SURROGATE-ASSISTED COEVOLUTIONARY SEARCH

YewSoon Ong, Andy J. Keane and Prasanth B. Nair

Computational Engineering and Design Centre
University of Southampton
Highfield, Southampton SO17 1BJ, United Kingdom
E-Mail: {Y.S.Ong, Andy.Keane, P.B.Nair}@soton.ac.uk

ABSTRACT

This paper is concerned with an experimental evaluation of coevolutionary optimization techniques, which are integrated with surrogate models of the fitness function. The motivation for this study arises from the fact that since coevolutionary search is based on the divide-and-conquer paradigm, it may be possible to circumvent the *curse of dimensionality* inherent in surrogate modeling techniques such as radial basis networks. We investigate the applicability of the algorithms presented in this paper to solve computationally expensive optimization problems on a limited computational budget via studies on a benchmark test function and a real world two-dimensional cantilevered space structure design problem. We show that by employing approximate models for the fitness, it becomes possible to converge to good solutions even for functions with a high degree of epistasis.

1. INTRODUCTION

In recent years, coevolutionary computation has been applied with a great degree of success to function optimization, neural network training, and concept learning [1,2]. Coevolutionary computation can be interpreted as an instance of the divide-and-conquer paradigm. For example, in the context of optimization, the variables in the original problem are decomposed into a number of subsets. Subsequently, species which independently handle each subset of variables are evolved simultaneously to locate the optima of the original problem; see Figure 1 for an illustration of how coevolutionary computation can be applied to function optimization. An important advantage of coevolutionary optimization techniques is their ability to converge faster to the optima as compared to standard evolutionary algorithms, particularly for problems with low epistasis.

In this paper, our focus is on optimization of computationally expensive functions. In recent years, many researchers have examined strategies for integrating surrogate models with evolutionary search techniques to tackle the computational cost issue associated with this class of problems [3-6]. Surrogate models are essentially metamodels of the original objective function, which are constructed using techniques in the machine learning literature such as neural networks and radial basis networks [7]. Even though promising results have been obtained for some problems, the curse of dimensionality poses a major obstacle to the successful application of surrogate modeling techniques to multimodal functions with large number of variables.

In this paper, we address the issue of how surrogate models can be integrated with coevolutionary optimization algorithms. The motivation for this study arises from the observation that since coevolutionary algorithms decompose the original optimization problem into subproblems with smaller number of variables, it may become possible to circumvent the curse of dimensionality in the surrogate model construction phase. However, while this divide-and-conquer approach enables us to tackle the curse of dimensionality, epistatic interactions between the variables can lead to a significant degradation of the convergence rate. The experimental studies by Potter [8] suggest that the convergence of coevolutionary techniques can be significantly slower than standard evolutionary algorithms for problems with high degrees of epistasis.

Currently, little work is available in the literature on the convergence of coevolutionary optimization when approximate fitness models are employed during the search. Even though, some studies were presented in Nair and Keane [9] for a space structure design problem using domain-specific approximation models, it is not clear whether the observations made there carry over to cases when more general surrogate modeling frameworks are employed. Here, we first examine the convergence behavior of surrogate-assisted coevolutionary optimization via detailed experimental studies on a benchmark function with random degrees of epistasis. Our results indicate that the noise in the fitness evaluations due to the use of surrogate models can cancel out the detrimental effects of epistasis on the convergence of coevolutionary optimization. Further, we proceed to show that the proposed general surrogate modeling framework arrives at better space structure designs over the time budget allocated for optimization when compared to conventional evolutionary searches.

```
BEGIN
 Initialize:
        Generate a population of individuals for s species.
 While (computational budget not exhausted)
    For species i = 1 to s
        Choose representatives from all the other species. The
        elite member of each species is often used.
        For each individual j in population i
          • Form collaboration between individual j with the
            representatives from other species.
          • Evaluate the new individual by applying it to the
            target problem and reward it with the resulting
            fitness value.
          • Implement Elitism.
        End For
        Apply standard EA operators to create a new
        population for species i.
    End For
 End While
END
```

Figure 1: Coevolutionary Optimization.

