EFFECTIVE ALLELE PRESERVATION BY OFFSPRING SELECTION: AN EMPIRICAL STUDY FOR THE TSP

Michael Affenzeller[(a)], Stefan Wagner[(b)], Stephan Winkler[(c)]

[(a),(b),(c)] Upper Austrian University of Applied Sciences,
School of Informatics, Communications and Media
Heuristic and Evolutionary Algorithms Laboratory
Softwarepark 11, 4232 Hagenberg, Austria

[(a)]michael.affenzeller@heuristiclab.com, [(b)]stefan.wagner@heuristiclab.com, [(c)]stephan.winkler@heuristiclab.com

ABSTRACT
The basic selection ideas of the different representatives of evolutionary algorithms are sometimes quite diverse. The selection concept of genetic algorithms (GAs) and genetic programming (GP) is basically realized by the selection of above-average parents for reproduction whereas evolution strategies (ES) use the fitness of newly evolved offspring as the basis for selection (survival of the fittest due to birth surplus). This contribution considers aspects of population genetics and Evolution Strategies in order to propose an enhanced and generic selection model for Genetic Algorithms which is able to preserve the alleles which are part of a high quality solution. Some selected aspects of these enhanced techniques are discussed exemplarily on the basis of travelling salesman benchmark (TSP) benchmark problem instances.

Keywords: softcomputing, evolutionary computation, selection, self adaptation

1. INTRODUCTION

As some kind of approximation for the gradient information which is not available for many optimization problems, neighborhood search aims to obtain information about the descent/increase of the objective function in the local neighborhood of a certain point. Conventional neighborhood search starts from an arbitrary point in the search space and iteratively moves to more and more promising points along a given neighborhood structure (w.r.t. the objective function) as long as no better solution can be detected in the local neighborhood.

The self-evident drawback of this method is that for more complex functions the algorithm converges and gets stuck in the next attracting local optimum which may be far away of a global optimum. It is a common feature of all methods based upon neighborhood search to counteract this essential handicap.

Simulated annealing (SA) for example also allows moves to worse neighborhood solutions with a certain probability which decreases as the search process progresses in order to scan the solution space broader at the beginning and to become more and more goal-oriented at the end. Tabu search on the other hand introduces some kind of memory in terms of a so-called tabu list which stores moves that are considered to lead to already visited areas of the search space. However, also evolution Strategies (ES), a well-known representative of Evolutionary Computation, have to be considered as some kind of parallel neighborhood search as asexual mutation (a local operator) is the only way to create new individuals (solution candidates) in the classical ES-versions. Therefore, in the case of multimodal search spaces, global optima are detected by an ES only if one of the starting values is located in the absorbing region (attracting basin) of a global optimum.

By considering recombination (crossover) as their main operator, GAs and also GP take a basically different approach compared to neighborhood-based techniques as recombination is a sexual operator, i.e. properties of individuals from different regions of the search space are combined in new individuals. Therefore, the advantage of applying GAs to hard problems of combinatorial optimization lies in their ability to scan broader regions of the solution space than heuristic methods based upon neighborhood search do. Nevertheless, also GAs are frequently faced with a problem which, at least in its impact, is quite similar to the problem of stagnating in a local but not global optimum. This drawback, called premature convergence in the terminology of GAs, occurs if the population of a Genetic Algorithm reaches such a suboptimal state that the genetic operators are no longer able to produce offspring that are able to outperform their parents (Fogel 1996; Affenzeller 2005). This happens if the genetic information stored in the individuals of a population does not contain that genetic information which would be necessary to further improve the solution quality. Therefore, in contrast to the present contribution, the topic of premature convergence is considered to be closely related to the loss of genetic variation in the entire population in GA-research.

In this contribution we do not identify the reasons for premature convergence in the loss of genetic

variation in general but more specifically in the loss of what we say essential genetic information, i.e. in the loss of alleles which are part of a global optimal solution. Therefore, we will denote the genetic information of the global optimal solution (which is usually unknown a priori) as essential genetic information in the following.

But what are the reasons for premature convergence, or in other words what are the reasons that this essential genetic information is not or no more available:

- Firstly, one reason for this loss of essential genetic information may be that these alleles are simply not represented in the initial population of the Genetic Algorithm.
- Then, especially in the earlier phase of genetic search it frequently happens that essential genetic information is hidden in individuals with bad total fitness and is therefore eliminated due to selection.
- Furthermore, for the majority of GA applications it is absolutely not guaranteed that the applied crossover operators are able to create new children in a way that the newly evolving child contains exactly the genetic information of its own parents. If this is not guaranteed this fact represents a further reason for a genetic algorithm to loose essential genetic information and therefore cause premature convergence.

