THE EVOLUTIONARY PROCESS OF RANDOMLY GROWING MUTATED DIGITAL STRUCTURES AS A MODEL OF EVOLUTION OP FIRST LIVING OHGANISMS :)

Konrad *R.* Fialkowski

Technical University of Warsaw,Poland

A b s t r a c t

The paper presents the results of research concerning the stability of an evolutionary process for randomly growing digital structures. The computer model for the mutated genotype case is considered. The genotype (in the models the set of rules for growth) changes in a random environmentally unoriented way, whereas the environment assumed as random, finite and repetitive in a cyclic way is disturbed. Prom the experiments it follows that stability could be achieved in a heuristic way after mutation of the genotype and disturbances in the environment without the built- in information concerning stability.

The model, results obtained from the computer and an interpretation of these results are presented.

Introduction

For the case without spontaneous mu tations it was shown in a previous paper (1) that with the assumption of random environment changing in a cyclic way, the digital structures,which could be a model of very simple organisms can grow, divide and achieve a stabile "shape" (phenotype) in spite of the fact that the built-in informstion (genotype) does not determine the phenotype in detail, but only gives "frame" information about the phenotype. Full Information to guarantee the stability is achievable from the cyclic environment.

In real life, however, it is not sufficient to achieve stability. Stability must be maintained in the presence of the mutations. Under these conditions the organlsms must survive and divide. The chain of survivals and divisions gives the evolutionary sequence (in each generation only one organism is taken into consideration). This sequence could in each generation be ended by the unsuccessful growth and death of the current element of sequence. That case is called the "degeneration*[1] of the sequence. Sometimes however the evolutionary sequence and the environment are in dynamic balance, and then degeneration does not happen. The balance (if it exists) could be broken by:

1. disturbance of the environment (the case discussed in (1))
2. mutation, which changes the genotype (the case discussed in this paper).

The second case is important, as the changable genotype means development, increase of complexity, and evolutionary progress.

M o d e l

1. Phenotype growth and division.

The phenotype is called in the model a Digital Structures DI6TRUC.(1).

Definition 1.

A Digital Structure,DISTRUC, is a sequence of positive, integer numbers

$$(c_i) \quad , \quad i = 1,2,\ldots,I$$

x) These computer experiments were supported by a grant from the ford Found.

The numer I is called the length of DISTRUC.

Definition 2.

The sequence of DISTRUCs created in the recursive growth and division process is called an Evolutionary Sequence: SB.

Each DISTRUC in £8 has its generation number g, used as an index. The Initial Basic Structure, IBSTRUC has g=o. IBSTRUCs are always the same for all ES (as the result of an assumption concerning the common origin of life).

Definition 3.

The growth process is the process of non- decreasing change of c_i values performed in the discrete moments of time t, t = 0,1,...,T (measured in discrete time units: DUT), according to the following equation:

$$c_{i,t+1,g} = c_{i,t,g} + F\,R \pmod 2 \qquad 1$$

for all g, g = 0,1,2,...
and i =1,2,...,I ,I+1

where: R is a random number, and F has value 1 or 0 according to the set of rules of growth shown further.

For all g, the growth process starts for t=o. For t = 0 DISTRUC is called Basic Structure: BASTRUC. IBSTRUC is BASTRUC for g = 0. IBSTRUC is denoted as:

$$c_{1,0,0} = 1$$
$$c_{2,0,0} = 2$$
$$c_{3,0,0} = 4$$

Generally BASTRUC is denoted as:

