SPARSE EXPERIMENTAL DESIGN: AN EFFECTIVE AND EFFICIENT WAY OF DISCOVERING BETTER GENETIC ALGORITHM STRUCTURES.

D.J. Stewardson[a], R.I. Whitfield[b], C. Hicks[c], P. Pongcharoen[c] P.M. Braiden[c].
[a] Industrial Statistics Research Unit, Department of Mechanical, Materials and Manufacturing Engineering, [c] Department of Mechanical, Materials and Manufacturing Engineering University of Newcastle upon Tyne, NE1 7RU, England, [b] CAD Centre, Department of Design, Manufacture and Engineering Management, University of Strathclyde, Glasgow, G11XQ, Scotland.

Key Words: Schedule Optimisation, Genetic Algorithm, Complex Assembly, Crossover, Mutation, Sequential Experimentation, Embedded Fractional Factorial.

ABSTRACT

The focus of this paper is the demonstration that sparse experimental design is a useful strategy for developing Genetic Algorithms. It is increasingly apparent from a number of reports and papers within a variety of different problem domains that the 'best' structure for a GA may be dependent upon the application. The GA structure is defined as both the types of operators and the parameters settings used during operation. The differences observed may be linked to the nature of the problem, the type of fitness function, or the depth or breadth of the problem under investigation. This paper demonstrates that advanced experimental design may be adopted to increase the understanding of the relationships between the GA structure and the problem domain, facilitating the selection of improved structures with a minimum of effort.

1 INTRODUCTION

An increasing amount of research has identified that the structure for a GA may be dependent upon the nature of the optimisation problem. Todd [22] investigated the use of several non-heuristic crossover and mutation operators within a GA applied to the travelling salesman problem. The purpose of the investigation was to find a good combinatorial operator for both crossover and mutation. Fourteen different crossover operators and five different mutation operators were explored within a series of experiments containing every combination of the two types of operators for problems containing 10, 20, 50, 100 and 200 cities. The crossover and mutation operators were evaluated whilst keeping the possibility of crossover and mutation at 0.8 and 0.01 throughout. Todd concluded that the enhanced edge recombination crossover and the adjacent two city mutation operators performed best under these test conditions. Todd also observed that the GA had difficulties solving problems involving a large number of cities, generating solutions which had good short local subtours linked together with long edges. Todd finally suggested that domain specific knowledge in the form of heuristics might be useful within the solution of these types of problems.

Todd and Sen [23] used a multi-criteria GA to determine the optimum loading sequence for a containership. Within this particular application, the authors used a multi-point crossover operator selecting mates contained only within the chromosome's niche. The operator was observed to produce infeasible sequences and a repair routine incorporating knowledge of the problem domain was incorporated. The authors also discovered that simple mutation operators did not provide sufficient diversity, converging early to non-optimal solutions, and subsequently utilised a heuristic mutation operator containing a number of rules that are used in the practice of containership loading.

Simpson et al. [19] demonstrated the use of a GA for calibrating friction factors within a piping network. Two different coding schemes were used: a discrete coding, and, a continuous coding which included new crossover and mutation operators. A mutation rate of zero was used throughout their experiments since it was previously discovered that the mutation rate had little effectiveness on the GA search within this particular application – Simpson & Goldberg [18]. The authors concluded that the discrete coding produced a more efficient search than the continuous representation, probably due to the significantly smaller search space.

Pongcharoen et al. [14] utilised experimental design to identify the appropriate values for the probabilities of crossover and mutation, the populations size and the number of generations. A full factorial experiment was created with five replications and from the results it was discovered that the number of generations and population size were both significant factors. The authors also concluded that high levels of mutation probability resulted in lower penalties, whilst crossover probability was discovered not to be statistically significant.

It is apparent from the work reviewed here that discrepancies exits with respect to the selection of the most appropriate structure for the GA. These discrepancies are most probably due to differences within the problem domain, indicating that the GA should be tailored to the problem to achieve efficient operation. Section 2 describes the GA used within this investigation, whilst Section 3 discusses the application of sparse experimental design for the selection of the most appropriate GA structure. Conclusions are drawn within Section 4.

2 GENETIC ALGORITHM STRUCTURE

Within this research, the general procedure for Genetic Algorithms developed by Goldberg [8] has been modified such that infeasible schedules may be repaired using an approach based upon precedence adjustments. The objective of the GA is to minimise the penalties due to the early supply of components and assemblies and the late delivery of final products, whilst simultaneously considering capacity utilisation. This ensures timeliness and appropriateness within the coordination of component manufacture and assembly operations. The algorithm is illustrated within Figure 1.

