K-MODEL – AN EVOLUTIONARY ALGORITHM WITH NEW SCHEMA OF REPRESENTATION

Halina Kwasnicka
Department of Computer Science, Wroclaw University of Technology
Wyspianskiego 27, 50-370 Wroclaw, Poland
tel. (48 71) 3202397, fax: (48 71)211018,
e-mail: kwasnicka@ci.pwr.wroc.pl, http://www.ci.pwr.wroc.pl/~kwasnick

ABSTRACT

The paper presents an evolutionary model with pleiotropy and polygene effects, named *K*-Model. Genes are integer numbers and the linear dependence between phenes and genes are assumed. The redundant genes, i.e., not active ones during some period of evolution, are also included. Such representation allows the evolutionary algorithm to escape from local optima (evolutionary traps) and evolve quicker especially for multi-modal and multi-variable fitness functions.

INTRODUCTION

From centuries people learn observing natural processes. Usually natural processes are perceived as very skilful, and researchers willingly imitate them to obtain effective techniques. Optimisation techniques called ***Genetic Algorithms*** (**GAs**), or more generally, ***Evolutionary Algorithms*** (**EAs**), base on biological evolution. They imitate the Darwinian paradigm *survival the fittest*, and use a vocabulary borrowed from genetics. **GAs** search large spaces efficiently without complete knowledge of an objective function. They require only a limited number of values of objective function to guide their search process. Operators analogical to genetic ones as crossover and mutation are used for achieving a new area in the search space (new solutions).

GAs act on the set of coded potential solutions (called chromosomes or individuals) and use the strategy of natural selection to achieve their aims. Proposed by J. Holland genetic algorithms, which we call here as *Classic Genetic Algorithm* (**CGA**) uses a binary alphabet for defining chromosomes. A population of individuals evolves in defined environment, which represents a problem to be solved. Individuals live one period (one generation), compete between themselves, and the better ones (fitter for the environment, what means, the better solutions) are reproduced. Offsprings differ from the parents' individuals because the genetic operators act during reproduction.

Before running a computer programme with **GAs** for solving some task we should:

- precisely set the problem (what has to be optimised, the scope of parameters, precision, constraints, etc.),
- define a chromosome (in a selected alphabet),
- define a fitness function (for evaluation of individuals).

This preparatory phase is often the most difficult, and it has to be done "manually". A choice of genetic operators and parameters of the **GA** suitable for the current problem also can be time consuming, we can make it use the trial and error method. Some knowledge about the influence of particular operators on the tempo and mode of evolution as well as some experience and intuition are desirable features of a **GA**'s designer.

BIOLOGICAL INSPIRATIONS OF THE *K*-MODEL

However artificial genetic algorithms imitate natural evolution, they are significantly differ from the nature. Some of the features of biological evolution being inspiration for designing of the *K*-Model are shortly described below.

Macromutations: A tempo of biological evolution is still strongly discussed by biologists (Hoffman, 1983). In a long perspective, gaining a new fitness niche, evolution of a population goes with a different rate and in different direction (Fogel, 1992, Hoffman, 1983). During the history of life, there could exist some sudden, stepping changes, singular macromutations affected essential differences between a child and its parents (Dawkins, 1995). *Macromutations* are seen as a process which gives radical phenotype effects. Biologists commonly agree that macromutations are present in the nature (Mayr, 1963, Gould, 1991).

Redundant genes: It is well known that only a small part of biological organisms' genotypes encodes the expressible features of the organisms. These genes we call *active* genes or *phenotype* genes (Gould, 1991, Szarski, 1986). Remaining part of genotypes does not influence on a phenotype. They are called *redundant* or *latent* genes. But genetic

operators, such mutation and crossover act also on them. Therefore, more changes in genome, at least in the first period, are neutral, or even potentially destructive. The partition of a genotype between the *active* and the *redundant* genes, is not fixed. As an effect of some processes a *redundant* gene can become the *active* (expressible) one, and it influences the organism's phenotype. Believing that nature is an efficient process, we can ask: Why redundant genes are not considered in artificial genetic (or widely, evolutionary) algorithms? What benefit can we have from redundant genes?

Pleiotropy and polygene effects: Usually, a simple relation: *one gene – one phene* is employed for representation of chromosomes in **GAs**. This means that each gene directly codes one parameter of an optimized function. But in biology, a single gene of individual can have an impact, at the same time, on several phenotype features. Such an effect is called a *pleiotropy*. On the other hand, each single phenotype's feature of an individual (its phene) can be determined by simultaneous influence of a number of genes – this effect is called a *polygene*. A model without above effects is a great simplification. We have designed three models to check how the coding scheme influence on tempo and mode of evolution:

- **CGA** (Classic GA), named *Version_1*,
- **GA** with small pleiotropy and polygene effects and real coded genes, and the third, named *Version_2*,
- **GA** acting directly on the parameters of optimized function (without a genotype level), named *Version_3*.

A comparison of simulations using the three above enumerated models shows that the tempo and mode of evolution depend on the used model, and evolution goes better with pleiotropy and polygene effects (Kwasnicka, 1997, 1998).

The alphabet of genes coding: Watson and Crick have discovered that genes are digital strings of information, but not binary as it is assumed in the classic genetic algorithm. It is difficult to justify the binary coding on the base of biological analogy. Lately, the thesis that binary coding is the most effective, has been discussed, and it is easy to find in the literature applications with different alphabets, suitable for the current task (Goldberg, 1991, Michalewicz, 1992, Kwasnicka 1998a).