$$(c_i)_{0,g}$$

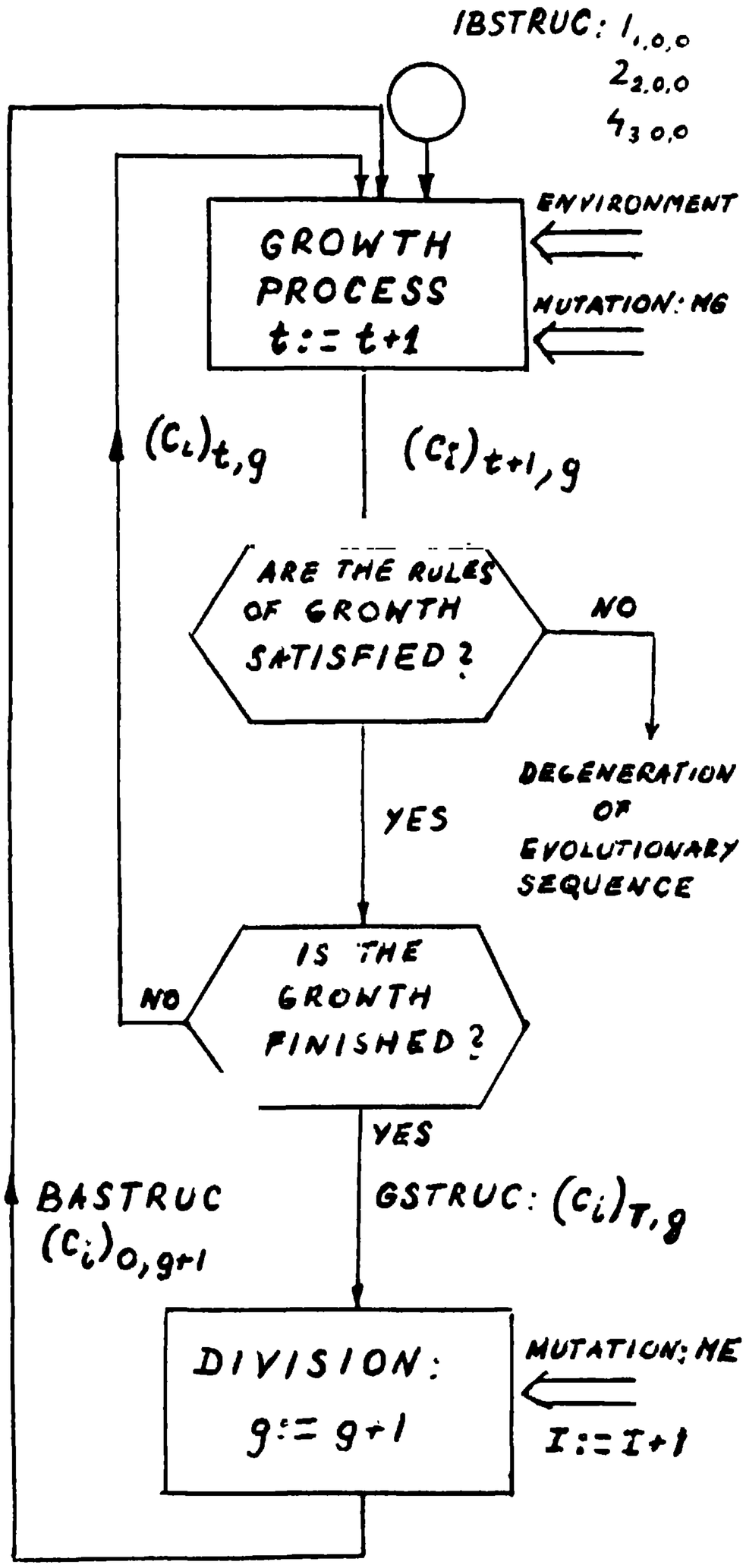

FIG. 1 : THE DISTRUC GROWTH AND DIVISION DIAGRAM

The growth could be finished in a one of two ways:

1. If the growing DISTRUC satisfies the rules of growth, the growth is successfully finished after T DUTs. Then DISTRUC is called the Grown structure: GSTRUC. T is called the life time of DISTRUC.

2. If the growing DISTRUC does not satisfy the rules of growth, the DISTRUC "dies" after T_d DUT. This is called "degeneration" and T_d - the degeneration time.

In case 1. GSTRUC divides into two BASTRUCs of the next generation (but only one of these two is taken into the further growth process and its history recorded)• The division is performed in the following way:

$$c_{i,0,g+1} = E\left(\frac{c_{i,T,g}}{2}\right) \qquad 2$$

for i = 1,2,3,..., I , where E is the Entier function (for positive numbers, the integral part of the number).

The above given definitions are summed up on the fig.1.

2. Genotype and the rules of growth

The genotype contains:

the Basic Genotype : BG
the Stability Treshold : ST
the Length of DISTRUC : I
the Interval of Regular Growth: D

Definition 4.

Basic Genotype: BG is a sequence of positive, real (in the ALGOL sense) numbers:

(G_i) , i = 1,2,...,I,I+1

Each BASTRUC of g's generation has BG determined as:

$(G_i)_{0,g}$

For IBSTRUC $(G_i)_{0,0}$ is assumed to be:

$G_{1,0,0} = 2.4$

$G_{2,0,0} = 4.4$

$G_{3,0,0} = 2.0$

for all experiments.

Definition 5.

