Proceedings of Adaptive Computing in Engineering Design and Control,
PEDC University of Plymouth, 21-22 September 1994

THE DEVELOPMENT OF A DIRECTED GENETIC SEARCH TECHNIQUE FOR HEAVILY CONSTRAINED DESIGN SPACES

I. C. Parmee* and G. Purchase**
*Plymouth Engineering Design Centre, University of Plymouth, UK.
email: ian@uk.ac.plym.cis
**Rolls Royce plc Filton, Bristol, UK.

ABSTRACT

The paper describes a genetic search methodology that allows the engineering designer to define a specific point within a complex design space and establish a search neighbourhood around that point. The technique utilises Schaffer's Vector Evaluated Genetic Algorithm to determine the parameter set / sets which describe a specific design solution.

This approach is then utilised as the initial activity within a tiered search process which, having initially defined a feasible point within a constrained design space, will generate an optimal hypercube around that point. The fitness of such a hypercube is defined by the degree of constraint violation measured at specific coordinates of that cube and the extent of the feasible subset thus created.

The engineering design relevance of these two approaches to global search / optimisation is explained and preliminary results are presented.

INTRODUCTION

A primary characteristic of a GA search is the algorithm's ability to sample widely varying areas of an extensive, high dimensional design space (1). In many cases this characteristic has the potential to offer the engineering designer a high level of decision support by rapidly providing a number of diverse, high performance solutions to any one design problem. This allows the exploration of regions that may offer innovative yet practical design solutions. Such solutions may otherwise be innaccessible within an engineer's design time and budget constraints (2). However, this approach will not be appropriate in all cases. In many situations decisions concerning certain design aspects will already be in place. For instance, the use of a specific material may be desirable to satisfy financial and manufacturing criteria. This upstream choice may limit the search process to specific regions of the overall design space and the bounds of these regions may be largely unknown.

Another consideration that must be taken into account is the reliability of a new design. Although a new product should exhibit improved performance at a competitive cost its reliability should be at least equal to the proven reliability of the previous model. In order to establish customer confidence in this respect the new design will be based to a large extent upon that of the previous model. Room for innovation may be restricted to slight improvement of each of the many components and systems that constitute the whole design. The designer may therefore require a technique that will direct the search to a specific point within a component / system design space before establishing a subset of the space around that specific point. A secondary search can then be implemented within this subset to locate the local optimum solution.

This ability to fix a specific point within n dimensional space can also be utilised to access feasible regions of a heavily constrained design space. Feasible values of the constrained parameters become the individual criteria of a multi-objective problem. The VEGA (3) solution of this problem provides the values of the related variables. A GA search can then be initiated to establish the optimal bounds of a hypercube about the point described by the variables. The optimal bounds are defined as the coordinates of those hyperplanes which best describe the extent of the feasible region. A final tertiary search is then initiated within the hypercube to locate an optimal design solution.

The development of this approach has stemmed from a dissatisfaction with the utilisation of penalty functions to draw a GA search process into feasible regions of a constrained space. The introduction of such functions to a design space can result in significant distortion of the fitness surface resulting in the generation of complex characteristics that require a deal of GA expertise to negotiate with any degree of success. Computational expense can become excessive as the chromosonal structures required to create a path to the feasible regions are constantly rejected.and recreated .

THE DESIGN DOMAIN

A high level of industrial collaboration is maintained at the Plymouth Engineering Design Centre (PEDC). The work described in the paper relates to one particular project involving the optimal design of gas turbine engine cooling hole geometries. This work is being carried out in association with Rolls Royce plc, UK.

The cooling hole geometry problem has involved the development of a simple, one-dimensional, steady-state model of a cooling passage situated at the leading edge of a high pressure gas turbine blade. The model represents a purely convective radial passage and provides a basis for extension to cover multiple passages and the inclusion of more comprehensive features and inter-dependencies. A constant cross-section is assumed from passage inlet to exit and the thermal design section is located at the aerofoil mid-span.

Typical variables include the coolant temperature at the passage inlet (TCI), the ratio of the passage inlet pressure to passage outlet pressure (RAT) and the ratio of the cooling passage perimeter to the effective gas-side perimeter of the blade (SGC). Various configurations (IGEOM) to promote mixing within the coolant passage have been included and these introduce variation to a heat transfer coefficient factor and coefficient of discharge. In order to introduce a variation in blade material and to investigate the significance of the application of insulative coatings, wall thickness (DTH) and thermal conductivity (KW) have also been included as independent variables. The primary objective function has been to minimize the coolant flow rate.

