Introduction.

Evolutionary methods are exceedingly popular with practitioners of many fields; more so than perhaps any optimization tool in existence. Historically Genetic Algorithms (GAs) led the way in practitioner popularity (Reeves 1997). However, in the last ten years Evolutionary Strategies (ESs) and Evolutionary Programs (EPS) have gained a significant foothold (Glover 1998). One partial explanation for this shift is the interest in using GAs to solve continuous optimization problems. The typical GA relies upon a cumbersome binary representation of the design variables. An ES or EP, however, works directly with the real-valued design variables. For detailed references on evolutionary methods in general and ES or EP in specific see Back (1996) and Dasgupta and Michalecisz (1997). We call our evolutionary algorithm BCB (bell curve based) since it is based upon two normal distributions.

BCB for continuous optimization, first presented in Sobieszczanski-Sobieski et al. (1998), is similar in spirit to ESs and EPs but has fewer parameters to adjust. A new generation in BCB is selected exactly the same as a $(\mu + \lambda)$-ES with $\lambda = \mu$. That is, the best $\mu$ individuals out of $\mu$ parents plus $\lambda$ children are selected for the next generation. Thus fit individuals may continue from one generation to the next. The recombination and mutation mechanisms are illustrated in Figure 1. Consider the line through two $n$-dimensional parent vectors $\vec{P}_1$ and $\vec{P}_2$ selected for mating. First, determine the weighted mean $\vec{M}$ of these two vectors where the weights are given by the fitness (KS value) of each parent. Next, sample from a normal distribution $N(0, \sigma_m)$. The resulting point $\vec{B} = \vec{M} + |\vec{P}_2 - \vec{P}_1| \ast N(0, \sigma_m)$ is the child, prior to mutation. Note that $\vec{B}$ is not restricted to lie on the line segment $\vec{P}_1 \vec{P}_2$. Mutation ensues by first generating a radius $r$ for an $n - 1$ dimensional hypersphere. The radius is a realization from a $N(0, \sigma_r)$. Typically $\sigma_r \gg \sigma_m$. Finally the mutated child $\vec{C}$ is selected by sampling uniformly on the surface of the
$n - 1$ dimensional hypersphere. Hence, there are two parameters $\sigma_r$ and $\sigma_m$ in addition to the traditional parameters of population size and number of generations.

![Figure 1. BCB Geometrical Construct in 3D Space](image)

**Summary of Research.**

The research effort over the grant period has resulted in two manuscripts—Kincaid et al. (2000a) and Kincaid et al. (2000b). Each of these manuscripts is available to the public at [http://www.math.wm.edu/rrkinc/heuristic.html](http://www.math.wm.edu/rrkinc/heuristic.html). In Kincaid et al. (2000a) we presented an example illustrating the ability of BCB to escape local optima. Following the example we examined the effect of parameter changes on the performance of BCB. Three rules of thumb were developed. First, we found that it was crucial to scale the values of the decision variables. Second, we found that, in general, $\sigma_r >> \sigma_m$ is best. This
is partially due to the influence of the Euclidean distance between parents on the placement of the center of the $n-1$ dimensional hypersphere in addition to $\sigma_m$. Third, we found that the performance of BCB is more sensitive to the value of penalty-2 than for penalty-1. Following the computational experiments of the parameters of BCB we explained and tested two variations of BCB. Both variations seek to identify clusters and outliers in a given population. It is clear that the extra work involved in identifying clusters and outliers is not beneficial in the serial version of BCB but provides some impetus for the development of a parallel version of BCB.

In Kincaid et al. (2000b) we demonstrated that BCB is similar in spirit to $(\mu + \mu)$ evolutionary strategies and evolutionary programs but with fewer parameters to adjust and no mechanism for self adaptation. The performance of BCB is shown to dominate a Classical Evolutionary Program (CEP) but not an improved Fast Evolutionary Program (FEP) on several standard test functions (see Yao 1999 for details of CEP and FEP). BCB does significantly better than FEP for unimodal test functions but loses out to FEP on three of the four multi-modal test functions. Rather than attempt to find ways to improve the performance of BCB we examined the utility of coupling BCB with local search procedures (both gradient based and non-gradient based). Here BCB's role is to identify high quality basins as opposed to determining high quality solutions.

We tested two couplings of BCB with local search heuristics. The first links BCB with a standard quasi-newton (gradient based) search—BCB-GS. The second links BCB with a pattern search—BCB-PS. We selected a pattern search based on the Hooke-Jeeves algorithm, whose full description can be found in Box et al. (1969). To summarize, the algorithm proceeds by performing several iterations of coordinate searches; that is, from a base solution, the points that lie one step size away along the coordinate vectors are tested for improvement. A list of temporary and permanent base points are maintained. When a coordinate search around a permanent base point fails to find improvement, the step size is halved, and another coordinate search is performed there. Otherwise, if improvement is found, a new temporary base point is constructed and used as the base of a coordinate search. If no improvement is found around the temporary base point, the search backtracks to the last permanent base point. Otherwise, if improvement is found around the temporary base point, the permanent base point is updated, and a new coordinate search starts there. A up to date assessment of the utility of pattern searches can be found in Lewis et al. (2000).
The BCB-GS and BCB-PS are promising techniques. Dramatic improvements, both in terms of solution quality and number of function evaluations, were obtained for three of the multimodal functions by starting with BCB and switching to either GS or PS both in terms of solution quality and number of function evaluations. The local optimization techniques for nonlinear optimization problems are quite mature while the pattern search scheme has good mathematical underpinnings and is easily ported to a parallel computing environment.

References.