OPTIMAL SEQUENCED MATROID BASES SOLVED BY GA'S WITH FEASIBLE SEARCH SPACE INCLUDING APPLICATIONS

Michael L. Gargano
Computer Science Dep't.
Pace University
New York, N.Y. 10038
mgargano@pace.edu

William Edelson
Computer Science Dep't.
Long Island University
Brooklyn , N.Y. 11201
edelson@liu.edu

Abstract

We consider an extension to the optimal matroid base problem [1] whereby the matroid element costs are not fixed, but are time dependent. We propose a genetic algorithm (GA) approach to solve the optimal sequenced matroid base problem (OSMBP) by employing efficient codes which are suffixed by a standard permutation code [2]. These novel encoding schemes *insure feasibility* after performing the classical operations of crossover and mutation and also ensure the feasibility of the initial randomly generated population (i.e., generation 0). This class of problems, where costs are not fixed but are time dependent, *embrace non-locality* which actually makes the GAs more efficient. A variety of practical matroid applications with time dependent costs will also be presented.

Introduction

A finite **matroid** M is a pair (E,β) where E is a non-empty finite set and β is a non-empty collection of subsets of E called **bases** satisfying the following properties:

i) no base properly contains another base, and

ii) if B_1 and B_2 are bases in β (i.e., $B_1, B_2 \in \beta$) and if $e \in B_1$ then there exists $f \in B_2$ such that $((B_1 - \{e\}) \cup \{f\}) \in \beta$ (i.e., $((B_1 - \{e\}) \cup \{f\})$ is also a base).

By repeatedly using ii) it is easily shown that any two bases $B_1, B_2 \in \beta$ have the same number of elements. This number is called the **rank of M**. Any subset of a base is called an independent set. The **dual matroid** $M^* = (E,\beta^*)$ of M over E has its bases in $\beta^* = \{ B^c \mid B \in \beta \}$, that is, the collection of the complements of β [1]. Given M = (E,β), the **optimal matroid base problem (OMBP)** is to find an optimal base (not necessarily unique)of M (i.e., a base B with a maximu (minimum) total cost (or weight)) where each element of $e \in E$ has a cost (or weight) which is a real number C(e). Kruskal's algorithm applied to matroids solves this problem as well as simultaneously solving the OMBP for M^* (which has an optimal base $(B)^c$ with a minimum (maximum) total cost (or weight)). Here we consider an extension to the optimal matroid base problem (OMBP) whereby the matroid element costs are not fixed, but are time dependent. This is called the **optimal sequenced matroid base problem (OSMBP).** Let M = (E,β) be a matroid of rank n such that each element $e \in E$ has assigned to it C(e,t) a real valued function of t having domain { 1, 2, …, n }. We refer to M as **a matroid with time dependent cost assignment C** and we say C(e,t) is the cost of selecting e at time t. A sequencing (permutation, ordering) of the elements of a base $B \in \beta$ is called a **sequenced (ordered) base**. It is assumed that the selection of any $e \in E$ takes exactly one unit of processing time. A sequenced base is an n-tuple $(e_1, e_2, \ldots, e_n)$ which shows the order in which the elements of the base are to be selected and determines the costs $C(e_1,1)$, $C(e_2,2)$, …, $C(e_n,n)$ associated with this sequence. The **total cost** $C(e_1,1) + C(e_2,2) + \ldots + C(e_n,n)$ is the objective function $T(e_1, e_2,\ldots, e_n)$ to be optimized. Since this problem is NP hard, we propose a genetic algorithm (GA) to solve the optimal sequence matroid base problem (OSMBP) using efficient codes which are suffixed by a permutation code [2]. These novel encoding schemes *insure feasibility* after performing the classical operations of crossover and mutation and also ensure the feasibility of the initial randomly generated population. These types of problems, where costs are not fixed but are time dependent, *embrace the non-locality* which is implied in our coding and which makes the GAs more efficient.

A wide variety of pragmatic matroid applications with time dependent costs will be presented (i.e., spanning tree, airline route sales (dual matroid), position assignment, node base communications, and bidding with slack time applications). (Note: A similar problem of less complexity has been solved in polynomial time in [3]).

References

1. Wilson, R.J. ,Graph Theory, Fourth Ed, Addison Wesley , 2000
2. Edelson, W. , and Gargano, M.L. , Minimal Edge-Ordered Spanning Trees Solved By a Genetic Algorithm With Feasible Search Space, Congressus Numerantium, 135 (1998) pp 37-45.
3. Gargano, M.L., Quintas, L.V. , and Friedrich, S.C., Matroid Bases With Optimal Sequencing, Congressus Numerantium 82, (1991) pp 65-77.

