Search-Based Software Engineering

Mark Harman	Bryan F. Jones
Brunel University	School of Computing
Uxbridge	University of Glamorgan
Middlesex	Pontypridd
UB8 3PH	CF37 1DL
United Kingdom	United Kingdom
Tel: +44 (0)1895 274 000	Tel: +44 (0)1443 482730
Fax: +44 (0)1895 251 686	Fax: +44 (0)1443 482715
mark.harman@brunel.ac.uk	bfjones@glam.ac.uk

Abstract

This paper claims that a new field of Software Engineering research and practice is emerging: Search-Based Software Engineering. The paper argues that Software Engineering is ideal for the application of metaheuristic search techniques, such as genetic algorithms, simulated annealing and tabu search. Such search-based techniques could provide solutions to the difficult problems of balancing competing (and sometimes inconsistent) constraints and may suggest ways of finding acceptable solutions in situations where perfect solutions are either theoretically impossible or practically infeasible.

In order to develop the field of Search-Based Software Engineering, a reformulation of classic Software Engineering problems as search problems is required. The paper briefly sets out key ingredients for successful reformulation and evaluation criteria for Search-Based Software Engineering.

1 Introduction

Software engineers often face problems associated with the balancing of competing constraints, trade-offs between concerns and requirement imprecision. Perfect solutions are often either impossible or impractical and the nature of the problems often makes the definition of analytical algorithms problematic.

Like other engineering disciplines, Software Engineering is typically concerned with near optimal solutions or those which fall within a specified acceptable tolerance. It is precisely these factors which make robust metaheuristic search-based optimization techniques readily applicable.

Metaheuristic algorithms, such as genetic algorithms [17], simulated annealing [37] and tabu search [16] have been applied successfully to a number of engineering problems. For example, a literature survey of genetic algorithms reveals applications to, among others:

- mechanical engineering [24, 33]
- chemical engineering [8, 22]
- medical and biomedical engineering [29, 32, 40]
- civil engineering [1, 6, 14, 19]
- electronic engineering [5, 11, 25]

Genetic algorithm (GA) research and researchers have even received interest from observers in the field of Social Science. Though GA practitioners may not agree with the findings of sociologists [18], it is an indication of the wide appreciation of the significance of these search-based technologies that they should have penetrated the collective consciousness of even 'non technical' disciplines such as Social Science.

However, the discipline of Software Engineering appears to be unique with regard to the application of genetic algorithms (and similar search-based, metaheuristic optimisation techniques); metaheuristic algorithms have received comparatively little attention from software engineers in comparison with that which they have received from researchers and practitioners in the more established fields of engineering.

A literature search on 'Software Engineering' and 'genetic algorithms' reveals work within the areas of testing [20, 21, 36, 28, 38, 39] and cost estimation [12, 13], but few others. Work has also taken place on automatic programming using genetic programming [23] and parallelization [31, 44], which could be thought of specific topics within coding, which in turn, could be considered to be a part of Software Engineering. Even with the inclusion of this work on coding, the application of search-based techniques to problems in Software Engineering has been, hitherto, somewhat patchy. The thesis underpinning the present paper is that search-based metaheuristic optimisation techniques are highly applicable to Software Engineering and that their investigation and application to Software Engineering is long overdue. It is time for Software Engineering to catch up with its more mature counterparts in traditional fields of engineering.

Software Engineering problems are often typified [4, 3, 10, 27, 26, 30, 35] by the observations paraphrased below:

> There is usually a need to balance **competing constraints**.
>
> Occasionally there is a need to cope with **inconsistency**.
>
> There are often **many potential solutions**.
>
> There is typically **no perfect answer**
> ... but **good ones** can be **recognised**.
>
> There are sometimes **no precise rules for computing the best solution**.

These properties of Software Engineering are also to be found in other engineering disciplines and they are *precisely* the attributes which make the application of search-based techniques so attractive. In order to apply metaheuristic search-based techniques to Software Engineering, the problems software engineers face must be reformulated as search problems. In section 2 the paper briefly sets out the key ingredients for such a reformulation. Section 3 provides some criteria against which to judge the success (or otherwise) of an application of search-based techniques to a Software Engineering problem. This paper is essentially a manifesto for Search-Based Software Engineering. Section 4 sets out some research goals for Search-Based Software Engineering, while section 5 concludes with a call for more work in this area.