The gene is encoded using an alphanumeric string containing two parts. The first part represents the operation number and the second represents the part code. Process times, date due and assembly relationships may be obtained from the part code.

The genes are then randomly combined to produce a population of chromosomes or candidate solutions. The chromosome is consequently divided into n sub-chromosomes which represent the sequence of activities for the n resources – see Figure 2.

The fitness function then evaluates each chromosome with respect to the total sum of earliness and tardiness costs. The probability of survival, and the number of replicates of a chromosome, is determined based upon its fitness using standard roulette wheel techniques – Goldberg [8]. Those chromosomes not selected using the roulette wheel are removed from the population.

Two chromosomes are then selected at random from the population and based upon the likelihood of the crossover probability have a crossover operator performed on them to produce two new permutations containing genetic material from the original chromosomes. Those chromosomes that do not have the crossover operator performed upon them are carried through.

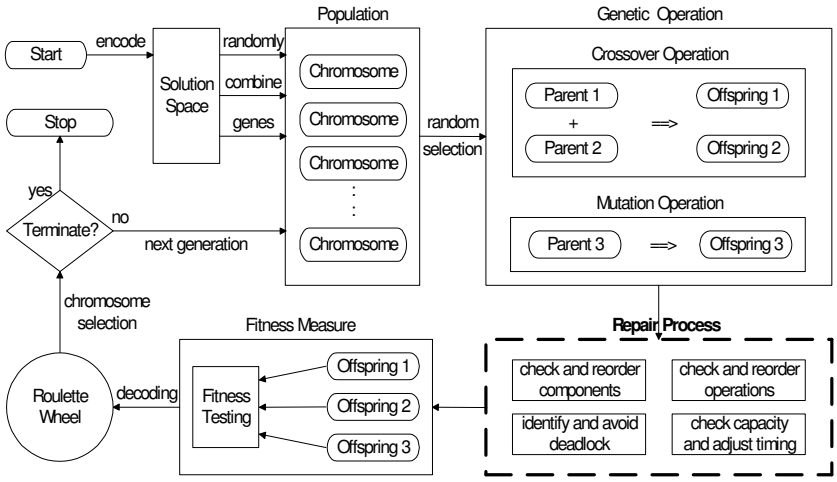

Figure 1. A general structure of Genetic Algorithms developed for production scheduling.

Similarly the mutation operator has a mutation probability likelihood to produce random changes within a randomly selected chromosome. Those chromosomes that do not have the mutation operator performed upon them are carried through.

Both the crossover and mutations operators when applied to chromosomes representing sequences of activities within a scheduling problem may, however, produce an impossible routing or assembly sequence since the operators are probabilistic and contain no domain-specific knowledge about the scheduling problem.

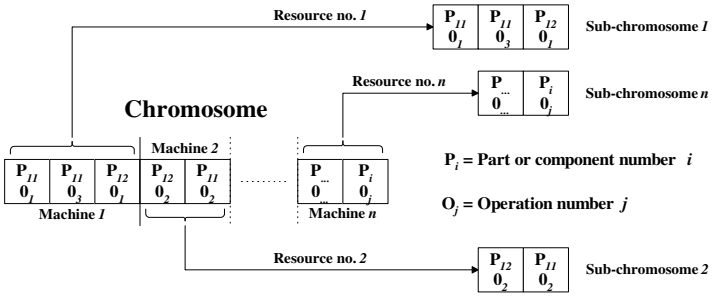

Figure 2. Sub-chromosome representation of resource sequence

Infeasible schedules are rectified using a four-stage repair process:

- Impossible routings are identified and converted into feasible sequences by reordering the operations.
- Ensuring that all components and sub-assemblies are correctly sequenced prior to their subsequent assembly by checking and reordering the precedence.
- Ensuring that all timing constraints are satisfied regarding the completion of previous operations. This may introduce a delay between operations.
- Identifying and avoiding deadlock through the cyclical dependencies of operations and resources.

Table 1 lists eleven crossover operators, whilst Table 2 lists eight different mutation operators. The operators indicated with one tick are implemented within the GA scheduling program, whilst those marked with two ticks are investigated within the

experimental programme described within this paper. A full description of these and other operators is given by Pongcharoen et al. [13].