Stability Treshold: ST is a sequence of positive, real (in the ALGOL sense) numbers:

(A_i) i = 1,2,...,1,1+1

For all experiments presented in this paper (A_i) has constant values:

$A_i = 0.8$ for i = 1,2,...,9

$A_{10} = 3.0$

The Length of Structure : I was defined before. For IBSTRUC, 1 = 3.

The value of D is constant for all experiments presented in this paper, D = 0.5

There are two indicators of the growth process: V and F (see eq. 1)• Both of them have value 0 or 1.

The values of indicators are determined for each t in the following way:

For t=0 $Q = 0$ and V=1.

For t>0 $Q = \sum_{k=1}^{I} \prod_{j=1}^{k} q_j$

where :

if $c_{1,t-1,g} \geq E(G_{1,t,g})$ 3

$q_1 = 1$, otherwise $q_1 = 0$

if $c_{2,t-1,g} \geq E(G_{2,t,g})$ 4

$q_2 = 1$, otherwise $q_2 = 0$

if $\frac{c_{j,t-1,g}}{c_{j-1,t-1,g}} > G_{j,t,g}$ 5

$q_j = 1$ otherwise $q_j = 0$

for j = 3,4,...,I, E is Entier func.

and $V = \prod_{j=0}^{Q} v_j$

where: $v_0 = 1$ and $v_1 = 1$

if $c_{2,t-1,g} \leqslant E(G_{2,t,g}) + 3$, 6

$v_2 = 1$ otherwise $v_2 = 0$

if $\frac{c_{3,t-1,g}}{c_{2,t-1,g}} \leqslant G_{3,t,g} + 1.5$ 7

$v_3 = 1$ otherwise $v_3 = 0$

if $\frac{c_{j,t-1,g}}{c_{j-1,t-1,g}} \leqslant G_{j,t,g}$ 8

$v_j = 1$ otherwise $v_j = 0$

for j=4,5,...,I, E is Entier func.

For each t, as long as V=1, the growth process is continued according to eq. 1

Then F is determined as follows:

$F = 0$ for $i - Q \leqslant 0$

$F = 1$ for $i - Q > 0$

If for certain t value V becomes equal to zero (V=0) the growth process is stopped, and it is the "degeneration" case.

3. Mutations

In the presented model there are two kinds of mutations:

1. BG's mutations: MG,which change the value of (G_i)

2. Expanding mutations: MB, which change the value I.

Both of them according to the general theory of evolution are neither environmentaly, or DISTRUCtly oriented.

MG occurs once for the determined period of time in given experiment (e.g. once for 15 000 DUTs). It changes one randomly choosen value of (G_i) during the growth process, and the process is continued according to the rules following from the new value of BG. This change is performed by adding to one randomly choosen G_i value the random number from the random generator. These random numbers are equally distributed from the range: -1.0, 1.0.

Sometimes the randomly choosen G_i(e.g. G_8 for 1=5) does not exist in the given genotype. Then the mutation does not occur.

MS can occur during each division (see fig.1). It occurs if, and only if:

$$A_{I+1} < P < G_{I+1,T,g} + D \qquad 9$$

where:

$$P = \frac{c_{I+1,T,g}}{c_{I,T,g}} + \frac{D}{2} R \pmod 2 \qquad 10$$

The value P is called the Expanding Coefficient: EXC.

If ME occurs, the DISTRUC changes its length and then:

$$I_{g+1} = I_g + 1 \qquad 11$$

$$G_{I,0,g+1} = G_{I,T,g} \qquad 12$$

The value of the newly created Ith place is determined as follows:

$$c_{I,0,g+1} = E\left(\frac{c_{I+1,T,g}}{2}\right) \qquad 13$$

3. Environment

The basic assumptions concerning the environment were given in (1). In all experiments the environment has been simulated as a sequence of 500 random numbers taken from the random number table used in the repetitive, cyclic way. The disturbances of the environment were simulated as skips of several (usualy one) numbers in the sequence.

In the model, for the given experiment, the disturbance occurs once for the determined period of time, on the

same basis as MGrs.

Exeriments and Results

The experiments presented here were the continuation of the research, which has been started from so called "version zero" presented in (1).

In contradistinction to the "version zero" all further versions were written in FORTRAN IV for IBM 360/50.

1. "first version" was the modification of "version zero" for IBM 360 with the extension of possible length of DISTRUC to I= 23.

2. "second version" was the introduction to "first version" of the timing and the automatic perturbation of the environment•

3. "third version" was the introduction to the "second version" of the mutations subroutines.

