Yugoslav Journal of Operations Research
6 (1996), Number 2, 245-256

RANDOM PHENOMENA AND SOME SYSTEMS GENERATING WORDS

Gabriel V. ORMAN

"Transylvania" University of Brasov, Romania

Abstract: The process of the generation of words by a generative system is considered from a stochastic point of view involving Markov chains. Because the sequences of intermediate words (called *derivations*) bu which the words are generated are finite, it results that finite Markov chain: will be connected to the process. In this paper a very general generative system from those constituing the *Chomsky hierarchy* is considered, frequently called a *phrase-structure grammar*. In Section 1 the basic definitions and notations relating to this type of generative system and some notions relating to Markov chains are given, according to [3] and [4]. Then, the random variable giving the number of derivations by which a eord can be generated is defined and its characteristic are determined according to [9]. Finally a new procedure to generate words is introduced and the property of invariance of the transition matrix is established: also a problem of the "reflecting barriers" type is discussed.

Keywords: Markov chain, random variable, transition matrix, alternating generation procedure.

1. INTRODUCTION

In order for our discussion to be as general as possible we consider generative systems free of any restrictions. The model of such systems is offered by the most general class of formal grammars from the so-called *Chomsky hierarchy, namely phrase-structure grammars*.

The novelty that we have proposed consists of organizing the process of word generation by considering the set of all the derivations according to such a system divided into equivalence classes, each of them containing sequences of intermediate words (that we shall call *derivations*) of the same lenght. In this way characterizations of the process up to an equivalence can be obtained.

For a good uderstanding of the facts considered below we shall first give the concept of a phrase-structure grammar.

A finite nonempty set is called an *alphabet* and will be denoted by Σ. A word over Σ is a finite sequence $u=u_1...u_k$ of elements Σ. The integer $k \geq 0$ is the length of the word u and is denoted by $|u|$. A word of length zero is called an *empty word* and is denoted by ε. If Σ is an alphabet let us denote by Σ^* the free semigroup, with identity, generated by Σ (Σ^* is considered in relation to the usual operation of concatenation).

A *phrase-structure grammar* (*psg*) is a system $G\ (V, \Sigma, P, \sigma)$ where:

(i) V is an alphabet called the *total alphabet*;
(ii) $\Sigma \subseteq V$ is an alphabet the elements of which are called *terminal symbols* (or *letters*);
(iii) P is a finite subset of the cartesian product $[\ V \setminus \Sigma)^* \setminus \{\varepsilon\}\ ] \times V^*$. Its elements are called *productions* (they are the rules of the grammar);
(iv) $\sigma \in (V \setminus \Sigma)$ is said to be the *initial symbol*. The elements of $V \setminus \Sigma$ are called *variables* (or nonterminals).

For y and z in V^*, it is said that y *directly generates* z, and one writes $y \Rightarrow z$. If the words t_1, t_2, u and v exists such that $y = t_1\ u\ t_2$, $z = t_1\ v\ t_2$ and $(u,v) \in P$ (alternatively written $u \rightarrow v$). Then, y is said to *generate* z, and one writes $y \Rightarrow^* z$, if either $y = z$ or a sequence $(w_0, w_1, \ldots, w_j)$ of words in V^*exists such that $y = w_0$, $z = w_j$ and $w_i \Rightarrow w_{i+1}$ for each i (we write $\Rightarrow^*$ for the reflexive-transitive closure of $\Rightarrow$). The sequence $(w_0, w_1, \ldots, w_j)$ is called a *derivation of length j* and will be denoted by $D(j)$. Obviosly, many derivations of the same length according to G may exist.

The subset of Σ^*, written $L(G)$, such that $L(G) = \{w \in \Sigma^* \mid \sigma \Rightarrow^* w\}$ is called the *phrase-structure language* (*psl*) generated by G. It is known that the family of *psl*'s coincides with the family of recursively enumerable sets studied in mathematical logic.

Because a derivation of length to 1 is just a production, from now on we shall suppose that the length of a derivation is $j \geq 2$.