THE *K*-MODEL – A VERBAL DESCRIPTION

The *K*-Model encompasses pleiotropy and polygene effects, redundant genes, and it bases on real coding chromosomes. An exemplary schema of an individual is shown in Figure 1. Each individual consists of a constant number phenotype genes *np*, and some number of redundant genes *nr*. Genes are placed in assumed number n_{ch} of chromosomes. A number of redundant genes can change during evolution, but the maximal length of a single chromosome n_{max} is assumed.

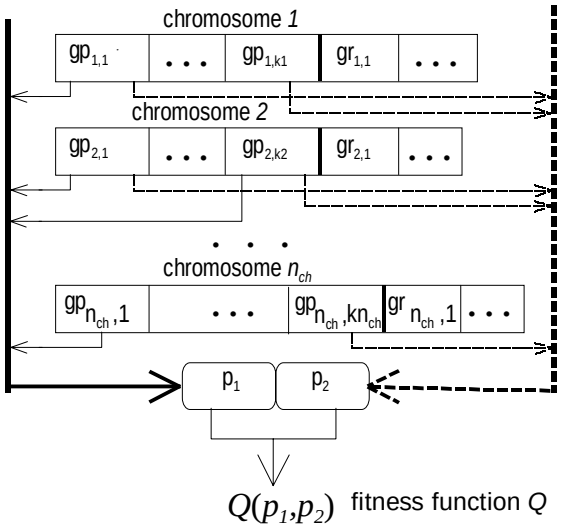

Figure 1. The general scheme of individual's representation in the *K*-Model

Depending of the *fitness function*, a number of phenes are coded by genes. Biology does not give us an intuition how phenes are coded by genes. We assume that each phene is a linear combination of a number of phenotype genes. We assume a pleiotropy matrix ***A*** as coefficients' matrix for the linear dependence. A gene can affect a phene in positive

or negative way, what depends on the value of coefficient $a_{i,j}$. Such representation allows us to model the process called *epistasis*[1].

Population: In general, we assume an average number of individuals in the population $\mathbb{N}$ (the level of environment's saturation). During evolution the size of the population $\mathbb{N}(t)$ changes around this value. It is possible to simplify the model by assuming a constant size of an evolving population $\mathbb{N}$.

Selection: For each individual in the population, the number of its offsprings is calculated according to some selection scheme. In the *K*-Model, the Poisson distribution, with expected value λ equal to the ratio of the quality of the individual Q to the average quality of the population Q_{av}.

Genetic operators:

Recombination: Reproduced individual can exchange its chromosome with randomly selected individual. The probability p_{rkm} for each chromosome is assumed.

Mutation: It is implemented as a random change of a single gene of an individual (with probability p_{mut}). The size of this change is randomly selected within the assumed maximum range (mut_{max}). We see that we have to state the two parameters connected with mutation process. Mutation of redundant genes is seen as the *neutral mutation*. Redundant genes can accumulate changes given by mutations during a number of generations.

Transposition: From time to time, one redundant gene exchanges its place with a randomly selected phenotype gene. The redundant gene becomes the phenotype one, and vice versa. Transposition occurs with probability p_{trs}. All genes have the same chance to be exchanged.

Transition: The process plays the role only if redundant genes are present. It causes that a single gene (in fact, its copy) can be moved (with probability p_{trz}) from one individual to another. Each gene of an individual can be moved and added to a randomly selected chromosome of a randomly selected individual, but always as a redundant one.

We introduce additional two operators for modelling the macromutations, namely recrudescence and crisis.

Recrudescence – in each generation of an evolving population, a number of individuals have enlarged probabilities of mutation, recombination transition and transposition processes. It gives a radical reorganization of genotypes of individuals (Mayr's "*loosing the cohesion of genotype*"). Most of them are eliminated, but randomly, there arise Goldschmidt's *hopeful monsters* – offsprings that survive and can enable the population to achieve a new fitness niche. In biology, internal stabilizing factors are responsible for that process. The parameter *rcd* describes the part of population assumed for the recrudescence process.

Crisis is also a radical reorganization of genotypes, but it concerns all individuals in an evolving population, and it cannot happen frequently. In biology, external factors are responsible for such processes. Crisis also models macromutations, but they occur rarely and affect all individuals in the population. In the model, a number of generations between crisis t_{kr} is assumed.

THE *K*-MODEL – A FORMAL DESCRIPTION

We consider evolution of a population in discrete time t, the generation is a unit of time. Let us assume:

$\mathcal{P}(t)$ – a population in time t consists of $\mathbb{N}(t)$ individuals $\mathbb{I}_i$: $\mathcal{P}(t) = \langle \mathbb{I}_1, \mathbb{I}_2, ..., \mathbb{I}_{\mathbf{N}(t)} \rangle$.

Each individual is represented as its genom: $\mathbb{I}_i \to G_i$, $i = 1, ..., \mathbb{N}(t)$.

A genom is a vector of chromosomes: $G_i = \left\langle C_i^1, ..., C_i^{n_{ch}} \right\rangle$.