2. SURROGATE-ASSISTED COEVOLUTIONARY OPTIMIZATION

First, we outline an approach for integrating surrogate models with coevolutionary optimization algorithms. For simplicity of presentation, consider the optimization problem with simple bound constraints given by:

Minimize: $f(x)$

subject to: $x_l \leq x \leq x_u$

where $x \in \Re^n$ is the vector of design variables, and x_l and x_u are the lower and upper bounds, respectively.

We are concerned here with problems where evaluation of $f(x)$ is computationally expensive and it is desired to converge close to the global optima on a limited computational budget. To reduce the computational cost, we seek to interleave a surrogate model for $f(x)$ along with the exact function in the coevolutionary search procedure. In particular, we consider the case when a surrogate model of the objective function is constructed using radial basis functions (RBF). Here, the RBF approximation for the original objective can be written as

Equation 1: $$y = \sum_{i=1}^{m} \alpha_i K(x, x_i) \qquad (1)$$

where $K(x, x_i)$ is a positive semi-definite kernel, α_i denotes the vector of weights, and $\{x_i, y_i, i=1,2,...m\}$ is the training dataset. Typical choices for the kernel $K(..,..)$ include linear splines, cubic splines, multiquadrics, thin-plate splines, and Gaussian functions [7]. m is the number of centers or the number of radial basis functions employed. The centers of the radial basis functions are chosen via the *k-means* clustering algorithm. Subsequently, the exact values of $f(x)$ are evaluated at these centers to generate training data for constructing a RBF approximation. Models are constructed on the fly at each ecosystem generation independently for each species. Note that by dividing the original problem variables among multiple species, we have reduced the number of inputs in the surrogate models since each species handles only a subset of the original variables. This procedure thus tackles the curse of dimensionality, which often limits the success of surrogate modeling on multimodal problems with many variables. The steps involved in the proposed surrogate-assisted coevolutionary optimization algorithm are outlined in Figure 2.

```
BEGIN
 Initialize:
        Generate a population of individuals for s species.
        Set fitness function:= Surrogate for all species.
 While (computational budget not exhausted)
    For species i = 1 to s
        Choose representatives from all the other species. The
        elite member of each species is used.
      If fitness function_i == Surrogate
        • Decompose the design subspace into m cluster
          centers using the k-means algorithm.
        • Form collaboration between cluster centers with the
          representatives from other species and evaluate
          them using the exact analysis model.
        • Build Surrogate_i based on the m exact points.
        For each individual j in population i
          • Form collaboration between individual j with the
            representatives from other species.
          • Evaluate new individual j using Surrogate_i.
        End For
      Else
        For each individual j in population i
          • Form collaboration between individual j with the
            representatives from other species.
          • Evaluate new individual j using the exact model.
        End For
      End If
      If (Surrogate Stalls)
        fitness function_i := Exact Model
      Else
        fitness function_i := Surrogate
      End If
        Implement Elitism.
        Apply standard EA operators to create a new
        population for species i.
    End For
 End While
END
```

Figure 2: Surrogate Assisted Coevolutionary Optimization

In the first step, we initialize a population of designs for s species either randomly or using design of experiments techniques such as Latin hypercube sampling. The search space of each species is then decomposed into clusters using the *k-means* algorithm and augmented with the values of representatives from the other species before fitness evaluations based on the exact analysis model is conducted. Here, the elite member of each species is chosen as the representative for the other species. We then use a linear spline RBF to construct the surrogate model since this approximation technique is capable of providing surrogate models with good generalization capability at a low computational cost. Subsequently, the surrogate of each species constructed using their cluster center members as training data are used to evaluate the individuals. The exact fitness of the elite individual from each species is evaluated to check whether any improvement in the actual fitness is achieved. This enables us to switch to the original exact objective function when the coevolutionary search on the surrogate stalls. Standard EA operators then proceed to create a new population with elitism to prevent loss of the fittest design in each species. This process of coevolutionary optimization is continued until a specified termination criterion is met.