The main measure in conventional GA-theory to counteract against this phenomenon is mutation (and migration in the parallel variants) and indeed - as will be shown in the empirical part of the paper - this works quite well and a lot of already lost essential genetic information can be recovered by mutation.

The main aim of the present work is to discuss, analyze and improve new generic theoretical concepts for avoiding or at least retarding premature convergence in a non-problem-specific way by taking the above stated considerations into account:

A very essential question about the general performance of a GA is, whether or not good parents are able to produce children of comparable or even better fitness (the building block hypothesis implicitly relies on this). In natural evolution, this is almost always true. For Genetic Algorithms this property is not so easy to guarantee. The disillusioning fact is that the user has to take care of an appropriate coding in order to make this fundamental property hold. In order to somehow overcome this strong requirement we try to get to the bottom of reasons for premature convergence from a technical as well as from a population genetics inspired point of view and draw some essential interconnections.

The basic idea of the new selection model, introduced as offspring selection (Affenzeller and Wagner 2004a) is to consider not only the fitness of the parents in order to produce a child for the ongoing evolutionary process. Additionally, the fitness value of the evenly produced child is compared with the fitness values of its own parents. The child is accepted definitely as a candidate for the further evolutionary process if and only if the reproduction operator was able to produce a child that could outperform the fitness of its own parents. This strategy guarantees that evolution is presumed mainly with crossover results that were able to mix the properties of their parents in an advantageous way. I.e. **survival of the fittest alleles is rather supported than survival of the fittest individuals** which is a very essential aspect for the preservation of essential genetic information stored in many individuals (which may not be the fittest in the sense of individual fitness).

The experimental part analyzes the characteristics of offspring selection on the basis of some rather small TSP benchmark problems: As commonly done when evaluating the capability of heuristic techniques, some main features are analyzed separately.

2. SELECTION

In terms of goal orientedness, selection is the driving force of GAs. In contrast to crossover and mutation, selection is completely generic, i.e. independent of the actually employed problem and its representation. A fitness function assigns a score to each individual in a population that indicates the 'quality' of the solution the individual represents. The fitness function is often given as part of the problem description or based upon the objective function.

In the Standard Genetic Algorithm the probability that a chromosome in the current population is selected for reproduction is proportional to its fitness (roulette wheel selection). However, there are also many other ways of accomplishing selection. These include for example linear-rank selection or tournament selection (Michalewicz 1996; Schöneburg, Heinzmann and Feddersen, 1994).

However, all evenly mentioned GA-selection principles have one thing in common:

They all just consider the aspect of sexual selection, i.e. mechanisms of selection only come into play for the selection of parents for reproduction. The enhanced selection model which will be described in the following section defies this limitation by considering selection in a more general sense.

2.1. Parent Selection vs. Offspring Selection

In the following we describe and aim to bring together technical as well as biologically motivated considerations in order to motivate the concepts proposed in this paper. The following listing itemizes the most essential aspect in the development phase of the new methods in the above mentioned sense

2.1.1. Selection and Selection Pressure

In the theory of Genetic Algorithms selection and selection pressure are predetermined by the so-called mating scheme and by the replacement strategy actually deployed. By that 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. This classical GA selection concept is known as sexual selection in the terminology of population genetics. In the population genetics view, sexual selection covers only a rather small aspect of selection which appears when individuals have to compete to attract mates for reproduction. The population genetics basic selection model basically considers the selection process in the following way:

Random mating → selection → random mating → selection → …….

I.e. selection is considered to depend mainly on the probability of surviving of newborn individuals until they reach pubescence which is called viability in the terminology of population genetics.

The essential aspect of offspring selection in the interpretation of selection is rarely considered in conventional GA selection. The classical (μ, λ) Evolution Strategy in contrast does this very well: Reconsidering the basic functioning of a (μ, λ)-ES in terms of selection μ parents produce λ ($\mu \leq \lambda$) offspring from which the best μ are selected as members of the next generation. In contrast to GAs where selection pressure is predetermined by the choice of the mating scheme and the replacement strategy, ES allow an easy steering of selection pressure by the ratio between λ and μ. The selection pressure steering model introduced in (Affenzeller 2001) and further developed in (Affenzeller and Wagner 2004) 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 advanced selection scheme allowing self-adaptive steering of selection pressure aims to transform the basic ideas for improving the performance of GAs. In doing so 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 (selection before reproduction) as in case of a standard GA does not drastically effect the qualitative or quantitative performance of the algorithm if being equipped with the newly defined offspring selection step (selection after reproduction). Even with random sexual selection (corresponding to the basic model of the population genetic's selection model) the results are 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 from the population
- The allele may converge to an equilibrium state
- No change in allele frequency