Each chromosome consists of a number of phenotype genes and a number of redundant (latent) genes, each gene is an integer value and is identified by its number in the chromosome, a number of chromosome j, and a number of individual i:

$$C_i^j = \left\langle g_1^{i,j}, g_2^{i,j}, ..., g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, ..., g_{np_j+nr_j^i}^{i,j} \right\rangle,$$

where: np_j – number of phenotype genes in the j-th chromosome (the same for all individuals),

nr_j^i – number of redundant genes in the j-th chromosome of i-th individual,

$g_1^{i,j}, g_2^{i,j}, ..., g_{np_j}^{i,j}$ – phenotype genes in the j-th chromosome of i-th individual,

---

[1] *epistasis* – a form of interaction between nonallelic genes in which one combination of such genes has a dominant effect over other combinations (Random House Webster's Electronic Dictionary and Thesaurus, College Edition, W. 1.0, Reference Software International, 1992). A gene called *epistatic* one, can suppress the phenotype expression of another gene, called *a hypostatic* gene.

$g^{i,j}_{np_j+1}, \ldots, g^{i,j}_{np_j+nr^i_j}$ – redundant genes in the *j*th chromosome of *i*th individual,

$i = 1, \ldots, \mathbb{N}(t)$, $j = 1, \ldots, n_{ch}$; $nr^i_j \geq 0$, $np_j + nr^i_j \leq n_{max}$, $g^{i,j}_k \in \mathbf{Z}$, $\sum_{k=1}^{n_{ch}} np_k = np$.

n_{max} – maximal number of genes in a single chromosome,
n_{ch} – number of chromosomes in the individuals' genotype,
np – number of phenotype genes in each individual in a population.

For each individual i we define a vector AG_i of its all phenotype genes:

$$AG_i = \left\langle gp^i_1, gp^i_2, \ldots, gp^i_{np} \right\rangle$$

Each phene of *i*th individual is a product of its genotype phenes and pleiotropy matrix $\boldsymbol{A}$ (we assume a linear dependency between phenotypes genes and phenes):

$$Ph^i = \left\langle p^i_1, p^i_2, \ldots, p^i_m \right\rangle = AG_i \times A ,$$

where: Ph^i – vector of phenes of *i*th individual,

p^i_k – *k*-th phene of *i*th individual, i=1,2, ..., m, m is a number of phenes (variables in optimization function),

$[A]_{np \times m}$ – pleiotropy matrix, $a_{kl} \neq 0$ if *k*-th gene influences *l*-th phene.

A fitness value if *i*th individual (Q^i) directly depends on its phenes:

$$Q^i = Q(\mathbf{I}_i) = Q(Ph^i) = Q(AG_i \times A) = Q(p^i_1, \ldots, p^i_m)$$

Genetic operators included in the *K*-Model

Selection: We assume that the number of offsprings is calculated according to the Poisson distribution. In each unit of time (generation) t, after calculation of fitness functions $\boldsymbol{Q_i}$ of all individuals in the population $\mathcal{P}(\boldsymbol{t})$, an expected number of offsprings λ is calculated:

$$\lambda_i = \frac{Q^i}{Q_{av}(t)} \cdot S(t) ,$$

where: λ_i – expected number of offsprings of *i*-th individual,

Q^i – quality (fitness) value of *i*-th individual,

Q_{av} – average quality of the population $\mathcal{P}(\boldsymbol{t})$; it is calculated as follows:

$$Q_{av} = \frac{1}{\mathbf{N}(t)} \sum_{j=1}^{N(t)} Q^j ,$$

$N(t)$ – size of the population in generation t,

$S(t)$ – actual rate of environment saturation, it is calculated as follow:

$$S(t) = 1 + a \cdot \frac{\mathbf{N} - N(t)}{\mathbf{N}} ,$$

$\mathbf{N}$ – assumed saturation of population's environment,

a – assumed coefficient, it responses for the rate of achieving the saturation level $\mathbf{N}$, usually we assume a=1.

Recombination: Recombination operates on chromosomes. Reproduced individual can exchange its chromosomes with selected individual. Each chromosome can be exchanged with assumed probability p_{rkm}. The second individual for reproduced one is selected randomly. Recombination of *j*th chromosome between *i*-th and *k*-th individual goes as follow:

$$G_i = [C^1_i, C^2_i, \ldots, C^j_i, \ldots, C^{n_{ch}}_i] \text{ and } G_k = [C^1_k, C^2_k, \ldots, C^j_k, \ldots, C^{n_{ch}}_k]$$

$$\overset{p_{rkm}}{\Rightarrow} : \quad G_i' = [C^1_i, C^2_i, \ldots, C^j_k, \ldots, C^{n_{ch}}_i] ,$$

where: G_i' – genom of *i*-th individual after recombination.

Mutation: Mutation operates on genes. It is a random change of a gene's value. Each gene mutates independently, with assumed probability p_{mut}. During mutation process, a value of mutated gene is increased or decreased about the random value from the assumed range [$-mut_{max}$, mut_{max}] according to the uniform distribution:

$$C_i^j = \left\langle g_1^{i,j},\ldots, g_k^{i,j}, \ldots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \ldots, g_{np_j+nr_j^i}^{i,j} \right\rangle \overset{p_{mut}}{\Rightarrow}$$

$$C_i^{j'} = \left\langle g_1^{i,j},\ldots, g_k^{i,j} + \Delta g, \ldots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \ldots, g_{np_j+nr_j^i}^{i,j} \right\rangle,$$

where: $C_i^{j'}$ is a chromosome after mutation of one gene,

Δg is a random number, and: $-mut_{\max} \le \Delta g \le mut_{\max}$.

