Časopis pro pěstování matematiky, roč. 103 (1978), Praha

A DETERMINISTIC SUBCLASS OF CONTEXT-FREE LANGUAGES

JAN PECKEL, Praha

(Received March 31, 1977)

INTRODUCTION

G. WECHSUNG in [1] has introduced a new complexity measure and has proved that the class of all context-free languages turns out to be a complexity class with respect to this measure for nondeterministic Turing machines.

We investigate the complexity class C given by the same bound and complexity measure for deterministic Turing machines in this paper. Namely, the relation of this complexity class to the class of all deterministic context-free languages is studied. It is proved that these two classes of languages are incomparable. Moreover, similar incomparability result is proved for the class C and the class of all linear languages.

WECHSUNG'S COMPLEXITY MEASURE

By a Turing machine (or simply TM) $M = (Q, X, d, q_0, F)$ we shall mean a deterministic one-tape, one-head model of Turing machine with the state space Q, the alphabet X, the next-state function d, the initial state q_0 and the accepting state space F. The alphabet of every TM will contain the blank symbol b. X_b will denote the set $X - \{b\}$.

By a computation of a TM $M = (Q, X, d, q_0, F)$ on a word $w \in X^*$ we shall mean the computation starting in the initial state q_0 on the leftmost symbol of w.

A TM $M = (Q, X, d, q_0, F)$ accepts a word $w \in X_b^*$ iff the computation of M on w halts in an accepting state.

A TM $M = (Q, X, d, q_0, F)$ recognizes a language $L \subseteq X_b^*$ iff for every word $w \in X_b^*$ the following condition holds: $w \in L \Leftrightarrow M$ accepts w.

STUDENTS' RESEARCH ACTIVITY AT THE FACULTY OF MATHEMATICS AND PHYSICS, CHARLES UNIVERSITY. Award the 1st prize in the National Students' Research Work Competition, section Theoretical Cybernetics, in the year 1977. Scientific adviser: Professor M. CHYTIL.

In case that during a computation the content of a tape square is changed, every visit of the head payed to this square after its first altering shall be called an active visit. For every word w accepted by a TM M the maximal number of all active visits on one tape square during the computation of M on w shall be denoted as $g_M(w)$.

Let k be a nonnegative integer.

A TM $M = (Q, X, d, q_0, F)$ recognizes a language $L \subseteq X_b^*$ with Wechsung's complexity k iff 1. M recognizes L and 2. for every word $w \in L$ it is $g_M(w) \leqq k$.

A language L is recognizable with Wechsung's complexity k iff there is a TM recognizing L with Wechsung's complexity k.

NOTATION AND DEFINITIONS

For every nonnegative integer k denote by $W(k)$ the class of all languages recognizable with Wechsung's complexity k. Then

$$\mathrm{C} =_{\mathrm{df}} \bigcup_{k=0}^{\infty} W(k),$$

CFL $=_{\mathrm{df}}$ the class of all context-free languages,

DCFL $=_{\mathrm{df}}$ the class of all deterministic context-free languages,

LIN $=_{\mathrm{df}}$ the class of all linear context-free languages,

$\overleftarrow{w} =_{\mathrm{df}}$ the "mirror image" of the word w,

$\Lambda =_{\mathrm{df}}$ the empty word.

By a numbering of the tape of a TM we shall understand a 1-1 mapping of the set of tape squares into the set of integers. So every tape square has a number, "square p" will denote "the square numbered by p".

Let $M = (Q, X, d, q_0, F)$ be a TM and let k be a nonnegative integer.

If $w \in X^+$ then the symbol $P(w)$ stands for "the part of the tape which was initially occupied by the characters of the input word w". If the tape of M is numbered in such a way that the square p_1 stands to the left from the square p_2, then the symbol $P(p_1, p_2)$ denotes the word formed by the sequence of characters in the squares between p_1 and p_2 ("between squares p_1 and p_2" will always implicitly include "excluding the squares p_1 and p_2").