The basic approaches for retarding premature convergence discussed in GA literature aim to maintain genetic diversity. The most common techniques for this purpose are based upon preselection (Cavicchio1970), crowding (DeJong 1975), or fitness-sharing (Goldberg 1989). The main idea of these techniques is to maintain genetic diversity by the preferred replacement of similar individuals (Cavicchio1970), (DeJong 1975) or by the fitness-sharing of individuals which are located in densely populated regions (Goldberg 1989). While methods based upon crowding (DeJong 1975) or fitness sharing (Goldberg 1989) require some kind of neighborhood measure depending on the problem representation, (Goldberg 1989) is additionally quite restricted to proportional selection. Moreover, these techniques have the common goal to maintain genetic diversity which is very important in natural evolution where a rich gene pool is the guarantor in terms of adaptiveness w.r.t. changing environmental conditions.

In case of artificial genetic search as being performed by a Genetic Algorithm the optimization goal does not change during the run of a GA and the fixing of alleles of high quality solutions is desirable in the same manner as the erasement of alleles which are definitely not part of a good solution in order to reduce the search space and make genetic search more goal-oriented. I.e. we claim that pure diversity maintenance mechanisms as commonly proposed in GA theory do not support goal-oriented genetic search w.r.t the locating of global optimal solutions.

2.1.2. Adjustable Selection Pressure

One interpretation of GA-selection similar to the concepts of a (μ, λ)-ES is to generate an intermediate population (what is called virtual population in our notation) of size $|POP| \cdot T$ (where $T \geq 1$) 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 $(1-T)\ |POP|$ 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 actually used operators, etc.

I.e. 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 may support 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 (Affenzeller 2001; Affenzeller 2002) 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 i.e. non problem specific way. The approach which we have developed for this reason will be described in the following.

3. OFFSPRING SELECTION

The basic idea to create and evaluate a certain amount (greater or equal population size) of offspring, to be considered for future members of the next generation, is adapted from Evolution Strategies. Self-adaption comes into play when considering the question which amount of offspring is necessary 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 the well-known way of Genetic Algorithms by proportional, 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 the proceeding of genetic search mainly with successful offspring in that way that the used crossover and mutation operators were able to create a child that surpasses its 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 (Rechenberg 1973) 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 have decided to introduce a cooling strategy similar to Simulated Annealing. Following the basic principle of Simulated Annealing 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 and while evolution proceeds the child has to be better than a fitness value continuously increasing between the fitness of the weaker and the better parent. Like in the case of Simulated Annealing, this strategy effects 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 ratio (*SuccRatio*) of the next generation with successful individuals in the above meaning, the rest of the next generation $((1\text{-}SuccRatio)\cdot|POP|)$ 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:

$$ActSelPress = \frac{|POP|SuccRatio + |POOL|}{|POP|}$$

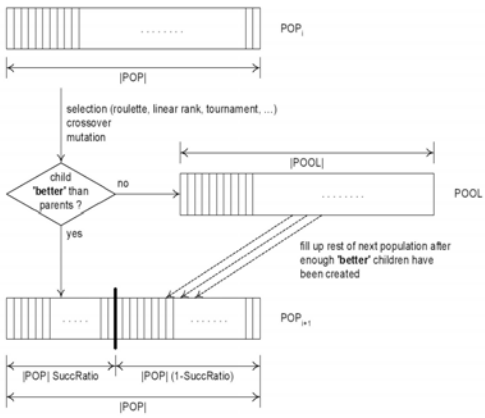

Figure 1: Flowchart for Embedding the new Selection Principle (Offspring Selection) into a Genetic Algorithm

Figure 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:

> *If it is no longer possible to find a sufficient number of $SuccRatio\cdot|POP|$ 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 as the preservation of fitter alleles is additionally supported and not only the preservation of fitter individuals.

Also it becomes possible within this model to state selection pressure in a very intuitive way that is quite similar to the notation of selection pressure in Evolution Strategies. Concretely, we define the actual selection pressure as the ratio of individuals that had to be generated in order to fulfill the success ratio to the population size.