Our algorithm has two user specified parameters – (1) number of species (s) and (2) number of cluster centers (m). Note that the accuracy of the surrogate model can be improved by increasing m. In the limiting case, when m equals the population size for a species, the fitness of all the individuals is evaluated exactly. Similarly when s is unity the process becomes simple evolutionary algorithm.

3. EXPERIMENTAL STUDIES USING BENCHMARK FUNCTION

In this section, we present experimental studies obtained by implementing the coevolutionary genetic algorithm (CGA) and the proposed Surrogate Assisted Coevolutionary Optimization (SCGA) within a real-coded genetic algorithm (GA) for evolutionary search. In the standard GA we have employed population size and truncation threshold of 50 and 0.8, respectively. Extended linear recombination and non-uniform mutation are used in the reproduction [10].

3.1. Rastrigin Test Function

The Rastrigin test function adopted for study is defined as:

Minimize: $$F_{Rastrigin} = (10 * n) + \sum_{i=1}^{n} \left(x_i^2 - 10\cos(2\pi x_i)\right)$$

for $-5.12 <= x_i <= 5.12, \quad i = 1, \ldots, n.$

A twenty dimensional ($n=20$) version of the function is used here. The function is highly multimodal, having many local minima surrounding the global minimum at zero. It is a separable function. A function of n variables is *separable* when there are no variables interdependencies or linkages between the variables. In the GA literature, this implies no epistasis interactions. This separability can be varied, however, by a simple rotation of the coordinate system [11]. Note that such rotation does not change the function's structure but only the epistatsis interactions.

3.2. Empirical Results

The average convergence trends of the standard GA, CGA and proposed SCGA when applied to the 20-variable Rastrigin function without any rotation are summarized in Figures 3-5.

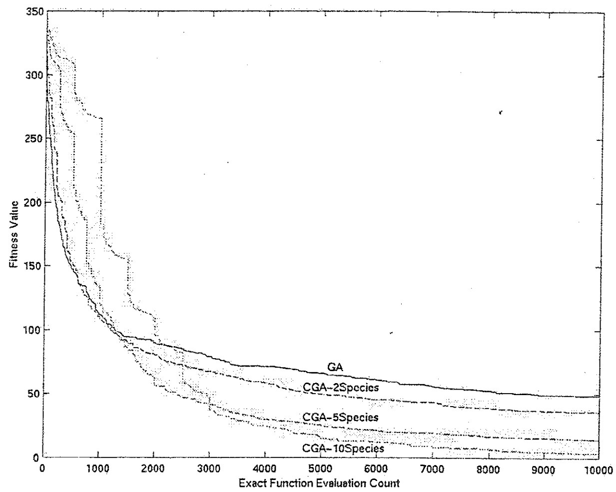

Figure 3: Convergence trends of GA and CGA with various species sizes of 2, 5 and 10 when applied on the Un-Rotated 20-Variable Rastrigin function.

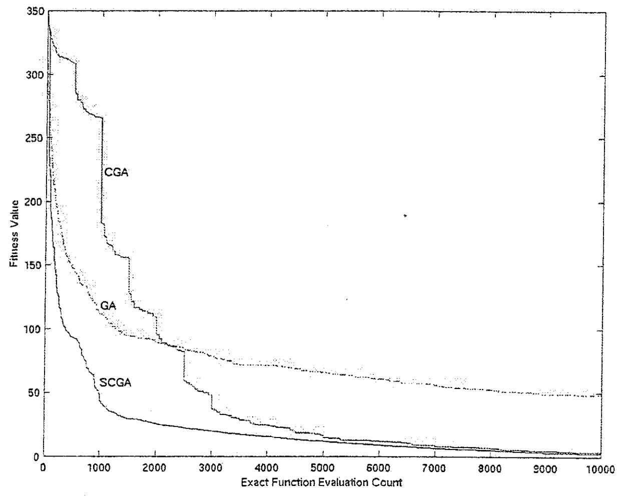

Figure 4: Convergence trends of GA, CGA with species sizes s=10 and SCGA with s=10 and m=5 when applied on the Un-Rotated 20-Variable Rastrigin function.