Definition 1. Two words $u, v \in X^+$ are said to be E_1-equivalent (notation $u \sim_{E_1} v$) iff for arbitrary states $q, q' \in Q$ the following conditions hold:

1. [If M starts in the state q on the leftmost (rightmost) symbol of the word u, then M changes the content of $P(u)$ without leaving it before]

$$\Leftrightarrow$$

[If M starts in the state q on the leftmost (rightmost) symbol of the word v, then M changes the content of $P(v)$ without leaving it before].

2. [If M starts in the state q on the leftmost (rightmost) symbol of the word u, then the first exit from $P(u)$ is made leftwards in the state q']

$$\Leftrightarrow$$

[If M starts in the state q on the leftmost (rightmost) symbol of the word v, then the first exist from $P(v)$ is made leftwards in the state q'].

3. [If M starts in the state q on the leftmost (rightmost) symbol of the word u, then the first exit from $P(u)$ is made rightwards in the state q']

$$\Leftrightarrow$$

[If M starts in the state q on the leftmost (rightmost) symbol of the word v, then the first exit from $P(v)$ is made rightwards in the state q'].

4. [If M starts in the state q on the leftmost (rightmost) symbol of the word u, then M enters an accepting state without leaving $P(u)$ before]

$$\Leftrightarrow$$

[If M starts in the state q on the leftmost (rightmost) symbol of the word v, then M enters an accepting state without leaving $P(v)$ before].

For any $u \in X^+$ and $q \in Q$, the symbols $(u)_{qL}$ and $(u)_{qR}$ will denote, respectively, the content of the tape segment $P(u)$ after the first exit from $P(u)$, provided the TM M has started on the leftmost or rightmost symbol of the word u in the state q. If M does not leave the segment, the meaning of the symbols is not defined.

Definition 2. Two words $u, v \in X^+$ are said to be E_2-equivalent iff for an arbitrary sequence $q_1, A_1, q_2, A_2, \ldots, q_j, A_j$
where $j \in N$, $j \leqq 2k$,

$$A_i = \text{either } L \text{ or } R \text{ for } i = 1, 2, \ldots, j,$$

$$q_i \in Q \text{ for } i = 1, 2, \ldots, j,$$

the following condition holds:

If at least one of the symbols $(\ldots((u)_{q_1A_1})_{q_2A_2}\cdots)_{q_jA_j}$ and $(\ldots((v)_{q_1A_1})_{q_2A_2}\cdots)_{q_jA_j}$ is meaningful, then both of them are meaningful and at the same time

$$(\ldots((u)_{q_1A_1})_{q_2A_2}\cdots)_{q_jA_j} \sim_{E_1} (\ldots((v)_{q_1A_1})_{q_2A_2}\cdots)_{q_jA_j}.$$

Remark. For $j = 0$ the last relation has the form $u \sim_{E_1} v$.

Both above defined equivalences have a finite number of classes.

Lemma 1. *Let a TM $M = (Q, X, d, q_0, F)$ have s states. Let a tape segment contain the word z^s, where $z \in X$. If M enters this tape segment from the left or right and passes through it rightwards or leftwards, respectively, without any rewriting, then the first rewriting of a tape square cannot be performed before scanning a symbol different from z.*

The proof is obvious and follows from the fact that M must reach (at least) twice the same state when scanning the word z^s.

Lemma 2. *Let a* TM *M recognize a language L with Wechsung's complexity k, where $k \in N$. Then there exists such a positive integer l that during the computation of M on any word $w \in L - \{\Lambda\}$ the head reaches maximally l-1 squares out of $P(w)$.*

For the proof cf. [1].

Theorem 1. DCFL *and* C *are incomparable, i.e.* DCFL $\nsubseteq$ C & C $\nsubseteq$ DCFL.