Initial	Description	Reference	BCGA
CX	Cycling crossover	Oliver et al. [12]	✓✓
EERX	Enhanced edge recombination crossover	Starkweather et al. [20]	✓✓
MPX	Maximal preservation crossover	Mühlenbein et al. [10]	✓✓
1PX	One point crossover	Murata & Ishibuchi [11]	✓✓
OX	Order crossover	Davis [2]	✓✓
PBX	Position based crossover	Syswerda [21]	✓✓
PMX	Partial matching crossover	Goldberg & Lingle [7]	✓✓
LOX	Linear order crossover	Falkenauer et al. [5]	✓
2PEX	Two point end crossover	Murata [11]	✓
2PCX	Two points centre crossover	Murata [11]	✓✓
DX	Diagonal (three parent) crossover	Eiben et al. [4]	✓

Table 1. Crossover operations

Initial	Description	Reference	BCGA
2OAS	Two operations adjacent swap	Murata [11]	✓✓
3OAS	Three operations adjacent swap	Murata [11]	✓✓
2ORS	Two operations random swap	Murata [11]	✓✓
3ORS	Three operations random swap	Murata [11]	✓✓
IM	Inverse mutation	Goldberg [7]	✓✓
SOM	Shift operation mutation	Murata [11]	✓✓
CIM	Centre Inverse mutation	Tralle [24]	✓✓
E2ORS	Enhanced two operations random swap	Tralle [24]	✓✓

Table 2. Mutation operations

3 APPLICATION OF DESIGNED EXPERIMENTS

3.1 Screening experiment

A sequential experimental strategy was adopted to identify the genetic operators and parameters that produce results with minimum total cost. The factors considered in the screening experiment were: the combination of population size and number of generations (P/G); the probabilities of crossover (%C) and mutation (%M), and, the crossover (COP) and mutation operators (MOP). An L_4 fractional factorial design embedded within an 8 level Latin Square was used for the screening experiment. The total number of possible combinations of factors at each level in this case is $8^2\%2^3 = 512$. A full factorial experiment of all these would take approximately 1280 hours to complete at 2.5 hours per run. However the screening experiment involved only a sub-set of 32 of the possible 512 runs, replicated using two random seeds, which took less than 80 hours in total. This is a saving of about 1200 hours, being approximately 94% of the potential spend.

Mutation operators	Crossover operators							
	CX	EERX	MPX	1PX	OX	PBX	PMX	2PCX
2OAS	A					D	C	B
3OAS	B	A					D	C
2ORS	C	B	A					D
3ORS	D	C	B	A				
IM		D	C	B	A			
SOM			D	C	B	A		
E2ORS				D	C	B	A	
CIM					D	C	B	A

Parameter settings			
Combine	P/G	%C	%M
A	60/20	0.9	0.18
B	60/20	0.3	0.02
C	20/60	0.3	0.18
D	20/60	0.9	0.02

Coded parameter settings			
Combine	P/G	%C	%M
A	1	1	1
B	1	-1	-1
C	-1	-1	1
D	-1	1	-1

Table 3. Screening stage design

3.2 Screening experiment results

The screening results were analysed using the general linear model form of ANOVA which is one of the most effective methods for analysing a balanced combination of categorical and non-categorical factors – Draper and Smith [3]. The analysis of variance for the two replicates within the screening experiment is shown within Table 4 which includes the estimates of the main effects as well as the two-way interactions between the random seed and the probability of probability of mutation, crossover and the P/G combination. The seed interaction was investigated because it is a potential nuisance factor, and it can be seen that there may be a difference between seeds with respect to the effect of P/G. Sexton et al. [17] advised not to ignore p values of less than 0.2 in screening experiments with lower power, however, many screening experiments consider effects with much bigger p values than this as potentially significant – Box and Liu [1].

Source	DF	SS	MS	F	P
P/G (%C-%M)	1	243.48	243.48	7.06	0.011
%C (P/G-%M)	1	195.78	195.78	5.68	0.022
%M (P/G-%C)	1	21.39	21.39	0.62	0.435
COP	7	232.58	33.23	0.96	0.470
MOP	7	164.23	23.46	0.68	0.688
Seed	1	4.96	4.96	0.14	0.706
Seed * %M	1	55.66	55.66	1.61	0.211
Seed * %C	1	43.82	43.82	1.27	0.266
Seed * P/G	1	76.15	76.15	2.21	0.145
Error	42	1448.25	34.48		
Total	63	2486.30			

Table 4. Analysis of variance for screening experiment

Table 4 also indicates that the population/generation (P/G) combination and the probability of crossover, or the interaction (P/G*%M) with which it is confounded, is significant. Earlier work, Pongcharoen et al. [15],[16], suggested that the P/G*%M interaction was likely to be active. All the two factor interactions are confounded with the main effects in this design, and hence cannot be estimated separately.
Standard ANOVA assumes that there is common residual variance across the design space. The possibility of the effect of some apparent non-homogeneity of variance was investigated via the application of a logarithmic transform to the response, a procedure described by Grove and Davis [9]. Fears that this may be a problem proved unfounded.

