

Aspects of Adaptation in Natural and Artificial Evolution

Michael Affenzeller
Upper Austrian University of
Applied Sciences,
College of Information
Technology at Hagenberg,
Softwarepark 11
4232 Hagenberg, Austria
michael@heuristiclab.com

Stefan Wagner
Upper Austrian University of
Applied Sciences,
College of Information
Technology at Hagenberg,
Softwarepark 11
4232 Hagenberg, Austria
stefan@heuristiclab.com

Stephan Winkler
Upper Austrian University of
Applied Sciences,
College of Information
Technology at Hagenberg,
Softwarepark 11
4232 Hagenberg, Austria
stephan@heuristiclab.com

ABSTRACT

This work addresses selected aspects of natural evolution, especially of the field of population genetics, that are considered to be meaningful for algorithmic further developments in the field of Genetic Algorithms (GAs) and Genetic Programming (GP) by the authors. In this connection special attention is devoted to selection and replacement strategies, as these are exactly the aspects that do not depend on certain problem representations and corresponding operators and therefore allow generic algorithmic further development. The concept of offspring selection is described as an example of such a problem independent further developed algorithmic concept, which allows to maintain the relevant genetic information stored in a population more efficiently. The potential of this new selection strategy is pointed out in terms of references to recent results achieved on the basis of well known benchmark problems in the field of GAs and GP.

Categories and Subject Descriptors

I.2.8 [Artificial Intelligence]: Problem Solving, Control Methods, and Search—*Heuristic Methods*

General Terms

Algorithms

Keywords

Evolutionary Computation, Population Genetics, Genetic Drift, Population Structure, Selection, Selection Pressure

1. INTRODUCTION

Research in population genetics was introduced into the scientific community in the context of genetics and evolution biology. The history of evolution biology as an independent discipline dates back to the 17th and 18th century. At that time the individual was considered as the object of evolution

which is also the outmoded concept of Lamarck's hypothesis [13] about the heredity transmission of attributes obtained by the development of the individual. In the middle of the 19th century, Charles Darwin and Alfred Russel Wallace formulated the idea of selection of individuals. A main difference to Lamarck's theory was that not the individuals but the populations and species evolve [10]. Darwin's theory of evolution also represents the theoretical basis of Genetic Algorithms. In contrast to Lamarck's postulate, Darwin's theory assumes no individual evolution apart from unfrequent mutations: As the theory of chromosomes and thus modern genetics evolved no more until around 1900, Darwin was only able to speculate about the principles of heredity transmission. Populations vary in their frequency of various genetic variants, due to differences in terms of success of individual reproduction which is determined by selection and by chance.

In contrast to the early years of research of evolution, where selection was considered as the more or less only driving force of evolution, modern population genetics also considers the effects of genetic drift - especially for rather small populations (as it is certainly the case in Evolutionary Computation). Population genetics is a well suited paradigm when it comes to describing the topology and temporal dynamics of genetic variation in natural populations with the goal to understand the evolutionary forces that act on populations.

Similar to Evolutionary Computation, population genetics also has an empirical as well as a theoretical component, and especially for scientists in the field of Evolutionary Computation, it should be a very fruitful approach to consider recent developments of population genetics, which should be kept in mind as the bionic role-model for further developments. Therefore, in the following we summarize some up to date considerations of population genetics which are relevant for the advanced GA concepts. In doing so we first state some basic principles of population genetics and give a short overview about the characteristics and the interactions of selection and genetic drift.

As a result a new problem independent selection model is presented which supports the survival of the fittest alleles rather than survival of the fittest individuals which is exemplarily demonstrated on the basis of a 130-city benchmark TSP. Finally, some references to recent related works that use the described selection mechanism are stated.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO '07, July 7–11, 2007, London, England, United Kingdom.
Copyright 2007 ACM 978-1-59593-698-1/07/0007 ...\$5.00.

2. GENETIC DRIFT

As a consequence of the Hardy-Weinberg law of population genetics [12], evolution will not occur if no mechanisms that can cause evolution are acting and the gene pool frequencies will remain unchanged. Even if no infinitely large population actually exists, this does not necessarily negate the Hardy-Weinberg Law. The effect that is of concern is genetic drift, which is a change in gene frequency that is the result of chance deviation from expected genotypic frequencies. This effect is minimal in moderate sized or larger populations, like is mostly the case in natural populations. But in small populations as used for Genetic Algorithms or for the certain demes of distributed Genetic Algorithms, genetic drift is of major influence and therefore of special interest for our purposes. Griffiths et al. ([11], p.704) give a quite easy to understand and intuitive explanation of genetic drift: "If a population is finite in size (as all populations are) and if a given pair of parents have only a small number of offspring, then even in the absence of all selective forces, the frequency of a gene will not be exactly reproduced in the next generation because of sampling error.

