Reconstructing the Shifting Balance Theory in a GA: Taking Sewall Wright Seriously

Franz Oppacher
Intelligent Systems Laboratory,
Carleton University
Ottawa, Ont., Canada
oppacher@scs.carleton.ca

Mark Wineberg
Intelligent Systems Laboratory,
Carleton University,
Ottawa, Ont., Canada
wineberg@scs.carleton.ca

Abstract- We attempt to reconstruct Sewall Wright's Shifting Balance Theory in order to address some of the major criticisms leveled against it. The resulting abstract process is applied to the GA forming the Shifting Balance Genetic Algorithm (SBGA), which is shown to behave as Wright intended. For example, the SBGA avoids local optima through a shifting balance between subpopulations, as is demonstrated in an experiment. The experiment also shows that the SBGA outperforms the classical GA in both stationary and changing environments.

1 Sewall Wright's Shifting Balance Theory

The population geneticist Sewall Wright (Wright, 1932) realized the problem of an evolutionary system getting trapped at local maxima back in the 1930's. Wright called the (hypothetical) function that was being optimized by the evolutionary system the "fitness landscape" and thought that each species was, over time, hill climbing through the process of mutation and natural selection. Having this view, he naturally became aware of the problem of local maxima:

> In a rugged field of this character [very many peaks], selection will easily carry the species to the nearest peak, but there may be innumerable other peaks which are higher but which are separated by "valleys." The problem of evolution as I see it is that of a mechanism by which the species may continually find its way from lower to higher peaks in such a field.[1]

Wright not only observed the problem, but also proposed a mechanism to cope with it. This mechanism was eventually called by Wright "the shifting balance theory" (SBT). Hartl and Clark present a good summary of the shifting balance theory in their textbook *Principles of Population Genetics*:

> In the **shifting balance theory**, a large population that is subdivided into a set of small, semi-isolated subpopulations (demes) has the best chance for the subpopulations to explore the full range of the adaptive topography … If the subpopulations are sufficiently small, and the migration rate between them is sufficiently small, then the subpopulations are susceptible to random genetic drift of allele frequencies, which allows them to explore their adaptive topography more or less independently. In any subpopulation, random genetic drift can result in a temporary reduction in fitness that would be prevented by selection in a larger population, and so a subpopulation can pass through a "valley" of reduced fitness and possibly end up "climbing" a peak of fitness higher than the original. Any lucky subpopulation that reaches a higher adaptive peak on the fitness surface increases in size and sends out more migrants to nearby subpopulations, and the favorable gene combinations are gradually spread throughout the entire set of subpopulations by means of interdeme selection[2].

Hartl and Clark then identify three distinct phases:

> 1. An exploratory phase, in which random genetic drift plays an important role in allowing small populations to explore their adaptive topography.
>
> 2. A phase of mass selection, in which favorable gene combinations created by chance in the random drift phase rapidly become incorporated into the genome of local subpopulations by the action of natural selection.
>
> 3. A phase of interdeme selection, in which the more successful demes increase in size and rate of migration; the excess migration shifts the allele frequencies of nearby population until they also come under the control of the higher fitness peak. The favorable genotypes thereby become spread throughout the entire population in an ever-widening distribution.[3]

In the above explanation the term *random drift* refers to a statistical process that occurs in populations because of the random sampling of alleles produced by random mating.

---

[1] (Wright 1932) as reprinted in (Provine 1986) pp. 162-164.

[2] (Hartl & Clark 1997) pg.259.

[3] (Hartl & Clark 1997) pp. 259-260.

This effectively causes a random walk in the relative proportions of two competing alleles; the smaller the population, the more erratic the random walk. Wild swings in the proportion of the two alleles can, in very small populations, even eliminate an allele independent of the selective value of that allele. Wright's visualization of the demes moving from one peak to another in a fitness landscape (based on a diagram from Wright (1932)) can be seen in Figure 1. However, when viewing the diagram, keep in mind that the landscape is a "fitness" landscape, not a geographical one, and the demes consist of genetically similar individuals, not necessarily spatially local individuals (although Wright views the first as being implied by the second).

2 The Deme Approach in the GA Literature as Inspired by the SBT

Wright's ideas have been popular in the Genetic Algorithm (GA) community, with many authors quoting his theory (often referred to as Sewall Wright's demes) when justifying the introduction of a novel feature in their GA implementation. This is particularly true in parallel GA research.