For example, if we work with a population size of say *100* and it would be necessary to generate *2000* individuals in order to fulfill the success ratio, the actual selection pressure would have a value of *20*. Via these means we are in a position to attack several reasons for premature convergence as illustrated in the following sections. Furthermore, this strategy has proven to act as a precise mechanism for self-adaptive selection pressure steering, which is of major importance in the migration phases of parallel evolutionary algorithms. The aspects of offspring selection w.r.t. parallel GAs are combined in the parallel SASEGASA-algorithm (Affenzeller and Wagner 2003; Affenzeller and Wagner 2004a).

4. EMPIRICAL DISCUSSION

The empirical section is subdivided into three parts:
The first subsection aims to highlight the main message of the paper (preservation of essential alleles) whereas the second subsection aims to touch on further subjects concerning the effects of self-adaptive offspring selection (Affenzeller 2004b). As the scope of the present work does not allow a deeper and more sophisticated analysis of different problem situations, the third part of the experimental discussion gives some references to related contributions which include a more detailed and statistically more relevant experimental discussion on the basis of several benchmark but also practical problems on which we have applied the new selection model recently.

4.1. Preservation of Essential Genetic Information

This subsection aims 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 the enhanced selection model discussed in this paper. 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 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 – every allele is fixed and premature convergence 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 in a quite typical range for a lot of GA applications - namely a mutation rate of 5%.

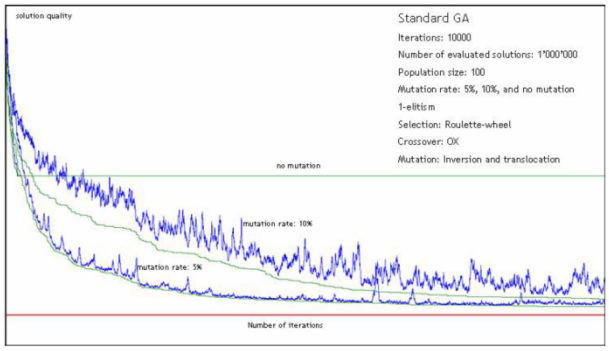

Figure 2: The Effect of Mutation in Case of a Standard GA for the ch130 benchmark TSP.

Considering a rather low dimensional standard benchmark problem like the ch130 with a unique and well known global optimal solution (a 130 city TSP taken from the TSPLib; (Reinelt 1991)) 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.

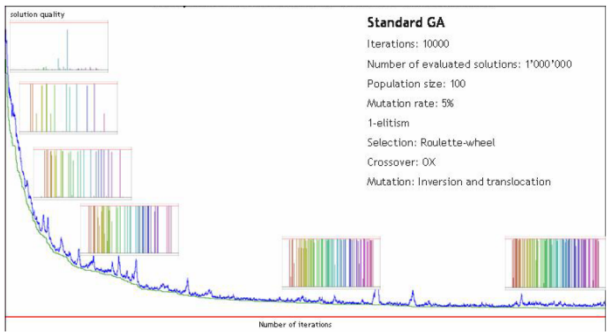

Figure 3: The Distribution of Essential Genetic Information in case of a Standard GA for the ch130 benchmark TSP

Figure 3 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 those 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 next curve (Figure 4) shows the quality curve and the distribution of essential alleles for 5% mutation which was able to achieve the best results in case of a standard GA.

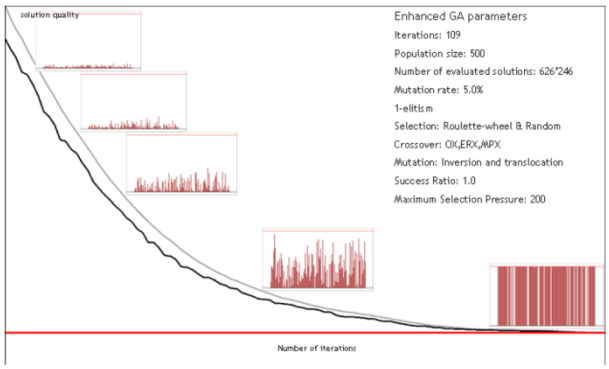

Figure 4: The Distribution of Essential Genetic Information when using the Enhanced Selection Concept considering the ch130 benchmark TSP

When applying the GA with the new offspring 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 - just without mutation.