Therefore, genetic drift can be perceived as an undirected change of allele frequency by stochastic processes and has some important properties:

- Genetic drift causes a loss of genetic variation in the population. The velocity of this loss is inversely proportional to the size of the population.
- Genetic drift has a strong effect on newly evolved mutants (this is also the case for advantageous mutants).

In natural populations the loss of genetic variation is compensated by mutation. The interdependency of both forces is called the mutation-drift equilibrium.

There are two reasons for the random effects which are considered as the reasons for genetic drift in natural populations [12]:

- **Mendel's Law of Segregation:** Alleles are randomly and with the same probability distributed on the gametes.
- **Demographic Stochasticity:** This may be traced back to the fact that individuals differ in the number of offspring. In nature the reasons for these differences are complex and are considered accidental.

In the theory of Genetic Algorithms, selection and selection pressure are among the most essential success criteria. But these factors support demographic stochasticity, i.e. the alleles of fitter individuals tend to dominate the population (as these individuals are selected for reproduction more often) whereas alleles which are stored only in individuals with lower fitness tend to get lost due to parent selection. Together with the fact that the size of the population is rather small in a typical GA, or in a certain deme of a distributed GA, indicates the importance of thinking about genetic drift for theoretical GA considerations. In population genetics, the loss of genetic diversity due to genetic drift is mainly considered adverse, as in nature genetic diversity in a population is essential for the ability to adapt to changing environmental conditions. But from an Evolutionary Computations point of view, the fixing of alleles containing genetic information of a global optimal or at least of a highly

qualified solution is desirable. Vice versa, the loss of definitely disadvantageous alleles also positively influences the Evolutionary Algorithm.

3. POPULATION STRUCTURE

In nature the fragmentation of the population of a certain species into unequally sized subpopulations is a commonly observable phenomenon. Many species have a great area of circulation of various geographical and environmental structure which leads to the formation of subpopulations. An important consequence of the population structure is the genetic differentiation of subpopulations, i.e. the shift of allele frequencies in the certain subpopulations. The reasons for genetic differentiation are:

- Local adaptation of different genotypes in different populations
- Genetic drift in the subpopulations
- Random differences in the allele frequency of individuals which build up a new subpopulation

The structure of the population is hierarchically organized in different layers (The concept of hierarchical population structures has been introduced by Wright [28]):

- Individual
- Subpopulation
- Local population
- Entire population (species)

An important goal of population genetics is the detection of population structures, the analyses of consequences and the location of the layer with the most diverse allele frequencies. In our context a deeper consideration of genetic drift and its consequences is of major interest. The aspect of local adaptation of different genotypes in different populations should give useful hints for multi-objective function optimization or for optimization in changing environments.

One consequence of the population structure is the loss of heterozygosity (genetic variation). The Swedish statistician and geneticist Wahlund described that genetic variation rises again if the structure is broken up and mating becomes possible in the entire population. The well established island model (cf. [9]) as the standard model of coarse grained parallel GAs as well as recent developments in SEGA [1] and SASEGASA [6] systematically take advantage of this effect.

Especially SASEGASA, which is a parallel GA based upon the selection concept described in this work utilizes the interactions between directed genetic drift and a new dynamic migration policy by means of self-adaptive selection pressure steering, very advantageously in terms of global solution quality:

In contrast to the island model, SEGA and especially SASEGASA allow dynamic and self-triggered migration in that sense that migration does not happen in fixed intervals like it is the case for island model parallel GAs; SASEGASA in particular initiates migration exactly when subpopulations begin to stagnate which is indicated by high selection pressure values.

As empirically demonstrated in recent conference and journal publications (cf. Section 7) the procedure mentioned

above makes it possible to scale up global solution quality by simply increasing the number of subpopulations, with basically just linearly increasing computational cost. With an increasing number of equally sized subpopulations the probability that essential alleles are bared from dying off increases due to the greater total population size (higher survival probability of the alleles which are not yet considered) as well as due to the finer scaling of reunification phases.

4. SELECTION

Since Darwin [10], selection is considered the most important (and only) evolutionary mechanism for adaptation to the environment. Population genetics considers the basic model of selection for hermaphroditic organisms in the following way:

Random mating \rightarrow selection \rightarrow random mating \rightarrow selection \rightarrow