Now the notion of Markov chain can be defined as follows. We can imagine that we have a sequence of trials in each of which one and only one of k mutually exclusive events $A_1^{(s)}, A_2^{(s)}, \ldots, A_k^{(s)}$ (where the superscript denotes the number of the trial) can occur. We say that the sequence of trials forms a *Markov chain*, or more precisely a *simple Markov chain*, if "the conditional probability that events $A_i^{(s)}$, $(i=1,2,...,k)$ will occur in the $(s+1)$st trial $(s=1,2,...)$ after a known event has occurred in the sth trial, depends solely on the event that occurred in the sth trial and is not modified by supplementary information about the events that occurred in earlier trials".

For Markov chains, the probability of passing to some state $A_i\ (i = 1, 2, \ldots, k)$ at time $\tau(t_s < \tau < t_{s+1})$ depends only on the state the system is in at time $t(t_{s-1} < t < t_s)$ and does not change if we learn what its states were at earlier times. *Homogeneous*

Markov chains are those in which "the conditional probability of the occurrence of an event $A_j^{(s+l)}$ in the $(s+l)$st trial, provided that in the sth trial the event $A_i^{(s)}$ occurred, does not depend on the number of the trial". This probability is called *the transition probability* and is denoted by p_{ij} ; in this notation the first subscript always denotes the result of the previous trial, and the second indicates the state into which the system passes in the subsequent instant of time.

The total probabilistic picture of possible changes that occur during a transition from one trial to the immediately following one is given by the matrix

$$\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1k} \\ p_{21} & p_{22} & \cdots & p_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ p_{k1} & p_{k2} & \cdots & p_{kk} \end{bmatrix}$$

compiled of the transition probabilities. This matrix is called *the transition matrix* (or *matrix of transition probabilities*). Its elements being probabilities, they must be nonnegative numbers, i.e. for all i and j

$$0 \leq p_{ij} \leq 1.$$

Also, from the fact that in the transition from $A_i^{(s)}$ prior to the $(s+l)$st trial the system must definitely pass to one and only one of the states $A_j^{(s+l)}$ after the $(s+l)$st trial, there follows the equation

$$\sum_{j=1}^{k} p_{ij} = 1, \quad (i = 1, 2, \ldots, k).$$

Thus, the sum of the elements of each row of the transition matrix is equal to unity. But the first problem in the theory of Markov chains consists of determining the transition probability from state $A_i^{(s)}$ in the sth trial to state $A_j^{(s+l)}$ after n trials. This probability is denoted by $p_{ij}(n)$ and is referred to as *the transition probability after* n *steps*. It is given by the formula

$$p_{ij}(n) = \sum_{h=1}^{k} p_{ih}(m)\, p_{hj}(n-m).$$

By means of this formula we shall obtain in Section 4 the two-step transition matrix in some special cases of word generation.

2. THE CONDITION FOR A WORD TO BE GENERATED INTO A CLASS OF DERIVATIONS

Let us denote by Ω the family of all the derivations accordding to G and let D_j be the class of all the derivations of length j. *Evidently*, Ω splits into equivalence clases, each of them being represented by one of its arbitrarily chosen elements. Let $D(j)$ be the representative of class D_j.

Now let μ_x be the number of derivations into the equivalence class in Ω of $D(x)$, $x \geq 2$, by which a word w is generated. Obviously, w can or cannot be generated into the equivalence class of $D(x)$. Thus, if w is generated into the class D_x then, the probability that it will be generated again into class D_{x+1} is denoted by γ; but should w not be generated into class D_x, the probability that it will be generated into the class D_{x+1} is denoted by β. Hence we are in the case when the equivalence classes of the derivations are connected into a simple Markov chain. We assume that both γ and β are different from 0 and 1 (these cases are of no particular interest). But each term of a derivation is the result of the application of an only production on the precedent term. Furthermore, w can or cannot be generated by a production (that is P can or cannot contain an element (σ, w)), such that the probability that w is or is not generated by a production is unknown to us. (In other words if we agree to denote the class of productions by D_1, it means that there is no equivalence class preceding class D_1). Thus, we denote by p_1 the probability that w will be generated into D_1 and by $q_1 = 1 - p_1$ the probability that w will not be generated into D_1 .