2 Reformulating Software Engineering as a Search Problem

The principal intention of this section is to demonstrate how conceptually simple is the reformulation of Software Engineering to Search-Based Software Engineering. It is hoped that the reader is convinced that, at least in principal, it will be possible to apply metaheuristic search to a large body of Software Engineering problems, where natural representations, fitness functions and operators suggest themselves.

In order to reformulate Software Engineering as a search problem, it will be necessary to define:

- a representation of the problem which is amenable to symbolic manipulation;
- a fitness function (defined in terms of this representation) and
- a set of manipulation operators.

In addition, the perception of the problem itself may need to shift, in order to recast it as a search-based problem. For example, in the case of testing, the problem becomes one of searching for test cases which satisfy some test-adequacy criterion.

These issues are covered in far more detail in the general literature on metaheuristic search [17, 41, 43]. They are only mentioned here to demonstrate that the reformulation suggested and consequent development of Search-Based Software Engineering, is conceptually feasible.

2.1 Representation

The representation of a candidate solution is critical to shaping the nature of the search problem. Representations which are frequently used in existing applications for problem parameters are floating point numbers and binary code. In the latter case, grey codes are generally preferred to 'pure binary' numbers [42], since successor decimals (for instance 7 and 8) are not near neighbours in a pure binary code (4 mutations are required to mutate the four differing digits, rather than one). Grey coding avoids this problem.

2.2 Fitness

The fitness function is the characterisation of what is considered to be a good solution. In measurement theoretical terms [34], the fitness function need merely impose an ordinal scale of measurement upon the individual solutions it is applied to. That is, it will generally be sufficient to know which of two candidate solutions is the better according to the properties (or set of properties) to be measured. The fitness function imposes a fitness landscape, the characteristics of which will both determine which search techniques are most applicable and will shed light on the nature of the problem and its candidate solutions in terms of their perceived fitness. Fitness landscapes should not be too flat, not should they have sharp maxima that can easily be missed. Often the fitness function which first occurs, requires tuning to avoid these problems and to help guide the search towards good solutions.

In some cases it may be hard to define a fitness function, because the artifact to be optimised may have aesthetic qualities which make the determination of an ordinal scale metric difficult. However, in such situations a search-based approach may still be applicable, as will be argued in Section 3.3.

2.3 Operators

Different search techniques use different operators. As a minimum requirement, it will be necessary to mutate an individual representation of a candidate solution to produce a representation of a different candidate solution. Clearly this is a very minimal requirement. It will make it possible to apply hill climbing approaches and certain forms of evolutionary computation. If it is also possible to determine the set of 'near neighbours' of a candidate solution (in terms of its representation) then simulated annealing and tabu search can be applied. If instead (or in addition), it is possible to sensibly cross-over two individuals (to produce a 'child' which retain characteristics of both 'parents') then genetic algorithms will be applicable.

This section has glossed over a lot of issues which underly the choice of representation, fitness function and operators. The intention was not to suggest that these issues are unimportant. On the contrary, such issues are likely to form the large body of research work to be undertaken to successfully move from speculation to application. However, this work will be highly domain dependent and so cannot be adequately covered in this overview paper.

3 Evaluation Criteria for Search-Based Software Engineering

This section sets out some guidelines against which an attempt to apply metaheuristics to Software Engineering can be validated. The base line criteria suggest essential goals for any search-based approach. For a more thorough validation, section 3.2 suggests ways in which Search-Based Software Engineering might be combined with Classical Software Engineering and the ways in which these combined approaches might be validated.

3.1 Base Line Validity

To achieve a measure of base line acceptability a metaheuristic technique must out-perform a purely random search. That is, metaheuristics should find better solutions or find them with less computational effort than random search. Metaheuristic techniques should be able to find values which compare well with known solutions and should tend to produce fitness values which are considered acceptable for a reasonable subclass of elements of the problem space. Of course, these latter two criteria are somewhat vague and depend upon the application domain. This section merely sets out some guidelines.