For the purpose of this work the initial model has been utilised to provide a practical fitness function that relates real engineering objectives to the search capabilities of the Genetic Algorithm at an acceptable level of computational expense. This allows extensive experimentation to take place which benefits from design-oriented analysis. The model therefore provides a compromise between the rapid processing of some mathematical function with little relation to any physical problem and a computationally intensive dynamic analysis that provides definitive information but severely restricts experimentation due to associated run time.

The initial model has been utilised in this manner for the development of the generic GA design strategies described in the following sections. It is envisaged that the directed search technique for constrained spaces will enable the achievement of optimal solutions from significantly more complex problems characterised by non-linearity of both function and constraints. The cooling hole problem is likely to present such complexity as the breadth and dimensionality of the model is increased. The recent addition of film cooling passages and flight cycle considerations is typical of the continuing model expansion. The development of an effective constrained search technique is therefore seen as a necessity to enable further progression of this and other projects within the PEDC.

In an unconstrained form the search space defined by the simple initial model presents little difficulty to the Genetic Algorithm. Although a degree of discontinuity is created by the varying internal configurations of the cooling passage, appropriate model coding ensures that such obstacles are easily negotiated by a simple GA. A global minimum cooling flow rate can be achieved within thirty five generations. The situation changes however with the introduction of metal temperature and cooling hole area constraints. The utilisation of appropriate penalty functions to draw the GA search into feasible space will result in a near optimal solution but at a significantly increased computational expense (>200 generations).

Fixing a Specific Point

In order to identify the parameter combination that defines a required solution a multi-criteria approach has been utilised. This enables the engineer to direct the search process to areas of particular interest. For example, in the case of the cooling hole geometry problem the following returned values from the mathematical model are particularly significant:

- **Metal temperature (TW)**: this indicates the required material type and/or the necessity for insulative coatings for the turbine blade.
- **Cooling hole area (AC):** in a multi-hole situation individual dimensions must be compatible with those from previous design iterations.
- **Coolant flow rate (CFR):** a minimal value should be the objective in most cases.

Each of the returned values is treated as an independant objective function and a multi-criteria selection strategy is introduced to achieve the eventual isolation of an acceptable parameter set. This solution will provide a minimal CFR at the required values of TW and AC. A simple Vector Evaluated Genetic Algorithm, VEGA [3] has proved effective. For example, suppose a design specification requires a metal temperature of 1260° K with a Cooling Hole area of 7.8mm^2 and the overall objective function is to minimise Coolant Flow Rate. The achievement of the parameter set which approximately defines these utopian values is illusrated in figure 1. A third criteria is the minimisation of CFR. The following acceptable approximation to the utopian values is achieved within the thirty-five generations :

TW = 1260.18 K
AC = 7.79 mm^2;
CFR = 0.00136 kg/s

Further improvement can be achieved but incurs excessive computational expense with little real return.
A local search region can be simply established about this point by setting the mid-point of any variation equal to the value of the parameter defined by the VEGA search. The global range of each parameter is

then modified by a limit factor which defines the degree of allowable variation about the VEGA value as shown below:

Global variation of Initial Coolant Temperature (TCI)	= 700-800° K
Value as defined by VEGA	= 740° K
Limit factor	= 0.2
therefore, local variation	= 720-760° K

This process is repeated for each of the parameters to create a hypercube around the point defined by the VEGA parameter values. A secondary GA search can then be initiated within this hypercube.

CONSTRAINED SEARCH

The efficiency of the GA search can be significantly reduced in heavily constrained design spaces. Constraints can be effectively removed from the design space by the introduction of appropriate penalty and reward functions. This allows the GA to search across the entire space but the fitness of each solution is penalised in proportion to any violation of a constraint boundary. The search will eventually be concentrated within feasible regions of the design space.

In complex constrained spaces, however, convergence can be extremely difficult and it is still possible for the search process to converge upon a high fitness, non-feasible solution due to the constant rejection of the intermediate structures required to establish a path to a feasible region. Although a number of more complex methodologies are available from the literature (See Michalewicz [4]) non-linearity in both the design problem and the constraints appears to present significult difficulties to the GA search process.