Reliability Optimization Design Using Hybrid NN-GA with Fuzzy Logic Controller

Mitsuo GEN
Dept. of Indust. & Inform. Sys. Engg.
Ashikaga Institute of Technology
Ashikaga, 326-8558 Japan
gen@ashitech.ac.jp

ChangYoon LEE
Dept. of Indust. & Inform. Sys. Engg.
Ashikaga Institute of Technology
Ashikaga, 326-8558 Japan
d99102cy@mail.ashitech.ac.jp

1 SUMMARY

This paper proposes an hybrid genetic algorithm/neural network with fuzzy logic controller (NN-flcGA) to find the global optimum of reliability assignment/redundant allocation problems which should be simultaneously determined two different types of decision variables. For the same problem, Dhingra employed the method combined an interior-penalty function and a heuristic procedure, and Yokota *et. al* applied a GA method [1]. Recently, in order to overcome the weak point of GA, Lee *et. al* considered a combined GA with a neural networks technique (NN-hGA) as one of approximation methods suitable for a continuous decision variable values [1]. We also simulated the same problem by the traditional Hopfield networks as one of NN techniques.

The efficacy and efficiency of NN-flcGA is demonstrated by comparing its results with those of other traditional works. The essential features of NN-flcGA, 1) combining with neural network (NN) technique to devise initial values for the GA, 2) the concept of fuzzy logic controller (FLC) when tuning strategy GA parameters dynamically [2], and 3) incorporation of the revised simplex search method, make it not only improve quality of solutions but also reduce computational cost.

2 RESULTS AND ANALISYS

Table 1 shows the comparison results among the fitness value and computation time of the proposed method (NN-flcGA) and those of the previous works. The results indicate that NN-flcGA and NN-hGA find solutions with better fitness values than those of the other methods, while the GA yields worse solutions than the Dhingra's method and the Hopfield mehtod. This means that our GA combined with the NN technique to devise initial values overcomes the weak point of GAs which have a lot of enumeration of feasible solutions on real search area. Combining with revised simplex search more careful deals with continuous decision variables as well as auto-tuning by FLC helps for genetic operators to search more efficient for integer decision variables. It permits for our GA to improve for the quality of solutions of NN-hGA. The computation time is obtained at the generation keeping the best fitness up as the same one during over 30 generations. While NN-hGA took about 2 times longer computation time than other traditional method, NN-flcGA achieved a satisfactory and acceptable results in proper computation time by the dynamic auto-tuning of GA strategy parameters and the two-stage termination in the simplex search method.

Table 1: Comparison Results

Method	j	x_j^*	m_j^*	f^* and $\boldsymbol{g}^*$	tiem [s]
Dhingra's method [1]	1	0.816040	6	0.999607	-
	2	0.803090	6	399.936000	
	3	0.983640	3	185.000000	
	4	0.803730	5	495.652000	
GA method [1]	1	0.965993	3	0.999468	8.3
	2	0.760592	6	372.416811	
	3	0.972646	3	158.000000	
	4	0.804660	5	470.733576	
NN method	1	0.901361	4	0.999563	10.8
	2	0.839251	5	242.843812	
	3	0.912679	4	164.000000	
	4	0.823424	5	397.096075	
NN-hGA method[1]	1	0.944341	3	0.999701	16.4
	2	0.855253	5	390.553924	
	3	0.944386	6	217.000000	
	4	0.855320	5	480.099368	
proposed method	1	0.812610	6	0.999806	14.6
	2	0.916615	4	359.662400	
	3	0.937032	5	193.000000	
	4	0.841160	5	488.355300	

References

[1] C. Y. Lee, M. Gen and W. Kuo, "Reliability optimization design using hybridized genetic algorithm with a neural network technique," IEICE Trans. Fundamental, vol. E84-A, no. 2, pp.627-637, 2001.

[2] M. Gen and R. Cheng, Genetic Algorithms and Engineering Optimization, John Wiley & Sons, New York, 2000.

Estimation of Mixture Parameters by a Hybrid GA/EM Method: Application to Radar Remote Sensing Data

Reinhold Huber
Aero Sensing Radarsysteme GmbH
c/o DLR Oberpfaffenhofen
D-82234 Wessling, Germany

Abstract

A method for maximum likelihood parameter estimation of normal mixture components parameters based on a genetic algorithm and a hybrid approach based on expectation maximum initialized by a genetic algorithm is presented. The method is evaluated for and remote sensing data.