We now refer to some main problems discussed in [9]. First, we propose to determine the probability that a word w will be generated into class D_x. Let p_x be the probability that w will be generated into the class D_x and then we have $q_x = 1 - p_x$. Clearly w can be generated into D_x in two mutually exclusive ways: 1° w will be generated into D_{x-1} and will be generated again into D_x; 2° w was not generated into D_{x-1} but will be generated into D_x.

Theorem 1. *The probability that a word be generated into class D_x , $x \geq 2$, is given by the formula*

$$p_x = (p_1 - p)\delta^{x-1} + p \tag{1}$$

where $\delta = \gamma - \beta$ and $p = \dfrac{\beta}{1-\delta}$.

Proof. By the above conditions we have

$$p_x = p_{x-1}\gamma + q_{x-1}\beta.$$

or

$$p_x = p_{x-1}(\gamma - \beta) + \beta.$$

Denoting now $\gamma - \beta = \delta$ we obtain $p_x = p_{x-1}\delta + \beta$. But p_x can be developed as follows

$$p_x = p_1\delta^{x-1} + \beta(1 + \delta + \delta^2 + ... + \delta^{x-2})$$

so that we get

$$p_x = (p_1 - \frac{\beta}{1-\delta})\delta^{x-1} + \frac{\beta}{1-\delta} \tag{2}$$

From the previous conditions imposed on γ and β it results that $|\delta| < 1$ so that $p_x \to \frac{\beta}{1-\delta}$ as $x \to \infty$. On the other hand x being the length of a derivation it must be finite such that this situation corresponds to the case when the word cannot be generated into class D_x . But the constant to which p_x tends does not depend on probability p_l ; because it plays the role of a limiting probability it is natural to introduce the notation $p = \frac{\beta}{1-\delta}$ and $q = 1-p = \frac{1-\gamma}{1-\delta}$. Now the theorem is proved.

3. THE MAIN CHARACTERISTICS OF THE RANDOM VARIABLE

We now return to μ_x giving the number of derivations into the class D_x , $x \geq 2$, by which a word is generated. It is a random variable that takes the values 1 and 0 with probabilities p_x and $q_x = 1-p_x$ respectively. Then, the number of derivations in $n-1$ equivalence classes, by which a word is generated is

$$\mu = \sum_{x=2}^{n} \mu_x \tag{3}$$

We propose to determine the expectation and the variance of μ. To this end we shall first recall an intermediary result from [9]. Let us denote by $p_j^{(i)}$ the probability that a word w will be generated into class D_j if it was generated into class D_i , $i<j$.We have

Lemma 1. *The probability that a word w will be generated into class D_j if it was generated into class D_i , $i<j$, $i \geq 2$, $j>2$, is given by the formula*

$$p_j^{(i)} = p + q\,\delta^{j-i} \tag{4}$$

Now the main result is

Theorem 2. *If among the equivalence classes of the derivations according to a psg, a Markov dependence exists then the expectation and the variance of the random variable giving the number of derivations by which a word is generated verify the following relations*

$$E\mu = (n-1)p + u_n \text{ and } D_\mu = pq[n\frac{1+\delta}{1-\delta} - 1] + v_n$$

where u_n and v_n are certain quantities that remain bounded as n increases.