3.1.1 Random Search

Any search-based technique must perform better than a purely random search in order to qualify as worthy of consideration as a successful application of Search-Based Software Engineering. Random search therefore provides a lowest benchmark for the application of metaheuristic search. Where a metaheuristic search does not consistently outperform random search it must be rejected. However, where this occurs, it may well indicate a poor choice of search technique (for example hill climbing in a multi-modal 'rugged' fitness landscape), and therefore may invalidate merely the choice of technique or its implementation, rather than the whole notion of applying metaheuristics to the Software Engineering problem under consideration.

3.1.2 Discovery of Known Solutions

Even in situations where there is no known general algorithm for solving a problem, there may be known examples, constructed by hand, where individual solutions are known for particular elements of the problem domain. One natural approach to validation is, therefore, to establish whether the metaheuristic search is able to find solutions that compare well with these known individual solutions.

Of course, 'compare well' requires some elucidation. If the known solutions are provided for a small set of simple problems, then metaheuristic search should perform as well as or better. However, where hand constructed solutions are used to illustrate 'difficult problems' for which solutions are hard to define, the praise 'compare well' might be interpreted less stringently.

3.1.3 Discovery of Desirable Solutions

Since the fitness function is a measure of 'goodness' of the solutions, another natural approach to validation consists of seeing just how much the fitness increases and gathering empirical data to provide evidence of the kinds of solution which may be obtained for typical problem elements.

3.1.4 Efficiency

In many (but by no means all) cases, a search-based approach may be slower than an existing analytical approach, because the search will involve repeated trials of the fitness function to evaluate candidate solutions. However, so long as the search based technique can produce better solutions, there will be many software engineering applications where the search-based approach is more appropriate because quality overrides speed.

Where there exists a technique for producing solutions which is always applicable and which produces consistently high-quality solutions, a search based approach will clearly still be preferable if it produces solutions of equal or better quality more speedily.

3.2 Validation with respect to Existing Analytical Techniques

In this section the phrase 'analytical techniques' is used to describe any non search-based algorithm for constructing solutions to the Software Engineering problem in question.

Where there exist analytic techniques for constructing reasonable solutions, search-based techniques will still be applicable. For instance, where analytical techniques are known to favour certain forms of solution or to have biases which affect their behaviour, a search-based technique may find solutions which cannot be discovered by the analytical technique. Where analytical techniques are applicable only to a subset of the problem space, search techniques may provide a mechanism for 'filling in the gaps'. Where analytical techniques are not consistently good at constructing solutions, search-based techniques may be used as a 'second guess'. Finally, where analytic techniques exist, but are known to produce sub-optimal solutions, these existing techniques may be used to seed the metaheuristic search with an initial population of reasonable solutions. It would then be hoped that the metaheuristic search would improve upon these initial seeds.

For each of these situations, analytic techniques produce only a part of the overall answer and metaheuristic search can be used to augment the existing approach.

3.2.1 Avoiding Bias in Analytical Techniques

Where metaheuristic search is used to address bias in existing analytic solutions, the evaluation criteria will measure the novelty of solutions relative to those produced (and their relative fitness) compared to those produced by the analytic approaches. In such situations, even where the metaheuristic technique does not outperform the analytic techniques on fitness, it may be validated in terms of the additional insight gained from considering solutions previously thought atypical or non-standard.

3.2.2 Filling Problem Space Gaps left by Analytical Solutions

Where existing analytic techniques are not applicable to the entire problem space, the use of metaheuristics may be partially validated, simply by showing that they 'fill in the gaps'. However, a stronger validation would clearly involve demonstrating comparable (or improved) fitness of metaheuristic solutions relative to those produced by analytic algorithms. However, where problem subspaces uncovered by analytic algorithms represent 'harder' problems to solve than those which are covered, this criterion may be too stringent. In this situation validation may simply show that the metaheuristic technique is capable of producing better solutions for previously uncovered areas of the problem space than can be found with random search alone.

3.2.3 Optimisation of Partial or Suboptimal Analytical Solutions

Where metaheuristic techniques are used to provide second guesses or where such algorithms are seeded by an initial population of results from an existing analytic algorithm, the metaheuristic technique should clearly be shown to provide improvement (i.e. increased fitness).

