Introducing a New Advantage of Crossover: Commonality-Based Selection

Stephen Chen
The Robotics Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
chens@ri.cmu.edu
http://www.cs.cmu.edu/~chens

Stephen F. Smith
The Robotics Institute
Carnegie Mellon University
5000 Forbes Avenue
Pittsburgh, PA 15213
sfs@ri.cmu.edu
http://www.cs.cmu.edu/~sfs

Abstract

The Commonality-Based Crossover Framework defines crossover as a two-step process: 1) preserve the maximal common schema of two parents, and 2) complete the solution with a construction heuristic. In these "heuristic" operators, the first step is a form of selection. This commonality-based form of selection has been isolated in GENIE. Using random parent selection and a non-elitist generational replacement scheme, GENIE does not include fitness-based selection. However, a theoretical analysis shows that "ideal" construction heuristics in GENIE can potentially converge to optimal solutions. Experimentally, results show that the effectiveness of practical construction heuristics can be amplified by commonality-based restarts. Overall, it is shown that the commonality hypothesis is valid--schemata common to above-average solutions are indeed above average. Since common schemata can only be identified by multi-parent operators, commonality-based selection is a unique advantage that crossover can enjoy over mutation.

1 INTRODUCTION

The three basic features of a genetic algorithm (GA) are a population of solutions, fitness-based selection, and crossover [Hol75][Gol89]. Fitness-based selection is responsible for increasing the proportion of fit schemata in the population. This process allows exploitation of existing knowledge. Crossover recombines these schemata into new solutions--thereby allowing exploration to occur.

Traditionally, combination has been viewed as the primary mechanism and advantage of crossover. "This is, after all, the overt purpose of crossover." [Sys89] However, there is no guarantee that crossover combines the correct schemata. Thus, problem specific heuristics have been incorporated into crossover operators to help enhance the selection/exploration of schemata, e.g. Nearest Neighbor for the Traveling Salesman Problem (TSP) [GGR85]. In these "heuristic" operators, fit schemata (e.g. short edges) can be selected directly during crossover.

The Commonality-Based Crossover Framework presents a new model for designing (heuristic) crossover operators [CS98][CS99]. It defines crossover as a two-step process: 1) preserve the maximal common schema of two parents, and 2) complete the solution with a construction heuristic. The model follows from the commonality hypothesis which suggests that schemata common to above-average solutions are above average. Essentially, it is believed that the common schemata of two (parent) solutions are most likely responsible for their (high) observed fitness.

The commonality hypothesis attempts to explicitly identify the good schemata that the offspring should inherit from its parents. Conversely, (random) combination can lead to "hitch-hiking" (poor schemata enter the offspring along with the good schemata). Since the quality of the uncommon schemata is unknown, a commonality-based operator preserves only the common (fit) schemata. The solution is then completed with (new) heuristically generated schemata.

When following the new design model, the actions of the first step cause common schemata to be *selected*. This commonality-based form of selection is most easily observed with heuristically constructed solutions. In these solutions, the common schemata (of two parents) should have a higher ratio of fit to unfit schemata than either of the complete (parent) solutions. When this higher ratio occurs, it can be beneficial to restart the construction heuristic from this partial solution of common schemata. Specifically, if the construction heuristic is as effective from this restart point as it is when starting from scratch, the proportion of fit schemata in the offspring should be higher than in the parents.

To isolate the above commonality-based form of selection (in heuristic operators), the GENIE algorithm has been developed. GENIE uses random parent selection and a non-elitist generational replacement scheme. The selection of neither parents nor offspring is fitness based, so only commonality-based selection can cause the proportion of fit schemata to increase in the population. With these commonality-based restarts, the effectiveness of the embedded construction heuristic can be improved. This effect has been called "heuristic amplification".