Each gene (phenotype and redundant) mutates independently.

Transposition: The transposition process acts on genes, it allows to use neutral mutation (accumulated during evolution). Obviously, the redundant gene has to be phenotype one to influence the individual's fitness. Described process allows for such modification. With assumed probability p_{trs} two genes, one active and ane redundant, exchange their places, so, the redundant gene becomes the active one:

$$C_i^j = \left\langle g_1^{i,j},\ldots, g_k^{i,j}, \ldots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \ldots, g_{np_j+r}^{i,j}, \ldots, g_{np_j+nr_j^i}^{i,j} \right\rangle \overset{p_{trs}}{\Rightarrow}$$

$$C_i^{j'} = \left\langle g_1^{i,j},\ldots, g_{np_j+r}^{i,j}, \ldots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \ldots, g_k^{i,j}, \ldots, g_{np_j+nr_j^i}^{i,j} \right\rangle.$$

Transition: This process acts on genes, it causes that a gen from one individual is copied to the second, randomly chosen individual. Each gene can be copied and it can be put on a random chromosome of randomly selected individual but only as the redundant gene. The probability p_{trz} is given as a model's parameter.

$$C_i^j = \left\langle g_1^{i,j},\ldots, g_k^{i,j}, \ldots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \ldots, g_{np_j+r}^{i,j}, \ldots, g_{np_j+nr_j^i}^{i,j} \right\rangle \text{ and}$$

$$C_l^v = \left\langle g_1^{l,vj},\ldots, g_k^{l,v}, \ldots, g_{np_j}^{l,v}, g_{np_j+1}^{l,v}, \ldots, g_{np_j+r}^{l,v}, \ldots, g_{np_j+nr_j^i}^{l,v} \right\rangle$$

$$\overset{p_{trz}}{\Rightarrow} C_i^{j'} = \left\langle g_1^{i,j},\ldots, g_{np_j+r}^{i,j}, \ldots, g_{np_j}^{i,j}, g_{np_j+1}^{i,j}, \ldots, g_k^{i,j}, \ldots, g_{np_j+nr_j^i}^{i,j}, g_k^{l,v} \right\rangle.$$

The number of redundant genes in the *j*-th chromosome of *i*-th individual is increased, because of a copy of *k*-th gene of *v*-th chromosome of *l*-th individual is put as the last gene of the *j*-th chromosome of *i*-th individual.

Specialized operators for macromutation modelling: Three methods of macromutations' modeling are possible in the *K*-Model:

Redundancy: By the accumulation of genes' changes caused by neutral mutation (in the part of chromosomes with redundant genes),

Recrudenscence[2]: The probabilities of described above genetic operators (p_{rkm}, p_{mut}, p_{trs}, and p_{trz}) are increased in the part of population. It occurs during each generation of evolution. Individuals for recrudescence are selected randomly with uniform distribution.

Crisis: It acts as recrudescence, but all individuals are the subject of a crisis after each t_{kr} generations.

SIMULATION STUDY

Many real problems solved by using *Genetic* or *Evolutionary Algorithms* are the optimization problems with many local optimums. The main aim of the study is to investigate the possibility of escaping from the lower peak and reaching the global optimum. During experiments we have observed tempo and mode of evolution. A number of generations in which population reaches the highest peak is a measure of tempo of evolution. Mode of evolution is the way in which population evolves: gradually or stepwise goes through the phenotype space; what diversity is observed in the evolving population (how close is the average fitness of a population to the fitness of the best individual), and so on. We have made wide simulation studies of *K*-Model. In this paper some of them are described, the others are only mentioned in the summary.

[2] recrudescence comes from Latin *recrudesco* (to break out, to open, to renew) and *recrudescere* (to become raw again), it means a new outbreak after a period of abatement, inactivity or after a dormant period.

During simulation study we have used a family of exponential functions as fitness (quality) functions. They enable us to define multivariable and multimodal functions, that is: a number of peaks, their coordinates and slope, as well as a number of variables in the function. The general equation of the function is:

$$Q(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{l_s} h_1 \cdot e^{-n_i \cdot \sum_{j=1}^{m} (x_j - x_j^i)^2},$$

where: x_i – ith coordinate of fitness function Q,

x_j^i – value of jth coordinate of i-th peak,

m – dimension of Q,

l_s – number of peaks,

h_i – height of i-th peak,

n_i – slope of i-th peak.

Other functions, including the de Jong test functions, were used for simulation study, too. Simultaneously, an influence of pleiotropy matrix $\boldsymbol{A}$ on tempo and mode of evolution was observed.[3] Below the results of active marcomutations in the K-Model are presented.

From the optimization point of view, the success of evolution lie in finding the good solution, so it is enough if a single individual is present near the global optimum. From dynamic of population point of view, the most important question is if an evolving population can settle the best niche (global optimum).

Macromutations in the K-Model

The first model of macromutation is redundancy of genotypes. Let us use the quality functions:
Q_1 with two peaks and two variables: h_1=1, in the point **x**=[5, 5], h_2=1.5, in the point **x**=[20, 20], n_1=n_2= 0.02;
Q_2 with two peaks and five variables: h_1=1, in the point **x**=[5, 5, 5, 5, 5], n_1=0.02, h_2=1.5, in the point **x**=[20, 20, 20, 20, 20], n_1=0.00225.
Function with dimension equal to 10 and 30 were also used for experiments. In all experiments the initial population was placed on the lower peak.