Proof. (1.1) Let us consider the language $L = \{w\overleftarrow{w};\ w \in \{a, c\}^+\}$. It follows from [2] that $L \notin$ DCFL. We can construct a TM $M = (Q, X, d, q_0, F)$ where $X = \{a, c, b\}$ so that the computing process of M on an arbitrary word $w \in X_b^+$ will proceed as follows:

1. M will check if the leftmost of the squares of $P(w)$ which have not been rewritten contains the same character as the rightmost of the squares of $P(w)$ which have not been rewritten and if moreover these two squares are not identical. If it is so the both squares will be rewritten by the character b and then
   - either the activity No. 1 will proceed, in case some squares of $P(w)$ have not been rewritten
   - or the activity No. 2 will proceed, in case all squares contain the character b.

   If it is not so the activity No. 3 will proceed.
2. M will reach an accepting state.
3. The computation will halt in a situation for which the next-state function is not defined.

It is obvious that the next-state function of such a TM can be defined in such a way that during the computation of M on an arbitrary word $w \in \{a, c\}^+$ there will not appear more than one active visit on any square. It follows from this fact that $L \in W(1)$.

(1.2) The converse will be proved by contradiction. Consider the language $L = \{a^m c^{m+n} a^n;\ m, n = 1, 2, \ldots\}$. It is obvious that $L \in$ DCFL. Assume that $\hat{M} = (\hat{Q}, X, \hat{d}, q_0, F)$ is such a TM which recognizes L with Wechsung's complexity k, where k is a nonnegative integer. Let $q_1, q_2 \notin \hat{Q}$. Define $M = (Q, X, d, q_0, F)$, where $Q = \hat{Q} \cup \{q_1, q_2\}$ and the next-state function d is defined in the following way:

$d(q, z) = \hat{d}(q, z)$ if $(q, z) \in \hat{Q} \times X$ and $\hat{d}(q, z)$ is defined,
$d(q_1, b) = (q_1, b, R)$,
$d(q_1, a) = (q_2, a, L)$,
$d(q_2, b) = (q_0, b, R)$,
d is not defined for other arguments.

Now introduce for the TM M and k the equivalences E_1 and E_2 on X^+ according to Definitions 1 and 2. E_1 and E_2 are of finite indices, say e_1 and e_2, respectively.

Remark. If $u = b^{n_1}au_1$, $v = b^{n_2}av_1$, where $n_1, n_2 \in N$, $u_1, v_1 \in X^*$ and $u \sim_{E_1} v$, then for any state $q' \in Q$ the points 1, 2, 3, and 4 of Definition 1 hold even if we replace the words "If M starts in the state q on the leftmost (rightmost) symbol of the word u" by the words "If M starts in the state q_0 on the leftmost of the nonblank symbols of u" and if we replace the words "If M starts in the state q on the leftmost (rightmost) symbol of the word v" by the words "If M starts in the state q_0 on the leftmost of the nonblank symbols of v". This fact has been used in the proof and for this reason the TM $\hat{M}$ was extended to the TM M.

Now let us enumerate the tape of M as indicated by Fig. 1.

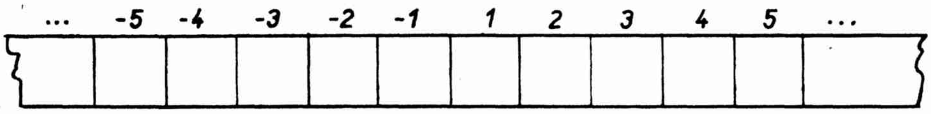


Figure 1.

Let positive integers m_1 and m_2, where $m_1 < m_2 \leqq e_1 + 1$, satisfy $a^{m_1} \sim_{E_1} a^{m_2}$ (such numbers can be found). Let l be a positive integer satisfying the assertion of Lemma 2 for the given TM M and for the language L. Define $s = \text{card } Q$.