1 INTRODUCTION

The presented approach is a hybrid method, i.e. a genetic algorithm (GA) provides initial guesses for expectation maximization (EM) [1] . The advantages of the hybrid GA/EM method are 1) the optimum solution is very likely to be found (at least experimentally this was the case), 2) the EM procedure shows convergence and 3) this convergence is achieved in less iteration steps.

2 PROBLEM STATEMENT

We assume that $\mathcal{S} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{|S|}\}$ is a finite set of $|\mathcal{S}|$ independent observations of a random vector $\mathcal{X}$ distributed according to some unknown probability density function $p(\boldsymbol{x})$ of the form:

$$p(\boldsymbol{x}) = \sum_{i \in \mathcal{M}} \alpha_i f(\boldsymbol{x}|\mu_i, \Sigma_i), \quad \boldsymbol{x} \in R^d, \qquad (1)$$

$$\sum_{i \in \mathcal{M}} \alpha_i = 1, \quad \alpha_i \geq 0, \quad \mathcal{M} = \{1, 2, \ldots, M\},$$

where $f(\boldsymbol{x}|\mu_i, \Sigma_i)$ are normal component densities with the mean vector μ_i and the covariance matrix Σ_i. The parameter d is the dimensionality of the data, in this study we restrict ourselves to $d = 1$. The task of mixture modeling is to estimate the parameters μ_i, Σ_i and α_i.

3 EXPERIMENT

A sample of 752 measurements was extracted from airborne AeS–1 X–Band Synthetic Aperture Radar (SAR) data. The EM method was initialized either by the GA solution (hybrid approach) or for each randomly chosen individual of the initial GA population EM was performed. The GA parameters were chosen as: population size 50, crossover rate 0.6, mutation rate 0.05, and binary tournament selection. Table 1 presents results using χ^2–test results, number of iterations/generations, and the log–likelihood.

Table 1: Comparison of Methods for $|\mathcal{M}| = 2, 3$.

Method	χ^2	p_{χ^2}	It.	$\log L$
$\|\mathcal{M}\| = 2$				
EM	17.12	0.45	133	1249.88
GA	38.59	0.00	90	1237.76
GA/EM	17.12	0.45	128	1249.88
$\|\mathcal{M}\| = 2$				
EM	15.00	0.60	498	1251.30
GA	12.83	0.75	71	1256.27
GA/EM	10.41	0.89	139	1257.79

4 DISCUSSION

A general observation is, that EM is the best algorithm for the estimation of normal components parameters. The drawback of EM is, that is was observed experimentally, that is is in approximately 60 percent of all cases, with randomly guessed initial guesses, it is not able to find the optimal solution.

References

[1] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum Likelihood from Incomplete Data via the EM Algorithm. *J. Royal Statist. Soc. Ser. B (methodological)*, 39:1–38, 1977.

Multidimensional Scaling Using Evolutionary Computation

Boris Lorenc

Department of Statistics
Stockholm University
106 91 Stockholm, Sweden
boris.lorenc@stat.su.se

Abstract

Possibility to apply simple genetic algorithm to the task of multidimensional scaling is explored. For metric scaling, Kruskal's Stress is used as the evaluation function, except that no prior transformation of data is performed. For nonmetric scaling, an evaluation function based on Spearman's rank correlation is devised. The proposed method, as compared to classical scaling and to a steepest descent iterative method, consequently yields solutions with better fit to data.

Multidimensional scaling (MDS) is the common name for a number of multivariate statistical techniques that, starting from a matrix of pairwise so-called dissimilarities between some entities, propose an appropriate number of dimensions and position the entities on them. This is in essence an optimization problem, apparently amenable for the evolutionary computation (ECO) approach.

The techniques group into classical scaling, which is based on Young-Householder theorem in matrix algebra and limited in that it assumes variables to be on a metric level, and into iterative approaches based on derivation of some function expressing misfit between the dissimilarities and a proposed *n*-dimensional configuration (methods of "steepest descent"/"negative gradient", pioneered in this context by Kruskal in 1964). Present paper offers a third approach, that of evolutionary computation. It is motivated by shortcomings of the iterative method. Kruskal and Wish (1978) sum these up as: inappropriate model assumptions, emergence of locally optimal solutions, and emergence of degenerate solutions. The present approach—termed ECOMDS—addresses the first two: no model (i.e., functional) assumptions are made, as they are not necessary for an evolutionary computation approach; and, using a genetic algorithm, it is taken that risk of entrapment into a local minimum is reduced as compared to a steepest descent method.