An alternative approach is to utilise Schaffer's VEGA to identify a specific point within feasible space. Using the directed search capabilities of VEGA it is possible to identify parameter combinations that define feasible points. For instance, both the values of metal temperature and cooling hole area returned from the turbine cooling hole model must lie within appropriate constraints. By selecting utopian values for both TW and AC that lie within these constraints and minimising coolant flow rate, VEGA will identify a relatively high fitness, feasible parameter combination.

Creating a Feasible Subset

Having defined a feasible point it is necessary to establish a hypercube around that point to create a local search region. The required objective is that a major proportion of this hypercube will lie within feasible space. The simplest approach is to significantly reduce the range of variation of each parameter in the manner previously described. However if the feasible point lies close to constraint boundaries a large percentage of the hypercube would lie in the non-feasible region. It is possible that an optimal hyperspace containing a high proportion of feasible solutions could be established if variable upper and lower bounds for each parameter are manipulated by a secondary GA search. The fitness of each hyperspace thus created would depend upon a combination of the degree of constraint violation at fixed points on the hypercube surface and the relative size of this local search space. Two approaches to this problem have been developed each depending upon

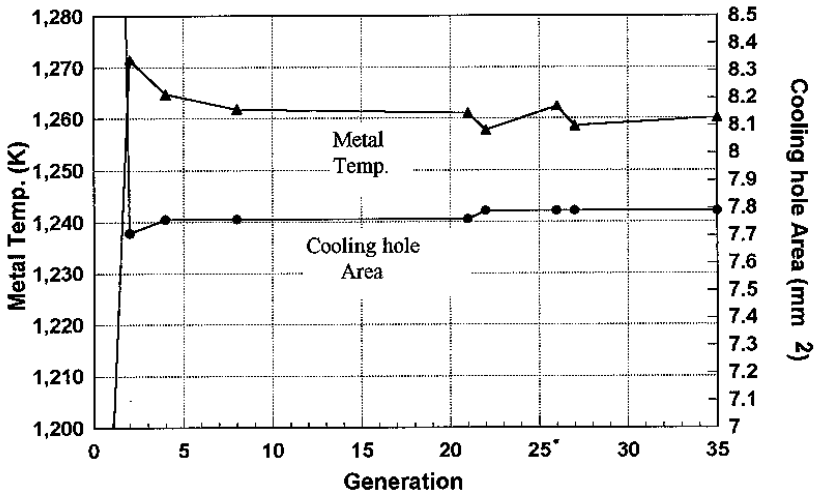

Figure 1: Convergence upon Utopian Values using VEGA

different points upon the hypersurface at which the degree of constraint violation is sampled.

Establishing the Degree of Constraint Violation

Let us assume that the following constraints apply:

$1210° K < TW < 1250° K$; $7.6mm^2 < AC < 8.6mm^2$

The results from a VEGA run searching for:

- $TW = 1230° K$
- $AC = 8.0mm^2$
- CFR = minimum

define the following Utopian Values for the five variable parameters:

TCI=703.91° K RAT=1.385 SGC=1.172

KW=29.25 DTH=1.61mm

It is now possible to introduce ten further variable parameters LB_1 .....LB_5 ; UB_1 .....UB_5 (where LB=lower bound; UB=upper bound) which define the upper and lower bounds of any variation of TCI, RAT, SGC, KW and DTH about their Utopian Values. This is achieved by discretising the ranges between each of the Utopian parameter values and their respective global upper and lower limits ie those limits that describe the global search space or the space defined by a limit factor.

A secondary GA search for the optimal upper and lower bounds of the five parameters can now be initiated. The objective function is to minimise the number of non-feasible solutions returned from these runs whilst maximising the range of each parameter within their respective initial limits. In the first instance the hypercube vertices were considered to be suitable points at which to test the degree of constraint violation of each of the GA generated hypercubes. To do this it is necessary to initiate thirty-two runs of the mathematical model, one for each of the combinations of LB_n and UB_n which define the hypercube vertices. The relative fitness of the hypercubes is determined from the degree of constraint violation and the magnitude of the allowable parameter variation. This results in conflicting criteria as violation is minimised by drawing in the hypercube vertices whilst increasing parameter variation is achieved by forcing the vertices apart.

