Convergence Characteristics of Keep-Best Reproduction

Kay Wiese and Scott D. Goodwin
Department of Computer Science, University of Regina
Regina, Saskatchewan, S4S 0A2, Canada
Phone: (306)-585-5210, Fax: (306)-585-4745
email: wiese@cs.uregina.ca, goodwin@cs.uregina.ca

Abstract

This paper presents some theoretical convergence characteristics of *Keep-Best Reproduction* (KBR), a selection strategy for genetic algorithms (GAs). We have previously introduced KBR and reported encouraging results in the traveling salesman domain (Wie98a) where KBR was compared with the standard replacement strategy of replacing the two parents by their two children (STDS). Here we demonstrate that in a non-operator environment as well as in the ONEMAX domain KBR has the same convergence characteristics as 2-tournament selection and elitist recombination (ELR) (Thi94a). We also show how a modification of ELR suggested in (Thi97) can be utilized to tune the selective pressure of KBR. These analytical models are fairly simplistic and cannot accurately model the convergence characteristics in more complex domains where building blocks are correlated, such as the TSP domain. We will give some empirical results of a comparison of KBR and ELR in this domain.

Topic Area: Selection Schemes, Family Competition

Keep-Best Reproduction: A Family Competition Scheme

We have previously developed a family competition scheme which we call *Keep-Best Reproduction* (KBR). KBR takes two parents, recombines them and then keeps the best parent and the best offspring in order to introduce good new genetic material into the population as well as to keep good old chromosomes. Intuitively this increases the selection pressure which should lead to faster convergence but by keeping the best parent we hope to maintain enough diversity to avoid overselection and premature convergence. Other researchers have used the idea of a competition (global or local) as well but their implementations all differ from KBR. A discussion of these other algorithms is not possible within the length of this short paper, but references are given at the end.

Selection Intensity and Convergence Models for KBR

The selection pressure of a selection scheme is usually quantified by its selection intensity I:

$$I(t) = \frac{S(t)}{\sigma(t)} = \frac{\overline{f^s(t)} - \overline{f(t)}}{\sigma(t)}$$

Here the selection differential $S(t)$ is the difference between the average fitness of the parent population at generation t, $\overline{f^s(t)}$, and the population mean fitness at generation t, $\overline{f(t)}$.

Assuming a standardized normal distribution of the initial population's fitnesses, i.e. $N(\overline{f}, \sigma) = N(0,1)$, the selection intensity I simply becomes the expected average fitness of the population after applying the selection scheme to the original population. This is exactly the model that Blickle and Thiele used to compute selection intensities (Bli95). They have derived the selection intensity for tournament selection of size s to be

$$I(s) = \int_{-\infty}^{\infty} s\ x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right)^{s-1} dx$$

According to (Bli95) these integral equations can be solved analytically for the cases $s = 1, \ldots, 5$. For example for a tournament of size 1 the selection intensity is $I(1) = 0$ and for a tournament of size 2 it is $I(2) = \frac{1}{\sqrt{\pi}}$. For tournaments of size $s > 5$, the integral equation has to be solved numerically. Alternatively for tournament sizes of $s > 5$ Blickle and Thiele derived an approximation formula with a relative error of less than 1%:

$$I(s) \approx \sqrt{2(\ln(s) - \ln(\sqrt{4.14 \ln(s)}))}$$

This approximation formula can also be used for $s \in [2, 5]$ with a relative error of less than 2.4%. Table 1 shows the selection intensities of tournament selection for various tournament sizes.

For a tournament of size 2, Thierens and Goldberg (Thi94b) derive the same selection intensity (in the form of the population average fitness increase from one generation to the next in the ONEMAX domain) but in a completely different manner as Blickle and

Table 1: Selection Intensities $I(s)$ for Tournaments of Size s

s	1	2	3	4	5
I(s)	0.00	0.56	0.85	1.03	1.16

Thiele. Thierens and Goldberg's formulation can not be extended to other tournament sizes. However, they also derive a convergence model for 2-tournament in the ONEMAX domain:

$$p(t) = \frac{1}{2}(1 + \sin(\frac{t}{\sqrt{\pi l}}))$$

$p(t)$ is the proportion of 1-bits in the total population at generation t, while l is the bitlenght of each chromosome. For a randomly initialized population $p(0) = 0.5$ can be assumed. To compute g_{conv}, the total number of generations the population needs to fully converge, we set $p(g_{conv}) = 1$, and solve for g_{conv}. We find

$$g_{conv} = \frac{\pi}{2}\sqrt{\pi l}$$

In the same paper, Thierens and Goldberg showed that their ELR algorithm has the exact same selection intensity and convergence characteristics as 2-tournament selection. In (Thi94a) they showed that when optimizing the ONEMAX function, the best parent will go to the next generation and the worst parent will be replaced by the best child. This is easy to understand if we consider that in the ONEMAX domain the total number of 1-bits before and after crossover remains the same.