I.e. Selection mainly depends on the probability of survival of newborn individuals until they reach pubescence, that is called viability in the terminology of population genetics. In population genetics, sexual selection, which considers selection as selection for reproduction (as it is the approach in Genetic Algorithm theory), represents just one of many facets of selection. Our advanced offspring selection scheme, which allows self-adaptive steering of selection pressure as described in Section 6, is therefore much more similar to selection as understood in the context of population genetics. In our new selection model the survival probability is determined by a comparison of the fitness of the newly generated individual, with the fitness values of its parents. Indeed, as demonstrated in the experimental part, it appears that the first sexual selection step (roulette-wheel, linear-rank, tournament) as in case of a standard GA does not drastically effect the qualitative or quantitative performance of the algorithm. Even with random sexual selection (corresponding to the basic model of the population genetic's selection model) the results were about the same or even better than with roulette-wheel or linear-rank, as the first selection step. A very important consequence of selection in population genetics, as well as in evolutionary computation, is its influence on certain alleles. As a matter of principle there are four possibilities for each allele in the population:

- The allele may be fixed in the population.
- The allele may disappear in the population.
- The allele may converge to an equilibrium state.
- No change in the allele frequency.

For the purpose of technical optimization it is desirable that alleles which are part of a global optimal or at least a highly qualified solution are fixed in the entire population, but on the other hand it is also desirable that alleles representing genetic information which is definitely not part of a global optimal solution are sieved out in a rather early stage of evolutionary optimization. By the help of the self-adaptive selection pressure model (offspring selection), the newly proposed SASEGASA (described in [6]) algorithm utilizes the interplay between genetic drift and migration in a very directed way: The demes fix a lot of alleles in a quite early

stage of evolution due to their rather small population sizes which promotes genetic drift. In the case of highly multimodal problem situations (which represent the interesting case for optimization) the self adaptive steering of selection pressure (w.r.t. the success of the crossover operator) enables the algorithm to combine the essential parts of genetic information (w.r.t. the global optimal solution) step by step during the recombination phases.

5. ADJUSTABLE AND SELF ADAPTIVE SELECTION PRESSURE STEERING

Selection and selection pressure are in the theory of Genetic Algorithms predetermined by the so-called mating scheme and by the replacement strategy actually deployed. The classical mating-scheme (roulette-wheel) selects the candidates for reproduction proportionally according to their fitness. Thus it should be achieved that offspring of highly fit individuals are represented in the next generation with a higher probability than offspring of average or below average individuals. The goal of this procedure is a continuous advancement of the population over the generations. Typical mating schemes are roulette wheel, linear rank or tournament, which are also often just called roulette wheel selection, linear rank selection or tournament selection in GA literature. This classical GA selection concept is known as sexual selection in the terminology of population genetics. From a population genetics point of view sexual selection only covers a rather small aspect of selection which appears when individuals have to compete to attract mates for reproduction.

As already stated before, the population genetics basic selection model considers the selection process as the repeated execution of random rating followed by a selection step; selection mainly depends on the probability of surviving of newborn individuals until they reach pubescence which is called viability in the terminology of population genetics. This essential aspect in the interpretation of selection is not considered in conventional GA selection. The classical (μ, λ) Evolution Strategy [16] in contrast does indeed: Reconsidering the basic functioning of a (μ, λ) ES in terms of selection, μ parents produce λ ($\lambda \geq \mu$) offspring from which the best μ are selected as members of the next generation. In contrast to GAs where selection pressure is widely predetermined by the choice of the mating scheme and the replacement strategy, ES allow an easy steering of selection pressure by the ratio between μ and λ . Offspring selection picks up this basic idea of ES and transforms these concepts for GAs in order to have an adjustable selection pressure (independent of the mating scheme and replacement strategy) at one's disposal.

Our interpretation of GA-selection similar to the concepts of a (μ, λ) Evolution Strategy is to generate an intermediate population (which is called virtual population in our notation) of size $|POP| \cdot T$ ($T \geq 1$) by sexual selection, crossover and mutation from the actual population of size $|POP|$. Then, similar to the interpretation of ES-selection, the best $|POP|$ members from the virtual population are chosen as members of the real next generation that contains the genetic information for the evolutionary process yet to come. The remaining $(1 - T) \cdot |POP|$ candidates can be seen as individuals that do not reach the age of sexual maturity. A practical problem in the technical appliance

of this technique is that it does not contain any indicator about the effectiveness of actual genetic search, the effectiveness of the operators actually used, etc. In other words, there is no information about the amount of selection pressure to be employed at a certain stage of genetic search. The aim is on the one hand to provide enough selection pressure for not losing essential building block information; on the other hand, too much selection pressure causes unwanted premature convergence to a suboptimal solution. Even if this concept of selection pressure steering has already proven to be very powerful in terms of stability and global solution quality [2], [2], it is a time consuming task to find an advantageous steering of (T) that requires an experienced user. These considerations already highly indicate the need for some kind of self-adaptation. The essential question is how to introduce self-adaptation into the GA-selection process in a generic way without losing the Markov property, which is an essential and general feature of GAs [18], [27]. The approach which we have developed for this reason will be described in the following subsection.