Proof. Starting from (3) with p_x given by relation (1) we get

$$E\mu = (n-1)p + \sum_{x=2}^{n}(p_1 - p)\delta^{x-1} - (n-1)p + (p_1 - p)\frac{\delta - \delta^n}{1-\delta} \tag{5}$$

Further, the variance of μ is

$$D\mu = E\Big[\sum_{x=2}^{n} (\mu_x - p_x) \Big]^2 = \sum_{x=2}^{n} E(\mu_x - p_x)^2 + 2 \sum_{i<j,\, i\ge 2} E(\mu_i - p_i)(\mu_j - p_j). \qquad (6)$$

But

$$E(\mu_x - p_x)^2 = D\mu_x = p_x q_x = \big[(p_1 - p)\delta^{x-1} + p\big]\big[1 - p_x\big] = pq + (p_1 - p)(q - p)\delta^{x-1} - (p_1 - p)^2 \delta^{2x-2}$$

and the second term in (6) becomes $E(\mu_i - p_i)(\mu_j - p_j) = E\mu_i\mu_j - p_i p_j$ where $\mu_i\mu_j$ is a random variable taking the values 1 and 0. The value 1 is taken with the probability $p_i p_j^{(i)}$ so that we have $E(\mu_i - p_i)(\mu_j - p_j) = pq^{j-i} - (p_1 - p)^2 \delta^{i+j-2} + (p_1 - p)(q - p)\delta^{j-1}$. Thus, the first term in (6) is obtained from the equality

$$\sum_{x=2}^{n} E(\mu_x - p_x)^2 = (n-1)pq + (p_1 - p)(q - p)\frac{\delta(1-\delta^{n-1})}{1-\delta} - (p_1 - p)^2 \frac{(1-\delta^{2n-2})}{1-\delta^2}$$

and the second from the equality

$$\sum_{i<j,\, i\ge 2} E(\mu_i - p_i)(\mu_j - p_j) = \sum_{i<j,\, i\ge 2} \big[pq\delta^{j-i} - (p_1 - p)^2 \delta^{i+j-2} + (p_1 - p)(q - p)\delta^{j-1} \big]$$

Now by observing the terms which are bounded as n increases, the theorem follows.

4. THE ALTERNATING GENERATION PROCEDURE AND THE ASSOCIATED TRANSITION MATRICES

Now we shall introduce a new procedure for generating words. To this end we consider the special case when a word can be generated into the equivalence class of a derivation on the following conditions: 1) It can be generated into the class D_x, $x \ge 2$, by more of its elements; 2) If it is not generated into the class D_x, $x \ge 2$, then it is generated into the preceding and the next class.

We refer to such a way of generating words as being *an alternating generation procedure* (cf. [5]). Four cases arise:

(i) A word will be generated by the derivations from the first class and the last (in brief: the word is generated by the first class and the last);
(ii) It will be generated by the first class but will be not generated by the last;
(iii) It will not be generated by the first class but will be generated by the last;
(iv) It will not be generated by either the first class or the last class.

With some supplemental conditions we have determined the probability that a word will be generated in n equivalence classes (see [9]).

Now we propose to determine the transition matrices corresponding to each of the above cases. We suppose that if word w is not generated by class D_x then it will be generated by class D_{x+l} with probability p and by class D_{x-l} with probability $q=1-p$. If w is not generated by the first class then, it will certainly be generated by the next, while if it is not generated by the last class then, it will certainly be generated by the preceding. Obviously, this procedure for generating a word is a typical Markov chain.

Now let us denote by A_l the event when the word is generated by class D_2, by A_2 when it is generated by class D_3 ,... by A_{n-1} when it is generated by class D_n .

1°. Let us first consider the case (i) Word w is not generated by class D_3 such that it will be generated by class D_2 and D_4 with probabilities q and p respectively; then it will not be generated by class D_5 but will be generated by classes D_4 and D_6 with probabilities q and p respectively, a.s.o. We have:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

There is an even number n of equivalence classes in the considered case. Because for our conditions of work we must have $n \geq 2$, it results that the above matrix is of the type $n-1$ (i.e. it has an odd number of rows and columns).

2°. In the second case word w is not generated by class D_3 , as in the first case, but it is also not generated by the last class such that it will be certainly generated by the last but one class. The transition matrix is now of the form:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \tag{8}$$

The number of the equivalence casses is an odd integer and the transition matrix has an even number of rows and columns equal to $n-1$.