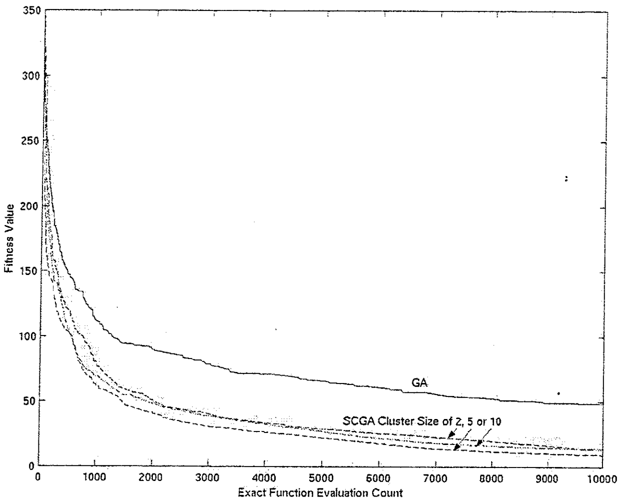

Figure 5: Convergence trends of GA and SCGA with *s*=10 and various cluster center size *m*=2,5 or 10 when applied on the Un-Rotated 20-Variable Rastrigin function.

All results presented were averaged over 20 runs. As shown in Figure 3, the search performance of the CGA is superior to the standard GA on this un-rotated function. The increase in number of species of the CGA is observed to result in better results. This however has the effect of slowing down the CGA's initial period of adaptation. The performance of the proposed SCGA algorithm may be established by comparison with the use of the exact analysis model in both the standard GA and CGA search. In figure 4, the results of the GA, CGA for *s*=10 and SCGA for *s*=10 and *m*=5 when applied on the un-rotated Rastrigin are compared. We see that the SCGA has brought about significant improvements in performance over both the standard GA and CGA. Note that here some Rastrigin surrogate calls are used alongside each exact function call.

We also studied the effect of varying the number of cluster centers, *m* (used to construct the surrogate model), and the species size, *s*, on the convergence behavior of SCGA. In common with the convergence behavior of CGA, increasing the number of species improves the SCGA convergence. A number of observations on varying *m* can be made from the results obtained in Figure 5. At first, it appears that the convergence rate of SCGA improves slightly with increasing *m*. This is because increasing *m* improves the accuracy of the surrogate model. However, it appears that further increase of *m* (to 10) slows down the SCGA's rate of convergence. In the limiting case, when *m* equals the population size for a species, both the CGA and SCGA would convergence similarly, thus losing the purpose of using surrogates. Nevertheless, it appears that employing small *m* seems appropriate considering that the dimensionality of each surrogate is generally low due to problem decomposition in SCGA. Moreover the gain in convergence rate obtained by increasing *m* does not appear to be very significant.

Subsequently, to determine how well the proposed SCGA can cope with problem of high epistasis, we alter the separability of the Rastrigin function using the coordinate system rotation discussed in section 3.1. Here, we use the random rotation algorithm proposed by Salomon [11] to create random epistasis interactions in the test function. The average convergence trends of the GA, CGA and SCGA when applied on the rotated 20-variable Rastrigin function are obtained in Figures 6-8. In Figure 6, we see that the CGA actually performs much worse than the standard GA on the rotated Rastrigin. The significant performance degradation is essentially what was expected and as described by Potter [8]. Compared to the CGA, the standard GA is less susceptible to the effect of the induced epistasis because the latter is able to perform changes on greater number of variables simultaneously.