This is exactly what KBR explicitly does. Therefore the selection intensity for KBR with random parent selection is the same as for 2-tournament selection and for elitist recombination, namely $I = \frac{1}{\sqrt{\pi}}$. The convergence model of KBR is the same as for 2-tournament selection and ELR in the ONEMAX domain:

$$p(t) = \frac{1}{2}(1 + \sin(\frac{t}{\sqrt{\pi l}}))$$

and

$$g_{conv} = \frac{\pi}{2}\sqrt{\pi l}$$

While the selection intensity of tournament selection can be tuned by changing the size of the tournament, both ELR and KBR have fixed selection intensities. Thierens (Thi97) proposed a modified elitist recombination that allows to tune the selection pressure of ELR, much in the same way as this can be done for tournament selection by modifying the size of the tournament. He has proposed to select one parent via a tournament of size s and to select the other parent randomly from the population. Then the local ELR family competition is applied. Thierens' model assumes a heritability of 1, which he achieves by not applying genetic operators. Since Thierens' model does not take into account genetic operators, children are simply copies of their parents. We have already argued that the two fittest individuals from the two parents and children are the fittest parent and its child clone, which are then inserted into the next generation by ELR. Again, this is exactly what KBR does explicitly. We can use Thierens' modified parent selection to tune selection intensities for KBR in the same way it is done for ELR.

The selection intensity of this modified KBR can be computed as the $(s+1)^{th}$ order statistics of a random sample of size $s+1$, which is also the population mean fitness increase since the standard deviation of the starting population is 1. Using the notation of Blickle and Thiele (Bli95) we can write the selection intensity of the modified KBR as:

$$I(s) = \int_{-\infty}^{\infty} (s+1)\, x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \left(\int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right)^s dx$$

This integral equation is analytically solvable for $s \leq 4$. For $s > 4$ it can be solved numerically or by using a modification of Blickle and Thiele's formula:

$$I(s) \approx \sqrt{2(\ln(s+1) - \ln(\sqrt{4.14 \ln(s+1)}))}$$

The relative error of this approximation is less then 2.4% for $s \in [2, 4]$ and less than 1% for $s > 4$. Table 2 shows the selection intensities of this modified KBR. We can conclude that the modified KBR with tournament size s has the same selection intensity as regular tournament selection with tournament size $s+1$.

Table 2: Selection intensities $I(s)$ for the modified KBR. One parent is selected by a tournament of Size s, while the other parent is selected at random

s	1	2	3	4	5
I(s)	0.56	0.85	1.03	1.16	1.27

Thierens and Goldberg (Thi94a) have performed experiments in the ONEMAX domain as well as on bounded fully deceptive functions and empirically compared tournament selection of size 2 and standard replacement with ELR. Their conclusions were that in the ONEMAX domain the two selection schemes show very little difference. In the deceptive function domain ELR performed slightly better than tournament selection. This became even more evident when the populations were undersized.

We have not performed empirical studies in the ONEMAX domain that compare KBR with ELR and tournament selection, since in this domain KBR and ELR are literally the same algorithms and one cannot expect that our results for KBR would differ from the ones found in (Thi94a). The above convergence analysis of KBR is limited to either the ONEMAX domain or a model with heritability one, meaning that the offspring are no different from the parents. We believe that the difference of KBR and ELR can only be shown in a domain where longer building blocks need to be processed and operators can have a disruptive effect on these building blocks. We choose to compare KBR and ELR in the TSP domain which is also a problem domain of practical interest.