The island or coarse-grained model, each processor runs a normal GA on its own population, where each population is called a deme. For all of the demes, the GAs act on the same fitness function during selection, although the various populations are different from one another because of the random process of chromosome creation. Furthermore, each deme may evolve in different directions due to the probabilistic nature of evolution. Periodically, migrants travel between processors bringing new - possibly better - genetic material from other demes. Usually the members of a population chosen to be migrants are the elite of the population. This basic mechanism has been independently "discovered" many times over; see (Cohoon et al. 1987), (Pettey et al. 1987), and (Tanese 1987).

In the second approach, which is run on a fine-grained machine, each processor holds only a single member of the population. Each member is only allowed to mate with others in their neighborhood. This design decision is motivated by the fact that the fastest communication happens among neighbors. The neighbors can be arranged in various topologies such as a chain (or circle), grid (or toroid), or a hypercube. Diffusion of genetic information occurs because neighborhoods overlap. These neighborhoods are identified in many publications as demes. As with the island model, the fine-grained model was also frequently "discovered" by various researchers: For example see (Gorges-Schleuter 1989), (Spiessens & Manderick 1991), (Mühlenbein 1989), (Hillis 1991), (Collins et al. 1991), and (Davidor 1991).

The use of demes in sequential GA are similar to the fine-grain model in parallelism. In (Pál 1994) the favored topology is a simple ring with mating restricted to contiguous members. The more common toroidal topology
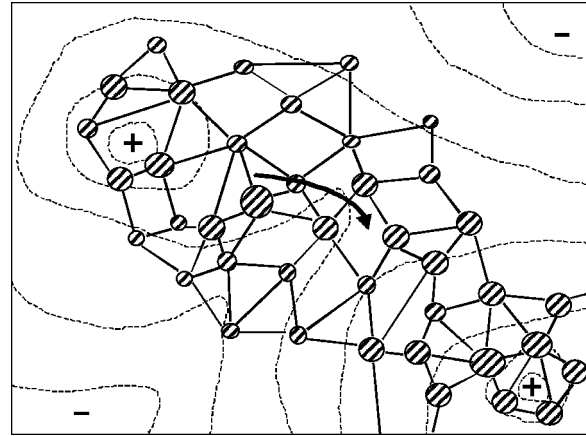

Figure 1: Demes shifting from one peak to another on a topographical representation of a fitness landscape (based on a diagram from Wright (1932)).

is used in (Sarma & De Jong 1999) and (Sumida & Hamilton 1994).

3 Problems With the SBT, both in Theory and in Practice

While all of the above deme models, both parallel and sequential, may be interesting in their own rights, they are not similar enough to the SBT in either form or purpose to claim direct descendance from it, as some of their authors purport. In fact, in view of the basic difference of the two approaches to parallelization, it is hard to see how the shifting balance theory could be invoked to support both simultaneously.

Of course not every implementor of a deme based GA makes such a claim. However, even those who directly mention Wright and the shifting balance theorem are really not making any attempt to hew closely to the theory and are just using it to support their system through analogy. For example most do not comment on the importance of random drift, and just mention that Wright proposed that interdemic selection was a powerful mechanism in natural evolutionary processes. To properly acknowledge the use of the mechanism one would have to ensure that the deme sizes are kept small enough for random drift to take effect and gain purchase on the footholds of a new maximum. Usually no special attention to the deme size is given for such a purpose and if the demes are in fact kept to a proper minimal size it is done by chance and not through experimentation or theoretical reasoning.

There is one paper that goes beneath the surface of Wright's theory. (Sumida and Hamilton 1994) makes reference to its details. They purposely keep the demes in their system small in order to facilitate the random drift that Wright postulated was necessary. They also attempt to model interplay between demes allowing for more distant

communication to periodically occur. It is questionable, however, whether these long distance communications are what Wright had in mind when he talked of interdeme selection and the shift in balance between random drift and selection. Furthermore, Sumida and Hamilton do not appear to have made any attempt to determine whether their deme size was small enough to promote random drift. Nor did they determine whether the interconnectedness fostered by long distance mating was effective enough to allow selection to shift the balance of the mass of the population onto the new hill and proceed to climb it.

However, even if the SBT were implemented exactly according to specification, there would still be a very large problem attendant on its use: the theory as it stands is inconsistent. The SBT depends quite heavily on balancing two evolutionary pressures against each other. Random drift is needed to cross the low fitness valleys, while selection is needed to climb the hills once the valleys have been crossed. Yet the two mechanisms cannot be invoked simultaneously. Random drift is dependent on small populations to operate while selection is effective in large populations. Indeed the very effects that are to be obtained through both are mutually exclusive since one cannot climb a hill and descend through a valley at the same time. Wright seemed to have the notion of segregating many small demes for the time necessary to fix various traits through random drift, then reintegrate them back into a larger population to shift the genetic makeup of the whole, but never made clear the details.