Consider the word $w = a^n c^{2n} a^n$, where $n \in N$, $n \geqq s((s + 1)(e_2 + l - 1) + l - 2) + \max\{s + 1, m_1\}$. It holds that $w \in L$, so during the computation of M on w, not more than k active visits on any square will appear and M will reach an accepting state.

Now place w on the tape in such a way that the leftmost character of the word w will be written in the square $-2n$.

Lemma 3. *Neither between the squares $-2n - 1$ and $-n$ nor between the squares n and $2n + 1$ there exist m_1 adjacent squares the contents of which would not be changed during the computation of M on w.*

Proof. By contradiction. Let there be m_1 squares of the above described property. Let us form a word u by replacing the word situated in the assumed m_1 squares by the word a^{m_2} in the word w. M accepts u, but $u \in \{a, c\}^+ - L$.

We shall choose tape squares p_i^L and p_i^R (for $i = 1, 2, \ldots, (s+1)(e_2 + l - 1) + l$) inductively as follows:

Define $p_1^L = -2n - l$ and $p_1^R = 2n + l$.

Let p_i^L and p_i^R be defined for a positive integer i, $i < (s+1)(e_2 + l - 1) + l$. Then by Lemmas 1 and 3, there is a tape square p_i such that during the computation of M on w, p_i is rewritten as the first of squares between p_i^L and p_i^R. Then define

$$\left.\begin{aligned} p_{i+1}^L &= p_i \\ p_{i+1}^R &= p_i^R \end{aligned}\right\} \text{ if } p_i \leqq \max\{p_i^L, -2n-1\} + s \text{ or } n - s < p_i \leqq n + s\,,$$

and

$$\left.\begin{aligned} p_{i+1}^L &= p_i^L \\ p_{i+1}^R &= p_i \end{aligned}\right\} \text{ otherwise.}$$

Lemma 4. *Let i be a positive integer, $i \leqq (s+1)(e_2 + l - 1) + l$. During the computation of M on w the head can enter the part of the tape between the squares p_i^L and p_i^R at most $2k+1$-times, after rewriting these two squares.*

Proof. Let the squares p_i^L and p_i^R, where $i \in N$, $0 < i \leqq (s+1)(e_2 + l - 1) + l$, be rewritten during the computation of M on w and let then the head enter the tape segment between the squares p_i^L and p_i^R more than $2k+1$-times. At the same time at least $k+1$ active visits on the square p_i^L or p_i^R must appear.

In the following paragraphs (1.2.1) and (1.2.2), we distinguish two possible situations. (1.2.2) is again decomposed into two parts. Each of the situations leads to a contradiction as shown in the paragraph (1.2.3).

(1.2.1) Assume that for $i = 1, 2, \ldots, e_2 + 2l - 1$ the condition $p_i^L \leqq n - s$ & & $p_i^R \geqq -n + s$ holds. Then among the words $P(p_1^L, p_1^R), P(p_2^L, p_2^R), \ldots, P(p_{e_2+2l-1}^L, p_{e_2+2l-1}^R)$ there exists a pair of E_2-equivalent words such that the difference between the number of the characters c and the number of the characters a in one word is smaller than the difference between the number of the characters c and the number of the characters a in the second word. Let such a pair be formed for instance by the words $P(p_i^L, p_i^R)$ and $P(p_j^L, p_j^R)$, where $i, j \in N$, $0 < i < j < e_2 + 2l$.

The proof continues at (1.2.3).

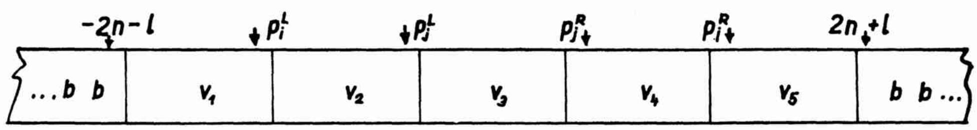
(1.2.2) Assume that $p_{e_2+2l-1}^L > n - s$. For $p_{e_2+2l-1}^R < -n + s$ the proof is quite analogous.