The potential benefits of commonality-based selection are only available through multi-parent operators that can identify and preserve common schemata (e.g. standard crossover operators). Although all evolutionary algorithms use fitness-based selection, crossover can allow genetic algorithms to additionally benefit from the newly introduced commonality-based form of selection. Theoretical and experimental results with GENIE demonstrate that commonality-based selection can usefully identify schemata to exploit. In particular, the commonality hypothesis is validated--schemata common to above-average solutions are indeed above average.

The remainder of this paper is presented as follows. In section 2, an intuitive argument for preserving common schemata is presented. In section 3, this argument is formalized for an ideal construction heuristic. In section 4, the GENIE algorithm is defined. In section 5, an ideal construction heuristic for the One Max problem is presented. In section 6, two construction heuristics for the Traveling Salesman Problem (and their associated commonality-based heuristic operators) are examined. In section 7, commonality-based selection is compared with random restart. In section 8, implications of the theoretical and experimental results are discussed. Lastly, final conclusions are presented in section 9.

2 A REASON TO PRESERVE COMMON SCHEMATA

A (greedy) construction heuristic incrementally builds a solution one step at a time. At each step, the heuristic can make a correct decision or an incorrect decision. Assuming that a correct decision causes correct (fit) schemata to be selected, the quality of the solution will vary with the number of correct/incorrect decisions. Thus, increasing the number of correct decisions should also improve the quality of the final solution.

Assume that a construction heuristic makes correct and incorrect decisions with a constant ratio. Then, the number of incorrect decisions *made by the construction heuristic* should decrease if it is started from a partial solution--there are fewer steps where the heuristic can make an incorrect decision. If the partial start solution has a higher proportion of correct decisions (fit schemata) than the construction heuristic normally produces, the final solution should also have a higher proportion of fit schemata than a solution constructed from scratch. Thus, construction heuristics may be more effective if they are (re)started from partial solutions with high proportions of fit schemata.

The common schemata of two heuristically constructed solutions is a partial solution that should have high proportions of fit schemata. For example, consider the Nearest Neighbor construction heuristic for the TSP. This heuristic starts at a random city and travels to the nearest unvisited city at each step. In this process, Nearest Neighbor first selects many short (fit) edges, but after myopically "painting itself into a corner", a long (unfit) edge must be

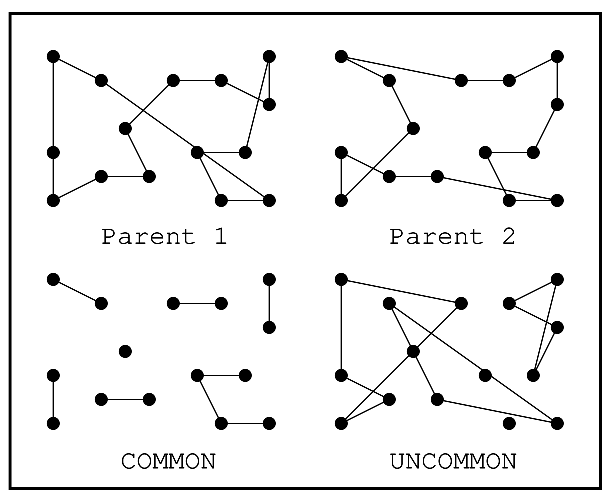

Figure 1: Example of two Nearest Neighbor (parent) solutions. Their uncommon edges tend to be the long/crossing edges.

selected. Compared to the selection of short edges, the selection of long edges is more dependent on the start city. Thus, two Nearest Neighbor tours are likely to have the same short edges, but different long edges. (See Figure 1.) Consequently, the partial solution of common edges likely has a higher ratio of short to long edges (fit to unfit schemata) than a complete Nearest Neighbor (parent) solution.

3 THEORETICAL PERFORMANCE

To measure the potential increase in effectiveness, assume that a construction heuristic selects correct schemata (i.e. schemata that are part of the optimal solution) with probability p and incorrect schemata with probability $1-p$. If this heuristic is used to generate an initial population, each (parent) solution is expected to have a proportion p of correct schemata, and a proportion $1-p$ of incorrect schemata. For random parent selection, Table 1 shows the expected distribution of correct/incorrect and common/uncommon schemata for parent pairs in the initial population.