Crossover Operators	Mean £k	Standard Deviation	Mutation Operators	Mean £k	Standard Deviation
EERX	103.2	2.28	2OAS	104.6	2.28
CX	104.5	2.28	IM	105.4	2.28
PBX	105.3	2.28	CIM	105.4	2.28
1PX	107.0	2.28	SOM	105.9	2.28
PMX	107.3	2.28	2ORS	106.1	2.28
MPX	107.6	2.28	E2ORS	107.7	2.28
OX	108.6	2.28	3ORS	108.0	2.28
2PCX	109.6	2.28	3OAS	110.0	2.28

Table 5. Relative performance of crossover and mutation operators

The result of the analysis of variance for the screening experiment did not find the effect of different crossover and mutation operators to be statistically significant. However, Table 5 shows that there are differences in the results obtained with different operators which are of practical significance. It can be seen that the enhanced edge recombination crossover (EERX) algorithm had the lowest mean penalty costs, whilst the two point centre crossover (2PCX) had the highest mean. Similarly, the two operation adjacent swap (2OAS) mutation operator produced the 'best' mean result, whilst the three operation adjacent swap (3OAS) produced the 'worst'. These differences could be established as truly significant if greater statistical power were applied by running more trials. This can also be established by picking the 'best' and 'worst' for further investigation.

3.3 Second stage experiment

A second experiment was designed to further explore the relative performance of some of the genetic operators. The experimental design for the second stage was a 2_{IV}^{5-1} design, which included the same levels of population/generation combination and probabilities of crossover and mutation used in the screening experiment. However, only the highest and lowest scoring crossover and mutation operators from the screening stage were considered. Note that in the first two (of four) replications of this L_{16} design only 14 of the 16 combinations were additional runs, as the other two had already been completed during the screening stage.

3.4 Second stage results

The analysis of variance for the factors and/or their two factor interactions that were found to be statistically significant is displayed within Table 6. The probability of mutation is

seen to be important since, results being achieved with higher probabilities, thus confirming the earlier findings from Pongcharoen et al. [15],[16]. Note that the screening experiment previously found the probability of mutation to be statistically insignificant. The crossover probability is not statistically significant within the range considered, confirming the findings of earlier studies, and it is now clear that the screening stage results may be due to the effect of the confounding. It is apparent that the significance of factors may be affected by either the specific GA application, or by the operators used.
The difference in performance between the chosen crossover and mutation operators is established as statistically significant in the second stage experiment. All the other factors are either significant on their own, or in conjunction with other factors (as interactions). Both the effect of different seeds and crossover operator needs to be considered when setting an optimum level for the probability of crossover and choice of P/G level respectively. These tests are more conclusive than the screening experiment partly because there are fewer factor levels considered and as a result there is greater power to distinguish between the effect of changed factor settings in relation to the number of tests.

Source	DF	SS	MS	F	P
%M	1	101.01	101.01	7.36	0.009
%C	1	27.38	27.38	1.99	0.164
MOP	1	116.42	116.42	8.84	0.005
COP	1	138.58	138.58	10.09	0.003
P/G	1	0.63	0.63	0.05	0.832
Seed	3	54.14	18.05	1.31	0.280
COP * P/G	1	56.91	56.91	4.14	0.047
Seed * %C	3	107.51	35.84	2.61	0.061
Error	51	700.30	13.73		
Total	63	1302.88			

Table 6. Analysis of variance for the investigation of the significance of genetic operations

In practice, additional replicates were run until the statistical power was sufficient to identify the statistical significance of the genetic operators. After two replicates, only the mutation operators had still not been established as statistically significant. Again a ln(y) data transformation was considered due to the possibility of non-homogenous variance, but was found to have no effect on results. Both histograms of residuals and probability plots produced satisfactory results. The relative performances of the crossover and mutation operators are shown in Table 7.