ECOMDS for metric scaling minimizes a variant of Kruskal's Stress: Stress, but with omission of a function transforming original dissimilarities. Thus, objective function minimized in metric ECOMDS is

$$Stress = \sqrt{\frac{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}\left(\delta_{ij} - d_{ij}\right)^2}{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} d_{ij}^2}}$$

where δ_{ij} denotes dissimilarity between entities i and j, and d_{ij} denotes distance between them in a configuration.

Nonmetric ECOMDS minimizes

$$Stress(r) = \frac{1 - r_s(\boldsymbol{\delta}_\mathbf{r}, \mathbf{d}_\mathbf{r})}{2}$$

where $r_s(\boldsymbol{\delta}_\mathbf{r}, \mathbf{d}_\mathbf{r})$ is Spearman's rank correlation between vectors of ranked dissimilarities and corresponding ranked distances.

ECOMDS was compared to classical scaling and to PROC MDS of SAS statistical package (the latter algorithm being a variant of Kruskal's original approach), with the result that ECOMDS in all the cases performed at least somewhat better than the other methods in terms of the obtained objective function values. These results are discussed in the light of computing cost, efficiency, and improvement with respect to the gradient descent approach.

Acknowledgments

I am indebted to Ove Frank and Tommy Johnsson for their comments and suggestions.

References

Kruskal, J.B., and Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: Sage.

Application of Genetic Algorithm to Determining Worst-Case Gust Loads and the Dynamic Response of Aircraft Structures

Rajiv Mehrotra
Graduate student
Department of Aerospace Engineering and Mechanics
Box 870280, University of Alabama
Tuscaloosa, AL 35486-0024

Charles L. Karr
Associate Professor
Department of Aerospace Engineering and Mechanics
Box 870280, University of Alabama
Tuscaloosa, AL 35487-0280

Thomas A. Zeiler
Assistant Professor
Department of Aerospace Engineering and Mechanics
Box 870280, University of Alabama
Tuscaloosa, AL 35487-0280

Abstract

Determining worst-case gust on the structural integrity of aircraft structures is an extremely complex yet terribly important problem. In this paper a genetic algorithm (GA) is shown to be a feasible approach to solving this problem. The GA outperforms a traditional method suitable for solving linear problems, then is extended to nonlinear problems for which there is no acceptable solution methodology.

1 PROBLEM STATEMENT AND SOLUTION METHODOLOGY

The primary structure of an aircraft must be designed to withstand all of the static and dynamic loads it is expected to encounter during the life of the aircraft. In general, the dynamic loads are more difficult to determine. The dynamic loads that must be accounted for include landing and taxi loads, maneuver loads, and gust loads. Of these, gust loads are by far the most difficult to account for due to the stochastic nature of gust loads as compared to other dynamic loads that at least tend to be deterministic.

This study employs a GA to determine the worst-case or critical gust profile for specific loading conditions on a specific aircraft model. This worst-case gust is one that results in maximum values of one of four responses [(1) wing root bending moment, (2) wing root torque, (3) engine lateral acceleration, or (4) c.g. normal acceleration] in a given aircraft model when the given gust is used as the forcing function in the model. The study uses the SDG representation of atmospheric turbulence (Jones, 1973) as a gust-loads analysis method.

The approach combines the SDG model with a dynamic GA (Hong et.al, 1999) to determine the worst-case gust. This formulation has at least two very desirable attributes. First, the approach is applicable to both linear and nonlinear aircraft models. Second, the method is easily adapted to account for a variety of definitions of "worst-case" gusts, e.g., the designer can easily incorporate a variety of criteria for determining exactly what a worst-case gust is. The fundamental methodology relies on both a model of the SDG method and on a dynamic GA. It is implemented with the genetic-based search procedure that maximizes the peak response in a given aircraft load quantity, thereby determining the associated critical gust profile. The flow diagram shown as Figure 1 gives an overview of the SDG constrained optimization loop, which has been implemented for this study.