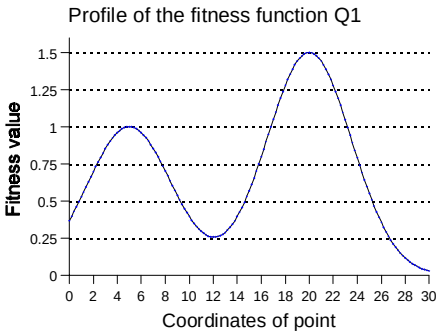

Figure 2. The profile of fitness function Q_1

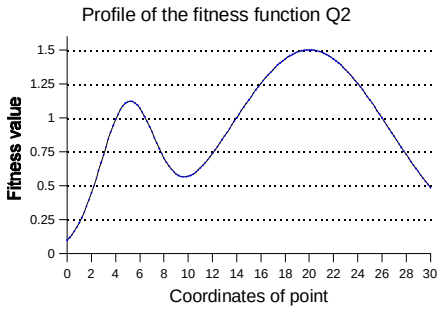

Figure 3. The profile of fitness value Q_2

The results obtained for 5-dimension function Q_2 are presented below. A solution with fitness value greater than 1.35 is seen as a success of evolution. All experiments were repeated minimum ten times.

The role of redundancy

1. Function Q_2, mut_{max}=10, values of all genes in an initial population: 1000, all phenes: 5. Without redundancy, population cannot get the higher peak, it evolves around the lower peak independent on the probabilities p_{rkm}, p_{mut}. The same result is observed with redundancy, independent of p_{trs} and p_{trz}.

[3] All simulation results and properties of the K-Model are shown in the forthcoming book *Obliczenia Ewolucyjne* (*Evolutionary Computation*, in Polish), chapters 5-9.

2. Function Q_2,. mut_{max}=100, p_{rkm}=0.25, p_{mut}=0,2, p_{trs} and p_{trz} are changed from 0.1 to 1. Evolution goes without success. Increasing probabilities recombination and mutation allows population to reach the higher peak. It occurs when p_{rkm}=0.5, p_{mut}=1, p_{trs}=p_{trz}=0.2, after 3300 generations. With so high recombination and mutation population loses the good solution and cannot occupy global optimum. The diversity of the population is high (see Figure 4).

It is worth to mention, that without redundant genes, population is not able to find a global optimum. With redundancy and greater p_{rkm} and p_{mut}, population finds a relatively good solution, but cannot reach a high average fitness.

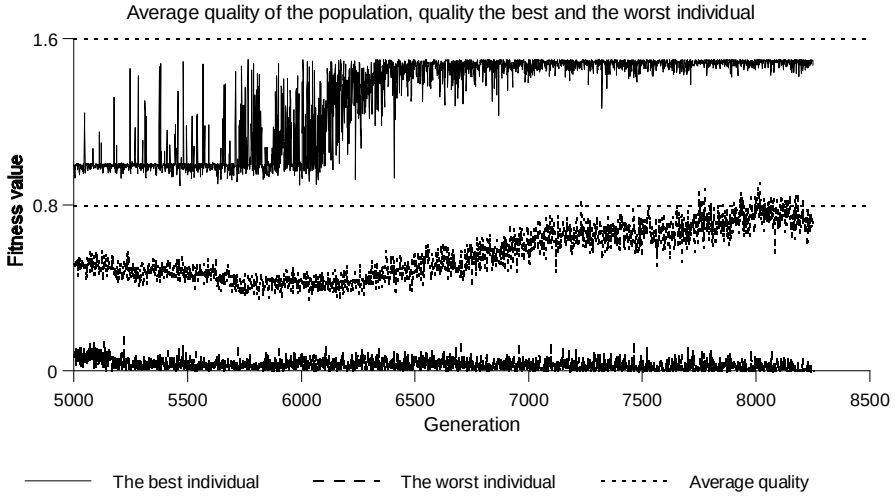

Figure 4. The average fitness of th population and the fitness of the best and the worst individual in this population: mut_{max}=100, p_{rkm}=0.5, p_{mut}=1, p_{trs}=p_{trz}=0.2 – a typical evolution

In the above simulations the best solutions appeared in the best area of phenotype space, but population is not able to exploit them. We can say, that our model can explore potential solutions space but it cannot exploit them. It is caused by frequent recombination and mutation.

3. Function Q_2, mut_{max} is stated to be equal 500. Population needs about 4500-5500 generation to get a global optimum. Population is more convergent. Increasing probabilities transposition and transition cause earlier finding the best solution, but the average quality is low (compare Figure 5 and Figure 6).

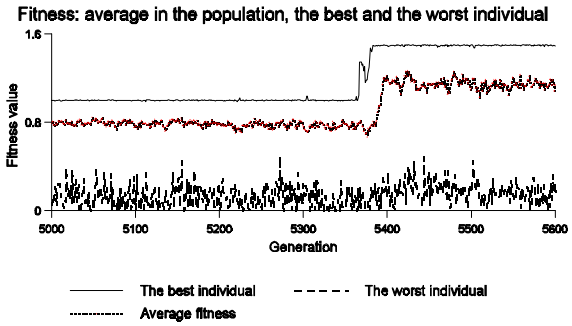

Figure 5. Average fitness and fitness of the best and the worst individuals during evolution with parameters: mut_{max}=500, p_{rkm}=0.25, p_{mut}=0.2, p_{trs}=p_{trz}=0,01