4. "fourth version" was the introduction to the "third version" of limitations for the expansion of structures (the way of saving CPU time) and automatic modification of the stability threshold (not used in experiments presented here).

For experiments presented here only the "fourth version" was used. The "fourth version" had two program realizations: Fourth A and Fourth B. Usualy Forth A was used as an experimental group, whereas Fourth B was a control group (in the same way as in the biological experiments)• The estimated number of observed generations is 50 000 . The average life time for one generation is app. 600 DUTs. The single experiment lasts 0.5×10^6 or 10^6 DUTs.

By analogy, if one assumes that 1DUT is equivalent of 1 second, then the life time of average DISTRUC is 10 minutes, which is a satisfactory approximation for very simple organisms. In this scale the single experiment lasts app. 6 days and all experiments are equivalent to over one year of constant observation.

The results are presented on the three hierarchical levels:

1. Leve one (the lowest)

The level one (presented on fig.2) is an example of direct graphic record of an experiment. The evolutionary sequence presented on fig.2 develops in the stabile environment (without the disturbances) and under the influence of both types of mutations (MG and MS).

The generation number g is on the abscissa (g=0 indicates IBSTRUC). The arrows under the abscissa indicate MG mutations. For the experiment the distance between two successive MG mutations is 10^4 DUT. The phenotype is presented as a dotted set of values for given g value. (Always BASTRUC $c_{i,o}(g)$ is presented on the diagrams. For these experiments max 1 = 9). The smooth line is T(g), (the life time). The phenotype values and the life time values are presented on the same scale. The P(g) (expanding coefficient, the upper part of the diagram) is presented in another scale and the stability tree hold (dashed, constant value line) is drawn there.The presented diagram is a part of the experiment named: 1106 Fourth B (for the detailed data of this experiment see level two).