(1.2.2.1) Let for an integer i_0 such that $e_2 + 2l - 2 \leqq i_0 \leqq (s+1) \cdot (e_2 + l - 1) - e_2$ the condition $p_{i_0+1}^L = p_{i_0+2}^L = \ldots = p_{i_0+e_2+l}^L$ hold. Then among the words $P(p_{i_0+1}^L, p_{i_0+1}^R), P(p_{i_0+2}^L, p_{i_0+2}^R), \ldots, P(p_{i_0+e_2+l}^L, p_{i_0+e_2+l}^R)$ there exists a pair of E_2-equivalent words such that the difference between the number of the characters c and the number of the characters a in one word is smaller than the difference between the number of the characters c and the number of the characters a in the second word. Let such a pair be formed for instance by the words $P(p_i^L, p_i^R)$ and $P(p_j^L, p_j^R)$, where $i, j \in N$, $i_0 < i < j \leqq i_0 + e_2 + l$.

The proof continues at (1.2.3).

(1.2.2.2) Let the introductory assumption of the paragraph (1.2.2.1) be not fulfilled. Define $r = s(e_2 + l - 1) + l$. Then $p_r^L \geqq n$. Among the words $P(p_r^L, p_r^R)$, $P(p_{r+1}^L, p_{r+1}^R), \ldots, P(p_{r+e_2+l-1}^L, p_{r+e_2+l-1}^R)$ there exists a pair of E_2-equivalent words such that the number of the characters a in one word is greater than the number of the characters a in the second word (these words do not contain the character c). Let such a pair be formed for instance by the words $P(p_i^L, p_i^R)$ and $P(p_j^L, p_j^R)$, where $i, j \in N$, $r \leqq i < j < r + e_2 + l$.

(1.2.3) Suppose now that on the tape of the TM M the word $w_1 = b^{l-1}wb^{l-1}$ is written in such a way that the leftmost character of the word w_1 is written in the square $-2n - l + 1$. Construct a word u by replacing the tape segment between p_i^L and p_i^R by the word $P(p_j^L, p_j^R)$ in the word w_1 (cf. Fig. 2). If we remove all blank characters b in the word u we shall obtain a word u_1 accepted by M although it holds that $u_1 \in \{a, c\}^+ - L$: a contradiction.



$w_1 = v_1v_2v_3v_4v_5$ $\qquad$ $u = v_1v_3v_5$

Figure 2.

Corollary. C *is a proper subclass of* CFL.

INCOMPARABILITY OF LIN AND C

Theorem 2. LIN *and* C *are incomparable, i.e.* LIN $\nsubseteq$ C & C $\nsubseteq$ LIN.

Proof. (2.1) Consider the language $L = \{a^n c^n a^i;\ i, n = 1, 2, \ldots\} \cup \{a^i c^n a^n;\ i, n = 1, 2, \ldots\}$. It holds that $L \in$ LIN. Assume that $\hat{M} = (\hat{Q}, X, \hat{d}, q_0, F)$ is such a TM which recognizes L with Wechsung's complexity k, where k is a nonnegative integer. Let $q_1, q_2 \notin \hat{Q}$. Define $M = (Q\ X, d, q_0, F)$, where $Q = \hat{Q} \cup \{q_1, q_2\}$ and the next-state function d is defined in the following way:

$d(q, z) = \hat{d}(q, z)$ if $(q, z) \in \hat{Q} \times X$ and $\hat{d}(q, z)$ is defined,
$d(q_1, b) = (q_1, b, R)$,
$d(q_1, a) = (q_2, a, L)$,
$d(q_2, b) = (q_0, b, R)$,
d is not defined for other arguments.

Now introduce the equivalences E_1 and E_2 on X^+ according to Definitions 1 and 2 and denote their indices e_1 and e_2, respectively.

Now enumerate the tape of M as indicated by Fig. 3.