3°. In the third case the novelty is that the word is not generated by the first class D_2 such that it will be certainly generated by D_3 . Also it is generated by the last class. The transition matrix is the following:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

Now the number n of the equivalence classes is again an odd integer such that the transition matrix has an even number of rows and columns equal to $n-1$.

4°. Finally, in the fourth case the word is not generated by either the first class D_2 or the last D_n such that it will be certainly generated by classes D_3 and D_{n-1}. The transition matrix is as follows:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \tag{10}$$

The number n of the equivalence classes is an even integer, as in the first case, such that the matrix has an odd number of rows and columns equal to $n-1$.

We now propose to determine the two-step transition matrix for each of the above cases. A surprising result that we shall call the *property of invariance of the transition matrix* will be obtained. It appears as a specific characteristic of the process of the generation of words by an alternating generation procedure. This is

Theorem 3. *The transition matrix for a word in a random process of generation by an alternating generation procedure is invariant to a two-step transition.*

Proof. 1°. In the first case we already know that $n-1$ is an odd integer. Then

$$p_{2i-1\,2i-1} = 1 \quad (i = 1,2,\dots,n/2), \quad p_{i\,2j} = 0 \quad (i = 1,2,\dots,(n-1),\ j = 1,2,\dots,(n-2)/2)$$

$$p_{2i\,2i-1} = q \quad (i = 1,2,\dots,(n/2)-1), \quad p_{2i\,2i+1} = p \quad (i = 1,2,\dots,(n/2)-1)$$

the other probabilities p_{ij} being all equal to 0 for $i \neq j$, $(i,j=1,2,\dots,(n-1))$. On the other hand, two consecutive rows are different in the following sense: if one of them contains the integer 1, the other contains p and q. The rows containing 1 are those of an odd rank, while the rows containing p and q are of an even rank.

We now compute the nonzero elements in two consecutive rows of the matrix (7). Let us consider the rows k and $k+1$. If k is an even integer then, by the above conditions we have $p_{k\ k+1} = q$ and $p_{k\ k+1} = p$, the other elements of this row being equal to 0. For the row of rank $k+1$ we get $p_{k+1\ k+1} = 1$, the other elements of this row being equal to 0; a similar result is obtained for the row of rank $k-1$.

Now we can compute the rows of rank k and $k+1$ of the two-step transition matrix. We find successively:

$$p_{k1} = \ldots = p_{k\,k-2} = 0, \quad p_{k\,k-1} = q, \quad p_{kk} = 0, \quad p_{k\,k+1} = p, \quad p_{k\,k+2} = \ldots = p_{k\,n-1} = 0.$$
$$p_{k+1\,1} = \ldots = p_{k+1\,k} = 0, \quad p_{k+1\,k+1} = 1, \quad p_{k+1\,k+2} = \ldots = p_{k+1\,n-1} = 0.$$

But this means that the obtained matrix is identical to matrix (7). For k being an odd integer the same result is found.

2°. In the second case, $n-1$ is an even integer and the transition probabilities are:

$$p_{2i-1\,2i-1} = 1 \quad (i = 1,2,\ldots,(n-1)/2), \quad p_{i\,2j} = 0 \quad (i = 1,2,\ldots,(n-1), \; j = 1,2,\ldots,(n-1)/2),$$
$$p_{2i\,2i-1} = q, \quad p_{2i\,2i+1} = p \quad (i = 1,2,\ldots,(n-3)/2)) \text{ and } p_{n-1\,n-2} = 1.$$

the other probabilities p_{ij} being all equal to 0 for $i \neq j$ $(i,j = 1,2,\ldots,(n-1))$.