Starting from IBSTRUC, after the single expansion (MSs effect) a nonperi-

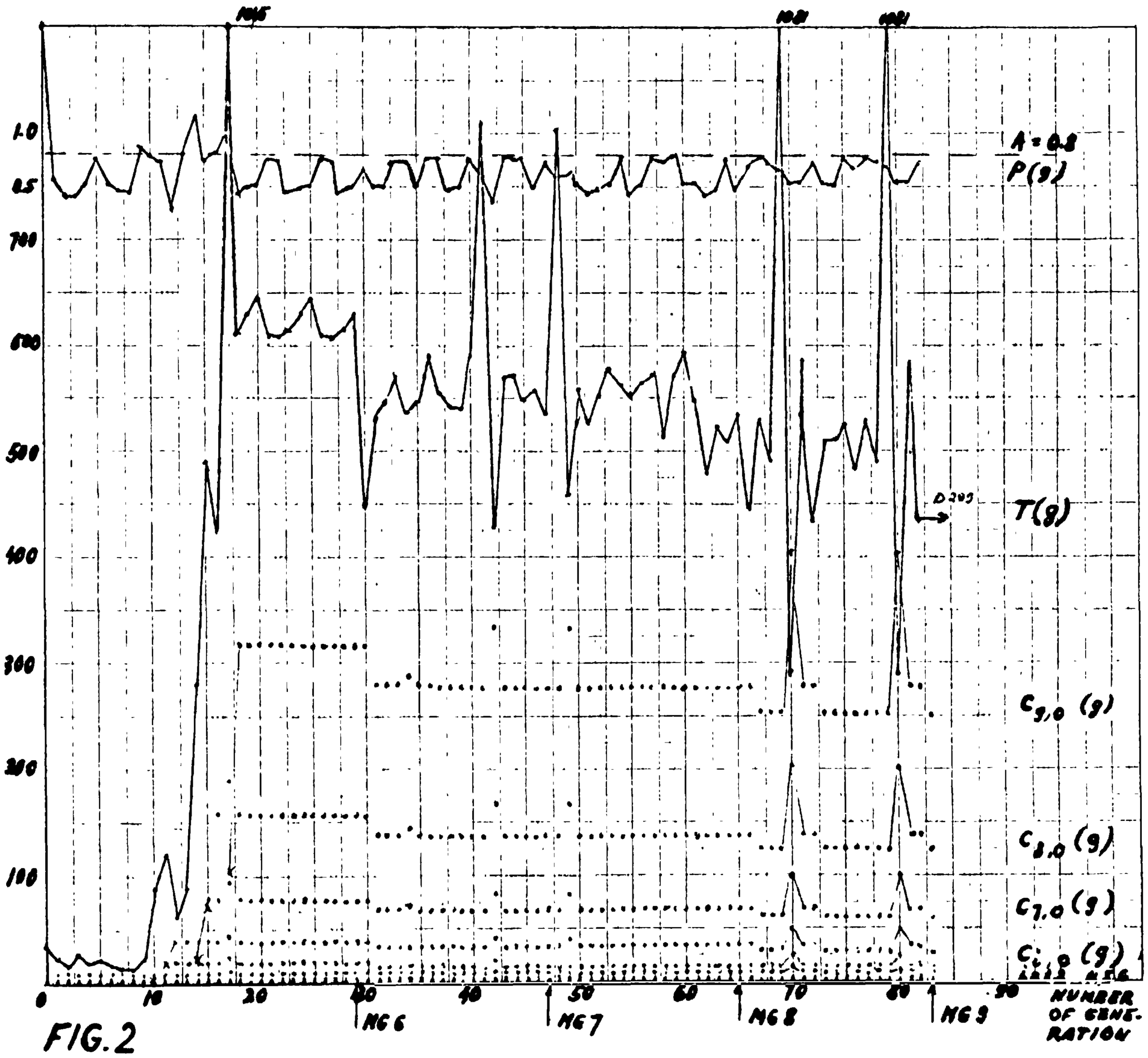

FIG. 2

odic sequence exists till the 8th generation. In the 9th and 10th there are two MEs (the effects are visible in the very next generation and marked "x" on the diagram). The I value becomes equal to 6. The next two MEs occur in 13th and 14th generations and one in 16th • After that I is equal to 9. From the 18th generation the periodic sequence, described in (1) starts, with a period equal to 10 (see periodic changes of T(18) to T(29)). All the phenotypes are the same, but they have different life times. The MG mutation occurs at the end of life of the 29th generation. According to the record it changes the G_5 value of BG from 2,00 to 1.76. It does not influence BASTRUC of the 30th generation (which has been already determined by GSTRUC nearly completely formed at the time of MG), but influences the growth of DISTRUC in 30th generation. As only BASTRUCs are recorded the effect of MG are visible in

the BASTRUC of 31st generation. The new, most probable phenotype is statisticaly smaller, than the previous one, but is not exactly determined and one can see other values of the phenotype (other "shapes") in 33rd, 43rd and 49th generations (the last two are larger than the phenotypes before MG6). The sequence is nonperiodic.

MG7 does not change anything. It changes G_9 value from 2.00 to 1.99 and that has no visible influence on the growth process.

MG8 changes (accidentally the same as MG6) the G_5 value from 1.76 to 1.61.

This causes a further statistical decrease of phenotype values, with some exceptions in the 70th, 71st and 72nd generation. From the 69th generation a periodic sequence starts with the period equal to 10. This dynamic balance is broken in the 84th generation by MG9. According to the experimental records, it changes the G_7 value from 2.00 to 1.21. The DISTRUC of 64th generation does not finish its growth. The MG9 degenerates the sequence.

From the record of the experiment it follows that the DISTRUC of the 84th generation "dies" with the "shape"

(c_i)= (2, 4, 9, 19, 31, 63,127,255,407)

with c_{10}= 171 after 299 DUTs.

2. Level two.

Fig.3 presents two histograms, which are the records of the two different experiments:

A) upper part (1006 Fourth B, lasted $0.5x10^6$ DUT/42 min. 37 sec. of IBM 360/50) CPU time).

B) lower part (1106 Fourth B, lasted $0.5x10^6$ DUTs/37min. 46 sec of IBM 360/50 CPU time)• The part of this experiment was presented as fig.2 for level one.

On fig.3 the numbers of MG's mutations are on the abscissa. The number of generations observed between two successive mutations are on the ordinate. The expanding sequence (the effect of MEs mutations), the periodic and the nonperiodic sequences are distinguished. In the experiments, immediately after the degeneration of one sequence a new sequence starts from the IBSTRUC what is marked on the diagrams by "_" . The time between the two succes MG's mutations is 10^4 DUTs (the abscissa is scaled with MGs or with the 10^4 DUTs).

The sequence of MGs mutations in the both cases (part A and B) is exactly the same, as well as the other conditions. But in case A the mutation occurs together with the disturbance of the environment. In case B the mutation occurs alone.

The fragment of B record between the MG5 and MG9 was presented in another form on fig.2. One can notice the expansions, the periodic and nonperiodic sequences described before on level one.