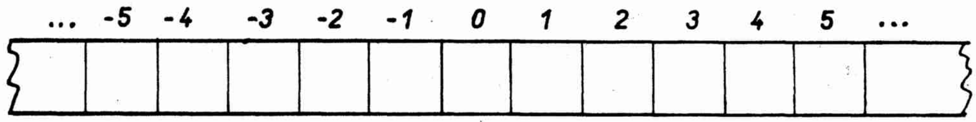


Figure 3.

Let for positive integers m_1, m_2, m_3 and m_4, where $m_1 < m_2 \leqq e_1 + 1 \;\&\; m_3 < m_4 \leqq e_1 + 1$, the condition $a^{m_1} \sim_{E_1} a^{m_2} \;\&\; c^{m_3} \sim_{E_1} c^{m_4}$ hold. Let l be a positive integer satisfying the assertion of Lemma 2 for the given TM M and the language L. Define $s = \text{card } Q$.

Take the word $w = a^n c^n a^n$, where $n \in N$, $n \geq s((e_2 + l - 1)^2 + e_2 + 2s + l - 4) + \max\{s + 1, m_1\}$. It holds that $w \in L$, hence during the computation of M on w at most k active visits on any square will occur and M will reach an accepting state.

Place w on the tape of the TM M in such a way that the leftmost character of the word w will be written in the square 1.

We shall construct inductively sequences

$$p_1^L, p_2^L, \ldots \quad \text{and} \quad p_1^R, p_2^R, \ldots .$$

Define $p_1^L = -l + 1$ and $p_1^R = 3n + l$.

Now let p_i^L and p_i^R for an i be defined. Then if there exists a square $p^{(i)}$ which is rewritten as the first of squares between p_i^L and p_i^R during the computation of M on w, define

$$\left.\begin{aligned} p_{i+1}^L &= p^{(i)} \\ p_{i+1}^R &= p_i^R \end{aligned}\right\} \quad \text{if} \quad p^{(i)} \leqq n - s \quad \text{and}$$

$$\left.\begin{aligned} p_{i+1}^L &= p_i^L \\ p_{i+1}^R &= p^{(i)} \end{aligned}\right\} \quad \text{if} \quad p^{(i)} > 2n + s\,,$$

p_{i+1}^L and p_{i+1}^R are not defined otherwise.

There are two possible cases which are studied in the paragraphs (2.1.1) and (2.1.2) in this proof. Each of this cases is decomposed into a number of subcases which are treated in the corresponding subparagraphs.

(2.1.1) Let the symbols p_i^L and p_i^R be meaningful for $i = (e_2 + l - 1)^2 + 1$.

(2.1.1.1) Let for a nonnegative integer i_0 such that $i_0 \leqq (e_2 + l - 1)(e_2 + l - 2)$, the condition $p_{i_0+1}^L = p_{i_0+2}^L = \ldots = p_{i_0+e_2+l}^L$ hold.

Consider the word $w_1 = a^{n+m_2-m_1} c^n a^n$. It holds that $w_1 \in L$. Place w_1 on the tape of M in such a way that the leftmost symbol of the word w_1 will be written in the square 1. It holds for $i = 1, 2, \ldots, (e_2 + l - 1)^2$ that during the computation of M on w_1, the first of squares between p_i^L and $m_2 - m_1 + p_i^R$ rewritten by M is

- the square p_{i+1}^L if $p_i^L \neq p_{i+1}^L$,
- the square $m_2 - m_1 + p_{i+1}^R$ otherwise.

Among the words $P(p^L_{i_0+l}, m_2 - m_1 + p^R_{i_0+l})$, $P(p^L_{i_0+l+1}, m_2 - m_1 + p^R_{i_0+l+1}), \ldots$ $\ldots, P(p^L_{i_0+e_2+l}, m_2 - m_1 + p^R_{i_0+e_2+l})$ there exists a pair of E_2-equivalent words. Let such a pair be formed by the words $P(p^L_{j_1}, m_2 - m_1 + p^R_{j_1})$ and $P(p^L_{j_2}, m_2 - m_1 + p^R_{j_2})$, where $j_1, j_2 \in N$, $i_0 + l \leqq j_1 < j_2 \leqq i_0 + e_2 + l$.