Table 1: Expected distribution of schemata for random parent pairs in the initial population.

	p correct	$1-p$ incorrect
p correct	p^2 common correct	$p(1-p)$ uncommon
$1-p$ incorrect	$p(1-p)$ uncommon	$(1-p)^2$ common incorrect

Among the common schemata, the ratio of correct to incorrect schemata is $p^2/(1-p)^2$. If the construction heuristic selects more correct schemata than incorrect schemata (i.e. $p > 0.5$), then $p/(1-p) > 1.0$ and

$$\frac{p^2}{(1-p)^2} > \frac{p}{1-p}.$$

The common-schema partial solutions are expected to have a higher ratio of correct to incorrect schemata than (parent) solutions of the initial population. The decision to exploit these schemata has been achieved *without fitness-based selection*.

Numerically, the proportion of correct schemata in the initial population is $c_0 = p$, and the proportion of incorrect schemata is $\bar{c}_o = 1-p$. To extend the analysis, assume that the above construction heuristic is embedded into a heuristic operator by following the Commonality-Based Crossover Framework[1]. Then, for a GA with random parent selection and generational replacement, the correct schemata in generation i are the common correct schemata from generation $i-1$ and the correct schemata selected by the construction heuristic during generation i. Assuming that a constant proportion p of correct schemata is generated (i.e. the construction heuristic is *ideal*), the expected proportion of correct schemata c_i in generation i is

$$c_i = c_{i-1} \cdot c_{i-1} + p[2 \cdot c_{i-1} \cdot \bar{c}_{i-1}].$$

Simplifying,

$$c_i = c_{i-1} \cdot (c_{i-1} + 2p \cdot \bar{c}_{i-1})$$

Representing $p \in [0,1]$ as $0.5(1+x)$, $x \in [-1,1]$;

$$\frac{c_i}{c_{i-1}} = c_{i-1} + 2p \cdot \bar{c}_{i-1}$$

$$\frac{c_i}{c_{i-1}} = c_{i-1} + (1+x) \cdot (1-c_{i-1})$$

$$\frac{c_i}{c_{i-1}} = c_{i-1} + 1 - c_{i-1} + x - xc_{i-1}$$

$$\frac{c_i}{c_{i-1}} = 1 + x(1-c_{i-1})$$

If $p > 0.5$, then $x > 0$ and $c_i > c_{i-1}$. The proportion of correct schemata increases with each generation when $p > 0.5$.

Further, at convergence, $c_* = c_i = c_{i-1}$. Therefore, c_* must satisfy:

$$c_* = c_*^2 + p[2 \cdot c_* \cdot (1-c_*)]$$

$$c_* = c_*^2 + 2pc_* - 2pc_*^2$$

$$c_*(1-2p) = c_*^2(1-2p)$$

This equality requires $c_* = 0$, $c_* = 1$, or $p = 0.5$.

For $p > 0.5$, c_i converges to 1 (all solutions in the population are optimal). However, c_i converges to 0 for $p < 0.5$. (See Figure 2.) Therefore, the effectiveness (or ineffectiveness) of construction heuristics is amplified by commonality-based selection. This effect is called *heuristic amplification*.

[1] Specifically, the construction heuristic is used to complete a solution that is started from the common schemata of two parents.

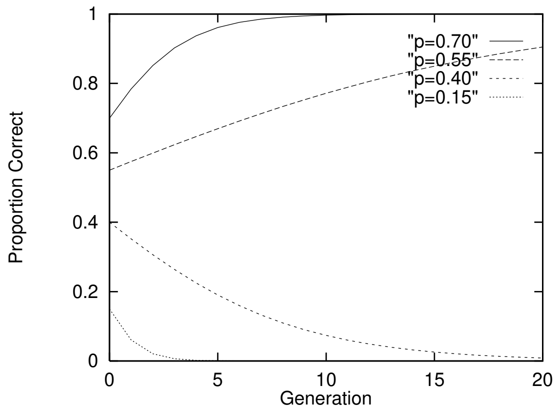

Figure 2: Expected proportion of correct schemata in generation `i` for commonality-based crossover operators encapsulating ideal construction heuristics with different `p`.