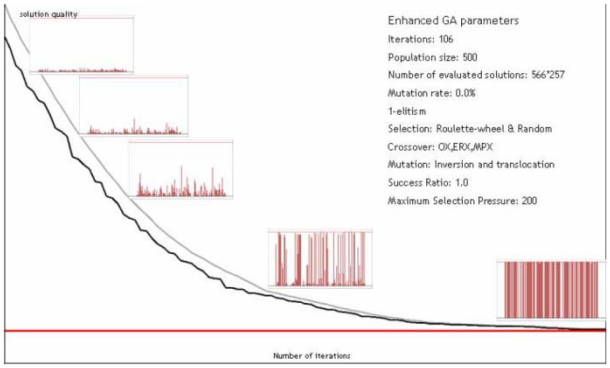

Figure 5: The Distribution of Essential Genetic Information when using Offspring Selection without Mutation

And indeed the assumption holds and also without mutation the algorithm finds a solution which is very close to the global optimum. 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.

4.2. Experimental Studies of Some More Aspects on the Basis of some TSP instances

The analysis of producible results considering various algorithms and benchmark test cases still denotes the most commonly used and possibly also the most objective way to analyze the potential of heuristic optimization techniques.

In our experiments, all computations are performed on Pentium 4 PCs with 2 GB RAM under Windows XP. The environment in which the algorithms are implemented and tested is HeuristicLab[1] (Wagner and Affenzeller 2005). For the tests shown here we have selected some rather small instances of the Travelling Salesman Problem (TSP) taken from the TSPLIB (Reinelt 1991). Reference to more sophisticated tests also on higher dimensional TSP instances are given in the next subsection. In all experiments, the results are

[1] More details can be found on the HeuristicLab homepage http://www.heuristiclab.com

represented as the relative difference to the best known solution defined as

$$relative\ Difference = \left(\frac{Result}{Optimal} - 1\right).$$

However, it is to be pointed out once again that the newly introduced methods are by no means restricted or somehow optimized to routing problems like the TSP. Similar effects could also be observed for other combinatorial optimization problems and especially for genetic programming applications. References are given in subsection 4.3.

In order to not dilute the effects of the different aspects too much, some selected aspects are pointed out separately in this section.

The first aspect to be considered is the effect of the enhanced selection model (offspring selection) to the quality improvement of different crossover operators. To visualize the positive effects of the new methods in a more obvious way we also present results that were generated by a classical GA with generational replacement and 1-elitism.

Remarkable in this context is the effect that also crossover operators that are considered as rather unsuitable for the TSP (Larranaga, Kuijpers, Murga and Dizdarevic 1999) achieve quite good results in combination with the new selection model. The reason for this behavior is given by the fact that in our selection principle only children that have emerged as a good combination of their parents' attributes are considered for the further evolutionary process; the success ratio levels off at a higher range. In combination with a higher upper value for the maximum selection pressure genetic search can therefore be guided advantageously also for poor crossover operators as the larger amount of handicapped offspring is simply not considered for the further evolutionary process.

Additionally to the already mentioned aspect that the enhanced selection mechanism is able to improve the performance of the certain crossover operators it is furthermore observable that the new self-adaptive selection model makes the performance of the GA almost independent of sexual selection in terms of qualitative performance.

Table 1: Parameter values used in the Test Runs of the several algorithms

Classical GA with several selection mechanisms (Tab.2, Tab. 3, Tab. 4)	
Generations	100.000
Population Size	120
Elitism Rate	1
Mutation Rate	0.05
Selection Operator	Roulette, Linear Rank, Random
Mutation Operator	Inversion & Translocation

GA with offspring selection in combination with several sexual selection mechanisms (Tab.5, Tab. 6, Tab. 7)	
Population Size	500
Elitism Rate	1
Mutation Rate	0.1
Selection Operator	Roulette, Linear Rank, Random
Mutation Operator	Inversion & Translocation
Success Ratio	0.7
Maximum Selection Pressure	250

In Tab. 2, 3, and 4 the results achieved with the conventional GA using either roulette-wheel (Tab. 2), linear-rank (Tab. 3) respectively random, i.e. no, (Tab. 4) selection are listed. On the other hand Tab. 5,6, and 7 show the results achieved with the enhanced self-adaptive selection concept using either roulette-wheel (Tab. 5), linear-rank (Tab. 6), or no (Tab. 7) sexual selection.

The fixed parameter values for all algorithms that were used in the different test runs of the present subsection are given in Tab. 1. All values presented in the following tables are the best resp. average values of twenty independent test runs executed for each test case.

Similar improvements of solution quality are also observable when comparing the GA using linear-rank selection with the enhanced GA using linear-ranking as the first selection step. So far the results underpin the crossover improvement not only for roulette-wheel selection but also for linear-ranking.

