n}$  may be constructed in a standard fashion. We will call it  $G(\mathfrak{A}_1 \cup \ldots \cup \mathfrak{A}_n)$ . We will also say that a DT  $\mathfrak{A}$  possesses finite delay if there exists a natural number C such that, if it begins to function in some arbitrary state, once it has read any C letters,  $\mathfrak{A}$  will output at least one letter. The language that is generated by an arbitrary context-free grammar K is denoted  $L_{K}$ . A grammar of the form  $G(\mathfrak{A}_1 \cup \ldots \cup \mathfrak{A}_n)$ , where  $\mathfrak{A}_1, \ldots, \mathfrak{A}_n$  are DT with finite delay, will be called a

diagrammed grammar.

<u>THEOREM 2.</u> Suppose that  $G = G(\mathfrak{A}_1 \cup \ldots \cup \mathfrak{A}_n)$  is an arbitrary diagrammed context-free grammar, and let K be an arbitrary context-free grammar. Then

- (a) the problem  $L_K = L_G$  is solvable;
- (b) the problem  $L_K \subseteq L_G$ ; is solvable;
- (c) if  $L_K \subseteq L_G$ , a diagrammed context-free grammar G that is equivalent to K may be constructed;
- (d) the set of essentially distinct diagrammed context-free grammars is not co-enumerable.

## Proof.

<u>LEMMA 2.1.</u> Given an arbitrary context-free grammar K and an arbitrary diagrammed context-free grammar G, it is possible effectively to either prove that  $L_K \subseteq L_G$ , or construct a c-linear context-free grammar K<sub>1</sub> that is equivalent to K.

<u>Proof.</u> Obviously, if  $L_K \not\subseteq L_G$ , it will not be possible for infinitely many segments of some word in  $L_K$  following (preceding) some marker to correspond to a fixed segment of this word preceding (following) this marker. Therefore, if  $L_K \subseteq L_G$ , at most one nonterminal symbol on the right side of every production in K will produce an infinite number of words in  $L_K$ . Since the finiteness problem of a context-free language is solvable (cf. [5, p. 305]), the latter condition may be checked and, if it holds, a linear context-free grammar K' may be constructed that is equivalent to K. For the reason given above, in K' only a finite number of words in  $L_{K'}$  may be produced from any nonterminal symbol on the right side of any production containing a marker. Once this condition has been checked, the desired c-linear context-free grammar K<sub>1</sub> is easily constructed. Lemma 2.1 is proved.

Steps (a), (b), and (c) of Theorem 2 follow immediately from Theorem 1 and Lemma 2.1. Assertion (d) of Theorem 2 was proved by Gladkii [6]. Theorem 2 is proved.

The next theorem contains an answer to the question posed in the introduction.

THEOREM 3. There does not exist an enumerable set of context-free grammars (not necessarily single-valued) whose elements generate precisely all the single-valued context-free languages. <u>Proof.</u> If such a set were to exist, it would follow from Theorem 2 (a) that all the diagrammed grammars that generate single-valued languages could be enumerated. But this contradicts assertion (d) of the theorem.

<u>Remark.</u> It has come to my attention that assertions (a) and (b) of Theorem 1 follow from results in [7] together with a result from [8] and the results of [9].

I would like to express my appreciation to An. A. Muchnik and N. K. Vereshchagin for the statement of the problem and for useful discussion.

## LITERATURE CITED

- A. L. Semenov, "Algorithmic problems for power series and context-free grammars," Dokl. Akad. Nauk SSSR, <u>212</u>, No. 1, 50-52 (1973).
- 2. An. A. Muchnik, "Application of Semenov's method to the analysis of the structure of context-free languages," in: Semiotic Aspects of Formalization of Mental Activity [in Russian], VINITI, Moscow (1985), pp. 212-214.
- 3. J. E. Hopcroft, "On the equivalence and containment problems for context-free languages," Math. Syst. Theor., <u>3</u>, No. 2, 119-124 (1969).
- 4. W. Brauer, Introduction to the Theory of Finite Automata [Russian translation], Radio i Svyaz', Moscow (1987).
- 5. H. Lalleman, Semigroups and Combinatorial Applications [Russian translation], Mir, Moscow (1985).
- 6. A. V. Gladkii, "The essential indeterminacy of context-free languages is not algorithmically recognizable," Algebra Logika, <u>4</u>, No. 4, 53-64 (1965).
- 7. K. Culik and J. Karhumäki, "The equivalence of finite-valued transducers (on HDTOL languages) is decidable," Theor. Comput. Sci., <u>47</u>, 71-84 (1986).
- 8. E. Gurari and O. Ibarra, "A note on finite-valued and finitely ambiguous transducers," Math. Systems Theory, 16, 61-66 (1983).
- 9. .A. Weber, "A decomposition theorem for finite valued transducers and application to the equivalence problem," unpub. ms. (1987).

## APPROXIMATIVE EVALUATION OF THE HEIGHT OF THE MAXIMAL UPPER ZERO

OF A MONOTONE BOOLEAN FUNCTION

A. Yu. Kitaev

UDC 517.1

Many numerical problems of the practical importance can be formulated in terms of a problem of finding the maximal upper zero of a certain monotone Boolean function f, computable according to a beforehand known algorithm. As an example, we may consider the search for the maximal solvable subsystem of a given system of inequalities [1]. In connection with this problem, there arises a question of the computational complexity of the height of the upper zero of a monotone Boolean function, defined with the help of an oracle. There exist two natural measurements of the complexity: the number of inquires to the oracle, and the full number of steps in the algorithm. The minimal number of inquiries to the oracle sufficient for defining the height of a maximal upper zero of an arbitrary monotone function of n variables is equal to  $\binom{n}{\lfloor n/2 \rfloor} + 1$  [2]. In the present paper there is investigated a question of the computational complexity of the height of the upper zero. A related problem was considered in [3]; however, in this paper there was used another criterion of the computational exactness (cf. also [4]).

Let  $B^n = \{0, 1\}^n$  be the n-dimensional Boolean cube, and  $F_n^+$  be the set of all monotone Boolean functions with n parameters, satisfying the condition  $f(0, \ldots, 0) = 0$ . For a Boolean vector  $u \in B^n$ , by h(u) we denote the number of ones in the vector  $(u_1, \ldots, u_n)$ . The set  $S_k = \{u \in B^n \mid h(u) = k\}$  is called a layer with height h in the Boolean cube. The height of the

L. D. Landau Institute of Theoretical Physics, Academy of Sciences of the USSR. Translated from Matematicheskie Zametki, Vol. 50, No. 1, pp. 41-45, July, 1991. Original article submitted April 18, 1989.