A simple-minded hybrid approach could simply try both the existing approach and a search-based approach and select the best solution. Fortunately, any such combined analytical and metaheuristic technique need only produce better results than the pure analytical approach on 'a few' occasions in order to be considered worth while. The interpretation of 'a few' will depend upon the domain of application.

3.2.4 Choice of Technique and Fitness Function

The general application of metaheuristic techniques to Software Engineering presents a large number of choices. The representation, fitness functions, operators and search technique to apply must all be carefully considered. For some applications, a simple hill-climbing approach may produce either good or adequate results. For other problems the multi-modal nature of the fitness landscape may demand a more robust technique. In general, it will be wise to experiment with a variety of techniques, fitness functions and operators. Fortunately, the highly generic nature of search-based techniques support just this form of experimentation.

Finally, for some applications, the fitness function itself may require some form of validation. Where there exist well-understood and widely accepted metrics for assessing elements of the solution space, the fitness function will not require validation; it can simply be appropriated from the software measurement literature. However, in the software measurement literature there are few uncontroversial metrics [34, 15]. Where there is disagreement about how to measure candidate solutions, the fitness function used and the landscape it imposes on the solution space will require validation. In all cases, the algorithmic complexity of the computation of the fitness function may be an issue. In some application areas a computationally demanding fitness evaluation may render the approach impractical, even where searches are often highly productive with respect to the solutions found.

3.3 Psychological Considerations

Software Engineering involves aesthetic as well as technical qualities. It may be found, as with more aesthetic applications of metaheuristic search [7, 9, 2], that a non-automatic, human-based fitness evaluation is required. In such cases, the psychological implications of repeated trials on the same human subject (and the judgment biases that this may induce) should be considered.

There is also a 'human dimension' to Search-Based Software Engineering research where the solution produced requires an explanation. In most cases the fitness function should be well-understood, so that the solutions can be explained purely in terms of their superior fitness. However, it may also be the case that the metaheuristic search approach is applied to give insight into the fitness function itself. Where a metric which captures the property (or properties) of interest is hard to define, the search may be used to explore the fitness landscapes created by several candidate metrics/fitness functions.

4 Future Work

The authors expect to see a dramatic development of the field of Search-Based Software Engineering. The special issue of Information and Software Technology, of which this paper forms a part, represents the embryonic state of a rapidly growing sub-area in Software Engineering research. This section sets out some goals and milestones for the emergent Search-Based Software Engineering research community.

It seems likely that within five years, metaheuristic search will have been applied to, at least, the following areas of Software Engineering. (To the authors' knowledge, work is already in progress in several of these areas.)

1. Requirements prioritisation
2. Finding good designs
3. Test data selection
4. Reverse and Re-engineering through transformation and re-factoring
5. Development of Software Measurement
6. User-based fitness evaluation for aesthetic aspects of Software Engineering

In addition, it is hoped that the field of Search-Based Software Engineering will provide:

1. Comparative studies and evaluations of different techniques as applied to Software Engineering problems. In an embryonic field of research, it is natural that many results will represent isolated advances in related but slightly different directions. As the field develops and matures, it is reasonable to hope for and expect that comparative studies and surveys will start to appear. The relative number of such studies might be thought of as a crude first approximation to an assessment of the maturity of Search-Based Software Engineering.

2. Theoretical development, in order to answer questions such as 'which technique/operators are likely to be fruitful for Software Engineering problem X?'.
   Comparative studies will help to answer questions like 'which technique works best for problem X?'. In addition, it is to be hoped that development of Search-Based Software Engineering will be deep as well as broad; that theoretical studies will emerge to complement the empirical work in the field.

3. Analyses of fitness landscapes for Software Engineering problems.
   This may yield deeper insight into the nature of the problems themselves.

4. Tailoring of fitness functions.
   Examination of the behaviours of metaheuristic search, may suggest modifications to the fitness landscape in order to improve the success of the search. Such modifications may produce useful feedback to the software measurement community. In any case, the wide-spread development of search-based Software Engineering is likely to inform and be informed by work on software measurement, since there is a clear and strong connection between software metrics and fitness evaluation.