The obvious disadvantage of this approach is the computational expense incurred by the thirty-two calls to the fitness function for each hypercube. To reduce the number of calls from a 2^n problem to a $2n$ problem (n = no of parameters) mid-point testing has been introduced. This involves the testing of the mid-points of each of the hypercube surfaces and requires ten calls to the fitness function. Figure 2 simply illustrates in two dimensions vertex and mid-point testing.

Either method will create a subset of the constrained region that lies to a greater extent within the feasible region. An optimal design solution can now be achieved by means of a tertiary search of this subset. The type of search required will depend upon the complexity of the fitness surface within the subset.

RESULTS

A simple Genetic Algorithm has been utilised with single point crossover, 0.6 crossover probability, 0.02 mutation probability, roulette wheel selection, fitness scaling and a population size of fifty chromosomes.

Average results have been calculated from fifty runs of the GA each consisting of 200 generations. An exhaustive search of the overall search space has been carried out in order to test the resulting GA-generated

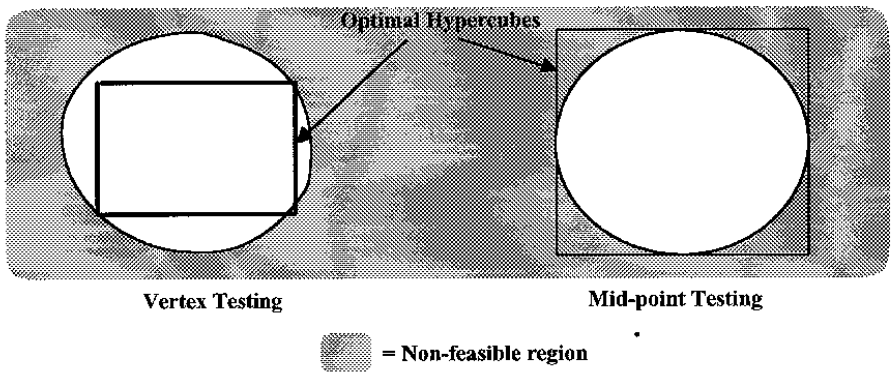

Figure 2: Vertex Testing versus Mid-Point

Constraints: $1210°K<TW<1250°K$ and $7.6mm^2<AC<8.6mm^2$
Utopian Values: TCI = 703.9°K; RAT = 1.39; SGC = 1.172; KW = 29.25; DTH = 1.61mm

Limit Factor: 1.0

	Vertex		Mid-Point		Overall	
	UB	**LB**	**UB**	**LB**	**UB**	**LB**
TCI (°K):	743.28	701.75	752.4	700.0	800.00	700.00
RAT:	1.50	1.34	1.57	1.34	1.60	1.05
SGC:	1.191	1.156	1.207	1.141	1.50	0.50
KW:	31.43	21.67	31.50	20.93	33.00	18.00
DTH (mm):	2.0	0.94	2.31	0.88	2.5	0.75
Violation:						
%age violation:	5.3%		18.8%		93.2%	
Max. TW (°K):	1261.63		1271.34		1402.00	
Min. TW (°K):	1202.182		1194.17		1062.56	
AMV TW(°K):	2.36		4.49		46.47	
Max AC (mm^2):	8.66		8.81		18.18	
Min AC (mm^2):	7.40		7.40		5.47	
AMV AC (mm^2):	0.028		0.079		0.0043	

Limit Factor: 0.4

	Vertex		Mid-Point		Overall	
	UB	**LB**	**UB**	**LB**	**UB**	**LB**
TCI (°K):	729.35	701.33	738.88	700.0	800.00	700.00
RAT:	1.49	1.35	1.57	1.32	1.60	1.05
SGC:	1.194	1.166	1.211	1.154	1.50	0.50
KW:	31.80	24.41	32.21	24.353	33.00	18.00
DTH (mm):	2.10	0.99	2.23	0.97	2.50	0.75
Violation:						
%age violation:	3.6%		14.91%		93.2%	
Max. TW (°K):	1260.00		1267.12		1402.00	
Min. TW (°K):	1206.40		1194.32		1062.56	
AMV TW(°K):	1.262		3.80		46.47	
Max AC (mm^2):	8.20		8.70		18.18	
Min AC (mm^2):	7.40		7.40		5.47	
AMV AC (mm^2):	0.00		0.04		0.0043	