Comparison of KBR and ELR in the TSP domain

We have compared KBR and ELR on a 100 city asymmetric travelling salesman problem. The cost between two cities was a random integer number between 0 and $maxcost$, where $maxcost$ was set to 100 times the number of cities. The parent selection was done via roulette wheel selection. The mutation operator was a simple swap operation that picks two random locations in the tour, and exchanges the two cities in those locations. The crossover operator we used was the partially mapped crossover (PMX). A detailed description of PMX can be found in (Gol85). The fitness function we used was $f_i = c_{max} - c_i$, where c_i is the actual tour cost of individual i and c_{max} is the maximum cost in the population.

Recombination Alone

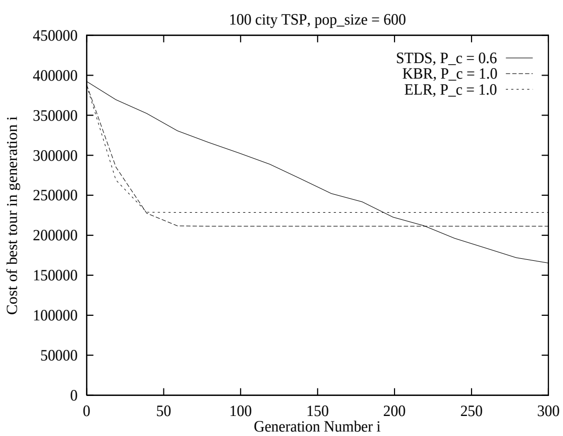

Figure 1: The x-axis shows the generation number i, while the y-axis shows the cost of the best tour after i generations for a 100 city random TSP. No mutation was used. The crossover probability was set to $P_c = 0.6$ for STDS and to $P_c = 1.0$ for KBR and ELR. The population size was 600

Figure 1 shows that both KBR and ELR converge prematurely. ELR more so than KBR. For comparison we have also depicted the convergence of STDS in Fig. 1. STDS simply replaces the two parents by the two offspring regardless of their fitness. The premature convergence is due to a loss of diversity. In order to reintroduce diversity we use higher mutation rates in combination with crossover.

Recombination and Mutation

With KBR we were able to speed up the convergence of the GA by using higher mutation rates. This should not come as a surprise. While with STDS, mutation is performed and the mutated chromosomes are inserted into the next generation, KBR only keeps the best child. In case mutation lowers the fitness of the offspring, there is always the good genetic material of the best parent that is kept. So higher mutation rates are not as disruptive as with STDS. On the other hand, without mutation, KBR very rapidly converges to local optima of low quality. The higher mutation rates reintroduce diversity and help steer the GA away from these inferior local optima. A similar argument can be made for ELR. According to Fig. 2, ELR converges even more rapidly, but fails to find better solutions than KBR. The cost of the cheapest tour found for a population size of 600 was 86,719 with ELR and 85,541 with KBR after 600 generations. In fact the population for ELR was fully converged after 220 generations, while for KBR after 600 generations there was still diversity in the population and room for exploration. Similar findings were made with other population sizes and problem sizes.

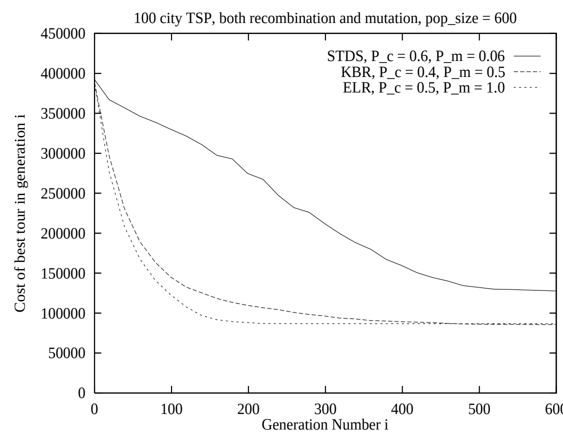

Figure 2: The x-axis shows the generation number i, while the y-axis shows the cost of the best tour after i generations for a 100 city random TSP. The population size was 600