From (1) it follows that a nonperiodic sequence, after sufficiently long time, can eather achieve periodicity or degenerate. Periodic sequences are not interesting for the experiments, as nothing could happen there till the nearest mutation or disturbance of environment. On the other side, under severe conditions a nonperiodic sequence lasts for a very short time. The conditions for the presented experiments were not too severe (see the long nonperiodic sequence at part A, fig. 3),

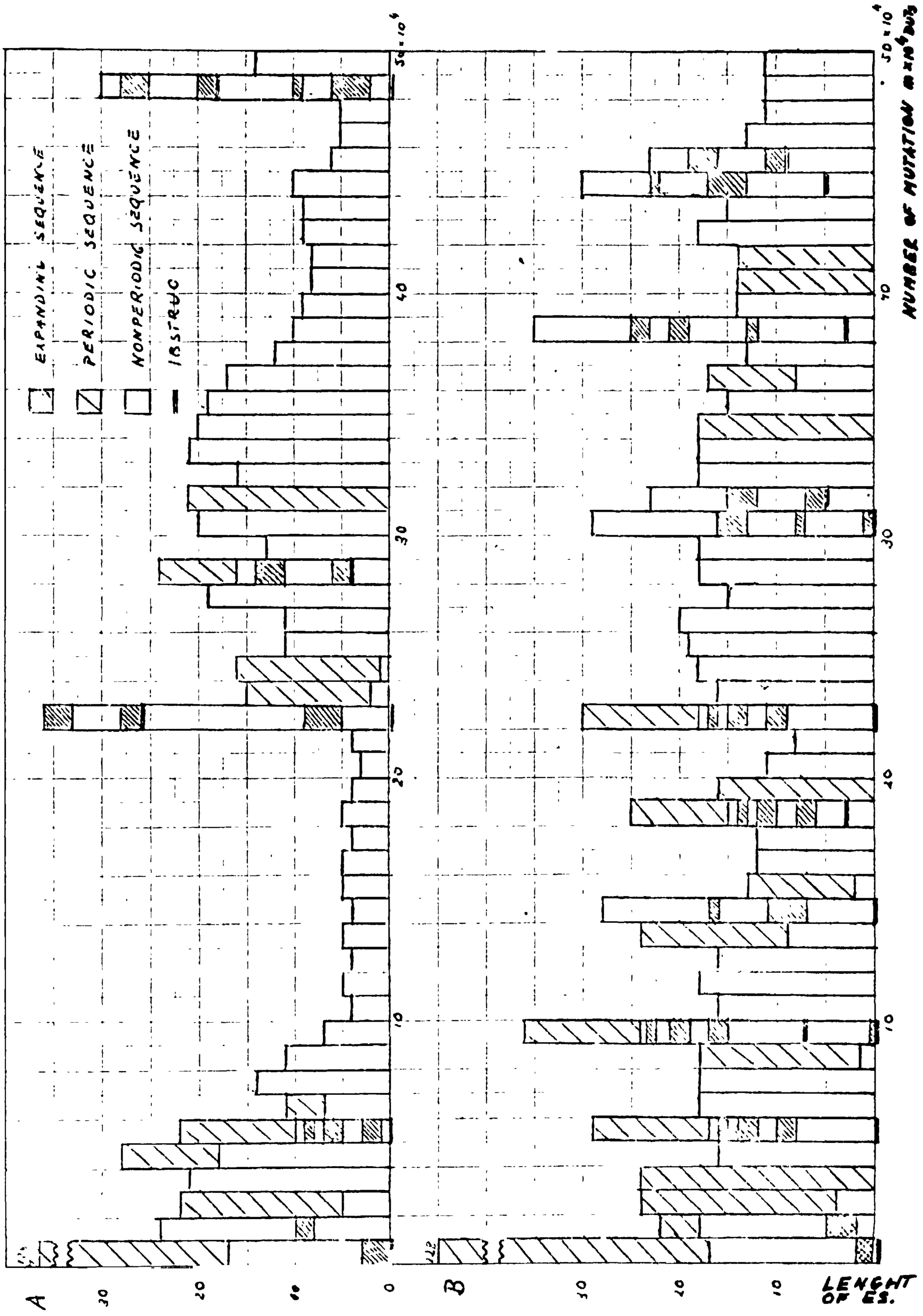

Fig.3. Lenght of evolutionary sequence (measured in number of generations) between the two successive MG mutations.

and the frequence of MGs repetition was chosen high enough for a periodic sequence not to be too probable. One can notice that the degenerations are more frequent in case B. To determine the probability of degeneration, the degeneration coefficient d will be introduced.