Especially notable is the comparison of Tab. 4 and Tab. 7. Firstly, it is barely remarkable that a GA with no (i.e. random) sexual selection is unable of producing high-quality results. So the results of Tab. 4 are in the region of random search which is caused by the 1-elitism (the only goal-oriented force under these settings). What is really remarkable is that the results of the new GA with enhanced selection and no sexual selection (Tab. 7) are about the same than the results obtained with roulette-wheel respectively with linear-rank as the first selection step. This observation supports the theory of population genetics that sexual selection really plays a rather inferior role in the natural selection process.

Table 2: Experimental Results achieved with the Classical GA using Roulette-Wheel Selection

Results for standard GA with proportional selection				
Problem	Crossover	Best	Av.	Eval. Sol.
berlin52	OX	0.00	3.76	12.000.000
	ERX	5.32	7.73	12.000.000
	MPX	21.74	26.52	12.000.000
ch130	OX	3.9	5.41	12.000.000
	ERX	142.57	142.62	12.000.000
	MPX	83.57	83.57	12.000.000
kroA200	OX	3.14	4.69	12.000.000
	ERX	325.92	336.19	12.000.000
	MPX	146.94	148.08	12.000.000

Table3: Experimental Results achieved with the Classical GA using Linear Rank Selection

Results for standard GA with linear rank selection				
Problem	Crossover	Best	Av.	Eval. Sol.
berlin52	OX	0.00	5.40	12.000.000
	ERX	2.52	4.58	12.000.000
	MPX	20.9	27.31	12.000.000
ch130	OX	5.60	8.88	12.000.000
	ERX	99.18	128.47	12.000.000
	MPX	85.78	97.46	12.000.000
kroA200	OX	8.58	12.24	12.000.000
	ERX	351.41	365.8	12.000.000
	MPX	144.25	150.34	12.000.000

Table 4: Experimental Results achieved with the Classical GA using Random Selection

Results for standard GA with proportional selection				
Problem	Crossover	Best	Av.	Eval. Sol.
berlin52	OX	25.07	31.85	12.000.000
	ERX	80.54	89.96	12.000.000
	MPX	52.24	78.52	12.000.000
ch130	OX	148.54	161.77	12.000.000
	ERX	397.46	406.94	12.000.000
	MPX	252.59	286.18	12.000.000
kroA200	OX	296.22	309.71	12.000.000
	ERX	667.71	692.22	12.000.000
	MPX	420.76	464.49	12.000.000

Table 5: Experimental Results achieved with Offspring Selection and Proportional Parent Selection

Results for offspring selection GA with proportional sexual selection				
Problem	Crossover	Best	Av.	Eval. Sol.
berlin52	OX	0.00	3.88	15.964.680
	ERX	0.00	3.10	16.337.700
	MPX	0.00	1.45	11.775.071
	OX, ERX, MPX	0.00	0.72	7.204.601
ch130	OX	3.88	5.40	15.602.824
	ERX	4.02	5.30	16.920.45
	MPX	1.83	3.53	13.994.68
	OX, ERX, MPX	0.00	2.71	7.702.818
kroA200	OX	2.25	5.72	10.814.980
	ERX	5.10	5.99	18.268.888
	MPX	5.21	7.65	12.296.581
	OX, ERX, MPX	0.00	2.78	6.647.256

Table 6: Experimental Results achieved with Offspring Selection and Linear Rank Parent Selection

Results for offspring selection GA with proportional sexual selection				
Problem	Crossover	Best	Av.	Eval. Sol.
berlin52	OX	2.29	4.94	7.448.762
	ERX	0.00	1.92	399.296
	MPX	0.00	3.92	8.199.592
	OX, ERX, MPX	0.00	1.59	60.920.006
ch130	OX	3.04	7.90	2.515.637
	ERX	4.36	5.36	1.245.727
	MPX	2.22	3.61	9.029.870
	OX, ERX, MPX	2.16	2.80	61.759.48
kroA200	OX	8.14	9.30	2.011.929
	ERX	6.28	8.12	4.822.588
	MPX	5.63	6.37	8.527.427
	OX, ERX, MPX	1.75	2.79	57.493.081

Table 7: Experimental results achieved with Offspring Oelection and Random Parent Selection