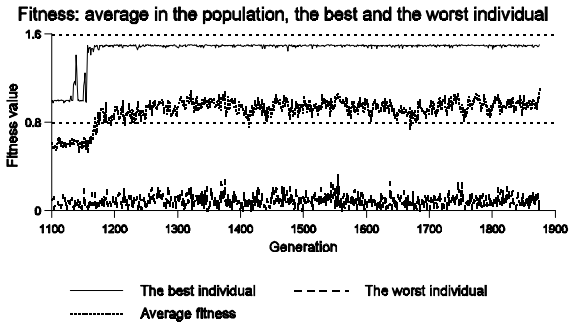

Figure 6. Average fitness and fitness of the best and the worst individuals during evolution with parameters: mut_{max}=500, p_{rkm}=0.25, p_{mut}=0.2, p_{trs}=p_{trz}=0,05

4. Experiment with 2-dimensional function Q_1. For evolving population this function is easier then Q_2. For mut_{max}=10, p_{rkm}=0.25 and p_{mut}=0.2, without redundancy population cannot find global solution during 1000 generations. Adding redundant genes (p_{trs}=p_{trz}=0.05) causes that in all ten simulation runs, 439 to 824 generations were enough to reach the higher peak.

The role of recrudenscence

5. Function Q_2, mut_{max}=10, 100, 500; p_{rkm}=0.25, 0.5, 0.8, 1; p_{mut}= 0.2, 0.5, 0.8, 1; *rec*=10%, 50% , without redundancy evolution fails. In all 10 simulation runs with above parameters, evolution stops without success (after assumed 5000 generations). We can say that under that conditions the recrudescence is not able to assure efficiency of evolution.

The role of crisis

6. Function Q_2. For based operators mut_{max}=100, p_{rkm}=0.25, p_{mut}=0,2, we add the crisis. We assume t_{kr}=20, and 50, and in crisis the probabilities are enlarged to p_{rkm}=0.25, 0.5, 0.8, 1, p_{mut}=0,2, 0.5, 0.8, 1. All simulation runs failed. Redundant genes were absent in these experiments.

It seems that the role of redundant genes is significant, they enable population to settle the best niche.

Effects of joined redundancy, recrudescence and crisis

7. Function Q_2, mut_{max}=200, p_{rkm}=0.25, p_{mut}=0.2, p_{trs}=p_{trz}=0.1, *rec*=5%, t_{kr}=20 (for individuals under the recrudescence and during crisis: p_{rkm}=0.5, p_{mut}=0.1, p_{trs}=p_{trz}=0.2). Success of evolution is possible during about 10000 generations. Average quality of the population is higher than the lower peak and we observe short fluctuations of average fitness. Fluctuations are caused by crisis, when the population explores larger area of phenotype space (see Figure 7, 8).

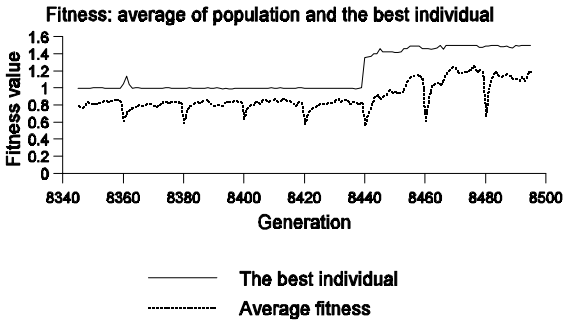

Figure 7. Average fitness and the best individual fitness: mut_{max}=200, p_{rkm}=0.25, p_{mut}=0.2, p_{trs}=p_{trz}=0.1, *rec*=5%, t_{kr}=20

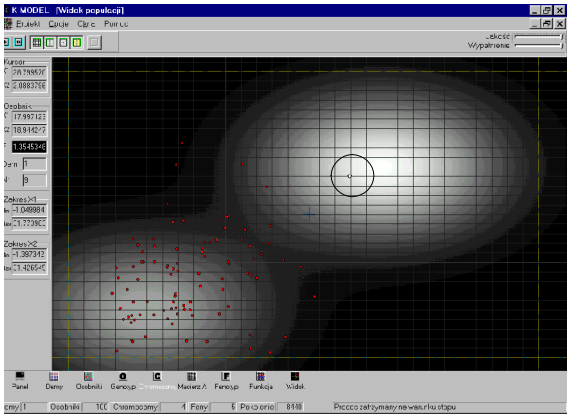

Figure 8. Generation 8440, parameters as in Figure 7, the first individual is seen on the higher peak (in circle)

Population goes on the global fitness niche, and after 20 generation average fitness is higher then the lower peak.

8. Function Q_1. We apply redundancy with crisis and redundancy with recrudescence. More effective is evolution with crisis than with recrudescence. Adding crisis causes that evolution goes quicker about 15%, but we have not observed this using recrudescence instead of crisis. Changing t_{kr} to 10 and to 50 does not affect significantly the tempo and mode of evolution in these experiments.

We must remember that crisis is a form of macromutation, it enables population to test wide area of phenotype space, but on the other hand, it decreasing average fitness of population.