The basic concept of the new model for self-adaptive selection pressure steering (cf. Section 6) is also abutted on the basic selection model of population genetics: The parent population generates a certain number of offspring (greater than the size of the parent population) which do not all survive until they reach the age of sexual maturity, i.e. not all generated offspring become members of the mating pool of the ongoing evolutionary process. Self adaptive selection pressure steering is thereby realized in a quite natural and intuitive way: The number of generated offspring (potential members of the next generation) depends on how easy or difficult it is for the algorithm in different phases of evolution to achieve improvements.

A quantitative measure of actual selection pressure s is given by the ratio of generated and evaluated individuals to the population size s :

$$s = \frac{|POP|}{|Individuals \text{ actually generated}|}$$

where a small value of s denotes high selection pressure and higher s indicates smaller selection pressure. In order to keep the mechanism for self adaptation generic (without using any problem- or solution space-specific knowledge or local search) we have decided to use the comparison of the fitness value of the evenly generated offspring and the fitness values of its own parents, as the basis for the decision whether or not the offspring is selected as a member of the next generation. Offspring that do not fulfill the criterion are not selected for the mating pool and can be interpreted as individuals that do not survive until they reach the age of sexual maturity, i.e. these unsuccessful individuals do not contain information which is relevant for the GA. Even if this procedure shows some similarities to Rechenberg's $\frac{1}{5}$ success rule for Evolution Strategies, it implicates a number of essential advantages especially for Genetic Algorithms:

- In contrast to natural selection, it very often occurs that when applying a specific GA crossover operator the offspring generated contain unwanted mutations. Unwanted mutations in a crossover result are newly emerged alleles which do not occur in any of the two parents. Very complex crossover operators (like the ERX [19] or the EERX [17] operator for the path rep-

resentation of the TSP) have been especially developed with the goal to keep the relative amount of unwanted mutations as low as possible. Particularly when applying sexual selection, as it is done in case of conventional GAs, this aspect is of major importance as genetic search is mainly directed by recombining the genetic properties of above average parents. As a consequence a high ratio of unwanted mutations implies that the demolition of essential building blocks during crossover counteracts the directedness of genetic search. Our self adaptive selection concept has already proven its potential especially in the case of proven disadvantageous crossover operators. For the path representation of the TSP, for example, all considered operators were able to achieve results in the range of the best operators when being equipped with the self adaptive selection model (cf. [4] and [6]). The reasoning for this behavior is rather simple: When using disadvantageous operators, selection pressure is self adaptively adjusted to a higher value, and more individuals are generated and evaluated - but in the end only advantageous crossover results become members of the next generation. Particularly in the development of new applications, where new representations and operators have to be introduced, this stabilization is essential because the theoretical properties of newly introduced operators are unknown in the majority of cases.

- The double selection of the same individual, or the reproduction of two equal individuals causes the same offspring in case of GAs and, is one of the reasons for rapid loss of diversity and therefore of premature convergence. When operating with the new selection principle these offspring are not accepted as they do not surpass the fitness value of any of the two parents.
- If more than one crossover operator is available for a certain problem representation, it is beneficial and recommendable to apply all contemplable crossover operators in parallel with the same probability of coming into operation. In this way all operators (even the worse) propagate only their advantageous crossover results into the next generation and as different crossover mechanisms tend to generate different offspring, this strategy increases the broadness of genetic search without counteracting the goal directedness, which also helps to retard premature convergence.
- The self adaptive selection mechanism which is closely related to the basic selection model of population genetics, allows almost complete independence of the sexual selection step for partner selection: Concretely it could be verified in various test series that the achievable solution quality is almost the same - independent of the mating scheme that is used in the first selection step (roulette-wheel, linear-rank, tournament). Moreover, it could be demonstrated that solution qualities in the same range could be achieved even under random partner selection with comparable computational cost (less selection pressure - more iterations).
- Premature convergence can be detected immediately after its occurrence: If it is no longer possible to generate a sufficient number of offspring that are able to

outperform their parents (with an upper limit for selection pressure) it becomes obvious that premature convergence has occurred which gives a precise automatic termination criterion. Particularly for the distributed GA variants, this criterion is very important as it indicates when new genetic information has to be inserted into a subpopulation by migration from other subpopulations, in order to restimulate genetic search by combining smaller building blocks that have evolved in different regions of the search space to larger building blocks.