5. Deeper appreciation of metaheuristics within the wider Software Engineering community.
   Currently, many software engineers are only tangentially familiar with metaheuristic search. It is hoped that the reformulation of many Software Engineering problems as search-based problems will widen awareness of metaheuristic search within the wider Software Engineering community.

6. Wider application of search-based Software Engineering in practice.
   Naturally, it is hoped that the next five years will witness a large increase in the application of metaheuristic search, building on the existing successes with test-data generation.

7. Provision of Surprising Solutions.
   One of the novel features which Search-Based Software Engineering adds to Software Engineering is the way in which the solutions found are occasionally surprising and not previously considered by analytic solutions. The metaheuristic search simply aims to optimise the value of the fitness function. Such a search-based approach is therefore entirely declarative; it is formulated in terms of what is required rather than how it is to be constructed. In other applications of metaheuristic search it has been observed that this declarative nature of the approach leads to unexpected solutions which tend to yield additional insight into the formulation of the problem.

5 Conclusion

Software Engineering problems are, by their very nature, ideal for the application of metaheuristic search techniques. This paper has argued that a new field of Software Engineering research, Search-Based Software Engineering is emerging. This sub-area of Software Engineering seeks to reformulate Software Engineering problems as search-based problems, facilitating the application of metaheuristic search-based techniques such as genetic algorithms, simulated annealing and tabu search.

The paper sets out key ingredients for this reformulation and provides criteria against which the success of individual approaches and the development of the field of research can be measured.

The authors believe that the development of Search-Based Software Engineering is long overdue. The aim of this paper (and part of the motivation for this special issue) is to encourage a 'research gold rush'. Many exciting and challenging problems in Software Engineering lie within reach of metaheuristic search.

Acknowledgements

This paper is essentially a manifesto for Search-Based Software Engineering. The authors have benefitted from extensive discussion within the auspices of the EPSRC-funded SEMINAL[1] network (GR/M78083). In particular, the authors would like to acknowledge the contribution made by Len Bottaci, Colin Burgess, John Clark, Jose Dolado, Gerdi Gross, Rob Hierons, Martin Lefley, Rudi Lutz, Mark Proctor, Vic Rayward Smith, Marc Roper, Nick Sharples, Martin Shepperd, Harman Sthamer, Nigel Tracey, Joachim Wegener and Darrell Whitley. Of course, all errors and omissions remain the sole responsibility of the authors.

[1] Software Engineering using Metaheuristic INnovative ALgorithms.

References

[1] Babovic, V. Mining sediment transport data with genetic programming. In *Proceedings of the First International Conference on New Information Technologies for Decision Making in Civil Engineering* (Montreal, Canada, 11-13 Oct. 1998), pp. 875–886.

[2] Baluja, S., Pomerleau, D., and Jochem, T. Simulating user's preferences: Towards automated artificial evolution for computer generated images. Tech. Rep. CMU-CS-93-198, Carnegie Mellon University, Pittsburgh, PA, USA, Oct. 1993.

[3] Balzer, R. Tolerating Inconsistency (Software Development). In *Proceedings of the 13th International Conference on Software Engineering* (May 1991), pp. 158–165.

[4] Barghouti, N. S., and Krishnamurthy, B. Using event contexts and matching constraints to monitor software processes effectively. In *Proceedings of the 17th International Conference on Software Engineering* (Apr. 1995), ACM Press, pp. 83–92.

[5] Bennett III, F. H., Keane, M. A., Andre, D., and Koza, J. R. Automatic synthesis of the topology and sizing for analog electrical circuits using genetic programming. In *Evolutionary Algorithms in Engineering and Computer Science* (Jyväskylä, Finland, 30 May - 3 June 1999), K. Miettinen, M. M. Mäkelä, P. Neittaanmäki, and J. Periaux, Eds., John Wiley & Sons, pp. 199–229.

[6] Bentley, P. J., and Wakefield, J. P. Generic representation of solid geometry for genetic search. *Microcomputers in Civil Engineering 11*, 3 (1996), 153–161.

[7] Biles, J. A. GenJam: A genetic algorithm for generating jazz solos. In *International Computer Music Conference* (San Francisco, CA, USA, 1994), International Computer Music Association, pp. 131–137.