This defect in the theory is well known in the biological community. From Hartl and Clark:

> For the theory to work as envisaged, ... [t]he population must be split up into smaller demes, which must be small enough for random genetic drift to be important but large enough for mass selection to fix favorable combinations of alleles. While migration between demes is necessary, neighboring demes must be sufficiently isolated for genetic differentiation to occur, but sufficiently connected for favorable gene combinations to spread.[4]

Consequently, while the SBT has been very useful in genetics as a metaphor to aid the visualization of the process of evolution, very little work has been done to verify whether it correctly describes evolution. This is primarily because much of it is untestable. Once again from Hartl and Clark: "Because of uncertainly (sic) about the applicability of these assumptions [mentioned in the previous quote], the shifting balance process remains a picturesque metaphor that is still largely untested."[5]

This balancing act of deme sizes, small enough for random drift but large enough for selection to occur, seems to be a giant impediment to implementation. What is the right size? Is there a right size; or are the two mechanisms of random drift and selection largely exclusive? These questions are not addressed by previous research.

4 Abstracting from the SBT: The Shifting Balance Genetic Algorithm

4.1 The First Level of Abstraction: Core Group and Colonies

One way to 'balance' the behavior of random drift and selection is to drop one in favor of the other. Of the two, keeping random drift would be a poor choice, after all random drift is just that – random. Selection on the other hand is more pliable. While selection does allow for the maximization of an objective function, it can be converted so that local maxima can be eschewed and valleys crossed. This can be accomplished by recognizing that maximizing the optimization function need not be the only trait selected for.

It is by the combination of isolating subpopulations and using different selection criteria in each that the ends of Wright's theory can be obtained without resorting to his conflicting means. The majority of the members continue to optimize the objective function. Smaller demes are created and given different environments in which to evolve. If these environments are different enough from that of the main population, the various local maxima will be in different positions in the gene space. Therefore the selection pressure in each deme will force the members within it to converge on these other maxima. Since the positions of these maxima are shifted from that of the main population, so too will be the members under its purview. Consequently, any members that migrate from the smaller demes to the large deme will no longer be in the same location in the gene space. If there are enough small demes, then migrants from those demes will be occasionally be on a larger hill when evaluated by the objective function of the large population. The genetic makeup of these members will begin to dominate in the large population, and shift its position in gene space to the newly found hill. Note the effect of the shifting balance is felt with only selection applied throughout. There is thus no need to balance two contrary forces to shift the balance of a population.

The following summarizes our first level of abstraction: A large central population, the *core*, experiences a single environment and consequently is under the selective pressure of a single fitness landscape. While adapting to this landscape the core can get stuck on a local maximum, at which point no evolutionary progress is occurring. However, the core is constantly sending out members to peripheral groups called *colonies*. The members within these colonies find themselves in a different environment; hence each experiences a fitness landscape different from that of the others and of the core group. Consequently the new landscape for each colony may not have the same local maximum as the landscape of the core group. The colonies are then forced to adapt through selection to these new landscapes, thus changing the genetic makeup of its

---

4   (Hartl & Clark 1997) pg. 260.

5   Ibid.

constituents. The modified colony members who migrate back to the core are now different from its members. Furthermore they may be on a better hill in the fitness landscape of the central area. Thus, through the use of the colonies, the species as a whole can now jump over valleys in the fitness landscape and continue to evolve.

Note again that the new model only depends on selection. While random drift may occur, it can only help add to the variation being produced by selection, and is not mandatory to the system anymore. Furthermore, the population sizes of the colonies, while much smaller than the core group, are set large enough to prevent undue effects due to random drift.

4.2 The Colonies as Exploratory Sub-Populations: A Further Abstraction

The reader will have noticed that the colonies are supposed to be in different fitness landscapes, yet no method exists in the model to determine how the environments should differ from that of the core. Specifically, these environments should be capable of moving the system off of local maxima. What property of the alternative fitness functions would allow them to push the system away from a local maximum?

The alternate fitness functions in the colonies encourage them to search in different areas of the search space. Any mechanism that encourages such an exploration will thus achieve the goal of moving the system away from a local maximum. Now, instead of trying to create new fitness functions to do this, perhaps we could directly determine whether the colony is searching in the same area as the core group. If the system can identify that such an overlap is occurring, then the colony is not performing its duty and should be moved elsewhere.