Population size in the K-Model

Numbers of experiments show that for Q_2 evolution can go with success with following parameters: mut_{max}=600, p_{rkm}=0.25, p_{mut}=0.1, p_{trs}=p_{trz}=0.05. These parameters values suit as the base for further experiments with this function. Assumed number of generation is 2000, we have evolved population consists of 10, 100 and 100 individuals. The results are in Table 1. The results surprise us, we see that smaller populations evolve quicker then the larger ones. It is even measured in generations, not the number of fitness calculations. Let us see the changes of average qualities of above populations (Figure 9). Population of 1000 individuals evolves gradually, after finding a good solution the population goes to the higher peak. Average quality rises but it is not high because of macromutations (Figure 10). Different situation we can see in Figure 11. Observed fluctuation are rather big. Evolution of such population is similar to the random walk through the phenotype space. Other experiments show, that such population is able to find quickly the highest peak when we have a multimodal fitness function (e.g., with four peaks). Small population can walk through the valley, and find a good area, but it cannot persist in this place. Experiments with other function and parameters were made, for example to test the influence of redundancy and dimension of fitness functions (10 and 30 variables). For Q_1 function, without redundancy, and mut_{max}=20, p_{rkm} =0.25, p_{mut}=0.2, we obtained a quite different results – 1000 individuals need from 314 to 1879 generations, 100 individuals – from 1532 to 22935 generations (about 6000 average), and 10 individuals cannot reach global optimum. Adding redundancy (p_{trs}=p_{trz}=0.05) changes above

results: population with 1000 individuals needs 42-205 generations, 100 individuals: 151-439 generations, 10 individuals: 81 to 121 generations to reach optimum.

Table 1. Results of evolution with Q_2 and different size of population

Size of population	Number of generations after which a solution with quality grater than 1.35 is achieved in particular simulation run									
	1	2	3	4	5	6	7	8	9	10
10	649	652	532	548	681	586	697	462	348	1330
100	1693	1041	1170	1132	1005	1380	1021	1269	1514	1330
1000	–	2598	2266	–	2049	–	2438	–	–	3863

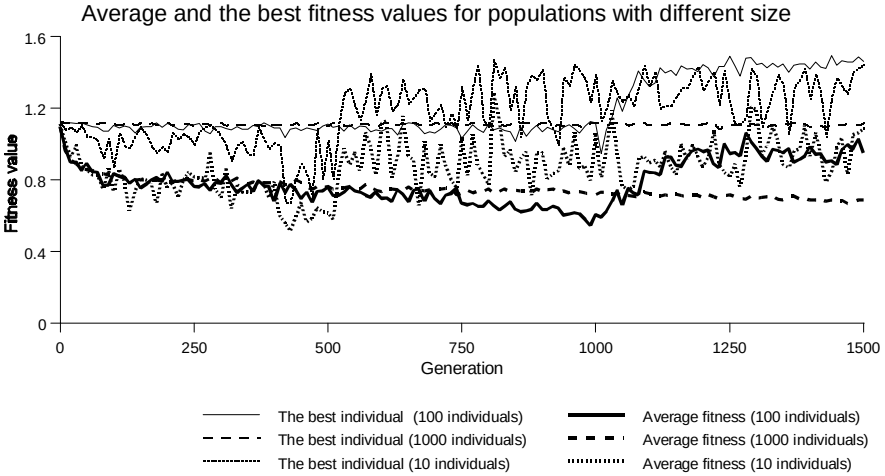

Figure 9. Changes of the best and average qualities in the population depending on the populations' size, the coordinates 'generation' is scaled after each 10 generations

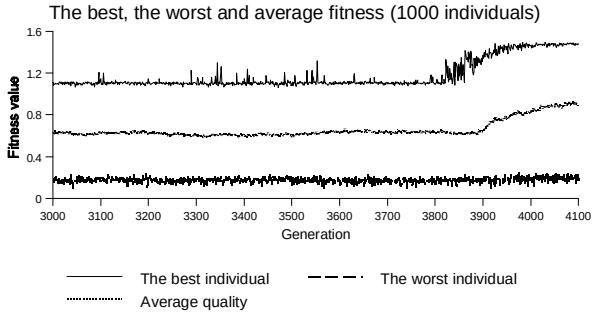

Figure 10. Fitness values of population consists of 1000 individuals (the best, the worst individual, and the average population's fitness)

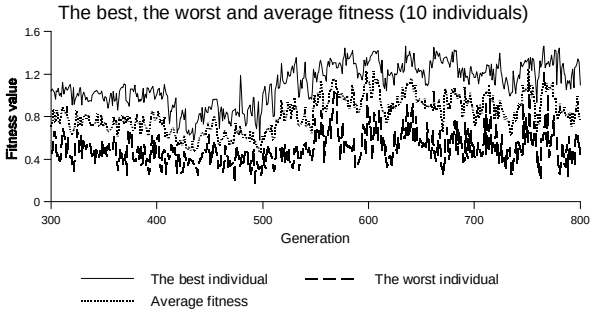

Figure 11. Fitness values of population consists of 10 individuals (the best, the worst individual, and the average population's fitness)

We have used function with 30 variables and two peaks, setting parameters as follow: mut_{max}=600, p_{rkm} =0.25, p_{mut}=0.2, p_{trs}=p_{trz}=0.05. All populations evolve with success (1000 individuals: 3181 to 3543 generations, 100 individuals: 3422-3775 generations, 10 individuals: 4505-6854 generations). The smallest populations usually reach place near the global optimum quickly, but it cannot find the top of the peak. We see that the dimension of fitness function changes the tempo of evolution, taking into account the size of a population. We can make a mental experiment. Let us assume that 10% of a single coordinate belongs to the optimal area. Probability to point the optimal area randomly is 10^{-1}. Next let we have a square, in which an optimal area is 10% of each coordinate. Probability that we randomly point this area is 10^{-2}. For 5th dimension space it is 10^{-5}, but for 30 dimension, the probability is very small: 10^{-30}. Population of 10 individuals walk more randomly then the bigger ones. Therefore the small populations are more effective for smaller dimension fitness functions.