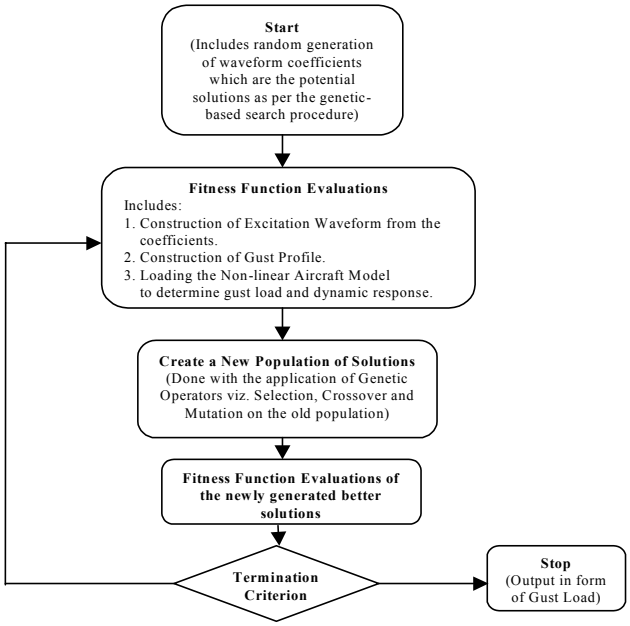


Figure 1: Schematic of the optimization loop.

References

Jones, J.G., (1973), "Statistical Discrete Gust Theory for Aircraft Loads", Royal Aircraft Establishment, Farnborough, England, UK, RAE TR-73167, 1973.

Hong, T.P., Wang, H.S. and Chen, W.C. (1999). "Simultaneously Applying Multiple Crossover and Mutation Operators". *Proceedings of the Genetic and Evolutionary Computation Conference*, vol. 1, pp.790, 1999.

Multilocus Pattern Recognition using Cellular Automata and Parallel Genetic Algorithms

Jason H. Moore, Ph.D.
Program in Human Genetics
519 Light Hall
Vanderbilt University
Nashville, TN 37232-0700
moore@phg.mc.vanderbilt.edu

Lance W. Hahn, Ph.D.
Program in Human Genetics
519 Light Hall
Vanderbilt University
Nashville, TN 37232-0700
hahn@phg.mc.vanderbilt.edu

1 OVERVIEW

We developed a strategy for identifying susceptibility genes for common complex human diseases that utilizes genetic algorithm-optimized cellular automata (CA) for pattern recognition and emergent computation. We have adapted a one-dimensional CA to accept an array of genotypes as input and produce an array of information as output that can be used to classify subjects as affected or unaffected. We implemented a parallel genetic algorithm for optimizing the selection of genotypes and CA parameters. Three different CA output encodings are compared. The primary advantage of our genetic algorithm-optimized CA approach is the ability to consider high-order interactions among genetic variations in small sample sizes without loss of power.

In this initial study, we fixed the number of cells presented to the CA to five. Assuming each genetic locus has only three possible genotypes, we used a binary encoding with '01' for the first genotype, '10' for the second genotype, '11' for the third genotype, and '00' for the wildcard. Thus, each array presented to the CA consisted of 10 bits with two bits encoding the state of each of the five cells. We allowed a maximum of 128 iterations for each CA. The combination of genotypes, the CA rule table, and the number of iterations selected were all optimized using a parallel genetic algorithm (two populations, population sizes of {10, 50, 100, 200, 500, 1000}, 200 iterations, migration every 25 iterations, recombination frequency of 0.6, and mutation frequency of 0.02). The GA used was a modification of the Parallel Virtual Machine (PVM) version of the Genetic ALgorithm Optimized for Portability and Parallelism System (GALLOPS) package for UNIX (http://garage.cps.msu.edu). This GA package was implemented in parallel using message passing on a 110 processor Beowulf-style parallel computer cluster.

Fitness of a particular CA model is based on the classification and prediction error estimated using 10-fold cross-validation. We evaluated three different CA output encodings by comparing the average classification and prediction errors obtained by analyzing 50 different simulated datasets. The first encoding we evaluated is based on counting the number of 1s in the binary encoded output array of the CA run on each set of genotypes for each affected and each unaffected sib in the sample. A classifier is formed by using a frequency histogram of the number of 1s among affected sibs and unaffected sibs. Each bin is labeled affected or unaffected depending on whether the number of 1s represented by that bin were more frequently observed among affected or unaffected sibs. For example, consider the case where 100 discordant sib pairs were evaluated. Suppose the number of CA arrays that contained three 1s was 20 for affected sibs and 10 for unaffected sibs. This bin would be labeled affected and thus the 10 unaffected sibs would be misclassified. This would contribute 0.05 to the overall misclassification rate. This is performed for each bin and a total classification error is estimated by summing together the individual rates for each bin. The second encoding evaluated is a simplification of the bit count encoding described above. Here, a special field within the binary output array is defined and the number of 1s within that field is used for classification and prediction. We defined the special field as the leftmost pair of bits. The third encoding evaluated transforms the binary output array from the CA into an integer value. The classification error is estimated from the overlap of the distributions of integer values for affected sibs and unaffected sibs.