4 THE GENIE ALGORITHM

To mimic the development of the theoretical results, the GENIE algorithm is defined to have random parent selection and a non-elitist generational replacement scheme. Further, to stay as close as possible to the expectations, each parent mates twice (with random partners) during each generation. However, the theoretical results imply an infinite population, and this is of course infeasible.

5 ONE MAX EXPERIMENTS

In One Max, the correct gene for each allele is a `1`. Therefore, a trivial heuristic for One Max is to select more `1`'s than `0`'s. For example, select a `1` for each allele with a (constant) probability of $p = 0.6$, and a `0` with a probability of $1 - p = 0.4$. This (construction) heuristic is ideal because the decision at each step has a constant (and independent) probability of being correct.

The above heuristic has been embedded into a commonality-based heuristic operator which has been implemented in GENIE. For a One Max problem of 100 bits, the experimental results with GENIE nearly match the theoretical expectations when a population size of 100 solutions is used. (See Figure 3.) The results are not surprising because this experiment trivially fits the previously derived equations--each decision is independent, and the construction heuristic is ideal.

6 TRAVELING SALESMAN EXPERIMENTS

The Traveling Salesman Problem is a benchmark combinatorial optimization problem. The objective is to find the shortest Hamiltonian cycle through a complete graph of `n` nodes. A feasible TSP solution has constraints (i.e. each node must be visited once and only once). Thus, each decision of a construction heuristic is not independent, so it is unlikely that a constant performance ratio can be maintained. In particular, the correct decision may be disallowed at a given step.

6.1 EDGE-BASED HEURISTIC

Nearest Neighbor has been embedded into a commonality-based heuristic operator--Common Sub-Tours/Nearest Neighbor (CST/NN) [CS98]. This operator has been implemented in GENIE. The results are disappointing as CST/NN in GENIE demonstrates almost no heuristic amplification--only 14% of the initial surplus is reduced from the best start solutions to the best final solutions. (See Table 2.)

For the TSP, it is not just the number of correct/incorrect edges that determines solution quality. The length (relative fitness) of the incorrect edges (schemata) is also important. Further, starting with a partial solution of short edges does not necessarily reduce the tendency of Nearest Neighbor to "paint itself into a corner". Overall, the results suggest that Nearest Neighbor is a weak heuristic (very far from an ideal heuristic).

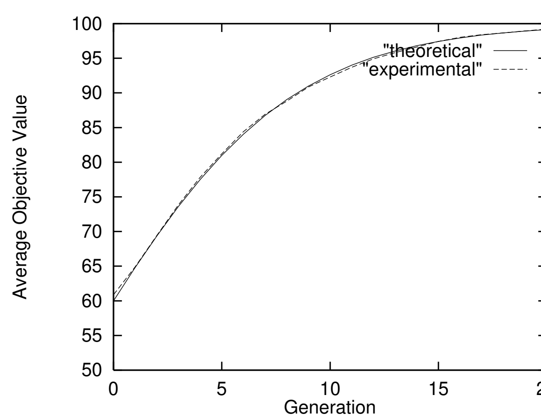

Figure 3: Expected and observed results for an ideal construction heuristic on One Max.

Table 2: Results for CST/NN in GENIE. Population size is equal to problem size. Initial population is NN started from each element. Values are percent surplus from known optimum for average of 5 runs (50 generations each).