[8] Birge, R. R. Protein-based optical computing and memories. *Computer 25*, 11 (Nov. 1992), 56–67.

[9] Caldwell, C., and Johnston, V. S. Tracking a criminal suspect through 'face-space' with a genetic algorithm. In *Fourth International Conference on Genetic Algorithms* (July 1991), Morgan Kaufmann, pp. 416–421.

[10] Cheung, S. C., and Kramer, J. Compositional Reachability Analysis of Finite-State Distributed Systems with User-Specified Constraints. In *Proceedings of SIGSOFT'95 Third ACM SIGSOFT Symposium on the Foundations of Software Engineering* (Oct. 1995), pp. 140–151.

[11] Cordón, O., Herrera, F., and Sánchez, L. Evolutionary learning processes for data analysis in electrical engineering applications. In *Genetic Algorithms and Evolution Strategy in Engineering and Computer Science*, D. Quagliarella, J. Périaux, C. Poloni, and G. Winter, Eds. John Wiley and Sons, Chichester, 1998, pp. 205–224.

[12] Dolado, J. J. A validation of the component-based method for software size estimation. *IEEE Transactions on Software Engineering 26*, 10 (2000), 1006–1021.

[13] Dolado, J. J. On the problem of the software cost function. *Information and Software Technology 43* (2001), 61–72.

[14] Feng, C.-W., Liu, L., and Burns, S. A. Using Genetic Algorithms to Solve Construction Time-Cost Trade-Off Problems. *Journal of Computing in Civil Engineering 10*, 3 (1999), 184–189.

[15] Fenton, N. E. Software measurement: A necessary scientific basis. *IEEE Transactions on Software Engineering 20*, 3 (1994), 199–206.

[16] Glover, F. Tabu search: A tutorial. *Interfaces 20* (1990), 74–94.

[17] Goldberg, D. E. *Genetic Algorithms in Search, Optimization & Machine Learning*. Addison-Wesley, Reading, MA, 1989.

[18] Helmreich, S. Recombination, rationality, reductionism and romantic reactions - culture, computers, and the genetic algorithm. *Social studies of Science 28*, 1 (1998), 39–71.

[19] Jenkins, W. M. The genetic algorithm-or can we improve design by breeding. In *IEE Colloquium on Artificial Intelligence in Civil Engineering* (London, UK, Jan., 16 1992), IEE, pp. 1/1–4.

[20] Jones, B., Sthamer, H.-H., and Eyres, D. Automatic structural testing using genetic algorithms. *The Software Engineering Journal 11* (1996), 299–306.

[21] Jones, B. F., Eyres, D. E., and Sthamer, H. H. A strategy for using genetic algorithms to automate branch and fault-based testing. *The Computer Journal 41*, 2 (1998), 98–107.

[22] Karr, C. L., Sharma, S. K., Hatcher, W. J., and Harper, T. R. Fuzzy control of an exothermic chemical reaction using genetic algorithms. *Engineering Applications of Artificial Intelligence 6 6* (1993), 575–582.

[23] Koza, J. R. *Genetic Programming: On the Programming of Computers by Means of Natural Selection.* MIT Press, Cambridge, MA, 1992.

[24] Labossiere, J. E., and Turrkan, N. On the optimization of the tensor polynomial failure theory with a genetic algorithm. *Transactions of the Canadian Society for Mechanical Engineering 16*, 3-4 (1992), 251–265.

[25] Miyamoto, Y., Miyatake, Y., Kurosaka, S., and Mori, Y. A parameter turning for dynamic simulation of power plants using genetic algorithms. *ELECTRICAL ENGINEERING IN JAPAN 115*, 1 (1995), 104–113.

[26] Ning, J., Miriyala, K., and Kozaczynski, W. An architecture-driven, business-specific, and component-based approach to software engineering. In *Proceedings of the Third International Conference on Software Reuse* (1994), pp. 84–93.

[27] Ou, Y. On using UML class diagram for object-oriented database design - specification of integrity constraints. In *The Unified Modeling Language, UML'98 - Beyond the Notation. First International Workshop, Mulhouse, France, June 1998* (1998), J. Bézivin and P.-A. Muller, Eds., pp. 147–156.