2 RESULTS AND CONCLUSIONS

The first main conclusion that can be drawn from this study is that the emergent computation features of cellular automata and the intelligent search features of parallel genetic algorithms can be used to identify patterns of genetic variations associated with disease when the genetic effects are primarily through interaction. The second main conclusion that can be drawn from this study is that encoding the CA output in binary form and then using a count of the number of 1s as the classifier (i.e. the bit count encoding strategy) is superior to using a special bit field or using an integer encoding.

Calculating Robot Trajectories with parallel Evolutionary Strategies

Matthias Ortmann
Department of Computer Science
Ruhr-University Bochum
44780 Bochum, Germany

Wolfgang Weber
Department of Computer Science
Ruhr-University Bochum
44780 Bochum, Germany

Abstract

Owing to the time-expensive calculation of evolutionary algorithms and especially of collision-detection algorithms, the calculation of robot trajectories mostly is not real timable. Therefore, this paper presents an approach for shortening the calculation time by parallelising the whole algorithm.

As evolutionary algorithms (Bäck, 1996) are stochastic search techniques they normally need a great amount of time to find the solution for the problem. A strategy for speeding up evolutionary algorithms is to parallelize them (Dorigo, 1993). Therefore, this paper deals with the calculation of robot trajectories in a nearly homogeneous 100 MBit TCP/IP-based cluster.

Without loss of generality, a manipulator with five degrees of freedom is used. To each joint a variable $GeneX \in \mathbb{N}$ is assigned. The point of time is added to the position of the robot. The whole vector is called *chromosome*. To get an *individual*, many chromosomes are lined up at a distance of 40 ms. Special user defined positions subdivide the path of the robot arm into many smaller parts called *regions*. Each region can be optimized independent of the others by a parallel evolutionary strategy of medium granularity.

To get on the one hand a smooth and short trajectory the fitness value F_{GeneX} of one gene is calculated by summing up the squared curvature values of each chromosome in the region. The fitness values F_{GeneX} of the genes are summed up to the fitness value F_{curv} of this part of the whole individual. This multi-criterion fitness value can be optimized Pareto-optimally (Ortmann, 2001). The second aim is on the other hand to avoid collisions between the manipulator and obstacles. The two criteria curvature and no collision mostly can not be satisfied Pareto-optimally because collisions have to be avoided resulting in possibly long and not smooth trajectories. Hence, if a collision occurs, the collision-fitness value F_{col} is added to F_{curv}. F_{col} is the weighted and squared degree of collision between the trajectory and the obstacles.

The speeding-up of the calculation in a cluster depends on the time needed for the optimization of the regions because in the case of short calculations, the time needed for the shipping of the results from the slaves back to the master can not be neglected. This time can be neglected with increasing time needed only for the optimization of one region leading to a speeding-up of the algorithm by the number of computers in the cluster. The number of regions should match the number of computers in the cluster to get an optimal speeding-up. In the case of collision, a homogeneous cluster is unsuitable because of the large differences in the loads. Here an inhomogeneous cluster is advantageous because then the regions without collision can be calculated on slower computers and the regions with collision can be calculated on fast computers, especially on multi-processor systems with an algorithm of high granularity.

References

T. Bäck (1996). Evolutionary Algorithms in Theory and Practice - Evolution Strategies, Evolutionary Programming, Genetic Algorithms. *Oxford University Press.*

M. Dorigo, and V. Maniezzo (1993). Parallel Genetic Algorithms: Introduction and Overview of Current Research. In J. Stender (ed.), *Parallel Genetic Algorithms - Theory and Applications*, IOS Press

M. Ortmann (2001). Multi-Criterion Optimization of Robot Trajectories with Evolutionary Strategies. to appear in *Facta Universitatis, April*, Department of Computer Science, Ruhr-University Bochum.