6. OFFSPRING SELECTION: AN ALTERNATIVE SELECTION PARADIGM FOR GENETIC ALGORITHMS AND GENETIC PROGRAMMING

As a logical consequence of these theoretical research achievements, our main goal in the development of advanced algorithmic aspects within GA theory was to find concepts for self-adaptive selection pressure steering. Self-adaption comes into play when considering the question which amount of offspring has to be created at each round, and which of these candidates are to be selected as members of the next generation, i.e. for the ongoing evolutionary process. In order to keep the concepts generic, no problem specific information about the solution space is allowed to be used for stating the self-adaptive model. Thus, it is desirable to systematically utilize just the fitness information of the individuals of the actual generation for building up the next generation of individuals, in order to keep the new concepts and methods generic. In principle, the new selection strategy acts in the following way:

The first selection step chooses the parents for crossover either randomly or in any well-known way of Genetic Algorithms like roulette-wheel, linear-rank, or some kind of tournament selection strategy. After having performed crossover and mutation with the selected parents, we introduce a further selection mechanism that considers the success of the apparently applied reproduction. In order to assure that the proceeding of genetic search occurs mainly with successful offspring, this is done in a way so that the used crossover and mutation operators are able to create a sufficient number of children that surpass their parents' fitness. Therefore, a new parameter, called success ratio ($SuccRatio \in [0, 1]$), is introduced. The success ratio gives the quotient of the next population members that have to be generated by successful mating in relation to the total population size. Our adaptation of Rechenberg's success rule [14] for Genetic Algorithms says that a child is successful if its fitness is better than the fitness of its parents, whereby the meaning of 'better' has to be explained in more detail: is a child better than its parents, if it surpasses the fitness of the weaker, the better, or is it in fact some kind of mean value of both?

For this problem, we claim that an offspring only has to surpass the fitness value of the worse parent in order to be considered as 'successful' at the beginning, while as evolution proceeds the child has to be better than a fitness value continuously increasing between the fitness of the weaker and the better parent. As in the case of Simulated Annealing, this strategy gives a broader search at the beginning, whereas at the end of the search process this operator acts in a more and more directed way. Having filled up the claimed

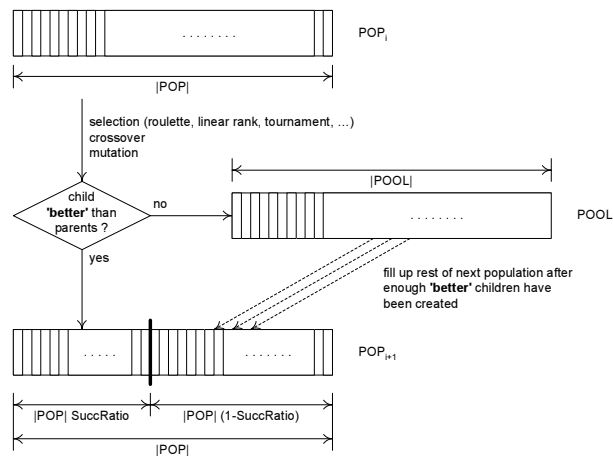


Figure 1: Flowchart for embedding the new offspring selection principle into a Genetic Algorithm.

ratio ($SuccRatio$) of the next generation with successful individuals according to the above meaning, the rest of the next generation ($(1 - SuccRatio) \cdot |POP|$) is simply filled up with individuals randomly chosen from the pool of individuals that were also created by crossover but did not reach the success criterion. The actual selection pressure $ActSelPress$ at the end of a single generation is defined by the quotient of individuals that had to be considered until the success ratio was reached, and the number of individuals in the population in the following way:

$$ActSelPress = \frac{|POP_{i+1}| + |POOL|}{|POP|}$$

Fig. 1 shows the operating sequence of the above described concepts.

With an upper limit of selection pressure ($MaxSelPress$) defining the maximum number of children considered for the next generation (as a multiple of the actual population size) that may be produced in order to fulfill the success ratio, this new model also functions as a precise detector of premature convergence:

This upper limit of selection pressure ($MaxSelPress$) defines the maximum number of offspring considered for the next generation (as a multiple of the actual population size) that may be produced in order to fulfill the success ratio. With a sufficiently high setting of ($MaxSelPress$), this new model also functions as a precise 'detector heuristics' for premature convergence:

If it is no longer possible to find a sufficient number ($SuccRatio \cdot |POP|$) of offspring outperforming their own parents even if ($MaxSelPress \cdot |POP|$) candidates have been generated, premature convergence has occurred.