For the given experiment and the given period of time the d value is counted by dividing the number of non-spontaneous degenerations by the number of MGs mutations. For the A and B case the d values are;

$d_A = 5/50 = 0.1$, $d_B = 10/50 = 0.2$

(in each case one spontaneous degeneration was observed, for A between MC48 and MG49, and for B between MG9 and MG10).

The experiments presented on fig.3 are rather typical. There are other records e.g. 306 Fourth A with the parameters of the process the same as for 1006 Fourth B, where d= 0.074.

From the computer experiments presented here and considered it follows that the probability of degeneration for evolutionary sequences increases when the mutations occur without the disturbance of the environment.

3. Level three.

On the fig.4 the degeneration cases for the five different experiments are presented. The experiments named t X, Y, Z,P and Q have the following characteristic:

x: 2805 Fourth A , lasted 10^6 DUTs
(1 hour, 11 min. 45s. of CPU time)

Y: 2905 Fourth B', lasted 10^6 DUTs
(1 hour, 14 min. 56s.of CPU time)

Z: 2705 Fourth B, lasted 10^6 DUTs
(1 hour, 12 min. 58s.of GPU time)

P: 2905 Fourth B, lasted $0.5x10^6$ DUTs
(40min. 27s.of CPU time)

Q: 5005 Fourth A, lasted $0.5x10^6$ DUTs
(58 min. 25s.of CPU time)