As in the previous case the rows containing 1 are those of an odd rank, while the rows containing p and q are those of an even rank, except the last row which also contains a 1 ($p_{n-1\,n-2} = 1$). Computing the nonzero elements in two consecutive rows k and $k+1$, we get (for an even integer k).

$$p_{k\,k-1} = q \text{ and } p_{k\,k+1} = p \;\; (k = 2,4,\ldots,(n-3)), \text{ the other elements of this row being } 0,$$
$$p_{k+1\,k+1} = 1 \;\; (k = 0,2,4,\ldots,(n-3), \text{ the other elements of this row being } 0.$$

Now we obtain the rows of rank k and $k+1$ of the two-step transition matrix:

$$p_{k1} = \ldots = p_{k\,k-2} = 0, \quad p_{k\,k-1} = q, \quad p_{kk} = 0, \quad p_{k\,k+1} = p, \quad p_{k\,k+2} = \ldots = p_{k\,n-1} = 0,$$
$$p_{k+1\,1} = \ldots = p_{k+1\,k} = 0, \quad p_{k+1\,k+1} = 1, \quad p_{k+1\,k+2} = \ldots = p_{k+1\,n-1} = 0.$$

but the last row is: $p_{n-1\,1} = \ldots = p_{n-1\,n-3} = 0, \quad p_{n-1\,n-2} = 1, \quad p_{n-1\,n-1} = 0$, which is just (8).

3°. Now n-1 is an even integer as in the preceding case. We obtain:

$$p_{12} = 1, \, p_{2i\,2i} = 1, \, (i = 1,2,\ldots,(n-1)/2), \, p_{i\,2j-1} = 0 \; (i = 1,2,\ldots,(n-1), \, j = 1,2,\ldots,(n-1)/2),$$
$$p_{2i+1\,2i} = q, \, p_{2i+1\,2i+2} = p \quad (i = 1,2,\ldots,(n-3)/2).$$

the other probabilities p_{ij} being all equal to 0 for $i \neq j \quad (i,j = 1,2,\ldots,(n-1))$.

The rows containing a 1 are those of an even rank, while the rows containing p and q are those of an odd rank, except the first row which contains a $1(p_{12} = 1)$.

The rows of rank k and $k+1$ (k being an even integer) of the two-step transition matrix are the following:

$$p_{k1} = \ldots = p_{k\,k-1} = 0,\ \ p_{kk} = 1,\ \ p_{k\,k+1} = \ldots = p_{k\,n-1} = 0.$$
$$p_{k+1\,1} = \ldots = p_{k+1\,k-1} = 0,\ \ p_{k+1\,k} = q,\ \ p_{k+1\,k+1} = 0,\ p_{k+1\,k+2} = p,\ p_{k+1\,k+3} = \ldots = p_{k+1\,n-1} = 0.$$

but the first row is: $p_{11} = 0,\ p_{12} = 1,\ p_{13} = \ldots = p_{1\,n-1} = 0$, which is just matrix (9).

4°. In this case n-1 is an odd integer. The transition probabilities are as follows

$$p_{12} = 0,\ p_{n-1\,n-2} = 1,\ p_{2i\,2i} = 1\ \ (i = 1,2,\ldots,(n-2)/2),$$
$$p_{i\,2j-1} = 0\ \ (i = 1,2,\ldots,(n-1),\ j = 1,2,\ldots,n/2),\ p_{2i+1\,2i} = q,\ p_{2i+1\,2i+2} = p\ \ (i = 1,2,\ldots,(n-4)/2)$$

the other probabilities p_{ij} being all equal to 0 for $i \neq j\ \ (i, j = 1, 2, \ldots, (n-1))$.

If k is an even integer, the rows of rank k and $k+1$ are as follows:

$p_{kk} = 1\ \ (k = 2, 4, \ldots, n-2))$, the other elements of this row being equal to 0. $p_{k+1\,k} = q$ and $p_{k+1\,k+2} = p\ \ (k = 2, 4, \ldots, (n-4))$, the other elements of this row being 0. But this is just matrix (10), and the theorem is proved.

As we have already emphasized, this is a property of invariance which characterizes the process of the generation of words up to an equivalence. Furthermore, because the generative systems that we have considered are free of any restrictions, this property has a sufficient by large character.