Crossover and mutation operators	Mean	Standard deviation
Enhanced Edge Recombination Crossover (EERX)	105.5	0.692
Two Points Centre Crossover (2PCX)	108.4	0.692
Two Operations Adjacent Swap (2OAS)	105.4	0.692
Three Operations Adjacent Swap (3OAS)	108.5	0.692

Table 7. Relative performance of crossover and mutation operators

The best results were obtained with the enhanced edge recombination crossover (EERX) and the two point adjacent swap (2OAS) mutation operator confirming both results

obtained by Todd [22] and from within the screening experiment. The standard deviation was reduced from 2.28 to 0.692, which was due to improved statistical power.
Table 8 shows the coefficients from the best regression model with operators coded as ±1. The interaction COP*P/G has a negative value. EERX was coded as –1 and the result indicates that a low value of P/G should be chosen to get the best result with this operator.

Predictor	Coefficient	Standard deviation	P value
Constant	106.975	0.485	0.000
COP	+1.4715	0.485	0.004
MOP	-1.5615	0.485	0.002
%M	-1.2563	0.485	0.012
COP * P/G	-0.9430	0.485	0.057

Table 8. Regression analysis

The interactions between the factors are shown in Figure 3. EERX produces slightly better results than 2PCX for a high value of P/G, however with a low value of P/G, the difference is more pronounced.
This information can be used to determine the optimum combination of the Genetic Algorithm operators and parameters. The levels of the factors which lead to the lowest penalty costs are: crossover operator EERX and mutation operator 2OAS, a low setting of P/G, a high probability of mutation and, based upon screening experiment findings, a low probability of crossover. A test with more statistical power might find that the probability of crossover is significant.

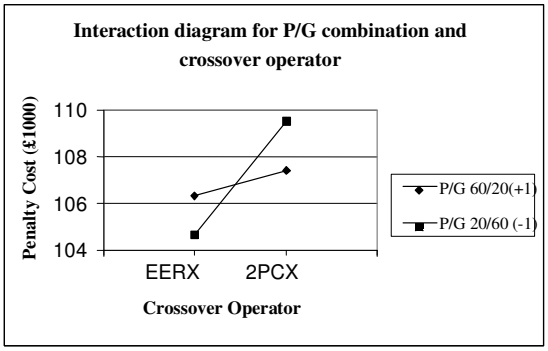

Figure 3. Interaction diagram for P/G combination and crossover operator

The population-generation and probability factors identified are the best over the range of operators tested. It also seems that several of the operators will perform well at these settings. However the model suggests that even better results may be obtained with a higher mutation probability than 0.18.
The findings differ from the earlier work with regards to the setting of P/G. In Pongcharoen et al. [15],[16] the best setting was a high P/G combination but in that work the operators 1PX and IM were used. In another study (Garzon et al. [6]) both the probability of mutation and crossover were found to be statistically significant. This is all further evidence that the importance of particular factors may be case, application or fitness function specific.

4 CONCLUSIONS

The performance of the Genetic Algorithm scheduling tool is influenced by a large number of factors. The investigation of these requires an efficient experimental design to enable the work to be performed within a reasonable time. A screening experiment was performed in which a fractional factorial was embedded within a half Latin Square. This is a novel experimental design. At this stage it was found that the population/generation (P/G) combination and the probability of crossover were statistically significant. Although the operators used were not statistically significant, differences in performance were obtained which were of practical importance.

The second stage experiment used a half-fraction design with a reduced number of GA operators. This increased the statistical power of the tests. It showed that the choice of operators was statistically significant. It also revealed interactions between the population/generation combination and the crossover operator used. The low level of P/G combination produced the best results when used with the Enhanced Edge Recombination crossover operator.

The use of experimental design has been very effective in minimising the amount of time and computational resources required. It has also enabled the discovery of good GA structures in the face of complex interactions that otherwise may have remained hidden. The indication of interactions between factors and the variety of different findings emerging from previous work suggests that appropriate GA operators and parameters may be case dependent. The use of efficient experimental designs to establish the best operators and parameters for particular applications appears to be a good strategy.

REFERENCES

[1] Box GEP & Liu PT, "Statistics as a Catalyst to Learning by Scientific Method Part I – An Example", *Journal of Quality Technology*, 31 pp1-15, 1999.

[2] Davis L, "Job shop scheduling and genetic algorithms", *Proceedings of the First International Conference on Genetic Algorithms and their Applications*, pp136-140, 1985.