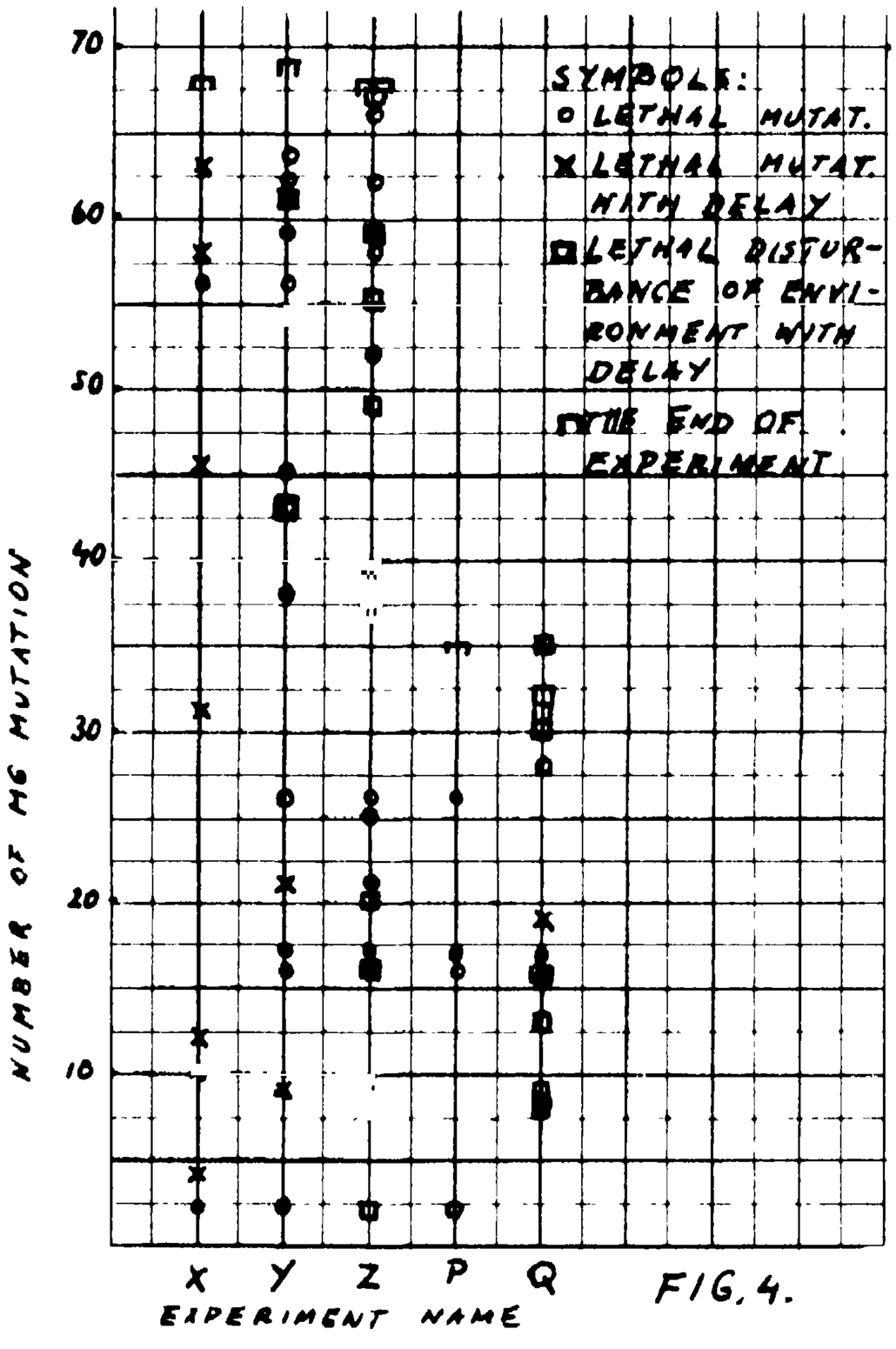

FIG. 4.

Fig.4. Distribution of lethal cases in five experiments.

For all these experiments the time period between the two successive MG's mutations, or two successive disturbances of the environment was $1.5x10^4$ DUTs.

From the detailed records of the experiments it follows that:

1. If the mutation is lethal for sequence it is usualy lethal in the very

next generation (see marks without delay on fig.4)

2. If the disturbance of the environment is lethal for the sequence it is usualy lethal after several generations (see fig.4).

From that it follows that the phase shift between the mutation and the disturbance of the environment could be an important factor, which influences the probability of degeneration.

The phase shift factor is counted as:

$$k = \frac{t(MG(j)) - t(EP(j))}{1.5 \times 10^4} 360^o \qquad 14$$

where: k is the phase shift factor
t(MG(j)),t(EP(j)) are the times when MG's mutation of number j, or disturbance of the environment of number j occur.

The d coefficient could be counted for all degeneration cases, or for the degenerations caused by MGs only. The second will be denoted as d_{MG}. . From fig.4 the k factor, d coefficient and d_{MG}, coefficient could be determine for the all five presented experiments. These values are presented in the table below:

Experiment	k	d_{MG}	d
X	0	?	0.132
y	72	0.216	0.246
Z	180	0.189	0.343
P	120	0.114	0.114
Q	-120 (or 240)	0.143	0.372

For X the d_{MG} value is not determined, as the mutation and the disturbance of the environment occur in the same time and the cause for degeneration is undetectable.

From the above table it follows that the probability of degeneration increases with the increase of the phase shift. The probability of the degeneration is highest for the disturbance of the environment being directly preceded by the mutation (Q experiment).

The degenerations caused by mutation comparable for each experiment. The probability of degeneration is lowest when MG and a disturbance of the environment occur in the same time (the experiment is considered as nontypical and shown only as an example of the existence of such records).

interpretation.

The presented computer model is a model of the evolution process. The DISTRUCs being the objects of simulated evolution are abstract and are not the image of structure of any living organisms. But their behavior simulates the behavior of organisms, i.e., they can grow, divide and mutate. The process of evolution, the real one as well as the simulated, has as a goal: the achievement of dynamic stability for an evolutionary sequence. In the model stability is achieved when the periodic sequence starts. *A* nonperiodic sequence is the phase of searching for stability. For the given genotype and the environment there are many possible stabile states (i.e. periodic sequences). The examples of different stabile sequences and the different ways of finding them one can see on fig.*3.*

Not all of possible stabile sequences are looked up. Only one (if any) is found in the heuristic search.

From the experiments follows that:

1. The stabile (the same in the different generations) phenotype could be achieved after mutation of the genotype and disturbances in the environment without the built-in information concerning stability. In conclusion, the relatively simple organisms, which could only grow and divide and had no built-in information about the final "shape" could be the first living organisms, able to develop and to complicate the structures.

2. The disturbances of the environment were not neccessarily the destructive factors causing the degeneration of the evolutionary sequences.They could be helpful to achieve stability if they have occured in the proper "phase shift" with the mutation.

The results presented in this paper were obtained from an abstract model of the evolution process simulated on the computer. They may be treated as indications of some problems, but for any extrapolations towards biology they should be tested in biological experiments•

References

1. K.R. Fialkowski, Cyclic Behavior of Randomly Growing Digital Structures in Finite Random Environment (Proc. of Int. Joint Conf. on Artificial Intelligence, Washington D.C. 1969, pp.347-359)
2. K.R. Fialkowski, Final Report for the Ford Foundation
3. L.J. Fogel, A.J. Owens, M.J. Walsh, Artificial Intelligence through Simulated Evolution (J.Wiley, 1966)