Limit Factor: 0.2

	Vertex		Mid-Point		Overall	
	UB	**LB**	**UB**	**LB**	**UB**	**LB**
TCI (°K):	720.32	700.97	721.90	700.0	800.00	700.00
RAT:	1.48	1.33	1.48	1.31	1.60	1.05
SGC:	1.191	1.168	1.262	1.149	1.50	0.50
KW:	31.90	26.80	32.01	26.57	33.00	18.00
DTH (mm):	1.92	1.29	1.95	1.29	2.50	0.75
Violation:						
%age violation:	0.1%		15.39%		93.2%	
Max. TW (°K):	1260.00		1264.94		1402.00	
Min. TW (°K):	1210.00		1207.57		1062.56	
AMV TW(°K):	0.638		3.684		46.47	
Max AC (mm^2):	8.20		8.73		18.18	
Min AC (mm^2):	7.40		7.40		5.47	
AMV AC (mm^2):	0.00		0.05		0.0043	

AMV = Average Magnitude of Violation
UT = Utopian Value as defined by Vega.

Table 1: Preliminary Results

hypercubes. The exhaustive search is executed between the upper and lower bounds of the optimal hypercubes generated from the GA search procedures using both vertex and mid-point testing. The number of constraint violations, their maximum and minimum values and their average magnitude can thus be ascertained. This allows a comparison of the effectiveness of the different techniques. The results are shown in Table 1 initially for the global search space and then for reduced spaces as created by limit factors of 0.4 and 0.2.

CONCLUSIONS

It is apparent that Schaffer's VEGA can provide a means of locating the parameter sets that define a specific design solution and this, in itself, provides the engineer with an alternative design approach.

The preliminary results shown in Table 1 illustrate that the GA can be utilised to define a feasible subset of a constrained space. The exponential relationship between the number of parameters and the number of hypercube vertices may restrict the application of the technique to problems of low dimension. The number of calls to the fitness function can be reduced by testing the mid-point of each surface of the GA generated hypercubes.

As would be expected the results show that vertex testing creates a hypercube that lies to a greater extent within the feasible region than those from mid-point testing. A comparison of the standard deviation of the results from fifty GA runs suggest that the search space described by mid-point testing is far more complex and sensitive than that for vertex testing.

Both approaches have particular characteristics. Vertex testing will create compact subsets with a high degree of feasibility. However the topology of the feasible region may be such that the resulting hypercube will define a very limited local search region. On the other hand mid-point testing within the same domain will result in a larger subset that contains a higher degree of non-feasible solutions. Both techniques, however, significantly reduce the global search space and will thus facilitate subsequent local search.

Further work will involve investigation of the overall alignment of the hypercube in relation to the topology of the feasible space. More sophisticated GA approaches will be introduced along with developments of the VEGA process. The application of the technique to more complex problems is essential for further development and for benchmarking against current constrained search techniques. The problem of accessing individual, isolated feasible regions within a single global space must also be addressed.

It should be stressed that the overall objective is to develop a technique that provides a means of entering feasible regions that otherwise would be difficult if not impossible to access.

ACKNOWLEDGEMENTS

The research described in the paper is being carried out at the Plymouth Engineering Design Centre at the University of Plymouth in collaboration with Rolls Royce plc.. The PEDC is one of the seven Engineering Design Centres that have been established by the UK Engineering and Physical Science Research Council (formerly the Science and Engineering Research Council). The authors wish to thank the EPSRC and Rolls Royce plc for their continuing support.

REFERENCES

1. Goldberg D. E., 1989, "Genetic Algorithms in Search, Optimisation & Machine Learning". Addison - Wesley Publishing Co., Reading, Massachusetts.

2 Parmee, I. C. , Denham M. J., 1994, "The Integration of Adaptive Search with Current Engineering Design Practice", Procs. Adaptive Computing in Engineering Design and Control, PEDC, University of Plymouth.

3. Schaffer J. D., 1985, "Multiple Objective Optimisation with Vector Evaluated Genetic Algorithms", Proceedings of International Conference on Genetic Algorithms and their Applications.

4. Michalewicz Z., Janikow C., 1991, "Handling Constraints in Genetic Algorithms", Proceedings of the 4th International Conference on Genetic Algorithms.