TSPLIB Instance	Size	Avg. Best NN Start Tour	Avg. Best CST/NN Tour
d198	198	+ 12.42 %	+ 8.67 %
lin318	318	+ 17.06 %	+ 16.30 %
fl417	417	+ 16.92 %	+ 13.37 %
pcb442	442	+ 15.17 %	+ 13.30 %
u574	574	+ 19.92 %	+ 18.40 %
average		+ 16.30 %	+ 14.01 %

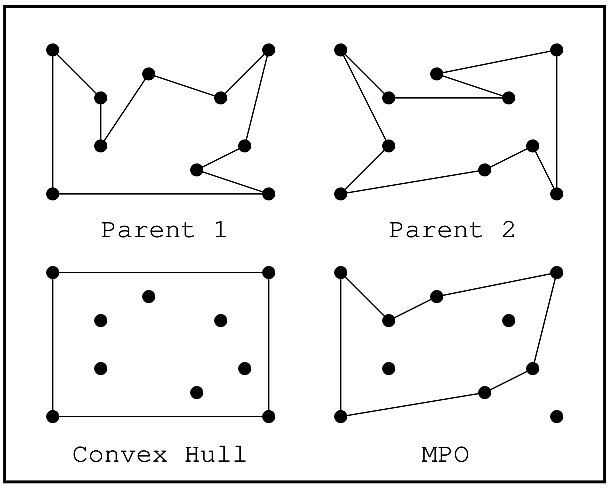

Figure 4: Example of "shape" from convex hull and MPO.

Table 3: Results for MPO/AI in GENIE. Population size is 400. Initial population is 400 AI solutions started from the convex hull. Values are percent surplus from known optimum for average of 5 runs (50 generations each).

TSPLIB Instance	Size	Avg. Best CH/AI Start Tour	Avg. Best MPO/AI Tour
d198	198	+ 3.05 %	+ 1.24 %
lin318	318	+ 6.04 %	+ 1.75 %
fl417	417	+ 1.91 %	+ 0.58 %
pcb442	442	+ 8.97 %	+ 3.48 %
u574	574	+ 8.45 %	+ 2.59 %
average		+ 5.68 %	+ 1.93 %

6.2 ORDER-BASED HEURISTIC

Arbitrary Insertion (AI) is an order-based construction heuristic for the TSP. Insertion heuristics are more effective when they are started from the convex hull than from three random nodes [Rei94]--the convex hull is a partial solution with more correct decisions. The Maximum Partial Order (MPO) is the largest partial solution that insertion can extend into both parents--it accumulates "shape" information. (See Figure 4.) Restarting Arbitrary Insertion from the MPO, the Maximum Partial Order/Arbitrary Insertion (MPO/AI) commonality-based heuristic operator has been developed [CS98].

Implementing MPO/AI in GENIE, the results are more promising--the surplus is reduced by 66% from the best start solutions to the best final solutions. (See Table 3.) Although Arbitrary Insertion is not an ideal construction heuristic (it does not perform at a constant ratio), commonality-based restarts have still amplified its effectiveness. Specifically, commonality-based selection has identified schemata of higher fitness than the convex hull.

7 INFORMATION ACCUMULATION IN GENIE

Without fitness-based selection, it may be difficult to see how GENIE does anything more than random restart. Indeed, it may be argued that GENIE does in fact do nothing--it lets crossover do everything. The difference between the random restart of a construction heuristic and a heuristic crossover operator is the use of a partial solution for restarts.

To demonstrate the advantage provided by a partial solution chosen through commonality-based selection, an experimental run of GENIE has been traced. For a run of MPO/AI on the lin318 TSP instance, the average length of the MPO used to generate the offspring was recorded for each generation. (See Figure 5.) The size of the partial solution preserved by commonality-based selection grows rapidly through generation 27. Similarly, the average quality of the MPO/AI solutions in each generation of GENIE also improves steadily until generation 28. After this phase of convergent search, a "drift" phase begins which appears to degrade solution quality.