As a basic principle of this selection model, a higher success ratio causes higher selection pressure. Nevertheless, higher settings of success ratio, and therefore of selection pressure, do not necessarily cause premature convergence. The reason for this is mainly, that the new selection step (after crossover) per definition does not accept clones that emanate from two identical parents. In conventional GAs such clones represent a major reason for premature convergence

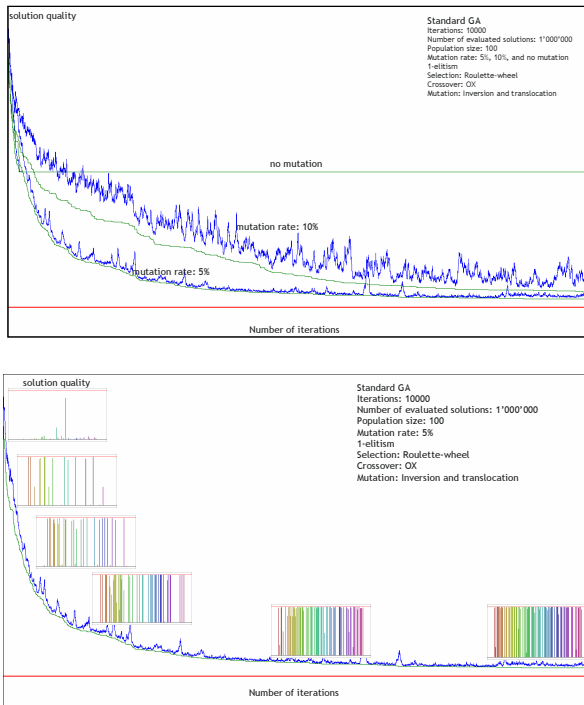


Figure 2: The effect of mutation for certain mutation rates (upper diagram) and the distribution of essential genetic information for a mutation rate of 5% (lower diagram) both in case of a standard GA for the ch130 benchmark TSP.

of the whole population around a suboptimal value, whereas the new selection step specifically counteracts against this phenomenon.

The following considerations aim to point out the importance of mutation for the recovery of essential genetic information in the case of conventional GAs in order to oppose these results with the results being achieved with offspring selection. By reasons of compactness, the results are mainly shown on the basis of diagrams and give only a brief description of introduced operators, parameter settings, and test environments. Furthermore, the chosen benchmark instance is of rather small dimension in order to allow the observation of essential alleles during the run of the algorithm.

The results displayed in Figure 2 (upper diagram) show the effect of mutation for reintroducing already lost genetic information. The horizontal line of the diagram shows the number of iterations and the vertical line stands for the solution quality. The bottom line indicates the global optimal solution which is known for this benchmark test case. The three curves of the diagram show the performance of a Genetic algorithm with no mutation, with a typical value of 5% mutation as well as a rather high mutation rate of 10%. For each of the three curves the lower line stands for the best solution of the actual population and the upper line shows the average fitness value of the population members. The results with no mutation are extremely weak and the quality curve stagnates very soon and far away from the global optimum. The best and average solution quality are the same and no further evolutionary process is possible - premature conver-

gence has occurred. As already stated before, mutation is a very essential feature of standard GAs in order to avoid premature convergence. But also a rather high mutation rate of 10% produces results which are not very satisfying and indeed the best results are achieved with a mutation rate which is very typical for GA applications - namely a mutation rate of 5%. Considering a standard benchmark problem like the ch130 (a 130 city TSP taken from the TSPLib [15]) with one single best solution allows to consider the edges of the shortest path as the essential alleles whose preservation during the run can be observed. The following figures indicate the spreading of essential alleles during the runs of the certain algorithms. This is visualized by inserting bar charts which have to be considered as snapshots after a certain number of iterations approximately corresponding to the position in the figure. The higher a certain bar (130 bars for a 130-city TSP) the higher the relative occurrence of the corresponding essential allele in the population.

The lower diagram of Figure 2 shows the distribution of essential alleles over the iterations for a standard GA with a mutation rate of 5%. The interesting thing is that some minor ratio of essential alleles is rapidly fixed in the population and the majority of essential alleles which are still missing have disappeared in the entire population. During the further run of the algorithm it is only mutation which can reintroduce this essential genetic information. As it could be seen in Figure 2, without mutation premature convergence would already have occurred at this early state of evolutionary search. But with an appropriate mutation rate (5% in this example) more and more essential alleles are discovered ending up with quite a good solution. But there is still a gap to the global optimum caused by that alleles which could not be recovered due to mutation. The next figures will show how the new selection concept is able to close this gap and make the algorithm much more independent of mutation.

So let us take a closer look at the distribution of essential genetic information in the population when using the enhanced selection concepts. The upper diagram of Figure 3 shows the quality curve and the distribution of essential alleles for a mutation rate of 5% (which was able to achieve the best results in case of a standard GA).

When applying the GA with the new selection principle to the same benchmark test case one can see that the global optimal solution is detected in only about 100 iterations. Nevertheless, the computational effort is comparable to the standard GA as much more individuals have to be evaluated at each iteration step due to the higher selection pressure. Considering the distribution of essential alleles we see a totally different situation. Almost no essential alleles get lost and the ratio of essential alleles continuously increases in order to end up with a final population that contains almost all pieces of essential genetic information and therefore achieving a very good solution. This shows that the essential alleles are preserved much more effectively and indicates that the influence of mutation should be much less. But is this really the case? In order to answer this question, let us consider the same example with the same settings - only without mutation. And indeed the assumption holds and also without mutation the algorithm finds a solution which is very close to the global optimum (see lower diagram of Figure 3). The essential alleles interfuse the population more and more and almost all of them are members of the final population. Reconsidering the standard GA without mutation

the algorithm was prematurely converging very soon with a very bad total quality.