The above analysis makes it clear that we need a mechanism to relocate the population to other places in the search space independently of the selection pressure, thus escaping from a local maximum. To relocate a population away from a local maximum, other researchers resort to a blind restart mechanism, see (Eshelman 1991), (Whitley et al. 1991), (Cobb et al. 1993), and (Maresky et al. 1995). We prefer not to resort to such a draconian measure.

Fortunately a more evolutionary alternative is available. If we had a method of determining whether the colony is searching in the same area as the core, then we could use this to determine whether a colony is searching in novel territory. This 'distance' to the core could then be used as a selection criterion for the members in the colony: the members will be selected for reproduction according to an amalgamation of their 'distance' from the core the objective function rather than according to the objective function alone. They will *evolve* into a new area of the search space.

If selection based on 'distance' to the core is used exclusively, the colonies will just move as far from the core as possible. This will most likely send them into unpromising territories. A balance must therefore be kept. The colonies must be allowed to search for good solutions (using the regular fitness function) when searching in areas

where the core is not; when searching in the part of gene-space near the core, they must be selected for distance from the core, thus moving them away.

While this new abstraction may seem to be a more radical departure from the SBT than the first, it is actually only an extension of it. The original modification of the SBT required environments in the colonies that were slightly different than the environment in the core. How those environments were to have become different was not specified. With the new abstraction, the new colony environment is differentiated from that of the core through the direct means of determining where the core is and changing the environment such that it rewards members that are further from the core as well as members that are maximizing the objective function. The colony environment is therefore slightly different from that of the core, with local maxima positioned away from the local maximum where the core is trapped.

4.3 Prevention versus Promotion of Inbreeding

Selecting for dissimilarity to core which is being done in the colonies has its analogy in nature apart from Sewall Wright's theory. If one selects among individuals that differ from a given group, one is preventing inbreeding. In the modified SBT, the colonies are specifically bred to be different from the core, and so when introduced back into the core, through migration, variety is generated. It is just this variety that the SBT wants to promote.

Indeed, mechanisms to prevent inbreeding are known in both biology and the evolutionary computation literature. Selective mating in mice was found to promote outbreeding (Potts et al. 1991). The dispersal of birth and breeding sites in birds is correlated to less inbreeding and higher diversity (Greenwood et al. 1978).

In evolutionary computation, there are three main techniques to prevent inbreeding: incest prevention, crowding and fitness sharing. Incest prevention was introduced by (Eshelman 1991) in their CHC implementation of the GA. Here chromosomes with Hamming distances below a certain threshold were prevented from mating. In crowding, mating takes place as usual, but the offspring is inserted back into the population replacing the individual most similar to it, again as measured by the Hamming distance, see (De Jong 1975) and (Goldberg 1989) for details. Finally, fitness sharing divides the fitness of similar individuals, again as measured by the Hamming distance, thus promoting the selection of individuals that are more distinct in the population (Goldberg & Richardson 1987). Both crowding and fitness sharing have been used to tackle multi-modal fitness functions.

The SBT, while focused on increasing variety, does not do so by preventing inbreeding – rather, it actively promotes it! The effect of having a small deme size to encourage random drift is to increase inbreeding. This is not an unwanted side effect, but rather an actively pursued feature of the model. Wright felt that the deleterious effects of inbreeding are offset by the ability of inbreeding to bring to

light yet undiscovered traits. Once this was accomplished, the crossbreeding between demes would 'restore vigor'. In Wright's own words:

> Progress by ordinary selection of individuals would thus be very slow or nil. A single unfortunate selection of a sire, good as an individual, but inferior in heredity, is likely at any time to undo all past progress. On the other hand, by starting a large number of inbred lines, important hereditary differences in these respects are brought clearly to light and fixed. Crosses among these lines ought to give a full recovery of whatever vigor has been lost by inbreeding, and particular crosses may be safely expected to show a combination of desired characters distinctly superior to the original stock. Thus crossbred stock can be developed which can be maintained at a higher level than the original stock, a level which could not have been reached by selection alone[6].

Once again Wright is interested in 'balance'. Instead of trying to enhance variety by discouraging similarity, he uses inbreeding to cause random drift. This promotes exploration. The resulting instability is balanced by integrating the small populations into a large one where selection, with increased variation, can take place. The problem is in regulating the proper balance.

With the modified and abstracted SBT this balance becomes unnecessary, as does the reliance on inbreeding. The exploration is done by the colonies while the exploitation is left to the core. Furthermore, unlike the inbreeding prevention mechanisms incorporated into many EC systems, the wanted variety is produced directly, by selecting for it, and not circuitously, by the prevention of similarity through the reduction of inbreeding.