Conclusion

We have derived the selection intensity of KBR in a simplistic non-operator environment as well as for the ONEMAX domain to be $I = \frac{1}{\sqrt{(\pi)}}$. Also we have derived a convergence model for KBR in those two domains. Both selection intensity and convergence model of KBR are identical to the selection intensity and convergence models for 2-tournament selection and ELR. We have demonstrated how an idea introduced by Thierens (Thi97) can be used to tune the selection intensity of KBR in the same way the selection intensity can be tuned for tournament selection by modifying the tournament size. We believe the difference between KBR and ELR can only be shown in a domain where building blocks are correlated, such as the TSP domain. In this domain both KBR and ELR show similar advantages when compared with the standard selection strategy of replacing both parents by their offspring, such as a more efficient and more effective search. Also both KBR and ELR work well with smaller population sizes. One of the differences between KBR and ELR in the TSP domain is that ELR converges more rapidly, but usually towards solutions of lesser quality than the ones found by KBR. For the tests that we have performed the tours found by ELR were on average about 3% more expensive than the tours found by KBR.

References

Altenberg, L., "The Evolution of Evolvability in Genetic Programming", *Advances in Genetic Programming*, pp. 47–74, 1994.

Baeck, T., "Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms", Oxford University Press, 1996.

Blickle, T., and Thiele, L., *A Comparison of Selection Schemes Used in Genetic Algorithms*, TIK Report Nr.11, 2nd Edition, Swiss Federal Institute of Technology, 1995.

Culberson J., "Crossover versus Mutation: Fueling the Debate: TGA versus GIGA", *Proceedings of the Fifth International Conference on Genetic Algorithms ICGA-93*, pp. 632, 1993.

Eshelman, L.J., "The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination", *Foundations of Genetic Algorithms, FOGA-I*, pp. 265–283, 1991.

Goldberg, D.E. and Lingle, R., Jr., "Alleles, Loci, and the Traveling Salesman Problem", *Proceedings of the 1st ICGA*, pp. 154–159, 1985.

Holland, J.H., *Adaptation in Natural and Artificial Systems*, Ann Arbor, MI, University of Michigan Press, 1975.

Mahfoud, S.W., "Crowding and Preselection Revisited", *Proceedings of Parallel Problem Solving from Nature PPSN-II*, pp. 27–36, 1992.

Oliver, I.M., Smith, D.J., and Holland, J.R.C.,, "A Study of Permutation Crossover Operators on the Traveling Salesman Problem", *Proceedings of the 2nd ICGA*, pp. 224-230, 1987.

Schwefel, H.P., "Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie", *Interdisciplinary Systems Research*, Volume 26, Birkhaeuser, Basel, 1977.

Starkweather, T., McDaniel, S., et. al, "A Comparison of Genetic Sequencing Operators", *Proceedings of the 4th ICGA*, pp. 69–76, 1991.

Syswerda, S., "Uniform Crossover in Genetic Algorithms", *Proceedings of the Third International Conference on Genetic Algorithms, ICGA-89*, pp. 2–9, 1989.

Thierens, D. and Goldberg, D.E., "Elitist Recombination: An Integrated Selection Recombination GA", *Proceedings of the 1st IEEE World Congress on Computational Intelligence*, pp. 508–512, 1994.

Thierens, D. and Goldberg, D.E., "Convergence Models of Genetic Algorithm Selection Schemes", *Parallel Problem Solving from Nature PPSN-III*, Lecture Notes in Computer Science Vol. 866, Springer Verlag, pp. 119-129, 1994.

Thierens, D., "Selection Schemes, Elitist Recombination and Selection Intensity", *Proceedings of the 7th International Conference on Genetic Algorithms ICGA-97*, pp. 152–159, 1997.

Wiese, K. and Goodwin, S.D., "Keep-Best Reproduction: A Selection Strategy for Genetic Algorithms", *Proceedings of the 1998 ACM Symposium on Applied Computing SAC'98*, pp. 343–348, 1998.

Wiese K. and Goodwin, S.D., "Parallel Genetic Algorithms for Constrained Ordering Problems", *Proceedings of the 11th International Florida Artificial Intelligence Research Symposium, FLAIRS'98*, in print, 1998.

Whitley, D., "The GENITOR Algorithm and Selection Pressure: Why Rank-Based Allocation of Reproductive Trials is Best", *Proceedings of the Third International Conference on Genetic Algorithms ICGA-89*, pp. 116-121, 1989.

Whitley, D., Starkweather, T., and D'Ann Fuquay, "Scheduling Problems and Traveling Salesman: The Genetic Edge Recombination Operator", *Proceedings of the 3rd ICGA*, pp. 133–140, 1989.