To baseline the MPO/AI results, 50 generations of random restarts (Arbitrary Insertion from the convex hull) are also plotted. Obviously, the convex hull (CH) is a static entity--it does not accumulate information from previous solutions. Without information accumulation, the solution

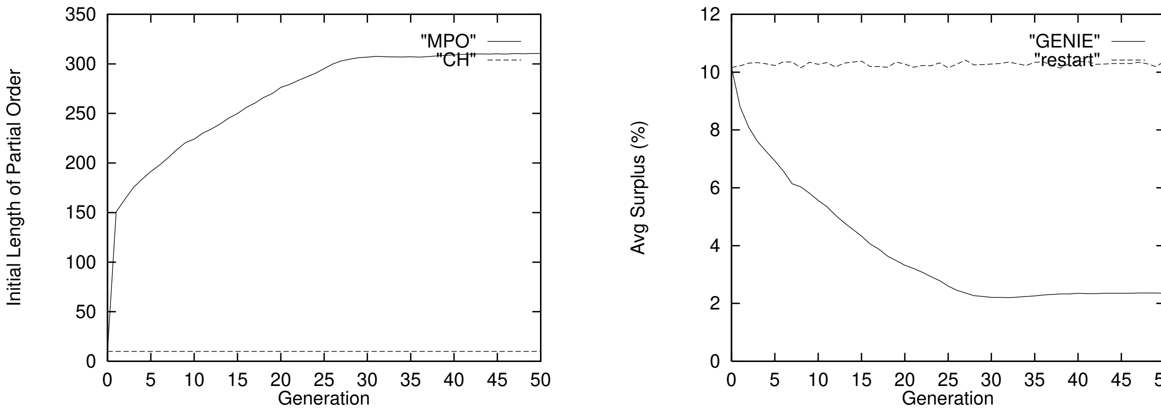

Figure 5: Comparison of commonality-based selection (in GENIE) and random restart.

quality cannot improve over time. These results demonstrate the difference between a random restart and a restart from a partial solution chosen by commonality-based selection.

8 DISCUSSION

In this paper, commonality-based selection has been isolated. This discovery introduces it as a new advantage for crossover in genetic algorithms. It also suggests a means to increase the effectiveness of practical construction heuristics.

8.1 THE ROLE OF COMMONALITY-BASED SELECTION IN STANDARD CROSSOVER

Many (evolutionary) algorithms use populations and/or fitness-based selection. However, only genetic algorithms use crossover. “[Crossover] is regarded as the distinguishing feature of [genetic] algorithms ... and as a critical accelerator of the search process” [Dav91]. Traditionally, the advantage provided by crossover has been attributed to the mechanism of combination. Unfortunately, it has been difficult to quantify this advantage in practice.

```
Parent 1:      1 0 1 1 0 1 0 1 1 1
Parent 2:      1 1 0 1 0 1 1 1 0 1

Common:        1     1 0 1   1   1

Uncommon 1:      0 1       0   1
Uncommon 2:      1 0       1   0
```

Figure 6: A beneficial mutation is more likely if mutation is restricted to uncommon schemata only.

However, it has previously been shown that (sequencing) operators which use only combination (e.g. Order Crossover [Dav85]) can be improved if they are redesigned to also preserve common schemata [CS98]. In fact, current guidelines for the design of recombination operators suggest that common components should be preserved [Rad91][EMS96]. This action allows crossover to employ commonality-based selection.

For example, in One Max, consider two parents with above-average fitness (i.e. both have more `1`’s than `0`’s). If a random bit is selected for mutation, it is more likely that a `1` will be mutated into a `0`, than vice-versa. Thus, a beneficial mutation is less likely than a deleterious mutation.

Conversely, if common schemata (bits) are preserved, the remaining uncommon bits should have as many `1`’s as `0`’s (see Table 1). Therefore, a beneficial mutation becomes as likely as a deleterious mutation. Specifically, beneficial mutations are more likely after common schemata have been preserved. (See Figure 6.) If crossover is viewed as “structured mutation”, it is structured mutation with the benefit of commonality-based selection.

With heuristic operators, the effects of commonality-based selection have been isolated in GENIE. These results validate the commonality hypothesis--schemata common to above-average solutions are indeed above average. To identify common schemata, multi-parent operators (e.g. crossover) must be employed. Thus, the ability to enhance fitness-based selection with commonality-based selection is a unique advantage that crossover can enjoy over mutation.