4.4 Migration

In the original shifting balance theory, migration happened in both directions, from core to colony and from colony to core. The migration from core to colony occurred to isolate members in a new environment so that they may be allowed to vary from the norm that exists in the main population. The back-migration from colony to core then in turn provided the core with the novel members, which, if those members demonstrated that they were better than the members of the original population, would shift the balance toward the new breed. The main population, now with a new composition, sends members out into new environments as new colonies, thus starting the process over again. This two-way migration between core and colony sets up a feedback loop where, with evolution as the driving mechanism, information processing can occur.

However, while the core gains information about the colonies through immigration of colony members, sending members from the core to the colony can actually create problems. Since the core has a very large population, most

of the time it will find far better solutions than a colony. We would expect these solutions, if allowed into the colony, to begin to dominate the colony population within a few generations, forcing the colonies to search in the same area as the core. This would interfere with the distance mechanism that prevents the colony from overlapping the core. Consequently, it may be preferable not to implement forward migration.

If migration from core to colony does not occur, one might think that the colonies, which are now in reproductive isolation, no longer have any knowledge about the core. This is not so. While forward migration does indeed provide information about the location of the core to the colonies, so too does the distance mechanism that prevents the colony from overlapping with the core.

Of course, we are not just trying to render the SBT as faithfully as possible; rather we are trying to abstract from it useful properties and methods to aid the GA when escaping from local maxima. Consequently, we must use procedures that are true to the purpose and goals of the SBT and not just the superficial mechanisms. As discussed previously, the two methods for providing information about the core to the colonies have completely opposite effects on the colony. Forward migration tries to make the colony more like the core. The distancing mechanism tries to make the colony less like the core. Since our purpose in designing the colonies is to promote exploration into areas of the search space where the core has not yet been, forward migration seems a counterproductive approach for providing a colony with information about the location of the core. Consequently, only migration from colony to core was implemented and investigated.

5 Overview of the SBGA

By extending the GA through the addition of the modified SBT mechanisms, we obtain what we call the Shifting Balance Genetic Algorithm (SBGA). An overview of this algorithm can be found in figure 2 (see next page).

In the algorithm we make reference to the distance from a member of a population to the entire population. This is a well-known concept in cluster analysis. In GA terminology such a distance is simply the average Hamming distance between the given member and all other members of the population. This concept is easily extended to encompass the distance from a single member of one population to an entire second population.

The details of the algorithm are in (Wineberg & Oppacher 2000).

6 The Behavior of the SBGA

6.1 Purpose
Wright's SBT has been criticized as untestable because of its overall looseness. One criticism of it is the unsatisfactory specification of the balance between random drift and selection. Another one is his ill-defined notion of 'fitness

---

6   As quoted in (Provine 1986), pg. 156.

```
Randomly initialize all populations
Eval. all populations using the fitness fⁿ
Loop Until (solution found) or (max gen. hit)
  For each colony
    If colony is ready to migrate
    (as determined by the migration interval)
      Extend the core size to hold the immigrants
      Send the colony's migrants to the core
  Create a new core (the size of a regular, unexpanded core)
  Populate the new core from members of the old core using
      reproduction; selection is based on fitness values
  For each colony
    Create a new colony
    Determine the extent to which the old colony's population is
        contained in the new core
    Divide the new colony into 2 section
    The "inside section";
        size = #(old colony members inside core)
    The "outside section"
        size = #(old colony members outside core)
        Populate the "inside" section through reproduction;
        selection is based on distance to the core
        Populate the "outside" section through reproduction;
        selection is based on fitness values
  Evaluate all populations using the fitness function
  Set the core and each colony to their new population
```

Figure 2: The Shifting Balance Genetic Algorithm (SBGA).

landscapes'. A third criticism is that Wright had envisioned the SBT as working on only stationary landscapes, while in biology the landscape is obvious far from stationary.

Our modified version of the SBT addresses these criticisms. The balance problem has been resolved in §4.1. The fuzziness of Wright's 'fitness landscape' is taken care of by the simplicity of the GA fitness functions. Finally, we feel that Wright's SBT will actually work well in a changing environment. Therefore, the SBT has become testable.

In our experimental design we have decided to test the SBGA in both stationary and changing environments.

6.2 Experimental Design

This experiment was done using the F8F2 function, described in (Whitley et al. 1996). This minimization function is non-symmetric, linearly separable, increases in difficulty as the dimension increases, and has a known minimum at (1, 1). The solutions to F8F2 were encoded using Gray coding.

We compared the SBGA to a standard GA on the F8F2 fitness function of dimensionality 5, for 750 generations. For the first 150 generations, F8F2 was held stationary and then the environment was changed for the next 600 generations.