7. RESULTS

Population genetics inspired further developed algorithmic concepts, as shown in the last sections, have been applied to several fields of application.

Firstly, offspring selection in its parallel formulation (SASEGASA) has been applied to several well known combinatorial optimization problems like the Traveling Salesman Problem (TSP), Job Shop Scheduling Problem (JSSP), Capacitated Vehicle Routing Problem (CVRPTW) and also to the optimization of various real-valued test functions which have been tackled in various dimensions. It has been shown empirically in [6], [3], [5], [8] that for combinatorial optimization problems it is possible to use the concept of multiple subpopulations in order to scale the achievable results in terms of achievable solution quality. Also for the analyzed real-valued test functions (Ackley, Schwefel, Rastrigin, Rosenbrock, Griewangk) highest quality results could be achieved just by integrating offspring into the general concept of a Genetic Algorithm with well-established operators for real-valued codification [6], [3], [7].

Very recently, offspring selection has been able to proof its high ability in the field of Genetic Programming based structure identification. In that context, further developed concepts of Genetic Programming with offspring selection as a key feature has been applied to classification, regression and time-series. We have shown in several articles this method's ability to efficiently and robustly detect highly nonlinear intercorrelations in benchmark as well as real-world datasets steaming from the fields of mechatronics, business-and financial engineering as well as from the field of medical data mining. In [20], [22], [21], [24], [23], [26] and [25] further algorithmic developments and results are shown. It has been shown that the use of these advanced developments yields results that outperform the results achievable by conventional Genetic Programming concepts as well as results achieved by techniques from the field of applied statistics and machine learning.

8. CONCLUSION

Some selected aspects of population genetics have been pointed out and possibilities how to integrate these aspects into the general concept of a Genetic Algorithm. Special attention has been given to those aspects that are independent of problem representations and therefore independent of the problem at hand. The realization of offspring selection with birth surplus and its integration in the algorithmic flow of a GA or GP has been shown. References to recent publications have been given in order to demonstrate the benefits of these further developed algorithmic variants.

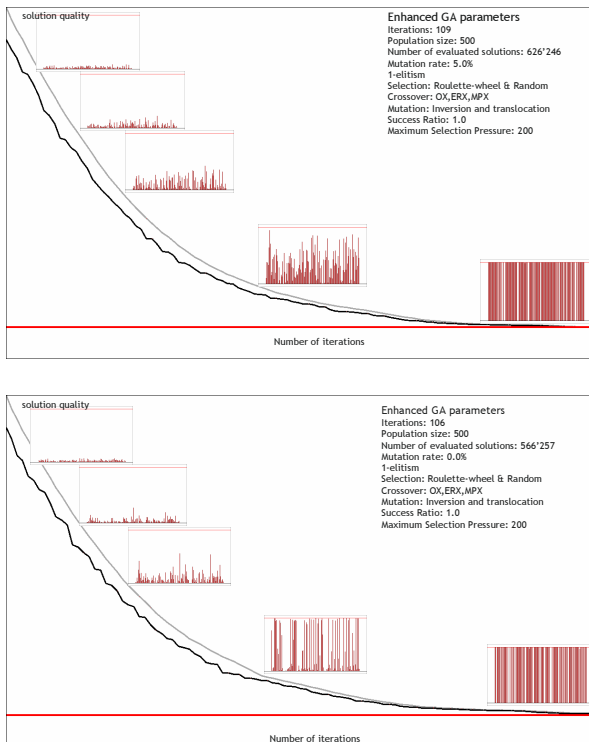


Figure 3: The distribution of essential genetic information when using the enhanced selection concept considering the ch130 benchmark TSP with 5% mutation (upper diagram) and with no mutation (lower diagram).