Results for offspring selection GA with proportional sexual selection				
Problem	Crossover	Best	Av.	Eval. Sol.
berlin52	OX	3.09	5.62	16.045.200
	ERX	0.00	1.35	16.938.904
	MPX	0.00	3.78	19.307.034
	OX, ERX, MPX	0.00	1.45	7.233.215
ch130	OX	2.24	4.59	15.281.04
	ERX	2.27	5.20	18.840.03
	MPX	3.60	4.77	23.164.73
	OX, ERX, MPX	0.00	2.05	6.797.867
kroA200	OX	3.77	6.49	13.188.469
	ERX	118.6	121.7	28.406.603
	MPX	3.13	4.04	22.728.010
	OX, ERX, MPX	0.00	2.72	6.171.308

4.3. References to Offspring Selection Applications

The basic concepts of the enhanced selection ideas as published in the present paper have already emerged a couple of years ago. As the actual focus (like also stated in the present contribution) is to study the properties of the new selection concepts systematically, the potential w.r.t. achievable advancements in global solution quality were obvious immediately. Therefore, the main aim of the first works in this area was to check the generality of the new algorithmic concepts by applying them to various theoretically as well as practically relevant problems. And indeed this worked out very well and it was possible to demonstrate similar effects and achievements in global solution quality in various areas of application under very different problem codifications with exactly that enhanced generic selection techniques as being proposed in this paper.

While the last subsections considered only relatively small TSP instances in order to illustrate some selected aspects, journal article (Affenzeller and Wagner 2004) includes a detailed and comprehensive empirical analysis also based on TSP instances of much higher dimension. Furthermore, (Affenzeller and Wagner 2003; Affenzeller and Wagner 2004; Affenzeller 2005) give a comprehensive solution analysis based on several real valued n-dimensional test functions (like the n-dimensional Rosenbrock, Rastrigin, Griewangk, Ackley, or Schwefel's sine root function). Also here it is possible to locate the global optimal solution in dimensions up to *n=2000* with exactly the same generic extensions of the selection

model as being stated here - only the crossover- and mutation-operators have been replaced with standard operators for real-valued encoding.

In (Winkler, Affenzeller and Wagner 2005; Winkler, Efendic, Affenzeller, Del Re and Wagner 2005) we report promising results achieved in the context of nonlinear structure identification based on time-series data of a diesel combustion engine. Concretely the aim of this project is the development of models for the NO_x emission. Already until now it has become possible with a GP-based approach equipped with offspring selection to identify models which are superior to the models achieved with conventional GP-techniques and also superior to machine learning techniques which have also been considered in earlier stages of this project. Such results including also extensive empirical comparisons between offspring selection GP, conventional GP and other machine learning techniques like artificial neural networks are reported for dynamical mechatronical systems as well as for well known medical benchmark data sets taken from the UCI machine learning library (Winkler, Affenzeller and Wagner 2006; Winkler, Affenzeller and Wagner 2007; Winkler, Affenzeller and Wagner 2008)[2].

5. CONCLUSION

Possibly the most important feature of the newly introduced concepts is that the achievable solution quality can be improved in a non-problem specific manner so that it can be applied to all areas of application for which the theory of Genetic Algorithms and Genetic Programming provides suitable operators. Further aspects worth mentioning concern the robustness and self-adaptiveness of the population genetics inspired measures:

Basically weak operators become powerful and the selection pressure is steered self-adaptively in a way that the amount of selection pressure actually applied is that high that further progress of evolutionary search can be achieved. Nevertheless the newly developed selection techniques are not problem specific at all (cf. subsection 4.3.). Possible future research topics in that area are certainly to open new areas of application due to the increased robustness and also more theoretical topics like the analysis of various aspects of population genetics and their interaction with concrete applications of evolutionary computation. Especially for the theory of parallel genetic algorithms the interactions between genetic drift and migration should be a very fruitful field of research.

[2] A complete list of publications is available at http://www.heuristiclab.com/publications/

REFERENCES

Affenzeller, M., 2001. 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.

Affenzeller, M., 2002. A generic evolutionary computation approach based upon genetic algorithms and evolution strategies. *Journal of Systems Science*; 28 (2), 59 – 72.

Affenzeller, M., 2005, *Population genetics and evolutionary computation: Theoretical and practical aspects*. Linz: Trauner Verlag.

Affenzeller, M., Wagner, S., 2003. SASEGASA: An evolutionary algorithm for retarding premature convergence by self-adaptive selection pressure steering. *Computational Methods in Neural Modeling, Lecture Notes of Computer Science*; 2686, 438 – 445.