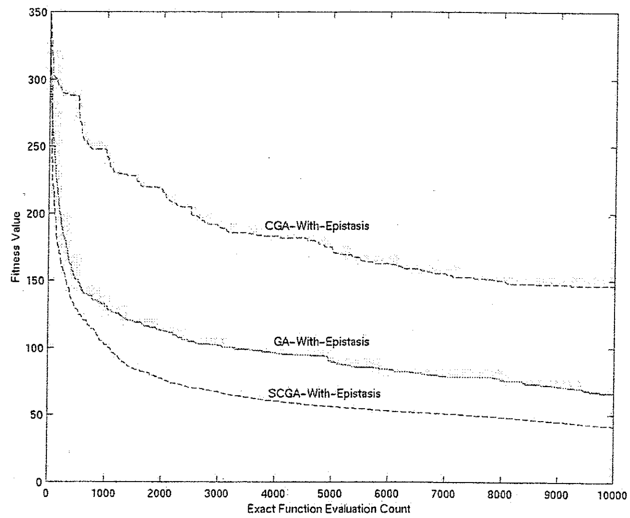

Figure 6: Convergence trends of GA, CGA with species sizes *s*=10 and SCGA with *s*=10 and *m*=2 when applied on the Rotated 20-Variable Rastrigin function.

On the other hand, the effect of surrogates or fitness uncertainties on the convergence of CGAs is an area that has not yet been fully investigated. Although preliminary studies by Potter indicate that CGAs are more sensitive to noise as compared to the standard GA, it is interesting to note the SCGA actually performs much better than both the GA and CGA on the rotated Rastrigin function. This surprising result appears to indicate that the errors in the fitness evaluations apparently circumvent the convergence difficulties encountered when coevolutionary optimization techniques are applied to functions with high epistasis. The effects of the cluster center size, *m* and species size, *s*, on the convergence behavior of SCGA on the rotated function are shown in Figures 7 and 8.

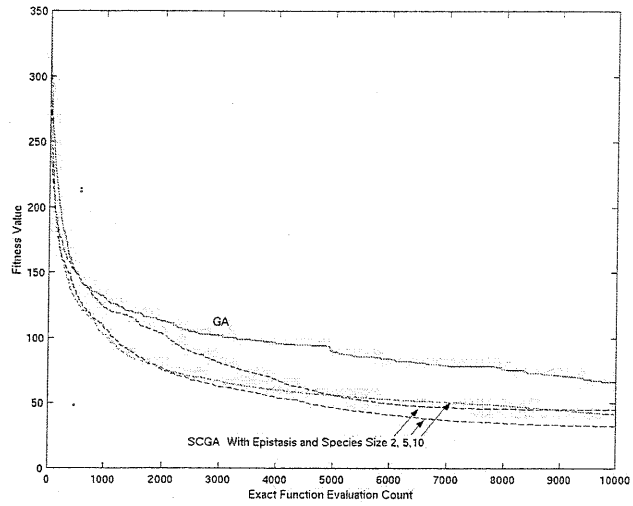

Figure 7: Convergence trends of GA and SCGA for s=2,5 or 10 and m=2 when applied on the Rotated 20-Variable Rastrigin function.

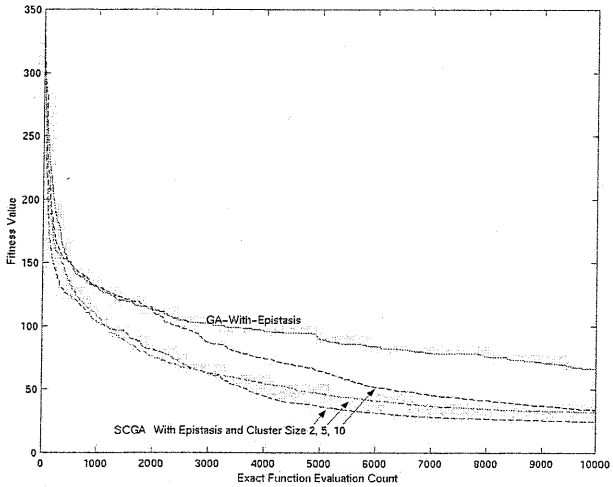

Figure 8: Convergence trends of GA and SCGA with s=5 and m=2,5 or 10 when applied on the Rotated 20-Variable Rastrigin function.

Unlike on the un-rotated Rastrigin, the performance of the SCGA on the rotated function degrades when species size is increased to 10, see Figure 7. In addition, the convergence behaviours for using different cluster center sizes in the SCGA (see Figure 8) appears to indicate that higher degrees of errors in the surrogate enable the SCGA to better circumvent the convergence difficulties associated with this high epistasis function.