In the experiment, both the GA and SBGA were given the following parameter settings: probability of mutation = 0.006 per bit, and probability of one-point crossover = 0.7. Linear rank selection with elitism was used. The slope was set to be as steep as possible (i.e. with Max = 2.0).
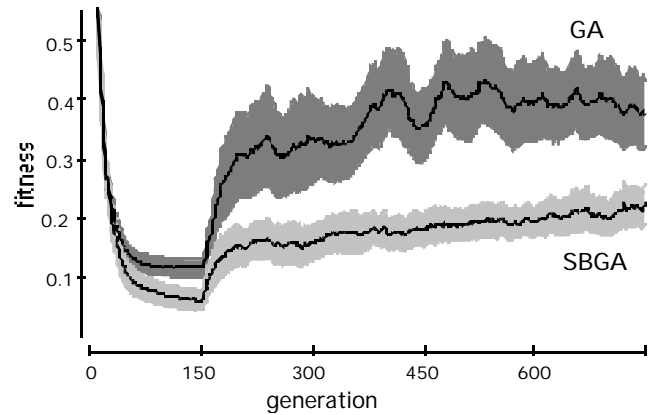

Figure 3: The Fitness of the Best Member Per Generation

The SBGA was given 10 colonies of 100 members each. The colony size was chosen to reduce the amount of random drift, yet keep the colonies relatively small in size. The core group size was set to 1000, an amount equal to the size of all the colonies combined. During migration 25 elite members of a colony are sent to the core group. The immigration interval was set to 5 generations (so in every generation, only 2 colonies are sending members to the core).

The GA was given a population of 2000, equal to the total population of the SBGA system.

The environment undergoes a simple translation in phenotype space along the hyper-diagonal. This motion is kept at a constant speed for 600 generations, at which point the global minimum has moved from the point (1,1,1,1,1) to the point (1.5, 1.5, 1.5, 1.5, 1.5). Since each dimension is encoded as a 12 bit Gray coded integer, the genome and consequently the environment changes, on average, at 0.8333 bits per dimension each generation, which produces a genotypic speed of 4.1667 bits per generation.

The best fitness value of the population is recorded every generation, as is the median of the population and the diversity[7] of the population.

6.3 Results and Discussion

6.3.1   Best of the Generation Response

In the stationary region one expects an evolutionary system to rapidly close in on the global optimum (at the price of decreasing diversity), eventually either finding it, or approximating it. Once the environment changes, the SBGA because of its design should begin to track the global optimum sooner and with a greater accuracy than the GA, which should have experienced an extreme loss of diversity.

Figure 3 shows the fitness of the best individual for each generation for both the GA and the SBGA. Confidence

---

7 The diversity of a population as used here is the average genetic entropy across the population, averaged across all loci, see (Wineberg & Oppacher 1996).

bands are displayed around the mean curves, the SBGA's confidence band is in light gray, the GA's is in dark gray.

Looking at the graph, during the first 150 generations, both the GA and SBGA converge rapidly to very low fitness values (remember F8F2 is a minimization problem) as expected. About generation 60, the two curves begin to diverge; the SBGA begins to outperform the GA. In fact the GA flatlines by generation 75 while the SBGA continues to improve even until the environment starts to move at generation 150.

Again, by looking at the graph after the environment has started to move, one can discern two stages. In the first stage, both systems drastically drop in fitness as a result of having converged. In the second stage, the systems recover somewhat, although they are still losing ground. However the slope of the GA is far steeper than that of the SBGA (until it plateaued) and the GA cannot track the global as closely. The fitness difference between the GA and the SBGA is far greater now than during the stationary stage.

6.3.2   Median Response

Figure 4 shows the median fitness value of the population for every generation (we have only recorded the median of the core population for the SBGA). Unlike figure 3 the 111 median fitness values at a given generation are not averaged, but rather the median is taken because we found that the results were not normally distributed; the distribution curves showed a large one sided tail which heavily skews the results of an average.