Redundancy of genotypes gives a population possibilities to test wide area of phenotype space. This effect is seen during evolution all populations but the redundancy hardly influences smaller populations.

SUMMARY

The new schema of individuals' representation and some new genetic operators are proposed in the paper. Wide simulation study of proposed *K*-Model have been made, but only a small part of them are presented here. The proposed algorithm seems to be very effective. The five de Jong function were tested and obtained results are satisfactory. The pleiotropy matrix $\boldsymbol{A}$ gives possibility to direct the course of search in evolution.

In numerous experiments we have found that partition of the whole population into a number of semi-isolated demes (subpopulations) allows for much quicker evolution. It is interesting that a population structure with one big deme and numerous small demes is the most efficient.

Applications of genetic algorithms to specific problems (e.g., a neural network design problem) do not give satisfactory results. Therefore, the investigation of dependence of evolutionary algorithms on applied genetic operators and parameters seems to be important. It can suggest an efficient approach to build good optimizing tools, applicable to a wide spectrum of problems. We have started to elaborate the system called INT-EVOL. The family of evolutionary algorithms will act, and the intelligent control module will analyse their results and control their parameters.

References

Altenberg L. (1994), *Evolving Better Representations Through Selective Genome Growth*, Proceedings of the IEEE 1994 World Congress on Computational Intelligence, pp. 182-187.

Back T. (1996), *Evolutionary Algorithms in Theory and Practice*, Oxford University Press, New York, Oxford.

Beasley D., Bull D.R., Martin R.R. (1994), *Research Topics in Genetic Algorithms*, Vivek, A Quarterly in Artificial Intelligence, April 1994, National Centre for Software Technology, Bombay, India.

Dawkins R. (1995), *Rzeka genów* (*River Out of Eden. A Darwinian View of Life*, translated), Wydawnictwo CIS, Oficyna Wydawnicza MOST, Warszawa, Poland.

Fogel D.B. (1992), *Evolving Artificial Intelligence*, University of California, SanDiego, A dissertation for the Phd.

Goldberg D.E. (1989), *Genetic Algorithms in Search, Optimization, and Machine Learning*, Addison-Wesley Publishing Com., Inc.

Goldberg, D.E., (1991), *Real-coded Genetic Algorithms, Virtual Alphabets, and Blocking*, Complex Systyems, 5(2), pp.139-168.

Goldschmidt R., (1940), *The Material Basis of Evolution*, New Haven: Yale University Press.

Gould S.J. (1991), *Niewczesny pogrzeb Darwina*, Państwowy Instytut Wydawniczy, Warszawa, Poland (essays from: *Ever Since Darwin*, *The Panda's Thumb*, *Hen's Teeth and Horses Toes*, *The Flamingo's Smile*).

Hoffman A. (1983) *Wokół ewolucji*, (*About evolution*, in Polish), Biblioteka Myśli Współczesnej, PIW, Warszawa, Poland.

Holland J.H. (1975), *Adaptation in Natural and Artificial Systems*, The University of Michigan, U.S.A.

Kwaśnicka H. (1997), *Pleiotropy and polygene effects in evolutionary algorithms*, Proceedings of the Workshop on Intelligent Systems, June 9-13, 1997, Zakopane, Polish Academy of Science, Poland.

Kwaśnicka H. (1997a), *The Role of Redundant Genes in Evolutionary Algorithms – Simulation Study*, 15th World Congress on Scientific Computation, Modelling and Applied Mathematics, Berlin, August 1997, vol. 2, Numerical Mathematics, pp. 645-650.

Kwaśnicka H. (1997b), *Redundancy of Genotypes as the Way for Some Advanced Operators in Evolutionary Algorithms – Simulation Study*, VIVEK A Quarterly in Artificial Intelligence, Vol. 10, No. 3, July 1997., pp. 2-11, National Centre for Software Technology, Mumbai, India.

Kwaśnicka H. (1998), *Efficiency Of Genetic Algorithms With Plejotropy and Polygene Effect – Simulation Study*, VI Proceedings of the Workshop on Intelligent Systems, June 9-13, 1997, Zakopane, Polish Academy of Science, Poland. Malbork.

Kwaśnicka H. (1998a), *Evolutionary Computation in Applications*, The invited lecture at the 7th International Colloquium on Numerical Analysis and Computer Science with Applications, slides in the home page: http://www.ci.pwr.wroc.pl/~kwasnick), August 13-17, 1998, Plovdiv, Bulgaria.

Michalewicz Z. (1992), *Genetic Algorithms + Data structure = Evolution Programs*, Springers Series Artificial Intelligence (second edition in Polish in 1996).

Mayr E. (1982), *Speciation and Macroevolution*, Evolution 36 (6).

Szarski H. (1986), *Mechanizmy ewolucji*, (*Mechanisms of Evolution*, in Polish) PWN, Warszawa.

Waddington C.H. (1977), *Stabilization in Systems. Chreods and Epigenetic Landscapes*, Futures, April 1977.