Affenzeller, M., Wagner, S., 2004a. SASEGASA: A new generic parallel evolutionary algorithm for achieving highest quality results. *Journal of Heuristics, Special Issue on New Advances on Parallel Meta-Heuristics for Complex Problems*; 10 (3), 239 – 263.

Affenzeller, M., Wagner, S., 2004b. Reconsidering the Selection Concept of Genetic Algorithms from a Population Genetics Inspired Point of View. *Cybernetics and Systems 2004*, 701-706.

Cavicchio, D., 1970. *Adaptive Search using Simulated Evolution*. PhD Thesis, University of Michigan.

DeJong, K., 1975. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*. PhD Thesis, University of Michigan.

Fogel, D., 1994. An introduction to simulated evolutionary optimization. *IEEE Trans. on Neural Network*; 5 (1), 3 – 14.

Goldberg, D., 1989. *Genetic Alogorithms in Search, Optimization and Machine Learning*. Addison Wesley Longman.

Larranaga, P., Kuijpers, C., Murga, R. and Dizdarevic S., 1999. Genetic algorithms for the travelling salesman problem: A review of representations and operators. *Artificial Intelligence Review*, 13 (2), 129 – 170.

Michalewicz, Z., 1996. *Genetic Algorithms + Data Structurs = Evolution Programs*. Berlin: Springer-Verlag, 3rd edition.

Rechenberg, I., 1973. *Evolutionsstrategie*. Stuttgart: Friedrich Frommann Verlag.

Reinelt, G., 1991. TSPLIB - A traveling salesman problem library. *ORSA Journal on Computing*; 3, 376 – 384.

Schoeneburg, E., Heinzmann, F. and Feddersen, S., 1994. *Genetische Algorithmen und Evolutionsstrategien*. Addison-Wesley.

Wagner, S. and Affenzeller, M., 2005. Heuristiclab: A generic and extensible optimization environment. Adaptiveand *Natural Computing Algorithms* –

Proceedings of ICANNGA 2005, 538 – 541. 21st – 23rd March, Coimbra, Portugal.

Winkler, S.; Affenzeller, M. and Wagner, S., 2005. New methods for the identification of nonlinear model structures based upon genetic programming techniques. *Journal of Systems Science*, 31, 5 – 13.

Winkler, S.; Affenzeller, M. and Wagner, S., 2006. Heuristic Modeler: A multi-purpose evolutionary machine learning algorithm and its applications in medical data analysis. *Proceedings of the International Mediterranean Modelling Multiconference I3M 2006*, 629 – 634. 4th Oct. – 6th Oct., Barcelona, Spain

Winkler, S., Affenzeller, M. and Wagner, S., 2007, Advanced genetic programming based machine learning. *Journal of Mathematical Modelling and Algorithms*, 6 (3), 455 – 480.

Winkler, S., Affenzeller, M. and Wagner, S., 2008, Offspring selection and its effects on genetic propagation in genetic programming based system identification. *Cybernetics and Systems 2008*, 2, 49-54.

Winkler, S.; Efendic,H.; Affenzeller, M.; Del Re, L. and Wagner, S., 2005, On-line modeling based on genetic programming. *Proceedings of the 1st International Workshop on Automatic Learning and Real-Time (ALaRT'05)*, 119 – 130. 7th – 8th Sept. 2005, Siegen; Germany.

AUTHORS BIOGRAPHY

MICHAEL AFFENZELLER has published several papers and journal articles dealing with theoretical aspects of evolutionary computation and genetic algorithms. In 1997 he received his MSc in mathematics and in 2001 his PhD in computer science, both from JKU Linz, Austria. He is professor at the Upper Austria University of Applied Sciences (Campus Hagenberg) and associate professor at the Institute of Formal Models and Verification at JKU Linz since his habilitation in 2004.

STEFAN WAGNER received his MSc in computer science in 2004 from Johannes Kepler University Linz, Austria. He now holds the position of an associate professor at the Upper Austrian University of Applied Sciences (Campus Hagenberg). His research interests include evolutionary computation and heuristic optimization, theory and application of genetic algorithms, machine learning and software development.

STEPHAN M. WINKLER received his MSc in computer science in 2004 and his PhD in technical sciences in 2008, both from JKU Linz, Austria. His research interests include genetic programming, nonlinear model identification and machine learning. Currently he is research associate at the Research Center Hagenberg of the Upper Austrian University of Applied Sciences, working on the research program L284-N04 "GP-Based Techniques for the Design of Virtual Sensors", a research project funded by the Austrian Science Fund (FWF).