Now suppose that on the tape of M the word $w_2 = b^{l-1}w_1$ is written in such a way that the leftmost character of the word w_2 is written in the square $-l + 2$. Construct a word u by replacing the tape segment between $p^L_{j_1}$ and $m_2 - m_1 + p^R_{j_1}$ by the word $P(p^L_{j_2}, m_2 - m_1 + p^R_{j_2})$ in the word w_2. If we remove all blank characters b in the word u we shall obtain a word u_1 accepted by M although it holds that $u_1 \in \in \{a, c\}^+ - L$:

$$u_1 = a^{n+m_2-m_1}c^n a^{n_1} \quad \text{where} \quad n_1 \in N\,, \quad n_1 < n\,.$$

(2.1.1.2) Let for a nonnegative integer i_0 such that $i_0 \leqq (e_2 + l - 1)(e_2 + l - 2)$, the condition $p^R_{i_0+1} = p^R_{i_0+2} = \ldots = p^R_{i_0+e_2+l}$ hold. Contradiction can be deduced analogously as in paragraph (2.1.1.1).

(2.1.1.3) Let neither the introductory assumption of the paragraph (2.1.1.1) nor the introductory assumption of the paragraph (2.1.1.2) be fulfilled. Define $r = e_2 + + l - 1$. Among the words $P(p^L_{r(l-1)+1}, p^R_{r(l-1)+1})$, $P(p^L_{r.l+1}, p^R_{r.l+1})$, $P(p^L_{r(l+1)+1}, p^R_{r(l+1)+1}), \ldots, P(p^L_{r.r+1}, p^R_{r.r+1})$ there certainly exists a pair of E_2-equivalent words. Let such a pair be formed by the words $P(p^L_{j_1}, p^R_{j_1})$ and $P(p^L_{j_2}, p^R_{j_2})$, where $j_1 = = ri_1 + 1$, $j_2 = ri_2 + 1$, $i_1, i_2 \in N$, $l - 1 \leqq i_1 < i_2 \leqq r$.

Now construct a word u by replacing the tape segment between $p^L_{j_1}$ and $p^R_{j_1}$ by the word $P(p^L_{j_2}, p^R_{j_2})$ in the word w. M accepts u but $u \in \{a, c\}^+ - L$: $u = a^{n_1}c^n a^{n_2}$ where $n_1, n_2 \in N$, $n_1 < n$, $n_2 < n$. This contradiction completes the paragraph (2.1.1). In the paragraph (2.1.2) the following lemma is used. The proof of the lemma is evident.

Lemma 5. *There do not exist m_3 adjacent squares between n and $2n + 1$ the contents of which would not be changed during the computation of M on w.*

(2.1.2) Let the symbols p^L_i and p^R_i be meaningful for $i = j$, where j is a positive integer such that $j \leqq (e_2 + l - 1)^2$, and not meaningful for $i = j + 1$.

Let p_1 be such a square which is rewritten as the first of the squares between p^L_j and p^R_j during the computation of M on w (by Lemma 5 such a square exists).

(2.1.2.1) Let $n - s < p_1 \leqq n + s$.

Consider the word $w_1 = a^n c^n a^{n+m_2-m_1}$. It holds that $w_1 \in L$. Place w_1 on the tape of M in such a way that the leftmost symbol of the word w_1 will be written in the square 1. It holds for $i = 1, 2, \ldots, j - 1$ that during the computation of M on w_1, the first of the squares between p^L_i and $m_2 - m_1 + p^R_i$ rewritten by M is

– the square p^L_{i+1} if $p^L_i \neq p^L_{i+1}$,

– the square $m_2 - m_1 + p^R_{i+1}$ otherwise.