[3] Draper NR & Smith H, *Applied Regression Analysis,* Wiley, New York, 1966.

[4] Eiben AE, Springhuizen-Kuyper IG & Thijssen BA, "Competing crossovers in an adaptive GA framework", *Proceedings of the IEEE Conference on Evolutionary Computation*, pp787-791, 1989.

[5] Falkenauer E & Bouffoix S, "A genetic algorithm for job shop", *Proceedings of the IEEE International Conference on Robotics and Automation*, 1991.

[6] Garzon IE, Taher MA & Anderson A, "Evaluating Taguchi tools through case studies", *Proceedings of the industrial statistics in action conference*, 2000.

[7] Goldberg DE & Lingle R, "Alleles, loci and the travelling salesman problem", *Proceedings of the First International Conference on Genetic Algorithms and Their Applications*, pp154-159, 1985.

[8] Goldberg DE, *Genetic Algorithms in Search, Optimisation and Machine Learning*, Addison-Wesley, Massachusetts, 1989.

[9] Grove DM & Davis TP, *Engineering, Quality & Experimental Design,* Longman, London, 1992.

[10] Mühlenbein H, Gorges-Schleuter M & Krämer O, "Evolution algorithms in combinatorial optimisation", *Parallel Computing*, 7, pp65-85, 1992.

[11] Murata T & Ishibuchi H, "Performance evaluation of genetic algorithms for flow shop scheduling problems", *Proceedings of the First IEEE Conference on Evolutionary Computation*, 2, pp812-817, 1994.

[12] Oliver IM, Smith CJ & Holland JRC, “A study of permutation crossovers on the travelling salesman problem”, *Proceedings of the Second International Conference on Genetic Algorithms and their Applications*, pp225-230, 1987.
[13] Pongcharoen P, Stewardson DJ, Hicks C & Braiden PM, "Applying designed experiments to optimise the performance of genetic algorithms used for scheduling products in the capital goods industry", *Journal of Applied Statistics*, 28(3), pp441-455.
[14] Pongcharoen P, Hicks C & Braiden PM, “The development of bi-criteria genetic algorithms for the finite capacity scheduling of complex products, with multiple levels of product structure”, *Submitted to European Journal of Operational Research*, 1999.
[15] Pongcharoen P, Hicks C, Braiden PM, Stewardson DJ & Metcalfe AV, “Using genetic algorithms for scheduling the production of capital goods”, *Proceedings of the Eleventh International Working Seminar on Production Economics*, February 21-25, Austria, 2000.
[16] Pongcharoen P, Hicks C, Braiden PM, Stewardson DJ & Metcalfe AV, “Using genetic algorithms for scheduling the production of capital goods”, *Submitted to International Journal of Production Economics*, 2000.
[17] Sexton CJ, Dunsmore W, Lewis SM, Please CP & Pitts G, “Semi-controlled experiment plans for improved mechanical engineering designs”, *Proceedings of the Institute of Mechanical Engineers*, 214b, pp95-105, 2000.
[18] Simpson AR & Goldberg DE, “Pipeline optimisation via Genetic Algorithms: from theory to practice”, *Water pipeline systems*, Miller DS MechEng Lon, pp309-320, 1994.
[19] Simpson AR, Vitkovsky J & Lambert M, “Transients for calibration of pipe roughnesses using genetic algorithms”, *8th International Conference on Pressure Surges*, Hague, Netherlands, pp587-597, 2000.
[20] Starkweather T, McDaniel S, Mathias K, Whitley D & Whitley C, “A comparison of genetic sequencing operators”, *Proceedings of the Third International Conference on Genetic Algorithms*, pp69-76, 1991.
[21] Syswerda G, “Scheduling optimisation using genetic algorithms”, *Handbook of Genetic Algorithms*, pp332-349, 1991.
[22] Todd D, “Multiple criteria genetic algorithms in engineering design and operation”, Ph.D. thesis, Engineering Design Centre, Department of Marine Technology, University of Newcastle upon Tyne, 1997.
[23] Todd D & Sen P, “A multiple criteria genetic algorithm for containership loading”, *Proceedings of the Seventh International Conference on Genetic Algorithms*, Morgan Kaufmann, ISBN 1-55860-487-1, pp674-681, 1997.
[24] Tralle D, “Analysing of genetic operations in genetic algorithms applied to optimisation of manufacturing systems”, *Undergraduate project*, Department of Mechanics, Materials and Manufacturing Engineering, University of Newcastle upon Tyne, 2000.