Designing A New Elitist Nondominated Sorted Genetic Algorithm For A Multiobjective Long Term Groundwater Monitoring Application

Patrick M. Reed
4129 NCEL, MC. 250
205 N. Mathews Ave.
Urbana, IL. 61801
preed@uiuc.edu
217-333-6979

Barbara S. Minsker
3230d NCEL, MC. 250
205 N. Mathews Ave.
Urbana, IL. 61801
minsker@uiuc.edu
217-333-9017

David E. Goldberg
117 Transportation, MC. 238
104 S. Mathews
Urbana, IL. 61801
deg@uiuc.edu
217-333-0897

Abstract

This study presents a niching-based elitist enhancement of the Non-dominated Sorted Genetic Algorithm (NSGA) and tests its performance in identifying the Pareto frontier for a groundwater monitoring application. The application utilizes historical data at a single snapshot in time to identify potential spatial redundancies within a monitoring network. The study combines nonlinear spatial interpolation with the elitist NSGA to identify the Pareto frontier for sampling costs and local concentration estimation errors. The Elitist NSGA nearly replicated the true front, finding representative solutions along the entire trade off between cost and estimation error.

1 INTRODUCTION

The major goal of this study is to present a new niching-based elitist enhancement of the NSGA. Additionally, an extension of the genetic algorithm design methodology presented by Reed et al. (2000) is demonstrated on a multiobjective long-term groundwater monitoring application.

2 MANAGEMENT MODEL

To identify monitoring wells that are redundant, this study employs a multiobjective approach with the intention of attaining the best-interpolated picture of the perchloroethylene (PCE) plume for the least cost.

3 ELITIST ENHANCEMENT

Conceptually, the elitist strategy proposed in this study is very similar to the sGA, in that the current best individual in a given niche at generation t, if not present in generation (t+1), is inserted into that subpopulation, ensuring that its traits are available for subsequent search for the Pareto front.

This strategy was implemented by defining σ_{elite} or the elite radius, which defines the distance beyond which members of the current nondominated set are considered independent from one another. Only independent members of the nondominated set are considered for insertion in the next generation. For this application, σ_{elite} was set equal to the niche radius which means that only one representative of each niche in the current nondominated set is considered for elitist reproduction into the next generation.

4 RESULTS

The NSGA found solutions near 18 of the 36 Pareto optimal solutions attained from enumeration. Without elitism, the algorithm only exactly identified 5 of the 36 solutions along the Pareto front. Utilizing the niching-based elitist strategy described in Section 3, the elitist NSGA was able to find solutions near 33 of the 36 Pareto optimal solutions. The Elitist NSGA exactly identified 20 of the 36 solutions along the Pareto front. The niching-based elitist strategy effectively rescaled the system, increasing selection pressure along the entire extent of the Pareto front and reducing the loss of niches. For more details on this study see Reed et al. (2001).

5 REFERENCES

Reed, P., Minsker B., & Goldberg, D. E. 2000 Designing a competent simple genetic algorithm for search and optimization. *Water Resources Research*, 36(12), 3757-3761.

Reed, P., Minsker B., & Goldberg, D. E. 2001 A multiobjective approach to cost effective long-term groundwater monitoring using an Elitist Nondominated Sorted Genetic Algorithm with historical data, *Journal of Hydroinformatics*, (in press).

Application Service Providing for Shop Floor Scheduling in Non-Hierarchical Regional Production Networks using Genetic Algorithms

Tobias Teich*
Dept. of Economics†
Technische Universitt Chemnitz
D-09107 Chemnitz
Germany

Horst Meier‡
Dept. of Engineering
Ruhr-Universitt Bochum
D-44801 Bochum
Germany

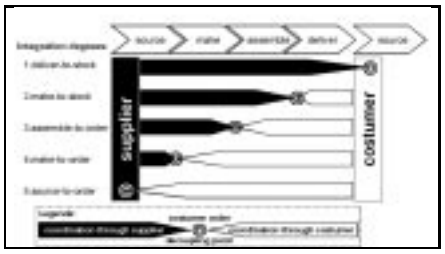
Harald Schallner‡
Dept. of Engineering
Ruhr-Universitt Bochum
D-44801 Bochum
Germany

SUMMARY

The global economy of the 21th century will be characterized by competing corporate networks rather than by competing enterprises. Competition outside of and cooperation within supply chains will be the important factors. The special research area 457 of the German Research Community and INTERKON, proposes an inter-organizational framework for collaborative manufacturing control in production networks, are dealing with these topics. In our work we combine both approaches.

The non-hierarchical structure of the production networks examined here demands a collaborative decision-making. A multitude of competence cells with different interests and objectives is taking part in these processes. Consequently, inter-organizational manufacturing control is carried out via collaborative agreements on detailed order conditions and across several manufacturing levels. However, identifying and quantifying the implications of each decision that must be taken in conjunction with all but the smallest manufacturing level is not humanly possible. This makes an ASP-functionality necessary that performs "what-if" analyses for each competence cell before making local production decisions.