8.2 HEURISTIC AMPLIFICATION

The (validated) commonality hypothesis provides a “confidence measure” on the performance of heuristics. The (partial) intelligence that is incorporated into heuristics is fallible. When failure occurs, the behavior of a heuristic is mistake-prone and unpredictable. However, this unpredictability also makes it unlikely that the same mistake is independently produced by separate applications. Thus, it is proposed that the uncommon decisions should be labelled as potential mistakes. Conversely, when common decisions have been made by the heuristic, it is reasonable to believe that the heuristic has acted as intended. Explicitly, common decisions are likely to be above average (good), and uncommon decisions are likely to be below average (bad).

This commonality-based confidence measure can be used to amplify the effectiveness of a construction heuristic. Unlike random restarts (from scratch), a heuristic operator can use commonality-based selection to identify highly-fit partial solutions. With this accumulated knowledge, a construction heuristic can ideally develop better final solutions when it is (re)started from these (common schema) partial solutions. For problem domains where the random restart of a (global search) construction heuristic is the primary optimization method, commonality-based restarts (through heuristic operators) should be more effective.

9 CONCLUSIONS

Fitness-based selection is fundamental to all evolutionary algorithms, including genetic algorithms. However, the Commonality-Based Crossover Framework suggests that common schemata should be preserved. This is a form of selection. Theoretical and experimental results demonstrate that this commonality-based form of selection is capable of identifying above-average schemata. In particular, the effects of commonality-based selection have been isolated in GENIE, a genetic algorithm that does not include fitness-based selection. Overall, commonality-based selection is presented as an advantage that multi-parent operators like crossover can have over single-parent operators like mutation.

Acknowledgments

The work described in this paper was sponsored in part by the Advanced Research Projects Agency and Rome Laboratory, Air Force Material Command, USAF, under grant numbers F30602-95-1-0018 and F30602-97-C-0227, and the CMU Robotics Institute. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Advanced Research Projects Agency and Rome Laboratory or the U.S. Government.

References

[CS98] S. Chen and S.F. Smith. (1998) "Experiments on Commonality in Sequencing Operators." In *Genetic Programming 1998: Proceedings of the Third Annual Conference*.

[CS99] S. Chen and S.F. Smith. (1999) "Putting the "Genetics" back into Genetic Algorithms (Reconsidering the Role of Crossover in Hybrid Operators)." To appear in *Foundations of Genetic Algorithms 5*, W. Banzhaf and C. Reeves, eds. Morgan Kaufmann.

[Dav85] L. Davis. (1985) "Applying Adaptive Algorithms to Epistatic Domains." In *Proc. Ninth International Joint Conference on Artificial Intelligence*.

[Dav91] L. Davis. (1991) *Handbook of Genetic Algorithms*. Van Nostrand Reinhold.

[EMS96] L.J. Eshelman, K.E. Mathias, and J.D. Schaffer. (1996) "Convergence Controlled Variation." In *Foundations of Genetic Algorithms 4*, R. Belew and M. Vose, eds. Morgan Kaufmann.

[GGR85] J. Grefenstette, R. Gopal, B. Rosmaita, and D. Van Gucht. (1985) "Genetic Algorithms for the Traveling Salesman Problem." In *Proc. of an International Conference on Genetic Algorithms and their Applications*.

[Gol89] D. Goldberg. (1989) *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley.

[Hol75] J. Holland. (1975) *Adaptation in Natural and Artificial Systems*. The University of Michigan Press.

[Rad91] N.J. Radcliffe. (1991) "Forma Analysis and Random Respectful Recombination." In *Proc. Fourth International Conference on Genetic Algorithms*.

[Rei94] G. Reinelt. (1994) *The Traveling Salesman: Computational Solutions for TSP Applications*. Springer-Verlag.

[Sys89] G. Syswerda. (1989) "Uniform Crossover in Genetic Algorithms." In *Proc. Third International Conference on Genetic Algorithms*.