4. TWO-DIMENSIONAL CANTILEVERED SPACE STRUCTURE DESIGN

In this section, the proposed Surrogate Assisted Coevolutionary Optimization algorithm is applied to the design of flexible space structures with non-periodic geometries to achieve passive vibration. The space structure considered in this paper is a two-dimensional cantilevered structure shown in Figure 9, subjected to transverse excitation at joint F near the fixed end. The objective of the design problem considered here is to suppress the vibration response at joint R over the frequency region of 100-200Hz. This isolates any instrumentation package mounted at joint R on the space structure from external vibrations arising in the main body of the satellite.

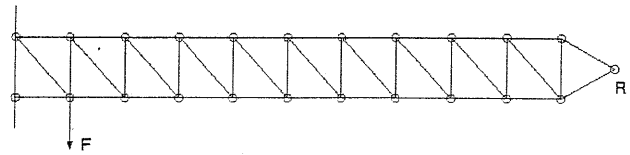

Figure 9: Two-Dimensional Space Structure.

The objective function is defined as the integral of the frequency responses at joint R. The performance measure of a candidate design is defined as:

$$J = -20\log_{10}\left\{\frac{1}{I_o}\int_{100}^{200}\left[u_R(w)+v_R(w)+L\theta_R(w)\right]dw\right\} \quad (2)$$

where u_R , v_R, and θ_R are the axial, transverse, and rotational components of the displacement response at joint R, respectively. I_0 is the integral of the frequency responses at R for the baseline periodic structure shown in Figure 9. The first 100 modes are used to compute the dynamic response of the structure in the region of 100–200Hz, with the integral computed at a resolution of 2.7 Hz. The design is parameterized in terms of the coordinates of the structural joints, which are allowed to vary between ±0.25 m from the baseline values, with the coordinates of joint R being kept fixed. This leads to a nonlinear multi-modal design problem of 40 geometric design variables with high epistatic linkages among some of the variables. Here, a finite element method is used to compute the free-vibration natural frequencies and mode shapes of the structure. The exact analysis method takes about 100 seconds to compute each objective function. In comparison, each RBF approximation takes less than a fraction of a second. These numerical studies were conducted on one processor of an SGI Power Challenge machine. For greater details of the space structure considered here in this paper, the reader is referred to [9].

Figure 10 shows the average convergence trends of the conventional GA, CGA and SCGA (for s=2 and m=15) over four runs as a function of number of exact analysis. Here, all the control parameters were set the same as in the previous experiment, but with stopping criteria at 5000 exact analyses. Therefore, each run took almost six days of computer time on a single processor. On this real world design problem, the results clearly indicate that the proposed SCGA arrives at a better design as compared to both the conventional GA and CGA when a constraint is imposed on the computational budget available for optimization. The optimized space structure using SCGA under limited computational budget is revealed in Figure 11.

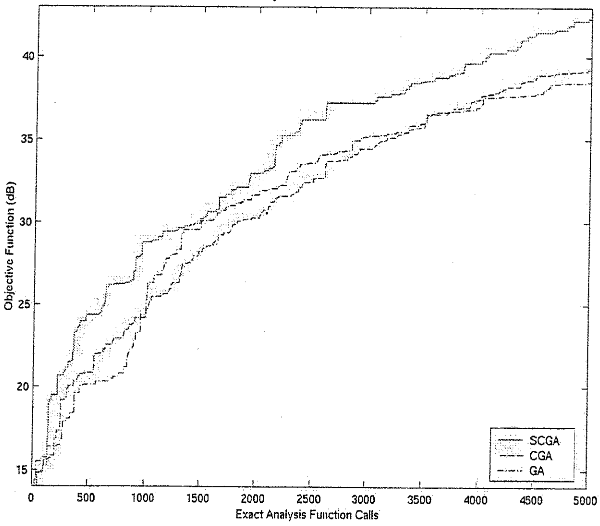

Figure 10: Convergence trends of GA, CGA and SCGA with s=2 and m=15 when applied for the design of the two-dimensional non-periodic cantilevered space structure to achieve passive vibration suppression.