9. REFERENCES

- [1] M. Affenzeller. Segregative genetic algorithms (SEGA): A hybrid superstructure upwards compatible to genetic algorithms for retarding premature convergence. *International Journal of Computers, Systems and Signals (IJCSS)*, 2(1):18–32, 2001.
- [2] M. Affenzeller. A generic evolutionary computation approach based upon genetic algorithms and evolution strategies. *Journal of Systems Science*, 28(2):59–72, 2002.
- [3] M. Affenzeller. *Population Genetics and Evolutionary Computation: Theoretical and Practical Aspects*. Trauner Verlag Linz, 2005.
- [4] M. Affenzeller and S. Wagner. A self-adaptive model for selective pressure handling within the theory of genetic algorithms. *Computer Aided Systems Theory: EUROCAST 2003, Lecture Notes in Computer Science*, 2809:384–393, 2003.
- [5] M. Affenzeller and S. Wagner. The influence of population genetics for the redesign of genetic algorithms. *Journal of Systems Science*, 30(4), 2004.
- [6] M. Affenzeller and S. Wagner. SASEGASA: A new generic parallel evolutionary algorithm for achieving highest quality results. In: *Journal of Heuristics, Special Issue on New Advances on Parallel Meta-Heuristics for Complex Problems*, 10(3):239–263, 2004.
- [7] M. Affenzeller and S. Wagner. Offspring selection: A new self-adaptive selection scheme for genetic algorithms. *Adaptive and Natural Computing Algorithms, Springer Computer Science*, pages 218–221, 2005.
- [8] M. Affenzeller, S. Wagner, and S. Winkler. GA-selection revisited from an ES-driven point of view. *Intelligence and Knowledge Engineering Applications: A Bioinspired Approach, Lecture Notes in Computer Science*, 3562:262–271, 2004.
- [9] E. Cantu-Paz and D. E. Goldberg. On the scalability of parallel genetic algorithms. *Evolutionary Computation*, 7(4):429–449, 1999.
- [10] C. Darwin. *The Origin of Species By Means of Natural Selection or the Preservation of Favoured Races in the Struggle for Life*. Penguin Books London, 1859/1985.
- [11] A. J. F. Griffiths, J. H. Miller, D. T. Suzuki, R. C. Lewontin, and W. M. Gelbart. *An Introduction to Genetic Analysis*. W H Freeman Co., 1989.
- [12] P. W. Hendrick. *Genetics of Populations, Second Edition*. Jones and Bartlett Publishers, 2000.
- [13] J. Lamarck. *Zoological Philosophy*. 1914.
- [14] I. Rechenberg. *Evolutionstrategie*. Friedrich Frommann Verlag, 1973.
- [15] G. Reinelt. TSPLIB - A traveling salesman problem library. *ORSA Journal on Computing* 3, 3:376–384, 1991.
- [16] H. Schwefel. *Numerische Optimierung von Computer-Modellen mittels Evolutionstrategie*. Birkhäuser Verlag. Basel, 1994.
- [17] T. Starkweather, S. McDaniel, K. Mathias, D. Whitley, and C. Whitley. A comparison of genetic scheduling operators. *Proceedings of the Fourth International Conference on Genetic Algorithms*, pages 69–76, 1991.
- [18] J. Suzuki. A markov chain analysis on simple genetic algorithms. *IEEE Trans. on Systems, Man, and Cybernetics*, 25(4):655–659, 1995.
- [19] D. Whitley, T. Starkweather, and D. Fuguay. Scheduling problems and traveling salesman: The genetic edge recombination operator. *Proceedings of the Third International Conference on Genetic Algorithms and their Application*, pages 133–140, 1989.
- [20] S. Winkler, M. Affenzeller, and S. Wagner. New methods for the identification of nonlinear model structures based upon genetic programming techniques. *Journal of Systems Science*, 31:5–13, 2005.
- [21] S. Winkler, M. Affenzeller, and S. Wagner. Advances in applying genetic programming to machine learning, focussing on classification problems. *Proceedings of the 20th IEEE International Parallel & Distributed Processing Symposium (IPDPS 2006)*, 2006.
- [22] S. Winkler, M. Affenzeller, and S. Wagner. Automatic data based patient classification using genetic programming. *Cybernetics and Systems 2006*, 1:251–256, 2006.
- [23] S. Winkler, M. Affenzeller, and S. Wagner. Heuristicmodeler: A multi-purpose evolutionary machine learning algorithm and its applications in medical data analysis. *Proceedings of the International Mediterranean Modelling Multiconference I3M 2006*, pages 629–634, 2006.
- [24] S. Winkler, M. Affenzeller, and S. Wagner. Using enhanced genetic programming techniques for evolving classifiers in the context of medical diagnosis - an empirical study. *Proceedings of the GECCO 2006 Workshop on Medical Applications of Genetic and Evolutionary Computation (MedGEC 2006)*, 2006.
- [25] S. Winkler, M. Affenzeller, and S. Wagner. Advanced genetic programming based machine learning. *Journal of Mathematical Modelling and Algorithms*, 2007.
- [26] S. Winkler, H. Efendic, M. Affenzeller, L. D. Re, and S. Wagner. On-line modeling based on genetic programming. *International Journal on Intelligent Systems Technologies and Applications*, 2:255–270, 2007.
- [27] A. H. Wright and Y. Zhao. Markov chain models of genetic algorithms. *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 734–741, 1999.
- [28] S. Wright. Isolation by distance. *Genetics*, 28:114–138, 1943.