Figure 4 shows that the median response of the GA population mirrors the curve for the best fitness found in figure 3. However, this is not the case for the SBGA: during the stationary phase, the fitness remains surprisingly high even though the best of the SBGA is lower than that of the GA. Furthermore, once the environment starts to move, the median fitness actually drops dramatically, even below that of the GA. This counterintuitive behavior (the SBGA seems to be doing worse on 'average' on the easier task) can be
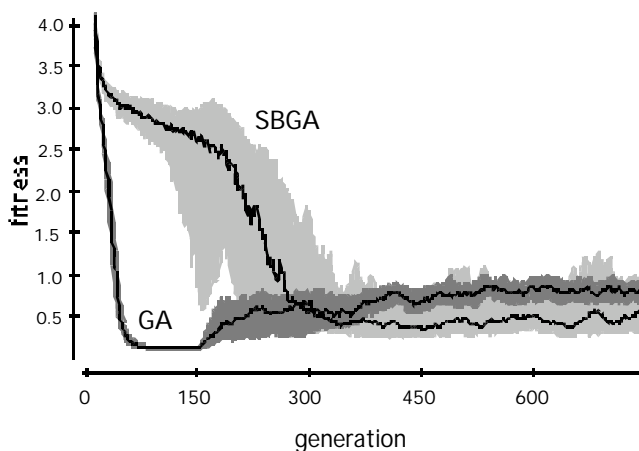

Figure 4: The Median Fitness of the
Core Population Per Generation

explained when one takes the effect of the shifting balance into account.

As the SBGA converges during the stationary phase, the core envelops the region in which the global minimum resides. The colonies, kept away from the core by the distancing mechanism, cannot but help to have poor members. These members are sent into the core, keeping the median very high. Once the environment moves, the core, due to its huge bulk, cannot respond quickly and so its members become very unfit. However, the small, quick moving colonies, which, by being outside the core are most likely already near the new location of the optimum, will send highly fit members back to the core. These members will begin to dominate the core, thus reducing the median.

It is this very fact that shows that the benefits of the shifting balance theory are indeed operating. It also shows that the SBGA thrives in a dynamic environment.

7 Conclusion

Starting with the fertile SBT metaphor, we address some of the main criticisms leveled against it (its loose formulation, possible inconsistencies, and its untestability) by converting it to a precisely formulated and testable theory.

The test results indicate that our version of the SBT, i.e. the SBGA, indeed allows the evolutionary system to escape from local optima. This claim is supported by the improved behavior of the SBGA over the GA in a stationary environment. Furthermore, the SBGA was shown to easily outperform the GA in a changing environment, thus defeating the criticism that Wright's version of the SBT can only cope with static environment. Finally, we were able to show, that the main population does actually shift in gene space as envisaged by Wright.

Bibliography

Cobb, H. G., Grefenstette, J. J. (1993). Genetic algorithms for tracking changing environments. In Forrest, S. (Ed.) *Proceedings of the Fifth International Conference on Genetic Algorithms* (pp. 523-530). San Mateo, California: Morgan Kaufmann.

Cohoon, J., Hegde, U., Martin, W., and Richards, D. (1987). Selection in Massively Parallel Genetic Algorithms. In J.J. Grefenstette (Ed.), *Proceedings of the Second International Conference on Genetic Algorithms* (pp. 148-154). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Collins, R. J., and Jefferson, D. R. (1991). Selection in Massively Parallel Genetic Algorithms. In R. K. Belew, and L. B. Booker (Eds.), *Proceedings of the Fourth International Conference on Genetic Algorithms* (pp. 249-256). San Mateo, California: Morgan Kaufmann.

Davidor, Y. (1991). A Naturally Occurring Niche & Species Phenomenon: The   Model and First Results.   In R. K. Belew, and L. B. Booker (Eds.), *Proceedings of the Fourth*

International Conference on Genetic Algorithms (pp. 257-263). San Mateo, California: Morgan Kaufmann.

De Jong, K. A. (1975). An analysis of the behavior of a class of genetic adaptive systems. Doctoral dissertation, University of Michigan.

Eshelman, L. J. (1991). The CHC Adaptive Search Algorithm: How to Have Safe Search When Engaging in Nontraditional Genetic Recombination. In Rawlins, G. J. E. (Ed.), *Foundations of Genetic Algorithms* (pp. 265-283). San Mateo, California: Morgan Kaufmann.

Goldberg, D. E., Richardson, J. (1987). Genetic algorithms with sharing for multimodal function optimization. In Grefenstette, J. (Ed.), *Genetic Algorithms and their Applications : Proceedings of the Second International Conference on Genetic Algorithms* (pp. 41-49). Hillsdale, NJ: Lawrence Erlbaum Assoc. Inc.

Goldberg, D. E. (1989). Genetic Algorithms in Search, Optimization & Machine Learning. Reading, Massachusetts: Addison-Wesley.

Gorges-Schleuter, M. (1989). ASPARAGOS An Asynchronous Parallel Genetic Optimization Strategy. In J. D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 422-427). San Mateo, California: Morgan Kaufmann.

Greenwood, P. J., Harvey, P. H., Perrins, C. M. (1978). Inbreeding and dispersal in the great tit. *Nature* 271:52-54.