[28] Pargas, R. P., Harrold, M. J., and Peck, R. R. Test-data generation using genetic algorithms. *The Journal of Software Testing, Verification and Reliability 9* (1999), 263–282.

[29] Poli, R., Cagnoni, S., and Valli, G. Genetic design of optimum linear and nonlinear QRS detectors. *IEEE Transactions on Biomedical Engineering 42*, 11 (Nov. 1995), 1137–41.

[30] Pressman, R. S. *Software Engineering: A Practitioner's Approach*, 4th ed. McGraw-Hill, New York, NY., 1997. European adaptation. Adapted by Darrel Ince. ISBN 0-07-052182-4.

[31] Ryan, C., and Walsh, P. The evolution of provable parallel programs. In *Genetic Programming 1997: Proceedings of the Second Annual Conference* (Stanford University, CA, USA, 13-16 July 1997), J. R. Koza, K. Deb, M. Dorigo, D. B. Fogel, M. Garzon, H. Iba, and R. L. Riolo, Eds., Morgan Kaufmann, pp. 295–302.

[32] Sanchez, E., Miyano, H., and Brachet, J. P. Optimization of fuzzy queries with genetic algorithms. applications to a data base of patents in biomedical engineering. *Proc. Sixth International Fuzzy Systems Association World Congress (IFSA'95) 2* (1995), 293–296. Sao Paulo.

[33] Sebag, M., Schoenauer, M., and Maitournam, H. Parametric and non-parametric identification of macro-mechanical models. In *Genetic Algorithms and Evolution Strategy in Engineering and Computer Science*, D. Quagliarella, J. Périaux, C. Poloni, and G. Winter, Eds. John Wiley and Sons, Chichester, 1998, pp. 327–340.

[34] Shepperd, M. J. *Foundations of software measurement.* Prentice Hall, 1995.

[35] Sommerville, I. *Software Engineering*, 6th ed. Addison-Wesley, 2001.

[36] Tracey, N., Clark, J., and Mander, K. Automated program flaw finding using simulated annealing. In *International Symposium on Software Testing and Analysis* (March 1998), ACM/SIGSOFT, pp. 73–81.

[37] van Laarhoven, P. J. M., and Aarts, E. H. L. *Simulated Annealing: Theory and Practice.* Kluwer Academic Publishers, Dordrecht, the Netherlands, 1987.

[38] Wegener, J., Grimm, K., Grochtmann, M., Sthamer, H., and Jones, B. F. Systematic testing of real-time systems. In *4th International Conference on Software Testing Analysis and Review (EuroSTAR 96)* (1996).

[39] Wegener, J., Sthamer, H., Jones, B. F., and Eyres, D. E. Testing real-time systems using genetic algorithms. *Software Quality 6* (1997), 127–135.

[40] Weiss, R., and Knight, J. T. F. Engineered communications for microbial robotics. In *Proceedings 6th DIMACS Workshop on DNA Based Computers, held at the University of Leiden, Leiden, The Netherlands, 13 - 17 June 2000* (2000), DIMACS: Series in Discrete Mathematics and Theoretical Computer Science., Leiden center for natural computing, pp. 5–19.

[41] Whitley, D. A genetic algorithm tutorial. *Statistics and Computing 4* (1994), 65–85.

[42] Whitley, D. A free lunch proof for gray versus binary encodings. In *Proceedings of the Genetic and Evolutionary Computation Conference* (Orlando, Florida, USA, 13-17 July 1999), W. Banzhaf, J. Daida, A. E. Eiben, M. H. Garzon, V. Honavar, M. Jakiela, and R. E. Smith, Eds., vol. 1, Morgan Kaufmann, pp. 726–733.

[43] Whitley, D. An overview of evolutionary algorithms: Practical issues and common pitfalls. *Information and Software Technology Special Issue on Software Engineering using Metaheuristic Innovative Algorithms* (2001). This volume.

[44] Williams, K. P. *Evolutionary Algorithms for Automatic Parallelization.* PhD thesis, University of Reading, UK, Department of Computer Science, Sept. 1998.