5. PROBLEM OF REFLECTING BARRIERS

In this section we shall discuss the fourth case in the alternating generation procedure in which a word is not generated by either the first class or the last. But now we consider only the equivalence classes of derivations in which a word is not generated. In this case the number of the equivalence classes of derivations is an even integer, say $n = 2k$. Because $n \geq 2$, the number of classes by which a word is not generated is equal to $n / 2$.

If word w is not generated by class D_x then, it will be generated by class D_{x+1} with probability p and by class D_{x-1} with probability $q = 1-p$. If w is not generated by the first class then it will certainly be generated by the next, and if it is not generated by the last class then, it will certainly be generated by the preceding. Evidently, this procedure for generating a word is a typical Markov chain.

In our case of study w is not generated by the classes D_{2k} $(k = 1,2,\ldots,n / 2)$, and we are interested only in these equivalence classes. Because it is not generated by either the first class D_2 or the last D_n, it will certainly be generated by D_3 and D_{n-1}. Let us denote by A_1 the event when the word is not generated by class D_2, by A_2 when it is not generated by class D_4 ..., by A_s, $s = n / 2$, when it is not generated by class D_n. The transition matrix corresponding to this case is the following:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ q & 0 & p & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 0 & q & 0 & p & 0 & \cdots & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & q & 0 & p \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 \end{bmatrix}$$

We now propose to determine the two-step transition matrix. Using the formula for the transition probabilities we obtain the following form:

$$\begin{bmatrix} q & 0 & p & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ 0 & q+pq & 0 & p^2 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ q^2 & 0 & 2pq & 0 & p^2 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & q^2 & 0 & qp+p & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & q & 0 & p \end{bmatrix}$$

This problem is of the type of *reflecting barriers* known in the theory of Markov chains. For example, let us consider that a particle located on a straight line moves along the line via random impacts occuring at times $t_1, t_2, t_3, \ldots$. The particle can be at points with integral coordinates a, $a+1$, ... , b. At points a and b there are reflecting barriers. Each impact displaces the particle to the right with probability p and to the left with probability $q = 1 - p$ so long as the particle is not located at a barrier. If the particle is at a barrier, any impact will transfer it one unit inside the gap between the barriers. Thus, our case of the generation of words by an alternating generation procedure becomes of a special interest. Its practical nature must also be emphasized and new interesting results may be obtained in the future.

REFERENCES

[1] Gainsburg, S., *Algebraic and Automata-Theoretic Properties of Formal Languages*, North-Holland Publ. Co., Amsterdam, 1975.

[2] Gnedenko, V.G., *The Theory of Probability*, Mir Publisher, Moscow, 1976.

[3] Loeve, M., *Probability Theory*, Springer-Verlag, New York, 1977.

[4] Orman, G.V., *Limbaje formale* (Formal Languages), Ed. Technica, Bucharest, 1982.

[5] Orman, G.V., "Alternating generation of words by a phrase-structure grammar; conceptual aspects", Bull. Univ. Brasov, Series C, 31 (1989).

[6] Orman, G.V., "Some general aspects of the markov dependence in the process of the generation of words", Twentieth Conference on Stochastic Processes and their Applications, Naharuya, June 9-14, 1991.

[7] Orman, G.V., "The probability of alternating generation of the words", Seminar on Probability Theory, Stochastic Processes and Applications, Transylvania Univ. of Brasov, Report No. 1, 1992.

[8] Orman, G.V., "On some numerical functions utilized in the study of a class of discrete systems", *Systems Science*, 19, (1) (1993).

[9] Orman, G.V., "Markov Chains Connected to the Process of the Generation of Words by a Generative System", in: *Proceedings of the 2th Balkan Conference on Operational Research*, Thessaloniki, 18-21 Oct. 1993.

[10] Orman, G.V., *On the Characterization of Some Generative Grammars by Numerical Functions*, Current Issues in Mathematical Linguistics, Elsevier Science B.V., North Holland, Amsterdam, 1994.

[11] Salomaa, A., and Soittola, M., *Automata-Theoretic Aspects of Formal Power Series*, Springer-Verlag, New York, 1978.