Hartl, Daniel L. and Andrew G. Clark (1997). *Principles of Population Genetics: Third Edition*. Sunderland, Massachusetts: Sinauer Associates.

Hillis, W. D. (1991). Co-Evolving Parasites Improve Simulated Evolution as an Optimization Procedure. In S. Forrest (Ed.), *Emergent Computation: Self-Organizing, Collective, and Cooperative Phenomena in Natural and Artificial Computing Networks* (228-234). Cambridge Massachusetts: MIT Press.

Maresky, J., Davidor, Y., Gitler, D., Aharoni, G., and Barak, A. (1995). Selectively Destructive Re-start. In L. J. Eshelman (Ed.), *Proceedings of the Sixth International Conference on Genetic Algorithms* (pp. 144-150). San Mateo, California: Morgan Kaufmann.

Mühlenbein, H. (1989). Parallel Genetic Algorithms, Population Genetics and Combinatorial Optimization. In J. D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 416-421). San Mateo, California: Morgan Kaufmann.

Pál, K. F. (1994). Selection Schemes with Spatial Isolation for Genetic Optimization. In Y. Davidor, H-P. Schwefel, R. Männer (Eds.) *Parallel Problem Solving from Nature - PPSN III* (pp. 170-179). Berlin: Springer-Verlag.

Paredis, J. (1995). The Symbiotic Evolution of Solutions and their Representations. In L. J. Eshelman (Ed.), *Proceedings of the Sixth International Conference on Genetic Algorithms* (pp. 359-365). San Mateo, California: Morgan Kaufmann.

Pettey, Leuze, & Grefenstette, (1987). A Parallel Genetic Algorithm. In J.J. Grefenstette (Ed.), *Proceedings of the Second International Conference on Genetic Algorithms* (pp. 148-154). Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Potts, W. K., Manning, C. J., Wakeland, E. K. (1991). Mating patterns in seminatural populations of mice influenced by MHC genotype. *Nature* 352:619-621.

Provine, W. B. (1986). *Sewall Wright and Evolutionary Biology*. Chicago Illinois: The University of Chicago Press.

Sarma, J., De Jong, K. (1999). The Behavior of Spatially Distributed Evolutionary Algorithms in Non-Stationary Environments. In Banzhaf, W. et al. (Eds.), *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'99)* (pp. 572-578). San Francisco, California: Morgan Kaufmann.

Spiessens, P., and, Manderick, B. (1991). A Massively Parallel Genetic Algorithm. In R. K. Belew, and L. B. Booker (Eds.), *Proceedings of the Fourth International Conference on Genetic Algorithms* (pp. 279-286). San Mateo, California: Morgan Kaufmann.

Sumida, B. H., and Hamilton, W. D. (1994). Both Wrightian and 'Parasite' Peak Shifts Enhance Genetic Algorithm Performance in the Traveling Salesman Problem. In R. Paton (Ed.), *Computing With Biological Metaphors* (pp. 264-279). London U.K.: Chapman & Hall.

Tanese, R. (1989). Distributed Genetic Algorithms. In J. D. Schaffer (Ed.), *Proceedings of the Third International Conference on Genetic Algorithms* (pp. 434-439). San Mateo, California: Morgan Kaufmann.

Whitley, D., Mathias, K., Fitzhorn, P. (1991). Delta Coding: An Iterative Search Strategy for Genetic Algorithms. In Belew R. K., Booker L. B. (Eds.), *Proceedings of the Fourth International Conference on Genetic Algorithms* (pp. 77-84). San Mateo, California: Morgan Kaufmann.

Whitley, D., Mathias, K., Rana, S., and Dzubera, J. (1996). Building Better Test Functions. In L. J. Eshelman (Ed.), *Proceedings of the Sixth International Conference on Genetic Algorithms* (pp. 239-246). San Mateo, California: Morgan Kaufmann.

Wineberg, M., and Oppacher, F. (1996). The Benefits of Computing with Ißntrons. In: J.K. Koza, et al. *Genetic Programming 1996: Proceedings of the First Annual Conference* (410-415). Cambridge, MA: MIT Press.

Wineberg, M., and Oppacher, F. (2000). Enhancing the GA's Ability to Cope with Dynamic Environments. Submitted to GECCO'2000.

Wright (1932). The Roles of Mutation, Inbreeding, Crossbreeding and Selection in Evolution. In, W. B. Provine (Ed.), *Sewall Wright Evolution; Selected Papers* (pp. 161-171). Chicago, Ill.: University of Chicago Press, 1986.