A developed model editor offers intelligent support to help companies model their manufacturing process and the features they require. This approach is based on the "Reference Model for Shop Floor Production" according to the ISO-Standard (Workgroup TC184/SC5/WG1) [1]. Each process step is modelled on a very general level as an activity with the input or output of information, material, resource and control information. The ASP-functionality checks such factors like manufacturing capacities and parts orders to decide whether a new order can be inserted into the production schedule and still satisfies the resource constraints and the order conditions. Thus we differentiate between the following five coordination patterns:



The next step is the integration of this software-system into an organizational framework. The authors carried out a study [2] to identify objectives used in companies for manufacturing control and to develop a ranking according to their importance. It has been argued by the authors that modern shop floor scheduling tools based on Evolutionary Algorithms prove their potential for solving complex benchmark scheduling problems but failed to be used widely in real-world applications throughout the last ten years. One of the reasons identified is the need for organizational changes in manufacturing control, caused by systems. The authors are proposing an organizational framework which includes bi-directional PDA-systems in which EA-based shop floor scheduling tools can be successfully implemented.

References

[1] ISO TC184/SC5/WG1: ISO Reference Model for Shop Floor Production Standards Part 1&2. ISO, ANSI-NEMA, Washington DC (USA) 1989, (1990)

[2] Käschel, J., Teich, T., Zacher, B.: An empirical study of manual shop floor scheduling in job shop environments. Working Report 30/2000. University of Technology of Chemnitz, Department of Economic Science, Chemnitz (2000)

*Email: t.teich@wirtschaft.tu-chemnitz.de
†phone: ++49-371-531 42 52
‡Email: meier@lps.ruhr-uni-bochum.de
‡Email: schallner@lps.ruhr-uni-bochum.de

Evolutionary Surface Reconstruction Using CSG-NURBS-Hybrids

Klaus Weinert, Tobias Surmann, and Jörn Mehnen
Dept. of Machining Technology
University of Dortmund
D-44227 Dortmund, Germany

Abstract

This work presents a surface reconstruction system which combines the concepts of *GP* and *ES* in one main algorithm as well as the concepts of *CSG* (**C**onstructive **S**olid **G**eometry) and *NURBS* (**N**on**U**niform **R**ational **B**-**S**plines) to represent arbitrarily sculptured surfaces.

1 INTRODUCTION

Surface reconstruction is a problem of discrete approximation and structural optimization. Since this problem is multidimensional and multiobjective, a *GP*/*ES* hybrid algorithm [1] has been developed to reconstruct surfaces represented by sampling points of real physical objects.

2 *GP*/*ES*-HYBRID

The genome of the individuals is represented by a *CSG* tree data structure with an extension of the set of primitives by *NURBS*-solids. When reconstructing a surface there are two major problems to be solved:

- Determination of the construction logic. I. e. the algorithm has to find the correct structure of the *CSG* tree regarding the number, position and type of nodes.
- Fitting the leaves into the sampling points by adjusting their parameters (position, size and tilt angles). For *NURBS*-solids the control vertices also have to be adjusted.

These two problems lead to a *GP*/*ES* hybrid algorithm. The objective of the *GP* is to find the structure of the *CSG* tree. Once the construction logic is found, the *ES* part adjusts the parameters of the primitive geometric objects. Therefore, the following genetic operators were applied to the *CSG* tree:

- Variation of the primitive's parameters.
- Variation of inner node functions.
- Deleting and inserting nodes.
- Replacing of a primitive by another one.
- Recombination by exchanging subtrees.

The multiobjective fitness function compares the distances and the normal vectors between each sampling point and the corresponding point of the individuals' surface. The number of nodes in the genome is also considered in the fitness calculation to ensure a low number of primitive objects in the reconstruction.

3 CONCLUSIONS

By combining two major concepts of geometric modelling, *CSG* and *NURBS*, into one common data structure it is possible to reconstruct surfaces that contain both regular geometric and sculptured parts using one evolutionary surface reconstruction system.

Acknowledgments

This research was supported by the Deutsche Forschungsgemeinschaft (SFB 531).

References

[1] R. E. Keller, J. Mehnen, W. Banzhaf, and K. Weinert. *Surface reconstruction from 3D point data with a genetic programming / evolution strategy hybrid, volume 3 of Advances in Genetic Programming*, pp 41–65. MIT Press, 1999