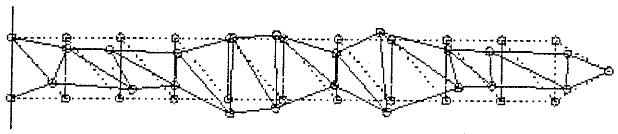

Figure 11: Optimized Structure Using SCGA.

5. CONCLUDING REMARKS

In this paper, we present an algorithm for integrating surrogate models with coevolutionary search procedures. Experimental studies are presented for the Rastrigin benchmark test function with varying degree of epistasis. In addition, the proposed algorithm is also applied to the optimal design of a flexible space structure to achieve passive vibration suppression. The empirical results were compared with those obtained using a conventional Genetic Algorithm and a coevolutionary Genetic Algorithm. This suggest that the surrogate-assisted coevolutionary search is capable of solving computationally expensive optimization problems with varying degrees of epistasis more efficiently than both the conventional GA and CGA under a limited computational budget. Nevertheless, there is evidence from the experiments conducted to warrant additional studies on adapting the two control parameters, number of species (s) and number of cluster centers (m) as well as reducing further the detrimental effects of epistasis on the performance of surrogate-assisted coevolutionary search. One possible way of doing so may involve adaptive linkage identification and decomposition [12]. Besides the degree of epistasis, it would be important to also identify the design problem and epistasis structure.

6. REFERENCES

[1] M. Potter and K.D. Jong, "A cooperative coevolutionary approach to function optimization", In Parallel Problem Solving from Nature - 3, pp. 249--257, 1994.

[2] M. Potter and K.D. Jong, "Evolving neural networks with collaborative species", In Proc. of Summer Computer Simulation Conference, pp. 340-345, 1995.

[3] A. Ratle, "Accelerating the convergence of evolutionary algorithms by fitness landscape approximation", Parallel Problem Solving from Nature V, pp. 87-96, 1998.

[4] M.A. El-Beltagy, P.B. Nair, and A.J. Keane, "Metamodelling Techniques For Evolutionary Optimization of Computationally Expensive Problems: Promises and Limitations",Proc. of the Genetic and Evolutionary Computation Conference (GECCO99), Morgan Kaufman, pp. 196-203, 1999.

[5] K.H. Liang, X. Yao and C. Newton, "Evolutionary search of approximated N-dimensional landscapes", International Journal of Knowledge-Based Intelligent Engineering Systems, 4(3):172-183, July 2000.

[6] Y.S. Ong, P.B. Nair, A.J. Keane, "Evolutionary Optimization of Computationally Expensive Problems via Surrogate Modeling", AIAA Journal, In Communication, 2002.

[7] C. Bishop, Neural Networks for Pattern Recognition, Oxford University Press, 1995.

[8] M. Potter, The Design and Analysis of a Computational Model of Cooperative Coevolution. Ph.D. thesis, George Mason University, Fairfax, VA, 1997.

[9] P.B. Nair, and A.J. Keane, "Passive Vibration Suppression of Flexible Space Structures via Optimal Geometric Redesign," AIAA Journal, 39(7), pp. 1338-1346, 2001.

[10] F. Herrera, M. Lozano and J.L. Verdegay, "Tackling real-coded genetic algorithms: Operators and tools for behavioral analysis", Artificial Intelligence Review 12, pp. 265-319, 1998.

[11] R. Salomon, "Reevaluating Genetic Algorithm Performance under Coordinate Rotation of Benchmark Functions; A survey of some theoretical and practical aspects of genetic algorithms". BioSystems, 39(3):263-278, Elsevier Science, 1996.

[12] K. Weicker and N. Weicker, "On the Improvement of Coevolutionary Optimizers by learning variable", In 1999 Congress on Evolutionary Computation, pp. 1627-1632, Piscataway, NJ, IEEE Service Center, 1999.